Generalized hyperbolic distributions: Theory and applications to CDO pricing

Dissertation zur Erlangung des Doktorgrades
der Fakultät für Mathematik und Physik
der Albert-Ludwigs-Universität Freiburg im Breisgau

vorgelegt von

Ernst August Frhr. v. Hammerstein

Dezember 2010

Dekan:	Prof. Dr. Kay Königsmann
Referenten:	Prof. Dr. Ernst Eberlein
	Prof. Dr. Rüdiger Frey

Datum der Promotion: 12. Januar 2011

Abteilung für Mathematische Stochastik
Fakultät für Mathematik und Physik
Albert-Ludwigs-Universität Freiburg
Eckerstr. 1
79104 Freiburg im Breisgau

Abstract

This thesis can essentially be split up into two parts.

The theoretical part (Chapters 1 and 2) is devoted to a thorough study of uni- and multivariate generalized hyperbolic (GH) distributions which are defined as normal mean-variance mixtures with generalized inverse Gaussian (GIG) mixing distributions. We provide moment formulas and analyze the tail behaviour of the distribution functions and their convolutions, including all possible limit distributions which are derived in detail. Univariate GH and GIG distributions are shown to belong to the class of (extended) generalized Γ-convolutions which allows an explicit derivation of their Lévy–Khintchine representation and the construction of weakly convergent approximation schemes for the associated Lévy processes. From the formulas of the Lévy measure of multivariate GH distributions we conclude that, in contrast to the univariate case, not all of them are selfdecomposable. Moreover, we take a closer look at their dependence structure and show that they are either tail independent or completely dependent.

In the applied part (Chapter 3) we give a detailed introduction to synthetic CDOs and discuss the deficiencies of the normal factor model which is used as a market standard to price the latter. We demonstrate how these can be remedied by implementing more flexible and advanced factor distributions. Extended models using GH distributions provide an excellent fit to market data, but remain numerically tractable as the calibration examples to DJ iTraxx Europe spread quotes show. We also discuss further possible applications of the developed algorithms.

Acknowledgements

IUCUNDI ACTI LABORES. Cicero, de finibus bonorum et malorum.

It is, however, not only a pleasure to finish the thesis itself, but also to thank all the people without their help this would not have been achieved successfully.

First of all, I would like to express my gratitude to my academic teacher and supervisor Prof. Dr. Ernst Eberlein for his confidence, patience, and valuable support. He introduced me to mathematical finance and gave me the opportunity to participate in numerous conferences on this field.

I would also like to thank my former and present colleagues from the Department of Mathematical Stochastics at the University of Freiburg for stimulating, serious as well as funny discussions, especially Jan Bergenthum, Kathrin Glau, Zorana Grbac, Wolfgang Kluge, Ilse Maahs, Antonis Papapantoleon, and Volker Pohl. Prof. Dr. Hans Rudolf Lerche and Prof. Dr. Peter Pfaffelhuber gave helpful comments which improved the manuscript. Special thanks go to Monika Hattenbach who has been an inexhaustible source of help for all kinds of LaTeX-problems and many more.

I cordially thank my family and friends for all the encouragement, support, and impetus through the years.

I also gratefully acknowledge financial support from the DFG-Graduiertenkolleg "Nichtlineare Differentialgleichungen: Modellierung, Theorie, Numerik, Visualisierung" as well as from the BMBF-Projekt "Risikomanagement in Finanz- und Versicherungsdiensten: Hochdimensionale Modelle im Kreditrisikomanagement II".

Contents

Chapter 1

Univariate GH distributions and their limits

In continuous time finance prices $S = (S_t)_{t\geq 0}$ of risky assets are often modeled by exponential processes of the form

$$S_t = S_0 e^{L_t}$$

where $L = (L_t)_{t\geq 0}$ is a Lévy process. This approach is based on the assumption that log returns from price series which are recorded along equidistant time grids are independent and identically distributed random variables. The model is completely specified by $\mathcal{L}(L_1)$, the distribution of the Lévy process at time 1. It is a crucial property of this model that the log returns $\ln(S_t) - \ln(S_{t-1}) = L_t - L_{t-1}$ are independent and, because of the stationarity of L, also $\mathcal{L}(L_1)$-distributed. Thus if $\mathcal{L}(L_1)$ is the (infinitely divisible) distribution derived from fitting data—say daily closing prices—the log returns from the model should have almost exactly the distribution which one sees in the data. The goodness of fit of theoretical to empirical densities shows if the model is able to reproduce the observed market movements reasonably well. In the classical case, the so-called geometric Brownian motion, the Lévy process is

$$L_t = \sigma B_t + \left(\mu - \frac{\sigma^2}{2}\right) t, \qquad \mu \in \mathbb{R},\, \sigma > 0, \tag{1.1}$$

where $(B_t)_{t\geq 0}$ denotes a standard Brownian motion, μ is a drift and σ a volatility parameter. In this case the distribution of daily log returns is a normal one.

This model has become the standard in financial mathematics, although its deficiencies are widely known. Comparing empirical densities with normal ones exhibits substantial deviations between them. Empirical densities are typically leptokurtic, that is, they have more mass around the origin and in the tails, but less in the flanks, and in addition they are often skewed. It is therefore a natural task to look for alternative classes of infinitely divisible distributions which provide a better fit to market data and induce a Lévy process L such that the model above admits more realistic log return distributions. We just name a few references from the large amount of literature on this line of research here, a more comprehensive overview can be obtained in the books of Schoutens (2003) and Cont and Tankov (2004). One of the first to investigate the deviations from normality was Mandelbrot (1963). He suggested to replace the Brownian motion by a symmetric α-stable Lévy process. Other proposals include the Student's t-distribution (Blattberg and Gonedes 1974), the Variance Gamma distribution (Madan and Seneta 1990; Madan, Carr, and Chang 1998), the Normal Inverse Gaussian distribution (Barndorff-Nielsen 1998; Barndorff-Nielsen and Prause 2001) and the CGMY distribution (Carr, Geman, Madan, and Yor 2002).

Most of the above mentioned distributions are, as we shall see in the following sections, either special subclasses or limiting cases of the class of generalized hyperbolic distributions which allows an almost perfect fit to financial data. The hyperbolic subclass was the first to be introduced to finance as a more realistic model for stock returns in Eberlein and Keller (1995) and Eberlein, Keller, and Prause (1998). The general case was considered in Eberlein (2001), Eberlein and Prause (2002) and Eberlein and Özkan (2003). Generalized hyperbolic distributions have further been successfully applied to interest rate theory where they allow for a very accurate pricing of caplets and other derivatives, see for example Eberlein and Raible (1999), Eberlein and Kluge (2006) and Eberlein and Kluge (2007), and on currency market data (Eberlein and Koval 2006). In Chapter 3 it will be shown that generalized hyperbolic distributions also enable significant improvements in credit portfolio modeling and CDO pricing. But first we want to provide a thorough investigation of this distribution class itself and give detailed proofs for some of its most useful and interesting properties.

The present chapter is devoted to the study of univariate generalized hyperbolic distributions. Section 1.1 contains some technical preliminaries and inroduces the concept of normal mean-variance mixtures. In Section 1.2 we take a closer look at a special class of mixture distributions, the generalized inverse Gaussian distributions, which lead to the class of generalized hyperbolic distributions. After presenting some basic facts of the latter in Section 1.3, we derive all possible limit distributions in Section 1.4. Section 1.5 gives a short introduction to the class of extended generalized Γ-convolutions. In Section 1.6 we show that this class contains both the generalized inverse Gaussian and the generalized hyperbolic distributions as subclasses which allows an explicit derivation of their Lévy–Khintchine representations. Moreover, this result entails the possibility to approximate these distributions and the corresponding Lévy processes by sums of suitably scaled and shifted Gamma variables which is discussed in Section 1.7. Greater parts of this chapter are based on Eberlein and v. Hammerstein (2004) where many of the results can also be found.

1.1 Infinitely divisible distributions and normal mean-variance mixtures

We want to provide some basic definitions and facts which will be referred to in later sections and start with

Definition 1.1 *A probability measure μ on $(\mathbb{R}, \mathcal{B})$ is* infinitely divisible *if for any integer $n \geq 1$ there exists a probability measure μ_n on $(\mathbb{R}, \mathcal{B})$ such that μ equals the n-fold convolution of μ_n, that is, $\mu = \mu_n * \cdots * \mu_n =: *_{i=1}^{n} \mu_n$.*

The characteristic function $\phi_\mu(u) := \int_{\mathbb{R}} e^{iux}\mu(\mathrm{d}x)$ of every infinitely divisible distribution μ can be repesented in a very special form, the *Lévy–Khintchine formula*:

$$\phi_\mu(u) = \exp\left(iub - \frac{1}{2}cu^2 + \int_{\mathbb{R}} \left(e^{iux} - 1 - iux\mathbb{1}_{[-1,1]}(x)\right)\nu(\mathrm{d}x)\right).$$

In this representation, the coefficients $b \in \mathbb{R}$, $c \geq 0$ and the *Lévy measure* $\nu(\mathrm{d}x)$ which satisfies $\nu(\{0\}) = 0$ and $\int_{\mathbb{R}}(x^2 \wedge 1)\,\nu(\mathrm{d}x) < \infty$ are unique (see for example Sato 1999, Theorem 8.1) and completely characterize μ.

Definition 1.2 *Assume a filtered probability space $(\Omega, \mathcal{F}, (\mathcal{F}_t)_{t\geq 0}, P)$ with increasing filtration to be given. An adapted process $L = (L_t)_{t\geq 0}$, with values in $\mathbb{R}$ and $L_0 = 0$ almost surely, is a* Lévy process *if the following conditions hold:*

1. *L has* independent increments*, that is, $L_t - L_s$ is independent of $\mathcal{F}_s$, $0 \leq s < t < \infty$.*
2. *L has* stationary increments*, that is, $L_t - L_s$ has the same distribution as L_{t-s}, $0 \leq s < t < \infty$.*
3. *L is* continuous in probability*, that is, $\lim_{s\to t} P(|L_s - L_t| > \epsilon) = 0$.*

The stationarity and independence of the increments $L_t - L_s$ imply that the distribution $\mathcal{L}(L_t)$ is infinitely divisible for all $t \in \mathbb{R}_+$, and its characteristic function ϕ_{L_t} fulfills $\phi_{L_t}(u) = \phi_{L_1}(u)^t$. Conversely, every infinitely divisible distribution μ induces a Lévy process L via $\phi_{L_t}(u) = \phi_\mu(u)^t$ (Sato 1999, Theorem 7.10).

An important subclass of infinitely divisible distributions are the selfdecomposable distributions (also called *(Lévy's) class L*) which are defined as follows:

Definition 1.3 *A probability measure μ on $(\mathbb{R}, \mathcal{B})$ is called* selfdecomposable *if for every $0 < s < 1$ there is a probability measure μ_s on $(\mathbb{R}, \mathcal{B})$ such that*

$$\phi_\mu(u) = \phi_\mu(su)\phi_{\mu_s}(u).$$

Equivalently, a real valued random variable X is said to have a selfdecomposable distribution if for every $0 < s < 1$ there exists a real valued random variable $X^{(s)}$ independent of X such that

$$X \overset{d}{=} sX + X^{(s)},$$

where $\overset{d}{=}$ means equality in distribution.

The selfdecomposable distributions are uniquely characterized by the following lemma which can be found in Sato (1999, Corollary 15.11).

Lemma 1.4 *A probability measure μ on $(\mathbb{R}, \mathcal{B})$ is selfdecomposable if and only if it is infinitely divisible and has a Lévy measure of the form $\nu(\mathrm{d}x) = \frac{k(x)}{|x|}\,\mathrm{d}x$ where $k(x)$ is non-decreasing on $(-\infty, 0)$ and non-increasing on $(0, \infty)$.*

In particular, the Lévy measure of every selfdecomposable distribution possesses a Lebesgue density that is strictly increasing on $\mathbb{R}_-$ and strictly decreasing on $\mathbb{R}_+$. The properties of this subclass are very useful in financial modeling; a discussion of the implications on option pricing can be found in Carr, Geman, Madan, and Yor (2007).

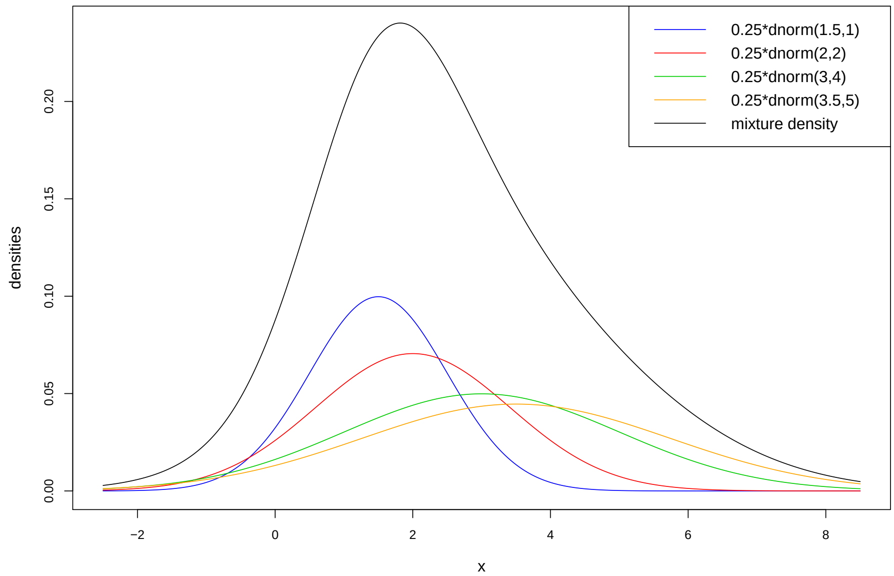


Figure 1.1: Discrete normal mixture of four equally weighted normal densities

As already mentioned in the introduction, realistic models for return distributions should also allow for skewness and kurtosis. For a distribution F having a finite fourth moment, the *skewness* $\gamma_1(F)$ and the *excess kurtosis* $\gamma_2(F)$ are defined as follows: $\gamma_1(F) := m_3 m_2^{-3/2}$ and $\gamma_2(F) := m_4 m_2^{-2} - 3$, where m_k denotes the kth central moment of F. For every normal distribution $N(\mu, \sigma^2)$ one has $\gamma_1(N(\mu, \sigma^2)) = 0 = \gamma_2(N(\mu, \sigma^2))$. Although normal distributions are neither skewed nor leptokurtic themselves, it is, however, fairly easy to construct new distributions from them which do have these properties: one just has to pass from single distributions to mixtures. In the simplest case, a normal mixture is a weighted average of several normal distributions with different means and variances which has the density $f_{\mathrm{mix}}(x) = \sum_{i=1}^n a_i d_{N(\mu_i, \sigma_i^2)}(x)$ where $a_i \geq 0$, $\sum_{i=1}^n a_i = 1$ and $d_{N(\mu,\sigma^2)}$ denotes the density of $N(\mu, \sigma^2)$. Figure 1.1 shows an example where four normal densities are mixed with equal weights $a_i = 0.25$. Apparently, the obtained mixture density f_{mix} (in black) is skewed and has positive kurtosis. Some simple calculations yield that for the above choice of parameters we have $\gamma_1(F_{\mathrm{mix}}) = 0.543$ and $\gamma_2(F_{\mathrm{mix}}) = 0.522$. This concept can be formalized by

Definition 1.5 *A real valued random variable X is said to have a* normal mean-variance mixture distribution *if*

$$X \stackrel{d}{=} \mu + \beta Z + \sqrt{Z} W,$$

where $\mu, \beta \in \mathbb{R}$, $W \sim N(0,1)$ and $Z \sim G$ is a real-valued, non-negative random variable which is independent of W.

Equivalently, a probability measure F on $(\mathbb{R}, \mathcal{B})$ is said to be a normal mean-variance mixture if

$$F(\mathrm{d}x) = \int_{\mathbb{R}_+} N(\mu + \beta y, y)(\mathrm{d}x)\, G(\mathrm{d}y),$$

where the mixing distribution G is a probability measure on $(\mathbb{R}_+, \mathcal{B}_+)$.
We shall use the short hand notation $F = N(\mu + \beta y, y) \circ G$. If $\mathbb{G}$ is a class of mixing distributions, then $N(\mu+\beta y, y) \circ \mathbb{G} := \{N(\mu+\beta y, y) \circ G \,|\, G \in \mathbb{G}, \mu \in \mathbb{R}\}$.

Remark: Sometimes random variables with normal mean-variance mixture distributions are defined more generally by $X \overset{d}{=} m(Z) + \sqrt{Z}W$ with an arbitrary measurable function $m : \mathbb{R}_+ \to \mathbb{R}$ (see for example McNeil, Frey, and Embrechts 2005, Definition 3.11), but for most purposes the above setting $m(z) = \mu + \beta z$ is completely sufficient. If $\beta = 0$, one obtains the so-called *normal variance mixtures* which are obviously symmetric around μ and therefore have no skewness.

Also note that the mixture distribution F in general does not possess a Lebesgue density. If for example $G = \mathrm{Pois}(\lambda)$ with $\lambda \in \mathbb{R}_+$, then $F(\{\mu\}) = e^{-\lambda} > 0$. Necessary conditions that ensure the existence of a density of F are that the mixing distribution G either possesses a density itself or has an at most countable support which is bounded away from 0.

The most important facts about normal mean-variance mixtures are summarized in the following lemma. It especially shows that properties like stability under convolutions, infinite divisibility and selfdecomposability are inherited from the mixing distributions.

Lemma 1.6 *Let $\mathbb{G}$ be a class of probability distributions on $(\mathbb{R}_+, \mathcal{B}_+)$ and $G, G_1, G_2 \in \mathbb{G}$.*

a) If G possesses a moment generating function $M_G(u) = \int_{\mathbb{R}_+} e^{ux} G(\mathrm{d}x)$ on some open interval (a, b) with $a < 0 < b$, then $F = N(\mu + \beta y, y) \circ G$ also possesses a moment generating function and $M_F(u) = e^{\mu u}\, M_G(\frac{u^2}{2} + \beta u)$, $a < \frac{u^2}{2} + \beta u < b$.

b) If $G = G_1 * G_2 \in \mathbb{G}$, then $\big(N(\mu_1 + \beta y, y) \circ G_1\big) * \big(N(\mu_2 + \beta y, y) \circ G_2\big) = N(\mu_1 + \mu_2 + \beta y, y) \circ G \in N(\mu + \beta y, y) \circ \mathbb{G}$.

c) If G is infinitely divisible, then so is $N(\mu + \beta y, y) \circ G$.

d) If G is selfdecomposable, then so is $N(\mu + \beta y, y) \circ G$.

Proof: a) Since the moment generating function of a normal distribution $N(\mu, \sigma^2)$ is given by $M_{N(\mu,\sigma^2)}(u) = e^{\frac{\sigma^2 u^2}{2} + \mu u}$, we get with the help of Fubini's theorem

$$\begin{aligned} M_F(u) &= \int_{\mathbb{R}} e^{ux} \int_{\mathbb{R}_+} N(\mu + \beta y, y)(\mathrm{d}x)\, G(\mathrm{d}y) = \int_{\mathbb{R}_+} G(\mathrm{d}y) \int_{\mathbb{R}} e^{ux} N(\mu + \beta y, y)(\mathrm{d}x) \\ &= \int_{\mathbb{R}_+} e^{\mu u}\, e^{(\frac{u^2}{2} + \beta u)y}\, G(\mathrm{d}y) = e^{\mu u}\, M_G\big(\tfrac{u^2}{2} + \beta u\big). \end{aligned}$$

b) Since G_1, G_2 are measures on $(\mathbb{R}_+, \mathcal{B}_+)$, their Laplace transforms $\mathfrak{L}_{G_i}(u) = \int_{\mathbb{R}_+} e^{-ux} G_i(\mathrm{d}x)$ are well defined for $u \in \mathbb{R}_+$ and can be extended to complex arguments $z \in \mathbb{C}$ with $Re(z) \geq 0$ because $\int_{\mathbb{R}_+} |e^{-zx}| G_i(\mathrm{d}x) = \mathfrak{L}_{G_i}(Re(z)) < \infty$. From Doetsch (1950, p. 123, Satz 3) we have $\mathfrak{L}_{G_1 * G_2}(z) = \mathfrak{L}_{G_1}(z)\mathfrak{L}_{G_2}(z)$ (see also Raible (2000, Theorem B.2)). Similar to a) we obtain for the characteristic function of $F := N(\mu + \beta y, y) \circ G$:

$$\begin{aligned}
\phi_F(u) &= \int_{\mathbb{R}} e^{iux} F(\mathrm{d}x) = \int_{\mathbb{R}} e^{iux} \int_{\mathbb{R}_+} N(\mu + \beta y, y)(\mathrm{d}x)\, G(\mathrm{d}y) \\
&= \int_{\mathbb{R}_+} G(\mathrm{d}y) \int_{\mathbb{R}} e^{iux} N(\mu + \beta y, y)(\mathrm{d}x) = \int_{\mathbb{R}_+} e^{iu\mu}\, e^{-\left(\frac{u^2}{2} - iu\beta\right)y}\, G(\mathrm{d}y) \\
&= e^{iu\mu}\, \mathfrak{L}_G\left(\tfrac{u^2}{2} - iu\beta\right).
\end{aligned}$$

Now if $F = N(\mu_1 + \mu_2 + \beta y, y) \circ G$, $F_i := N(\mu_i + \beta y, y) \circ G_i$, $i = 1, 2$, and $G_1 * G_2 = G$, then

$$\begin{aligned}
\phi_{F_1}(u)\phi_{F_2}(u) &= e^{iu\mu_1}\, \mathfrak{L}_{G_1}\left(\tfrac{u^2}{2} - iu\beta\right) e^{iu\mu_2}\, \mathfrak{L}_{G_2}\left(\tfrac{u^2}{2} - iu\beta\right) \\
&= e^{iu(\mu_1+\mu_2)} \mathfrak{L}_G\left(\tfrac{u^2}{2} - iu\beta\right) = \phi_F(u)
\end{aligned}$$

which proves $F_1 * F_2 = F$.

c) If G is infinitely divisible, then by Definition 1.1 for every $n \geq 1$ there exists a probability measure G_n on $(\mathbb{R}_+, \mathcal{B}_+)$ (which not necessarily is an element of $\mathbb{G}$ itself) such that $G = *_{i=1}^n G_n$. Analogously to part b) it follows

$$N(\mu + \beta y, y) \circ G = *_{i=1}^n \left(N\left(\tfrac{\mu}{n} + \beta y, y\right) \circ G_n\right),$$

hence again by Definition 1.1 also $N(\mu + \beta y, y) \circ G$ is infinitely divisible.

d) Here we give a short proof for the special case that $F = N(\mu, y) \circ G$ is a normal variance mixture, using the idea of Halgreen (1979). The general result was established in Sato (2001, Theorem 1.1).
Because probability measures on $(\mathbb{R}_+, \mathcal{B}_+)$ are uniquely determined by their Laplace transforms (see for example Feller 1971, Chapter XIII.1, Theorem 1), we may rewrite the first part of Definition 1.3 for G in the following way: For every $0 < s < 1$ we have $\mathfrak{L}_G(u) = \mathfrak{L}_G(su)\mathfrak{L}_{G_s}(u)$, where G_s is some probability measure on $(\mathbb{R}_+, \mathcal{B}_+)$, too. If $0 < s < 1$, then also $0 < s^2 < 1$, so it follows from the proof of part b) that

$$\frac{\phi_F(u)}{\phi_F(su)} = e^{i\mu(1-s)u}\, \frac{\mathfrak{L}_G\left(\frac{u^2}{2}\right)}{\mathfrak{L}_G\left(\frac{s^2u^2}{2}\right)} = e^{i\mu(1-s)u}\, \mathfrak{L}_{G_{s^2}}\left(\tfrac{u^2}{2}\right) = \phi_{F_s}(u)$$

where $F_s := N((1-s)\mu, y) \circ G_{s^2}$ is a probability measure on $(\mathbb{R}, \mathcal{B})$, hence F is selfdecomposable by Definition 1.3. □

Lemma 1.7 *Let $(\mu_n)_{n\geq 1}$ and $(\beta_n)_{n\geq 1}$ be convergent sequences of real numbers with finite limits $\mu, \beta < \infty$, and $(G_n)_{n\geq 1}$ be a sequence of mixing distributions with $G_n \xrightarrow{w} G$. Then $N(\mu_n + \beta_n y, y) \circ G_n \xrightarrow{w} N(\mu + \beta y, y) \circ G$.*

PROOF: Let $F_n := N(\mu_n + \beta_n y, y) \circ G_n$ and $F := N(\mu + \beta y, y) \circ G$. According to the proof of part b) of the previous lemma we have to show that for an arbitrarily fixed $u \in \mathbb{R}$

$$\phi_{F_n}(u) = e^{iu\mu_n}\, \mathfrak{L}_{G_n}\big(\tfrac{u^2}{2} - iu\beta_n\big) \underset{n\to\infty}{\longrightarrow} e^{iu\mu}\, \mathfrak{L}_G\big(\tfrac{u^2}{2} - iu\beta\big) = \phi_F(u).$$

Because $e^{iu\mu_n} \to e^{iu\mu}$ holds trivially, it suffices to prove convergence of the above Laplace transforms. From Doetsch (1950, pp. 71/72 and 156) it follows that the Laplace transform $\mathfrak{L}_H(z)$ of every finite measure H on $(\mathbb{R}_+, \mathcal{B}_+)$ is holomorphic on the open complex half-plane $\mathbb{C}_+^o := \{z \in \mathbb{C} \,|\, Re(z) > 0\}$, and the weak convergence of the sequence $(G_n)_{n\geq 1}$ implies $\mathfrak{L}_{G_n}(v) \to \mathfrak{L}_G(v)$ pointwise for all $v \in \mathbb{R}_+$. Moreover, $|\mathfrak{L}_{G_n}(z)| \leq \mathfrak{L}_{G_n}\big(Re(z)\big) \leq 1$. In particular $(\mathfrak{L}_{G_n})_{n\geq 1}$ is a locally bounded sequence of holomorphic functions on $\mathbb{C}_+^o$ which converges on a subset that has an accumulation point in $\mathbb{C}_+^o$. Thus by Jänich (1996, Korollar on p. 96) the sequence converges uniformly on every compact subset of $\mathbb{C}_+^o$. Since $\beta_n \to \beta < \infty$ by assumption, $\{\frac{u^2}{2} - iu\beta_n \,|\, n \in \mathbb{N}\} \cup \{\frac{u^2}{2} - iu\beta\}$ is a compact subset of $\mathbb{C}_+^o$ (remember u is fixed), hence we have $\mathfrak{L}_{G_n}\big(\frac{u^2}{2} - iu\beta_n\big) \to \mathfrak{L}_G\big(\frac{u^2}{2} - iu\beta\big)$ which completes the proof. □

As pointed out before, the Lévy processes induced by infinitely divisible normal mean-variance mixtures provide a natural and more realistic generalization of the classical model (1.1). The next proposition shows that they can be represented as subordinated Brownian motions where the subordinator is generated by the mixing distribution.

Proposition 1.8 *Let $F = N(\mu + \beta y, y) \circ G$ be a normal mean-variance mixture with infinitely divisible mixing distribution G and $(X_t)_{t\geq 0}$, $(\tau(t))_{t\geq 0}$ be two Lévy processes with $\mathcal{L}(X_1) = F$ and $\mathcal{L}(\tau(1)) = G$. Then $(Y_t)_{t\geq 0}$, defined by*

$$Y_t := \mu t + \beta\tau(t) + B_{\tau(t)},$$

where $(B_t)_{t\geq 0}$ is a standard Brownian motion independent of $(\tau(t))_{t\geq 0}$, is a Lévy process that is identical in law to $(X_t)_{t\geq 0}$.

PROOF: First observe that the process $(\tau(t))_{t\geq 0}$ is increasing because the fact that $\mathcal{L}(\tau(1)) = G$ is a measure concentrated on $\mathbb{R}_+$ implies that for every $t > 0$ the same holds for $\mathcal{L}(\tau(t))$ (Sato 1999, Theorem 24.11). By stationarity we have $\mathcal{L}(\tau(t+\epsilon) - \tau(t)) = \mathcal{L}(\tau(\epsilon))$, hence for arbitrary $t \geq 0$ and $\epsilon > 0$ the increment $\tau(t+\epsilon) - \tau(t)$ is almost surely non-negative.

If $(B_t)_{t\geq 0}$ is a standard Brownian motion independent of $(\tau(t))_{t\geq 0}$, then $(\beta\tau(t) + B_{\tau(t)})_{t\geq 0}$ is a Lévy process according to Sato (1999, Theorem 30.1) and hence so is $(Y_t)_{t\geq 0}$. The characteristic function of Y_1 is given by

$$\begin{aligned}\phi_{Y_1}(u) &= \mathrm{E}\big[e^{iuY_1}\big] = e^{iu\mu}\, \mathrm{E}\big[e^{iu\beta\tau(1)}\mathrm{E}[e^{iuB_{\tau(1)}}|\tau(1)]\big] \\ &= e^{iu\mu}\, \mathrm{E}\Big[e^{-(\frac{u^2}{2} - iu\beta)\tau(1)}\Big] = e^{iu\mu}\, \mathfrak{L}_G\big(\tfrac{u^2}{2} - iu\beta\big),\end{aligned}$$

because $\mathcal{L}(\tau(1)) = G$. From the proof of Lemma 1.6 b) we know that $\phi_{Y_1}(u) = \phi_F(u) = \phi_{X_1}(u)$, so $(X_t)_{t\geq 0}$ and $(Y_t)_{t\geq 0}$ are Lévy processes with $\mathcal{L}(X_1) = \mathcal{L}(Y_1)$. The assertion now follows from Sato (1999, Theorem 7.10 (iii)). □

Remark: By Jacod and Shiryaev (2003, Corollary VII.3.6) a sequence of Lévy processes $(X_t^n)_{t\geq 0}$ converges in law to a Lévy process $(X_t)_{t\geq 0}$ if and only if $X_1^n \xrightarrow{\mathcal{L}} X_1$. Combining Lemma 1.7 and Proposition 1.8 we see that if $(\tau_n(t))_{t\geq 0}$ are Lévy processes with $\mathcal{L}(\tau_n(1)) = G_n$, then the sequence $(Y_t^n)_{t\geq 0}$ with $Y_t^n := \mu_n t + \beta_n \tau_n(t) + B_{\tau_n(t)}$ converges in law to the process $(Y_t)_{t\geq 0}$. This can easily be extended to general subordinated Lévy processes; a corresponding result can be found in Küchler and Tappe (2008, Lemma 3.2). The proof given there is similar to ours of Lemma 1.7, but—in our opinion—incomplete since they just claim the uniform convergence of the Laplace transforms would hold and could be proven analogously to Lévy's continuity theorem. They do not seem to be aware that this line of argumentation only yields uniform convergence on compact subsets of $\mathbb{R}$, whereas in the present case the domain of the Laplace transforms is $\mathbb{C}_+$, and the extension of the well-known convergence result to complex arguments requires a more precise justification.

Comparing the Lévy processes $(Y_t)_{t\geq 0}$ defined in the proposition above and $(L_t)_{t\geq 0}$ from (1.1) one might also conclude that more realistic models emerge from the classical one via suitable time changes $t \rightsquigarrow \tau(t)$. This new time $\tau(t)$ is often called operational or business time and can be regarded as a measure of economic activity. Since the latter is obviously not evolving uniformly, the introduction of a "random clock" seems to be quite natural also from this perspective. An extensive discussion on this topic can be found in Geman, Madan, and Yor (2001).

1.2 Generalized inverse Gaussian distributions

In this section we are concerned with a special class of mixing distributions: the generalized inverse Gaussian distributions (henceforth GIG). This class was introduced more than 50 years ago (one of the first papers where its densities are mentioned is Good (1953)) and rediscovered by Sichel (1973, 1974) and Barndorff-Nielsen (1977). An extensive survey with statistical applications can be found in Jørgensen (1982). The density of a GIG distribution is as follows:

$$d_{GIG(\lambda,\delta,\gamma)}(x) = \left(\frac{\gamma}{\delta}\right)^{\lambda} \frac{1}{2K_\lambda(\delta\gamma)}\, x^{\lambda-1} e^{-\frac{1}{2}(\delta^2 x^{-1}+\gamma^2 x)}\, \mathbb{1}_{(0,\infty)}(x), \tag{1.2}$$

where $K_\lambda(x)$ denotes the modified Bessel function of third kind with index λ (see Appendix A for further details). Permitted parameters are

$$\begin{aligned} &\delta \geq 0,\ \gamma > 0, \quad \text{if } \lambda > 0,\\ &\delta > 0,\ \gamma > 0, \quad \text{if } \lambda = 0,\\ &\delta > 0,\ \gamma \geq 0, \quad \text{if } \lambda < 0. \end{aligned}$$

Parametrizations with $\delta = 0$ or $\gamma = 0$ have to be understood as limiting cases.

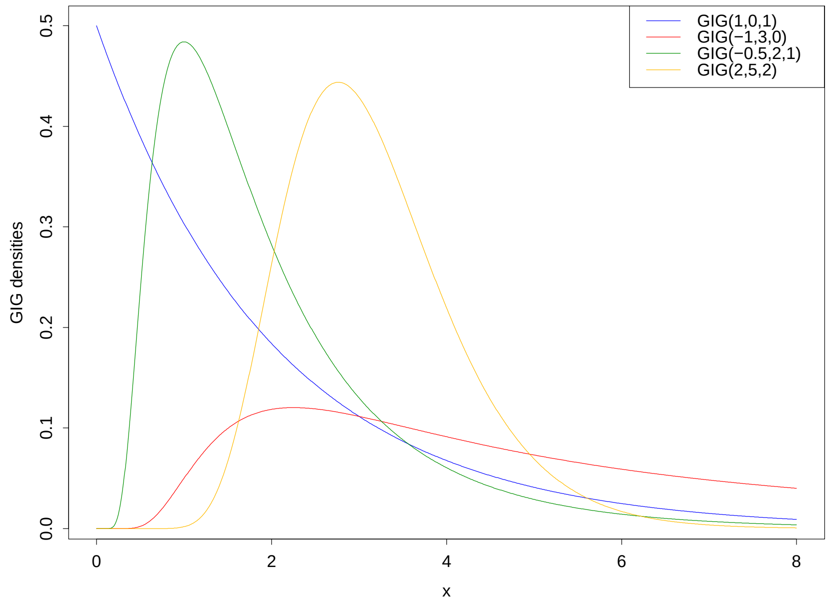


Figure 1.2: Densities of various GIG distributions

By equation (A.8) we have

$$\left(\frac{\gamma}{\delta}\right)^{\lambda} \frac{1}{2K_\lambda(\delta\gamma)} \longrightarrow \begin{cases} \left(\frac{\gamma^2}{2}\right)^{\lambda} \frac{1}{\Gamma(\lambda)}, & \lambda > 0,\ \delta \to 0, \\ 0, & \lambda > 0,\ \gamma \to 0, \\ 0, & \lambda < 0,\ \delta \to 0, \\ \left(\frac{2}{\delta^2}\right)^{\lambda} \frac{1}{\Gamma(-\lambda)}, & \lambda < 0,\ \gamma \to 0, \end{cases}$$

hence in the second and third case one obtains no probability density in the limit. (Here and in the following $\Gamma(x)$ denotes the Gamma function.) In the two other cases the limits are given by

$$d_{GIG(\lambda,0,\gamma)}(x) = \left(\frac{\gamma^2}{2}\right)^{\lambda} \frac{x^{\lambda-1}}{\Gamma(\lambda)} e^{-\frac{\gamma^2}{2}x} \mathbb{1}_{(0,\infty)}(x), \qquad \lambda > 0, \tag{1.3}$$

$$d_{GIG(\lambda,\delta,0)}(x) = \left(\frac{2}{\delta^2}\right)^{\lambda} \frac{x^{\lambda-1}}{\Gamma(-\lambda)} e^{-\frac{\delta^2}{2x}} \mathbb{1}_{(0,\infty)}(x), \qquad \lambda < 0, \tag{1.4}$$

which are the densities of a *Gamma distribution* $G\big(\lambda, \frac{\gamma^2}{2}\big)$ and an *inverse Gamma distribution* $iG\big(\lambda, \frac{\delta^2}{2}\big)$, respectively. The blue and red densities in Figure 1.2 show an example of each case.
If $\lambda = 0$, then equation (A.9) implies

$$\left(\frac{\gamma}{\delta}\right)^{\lambda} \frac{1}{2K_\lambda(\delta\gamma)} = \frac{1}{2K_0(\delta\gamma)} \sim \frac{1}{-2\ln(\delta\gamma)} \to 0 \qquad \text{if } \ \delta\gamma \to 0,$$

therefore a $GIG(0,\delta,\gamma)$-distribution does not converge weakly if $\delta \to 0$ or $\gamma \to 0$. If $\lambda = -\frac{1}{2}$, we get with the help of (A.7)

$$d_{GIG(-\frac{1}{2},\delta,\gamma)}(x) = \frac{\gamma}{\sqrt{2\pi x^3}}\, e^{-\frac{1}{2x}(\delta+\gamma x)^2}\, \mathbb{1}_{(0,\infty)}(x),$$

which equals the density of an *inverse Gaussian distribution* $IG(\delta,\gamma)$, so the GIG distributions are in fact a natural extension of this subclass.

Proposition 1.9 *The Laplace transforms of GIG distributions are given by*

$$\begin{aligned}
\mathfrak{L}_{GIG(\lambda,\delta,\gamma)}(u) &= \left(\frac{\gamma}{\sqrt{\gamma^2+2u}}\right)^{\lambda} \frac{K_\lambda\big(\delta\sqrt{\gamma^2+2u}\big)}{K_\lambda(\delta\gamma)}, && \delta,\gamma>0,\\
\mathfrak{L}_{GIG(\lambda,0,\gamma)}(u) &= \left(1+\frac{2u}{\gamma^2}\right)^{-\lambda}, && \lambda>0,\\
\mathfrak{L}_{GIG(\lambda,\delta,0)}(u) &= \left(\frac{2}{\delta\sqrt{2u}}\right)^{\lambda} \frac{2K_\lambda\big(\delta\sqrt{2u}\big)}{\Gamma(-\lambda)}, && \lambda<0.
\end{aligned}$$

Proof: Let $n_1(\lambda,\delta,\gamma) := \big(\frac{\gamma}{\delta}\big)^{\lambda} \frac{1}{2K_\lambda(\delta\gamma)}$ denote the norming constant of the GIG density, then

$$\begin{aligned}
\mathfrak{L}_{GIG(\lambda,\delta,\gamma)}(u) &= \int_0^\infty n_1(\lambda,\delta,\gamma)\, e^{-ux} x^{\lambda-1} e^{-\frac{1}{2}\left(\delta^2 x^{-1}+\gamma^2 x\right)}\, \mathrm{d}x\\
&= \int_0^\infty n_1(\lambda,\delta,\gamma)\, x^{\lambda-1} e^{-\frac{1}{2}\left(\delta^2 x^{-1}+(\gamma^2+2u)x\right)}\, \mathrm{d}x\\
&= \frac{n_1(\lambda,\delta,\gamma)}{n_1\big(\lambda,\delta,\sqrt{\gamma^2+2u}\big)} = \left(\frac{\gamma}{\sqrt{\gamma^2+2u}}\right)^{\lambda} \frac{K_\lambda\big(\delta\sqrt{\gamma^2+2u}\big)}{K_\lambda(\delta\gamma)}.
\end{aligned}$$

Similarly, we get with $n_2\big(\lambda,\frac{\gamma^2}{2}\big) := \big(\frac{\gamma^2}{2}\big)^{\lambda}\Gamma(\lambda)^{-1}$

$$\mathfrak{L}_{GIG(\lambda,0,\gamma)}(u) = \frac{n_2\big(\lambda,\frac{\gamma^2}{2}\big)}{n_2\big(\lambda,\frac{\gamma^2}{2}+u\big)} = \left(1+\frac{2u}{\gamma^2}\right)^{-\lambda}.$$

For the inverse Gamma limit we set $n_3\big(\lambda,\frac{\delta^2}{2}\big) := \big(\frac{2}{\delta^2}\big)^{\lambda}\Gamma(-\lambda)^{-1}$ and calculate

$$\begin{aligned}
\mathfrak{L}_{GIG(\lambda,\delta,0)}(u) &= \int_0^\infty n_3\big(\lambda,\tfrac{\delta^2}{2}\big)\, e^{ux}\, x^{\lambda-1}\, e^{-\frac{\delta^2}{2x}}\, \mathrm{d}x\\
&= \int_0^\infty n_3\big(\lambda,\tfrac{\delta^2}{2}\big)\, x^{\lambda-1}\, e^{-\frac{1}{2}\left(\delta^2 x^{-1}+2ux\right)}\, \mathrm{d}x\\
&= \frac{n_3\big(\lambda,\frac{\delta^2}{2}\big)}{n_1\big(\lambda,\delta,\sqrt{2u}\big)} = \left(\frac{2}{\delta\sqrt{2u}}\right)^{\lambda} \frac{2K_\lambda\big(\delta\sqrt{2u}\big)}{\Gamma(-\lambda)}.
\end{aligned}$$

□

Remark: The proof of Lemma 1.6 b) also implies that the characteristic functions of GIG distributions can be obtained from the above via the relation $\phi_{GIG(\lambda,\delta,\gamma)}(u) = \mathfrak{L}_{GIG(\lambda,\delta,\gamma)}(-iu)$.

Corollary 1.10 *Let ϵ_a denote the degenerate distribution (or unit mass) located at $a \in \mathbb{R}$. If $\delta \to \infty$, $\gamma \to \infty$ and $\frac{\delta}{\gamma} \to \sigma \geq 0$, then $GIG(\lambda, \delta, \gamma) \xrightarrow{w} \epsilon_\sigma$.*

Proof: From equation (A.10) it follows that for $\delta, \gamma \to \infty$ we have

$$\mathfrak{L}_{GIG(\lambda,\delta,\gamma)}(u) \sim \left(\frac{\gamma}{\sqrt{\gamma^2+2u}}\right)^{\lambda} \sqrt{\frac{\gamma}{\sqrt{\gamma^2+2u}}}\, e^{\delta\gamma-\delta\gamma\sqrt{1+2u/\gamma^2}} \sim e^{\delta\gamma-\delta\gamma\sqrt{1+2u/\gamma^2}}.$$

Using the Taylor series expansion $\sqrt{1+2x} = 1 + x + o(x)$, $x \to 0$, we conclude

$$\lim_{\substack{\delta,\gamma\to\infty\\ \delta/\gamma\to\sigma}} \mathfrak{L}_{GIG(\lambda,\delta,\gamma)}(u) = \lim_{\substack{\delta,\gamma\to\infty\\ \delta/\gamma\to\sigma}} e^{\delta\gamma-\delta\gamma(1+u/\gamma^2)} = e^{-\sigma u} = \mathfrak{L}_{\epsilon_\sigma}(u)$$

which proves the assertion (see Feller 1971, Chapter XIII.1, Theorem 2). □

Because the densities of all $GIG(\lambda, \delta, \gamma)$-distributions with $\gamma > 0$ decay at an exponential rate for $x \to \infty$, they possess moments of arbitrary order, and the moment generating functions are given by

$$M_{GIG(\lambda,\delta,\gamma)}(u) = \int_0^\infty e^{ux}\, d_{GIG(\lambda,\delta,\gamma)}(x)\,\mathrm{d}x = \mathfrak{L}_{GIG(\lambda,\delta,\gamma)}(-u), \quad u \in \left(-\infty, \tfrac{\gamma^2}{2}\right).$$

The rth moments can easily be derived with the same technique and notation used in the proof of Proposition 1.9. If $X \sim GIG(\lambda, \delta, \gamma)$, then

$$\mathrm{E}[X^r] = \int_0^\infty n(\lambda,\delta,\gamma)\, x^r x^{\lambda-1} e^{-\frac{1}{2}(\delta^2 x^{-1}+\gamma^2 x)}\,\mathrm{d}x = \frac{n(\lambda,\delta,\gamma)}{n(\lambda+r,\delta,\gamma)},$$

where $n(\lambda, \delta, \gamma)$ again denotes the norming constant of the corresponding GIG density. Exploiting this relation we get the following expressions:

$$\mathrm{E}[X^r] = \frac{K_{\lambda+r}(\delta\gamma)}{K_\lambda(\delta\gamma)}\left(\frac{\delta}{\gamma}\right)^r, \qquad \text{if } \lambda \in \mathbb{R},\ \delta,\gamma > 0,$$

$$\mathrm{E}[X^r] = \begin{cases} \dfrac{\Gamma(\lambda+r)}{\Gamma(\lambda)}\left(\dfrac{2}{\gamma^2}\right)^r, & \text{if } r > -\lambda \\ \infty, & \text{if } r \leq -\lambda \end{cases} \qquad \text{and } \lambda,\gamma > 0,\ \delta = 0,$$

$$\mathrm{E}[X^r] = \begin{cases} \dfrac{\Gamma(-\lambda-r)}{\Gamma(-\lambda)}\left(\dfrac{\delta^2}{2}\right)^r, & \text{if } r < -\lambda \\ \infty, & \text{if } r \geq -\lambda \end{cases} \qquad \text{and } \lambda < 0, \delta > 0, \gamma = 0.$$

We close this section with an examination of convolution formulas for GIG distributions which by Lemma 1.6 b) transfer to the corresponding normal mean-variance mixtures derived from them.

Proposition 1.11 *Within the class of GIG distributions the following convolution properties hold:*

a) $GIG\left(-\tfrac{1}{2}, \delta_1, \gamma\right) * GIG\left(-\tfrac{1}{2}, \delta_2, \gamma\right) = GIG\left(-\tfrac{1}{2}, \delta_1+\delta_2, \gamma\right)$,

b) $GIG\left(-\tfrac{1}{2}, \delta_1, \gamma\right) * GIG\left(\tfrac{1}{2}, \delta_2, \gamma\right) = GIG\left(\tfrac{1}{2}, \delta_1+\delta_2, \gamma\right)$,

c) $GIG(-\lambda, \delta, \gamma) * GIG(\lambda, 0, \gamma) = GIG(\lambda, \delta, \gamma)$, $\quad \lambda > 0$,

d) $GIG(\lambda_1, 0, \gamma) * GIG(\lambda_2, 0, \gamma) = GIG(\lambda_1+\lambda_2, 0, \gamma)$, $\quad \lambda_1, \lambda_2 > 0$.

Proof: We use the Laplace transforms derived in Proposition 1.9 and the fact that $\mathfrak{L}_{G_1}(u)\mathfrak{L}_{G_2}(u) = \mathfrak{L}_G(u)$ implies $G_1 * G_2 = G$.

a) In case of the inverse Gaussian distribution the Laplace transforms simplify considerably using equation (A.7):

$$\mathfrak{L}_{GIG(-\frac{1}{2},\delta,\gamma)}(u) = \left(\frac{\gamma}{\sqrt{\gamma^2+2u}}\right)^{-\frac{1}{2}} \frac{K_{-\frac{1}{2}}\big(\delta\sqrt{\gamma^2+2u}\big)}{K_{-\frac{1}{2}}(\delta\gamma)} = e^{\delta\gamma-\delta\sqrt{\gamma^2+2u}}.$$

Thus

$$\begin{aligned}\mathfrak{L}_{GIG(-\frac{1}{2},\delta_1,\gamma)}(u)\mathfrak{L}_{GIG(-\frac{1}{2},\delta_2,\gamma)}(u) &= e^{\delta_1(\gamma-\sqrt{\gamma^2+2u})}\, e^{\delta_2(\gamma-\sqrt{\gamma^2+2u})}\\ &= e^{(\delta_1+\delta_2)(\gamma-\sqrt{\gamma^2+2u})} = \mathfrak{L}_{GIG(-\frac{1}{2},\delta_1+\delta_2,\gamma)}(u).\end{aligned}$$

b) Again with equation (A.7) it follows

$$\mathfrak{L}_{GIG(\frac{1}{2},\delta,\gamma)}(u) = \frac{\gamma}{\sqrt{\gamma^2+2u}}\, e^{\delta\gamma-\delta\sqrt{\gamma^2+2u}}$$

and hence

$$\begin{aligned}\mathfrak{L}_{GIG(-\frac{1}{2},\delta_1,\gamma)}(u)\mathfrak{L}_{GIG(\frac{1}{2},\delta_2,\gamma)}(u) &= \frac{\gamma}{\sqrt{\gamma^2+2u}}\, e^{(\delta_1+\delta_2)(\gamma-\sqrt{\gamma^2+2u})}\\ &= \mathfrak{L}_{GIG(\frac{1}{2},\delta_1+\delta_2,\gamma)}(u).\end{aligned}$$

c) Using $K_{-\lambda}(x) = K_\lambda(x)$ (see equation (A.2)) we calculate

$$\begin{aligned}&\mathfrak{L}_{GIG(-\lambda,\delta,\gamma)}(u)\mathfrak{L}_{GIG(\lambda,0,\gamma)}(u) =\\ &\qquad= \left(\frac{\gamma}{\sqrt{\gamma^2+2u}}\right)^{-\lambda} \frac{K_{-\lambda}\big(\delta\sqrt{\gamma^2+2u}\big)}{K_{-\lambda}(\delta\gamma)} \left(1+\frac{2u}{\gamma^2}\right)^{-\lambda}\\ &\qquad= \left(\frac{\gamma}{\sqrt{\gamma^2+2u}}\,\frac{\gamma^2+2u}{\gamma^2}\right)^{-\lambda} \frac{K_\lambda\big(\delta\sqrt{\gamma^2+2u}\big)}{K_\lambda(\delta\gamma)}\\ &\qquad= \left(\frac{\gamma}{\sqrt{\gamma^2+2u}}\right)^{\lambda} \frac{K_\lambda\big(\delta\sqrt{\gamma^2+2u}\big)}{K_\lambda(\delta\gamma)} = \mathfrak{L}_{GIG(\lambda,\delta,\gamma)}(u).\end{aligned}$$

d) It is easily seen from Proposition 1.9 that

$$\mathfrak{L}_{GIG(\lambda_1,0,\gamma)}(u)\mathfrak{L}_{GIG(\lambda_2,0,\gamma)}(u) = \left(1+\frac{2u}{\gamma^2}\right)^{-(\lambda_1+\lambda_2)} = \mathfrak{L}_{GIG(\lambda_1+\lambda_2,0,\gamma)}(u).$$

□

Remark: The convolution formulas a) and d) of Proposition 1.11 imply the well-known fact that

$$IG(\delta,\gamma) = *_{i=1}^n\, IG\big(\tfrac{\delta}{n},\gamma\big) \quad\text{and}\quad G\big(\lambda,\tfrac{\gamma^2}{2}\big) = *_{i=1}^n\, G\big(\tfrac{\lambda}{n},\tfrac{\gamma^2}{2}\big),$$

so all inverse Gaussian and Gamma distributions are infinitely divisible according to Definition 1.1. But this property is not restricted to these two subclasses. Actually every GIG distribution is not only infinitely divisible, but even selfdecomposable. This was shown in Barndorff-Nielsen and Halgreen (1977) and Halgreen (1979) and also follows from Propositions 1.20 and 1.23 later on.

1.3 Generalized hyperbolic distributions

Generalized hyperbolic distributions (henceforth GH) have been introduced in Barndorff-Nielsen (1977) in connection with the modeling of aeolian sand deposits and dune movements. They are defined as a normal mean-variance mixture with a GIG mixing distribution as follows:

$$GH(\lambda,\alpha,\beta,\delta,\mu) := N(\mu+\beta y, y)\circ GIG\big(\lambda,\delta,\sqrt{\alpha^2-\beta^2}\big). \tag{1.5}$$

The parameter restrictions for GIG distributions (see p. 8) immediately imply that the GH parameters have to fulfill the constraints

$$\lambda,\mu\in\mathbb{R} \quad\text{and}\quad \begin{array}{ll} \delta\ge 0,\ 0\le|\beta|<\alpha, & \text{if } \lambda>0,\\ \delta>0,\ 0\le|\beta|<\alpha, & \text{if } \lambda=0,\\ \delta>0,\ 0\le|\beta|\le\alpha, & \text{if } \lambda<0. \end{array}$$

As before, parametrizations with $\delta = 0$ and $|\beta| = \alpha$ have to be understood as limiting cases which by Lemma 1.7 equal normal mean-variance mixtures with the corresponding GIG limit distributions. We defer a thorough study of these limits to the next section and assume in the following, if not stated otherwise, that $\delta > 0$ and $|\beta| < \alpha$. Note that by (1.5) and Lemma 1.6 c) all GH distributions (and their limits) inherit the property of infinite divisibility from the GIG distributions. For the Lebesgue densities one obtains

$$\begin{aligned} &d_{GH(\lambda,\alpha,\beta,\delta,\mu)}(x) = \\ &= \int_0^\infty d_{N(\mu+\beta y,y)}(x)\, d_{GIG\left(\lambda,\delta,\sqrt{\alpha^2-\beta^2}\right)}(y)\,\mathrm{d}y \\ &= a(\lambda,\alpha,\beta,\delta,\mu)\big(\delta^2+(x-\mu)^2\big)^{(\lambda-\frac12)/2} K_{\lambda-\frac12}\big(\alpha\sqrt{\delta^2+(x-\mu)^2}\big)\, e^{\beta(x-\mu)} \end{aligned} \tag{1.6}$$

with the norming constant

$$a(\lambda,\alpha,\beta,\delta,\mu) = \frac{\big(\alpha^2-\beta^2\big)^{\frac{\lambda}{2}}}{\sqrt{2\pi}\,\alpha^{\lambda-\frac12}\delta^\lambda K_\lambda\big(\delta\sqrt{\alpha^2-\beta^2}\,\big)}. \tag{1.7}$$

(A more detailed derivation will be given in Chapter 2 for the multivariate case.) A closer look at the densities reveals that the influence of the parameters is as follows: α determines the shape, β the skewness, μ is a location parameter and δ serves for scaling. λ characterizes certain subclasses and considerably influences the size of mass contained in the tails. See also Figure 1.3 for an illustration. α and β can be replaced by the alternative parameters

$$\rho := \frac{\beta}{\alpha},\quad \zeta := \delta\sqrt{\alpha^2-\beta^2} \qquad\text{or}\qquad \chi := \rho\xi,\quad \xi := \frac{1}{\sqrt{1+\zeta}}. \tag{1.8}$$

From Theorem 2.11 c) in Chapter 2 it follows that if $X \sim GH(\lambda,\alpha,\beta,\delta,\mu)$, then $\tilde{X} = aX + b \sim GH\big(\lambda, \frac{\alpha}{|a|}, \frac{\beta}{a}, \delta|a|, a\mu+b\big)$, hence for $a > 0$ the last two parametrizations are scale- and location-invariant. Moreover, the above mentioned parameter restrictions imply $0 \le |\chi| \le \xi \le 1$, so all possible values for χ and ξ lie within a triangle with upper corners $(-1, 1), (1, 1)$ and lower corner $(0, 0)$, the so-called shape triangle.

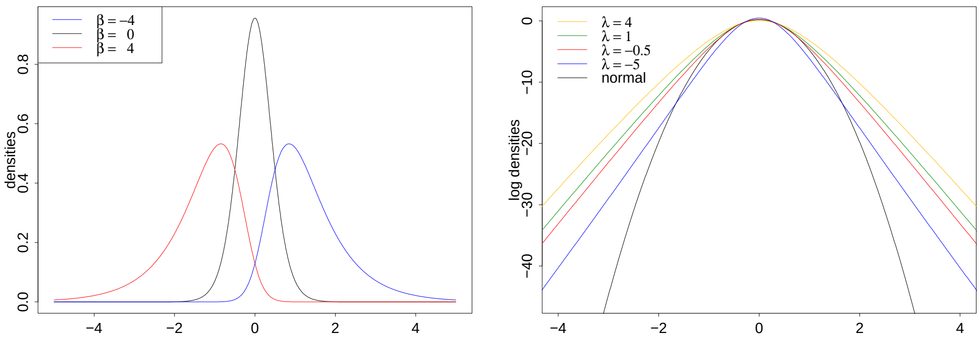


Figure 1.3: Influence of the GH parameters β (left) and λ (right), where on the right hand side log densities are plotted.

Let us mention two important subclasses of GH distributions which will also be used in the calibration procedures described in Chapter 3. For $\lambda = 1$ one obtains the class of *hyperbolic distributions* (HYP) whose densities have, combining equations (1.6) and (A.7), a much simpler form:

$$d_{HYP(\alpha,\beta,\delta,\mu)}(x) = \frac{\sqrt{\alpha^2 - \beta^2}}{2\alpha\delta K_1\big(\delta\sqrt{\alpha^2 - \beta^2}\big)}\, e^{-\alpha\sqrt{\delta^2+(x-\mu)^2}+\beta(x-\mu)}.$$

Its name stems from the fact that the exponent $-\alpha\sqrt{\delta^2 + (x-\mu)^2} + \beta(x-\mu)$ describes a scaled and shifted hyperbola, or, in other words, the graphs of the log densities are hyperbolas with asymptotes having the slopes $\alpha+\beta$ and $-(\alpha-\beta)$. The green log density on the right hand side of Figure 1.3 shows an example with parameters $(\alpha, \beta, \delta, \mu) = (10, 0, 1, 0)$.

Setting $\lambda = -\frac{1}{2}$ leads to the subclass of *normal inverse Gaussian distributions* (NIG). By (1.5) these are the normal mean-variance mixtures arising from inverse Gaussian mixing distributions which explains their name. Using again equations (1.6) and (A.7), its densities are given by

$$d_{NIG(\alpha,\beta,\delta,\mu)}(x) = \frac{\alpha\delta}{\pi}\,\frac{K_1\big(\alpha\sqrt{\delta^2 + (x-\mu)^2}\big)}{\sqrt{\delta^2 + (x-\mu)^2}}\, e^{\delta\sqrt{\alpha^2-\beta^2}+\beta(x-\mu)}.$$

Lemma 1.6 b) and Proposition 1.11 imply the following convolution properties of the GH family:

$$\begin{aligned}
&NIG(\alpha,\beta,\delta_1,\mu_1) * NIG(\alpha,\beta,\delta_2,\mu_2) = NIG(\alpha,\beta,\delta_1+\delta_2,\mu_1+\mu_2),\\
&NIG(\alpha,\beta,\delta_1,\mu_1) * GH\big(\tfrac{1}{2},\alpha,\beta,\delta_2,\mu_2\big) = GH\big(\tfrac{1}{2},\alpha,\beta,\delta_1+\delta_2,\mu_1+\mu_2\big),\\
&GH(-\lambda,\alpha,\beta,\delta,\mu_1) * GH(\lambda,\alpha,\beta,0,\mu_2) = GH(\lambda,\alpha,\beta,\delta,\mu_1+\mu_2),\\
&GH(\lambda_1,\alpha,\beta,0,\mu_1) * GH(\lambda_2,\alpha,\beta,0,\mu_2) = GH(\lambda_1+\lambda_2,\alpha,\beta,0,\mu_1+\mu_2),
\end{aligned} \tag{1.9}$$

where in the last two equations of course $\lambda, \lambda_1, \lambda_2 > 0$.

Remark: Inspecting the Laplace transforms of GIG distributions given in Proposition 1.9 more closely one can deduce that the list of convolution formulas in Proposition 1.11 is complete, that is, no other convolution of two GIG distributions will yield a distribution that itself is contained in the GIG class. Consequently there do not exist more than the four convolution formulas (1.9) for the GH family either, and the NIG subclass is, apart from the limiting distributions with $\delta = 0$, the only one which forms a semigroup under convolution.

From Lemma 1.6 a) and Proposition 1.9 we conclude that all GH distributions (except for some of their limits) possess a moment generating function of the following form:

$$\begin{aligned} M_{GH(\lambda,\alpha,\beta,\delta,\mu)}(u) &= e^{\mu u}\, M_{GIG(\lambda,\delta,\sqrt{\alpha^2-\beta^2})}\big(\tfrac{u^2}{2} + \beta u\big) \\ &= e^{\mu u}\, \mathfrak{L}_{GIG(\lambda,\delta,\sqrt{\alpha^2-\beta^2})}\big(-\tfrac{u^2}{2} - \beta u\big) \qquad (1.10) \\ &= e^{\mu u} \left(\frac{\alpha^2-\beta^2}{\alpha^2-(\beta+u)^2}\right)^{\frac{\lambda}{2}} \frac{K_\lambda\big(\delta\sqrt{\alpha^2-(\beta+u)^2}\big)}{K_\lambda\big(\delta\sqrt{\alpha^2-\beta^2}\big)}, \end{aligned}$$

which is defined for all $u \in (-\alpha-\beta, \alpha-\beta)$.

Remark: The characteristic functions of GH distributions are easily obtained via the relation

$$\phi_{GH(\lambda,\alpha,\beta,\delta,\mu)}(u) = e^{iu\mu}\, \mathfrak{L}_{GIG(\lambda,\delta,\sqrt{\alpha^2-\beta^2})}\big(\tfrac{u^2}{2} - iu\beta\big) = M_{GH(\lambda,\alpha,\beta,\delta,\mu)}(iu)$$

which follows from the proof of Lemma 1.6 b) and equation (1.10) above.

The existence of a moment generating function implies that GH distributions possess moments of arbitrary order which can be obtained by calculating the derivatives of $M_{GH(\lambda,\alpha,\beta,\delta,\mu)}(u)$ at $u = 0$. With the help of equation (A.4) we get the following expressions for mean and variance:

$$\begin{aligned} \mathrm{E}[GH(\lambda,\alpha,\beta,\delta,\mu)] &= \mu + \frac{\beta\delta^2}{\zeta}\frac{K_{\lambda+1}(\zeta)}{K_\lambda(\zeta)}, \\ \mathrm{Var}[GH(\lambda,\alpha,\beta,\delta,\mu)] &= \frac{\delta^2}{\zeta}\frac{K_{\lambda+1}(\zeta)}{K_\lambda(\zeta)} + \beta^2\frac{\delta^4}{\zeta^2}\left(\frac{K_{\lambda+2}(\zeta)}{K_\lambda(\zeta)} - \frac{K^2_{\lambda+1}(\zeta)}{K^2_\lambda(\zeta)}\right). \end{aligned} \qquad (1.11)$$

Skipping the tedious details of the differentiation, we arrive at the following formulas for skewness and kurtosis:

$$\begin{aligned} \gamma_1(GH) = \mathrm{Var}[GH]^{-\frac{3}{2}} \Bigg[&\frac{\beta^3\delta^6}{\zeta^3}\left(\frac{K_{\lambda+3}(\zeta)}{K_\lambda(\zeta)} - \frac{3K_{\lambda+2}(\zeta)K_{\lambda+1}(\zeta)}{K^2_\lambda(\zeta)} + \frac{2K^3_{\lambda+1}(\zeta)}{K^3_\lambda(\zeta)}\right) \\ &+ \frac{3\beta\delta^4}{\zeta^2}\left(\frac{K_{\lambda+2}(\zeta)}{K_\lambda(\zeta)} - \frac{K^2_{\lambda+1}(\zeta)}{K^2_\lambda(\zeta)}\right)\Bigg] \end{aligned}$$

$$\begin{aligned} \gamma_2(GH) &= -3 + \mathrm{Var}[GH]^{-2} \cdot \\ &\cdot \Bigg[\frac{\delta^8\beta^4}{\zeta^4}\left(\frac{K_{\lambda+4}(\zeta)}{K_\lambda(\zeta)} - \frac{4K_{\lambda+3}(\zeta)K_{\lambda+1}(\zeta)}{K^2_\lambda(\zeta)} + \frac{6K_{\lambda+2}(\zeta)K^2_{\lambda+1}(\zeta)}{K^3_\lambda(\zeta)} - \frac{3K^4_{\lambda+1}(\zeta)}{K^4_\lambda(\zeta)}\right) \\ &+ \frac{\delta^6\beta^2}{\zeta^3}\left(\frac{6K_{\lambda+3}(\zeta)}{K_\lambda(\zeta)} - \frac{12K_{\lambda+2}(\zeta)K_{\lambda+1}(\zeta)}{K^2_\lambda(\zeta)} + \frac{6K^3_{\lambda+1}(\zeta)}{K^3_\lambda(\zeta)}\right) + \frac{3\delta^4}{\zeta^2}\frac{K_{\lambda+2}(\zeta)}{K_\lambda(\zeta)}\Bigg] \end{aligned}$$

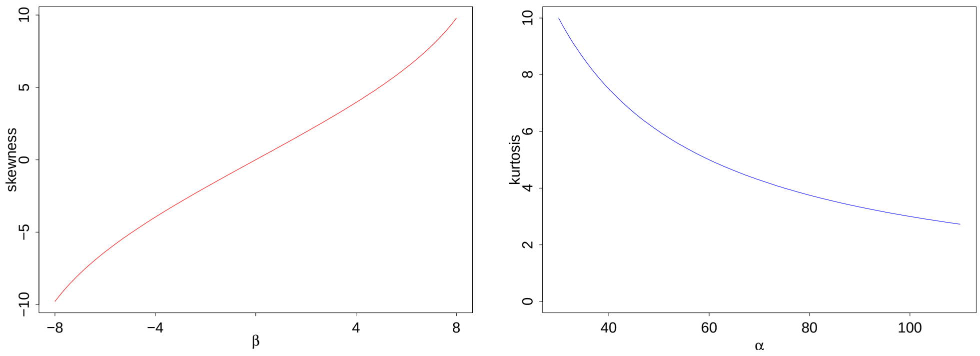


Figure 1.4: Dependence of kurtosis and skewness on GH parameters α and β. Left: skewness of $NIG(10, \beta, 0.01, 0)$. Right: kurtosis of $NIG(\alpha, 0, 0.01, 0)$.

It was already mentioned in Section 1.1 that mixtures of normal distributions are capable of having non-zero skewness and kurtosis. Figure 1.4 shows that for GH distributions the range of attainable values of each quantity can be fairly large. Moreover, it clarifies the role of α as a shape parameter which considerably influences the kurtosis of a GH distribution.

From the existence of a moment generating function one may also conclude that the tails of the GH densities decay at an exponential rate. More precisely, for $|x| \to \infty$ we have $\delta^2 + (x-\mu)^2 \sim x^2$ and by equation (A.10) $K_{\lambda-\frac{1}{2}}\big(\alpha\sqrt{\delta^2+(x-\mu)^2}\big) \sim \sqrt{\frac{\pi}{2\alpha}}\,|x|^{-1/2}\,e^{-\alpha|x|}$, consequently

$$d_{GH(\lambda,\alpha,\beta,\delta,\mu)}(x) \sim c\,|x|^{\lambda-1}\,e^{-\alpha|x|+\beta x}, \quad x \to \pm\infty, \tag{1.12}$$

where $c = \sqrt{\frac{\pi}{2\alpha}}a(\lambda,\alpha,\beta,\delta,\mu)$ and $a(\lambda,\alpha,\beta,\delta,\mu)$ is the norming constant from (1.7). Thus the GH densities have semi-heavy tails in the sense of the following

Definition 1.12 *A probability density f with support $\mathbb{R}$ has* semi-heavy tails *if there exist some constants $a_1, a_2 \in \mathbb{R}$ and $b_1, b_2, c_1, c_2 > 0$ such that*

$$f(x) \sim c_1\,|x|^{a_1}\,e^{-b_1|x|}, \quad x \to -\infty, \quad \text{and} \quad f(x) \sim c_2\,x^{a_2}\,e^{-b_2 x}, \quad x \to +\infty.$$

Remark: From the above definition it can be easily deduced that every probability distribution F having a Lebesgue density f with semi-heavy tails also possesses a moment generating function which is defined at least on the open interval $(-b_1, b_2)$. In case of the GH distributions we have $a_1 = a_2 = \lambda - 1$, $b_1 = \alpha + \beta$, $b_2 = \alpha - \beta$ and $c_1 = c_2 = c$.

A remarkable and probably surprising property of densities with semi-heavy tails is that the tail behaviour of the corresponding distribution functions is the same up to a multiplicative constant, which is shown in the next proposition.

Proposition 1.13 *Let f be a probability density with semi-heavy tails characterized by $a_1, a_2, b_1, b_2, c_1, c_2$, $F(x) := \int_{-\infty}^{x} f(y)\,dy$ be the associated distribution function and $\bar{F}(x) := 1 - F(x)$. Then $f(x) \sim b_1\,F(x)$ as $x \to -\infty$ and $f(x) \sim b_2\,\bar{F}(x)$ as $x \to +\infty$.*

Proof: Let us consider the right tail $\bar{F}(x)$ first. From the assumptions we get, using partial integration,

$$\bar{F}(x) = \int_x^\infty f(y)\,\mathrm{d}y \sim c_2 \int_x^\infty y^{a_2}\, e^{-b_2 y}\,\mathrm{d}y = \frac{c_2}{b_2}\, x^{a_2}\, e^{-b_2 x} + \frac{c_2 a_2}{b_2} \int_x^\infty y^{a_2 - 1}\, e^{-b_2 y}\,\mathrm{d}y.$$

The claim now follows if we can show that $\left(\int_x^\infty y^{a_2-1}\, e^{-b_2 y}\,\mathrm{d}y\right)\left(x^{a_2} e^{-b_2 x}\right)^{-1} \to 0$ as $x \to \infty$. But the latter quotient equals

$$\frac{1}{x}\int_x^\infty \left(\frac{y}{x}\right)^{a_2-1} e^{-b_2(y-x)}\,\mathrm{d}y = \frac{1}{x}\int_0^\infty \left(\frac{y+x}{x}\right)^{a_2-1} e^{-b_2 y}\,\mathrm{d}y$$

and thus converges to zero as $x \to \infty$ because the existence of an integrable majorant ensures that the integral on the right hand side remains bounded. Possible majorants are $g(y) = (y+1)^{a_2-1} e^{-b_2 y}$ if $a_2 > 1$ and $g(y) = e^{-b_2 y}$ if $a_2 \le 1$. Using the change of variables $z = -y$ we see that for $x \to -\infty$

$$F(x) \sim c_1 \int_{-\infty}^x |y|^{a_1}\, e^{-b_1 |y|}\,\mathrm{d}y = c_1 \int_{|x|}^\infty z^{a_1}\, e^{-b_1 z}\,\mathrm{d}z,$$

hence the assertion for the left tail immediately follows from what we have proven above. □

It seems worthwhile to be noticed that distributions with semi-heavy tails form a subclass of $\mathscr{L}_{a,b}$, the class of distributions with exponential tails with rates a and b, which we define as follows:

Definition 1.14 *A distribution function F is said to have exponential tails with rates $a > 0$ and $b > 0$ ($F \in \mathscr{L}_{a,b}$) if for all $y \in \mathbb{R}$*

$$\lim_{x\to-\infty} \frac{F(x-y)}{F(x)} = e^{-ay} \quad \textit{and} \quad \lim_{x\to\infty} \frac{\bar{F}(x-y)}{\bar{F}(x)} = e^{by}.$$

Remark: Most definitions of exponential tails only use one index which characterizes the behaviour of the right tail $\bar{F}(x)$. This is due to the fact that these arose from extreme value theory or more generally actuarial sciences where one typically works with probability distributions on $\mathbb{R}_+$. The above is a natural generalization to distributions having support $\mathbb{R}$ we are concerned with.

The class $\mathscr{L}_{a,b}$ is closely related to the class $\mathscr{R}_p$ of regularly varying functions to be introduced in

Definition 1.15 *A measurable function g is regularly varying with exponent $p \in \mathbb{R}$ ($g \in \mathscr{R}_p$) if $\lim_{t\to\infty} \frac{g(st)}{g(t)} = s^p$ for all $s > 0$.*

Remark: We have $F \in \mathscr{L}_{a,b}$ iff $F(-\ln(x)) \in \mathscr{R}_{-a}$ and $\bar{F}(\ln(x)) \in \mathscr{R}_{-b}$. To see this, put $s = e^y$ and $t = e^{-x}$, then

$$e^{-ay} = \lim_{x\to-\infty} \frac{F(x-y)}{F(x)} \iff s^{-a} = \lim_{t\to\infty} \frac{F(-\ln(t) - \ln(s))}{F(-\ln(t))} = \lim_{t\to\infty} \frac{F(-\ln(st))}{F(-\ln(t))},$$

and the assertion for the right tails follows analogously with $s = e^{-y}$ and $t = e^x$.

Using Definition 1.12 and Proposition 1.13, it is immediately seen that for a probability distribution F possessing a density f with semi-heavy tails we have

$$\lim_{x\to-\infty}\frac{F(x-y)}{F(x)}=\lim_{x\to-\infty}\frac{f(x-y)}{f(x)}=\lim_{x\to-\infty}\left(\frac{|x-y|}{|x|}\right)^{a_1}e^{-b_1(|x-y|-|x|)}=e^{-b_1 y}$$

and an analogous limit is obtained for the right tails, hence $F\in\mathscr{L}_{b_1,b_2}$. For practical purposes, especially in risk management, also the behaviour of convolution tails is of some interest. An easy solution occurs if the factors of the convolution have semi-heavy tails which decay at different rates: the convolution tails are determined by the factor with the heavier left (respectively right) tail. This seems to be a well-known result for distributions on $\mathbb{R}_+$ with exponential tails which is stated, for example, in Cline (1986, Lemma 1), but since we could not find an explicit proof in the literature, we provide one here.

Proposition 1.16 *Let $F_1\in\mathscr{L}_{b_1,b_2}$, $F_2\in\mathscr{L}_{\tilde b_1,\tilde b_2}$ with moment generating functions $M_{F_1}(u)$ and $M_{F_2}(u)$. If $b_1<\tilde b_1$ and $b_2<\tilde b_2$, then $F_1*F_2\in\mathscr{L}_{b_1,b_2}$ and*

$$\lim_{x\to-\infty}\frac{(F_1*F_2)(x)}{F_1(x)}=M_{F_2}(-b_1),\qquad \lim_{x\to\infty}\frac{\overline{(F_1*F_2)}(x)}{\bar F_1(x)}=M_{F_2}(b_2).$$

Proof: Suppose X and Y are independent random variables with distribution functions F_1 and F_2, respectively. For $x\in\mathbb{R}$ and $s>1$ we have

$$\begin{aligned}P(X+Y>x)=P\big(X+Y>x,\,X\le\tfrac{x}{s}\big)&+P\big(X+Y>x,\,Y\le x-\tfrac{x}{s}\big)\\&+P\big(X>\tfrac{x}{s}\big)P\big(Y>x-\tfrac{x}{s}\big).\end{aligned}$$

With the help of Fubini's theorem, the first summand can be written as

$$P\big(X+Y>x,\,X\le\tfrac{x}{s}\big)=\int_x^\infty\int_{-\infty}^{\frac{x}{s}}\mathrm{d}F_2(z-y)\,F_1(\mathrm{d}y)=\int_{-\infty}^{\frac{x}{s}}\bar F_2(x-y)\,F_1(\mathrm{d}y),$$

and the second one can be represented analogously, thus

$$\begin{aligned}\frac{\overline{(F_1*F_2)}(x)}{\bar F_1(x)}=\int_{-\infty}^{\frac{x}{s}}\frac{\bar F_2(x-y)}{\bar F_1(x)}\,F_1(\mathrm{d}y)&+\int_{-\infty}^{x-\frac{x}{s}}\frac{\bar F_1(x-y)}{\bar F_1(x)}\,F_2(\mathrm{d}y)\\&+\frac{\bar F_1\big(\frac{x}{s}\big)\bar F_2\big(x-\frac{x}{s}\big)}{\bar F_1(x)}\,.\end{aligned}\tag{1.13}$$

Since $b_2<\tilde b_2$ by assumption, we can find some $s>1$ such that $\tilde b_2(1-\frac{1}{s})-\frac{2}{s}>b_2$ which is kept fix for the arguments to come. Observing that $\bar F_2(x-y)$ is increasing in y, we have

$$\int_{-\infty}^{\frac{x}{s}}\frac{\bar F_2(x-y)}{\bar F_1(x)}\,F_1(\mathrm{d}y)\le\int_{-\infty}^{\frac{x}{s}}\frac{\bar F_2\big(x-\frac{x}{s}\big)}{\bar F_1(x)}\,F_1(\mathrm{d}y)\le\frac{\bar F_2\big(x-\frac{x}{s}\big)}{\bar F_1(x)}\,.$$

From Definition 1.12 and Proposition 1.13 we conclude that for sufficiently large x_0 and $x\ge x_0$ the inequalities $e^{-(b_2+\frac{1}{s})x}\le\bar F_1(x)\le e^{-(b_2-\frac{1}{s})x}$ and $\bar F_2(x)\le e^{-(\tilde b_2-\frac{1}{s})x}$ hold. Consequently

$$\lim_{x\to\infty}\int_{-\infty}^{\frac{x}{s}}\frac{\bar F_2(x-y)}{\bar F_1(x)}\,F_1(\mathrm{d}y)\le\lim_{x\to\infty}\frac{\bar F_2\big(x-\frac{x}{s}\big)}{\bar F_1(x)}\le\lim_{x\to\infty}e^{-(\tilde b_2-\frac1s)x(1-\frac1s)+(b_2+\frac1s)x}$$
$$=\lim_{x\to\infty}e^{-x\left(\tilde b_2(1-\frac1s)-\frac2s-b_2+\frac{1}{s^2}\right)}=0,$$

because the term in brackets in the last exponent is positive by the above choice of s. To determine the limit behaviour of the second and third summand on the right hand side of (1.13), we first derive an upper bound for $\frac{\bar F_1(z)}{\bar F_1(x)}$. As pointed out before, $F_1\in\mathscr{L}_{b_1,b_2}$ implies that $\bar F_1\big(\ln(x)\big)\in\mathscr{R}_{-b_2}$. By Bingham, Goldie, and Teugels (1987, Potter's Theorem 1.5.6 iii)) there exists some $x_1>1$ such that $\frac{\bar F_1(\ln(\bar z))}{\bar F_1(\ln(\bar x))}\le 2\left(\frac{\bar z}{\bar x}\right)^{-b_2-\frac1s}$ if $\bar x,\bar z\ge x_1$ and $\bar z\le\bar x$. Setting $\bar z:=e^z$, $\bar x:=e^x$ gives

$$\frac{\bar F_1(z)}{\bar F_1(x)}\le 2e^{-(b_2+\frac1s)(z-x)}\quad\text{if } z,x\ge\ln(x_1)\text{ and } z\le x.$$

Therewith we obtain

$$\lim_{x\to\infty}\frac{\bar F_1\big(\frac{x}{s}\big)}{\bar F_1(x)}\,\bar F_2\big(x-\tfrac{x}{s}\big)\le\lim_{x\to\infty}2e^{-(b_2+\frac1s)(\frac{x}{s}-x)-(\tilde b_2-\frac1s)(x-\frac{x}{s})}$$
$$=\lim_{x\to\infty}2e^{-(x-\frac{x}{s})(\tilde b_2-b_2-\frac2s)}=0,$$

hence by (1.13) we have

$$\lim_{x\to\infty}\frac{\overline{(F_1*F_2)}(x)}{\bar F_1(x)}=\lim_{x\to\infty}\int_{-\infty}^{x-\frac{x}{s}}\frac{\bar F_1(x-y)}{\bar F_1(x)}\,F_2(\mathrm{d}y)\,.$$

If $y\le 0$, then $\frac{\bar F_1(x-y)}{\bar F_1(x)}\le 1$ for all x. If $0\le y\le x-\frac{x}{s}$, then $\frac{x}{s}\le x-y\le x$, and the above considerations imply

$$\frac{\bar F_1(x-y)}{\bar F_1(x)}\le 2e^{-(b_2+\frac1s)(x-y-x)}=2e^{(b_2+\frac1s)y}\quad\text{if } x\ge s\ln(x_1).$$

Since the moment generating function $M_{F_2}(u)$ of F_2 is defined on $(-\tilde b_1,\tilde b_2)$ and $0<b_2+\frac1s<\tilde b_2$, the function $e^{(b_2+\frac1s)y}$ is integrable with respect to F_2, and the dominated convergence theorem yields

$$\lim_{x\to\infty}\frac{\overline{(F_1*F_2)}(x)}{\bar F_1(x)}=\int_{-\infty}^{+\infty}\lim_{x\to\infty}\frac{\bar F_1(x-y)}{\bar F_1(x)}\,F_2(\mathrm{d}y)=\int_{-\infty}^{+\infty}e^{b_2y}\,F_2(\mathrm{d}y)=M_{F_2}(b_2).$$

Analogously it can be verified that $\lim_{x\to-\infty}\frac{(F_1*F_2)(x)}{F_1(x)}=M_{F_2}(-b_1)$, and both limit equations together entail $F_1*F_2\in\mathscr{L}_{b_1,b_2}$ which completes the proof. □

Remark: The assumption above that both tails of F_1 are heavier than those of F_2 was just made for notational convenience. As it is easily seen, in general we have $F_1*F_2=\mathscr{L}_{b_1\wedge\tilde b_1,b_2\wedge\tilde b_2}$, that is, one factor may determine the left tail of the convolution and the other one the right tail. Embrechts and Goldie (1980, Theorem 3 b)) have shown that if the right tails of F_1 and F_2 are both exponential with the same rate a, then the right tail of F_1*F_2 is also exponential with rate a, so we may conclude that $F_1 * F_2 = \mathscr{L}_{b_1\wedge\tilde{b}_1, b_2\wedge\tilde{b}_2}$ remains valid if $b_1 = \tilde{b}_1$ and/or $b_2 = \tilde{b}_2$.

Since GH distributions possess densities with semi-heavy tails as pointed out before, an application of Propositions 1.13 and 1.16 to this class yields

Corollary 1.17 *Let F be the distribution function of $GH(\lambda,\alpha,\beta,\delta,\mu)$, then*

$$F(x) \underset{x\to-\infty}{\sim} \frac{c}{\alpha+\beta}\,|x|^{\lambda-1}\,e^{-\alpha|x|+\beta x} \quad \text{and} \quad \bar{F}(x) \underset{x\to\infty}{\sim} \frac{c}{\alpha-\beta}\,x^{\lambda-1}\,e^{-\alpha x+\beta x},$$

where $c = \sqrt{\frac{\pi}{2\alpha}}\,a(\lambda,\alpha,\beta,\delta,\mu)$ and $a(\lambda,\alpha,\beta,\delta,\mu)$ is given by (1.7).
Moreover, $GH(\lambda_1,\alpha_1,\beta_1,\delta_1,\mu_1) * GH(\lambda_2,\alpha_2,\beta_2,\delta_2,\mu_2) \in \mathscr{L}_{b_1,b_2}$ where $b_1 = \min(\alpha_1+\beta_1,\alpha_2+\beta_2)$ and $b_2 = \min(\alpha_1-\beta_1,\alpha_2-\beta_2)$.

1.4 Limits of generalized hyperbolic distributions

This section is devoted to a thorough study of possible weak limits of GH distributions. As was already mentioned in the last section, most of them can be obtained as normal mean-variance mixtures where, according to Lemma 1.7, the mixing distribution is the corresponding GIG limit. However, in the following we shall derive the limit distributions by investigating the pointwise convergence of the GH densities (1.6) instead of calculating lots of mixture integrals for two reasons: firstly, the latter method is often more lengthy and extensive, and in addition it can not capture all possible limiting cases.

To determine the limits of GH densities, we will frequently use some asymptotic properties of the Bessel functions K_λ which are collected in Appendix A, so the careful reader is encouraged to take a look out there before continuing here. As we shall see, some limiting cases occur when one ore more GH parameters tend to certain finite values, whereas for other limits some GH parameters necessarily have to tend to infinity. Because the limit distributions obtained in the first case are much more interesting and will also be used in Chapter 3 for CDO pricing, they will be examined first within the next subsection, thereafter we consider the limits with infinite parameters.

1.4.1 Limits with finite parameters

From the mixture representation (1.5) of GH distributions and the properties of the GIG distributions described on page 9 one can deduce that limiting cases with finite parameters can only be obtained if $\lambda > 0, \delta = 0$ or $\lambda < 0, |\beta| = \alpha$. Indeed, if $\lambda = 0$, then by (A.9) the behaviour of the norming constant (1.7) is

$$\frac{\sqrt{\alpha}}{\sqrt{2\pi}\,K_\lambda\big(\delta\sqrt{\alpha^2-\beta^2}\big)} \sim -\frac{\sqrt{\alpha}}{\sqrt{2\pi}\,\ln\big(\delta\sqrt{\alpha^2-\beta^2}\big)} \to 0 \quad \text{if } \delta\to 0 \ \text{ or } |\beta|\to\alpha,$$

consequently no weak limits exist, and we can concentrate on the case $\lambda \neq 0$.

Positive $\boldsymbol{\lambda}$

By equation (A.8), the asymptotic behaviour of the norming constant (1.7) for $\lambda > 0$ is given by

$$a(\lambda,\alpha,\beta,\delta,\mu) \sim \frac{(\alpha^2-\beta^2)^\lambda}{\sqrt{2\pi}\,\alpha^{\lambda-\frac{1}{2}}2^{\lambda-1}\Gamma(\lambda)} \quad \text{if } \delta \to 0 \ \text{ or } |\beta| \to \alpha,$$

hence a non-degenerate limit can only be obtained for $\delta \to 0$. In this case we have $\sqrt{\delta^2+(x-\mu)^2} \to |x-\mu|$, and inserting into (1.6) yields for $x-\mu \neq 0$

$$\begin{aligned}\lim_{\delta\to 0} d_{GH(\lambda,\alpha,\beta,\delta,\mu)}(x) &= \frac{(\alpha^2-\beta^2)^\lambda}{\sqrt{\pi}\,(2\alpha)^{\lambda-\frac{1}{2}}\Gamma(\lambda)}\,|x-\mu|^{\lambda-\frac{1}{2}}\,K_{\lambda-\frac{1}{2}}(\alpha|x-\mu|)\,e^{\beta(x-\mu)} \\ &=: d_{VG(\lambda,\alpha,\beta,\mu)}(x)\end{aligned} \tag{1.14}$$

(if $\lambda > 0.5$, convergence also holds for $x-\mu = 0$) which equals the density of a *Variance-Gamma distribution* (henceforth VG). This class was introduced in Madan and Seneta (1990) (symmetric case $\beta = \theta = 0$) and Madan, Carr, and Chang (1998) (general case), but with a different parametrization $VG(\sigma,\nu,\theta,\tilde{\mu})$. The latter is obtained by

$$\sigma^2 = \frac{2\lambda}{\alpha^2-\beta^2}, \qquad \nu = \frac{1}{\lambda}, \qquad \theta = \beta\sigma^2 = \frac{2\beta\lambda}{\alpha^2-\beta^2}, \qquad \tilde{\mu} = \mu.$$

Variance-Gamma distributions themselves are a subclass of *CGMY-distributions* introduced in Carr, Geman, Madan, and Yor (2002) which corresponds to the setting $Y = 0$. The other parameters are related as follows:

$$C = \frac{1}{\nu} = \lambda, \quad \frac{G-M}{2} = \frac{\theta}{\sigma} = \beta, \quad \frac{G+M}{2} = \frac{\sqrt{\frac{2}{\nu}+\frac{\theta^2}{\sigma^2}}}{\sigma} = \alpha.$$

If $\lambda = 1$ (hyperbolic limiting case), (1.14) simplifies using (A.7) to

$$d_{VG(1,\alpha,\beta,\mu)}(x) = \frac{\alpha^2-\beta^2}{2\alpha}\,e^{-\alpha|x-\mu|+\beta(x-\mu)}$$

which is the density of a skewed and shifted *Laplace distribution*.

Now we can reformulate the fourth line of the convolution properties (1.9) of GH distributions as follows:

$$VG(\lambda_1,\alpha,\beta,\mu_1) * VG(\lambda_2,\alpha,\beta,\mu_2) = VG(\lambda_1+\lambda_2,\alpha,\beta,\mu_1+\mu_2), \tag{1.15}$$

and by Lemma 1.7 all VG distributions are normal mean-variance mixtures with

$$VG(\lambda,\alpha,\beta,\mu) = N(\mu+\beta y, y) \circ G\big(\lambda, \tfrac{\alpha^2-\beta^2}{2}\big).$$

Lemma 1.6 a) and Proposition 1.9 then imply that all VG distributions possess a moment generating function of the following form:

$$M_{VG(\lambda,\alpha,\beta,\mu)}(u) = e^{\mu u}\mathfrak{L}_{GIG(\lambda,0,\sqrt{\alpha^2-\beta^2})}\big(-\tfrac{u^2}{2}-\beta u\big) = e^{\mu u}\left(\frac{\alpha^2-\beta^2}{\alpha^2-(\beta+u)^2}\right)^\lambda.$$

The characteristic functions are easily obtained via the relation $\phi_{VG(\lambda,\alpha,\beta,\mu)}(u) = M_{VG(\lambda,\alpha,\beta,\mu)}(iu)$ which can be justified analogously as in the GH case. Calculating the derivatives of $M_{VG(\lambda,\alpha,\beta,\mu)}(u)$ at $u = 0$, we get the following expressions for mean and variance of VG distributions:

$$\begin{aligned} \mathrm{E}[VG(\lambda,\alpha,\beta,\mu)] &= \mu + \frac{2\lambda\beta}{\alpha^2-\beta^2}, \\ \mathrm{Var}[VG(\lambda,\alpha,\beta,\mu)] &= \frac{2\lambda}{\alpha^2-\beta^2} + \frac{4\lambda\beta^2}{(\alpha^2-\beta^2)^2}. \end{aligned} \tag{1.16}$$

Some more tedious and lengthy calculations yield the following formulas for skewness and kurtosis:

$$\begin{aligned} \gamma_1(VG) &= \left(\frac{12\lambda\beta}{(\alpha^2-\beta^2)^2} + \frac{16\lambda\beta^3}{(\alpha^2-\beta^2)^3}\right)\cdot \mathrm{Var}[VG]^{-\frac{3}{2}} \\ \gamma_2(VG) &= \frac{12\lambda}{(\alpha^2-\beta^2)^2}\left(\frac{(4\lambda+8)\beta^4}{(\alpha^2-\beta^2)^2} + \frac{(4\lambda+8)\beta^2}{(\alpha^2-\beta^2)} + \lambda + 1\right)\cdot \mathrm{Var}[VG]^{-2} - 3 \end{aligned}$$

Comparing GH and VG densities it is obvious that the latter show an identical behaviour for large arguments:

$$d_{VG(\lambda,\alpha,\beta.\mu)}(x) \sim \tilde{c}\,|x|^{\lambda-1}\,e^{-\alpha|x|+\beta x}, \quad x \to \pm\infty,$$

where $\tilde{c} = \frac{(\alpha^2-\beta^2)^\lambda}{(2\alpha)^{\lambda+\frac{1}{2}}\Gamma(\lambda)}$. Hence also VG distributions possess densities with semi-heavy tails, and the assertions of Corollary 1.17 remain valid also in the VG case (only c has to be replaced by $\tilde{c}$).

Negative $\boldsymbol{\lambda}$

If $\delta \to 0$, then equation (A.8) implies

$$a(\lambda,\alpha,\beta,\delta,\mu) = \frac{(\alpha^2-\beta^2)^{\frac{\lambda}{2}}}{\sqrt{2\pi}\,\alpha^{\lambda-\frac{1}{2}}\,\delta^\lambda\,K_\lambda\big(\delta\sqrt{\alpha^2-\beta^2}\big)} \sim \frac{2^{\lambda+\frac{1}{2}}\,\delta^{-2\lambda}}{\sqrt{\pi}\,\alpha^{\lambda-\frac{1}{2}}\,\Gamma(-\lambda)} \to 0$$

and thus $\lim_{\delta\to 0} d_{GH(\lambda,\alpha,\beta,\delta,\mu)}(x) = 0$ if $\lambda < 0$. In the limiting cases where $|\beta| \to \alpha$ there are two possibilities: $|\beta| = \alpha > 0$ or $\beta = \alpha = 0$. We investigate the latter one first.

If $\alpha, \beta \to 0$, then we conclude from (A.8)

$$\frac{(\alpha^2-\beta^2)^{\frac{\lambda}{2}}\,K_{\lambda-\frac{1}{2}}\big(\alpha\sqrt{\delta^2+(x-\mu)^2}\big)}{\sqrt{2\pi}\,\alpha^{\lambda-\frac{1}{2}}\,\delta^\lambda\,K_\lambda\big(\delta\sqrt{\alpha^2-\beta^2}\big)} \to \frac{\Gamma\big(-\lambda+\frac{1}{2}\big)\big(\delta^2+(x-\mu)^2\big)^{(\lambda-\frac{1}{2})/2}}{\sqrt{\pi}\,\delta^{2\lambda}\,\Gamma(-\lambda)},$$

hence

$$\begin{aligned} \lim_{\alpha,\beta\to 0} d_{GH(\lambda,\alpha,\beta,\delta,\mu)}(x) &= \frac{\Gamma\big(-\lambda+\frac{1}{2}\big)}{\sqrt{\pi}\,\delta^{2\lambda}\,\Gamma(-\lambda)}\,\big(\delta^2+(x-\mu)^2\big)^{\lambda-\frac{1}{2}} \\ &= \frac{\Gamma\big(-\lambda+\frac{1}{2}\big)}{\sqrt{\pi\delta^2}\,\Gamma(-\lambda)}\left(1+\frac{(x-\mu)^2}{\delta^2}\right)^{\lambda-\frac{1}{2}} =: d_{t(\lambda,\delta,\mu)}(x), \end{aligned} \tag{1.17}$$

which is the density of a scaled and shifted *t distribution* with $f = -2\lambda$ degrees of freedom (the usual Student's t-distribution is obtained with $\delta^2 \equiv -2\lambda$). Using $\Gamma(1) = 1$ and $\Gamma\big(\frac{1}{2}\big) = \sqrt{\pi}$, (1.17) reduces in the NIG limiting case ($\lambda = -\frac{1}{2}$) to

$$d_{GH(-\frac{1}{2},0,0,\delta,\mu)}(x) = \frac{\delta}{\pi\big(\delta^2 + (x-\mu)^2\big)},$$

the density of a scaled and shifted *Cauchy distribution*.

Again by Lemma 1.7, all t distributions are normal variance mixtures with

$$t(\lambda,\delta,\mu) = N(\mu, y) \circ iG\big(\lambda, \tfrac{\delta^2}{2}\big).$$

The pointwise convergence of the densities of course implies weak convergence of the corresponding probability measures and hence convergence of the characteristic functions as well. From equation (1.10), the remark thereafter and (A.8), the characteristic function of $t(\lambda,\delta,\mu)$ is found to be

$$\begin{aligned}\phi_{t(\lambda,\delta,\mu)}(u) &= \lim_{\alpha,\beta\to 0} e^{iu\mu}\left(\frac{\alpha^2-\beta^2}{\alpha^2-(\beta+iu)^2}\right)^{\frac{\lambda}{2}} \frac{K_\lambda\big(\delta\sqrt{\alpha^2-(\beta+iu)^2}\big)}{K_\lambda\big(\delta\sqrt{\alpha^2-\beta^2}\big)} \\ &= \left(\frac{2}{\delta}\right)^{\lambda} \frac{2K_\lambda(\delta|u|)}{\Gamma(-\lambda)|u|^\lambda}\, e^{iu\mu}.\end{aligned} \tag{1.18}$$

In the NIG limiting case, (1.18) simplifies using (A.7) and $\Gamma(0.5) = \sqrt{\pi}$ to $\phi_{t(-1/2,\delta,\mu)}(u) = e^{iu\mu-\delta|u|}$, the well-known characteristic function of a Cauchy distribution. Specializing $\delta^2 \equiv -2\lambda = f$ we get the characteristic function of a *Student's t-distribution* with $f > 0$ degrees of freedom:

$$\phi_{t(f,\mu)}(u) = \left(\frac{f}{4}\right)^{\frac{f}{4}} \frac{2K_{f/2}\big(\sqrt{f}\,|u|\big)}{\Gamma\big(\frac{f}{2}\big)}\, |u|^{\frac{f}{2}}\, e^{iu\mu}.$$

If $f = m$ is an odd integer, by equation (A.6) this coincides with the formulas given in Johnson, Kotz, and Balakrishnan (1995, p. 367).

It is immediately seen from (1.17) that the asymptotic behaviour of the densities is given by $d_{t(\lambda,\delta,\mu)}(x) \sim \bar{c}\,|x|^{2\lambda-1}$, $x \to \pm\infty$, consequently rth moments exist only if $r < -2\lambda = f$ or, in other words, t distributions only possess moments of orders smaller than the degrees of freedom. The symmetry of the densities implies $\mathrm{E}[t(\lambda,\delta,\mu)] = \mu$ if $\lambda < -\frac{1}{2}$, and for $\lambda < -1$ the variance can be calculated as follows:

$$\begin{aligned}\mathrm{Var}[t(\lambda,\delta,\mu)] &= \int_{-\infty}^{+\infty} (x-\mu)^2\, d_{t(\lambda,\delta,\mu)}(x)\,\mathrm{d}x \\ &= \int_{-\infty}^{+\infty} \big((x-\mu)^2 + \delta^2 - \delta^2\big) \frac{\Gamma\big(-\lambda+\frac{1}{2}\big)}{\sqrt{\pi}\,\delta^{2\lambda}\,\Gamma(-\lambda)} \big(\delta^2 + (x-\mu)^2\big)^{\lambda-\frac{1}{2}}\,\mathrm{d}x \\ &= \int_{-\infty}^{+\infty} \frac{\Gamma\big(-\lambda+\frac{1}{2}\big)}{\sqrt{\pi}\,\delta^{2\lambda}\,\Gamma(-\lambda)} \big(\delta^2 + (x-\mu)^2\big)^{\lambda+\frac{1}{2}}\,\mathrm{d}x - \delta^2 \\ &= \frac{\Gamma\big(-\lambda+\frac{1}{2}\big)}{\sqrt{\pi}\,\delta^{2\lambda}\,\Gamma(-\lambda)} \frac{\sqrt{\pi}\,\delta^{2\lambda+2}\,\Gamma(-\lambda-1)}{\Gamma\big(-\lambda-\frac{1}{2}\big)} - \delta^2 \\ &= \delta^2\left(\frac{\lambda+\frac{1}{2}}{\lambda+1} - 1\right) = \frac{\delta^2}{-2\lambda-2}\,.\end{aligned} \tag{1.19}$$

If $\delta^2 = -2\lambda = f$, then the above expression becomes $\frac{f}{f-2}$, the familiar formula for the variance of a Student's t-distribution with $f > 2$ degrees of freedom. For skewness and kurtosis we obtain by similar considerations and calculations

$$\gamma_1\big(t(\lambda,\delta,\mu)\big) = 0, \qquad \gamma_2\big(t(\lambda,\delta,\mu)\big) = 3\left(\frac{\lambda+1}{\lambda+2} - 1\right), \quad \lambda < -2.$$

Note that the kurtosis γ_2 is always positive and tends to zero if $\lambda \to -\infty$, reflecting the fact that t distributions become approximately normal if the number of degrees of freedom is increasing.

The above mentioned asymptotic behaviour of the density $d_{t(\lambda,\delta,\mu)}$ implies $d_{t(\lambda,\delta,\mu)}(\pm x) \in \mathscr{R}_{2\lambda-1}$ and $F_{t(\lambda,\delta,\mu)}(-x), \bar{F}_{t(\lambda,\delta,\mu)}(x) \in \mathscr{R}_{2\lambda}$. Applying Bingham, Goldie, and Omey (2006, Theorem 1.1 and the Theorem on p. 54) we get

Corollary 1.18 *Let F_1, F_2 be the distribution functions of $t(\lambda_1,\delta_1,\mu_1)$ and $t(\lambda_2,\delta_2,\mu_2)$ with corresponding densities f_1, f_2, then*

$$\lim_{|x|\to\infty} \frac{(f_1 * f_2)(x)}{f_1(x)+f_2(x)} = 1 \quad \textit{and} \quad \lim_{x\to-\infty} \frac{(F_1 * F_2)(x)}{F_1(x)+F_2(x)} = \lim_{x\to\infty} \frac{\overline{(F_1 * F_2)}(x)}{\bar{F}_1(x)+\bar{F}_2(x)} = 1.$$

Remark: If $\lambda_1 < \lambda_2$, then with the notations of the preceeding corollary we have $f_1(x) = o\big(f_2(x)\big)$ as $|x| \to \infty$ and $F_1(x) = o\big(F_2(x)\big), x \to -\infty$, as well as $\bar{F}_1(x) = o\big(\bar{F}_2(x)\big), x \to \infty$, consequently

$$\lim_{|x|\to\infty} \frac{(f_1 * f_2)(x)}{f_2(x)} = 1 \quad \text{and} \quad \lim_{x\to-\infty} \frac{(F_1 * F_2)(x)}{F_2(x)} = \lim_{x\to\infty} \frac{\overline{(F_1 * F_2)}(x)}{\bar{F}_2(x)} = 1$$

(see also Bingham, Goldie, and Omey 2006, Theorem 2.1). Hence also in this case the tail behaviour of the convolution and the asymptotic behaviour of the convolution density is determined by the factor with the heavier tails.

Now suppose that $|\beta| \to \alpha > 0$, then by equation (A.8) we have

$$a(\lambda,\alpha,\beta,\delta,\mu) = \frac{(\alpha^2-\beta^2)^{\frac{\lambda}{2}}}{\sqrt{2\pi}\,\alpha^{\lambda-\frac12}\,\delta^\lambda\,K_\lambda\big(\delta\sqrt{\alpha^2-\beta^2}\big)} \to \frac{2^{\lambda+\frac12}}{\sqrt{\pi}\,\alpha^{\lambda-\frac12}\,\delta^{2\lambda}\,\Gamma(-\lambda)},$$

and inserting into (1.6) yields

$$\begin{aligned}\lim_{|\beta|\to\alpha>0} d_{GH(\lambda,\alpha,\beta,\delta,\mu)}(x) = {} & \frac{2^{\lambda+\frac12}}{\sqrt{\pi}\,\alpha^{\lambda-\frac12}\,\delta^{2\lambda}\,\Gamma(-\lambda)}\big(\delta^2+(x-\mu)^2\big)^{(\lambda-\frac12)/2}\,\cdot \\ & \cdot K_{\lambda-\frac12}\big(\alpha\sqrt{\delta^2+(x-\mu)^2}\big)\,e^{\pm\alpha(x-\mu)}\end{aligned} \tag{1.20}$$

which is the density of a normal mean-variance mixture $N(\mu\pm\alpha y, y)\circ iG\big(\lambda, \frac{\delta^2}{2}\big)$ by Lemma 1.7. This was called *generalized hyperbolic skew Student t distribution* and applied to financial data in Aas and Haff (2006). Its characteristic function is obtained similarly as before:

$$\begin{aligned}\phi_{GH(\lambda,\alpha,\pm\alpha,\delta,\mu)}(u) &= \lim_{\beta\to\pm\alpha} e^{iu\mu}\left(\frac{\alpha^2-\beta^2}{\alpha^2-(\beta+iu)^2}\right)^{\frac{\lambda}{2}} \frac{K_\lambda\big(\delta\sqrt{\alpha^2-(\beta+iu)^2}\big)}{K_\lambda\big(\delta\sqrt{\alpha^2-\beta^2}\big)} \\ &= \left(\frac{2}{\delta}\right)^\lambda \frac{2K_\lambda\big(\delta\sqrt{u^2\mp 2iu\alpha}\big)}{\Gamma(-\lambda)(u^2\mp 2iu\alpha)^{\frac{\lambda}{2}}}\,e^{iu\mu}.\end{aligned}$$

The tails of the density (1.20) behave completely different for large arguments. If $\beta = \alpha$, then by (A.10) the asymptotic behaviour is as follows:

$$\begin{aligned} d_{GH(\lambda,\alpha,\alpha,\delta,\mu)}(x) &\sim \tilde{c}_1 \, |x|^{\lambda-1} \, e^{-2\alpha|x|}, \quad && x \to -\infty, \\ d_{GH(\lambda,\alpha,\alpha,\delta,\mu)}(x) &\sim \tilde{c}_2 \, |x|^{\lambda-1}, && x \to +\infty, \end{aligned}$$

and the other way round if $\beta = -\alpha$, hence rth moments exist only if $r < -\lambda$. Consider a sequence of random variables $X_n \sim GH(\lambda, \alpha, \beta_n, \delta, \mu)$ with $|\beta_n| < \alpha$ and $\beta_n \overset{n\to\infty}{\longrightarrow} \pm\alpha$, then (1.20) implies $X_n \overset{\mathcal{L}}{\longrightarrow} X \sim GH(\lambda, \alpha, \pm\alpha, \delta, \mu)$ and hence also $X_n^r \overset{\mathcal{L}}{\longrightarrow} X^r$. Further it follows from the convergence of the norming constants $a(\lambda, \alpha, \beta_n, \delta, \mu) \to a(\lambda, \alpha, \pm\alpha, \delta, \mu)$, (1.12) and the above asymptotics that there exist some constants $\hat{c}, x_0 > 0$ such that $d_{GH(\lambda,\alpha,\beta_n,\delta,\mu)}(x) \le \hat{c}\, |x|^{\lambda-1}$ for all n and $|x| > x_0$, consequently the sequence $(X_n^r)_{n\ge1}$ is uniformly integrable if $r < -\lambda$, and therefore we also have $\mathrm{E}[X_n^r] \to \mathrm{E}[X^r]$. Expressions for the rth moments of the limit distributions can thus be obtained by determining the limits of the corresponding formulas for ordinary GH distributions for $|\beta| \to \alpha$ and hence $\zeta = \delta\sqrt{\alpha^2 - \beta^2} \to 0$. If $\lambda < -1$, then by equations (1.11) and (A.8) we get

$$\begin{aligned} &\mathrm{E}[GH(\lambda, \alpha, \pm\alpha, \delta, \mu)] = \\ &\quad = \lim_{\beta\to\pm\alpha} \mu + \frac{\beta\delta^2}{\zeta} \frac{K_{\lambda+1}(\zeta)}{K_\lambda(\zeta)} = \lim_{\beta\to\pm\alpha} \mu + \frac{\beta\delta^2}{\zeta} \frac{\Gamma\big(-(\lambda+1)\big)}{\Gamma(-\lambda)} \left(\frac{\zeta}{2}\right)^{\lambda+1-\lambda} \\ &\quad = \mu \pm \frac{\alpha\delta^2}{2} \frac{\Gamma\big(-(\lambda+1)\big)}{\Gamma(-\lambda)} = \mu \mp \frac{\alpha\delta^2}{2(\lambda+1)}\,. \end{aligned}$$

Similarly, we find for $\lambda < -2$

$$\begin{aligned} &\mathrm{Var}[GH(\lambda, \alpha, \pm\alpha, \delta, \mu)] = \\ &\quad = \lim_{\beta\to\pm\alpha} \left[\frac{\delta^2}{\zeta} \frac{K_{\lambda+1}(\zeta)}{K_\lambda(\zeta)} + \beta^2 \frac{\delta^4}{\zeta^2} \left(\frac{K_{\lambda+2}(\zeta)}{K_\lambda(\zeta)} - \frac{K^2_{\lambda+1}(\zeta)}{K^2_\lambda(\zeta)} \right) \right] \\ &\quad = \lim_{\beta\to\pm\alpha} \Bigg[\frac{\delta^2}{\zeta} \frac{\Gamma\big(-(\lambda+1)\big)}{\Gamma(-\lambda)} \left(\frac{\zeta}{2}\right)^{\lambda+1-\lambda} + \frac{\beta^2\delta^4}{\zeta^2} \Bigg(\frac{\Gamma\big(-(\lambda+2)\big)}{\Gamma(-\lambda)} \left(\frac{\zeta}{2}\right)^{\lambda+2-\lambda} \\ &\qquad\qquad - \frac{\Gamma\big(-(\lambda+1)\big)^2}{\Gamma(-\lambda)^2} \left(\frac{\zeta}{2}\right)^{2(\lambda+1-\lambda)} \Bigg) \Bigg] \\ &\quad = \frac{\delta^2}{-2\lambda-2} + \frac{\alpha^2\delta^4}{4(\lambda+1)} \left(\frac{1}{(\lambda+2)} - \frac{1}{(\lambda+1)} \right), \end{aligned}$$

which is strictly greater than $\mathrm{Var}[t(\lambda, \delta, \mu)]$, but converges to the latter if $\alpha \to 0$. Some analogous, but longer calculations yield the following formulas for skewness and kurtosis:

$$\begin{aligned} \gamma_1\big(GH(\lambda, \alpha, \pm\alpha, \delta, \mu)\big) &= v^{-\frac{3}{2}} \left(\mp \frac{\alpha^3\delta^6}{4(\lambda+3)(\lambda+2)(\lambda+1)^3} \mp \frac{3\alpha\delta^4}{4(\lambda+2)(\lambda+1)^2} \right), \\ \gamma_2\big(GH(\lambda, \alpha, \pm\alpha, \delta, \mu)\big) &= \Bigg(\frac{\delta^8\alpha^4(3\lambda-15)}{16(\lambda+4)(\lambda+3)(\lambda+2)(\lambda+1)^3} + \frac{3\delta^4}{4(\lambda+2)(\lambda+1)} \\ &\qquad + \frac{3\delta^6\alpha^2(\lambda-1)}{4(\lambda+3)(\lambda+2)(\lambda+1)^3} \Bigg) v^{-2} - 3, \end{aligned}$$

where $v = \mathrm{Var}[GH(\lambda, \alpha, \pm\alpha, \delta, \mu)]$ and $\lambda < -4$. From the asymptotic behaviour of the densities we conclude that $GH(\lambda, \alpha, \pm\alpha, \delta, \mu)$-distributions possess one semi-heavy and one regularly varying tail, so the tail behaviour of their convolutions can be determined by combining Corollaries 1.17 and 1.18.

1.4.2 Limits with infinite parameters

Now we turn to the limiting cases arising if α, β or δ tend to infinity. As we shall see, at most two of the three parameters can do so, whereas the third one necessarily has to remain finite to obtain a well-defined weak limit. We first consider the case where $\alpha, \beta \to \infty$. More precisely we assume

$$\beta = \alpha - \frac{\eta^2}{2}, \qquad \alpha \to \infty, \qquad \delta \to 0, \qquad \alpha\delta^2 \to \psi^2, \qquad \text{and } \eta, \psi > 0.$$

By equation (A.10), for sufficiently large α we have

$$\frac{(\delta^2 + (x-\mu)^2)^{(\lambda-\frac{1}{2})/2}}{\sqrt{2\pi}\,\alpha^{\lambda-\frac{1}{2}}\,\delta^\lambda} K_{\lambda-\frac{1}{2}}\big(\alpha\sqrt{\delta^2 + (x-\mu)^2}\big) \sim \frac{(\delta^2 + (x-\mu)^2)^{\frac{\lambda-1}{2}}}{2\,\alpha^\lambda\,\delta^\lambda\, e^{\alpha\sqrt{\delta^2+(x-\mu)^2}}},$$

and the above assumptions on the parameters imply $\sqrt{\delta^2 + (x-\mu)^2} \to |x-\mu|$, $\alpha^{\frac{\lambda}{2}}\delta^\lambda \to \psi^\lambda$, $\delta\sqrt{\alpha^2 - \beta^2} \to \eta\psi$ and

$$\frac{(\alpha^2-\beta^2)^{\frac{\lambda}{2}}}{\alpha^{\frac{\lambda}{2}}} = \frac{\Big(\alpha^2 - \big(\alpha - \frac{\eta^2}{2}\big)^2\Big)^{\frac{\lambda}{2}}}{\alpha^{\frac{\lambda}{2}}} = \left(\eta^2 - \frac{\eta^4}{4\alpha}\right)^{\frac{\lambda}{2}} \to \eta^\lambda.$$

Collecting these results we find

$$\begin{aligned} &d_{GH(\lambda,\alpha,\beta,\delta,\mu)}(x) = \\ &= \frac{(\alpha^2-\beta^2)^{\frac{\lambda}{2}}}{\alpha^{\frac{\lambda}{2}}} \frac{(\delta^2 + (x-\mu)^2)^{(\lambda-\frac{1}{2})/2}}{\sqrt{2\pi}\,\alpha^{\frac{\lambda-1}{2}}\,\delta^\lambda\, K_\lambda\big(\delta\sqrt{\alpha^2-\beta^2}\big)} K_{\lambda-\frac{1}{2}}\big(\alpha\sqrt{\delta^2+(x-\mu)^2}\big)\, e^{\beta(x-\mu)} \\ &\sim \left(\frac{\eta}{\psi}\right)^\lambda \frac{|x-\mu|^{\lambda-1}}{2K_\lambda(\eta\psi)}\, e^{\beta(x-\mu)-\alpha\sqrt{\delta^2+(x-\mu)^2}}. \end{aligned}$$

Comparing the last expression with the GIG densities (1.2) the conjecture of weak convergence to a shifted GIG distribution is obvious. To prove the latter, it remains to show that

$$\begin{aligned} e^{\beta(x-\mu)-\alpha\sqrt{\delta^2+(x-\mu)^2}} &= e^{\left(\alpha-\frac{\eta^2}{2}\right)(x-\mu)-\alpha\sqrt{\delta^2+(x-\mu)^2}} \\ &\to \begin{cases} e^{-\frac{1}{2}\left(\psi^2(x-\mu)^{-1}+\eta^2(x-\mu)\right)}, & x-\mu > 0, \\ 0, & x-\mu < 0. \end{cases} \end{aligned}$$

From the Taylor series expansion $\sqrt{1+x^2} = 1 + \frac{x^2}{2} + o(x^2)$, $x \to 0$, we conclude

$$\sqrt{\delta^2 + (x-\mu)^2} = |x-\mu|\sqrt{1 + \left(\frac{\delta}{x-\mu}\right)^2} \sim |x-\mu|\left(1 + \frac{\delta^2}{2(x-\mu)^2} + o\big(\delta^2\big)\right),$$

consequently if $x-\mu>0$, then

$$\left(\alpha-\frac{\eta^2}{2}\right)(x-\mu)-\alpha\sqrt{\delta^2+(x-\mu)^2}\to-\frac{1}{2}\left(\frac{\psi^2}{x-\mu}-\eta^2(x-\mu)\right),$$

and if $x-\mu<0$,

$$\left(\alpha-\frac{\eta^2}{2}\right)(x-\mu)-\alpha\sqrt{\delta^2+(x-\mu)^2}\sim-2\alpha|x-\mu|\to-\infty,$$

thus $d_{GH(\lambda,\alpha,\beta,\delta,\mu)}(x)\to d_{GIG(\lambda,\psi,\eta)}(x-\mu)$ pointwise for all $x\in\mathbb{R}$ under the above assumptions on α, β and δ. Setting $\beta=-\alpha+\frac{\eta^2}{2}$, analogous calculations yield $d_{GH(\lambda,\alpha,\beta,\delta,\mu)}(x)\to d_{GIG(\lambda,\psi,\eta)}\big(-(x-\mu)\big)$, so in this case the GH distributions converge weakly to a shifted GIG distribution on $\mathbb{R}_-$. If in addition we let $\eta\to0$ or $\psi\to0$, we may also obtain Gamma and inverse Gamma distributions as possible limits.

If δ tends to infinity instead of β and we further assume

$$\alpha\to\infty,\qquad\delta\to\infty,\qquad\frac{\delta}{\alpha}\to\sigma^2,$$

then $GIG\big(\lambda,\delta,\sqrt{\alpha^2-\beta^2}\big)\xrightarrow{w}\epsilon_{\sigma^2}$ by Corollary 1.10, and Lemma 1.7 entails

$$GH(\lambda,\alpha,\beta,\delta,\mu)\xrightarrow{w}N(\mu+\beta y,y)\circ\epsilon_{\sigma^2}=N\big(\mu+\beta\sigma^2,\sigma^2\big).\tag{1.21}$$

This is probably the easiest way to prove weak convergence to the normal distribution, but since we announced to show pointwise convergence of the densities before, we also do this here for the sake of completeness. Again by equation (A.10), we have for α,δ large enough similar as before

$$\begin{aligned}&\frac{(\alpha^2-\beta^2)^{\frac{\lambda}{2}}}{\alpha^{\lambda-\frac12}\,\delta^\lambda\,K_\lambda\big(\delta\sqrt{\alpha^2-\beta^2}\big)}\big(\delta^2+(x-\mu)^2\big)^{(\lambda-\frac12)/2}K_{\lambda-\frac12}\big(\alpha\sqrt{\delta^2+(x-\mu)^2}\big)\\&\quad\sim\frac{(\alpha^2-\beta^2)^{\frac{\lambda}{2}}}{\alpha^\lambda}\,\frac{(\delta^2+(x-\mu)^2)^{\frac{\lambda}{2}}}{\delta^\lambda}\,\frac{\sqrt{\delta}(\alpha^2-\beta^2)^{\frac14}}{\sqrt{\delta^2+(x-\mu)^2}}\,e^{-\alpha\sqrt{\delta^2+(x-\mu)^2}+\delta\sqrt{\alpha^2-\beta^2}}.\end{aligned}$$

The above assumptions on α and δ imply

$$\begin{gathered}\frac{(\alpha^2-\beta^2)^{\frac{\lambda}{2}}}{\alpha^\lambda}\to1,\qquad\frac{(\delta^2+(x-\mu)^2)^{\frac{\lambda}{2}}}{\delta^\lambda}\to1,\\\frac{\sqrt{\delta}\,(\alpha^2-\beta^2)^{\frac14}}{\sqrt{\delta^2+(x-\mu)^2}}=\Big(\frac{\alpha}{\delta}\Big)^{\frac12}\frac{\big(1-\frac{\beta^2}{\alpha^2}\big)^{\frac14}}{\sqrt{1+\big(\frac{x-\mu}{\delta}\big)^2}}\to\frac{1}{\sqrt{\sigma^2}},\end{gathered}$$

and together it follows

$$d_{GH(\lambda,\alpha,\beta,\delta,\mu)}(x)\sim\frac{1}{\sqrt{2\pi\sigma^2}}\,e^{-\alpha\sqrt{\delta^2+(x-\mu)^2}+\delta\sqrt{\alpha^2-\beta^2}+\beta(x-\mu)}.$$

Again from the Taylor series expansion $\sqrt{1 \pm x^2} = 1 \pm \frac{x^2}{2} + o\big(x^2\big)$, $x \to 0$, the convergence of the above exponent is found to be

$$
\begin{aligned}
&-\alpha\sqrt{\delta^2 + (x-\mu)^2} + \delta\sqrt{\alpha^2 - \beta^2} + \beta(x-\mu) \\
&\quad = -\alpha\delta\sqrt{1 + \frac{(x-\mu)^2}{\delta^2}} + \alpha\delta\sqrt{1 - \frac{\beta^2}{\alpha^2}} + \beta(x-\mu) \\
&\quad = -\alpha\delta\left[1 + \frac{1}{2}\,\frac{(x-\mu)^2}{\delta^2} + o\left(\frac{1}{\delta^2}\right)\right] + \alpha\delta\left[1 - \frac{1}{2}\,\frac{\beta^2}{\alpha^2} + o\left(\frac{1}{\alpha^2}\right)\right] + \beta(x-\mu) \\
&\quad \to -\frac{1}{2}\,\frac{(x-\mu)^2}{\sigma^2} - \frac{\sigma^2\beta^2}{2} - \beta(x-\mu) = -\frac{1}{2\sigma^2}\big(x - (\mu + \beta\sigma^2)\big)^2
\end{aligned}
$$

which shows $d_{GH(\lambda,\alpha,\beta,\delta,\mu)}(x) \to d_{N(\mu+\beta\sigma^2,\sigma^2)}(x)$ pointwise for all $x \in \mathbb{R}$.

Remark: Note that a necessary condition for normal convergence is that α and δ grow with the same rate. If δ is growing much slower than α or even kept fixed instead, then $\sigma^2 = 0$, and from (1.21) we conclude $GH(\lambda,\alpha,\beta,\delta,\mu) \xrightarrow{w} \epsilon_\mu$. This can alternatively be deduced from

$$
\lim_{\alpha\to\infty} \phi_{GH(\lambda,\alpha,\beta,\delta,\mu)}(u) = \lim_{\alpha\to\infty} M_{GH(\lambda,\alpha,\beta,\delta,\mu)}(iu) = e^{iu\mu}
$$

which immediately follows from equation (1.10).

Because we always have to have $|\beta| \le \alpha$, β cannot tend to infinity if α remains bounded, and if $\delta \to \infty$ while α, β are kept fixed, then equation (A.10) implies

$$
\frac{(\delta^2 + (x-\mu)^2)^{(\lambda-\frac{1}{2})/2} K_{\lambda-\frac{1}{2}}\big(\alpha\sqrt{\delta^2 + (x-\mu)^2}\big)}{\delta^\lambda\, K_\lambda\big(\delta\sqrt{\alpha^2-\beta^2}\big)} \sim \frac{e^{-\delta\alpha+\delta\sqrt{\alpha^2-\beta^2}}}{\sqrt{\alpha\delta}\,(\alpha^2-\beta^2)^{-1/4}} \to 0,
$$

hence in the latter case the GH densities degenerate. Therefore the considerations in this section are sufficiently general and cover all possible limiting cases.

1.5 Generalized and extended generalized Γ-convolutions

In this section we give a short introduction to the families of generalized and extended generalized Γ-convolutions. They provide a unified framework which allows an easy derivation of many important properties of GH and GIG distributions. Our presentation follows Thorin (1977a), Thorin (1977b) and Thorin (1978). A thorough investigation of generalized Γ-convolutions with many further examples can be found in the book of Bondesson (1992).

Gamma distributions $G(\lambda,\sigma)$ have already been encountered in Section 1.2 as possible limits of GIG distributions (see page 9). Here we are concerned with the slightly more general case of right-shifted Gamma distributions $G(\lambda,\sigma,a)$ which are defined as follows: If $X \sim G(\lambda,\sigma)$ and $Y = X + a$ with $a \ge 0$, then $G(\lambda,\sigma,a) := \mathcal{L}(Y) = G(\lambda,\sigma) * \epsilon_a$. From Proposition 1.11 d) we have $G(\lambda_1,\sigma,a_1) * G(\lambda_2,\sigma,a_2) = G(\lambda_1+\lambda_2,\sigma,a_1+a_2)$ which implies the infinite divisibility of $G(\lambda, \sigma, a)$. By Proposition 1.9 and the remark thereafter the characteristic functions are given by

$$\phi_{G(\lambda,\sigma,a)}(u) = \mathfrak{L}_{G(\lambda,\sigma)}(iu)\, e^{iua} = \left(1 - \frac{iu}{\sigma}\right)^{-\lambda} e^{iua}$$

which can also be represented in the following form:

$$\phi_{G(\lambda,\sigma,a)}(u) = \exp\left[iua - \int_0^\infty \ln\left(1 - \frac{iu}{y}\right) \lambda\epsilon_\sigma(\mathrm{d}y)\right], \tag{1.22}$$

where $\ln(z)$ denotes the main branch of the complex logarithm. If more generally

$$U_n(x) = \sum_{i=1}^n \lambda_i \mathbb{1}_{[\sigma_i,\infty)}(x), \quad \lambda_i > 0,\, 1 \le i \le n, \quad 0 < \sigma_1 < \sigma_2 < \cdots < \sigma_n < \infty,$$

$$\text{and} \quad 0 \le a = \sum_{i=1}^n a_i, \quad a_i \ge 0,\, 1 \le i \le n,$$

then it follows from (1.22) and the above mentioned convolution property that

$$\phi_{\overset{n}{\underset{i=1}{*}} G(\lambda_i,\sigma_i,a_i)}(u) = \exp\left[iua - \int_0^\infty \ln\left(1 - \frac{iu}{y}\right) U_n(\mathrm{d}y)\right].$$

These considerations lead to the following

Definition 1.19 *The class* Γ_0 *of* generalized Γ-convolutions *consists of all probability distributions* G *on* $(\mathbb{R}_+, \mathcal{B}_+)$ *whose characteristic functions* ϕ_G *can be represented in the following form:*

$$\phi_G(u) = \exp\left[iua - \int_0^\infty \ln\left(1 - \frac{iu}{y}\right) U(\mathrm{d}y)\right], \tag{1.23}$$

$$\left.\begin{aligned} &a \ge 0, \quad U : \mathbb{R}_+ \to \mathbb{R}_+ \textit{ non-decreasing with } U(0) = 0, \\ &\int_0^1 |\ln(y)|\, U(\mathrm{d}y) < \infty, \\ &\int_1^\infty \frac{1}{y}\, U(\mathrm{d}y) < \infty. \end{aligned}\right\} \tag{1.24}$$

The last two conditions of (1.24) ensure the finiteness of the integral term in (1.23) such that $|\phi_G(u)| > 0$ for all $u \in \mathbb{R}$. The above definition suggests that every distribution given by (1.23) and (1.24) is infinitely divisible. In fact, this holds true as we will prove below by showing that the characteristic functions of all elements $G \in \Gamma_0$ also possess a *Khintchine representation*

$$\phi_G(u) = \exp\left(ia_G u + \int_{-\infty}^{+\infty} \left(e^{iux} - 1 - \frac{iux}{1+x^2}\right) \frac{1+x^2}{x^2}\, \psi_G(\mathrm{d}x)\right).$$

This is an alternative representation of infinitely divisible distributions (see, for example, Loève (1977, pp. 310–313 and 344)) which for technical reasons is sometimes more convenient than the equivalent Lévy–Khintchine formula

$$\phi_G(u) = \exp\left(iub_G - \frac{1}{2}c_G u^2 + \int_{\mathbb{R}} \left(e^{iux} - 1 - iux\mathbb{1}_{[-1,1]}(x)\right)\nu_G(\mathrm{d}x)\right).$$

Both formulas can be transformed into each other via the relations

$$\begin{aligned} b_G &= a_G + \int_{\mathbb{R}} \big(x\mathbb{1}_{[-1,1]}(x) - x^{-1}1_{\mathbb{R}\setminus[-1,1]}(x)\big)\,\psi_G(\mathrm{d}x), \\ c_G &= \Delta\psi_G(0) = \psi_G(0) - \psi_G(0-), \\ \nu_G(\mathrm{d}x) &= \frac{1+x^2}{x^2}\,\psi_G(\mathrm{d}x), \end{aligned} \tag{1.25}$$

and since b_G, c_G and ν_G are uniquely determined, so are a_G and ψ_G.

Remark: Wolfe (1971) has shown that for every infinitely divisible distribution F with Lévy measure ν_F one has the equivalence $\int_{\mathbb{R}} |x|^r\,F(\mathrm{d}x) < \infty \Leftrightarrow \int_{\mathbb{R}\setminus[-1,1]} |x|^r\,\nu_F(\mathrm{d}x) < \infty$. Thus if the distribution G has finite first moments, we can omit the truncation function within the integral of the Lévy–Khintchine formula, and the first equation of (1.25) simplifies to

$$b_G = a_G + \int_{-\infty}^{+\infty} x\,\psi_G(\mathrm{d}x) = \mathrm{E}[G].$$

Proposition 1.20 *A generalized Γ-convolution is uniquely determined by the pair (a,U) defined in equations (1.23) and (1.24).*
Moreover, every $G \in \Gamma_0$ is selfdecomposable, and the characteristic pair (a_G,ψ_G) of its Khintchine representation is given by

$$a_G = a + \int_0^\infty\!\!\int_0^\infty \frac{e^{-yx}}{1+x^2}\,\mathrm{d}x\,U(\mathrm{d}y),$$

$$\psi_G(x) = \begin{cases} 0, & x \le 0, \\ \displaystyle\int_0^x \psi_G'(y)\,\mathrm{d}y, & x > 0, \end{cases} \qquad \psi_G'(y) = \frac{y}{1+y^2}\int_0^\infty e^{-yt}\,U(\mathrm{d}t).$$

Proof: Inserting the above expressions for a_G and ψ_G into the Khintchine formula yields

$$\begin{aligned} \ln\big(\phi_G(u)\big) &= iua + iu\int_0^\infty\!\!\int_0^\infty \frac{e^{-yx}}{1+x^2}\,\mathrm{d}x\,U(\mathrm{d}y) \\ &\quad + \int_0^\infty \left(e^{iux} - 1 - \frac{iux}{1+x^2}\right)\frac{1+x^2}{x^2}\,\frac{x}{1+x^2}\int_0^\infty e^{-xt}\,U(\mathrm{d}t)\,\mathrm{d}x \\ &= iua + iu\int_0^\infty\!\!\int_0^\infty \frac{e^{-yx}}{1+x^2}\,\mathrm{d}x\,U(\mathrm{d}y) - iu\int_0^\infty\!\!\int_0^\infty \frac{e^{-xt}}{1+x^2}\,\mathrm{d}x\,U(\mathrm{d}t) \\ &\quad + \int_0^\infty\!\!\int_0^\infty \frac{e^{iux}-1}{x}\,e^{-xt}\,U(\mathrm{d}t)\,\mathrm{d}x \\ &= iua - \int_0^\infty\!\!\int_0^\infty \frac{e^{-xt} - e^{-x(t-iu)}}{x}\,\mathrm{d}x\,U(\mathrm{d}t) \\ &= iua - \int_0^\infty\!\!\int_0^\infty\!\!\int_t^{t-iu} e^{-xz}\,\mathrm{d}z\,\mathrm{d}x\,U(\mathrm{d}t) \\ &= iua - \int_0^\infty\!\!\int_t^{t-iu} \frac{1}{z}\,\mathrm{d}z\,U(\mathrm{d}t) = iua - \int_0^\infty \ln(t-iu) - \ln(t)\,U(\mathrm{d}t) \\ &= iua - \int_0^\infty \ln\left(1 - \frac{iu}{t}\right)U(\mathrm{d}t) \end{aligned}$$

which equals the exponent on the right hand side of (1.23) as desired. Since the characteristic pair (a_G, ψ_G) of the Khintchine representation is uniquely determined, the uniqueness of (a, U) immediately follows from the formulas verified above (observe that the density of the Khintchine measure is $\psi_G'(y) = \frac{y}{1+y^2}\,\mathfrak{L}_U(y)$, so the Laplace transform $\mathfrak{L}_U(y)$ and thus U are uniquely characterized, hence so is a). Together with the last equation of (1.25) we have

$$\nu_G(\mathrm{d}x) = \frac{1+x^2}{x^2}\,\psi_G(\mathrm{d}x) = \mathbb{1}_{(0,\infty)}(x)\,\frac{1+x^2}{x^2}\,\frac{x}{1+x^2}\int_0^\infty e^{-xt}\,U(\mathrm{d}t)\,\mathrm{d}x,$$

thus the Lévy measure of every generalized Γ-convolution has a density of the form $\frac{k(x)}{|x|}$ where $k(x) = \mathbb{1}_{(0,\infty)}(x)\int_0^\infty e^{-xt}\,U(\mathrm{d}t)$ is non-decreasing on $(-\infty, 0)$ and non-increasing (actually decreasing) on $(0,\infty)$, so all generalized Γ-convolutions are selfdecomposable by Lemma 1.4. □

The class Γ_0 is closed under weak limits (see Theorem 1.22 below). If it is enlarged by permitting translations to the left and thus canceling the condition $a \geq 0$ in (1.24), one obtains a class $\Gamma_{-\infty}$ which is not closed under passages to the limit: Take for example $a_n = -n$ and $U_n(x) = n^2\,\mathbb{1}_{[n,\infty)}(x)$, then we have for n sufficiently large

$$\ln\big(\phi_n(u)\big) = -iun - n^2\ln\left(1 - \frac{iu}{n}\right) = -iun + iun + \frac{(iu)^2}{2} + n^2 o(n^2) \underset{n\to\infty}{\longrightarrow} -\frac{u^2}{2},$$

hence normal distributions are in the closure of $\Gamma_{-\infty}$, but obviously not in $\Gamma_{-\infty}$ itself. Also note that the weaker condition $a \in \mathbb{R}$ allows, according to (1.24), the decomposition

$$a = \tilde{a} - \int_0^\infty \frac{y}{1+y^2}\,U(\mathrm{d}y).$$

The closure of $\Gamma_{-\infty}$ is called the class of *left-extended generalized* Γ*-convolutions* and denoted by Γ_L. With the above considerations, its elements G can be defined as probability distributions on $(\mathbb{R}, \mathcal{B})$ whose characteristic functions are uniquely given by

$$\phi_G(u) = \exp\left[iub - \frac{cu^2}{2} - \int_0^\infty \left(\ln\left(1 - \frac{iu}{y}\right) + \frac{iuy}{1+y^2}\right) U(\mathrm{d}y)\right],$$

$$b \in \mathbb{R},\ c \geq 0,$$
$$U : \mathbb{R}_+ \to \mathbb{R}_+ \text{ non-decreasing with } U(0) = 0,$$
$$\int_0^1 |\ln(y)|\,U(\mathrm{d}y) < \infty \quad\text{and}\quad \int_1^\infty \frac{1}{y^2}\,U(\mathrm{d}y) < \infty.$$

Remark: As before, the two integrability conditions ensure the finiteness of the characteristic exponent such that $|\phi_G(u)| > 0$. They immediately follow from the asymptotic relations $\ln\big(1 - \frac{iu}{y}\big) = \ln(y - iu) - \ln(y) \sim -\ln(y)$, $y \to 0$, and $\frac{iuy}{1+y^2} + \ln\big(1 - \frac{iu}{y}\big) \sim \frac{iu}{y} - \frac{iu}{y} + \frac{u^2}{y^2} + o\big(\frac{u^2}{y^2}\big)$, $y \to \infty$. Note that the additional summand $\frac{iuy}{1+y^2}$ allows a weakening of the last condition compared to (1.24).

Analogously to Proposition 1.20 it can be shown that all elements $G \in \Gamma_L$ are selfdecomposable and possess a Khintchine representation obtained by

$$a_G = b + \int_0^\infty \left(\int_0^\infty \frac{e^{-yx}}{1+x^2}\, \mathrm{d}x \,-\, \frac{y}{1+y^2} \right) U(\mathrm{d}y),$$

$$\psi_G(x) = \begin{cases} 0, & x < 0, \\ c, & x = 0, \\ c + \displaystyle\int_0^x \psi_G'(y)\, \mathrm{d}y, & x > 0, \end{cases} \qquad \psi_G'(y) = \frac{y}{1+y^2} \int_0^\infty e^{-yt}\, U(\mathrm{d}t).$$

Of course we could have started our investigations with negative Gamma variables leading to the counterpart Γ'_0 of Γ_0 consisting of distributions on $(\mathbb{R}_-, \mathcal{B}_-)$. In the same way as above one obtains the class Γ_R as the closure of right-shifts of Γ_0'. Following Thorin (1978), we define the class Γ by

Definition 1.21 *The class Γ of convolutions $\Gamma_L * \Gamma_R$, called* extended generalized Γ-convolutions*, consists of all probability distributions F on $(\mathbb{R}, \mathcal{B})$ whose characteristic functions ϕ_F are of the form*

$$\phi_F(u) = \exp\left[iub - \frac{cu^2}{2} - \int_{-\infty}^{+\infty} \left(\ln\left(1 - \frac{iu}{y}\right) + \frac{iuy}{1+y^2} \right) U(\mathrm{d}y) \right], \tag{1.26}$$

$$\left. \begin{array}{l} b \in \mathbb{R},\ c \geq 0, \quad U : \mathbb{R} \to \mathbb{R} \ \textit{non-decreasing with } U(0) = 0, \\ \displaystyle\int_{-1}^{1} \big| \ln(|y|) \big|\, U(\mathrm{d}y) < \infty, \\ \displaystyle\int_{-\infty}^{-1} \frac{1}{y^2}\, U(\mathrm{d}y) + \int_1^{+\infty} \frac{1}{y^2}\, U(\mathrm{d}y) < \infty. \end{array} \right\} \tag{1.27}$$

Again, this representation is unique, and all elements of Γ are also selfdecomposable. Their Khintchine representations are

$$\begin{aligned} a_F = b &+ \int_0^\infty \left(\int_0^\infty \frac{e^{-yx}}{1+x^2}\, \mathrm{d}x - \frac{y}{1+y^2} \right) U(\mathrm{d}y) \\ &- \int_{-\infty}^0 \left(\int_{-\infty}^0 \frac{e^{-yx}}{1+x^2}\, \mathrm{d}x + \frac{y}{1+y^2} \right) U(\mathrm{d}y), \end{aligned} \tag{1.28}$$

$$\psi_F(x) = \begin{cases} \displaystyle\int_{-\infty}^x \psi_F'(y)\, \mathrm{d}y, & x < 0, \\ c + \displaystyle\int_{-\infty}^0 \psi_F'(y)\, \mathrm{d}y & x = 0, \\ \psi_F(0) + \displaystyle\int_0^x \psi_F'(y)\, \mathrm{d}y, & x > 0, \end{cases} \tag{1.29}$$

$$\psi_F'(y) = -\mathbb{1}_{(-\infty,0)}(y)\, \frac{y}{1+y^2} \int_{-\infty}^0 e^{-yt}\, U(\mathrm{d}t) + \mathbb{1}_{(0,\infty)}(y)\, \frac{y}{1+y^2} \int_0^\infty e^{-yt}\, U(\mathrm{d}t).$$

Remark: Equations (1.25) remain of course valid if G is replaced by F. Further note that not only every normal distribution belongs to Γ (take $U(y) \equiv 0$), but also all α-stable distributions with $0 < \alpha < 2$. The latter are obtained from the generating triplet $(b, 0, U)$ with

$$U(y) = \begin{cases} -\dfrac{c_1}{\Gamma(\alpha+1)}\,|y|^\alpha, \ y < 0, \\ \dfrac{c_2}{\Gamma(\alpha+1)}\,y^\alpha, \quad y \geq 0, \end{cases} \qquad c_1, c_2 \geq 0,\ c_1 + c_2 > 0.$$

It is easily seen that U fulfills the conditions (1.27), and from equations (1.28) and (1.25) it follows that for the proof of the assertion it is sufficient to show that the above choice of U yields the correct Lévy measure. By (1.29) and (1.25), for $x < 0$ the density $d_{\nu_F}(x)$ of the Lévy measure is given by

$$\begin{aligned} d_{\nu_F}(x) &= -\frac{1+x^2}{x^2}\,\frac{x}{1+x^2}\int_{-\infty}^{0} e^{-xt}\,U(\mathrm{d}t) = \frac{1}{|x|}\int_{-\infty}^{0} e^{|x|t}\,\frac{c_1}{\Gamma(\alpha+1)}\,\alpha|t|^{\alpha-1}\,\mathrm{d}t \\ &= -\frac{c_1}{\Gamma(\alpha+1)}\frac{1}{|x|}\left(-e^{|x|t}\,|t|^\alpha\Big|_{-\infty}^{0} + \int_{-\infty}^{0}|x|e^{|x|t}\,|t|^\alpha\,\mathrm{d}t\right) \\ &= \frac{c_1}{\Gamma(\alpha+1)}\frac{1}{|x|^{1+\alpha}}\int_0^\infty e^{-y}y^\alpha\,\mathrm{d}y = c_1|x|^{-1-\alpha} \end{aligned}$$

and analogously we get $d_{\nu_F}(x) = c_2 x^{-1-\alpha}$ for $x > 0$, which together equals the density of the Lévy measure of an α-stable distribution (see Sato 1999, p. 80).

The next theorem, taken from Thorin (1977b) and Thorin (1978), shows the closedness of the classes Γ_0 and Γ.

Theorem 1.22 (Continuity Theorem) *If a sequence $(G_n)_{n\geq 1}$ of generalized Γ-convolutions generated by $(a_n, U_n)_{n\geq 1}$ converges weakly to a distribution function G, then G is also a generalized Γ-convolution generated by (a, U) where*

$$\begin{aligned} U(x) &= \lim_{n\to\infty} U_n(x) \quad \textit{in every continuity point } x \textit{ of } U, \\ a &= \lim_{M\to\infty}\lim_{n\to\infty}\left[a_n + \int_M^\infty \frac{1}{x}\,U_n(\mathrm{d}x)\right]. \end{aligned}$$

If instead $(F_n)_{n\geq 1}$ is a sequence of extended generalized Γ-convolutions generated by $(b_n, c_n, U_n)_{n\geq 1}$ which converges weakly to a distribution function F, then F is also an extended generalized Γ-convolution generated by (b, c, U) where

$$\begin{aligned} &U(x) = \lim_{n\to\infty} U_n(x) \quad \textit{in every continuity point } x \textit{ of } U, \\ &b = \lim_{n\to\infty} b_n, \qquad c = \lim_{M\to\infty}\lim_{n\to\infty}\left[c_n + \int_{\mathbb{R}\setminus[-M,M]}\frac{1}{x^2}\,U_n(\mathrm{d}x)\right]. \end{aligned}$$

By construction every sum of a finite number of positive (and negative) Gamma variables is an (extended) generalized Γ-convolution. Thus it follows as a special case from the continuity theorem that if a sequence of such sums converges in distribution to a random variable X, then $\mathcal{L}(X)$ is an (extended) generalized Γ-convolution as well. Conversely, every extended generalized Γ-convolution generated by (b, c, U) may be approximated arbitrarily well in distribution by sums of suitably scaled and shifted independent Gamma variables. As it is easily seen from our considerations on p. 29, the main task in the latter case is to choose the parameters λ_i and σ_i of the summands in such a way that the functions U_n corresponding to the finite sums converge pointwise to U.

1.6 Representations of GIG and GH distributions as subclasses of Γ

We now show that all GIG and GH distributions belong to Γ_0 and Γ, respectively. The first statement was already proven in Halgreen (1979), the second was indicated in Thorin (1978). We give detailed proofs of both and extend the results with the help of the Continuity Theorem 1.22 to the limiting cases which allows to explicitly compute the Lévy–Khintchine representations for all distributions.

1.6.1 GIG distributions and their limits

The results of the present subsection are summarized in the following

Proposition 1.23 *Every $GIG(\lambda,\delta,\gamma)$-distribution is a generalized Γ-convolution with generating pair (a_{GIG}, U_{GIG}) as follows:*

a) If $\delta,\gamma > 0$, then

$$\begin{aligned} a_{GIG(\lambda,\delta,\gamma)} &= 0, \\ U_{GIG(\lambda,\delta,\gamma)}(x) &= \mathbb{1}_{[\gamma^2/2,\infty)}(x)\left(\max(0,\lambda) + \delta^2 \int_{\frac{\gamma^2}{2}}^{x} g_{|\lambda|}\big(2\delta^2 y - \delta^2\gamma^2\big)\,\mathrm{d}y\right) \end{aligned}$$

where g_ν is defined by

$$g_\nu(x) := \frac{2}{\pi^2 x\big[J_\nu^2\big(\sqrt{x}\big) + Y_\nu^2\big(\sqrt{x}\big)\big]}, \qquad \nu \geq 0.$$

b) If $\lambda > 0$ and $\delta = 0$ (Gamma limiting case), we have

$$a_{GIG(\lambda,0,\gamma)} = 0, \quad U_{GIG(\lambda,0,\gamma)}(x) = \lambda \mathbb{1}_{[\gamma^2/2,\infty)}(x).$$

c) If $\lambda < 0$ and $\gamma = 0$ (inverse Gamma limiting case), then

$$a_{GIG(\lambda,\delta,0)} = 0, \quad U_{GIG(\lambda,\delta,0)}(x) = \delta^2 \mathbb{1}_{[0,\infty)}(x) \int_0^x g_{|\lambda|}\big(2\delta^2 y\big)\,\mathrm{d}y.$$

d) The degenerate limiting case ($\delta,\gamma \to \infty$, $\frac{\delta}{\gamma} \to \sigma \geq 0$) is characterized by $a = \sigma$, $U(x) \equiv 0$.

Proof: a) From Proposition 1.11 c) and equation (1.3) of Section 1.2 we have

$$GIG(-\lambda,\delta,\gamma) * GIG(\lambda,0,\gamma) = GIG(-\lambda,\delta,\gamma) * G\big(\lambda, \tfrac{\gamma^2}{2}\big) = GIG(\lambda,\delta,\gamma),\ \lambda > 0.$$

Since the class Γ_0 is closed under convolutions with Gamma distributions by construction, it suffices to prove that all $GIG(\lambda,\delta,\gamma)$-distributions with $\lambda \leq 0$ belong to Γ_0 to establish the general result. Therefore we suppose $-\lambda =: \nu \geq 0$ for the moment. By Proposition 1.9 and equation (A.2) the Laplace transform of $GIG(\lambda,\delta,\gamma)$ is given by

$$\mathfrak{L}_{GIG(\lambda,\delta,\gamma)}(u) = \left(\frac{\gamma^2}{\gamma^2 + 2u}\right)^{\frac{\lambda}{2}} \frac{K_\nu\big(\delta\sqrt{\gamma^2 + 2u}\big)}{K_\nu(\delta\gamma)}.$$

Combining equations (A.3) and (A.4) we get $K'_\lambda(x) = -K_{\lambda-1}(x) - \frac{\lambda}{x}K_\lambda(x)$. With the help of this we find

$$
\begin{aligned}
\ln\big(\mathfrak{L}_{GIG(\lambda,\delta,\gamma)}(u)\big)' &= \frac{\mathfrak{L}'_{GIG(\lambda,\delta,\gamma)}(u)}{\mathfrak{L}_{GIG(\lambda,\delta,\gamma)}(u)} \\
&= \frac{-\lambda\gamma^\lambda\big(\gamma^2+2u\big)^{-\frac{\lambda}{2}-1}\,\frac{K_\nu(\delta\sqrt{\gamma^2+2u})}{K_\nu(\delta\gamma)} + \delta\gamma^\lambda\big(\gamma^2+2u\big)^{-\frac{\lambda+1}{2}}\,\frac{K'_\nu(\delta\sqrt{\gamma^2+2u})}{K_\nu(\delta\gamma)}}{\gamma^\lambda\big(\gamma^2+2u\big)^{-\frac{\lambda}{2}}\,\frac{K_\nu(\delta\sqrt{\gamma^2+2u})}{K_\nu(\delta\gamma)}} \\
&= \frac{\nu}{\gamma^2+2u} + \frac{\delta}{\sqrt{\gamma^2+2u}}\,\frac{K'_\nu\big(\delta\sqrt{\gamma^2+2u}\big)}{K_\nu\big(\delta\sqrt{\gamma^2+2u}\big)} \\
&= \frac{\nu}{\gamma^2+2u} - \frac{\delta}{\sqrt{\gamma^2+2u}}\,\frac{K_{\nu-1}\big(\delta\sqrt{\gamma^2+2u}\big) + \nu\big(\delta\sqrt{\gamma^2+2u}\big)^{-1}K_\nu\big(\delta\sqrt{\gamma^2+2u}\big)}{K_\nu\big(\delta\sqrt{\gamma^2+2u}\big)} \\
&= -\delta^2\,\Phi_\nu\big[\delta^2\big(\gamma^2+2u\big)\big], \qquad \text{where}\;\; \Phi_\nu(t) := \frac{K_{\nu-1}\big(\sqrt{t}\big)}{\sqrt{t}\,K_\nu\big(\sqrt{t}\big)}.
\end{aligned}
$$

Using the integral representation of Grosswald (1976),

$$
\Phi_\nu(t) = \int_0^\infty \frac{1}{t+x}\,g_\nu(x)\,\mathrm{d}x \;\;\text{with}\;\; g_\nu(x) = \frac{2}{\pi^2 x\big[J_\nu^2\big(\sqrt{x}\big) + Y_\nu^2\big(\sqrt{x}\big)\big]}, \tag{1.30}
$$

($J_\nu(x)$ and $Y_\nu(x)$ denote the Bessel functions of first and second kind with index ν, see Appendix A for further information) we obtain with Fubini's theorem

$$
\begin{aligned}
\ln\big(\mathfrak{L}_{GIG(\lambda,\delta,\gamma)}(u)\big) &= -\int_0^u\int_0^\infty \frac{\delta^2}{\delta^2\gamma^2+2\delta^2 t+x}\;g_\nu(x)\;\mathrm{d}x\,\mathrm{d}t \\
&= -\int_{\frac{\gamma^2}{2}}^\infty\int_0^u \frac{\delta^2}{t+y}\;g_\nu\big(2\delta^2 y-\delta^2\gamma^2\big)\;\mathrm{d}t\;\mathrm{d}y \\
&= -\int_{\frac{\gamma^2}{2}}^\infty \ln\Big(1+\frac{u}{y}\Big)\;\delta^2 g_\nu\big(2\delta^2 y-\delta^2\gamma^2\big)\;\mathrm{d}y.
\end{aligned}
$$

Since the corresponding characteristic functions are given by $\mathfrak{L}_{GIG(\lambda,\delta,\gamma)}(-iu)$, it immediately follows from (1.23) in Definition 1.19 that all $GIG(\lambda,\delta,\gamma)$-distributions with $\lambda \le 0$ and $\delta,\gamma > 0$ are generalized Γ-convolutions with characteristics

$$
a_{GIG} = 0, \quad U_{GIG}(x) = \delta^2\mathbb{1}_{[\gamma^2/2,\infty)}(x)\int_{\frac{\gamma^2}{2}}^x g_{|\lambda|}\big(2\delta^2 y-\delta^2\gamma^2\big)\,\mathrm{d}y,
$$

if U_{GIG} fulfills the integrability conditions (1.24) which is shown below. Since a Gamma distribution $G\big(\lambda,\frac{\gamma^2}{2}\big)$ is characterized by a jump of U of height λ at $x = \frac{\gamma^2}{2}$ (that is, $\lambda = U\big(\frac{\gamma^2}{2}\big) - U\big(\frac{\gamma^2}{2}-\big)$, see also equation (1.22) and the considerations thereafter), the general representation of $GIG(\lambda,\delta,\gamma)$-distributions with arbitrary $\lambda \in \mathbb{R}$ and $\delta,\gamma > 0$ immediately follows from what we have shown so far and the above mentioned convolution property of Proposition 1.11 c).

To verify conditions (1.24), note that $\int_0^1 |\ln(y)|\, U_{GIG}(\mathrm{d}y) < \infty$ holds trivially if $\gamma > 0$ because $U_{GIG}(y) \equiv 0$ on $\big[0, \frac{\gamma^2}{2}\big)$. Equations (A.16) and (A.17) imply the following asymptotic behavior of the Bessel functions $J_\lambda(x)$ and $Y_\lambda(x)$ for $x \to \infty$:

$$J_\lambda(x) \sim \sqrt{\frac{2}{\pi x}} \cos\left(x - \frac{\lambda\pi}{2} - \frac{\pi}{4}\right), \qquad Y_\lambda(x) \sim \sqrt{\frac{2}{\pi x}} \sin\left(x - \frac{\lambda\pi}{2} - \frac{\pi}{4}\right),$$

therefore $g_{|\lambda|}(x) \sim (\pi^2 x)^{-\frac{1}{2}}$, $x \to \infty$, and consequently $\int_1^\infty y^{-1}\, U_{GIG}(\mathrm{d}y) < \infty$.

b) Observing that $GIG(\lambda, 0, \gamma) = G\big(\lambda, \frac{\gamma^2}{2}\big)$, the corresponding result has already been shown implicitly in the proof of part a). However, we also provide an alternative proof here which is based on the Continuity Theorem 1.22 since we need some of the enclosed calculations later anyway.

First we require the asymptotics of the Bessel functions near the origin. According to equations (A.13) and (A.14), for $|\lambda| > 0$ these are given by

$$J_{|\lambda|}(x) \sim \left(\frac{x}{2}\right)^{|\lambda|} \big(\Gamma(|\lambda|)\big)^{-1}, \qquad Y_{|\lambda|}(x) \sim \frac{\Gamma(|\lambda|)}{\pi} \left(\frac{x}{2}\right)^{-|\lambda|}, \quad x \to 0.$$

For an arbitrary but fixed $x > \frac{\gamma^2}{2}$ we get combining the results of part a), equation (1.30), and the above mentioned asymptotics

$$\lim_{\delta\to 0} U'_{GIG(\lambda,\delta,\gamma)}(x) = \lim_{\delta\to 0} \frac{2}{\pi^2(2x-\gamma^2)\big[J_\lambda^2\big(\delta\sqrt{2x-\gamma^2}\big) + Y_\lambda^2\big(\delta\sqrt{2x-\gamma^2}\big)\big]} = 0,$$

hence $\lim_{\delta\to 0} U_{GIG(\lambda,\delta,\gamma)}(x) = \lambda \mathbb{1}_{[\gamma^2/2,\infty)}(x) = U_{GIG(\lambda,0,\gamma)}(x)$ by the Continuity Theorem. Further we conclude from the asymptotic behaviour of the Bessel functions that

$$\frac{1}{x}\, U'_{GIG(\lambda,\delta,\gamma)}(x) \sim \frac{\delta^{2\lambda}(2x-\gamma^2)^{\lambda-1}}{\Gamma(\lambda)^2\, 2^{2\lambda-1} x}, \quad \delta\sqrt{x} \to 0,$$

$$\frac{1}{x}\, U'_{GIG(\lambda,\delta,\gamma)}(x) \sim \frac{\delta(2x-\gamma^2)^{-\frac{1}{2}}}{\pi x}, \quad \delta\sqrt{x} \to \infty,$$

consequently $x^{-1}U'_{GIG}(x)$ is bounded on $[M,\infty)$ by an integrable majorant for sufficiently large M and $0 < \delta \le 1$. The Continuity Theorem and the dominated convergence theorem then yield

$$\begin{aligned} a_{GIG(\lambda,0,\gamma)} &= \\ &= \lim_{M\to\infty} \lim_{\delta\to 0} \left[a_{GIG(\lambda,\delta,\gamma)} + \int_M^\infty \frac{1}{x}\, U_{GIG(\lambda,\delta,\gamma)}(\mathrm{d}x)\right] \\ &= \lim_{M\to\infty} \lim_{\delta\to 0} \int_M^\infty \frac{1}{x}\, U'_{GIG(\lambda,\delta,\gamma)}(x)\,\mathrm{d}x = \lim_{M\to\infty} \int_M^\infty \frac{1}{x}\, U'_{GIG(\lambda,0,\gamma)}(x)\,\mathrm{d}x = 0, \end{aligned}$$

where the last step follows from $U'_{GIG(\lambda,0,\gamma)}(x) = 0$ for sufficiently large x as shown above.

c) The asymptotics of $U'_{GIG(\lambda,\delta,\gamma)}(x)$ derived in the proof of part b) imply

$$\lim_{M\to\infty} \lim_{\gamma\to 0} \left[a_{GIG(\lambda,\delta,\gamma)} + \int_M^\infty \frac{1}{x}\, U_{GIG(\lambda,\delta,\gamma)}(\mathrm{d}x)\right] = \lim_{M\to\infty} \int_M^\infty \frac{\delta}{\pi\sqrt{2}}\, x^{-\frac{3}{2}}\,\mathrm{d}x = 0,$$

hence by the Continuity Theorem every inverse Gamma distribution $iG\big(\lambda, \frac{\delta^2}{2}\big) = GIG(\lambda,\delta,0) \in \Gamma_0$ with generating pair $a_{GIG(\lambda,\delta,0)} = 0$ and, as is obvious from part a), $U_{GIG(\lambda,\delta,0)}(x) = \lim_{\gamma\to 0} U_{GIG(\lambda,\delta,\gamma)}(x) = \delta^2 \mathbb{1}_{[0,\infty)}(x) \int_0^x g_{|\lambda|}\big(2\delta^2 y\big)\,\mathrm{d}y$.

d) The generating pair (a,U) of the degenerate distribution ϵ_σ can of course immediately be obtained by comparing the characteristic function $\phi_{\epsilon_\sigma}(u) = e^{iu\sigma}$ with the general representation (1.23) in Definition 1.19. We derive it here again with the help of the Continuity Theorem which also provides an alternative proof of Corollary 1.10.

By part a) we have $U_{GIG(\lambda,\delta,\gamma)}(x) \equiv 0$ for $x \in \big[0, \frac{\gamma^2}{2}\big)$, consequently $U(x) = \lim_{\delta,\gamma\to\infty} U_{GIG(\lambda,\delta,\gamma)}(x) = 0$ for all $x \geq 0$. The first further implies

$$\lim_{M\to\infty} \lim_{\delta,\gamma\to\infty} \left[a_{GIG(\lambda,\delta,\gamma)} + \int_M^\infty \frac{1}{x}\, U_{GIG(\lambda,\delta,\gamma)}(\mathrm{d}x)\right] = \lim_{\delta,\gamma\to\infty} \int_{\frac{\gamma^2}{2}}^\infty \frac{1}{x}\, U_{GIG(\lambda,\delta,\gamma)}(\mathrm{d}x)$$

and from the Continuity Theorem and the asymptotics of $x^{-1}U'_{GIG(\lambda,\delta,\gamma)}(x)$ derived in b) we finally conclude

$$\begin{aligned} a &= \lim_{\delta,\gamma\to\infty} \int_{\frac{\gamma^2}{2}}^\infty \frac{1}{x}\, U_{GIG(\lambda,\delta,\gamma)}(\mathrm{d}x) = \lim_{\delta,\gamma\to\infty} \int_{\frac{\gamma^2}{2}}^\infty \frac{\delta(2x-\gamma^2)^{-\frac{1}{2}}}{\pi x}\,\mathrm{d}x \\ &= \lim_{\delta,\gamma\to\infty} \frac{\delta}{\pi}\,\frac{2}{\gamma} \arctan\left(\frac{\sqrt{2x-\gamma^2}}{\gamma}\right)\Bigg|_{\frac{\gamma^2}{2}}^\infty = \lim_{\delta,\gamma\to\infty} \frac{\delta}{\gamma} = \sigma. \end{aligned}$$

□

Remark: The proof of part d) also shows that the limits occuring in the Continuity Theorem must not be interchanged. By (1.24), the measure induced by the function U of every generalized Γ-convolution has to fulfill $\int_1^\infty x^{-1} U(\mathrm{d}x) < \infty$, so a swap of the limits would imply

$$a = \lim_{n\to\infty} \lim_{M\to\infty} \left[a_n + \int_M^\infty \frac{1}{x}\, U_n(\mathrm{d}x)\right] = \lim_{n\to\infty} a_n.$$

Therewith one would obtain in the situation of Proposition 1.23 d) that $a = \lim_{\delta,\gamma\to\infty} a_{GIG(\lambda,\delta,\gamma)} = 0$ which is obviously false in general.

1.6.2 Lévy–Khintchine representations of GIG distributions

With the above characteristics (a_{GIG}, U_{GIG}) of GIG distributions, their Lévy–Khintchine representation can now easily be derived using Proposition 1.20 and formulas (1.25). Again we summarize the results in

Proposition 1.24 *The characteristic functions of $GIG(\lambda,\delta,\gamma)$-distributions can be represented as follows:*

a) If $\delta,\gamma > 0$, then

$$\phi_{GIG(\lambda,\delta,\gamma)}(u) = \exp\left(iu\frac{\delta K_{\lambda+1}(\delta\gamma)}{\gamma K_\lambda(\delta\gamma)} + \int_0^\infty \big(e^{iux} - 1 - iux\big)\, g_{GIG(\lambda,\delta,\gamma)}(x)\,\mathrm{d}x\right)$$

where the density of the Lévy measure is defined for $x > 0$ by

$$g_{GIG(\lambda,\delta,\gamma)}(x) = \frac{e^{-x\frac{\gamma^2}{2}}}{x}\left[\int_0^\infty \frac{e^{-xy}}{\pi^2 y\big[J_{|\lambda|}^2\big(\delta\sqrt{2y}\big)+Y_{|\lambda|}^2\big(\delta\sqrt{2y}\big)\big]}\,\mathrm{d}y + \max(0,\lambda)\right].$$

b) If $\lambda > 0$ and $\delta = 0$ (Gamma limiting case), we have

$$\phi_{GIG(\lambda,0,\gamma)}(u) = \exp\left(iu\,\frac{2\lambda}{\gamma^2} + \int_0^\infty \big(e^{iux}-1-iux\big)\,g_{GIG(\lambda,0,\gamma)}(x)\,\mathrm{d}x\right),$$

and the density of the Lévy measure is $g_{GIG(\lambda,0,\gamma)}(x) = \mathbb{1}_{(0,\infty)}(x)\,\frac{\lambda}{x}\,e^{-\frac{\gamma^2}{2}x}$.

c) If $\lambda < 0$ and $\gamma = 0$ (inverse Gamma limiting case), then

$$\begin{aligned}\phi_{GIG(\lambda,\delta,0)}(u) = \exp\Bigg(& iu\delta^2\int_0^\infty \frac{1-e^{-x}}{x}\,g_{|\lambda|}\big(2\delta^2 x\big)\,\mathrm{d}x \\ &+ \int_0^\infty \big(e^{iux}-1-iux\mathbb{1}_{[0,1]}(x)\big)\,g_{GIG(\lambda,\delta,0)}(x)\,\mathrm{d}x\Bigg)\end{aligned}$$

with corresponding Lévy density

$$g_{GIG(\lambda,\delta,0)}(x) = \frac{1}{x}\int_0^\infty \frac{e^{-xy}}{\pi^2 y\big[J_{|\lambda|}^2\big(\delta\sqrt{2y}\big)+Y_{|\lambda|}^2\big(\delta\sqrt{2y}\big)\big]}\,\mathrm{d}y, \quad x > 0.$$

PROOF: a) Since all moments of $GIG(\lambda,\delta,\gamma)$ distributions with $\delta,\gamma > 0$ exist (see p. 11), according to the remark on p. 30 the truncation function within the integral of the Lévy–Khintchine formula can be omitted, and the drift term b_G then equals the mean of the distribution. From the moment formulas given on p. 11 we get $b_{GIG(\lambda,\delta,\gamma)} = \mathrm{E}[GIG(\lambda,\delta,\gamma)] = \frac{K_{\lambda+1}(\delta\gamma)}{K_\lambda(\delta\gamma)}\frac{\delta}{\gamma}$. Thus it only remains to prove the above formula for the Lévy density. By Proposition 1.20, the last equation of (1.25) and Proposition 1.23 a) we have

$$\begin{aligned}g_{GIG(\lambda,\delta,\gamma)}(x) &= \frac{1+x^2}{x^2}\,\frac{x}{1+x^2}\int_0^\infty e^{-xy}\,U_{GIG(\lambda,\delta,\gamma)}(\mathrm{d}y)\\ &= \frac{1}{x}\Bigg[\int_{\frac{\gamma^2}{2}}^\infty \frac{2\delta^2 e^{-xy}}{\pi^2\big(2\delta^2y-\delta^2\gamma^2\big)\big[J_{|\lambda|}^2\big(\sqrt{2\delta^2y-\delta^2\gamma^2}\big)+Y_{|\lambda|}^2\big(\sqrt{2\delta^2y-\delta^2\gamma^2}\big)\big]}\,\mathrm{d}y\\ &\qquad\qquad + \max(0,\lambda)\,e^{-x\frac{\gamma^2}{2}}\Bigg]\\ &= \frac{e^{-x\frac{\gamma^2}{2}}}{x}\left[\int_0^\infty \frac{e^{-xy}}{\pi^2 y\big[J_{|\lambda|}^2\big(\delta\sqrt{2y}\big)+Y_{|\lambda|}^2\big(\delta\sqrt{2y}\big)\big]}\,\mathrm{d}y + \max(0,\lambda)\right],\ x > 0.\end{aligned}$$

b) Because Gamma distributions $G\big(\lambda,\frac{\gamma^2}{2}\big) = GIG(\lambda,0,\gamma)$ also possess moments of arbitrary (positive) order, analogously to part a) we conclude that a truncation function within the Lévy–Khintchine representation is dispensable, $b_{GIG(\lambda,0,\gamma)} = \mathrm{E}[GIG(\lambda,0,\gamma)] = \frac{2\lambda}{\gamma^2}$, and the Lévy density is given by

$$\begin{aligned} g_{GIG(\lambda,0,\gamma)}(x) &= \frac{1+x^2}{x^2}\,\frac{x}{1+x^2}\int_0^\infty e^{-xy}\,U_{GIG(\lambda,0,\gamma)}(\mathrm{d}y) \\ &= \frac{\lambda}{x}\int_0^\infty e^{-xy}\,\epsilon_{\frac{\gamma^2}{2}}(\mathrm{d}y) = \frac{\lambda}{x}\,e^{-\frac{\gamma^2}{2}x}, \qquad x>0, \end{aligned}$$

because $U_{GIG(\lambda,0,\gamma)}(x) = \lambda \mathbb{1}_{[\gamma^2/2,\infty)}(x)$.

c) Contrary to a) and b), the limiting reciprocal Gamma distributions have finite first moments only if $\lambda < -1$ (see p. 11), so in general we have to use a truncation function and determine the drift term $b_{GIG(\lambda,\delta,0)}$ according to (1.25). By Proposition 1.23 c) the measure induced by $U_{GIG(\lambda,\delta,0)}$ is concentrated on $\mathbb{R}_+$ and $a_{GIG(\lambda,\delta,0)} = 0$, so Proposition 1.20 and (1.25) imply

$$\begin{aligned} b_{GIG(\lambda,\delta,0)} &= \int_0^\infty\!\!\int_0^\infty \frac{e^{-xy}}{1+x^2}\,\mathrm{d}x\,U_{GIG(\lambda,\delta,0)}(\mathrm{d}y) \\ &\quad + \int_0^\infty\!\!\int_0^\infty \left(\frac{x^2e^{-xy}}{1+x^2}\,\mathbb{1}_{[0,1]}(x) - \frac{e^{-xy}}{1+x^2}\,\mathbb{1}_{\mathbb{R}\setminus[0,1]}(x)\right)\mathrm{d}x\,U_{GIG(\lambda,\delta,0)}(\mathrm{d}y) \\ &= \int_0^\infty\!\!\int_0^1 e^{-xy}\,\mathrm{d}x\,U_{GIG(\lambda,\delta,0)}(\mathrm{d}y) = \delta^2\int_0^\infty \frac{1-e^{-y}}{y}\,g_{|\lambda|}\big(2\delta^2 y\big)\,\mathrm{d}y, \end{aligned}$$

and similarly to part a) the Lévy density is obtained to be

$$g_{GIG(\lambda,\delta,0)}(x) = \frac{1}{x}\int_0^\infty \frac{e^{-xy}}{\pi^2 y\big[J^2_{|\lambda|}\big(\delta\sqrt{2y}\big) + Y^2_{|\lambda|}\big(\delta\sqrt{2y}\big)\big]}\,\mathrm{d}y, \quad x>0.$$

□

Remark: If the GIG distributions arise as a limit of GH distributions studied in Section 1.4.2, and the parameter μ of the converging sequence is not equal to 0, all characteristic functions have an additional factor $e^{iu\mu}$. The corresponding formulas for the "negative" GIG distributions on $(\mathbb{R}_-,\mathcal{B}_-)$ are obtained from the above by changing $b_{GIG(\lambda,\delta,\gamma)}$ to $-b_{GIG(\lambda,\delta,\gamma)}$, the integration interval from $\mathbb{R}_+$ to $\mathbb{R}_-$ and the truncation function from $\mathbb{1}_{[0,1]}$ to $\mathbb{1}_{[-1,0]}$. In the expressions for the Lévy densities x has to be replaced by $|x|$.

Observe that the Lévy densities of $GIG(\lambda,\delta,\gamma)$-distributions with $\delta>0$ are essentially Laplace transforms of the function g_ν defined in Proposition 1.23:

$$g_{GIG(\lambda,\delta,\gamma)}(x) = \frac{e^{-x\frac{\gamma^2}{2}}}{x}\left[\max(0,\lambda) + \int_0^\infty e^{-xy}\,\delta^2 g_{|\lambda|}\big(2\delta^2 y\big)\,\mathrm{d}y\right].$$

This fact allows us to derive their asymptotic behaviour near the origin with the help of the following Tauberian theorem which can be found in Feller (1971, p. 446, and Problem 16 on p. 464).

Theorem 1.25 *Let G be a measure concentrated on $\overline{\mathbb{R}}_+$ with density g and existing Laplace transform $\mathfrak{L}_G(x)$ for every $x > 0$. Suppose $g(y) \sim v(y)$ for $y \to \infty$ and v is monotone on some interval (y_0, ∞). Let $0 < \rho < \infty$, then as $x \to 0$ and $y \to \infty$, respectively,*

$$\mathfrak{L}_G(x) \sim \frac{1}{x^\rho} L\left(\frac{1}{x}\right) \qquad \textit{iff} \quad v(y) \sim \frac{y^{\rho-1}}{\Gamma(\rho)} L(y)$$

for some positive function L defined on $\mathbb{R}_+$ and varying slowly at ∞ (that is, for every fixed $x > 0$ and $t \to \infty$ we have $\frac{L(tx)}{L(x)} \to 1$).

As we have already seen on p. 36, the asymptotic behaviour of the Bessel functions J_λ and Y_λ implies $\delta^2 g_{|\lambda|}(2\delta^2 y) \sim \delta(2\pi^2 y)^{-\frac{1}{2}}$, $y \to \infty$. The assumptions of Theorem 1.25 are thus fulfilled with $\rho = 0.5$ and $L(y) \equiv \frac{\delta\Gamma(0.5)}{\sqrt{2\pi}} = \delta(2\pi)^{-\frac{1}{2}}$, hence for $x \to 0$ the asymptotics of the Laplace transform above are $\delta(2\pi x)^{-\frac{1}{2}}$. Using the expansion $e^{-x\frac{\gamma^2}{2}} = 1 - x\frac{\gamma^2}{2} + o(x)$, $x \to 0$, we see that the behaviour of $g_{GIG(\lambda,\delta,\gamma)}$ near the origin is dominated by the integral term, multiplied with the preceeding factor x^{-1}, which gives

$$g_{GIG(\lambda,\delta,\gamma)}(x) \sim \frac{\delta}{\sqrt{2\pi}} x^{-\frac{3}{2}}, \qquad x \downarrow 0. \tag{1.31}$$

Only the Lévy densities of the limiting Gamma distributions show a different behaviour; in this case we have $g_{GIG(\lambda,0,\gamma)}(x) = \frac{\lambda}{x} e^{-x\gamma^2/2}$ and consequently

$$g_{GIG(\lambda,0,\gamma)}(x) \sim \frac{\lambda}{x}, \qquad x \downarrow 0.$$

1.6.3 GH distributions and their limits

The fact that GH distributions are a subclass of Γ is an immediate consequence of the more general result to be shown below that every normal mean-variance mixture is an extended generalized Γ-convolution if the mixing distribution belongs to Γ_0. This was already mentioned, but not rigorously proven in Thorin (1978).

Proposition 1.26 *If $F = N(\mu + \beta y, y) \circ G$ is a normal mean-variance mixture where the mixing distribution $G \in \Gamma_0$ is a generalized Γ-convolution with characteristic pair (a_G, U_G), then F is an extended generalized Γ-convolution ($F \in \Gamma$) generated by (b_F, c_F, U_F) with*

$$b_F = \mu + \beta a_G + \int_0^\infty \frac{2\beta(2y-1)}{(2y+1)^2 + 4\beta^2} U_G(\mathrm{d}y), \qquad c_F = a_G,$$

$$U_F(y) = \begin{cases} -U_G\big(\frac{y^2}{2} + \beta y\big), & y \le -|\beta| - \beta, \\ 0, & -|\beta| - \beta < y < |\beta| - \beta, \\ U_G\big(\frac{y^2}{2} + \beta y\big), & y \ge |\beta| - \beta. \end{cases}$$

PROOF: At first we note that equation (1.26) is equivalent to

$$\phi_\Gamma(0) = 1, \qquad \ln\big(\phi_\Gamma(u)\big)' = ib - cu + i\int_{-\infty}^{+\infty}\left(\frac{1}{y-iu} - \frac{y}{1+y^2}\right)U(\mathrm{d}y) \quad (1.32)$$

(conditions (1.27) justify the interchange between differentiation and integration). Because $G \in \Gamma_0$ by assumption, equation (1.23) implies

$$\mathfrak{L}_G(u) = \phi_G(iu) = \exp\left[-a_G u - \int_0^\infty \ln\left(1+\frac{u}{y}\right)U_G(\mathrm{d}y)\right].$$

From the proof of Lemma 1.6 b) we further know that the characteristic function of $F = N(\mu+\beta y, y)\circ G$ is given by

$$\begin{aligned}\phi_F(u) &= e^{iu\mu}\,\mathfrak{L}_G\big(\tfrac{u^2}{2} - iu\beta\big)\\ &= \exp\left[iu(\mu+\beta a_G) - a_G\frac{u^2}{2} - \int_0^\infty \ln\left(1+\frac{\frac{u^2}{2}-iu\beta}{y}\right)U_G(\mathrm{d}y)\right]\end{aligned}$$

and therefore

$$\ln\big(\phi_F(u)\big)' = i(\mu+\beta a_G) - a_G u + i\int_0^\infty \frac{iu+\beta}{y+\big(\frac{u^2}{2}-iu\beta\big)}\,U_G(\mathrm{d}y).$$

We thus have to show that the right hand side of the last equation admits a representation in the form of (1.32). A first calculation yields

$$\begin{aligned}\int_0^\infty &\frac{iu+\beta}{y+\big(\frac{u^2}{2}-iu\beta\big)}\,U_G(\mathrm{d}y)\\ &= \int_0^\infty \frac{2(iu+\beta)}{\big(\sqrt{2y+\beta^2}+\beta+iu\big)\big(\sqrt{2y+\beta^2}-\beta-iu\big)}\,U_G(\mathrm{d}y)\\ &= \int_{|\beta|-\beta}^\infty \frac{2(iu+\beta)}{(x+2\beta+iu)(x-iu)}\,\bar U(\mathrm{d}x)\end{aligned}$$

with $x = \sqrt{2y+\beta^2}-\beta$ and $\bar U(x) = U_G\big(\frac{x^2}{2}+\beta x\big)$. Extending the function $\bar U$ to the whole real line requires some care because by (1.24) the domain of U_G is $\mathbb{R}_+$, but $\frac{x^2}{2}+\beta x < 0$ if $\beta < 0$ and $0 < x < -2\beta$. Therefore we set

$$\bar U(x) := \begin{cases} 0, & 0 \le x < |\beta|-\beta,\\ U_G\big(\frac{x^2}{2}+\beta x\big), & x \ge |\beta|-\beta,\end{cases}$$

and complete it to a non-decreasing function on $\mathbb{R}$ symmetric around $-\beta$ by

$$\bar U(x) := \begin{cases} 0, & -|\beta|-\beta < x < 0,\\ -U_G\big(\frac{x^2}{2}+\beta x\big), & x \le -|\beta|-\beta.\end{cases}$$

Continuing the calculation (and writing y again instead of x) we find

$$\begin{aligned}
\int_0^\infty \frac{iu+\beta}{y+\left(\frac{u^2}{2}-iu\beta\right)}\, U_G(\mathrm{d}y) &= \int_0^\infty \frac{2(iu+\beta)}{(y+2\beta+iu)(y-iu)}\, \bar{U}(\mathrm{d}y) \\
&= \int_0^\infty \frac{2(iu+\beta)}{(y+\beta)^2+(u-i\beta)^2}\, \bar{U}(\mathrm{d}y) = \int_{-\infty}^{+\infty} \frac{iu+\beta}{(y+\beta)^2+(u-i\beta)^2}\, \bar{U}(\mathrm{d}y) \\
&= \int_{-\infty}^{+\infty} \frac{iu+\beta}{(y+\beta)^2+(u-i\beta)^2} + \frac{y+\beta}{(y+\beta)^2+(u-i\beta)^2}\, \bar{U}(\mathrm{d}y) \\
&= \int_{-\infty}^{+\infty} \frac{1}{y-iu}\, \bar{U}(\mathrm{d}y)
\end{aligned}$$

where the second and third lines follow from the symmetry of $\bar{U}(y)$ around $-\beta$. Summing up we have $\ln\big(\phi_F(u)\big)' = i(\mu+\beta a_G) - a_G u + i\int_{-\infty}^{+\infty} \frac{1}{y-iu}\, \bar{U}(\mathrm{d}y)$, and a comparison with equation (1.32) reveals that F is an extended generalized Γ-convolution with

$$b_F = \mu + \beta a_G + \int_{-\infty}^{+\infty} \frac{y}{1+y^2}\, U_F(\mathrm{d}y), \quad c_F = a_G, \quad U_F(y) = \bar{U}(y).$$

To complete the proof we must verify conditions (1.27) and show that the summand $\int_{-\infty}^{+\infty} \frac{y}{1+y^2}\, U_F(\mathrm{d}y)$ of b_F is finite and admits the desired representation. The definition of U_F implies

$$\begin{aligned}
\int_0^1 |\ln(x)|\, U_F(\mathrm{d}x) &= \int_{(|\beta|-\beta)\wedge 1}^1 |\ln(x)|\, U_F(\mathrm{d}x) \\
&= \begin{cases} \displaystyle\int_0^{\frac{1}{2}+\beta} \left|\ln\left(\sqrt{2y+\beta^2}-\beta\right)\right|\, U_G(\mathrm{d}y), & \beta > -\frac{1}{2}, \\ 0, & \beta \le -\frac{1}{2}. \end{cases}
\end{aligned}$$

If $\beta > -\frac{1}{2}$ and $y \in \left[0, \frac{1}{2}+\beta\right]$, then $1 \ge \sqrt{2y+\beta^2}-\beta \ge \frac{1-|\beta|+\beta}{\beta+1/2}\, y + |\beta| - \beta$, thus

$$\int_0^{\frac{1}{2}+\beta} \left|\ln\left(\sqrt{2y+\beta^2}-\beta\right)\right| U_G(\mathrm{d}y) \le \int_0^{\frac{1}{2}+\beta} \left|\ln\left(\tfrac{1-|\beta|+\beta}{\beta+1/2}\, y + |\beta| - \beta\right)\right| U_G(\mathrm{d}y) < \infty$$

because U_G fulfills the integrability conditions (1.24) since $G \in \Gamma_0$. Analogously it can be verified that $\int_{-1}^0 |\ln(|x|)|\, U_F(\mathrm{d}x) < \infty$. Using similar arguments we conclude, again with the help of (1.24), that

$$\int_1^\infty \frac{1}{x^2}\, U_F(\mathrm{d}x) = \int_{(|\beta|-\beta)\vee 1}^\infty \frac{1}{x^2}\, U_F(\mathrm{d}x) = \int_{0\vee(\frac{1}{2}+\beta)}^\infty \frac{1}{(\sqrt{2y+\beta^2}-\beta)^2}\, U_G(\mathrm{d}y) < \infty$$

and, with almost the same reasoning, $\int_{-\infty}^{-1} \frac{1}{x^2}\, U_F(\mathrm{d}x) < \infty$.

Since U_F is symmetric around $-\beta$ and $U_F(x) \equiv 0$ for $x \in [-|\beta|-\beta, |\beta|-\beta]$ by definition, it follows that for some $x_2 > |\beta| - \beta \ge 0$ and $x_1 = -x_2 - 2\beta < 0$

we have $-U_F(x_1) = U_F(x_2)$, consequently

$$\begin{aligned}\int_{-\infty}^{+\infty} \frac{x}{1+x^2}\, U_F(\mathrm{d}x) &= \int_{|\beta|-\beta}^{\infty} \frac{x}{1+x^2} - \frac{x+2\beta}{1+(x+2\beta)^2}\, U_F(\mathrm{d}x)\\ &= \int_{|\beta|-\beta}^{\infty} \frac{2\beta(x^2+2\beta x-1)}{(x^2+2\beta x+1)^2+4\beta^2}\, U_F(\mathrm{d}x)\\ &= \int_0^{\infty} \frac{2\beta(2y-1)}{(2y+1)^2+4\beta^2}\, U_G(\mathrm{d}y) < \infty\end{aligned}$$

where the finiteness of the last integral follows again from (1.24). □

REMARK: An alternative proof of the statement of the previous proposition can be found in Bondesson (1992, Theorem 7.3.2). It uses a different technique and therefore is much shorter, but does not provide any information about the generating triplet (b_F, c_F, U_F) of F and its connection to (a_G, U_G) we are interested in to derive the Lévy–Khintchine representations thereof.

With the help of Proposition 1.26 we now can easily derive the generating triplet (b_{GH}, c_{GH}, U_{GH}) of GH distributions using the mixture representation $GH(\lambda,\alpha,\beta,\delta,\mu) = N(\mu+\beta y, y) \circ GIG\big(\lambda,\beta,\sqrt{\alpha^2-\beta^2}\big)$ and the characteristic pairs (a_{GIG}, U_{GIG}) of the corresponding GIG distributions given in Proposition 1.23. We obtain

Corollary 1.27 *All GH distributions are extended generalized* Γ*-convolutions with generating triplets as follows:*

a) If $\delta > 0$ *and* $|\beta| < \alpha$*, then*

$$\begin{aligned}b_{GH(\lambda,\alpha,\beta,\delta,\mu)} &= \mu + \int_0^{\infty} \frac{2\beta(2y-1)}{(2y+1)^2+4\beta^2}\, U_{GIG(\lambda,\delta,\sqrt{\alpha^2-\beta^2})}(\mathrm{d}y),\\ c_{GH(\lambda,\alpha,\beta,\delta,\mu)} &= 0,\\ U_{GH(\lambda,\alpha,\beta,\delta,\mu)}(x) &= \big(\mathbb{1}_{[\alpha-\beta,\infty)}(x) - \mathbb{1}_{(-\infty,-\alpha-\beta]}(x)\big)\\ &\quad\cdot\left(\max(0,\lambda) + \delta^2 \int_{\frac{\alpha^2-\beta^2}{2}}^{\frac{x^2}{2}+\beta x} g_{|\lambda|}\big(2\delta^2 y - \delta^2(\alpha^2-\beta^2)\big)\,\mathrm{d}y\right).\end{aligned}$$

b) If $\lambda > 0$ *and* $\delta = 0$ *(Variance-Gamma limit), we have*

$$\begin{aligned}b_{VG(\lambda,\alpha,\beta,\mu)} &= \mu + \frac{2\lambda\beta(\alpha^2-\beta^2-1)}{(\alpha^2-\beta^2+1)^2+4\beta^2}, \qquad c_{VG(\lambda,\alpha,\beta,\mu)} = 0,\\ U_{VG(\lambda,\alpha,\beta,\mu)}(x) &= \lambda\big(\mathbb{1}_{[\alpha-\beta,\infty)}(x) - \mathbb{1}_{(-\infty,-\alpha-\beta]}(x)\big).\end{aligned}$$

c) If $\lambda < 0$ *and* $\alpha = \beta = 0$ *(t limiting case), then*

$$\begin{aligned}b_{t(\lambda,\delta,\mu)} &= \mu, \qquad c_{t(\lambda,\delta,\mu)} = 0,\\ U_{t(\lambda,\delta,\mu)}(x) &= \big(\mathbb{1}_{\mathbb{R}_+}(x) - \mathbb{1}_{\mathbb{R}_-}(x)\big)\int_0^{\frac{x^2}{2}} \delta^2 g_{|\lambda|}\big(2\delta^2 y\big)\,\mathrm{d}y.\end{aligned}$$

d) In the limit case with $\lambda < 0$ and $|\beta| = \alpha > 0$ we have

$$\begin{aligned} b_{GH(\lambda,\alpha,\pm\alpha,\delta,\mu)} &= \mu \pm \int_0^\infty \frac{2\alpha(2y-1)\,\delta^2}{(2y+1)^2+4\alpha^2}\, g_{|\lambda|}\big(2\delta^2 y\big)\,\mathrm{d}y, \\ c_{GH(\lambda,\alpha,\pm\alpha,\delta,\mu)} &= 0, \\ U_{GH(\lambda,\alpha,\pm\alpha,\delta,\mu)}(x) &= \big(\mathbb{1}_{[\alpha\mp\alpha,\infty)}(x) - \mathbb{1}_{(-\infty,-\alpha\mp\alpha]}(x)\big) \int_0^{\frac{x^2}{2}\pm\alpha x} \delta^2 g_{|\lambda|}\big(2\delta^2 y\big)\,\mathrm{d}y. \end{aligned}$$

PROOF: a) The representation follows almost immediately by combining Propositions 1.23 a) and 1.26. A direct application of the latter yields that the first factor of U_{GH} (which constitutes the symmetry around $-\beta$) has the following form:

$$\mathbb{1}_{[(\alpha^2-\beta^2)/2,\infty)}\big(\tfrac{x^2}{2}+\beta x\big)\,\big(\mathbb{1}_{[|\beta|-\beta,\infty)}(x) - \mathbb{1}_{(-\infty,-|\beta|-\beta]}(x)\big)$$

But since the solutions of $\frac{x^2}{2}+\beta x = \frac{\alpha^2-\beta^2}{2}$ are given by $-\alpha-\beta$ and $\alpha-\beta$, and $-\alpha-\beta < -|\beta|-\beta$ as well as $\alpha-\beta > |\beta|-\beta$ because $\alpha > |\beta|$, the above expression can be simplified to $\big(\mathbb{1}_{[\alpha-\beta,\infty)}(x) - \mathbb{1}_{(-\infty,-\alpha-\beta]}(x)\big)$.

The generating triplets (b, c, U) of the GH limit distributions could be obtained from part a) using the Continuity Theorem 1.22, but a careful determination of the limit expressions would require lengthy calculations and estimations. However, by Lemma 1.7 we know that the GH limits considered in b)–d) can also be represented as normal mean-variance mixtures, so we can apply Proposition 1.26 directly to the pairs (a, U) of the corresponding GIG limits derived in Proposition 1.23 b) and c). For the rest of the proof we follow this approach.

b) As seen on p. 21, the mixture representation of Variance-Gamma distributions is $VG(\lambda,\alpha,\beta,\mu) = N(\mu+\beta y, y)\circ G(\lambda,(\alpha^2-\beta^2)/2)$, and according to Proposition 1.23 b) the generating pair of the mixing Gamma distribution is given by $a_{G(\lambda,(\alpha^2-\beta^2)/2)} = 0$ and $U_{G(\lambda,(\alpha^2-\beta^2)/2)}(y) = \lambda\mathbb{1}_{[(\alpha^2-\beta^2)/2,\infty)}(y)$. By Proposition 1.26 we thus have $c_{VG(\lambda,\alpha,\beta,\mu)} = 0$ and

$$\begin{aligned} U_{VG(\lambda,\alpha,\beta,\mu)}(x) &= \lambda\mathbb{1}_{[(\alpha^2-\beta^2)/2,\infty)}\big(\tfrac{x^2}{2}+\beta x\big)\,\big(\mathbb{1}_{[|\beta|-\beta,\infty)}(x) - \mathbb{1}_{(-\infty,-|\beta|-\beta]}(x)\big) \\ &= \lambda\big(\mathbb{1}_{[\alpha-\beta,\infty)}(x) - \mathbb{1}_{(-\infty,-\alpha-\beta]}(x)\big) \end{aligned}$$

where the last equation follows with exactly the same arguments as in part a). With $U_{G(\lambda,(\alpha^2-\beta^2)/2)}(y) = \lambda\mathbb{1}_{[(\alpha^2-\beta^2)/2,\infty)}(y)$ we further obtain

$$\begin{aligned} b_{VG(\lambda,\alpha,\beta,\mu)} &= \mu + \int_0^\infty \frac{2\beta(2y-1)}{(2y+1)^2+4\beta^2}\, U_{G(\lambda,(\alpha^2-\beta^2)/2)}(\mathrm{d}y) \\ &= \mu + \frac{2\lambda\beta(\alpha^2-\beta^2-1)}{(\alpha^2-\beta^2+1)^2+4\beta^2}. \end{aligned}$$

c) The limiting t distributions are normal variance mixtures with an inverse Gamma distribution: $t(\lambda,\delta,\mu) = N(\mu,y)\circ iG(\lambda,\frac{\delta^2}{2})$. By Proposition 1.23 c) we have $a_{iG(\lambda,\delta^2/2)} = 0$ and $U_{iG(\lambda,\delta^2/2)}(x) = \delta^2\mathbb{1}_{[0,\infty)}(x)\int_0^x g_{|\lambda|}\big(2\delta^2 y\big)\,\mathrm{d}y$. Thus the generating triplet can be obtained from the formulas derived in part a) by inserting $\alpha=\beta=0$ and observing that $\max(0,\lambda) = 0$ in this case.

d) This case is very similar to the latter one. Here we have $GH(\lambda, \alpha, \pm\alpha, \delta, \mu) = N(\mu \pm \alpha y, y) \circ iG(\lambda, \frac{\delta^2}{2})$, so we get the generating triplet analogously as before by inserting $\beta = \pm\alpha$ and $\max(0, \lambda) = 0$ in the formulas of part a). $\square$

Observe that the function $U_{VG(\lambda,\alpha,\beta,\mu)}$ has only two jumps of height λ at $-\alpha-\beta$ and $\alpha-\beta$, and is constant elsewhere. From the definition of (extended) generalized Γ-convolutions (see pp. 29–32) we infer that these jumps correspond to the Gamma distributions $G(\lambda, \alpha-\beta)$ and $-G(\lambda, \alpha+\beta)$ (where the latter denotes a Gamma distribution on $(\mathbb{R}_-, \mathcal{B}_-)$ with density $d_{-G(\lambda,\alpha+\beta)}(x) = d_{G(\lambda,\alpha+\beta)}(-x)$) and conclude that a VG distributed random variable equals in probability the shifted difference of two independent Gamma variables. More precisely we have

Corollary 1.28 *For every VG distribution we have the decomposition*

$$VG(\lambda, \alpha, \beta, \mu) = -G(\lambda, \alpha+\beta) * G(\lambda, \alpha-\beta) * \epsilon_\mu.$$

Equivalently, let $X \sim VG(\lambda, \alpha, \beta, \mu)$ and X_1, X_2 be independent random variables with $\mathcal{L}(X_1) = G(\lambda, \alpha-\beta)$ and $\mathcal{L}(X_2) = G(\lambda, \alpha+\beta)$, then $X \stackrel{d}{=} X_1 - X_2 + \mu$.

Proof: From the proof of Proposition 1.26 we know, using $a_{G(\lambda,(\alpha^2-\beta^2)/2)} = 0$, that $b_{VG(\lambda,\alpha,\beta,\mu)} = \mu + \int_{-\infty}^{\infty} \frac{y}{1+y^2}\, U_{VG(\lambda,\alpha,\beta,\mu)}(\mathrm{d}y)$. Thus inserting the generating triplet given in Corollary 1.27 b) into the general equation (1.26) yields

$$\begin{aligned}
\phi_{VG(\lambda,\alpha,\beta,\mu)}(u) &= \exp\left[iub_{VG} - \int_{-\infty}^{+\infty} \left(\ln\left(1 - \frac{iu}{y}\right) + \frac{iuy}{1+y^2}\right) U_{VG}(\mathrm{d}y)\right] \\
&= \exp\left[iu\mu - \int_{-\infty}^{+\infty} \ln\left(1 - \frac{iu}{y}\right) U_{VG(\lambda,\alpha,\beta,\mu)}(\mathrm{d}y)\right] \\
&= \exp\left[iu\mu - \lambda \ln\left(1 + \frac{iu}{\alpha+\beta}\right) - \lambda \ln\left(1 - \frac{iu}{\alpha-\beta}\right)\right] \\
&= \left(1 + \frac{iu}{\alpha+\beta}\right)^{-\lambda} \left(1 - \frac{iu}{\alpha-\beta}\right)^{-\lambda} e^{iu\mu} \\
&= \phi_{-G(\lambda,\alpha+\beta)}(u)\, \phi_{G(\lambda,\alpha-\beta)}(u)\, \phi_{\epsilon_\mu}(u)
\end{aligned}$$

where the last line follows from Proposition 1.9 and the remark thereafter (see also p. 29). The reformulation in terms of random variables is obvious. $\square$

Remark: The fact that all GH distributions and their limits discussed above are extended generalized Γ-convolutions in particular implies that all of them are selfdecomposable. This can be shown analogously as in the proof of Proposition 1.20: Combining equations (1.29) and (1.25) one easily sees that the Lévy measure of every extended generalized Γ-convolution possesses a density that is increasing on $(-\infty, 0)$ and decreasing on $(0, \infty)$, hence the corresponding distributions are selfdecomposable by Lemma 1.4. Alternatively this can also be proven with the help of the mixture representations: Since GIG distributions and, by the Continuity Theorem 1.22, all of its weak limits are contained in Γ_0, they are selfdecomposable according to Proposition 1.20, and by Lemma 1.6 d) this property transfers to every normal mean-variance mixture generated from them.

1.6.4 Lévy–Khintchine representations of GH distributions

Similarly as in the GIG case, the Lévy–Khintchine representations of GH distributions can easily be derived from the generating triplets (b_{GH}, c_{GH}, U_{GH}) given in Corollary 1.27 and equations (1.29) and (1.25). The results are merged in the following

Proposition 1.29 *The characteristic functions of $GH(\lambda,\alpha,\beta,\delta,\mu)$-distributions can be represented as follows:*

a) If $\delta > 0$ and $|\beta| < \alpha$, then

$$\phi_{GH(\lambda,\alpha,\beta,\delta,\mu)}(u) = \exp\left[iu\mathrm{E}[GH] + \int_{-\infty}^{+\infty} \left(e^{iux} - 1 - iux\right) g_{GH(\lambda,\alpha,\beta,\delta,\mu)}(x)\,\mathrm{d}x\right]$$

where the mean $\mathrm{E}[GH]$ of $GH(\lambda,\alpha,\beta,\delta,\mu)$ is given by (1.11), and the density of the Lévy measure is

$$g_{GH(\lambda,\alpha,\beta,\delta,\mu)}(x) = \\ = \frac{e^{\beta x}}{|x|}\left(\int_0^\infty \frac{e^{-|x|\sqrt{2y+\alpha^2}}}{\pi^2 y\left[J_{|\lambda|}^2\left(\delta\sqrt{2y}\right) + Y_{|\lambda|}^2\left(\delta\sqrt{2y}\right)\right]}\,\mathrm{d}y + \max(0,\lambda)\,e^{-\alpha|x|}\right). \quad (1.33)$$

b) If $\lambda > 0$ and $\delta = 0$ (Variance-Gamma limit), we have

$$\phi_{VG(\lambda,\alpha,\beta,\mu)}(u) = \exp\left[iu\mathrm{E}[VG] + \int_{-\infty}^{+\infty} \left(e^{iux} - 1 - iux\right) g_{VG(\lambda,\alpha,\beta,\mu)}(x)\,\mathrm{d}x\right].$$

The mean $\mathrm{E}[VG]$ of $VG(\lambda,\alpha,\beta,\mu)$ is given by (1.16), and the Lévy density is

$$g_{VG(\lambda,\alpha,\beta,\mu)}(x) = \frac{\lambda}{|x|}\,e^{\beta x - \alpha|x|}. \quad (1.34)$$

c) If $\lambda < 0$ and $\alpha = \beta = 0$ (t limiting case), then

$$\phi_{t(\lambda,\delta,\mu)}(u) = \exp\left(iu\mu + \int_{-\infty}^{+\infty} \left(e^{iux} - 1 - iux\mathbb{1}_{[-1,1]}(x)\right) g_{t(\lambda,\delta,\mu)}(x)\,\mathrm{d}x\right),$$

and the density of the Lévy measure is

$$g_{t(\lambda,\delta,\mu)}(x) = \frac{1}{|x|}\int_0^\infty \frac{e^{-|x|\sqrt{2y}}}{\pi^2 y\left[J_{|\lambda|}^2\left(\delta\sqrt{2y}\right) + Y_{|\lambda|}^2\left(\delta\sqrt{2y}\right)\right]}\,\mathrm{d}y\,. \quad (1.35)$$

d) In the limit case with $\lambda < 0$ and $|\beta| = \alpha > 0$ we have

$$\phi_{GH(\lambda,\alpha,\pm\alpha,\delta,\mu)}(u) = \\ = \exp\left[iu\left(\mu + \int_{-\infty}^{+\infty}\left(\frac{1-e^{-|y|}}{y} - \frac{y}{1+y^2}\right) U_{GH(\lambda,\alpha,\pm\alpha,\delta,\mu)}(\mathrm{d}y)\right) \right. \\ \left. + \int_{-\infty}^{+\infty} \left(e^{iux} - 1 - iux\mathbb{1}_{[-1,1]}(x)\right) g_{GH(\lambda,\alpha,\pm\alpha,\delta,\mu)}(x)\,\mathrm{d}x\right],$$

and the Lévy density is obtained from (1.33) by inserting $\beta = \pm\alpha$.

Proof: a) It follows from (1.25) and (1.29) that the parameter c within the representation (1.26) of the characteristic function of an extended generalized Γ-convolution equals the Gaussian coefficient c of the Lévy–Khintchine formula (which justifies to use the same letter in both cases). By Corollary 1.27 a) we have $c_{GH(\lambda,\alpha,\beta,\delta,\mu)} = 0$, consequently the Gaussian part of the corresponding Lévy–Khintchine representation vanishes. Moreover, a truncation function within the integral term of the Lévy–Khintchine representation can be omitted since GH distributions possess moments of arbitrary orders, and the drift term then is given by the mean of the distribution (see the remark on p. 30). Hence it only remains to show that the density of the Lévy measure has the desired form. Equations (1.25) and (1.29) imply

$$g_{GH}(x) = -\mathbb{1}_{(-\infty,0)}(x)\,\frac{1}{x}\int_{-\infty}^{0} e^{-xy}\,U_{GH}(\mathrm{d}y) + \mathbb{1}_{(0,\infty)}(x)\,\frac{1}{x}\int_{0}^{\infty} e^{-xy}\,U_{GH}(\mathrm{d}y)$$

with U_{GH} as given in Corollary 1.27 a). The measure induced by U_{GH} has two point masses of size $\max(0,\lambda)$ at $-\alpha-\beta$ and $\alpha-\beta$ and the density $\mathrm{sign}(x)U'_{GH(\lambda,\alpha,\beta,\delta,\mu)}(x) = \mathrm{sign}(x)\delta^2 g_{|\lambda|}\big(\delta^2x^2 + 2\delta^2\beta x - \delta^2(\alpha^2-\beta^2)\big)(x+\beta)$ on $\mathbb{R}\setminus[-\alpha-\beta,\alpha-\beta]$, thus we get

$$\begin{aligned} g_{GH(\lambda,\alpha,\beta,\delta,\mu)}(x) &= \\ = -\mathbb{1}_{(-\infty,0)}(x)\,\frac{1}{x}\Bigg(&\max(0,\lambda)\,e^{-(-\alpha-\beta)x} \\ &- \int_{-\infty}^{-\alpha-\beta} \delta^2 e^{-xy}(y+\beta)g_{|\lambda|}\big(\delta^2y^2 + 2\delta^2\beta y - \delta^2(\alpha^2-\beta^2)\big)\,\mathrm{d}y\Bigg) \\ +\mathbb{1}_{(0,\infty)}(x)\,\frac{1}{x}\Bigg(&\max(0,\lambda)\,e^{-(\alpha-\beta)x} \\ &+ \int_{\alpha-\beta}^{\infty} \delta^2 e^{-xy}(y+\beta)g_{|\lambda|}\big(\delta^2y^2 + 2\delta^2\beta y - \delta^2(\alpha^2-\beta^2)\big)\,\mathrm{d}y\Bigg), \end{aligned}$$

and with the substitution $z = \frac{y^2}{2} + \beta y$ we finally obtain

$$\begin{aligned} g_{GH(\lambda,\alpha,\beta,\delta,\mu)}(x) &= \\ = \mathbb{1}_{(-\infty,0)}(x)\,\frac{e^{\beta x}}{|x|}\Bigg(&\max(0,\lambda)\,e^{-\alpha|x|} \\ &- \int_{\infty}^{\frac{\alpha^2-\beta^2}{2}} \delta^2 e^{x\sqrt{2z+\beta^2}} g_{|\lambda|}\big(2\delta^2z - \delta^2(\alpha^2-\beta^2)\big)\,\mathrm{d}z\Bigg) \\ +\mathbb{1}_{(0,\infty)}(x)\,\frac{e^{\beta x}}{|x|}\Bigg(&\max(0,\lambda)\,e^{-\alpha|x|} \\ &+ \int_{\frac{\alpha^2-\beta^2}{2}}^{\infty} \delta^2 e^{-x\sqrt{2z+\beta^2}} g_{|\lambda|}\big(2\delta^2z - \delta^2(\alpha^2-\beta^2)\big)\,\mathrm{d}z\Bigg) \end{aligned}$$

$$= \frac{e^{\beta x}}{|x|}\left(\int_0^\infty \delta^2 e^{-|x|\sqrt{2y+\alpha^2}} g_{|\lambda|}\big(2\delta^2 y\big)\,\mathrm{d}y + \max(0,\lambda)\,e^{-\alpha|x|}\right)$$
$$= \frac{e^{\beta x}}{|x|}\left(\int_0^\infty \frac{e^{-|x|\sqrt{2y+\alpha^2}}}{\pi^2 y\big[J^2_{|\lambda|}\big(\delta\sqrt{2y}\big)+Y^2_{|\lambda|}\big(\delta\sqrt{2y}\big)\big]}\,\mathrm{d}y + \max(0,\lambda)\,e^{-\alpha|x|}\right).$$

b) Similarly as before, VG distributions possess arbitrary moments, and by Corollary 1.27 b) $c_{VG(\lambda,\alpha,\beta,\mu)} = 0$, so again it only remains to verify the formula for the Lévy density. With $U_{VG(\lambda,\alpha,\beta,\delta,\mu)}(x) = \lambda\big(\mathbb{1}_{[\alpha-\beta,\infty)}(x) - \mathbb{1}_{(-\infty,-\alpha-\beta]}(x)\big)$, it is found to be

$$g_{VG(\lambda,\alpha,\beta,\mu)}(x) =$$
$$= -\mathbb{1}_{(-\infty,0)}(x)\,\frac{1}{x}\int_{-\infty}^0 e^{-xy}\,U_{VG}(\mathrm{d}y) + \mathbb{1}_{(0,\infty)}(x)\,\frac{1}{x}\int_0^\infty e^{-xy}\,U_{VG}(\mathrm{d}y)$$
$$= -\mathbb{1}_{(-\infty,0)}(x)\,\frac{\lambda}{x}\,e^{-(-\alpha-\beta)x} + \mathbb{1}_{(0,\infty)}(x)\,\frac{\lambda}{x}\,e^{-(\alpha-\beta)x}$$
$$= \frac{\lambda}{|x|}\,e^{\beta x-\alpha|x|}.$$

c) Again the Gaussian part within the Lévy–Khintchine representation vanishes because $c_{t(\lambda,\delta,\mu)} = 0$ by Corollary 1.27 c). However, the t distributions have finite means iff $\lambda < -\frac{1}{2}$, so in general the truncation function within the integral term of the Lévy–Khintchine formula is indispensable, and the drift coefficient has to be determined according to (1.25). Together with (1.28) and (1.29) we obtain

$$\begin{aligned}
b = \mu &+ \int_0^\infty \left[\int_0^\infty \frac{e^{-xy}}{1+x^2}\,\mathrm{d}x - \frac{y}{1+y^2}\right] U_{t(\lambda,\delta,\mu)}(\mathrm{d}y)\\
&- \int_{-\infty}^0 \left[\int_{-\infty}^0 \frac{e^{-xy}}{1+x^2}\,\mathrm{d}x + \frac{y}{1+y^2}\right] U_{t(\lambda,\delta,\mu)}(\mathrm{d}y)\\
&+ \int_0^\infty \left(x\mathbb{1}_{[0,1]}(x) - \frac{1}{x}\,\mathbb{1}_{\mathbb{R}\setminus[0,1]}(x)\right)\frac{x}{1+x^2}\int_0^\infty e^{-xy}\,U_{t(\lambda,\delta,\mu)}(\mathrm{d}y)\,\mathrm{d}x\\
&- \int_{-\infty}^0 \left(x\mathbb{1}_{[-1,0]}(x) - \frac{1}{x}\,\mathbb{1}_{\mathbb{R}\setminus[-1,0]}(x)\right)\frac{x}{1+x^2}\int_{-\infty}^0 e^{-xy}\,U_{t(\lambda,\delta,\mu)}(\mathrm{d}y)\,\mathrm{d}x\\
= \mu &+ \int_0^\infty \left(\int_0^1 e^{-xy}\,\mathrm{d}x - \frac{y}{1+y^2}\right) U_{t(\lambda,\delta,\mu)}(\mathrm{d}y)\\
&- \int_{-\infty}^0 \left(\int_{-1}^0 e^{-xy}\,\mathrm{d}x + \frac{y}{1+y^2}\right) U_{t(\lambda,\delta,\mu)}(\mathrm{d}y)\\
= \mu &+ \int_{-\infty}^{+\infty} \left(\frac{1-e^{-|y|}}{y} - \frac{y}{1+y^2}\right) U_{t(\lambda,\delta,\mu)}(\mathrm{d}y) = \mu,
\end{aligned}$$

because the measure induced by $U_{t(\lambda,\delta,\mu)}$ is symmetric around the origin (see Corollary 1.27 c)) but the integrand is antisymmetric. The density of the corresponding Lévy measure can be derived in exactly the same way as in part a).

d) This case is very similar to the previous one: By Corollary 1.27 d) we have $c_{GH(\lambda,\alpha,\pm\alpha,\delta,\mu)} = 0$, and the Lévy density can again be derived analogously as in part a). In general, the truncation function within the integral term of the Lévy–Khintchine formula cannot be omitted either because the limit distributions have finite first moments only if $\lambda < -1$. The drift term of the Lévy–Khintchine representation is obtained from almost the same calculation as in c) if one replaces $U_{t(\lambda,\delta,\mu)}$ by $U_{GH(\lambda,\alpha,\pm\alpha,\delta,\mu)}$, but because of the asymmetry of the latter function one does not get closed expressions in the end. □

REMARK: The Lévy measure of a Gamma distribution $G(\lambda,\sigma)$ has the density $g_{G(\lambda,\sigma)}(x) = \frac{\lambda}{x}\, e^{-\sigma x}$ (see Proposition 1.24 b)), so equation (1.34) shows that g_{VG} is just the sum of the Lévy densities $g_{-G(\lambda,-\alpha-\beta)}$ and $g_{G(\lambda,\alpha-\beta)}$. This agrees with (and is in fact equivalent to) the statement of Corollary 1.28.

Also note that in the t limiting case the truncation function within the integral of the Lévy–Khintchine representation can be omitted without further changes if $\lambda < -\frac{1}{2}$, because the Lévy measure is symmetric around the origin. In the Student's t limiting case ($\delta^2 = -2\lambda = f$), we can rewrite (1.35) in the following form:

$$g_{t(f,\mu)}(x) = \frac{1}{|x|}\int_0^\infty \frac{e^{-|x|\sqrt{2y}}}{\pi^2 y\big[J^2_{f/2}\big(\sqrt{2fy}\big) + Y^2_{f/2}\big(\sqrt{2fy}\big)\big]}\,\mathrm{d}y,$$

which is the density of the Lévy measure of a Student's t-distribution with f degrees of freedom.

The asymptotics of the Lévy densities near the origin can be derived similarly as in the GIG case. By equation (1.33) we have

$$g_{GH(\lambda,\alpha,\beta,\delta,\mu)}(x) = \frac{e^{\beta x}}{|x|}\left(\int_0^\infty \delta^2 e^{-|x|\sqrt{2y+\alpha^2}} g_{|\lambda|}\big(2\delta^2 y\big)\,\mathrm{d}y + \max(0,\lambda)\, e^{-\alpha|x|}\right),$$

which remains also valid for all limit distributions with negative λ as seen above. Applying the substitution $z = \sqrt{2y+\alpha^2} - \alpha$ we get

$$g_{GH}(x) = \frac{e^{\beta x-\alpha|x|}}{|x|}\left(\int_0^\infty e^{-|x|z}\delta^2(z+\alpha) g_{|\lambda|}\big(\delta^2(z^2+2\alpha z)\big)\,\mathrm{d}z + \max(0,\lambda)\right).$$

Clearly, the integral term now is a Laplace transform, hence its asymptotic behaviour can also be determined with the help of Theorem 1.25. On page 40 we already saw that $\delta^2 g_{|\lambda|}(2\delta^2 y) \sim \delta(2\pi^2 y)^{-\frac{1}{2}}$, $y \to \infty$. Replacing y by $\frac{z^2}{2}$ and multiplying the result by z we get $\delta^2(z+\alpha)g_{|\lambda|}(\delta^2(z^2+2\alpha z)) \sim \frac{\delta}{\pi}$, $z \to \infty$, so the assumptions on v in Theorem 1.25 are fulfilled with $\rho = 1$ and $L(y) \equiv \frac{\delta}{\pi}$. Consequently the asymptotic behaviour of the Laplace transform near the origin is given by $\frac{\delta}{\pi}|x|^{-1}$, and for the preceeding factor obviously holds $\frac{e^{\beta x-\alpha|x|}}{|x|} \sim |x|^{-1}$, $x \to 0$, so altogether we have

$$g_{GH(\lambda,\alpha,\beta,\delta,\mu)}(x) \sim \frac{\delta}{\pi}\,x^{-2}, \qquad x \to 0, \tag{1.36}$$

including the limits $GH(\lambda,0,0,\delta,\mu) = t(\lambda,\delta,\mu)$ and $GH(\lambda,\alpha,\pm\alpha,\delta,\mu)$ with $\lambda < 0$. Only the Lévy densities of VG distributions behave differently, namely

$$g_{VG(\lambda,\alpha,\beta,\mu)}(x) = \frac{\lambda}{|x|}\,e^{\beta x-\alpha|x|} \sim \frac{\lambda}{|x|}, \qquad x \to 0.$$

Remark: Apart from the VG limiting case, the asymptotic behaviour of the Lévy densities implies $\int_{[-1,1]} |x|\, \nu_{GH}(\mathrm{d}x) = \int_{[-1,1]} |x|\, g_{GH}(x)\, \mathrm{d}x = \infty$, consequently the sample paths $X_t(\omega)$ of all Lévy processes induced by GH distributions and their limits with $\lambda < 0$ considered above almost surely have infinite variation on $(0,t]$ for every $t > 0$ (Sato (1999, Theorem 21.9)). The sample paths of a VG Lévy process, however, have finite variation on every interval $(0,t]$ almost surely because $\int_{[-1,1]} |x|\, \nu_{VG}(\mathrm{d}x) = \int_{[-1,1]} \lambda e^{\beta x - \alpha|x|}\, \mathrm{d}x < \infty$. The latter can alternatively be deduced from Corollary 1.28: it implies that a VG process equals in law the difference of two Gamma processes which have increasing paths almost surely (see, for example, the proof of Proposition 1.8).

One may ask if the above method could be exploited further to obtain higher-order asymptotics of the Lévy densities around the origin. Unfortunately this is not the case. Equations (A.16) and (A.17) imply that $\delta^2(z+\alpha)g_{|\lambda|}(\delta^2(z^2+2\alpha z))$ has an asymptotic expansion of the form $\frac{\delta}{\pi} + \sum_{n\geq 1} a_n z^{-n}$ for $z \to \infty$, but powers x^r with $r \leq -1$ do not provide any information about the behaviour of the Laplace transform at the origin since the bound $\rho > 0$ in Theorem 1.25 cannot be lowered. Thus higher-order terms can only be obtained with different approaches.

In Raible (2000, Proposition 2.18), the second-order terms have been derived with the help of the Fourier transform of the modified Lévy measure $\tilde{\nu}_{GH}(\mathrm{d}x) = x^2 \nu_{GH}(\mathrm{d}x)$. For GH distributions with $0 \leq |\beta| < \alpha$ and $\delta > 0$, he found the asymptotic behaviour

$$g_{GH(\lambda,\alpha,\beta,\delta,\mu)}(x) = \frac{\delta}{\pi}\, x^{-2} + \frac{\lambda + \frac{1}{2}}{2}\, |x|^{-1} + \frac{\delta\beta}{\pi}\, x^{-1} + o\big(|x|^{-1}\big), \qquad x \to 0.$$

Since the Lévy densities $g_{t(\lambda,\delta,\mu)}$ and $g_{GH(\lambda,\alpha,\pm\alpha,\delta,\mu)}$ equal the pointwise limits of $g_{GH(\lambda,\alpha,\beta,\delta,\mu)}$ for $\alpha, \beta \to 0$ and $|\beta| \to \alpha$, respectively, it is tempting to infer that the above asymptotics are also valid for these limiting cases, but in general this is not necessarily true: A crucial assumption in the proof of Raible (2000, Proposition 2.18) is that the modified Lévy measure $\tilde{\nu}_{GH}(\mathrm{d}x)$ is a finite measure on $(\mathbb{R}, \mathcal{B})$, that is, $\int_{\mathbb{R}} x^2 g_{GH}(x)\, \mathrm{d}x < \infty$. By the remark on p. 30, this is equivalent to the existence of finite second moments of the corresponding GH distribution, but as we have seen before (cf. pp. 23 and 25), for the limit distributions $t(\lambda,\delta,\mu)$ and $GH(\lambda,\alpha,\pm\alpha,\delta,\mu)$ this imposes an additional constraint on λ, so we arrive at the following

Conjecture 1.30

a) For every t distribution $t(\lambda,\delta,\mu)$ with $\lambda < -1$, the asymptotic behaviour of the corresponding Lévy density near the origin is given by

$$g_{t(\lambda,\delta,\mu)}(x) = \frac{\delta}{\pi}\, x^{-2} + \frac{\lambda + \frac{1}{2}}{2}\, |x|^{-1} + o\big(|x|^{-1}\big), \qquad x \to 0.$$

b) For every $GH(\lambda,\alpha,\pm\alpha,\delta,\mu)$-distribution with $\lambda < -2$, the asymptotics of the corresponding Lévy densitiy are

$$g_{GH(\lambda,\alpha,\pm\alpha,\delta,\mu)}(x) = \frac{\delta}{\pi}\, x^{-2} + \frac{\lambda + \frac{1}{2}}{2}\, |x|^{-1} \pm \frac{\delta\alpha}{\pi}\, x^{-1} + o\big(|x|^{-1}\big), \qquad x \to 0.$$

The following example shows that in some cases the asymptotic behaviour of the Lévy density is still correctly described by the conjecture above although the constraints on λ are not fulfilled, so it might be possible that the latter could either be weakened or be replaced by some other condition.

By (1.35), the Lévy densitiy of a Cauchy distribution $t(-\frac{1}{2}, \delta, \mu)$ is given by

$$g_{t(-\frac{1}{2},\delta,\mu)}(x) = \frac{1}{|x|}\int_0^\infty \frac{e^{-|x|\sqrt{2y}}}{\pi^2 y\big[J^2_{\frac{1}{2}}\big(\delta\sqrt{2y}\big) + Y^2_{\frac{1}{2}}\big(\delta\sqrt{2y}\big)\big]}\,\mathrm{d}y,$$

and from equation (A.12) it follows that $J^2_{\frac{1}{2}}\big(\delta\sqrt{2y}\big) + Y^2_{\frac{1}{2}}\big(\delta\sqrt{2y}\big) = \frac{2}{\pi\delta\sqrt{2y}}$, hence

$$g_{t(-\frac{1}{2},\delta,\mu)}(x) = \frac{\delta}{\pi|x|}\int_0^\infty \frac{e^{-|x|\sqrt{2y}}}{\sqrt{2y}}\,\mathrm{d}y \underset{z=\sqrt{2y}}{=} \frac{\delta}{\pi|x|}\int_0^\infty e^{-|x|z}\,\mathrm{d}z = \frac{\delta}{\pi}\,x^{-2},$$

so part a) of Conjecture 1.30 still applies, but $\lambda = -\frac{1}{2} > -1$. Finally we take a closer look at the NIG distributions from which the Cauchy distributions arise as weak limits if the parameters α and β both tend to zero. In this case we have

$$\begin{aligned} g_{NIG(\alpha,\beta,\delta,\mu)}(x) &= \frac{e^{\beta x}}{|x|}\int_0^\infty \frac{e^{-|x|\sqrt{2y+\alpha^2}}}{\pi^2 y\big[J^2_{\frac{1}{2}}\big(\delta\sqrt{2y}\big) + Y^2_{\frac{1}{2}}\big(\delta\sqrt{2y}\big)\big]}\,\mathrm{d}y \\ &= \frac{\delta e^{\beta x}}{\pi|x|}\int_0^\infty \frac{e^{-|x|\sqrt{2y+\alpha^2}}}{\sqrt{2y}}\,\mathrm{d}y = \frac{\delta e^{\beta x}}{\pi|x|}\int_0^\infty e^{-|x|\sqrt{z^2+\alpha^2}}\,\mathrm{d}z \\ &= \frac{\delta\alpha}{\pi|x|}\,e^{\beta x}K_1\big(\alpha|x|\big), \end{aligned}$$

where the last equality follows from the fact that $e^{-|x|\sqrt{z^2+\alpha^2}}$ equals a non-normalized density of a symmetric hyperbolic distribution with parameters $\alpha = |x|$, $\beta = 0$, $\delta = \alpha$ and $\mu = 0$, consequently the value of the integral in the last but one line must be $\frac{1}{2}\,a(1, |x|, 0, \alpha, 0)^{-1}$ where $a(1, |x|, 0, \alpha, 0) = \frac{1}{2\alpha K_1(\alpha|x|)}$ is the norming constant of a $HYP(|x|, 0, \alpha, 0)$-distribution (see p. 14). By equation (A.8) $K_1(y) \sim \frac{1}{y}$ for $y \downarrow 0$, thus

$$\lim_{\alpha,\beta\to 0} g_{NIG(\alpha,\beta,\delta,\mu)}(x) = \lim_{\alpha,\beta\to 0}\frac{\delta\alpha}{\pi|x|}\,e^{\beta x}K_1\big(\alpha|x|\big) = \frac{\delta}{\pi}\,x^{-2} = g_{t(-\frac{1}{2},\delta,\mu)}(x),$$

so the Lévy density of a Cauchy distribution can in fact be obtained as pointwise limit of the Lévy density of an NIG distribution.

1.7 Approximations based on Gamma and Normal variables

At the beginning of the present chapter we pointed out that exponential Lévy processes $S_t = S_0 e^{L_t}$ provide a flexible and accurate model for asset prices. A central application of such models is the pricing of options and other derivatives. We are not going to discuss option pricing theory and methods in greater detail here (an overview of option pricing in Lévy models with many references can be obtained from Schoutens (2006)), but simply want to make the following point: If the option to be considered is of European type (that is, it can only be exercised at some fixed future date T) and its payoff only depends on the asset price S_T at the maturity time T, then it is usually possible to derive an explicit formula for the option price which can be evaluated numerically in an efficient way. If, however, the payoff of the option depends on the whole path $(S_t)_{0\leq t\leq T}$ and/or the option is of American type and thus can be exercised at an arbitrary point of time between the present date $t=0$ and the maturity date T, closed-form solutions of the option pricing problem typically do not exist. In these cases a fair option price can only be determined by either trying to solve the associated partial (integro) differential equation numerically (if an appropriate algorithm for this purpose is known at all) or by Monte Carlo-methods. The main task of the latter approach is to find an approximation scheme for the driving Lévy process L which enables a reasonable fast and accurate generation of sample paths $(L_t^n)_{0\leq t\leq T}$ and hence $(S_t^n)_{0\leq t\leq T}$ which converge in law to $(L_t)_{0\leq t\leq T}$ and $(S_t)_{0\leq t\leq T}$, respectively.

Remark: We shall use the notation $\xrightarrow{\mathcal{L}}$ to indicate weak convergence of the laws of real valued random variables as well as of laws of stochastic processes. In the latter case $(L_t^n)_{t\geq 0} \xrightarrow{\mathcal{L}} (L_t)_{t\geq 0}$ has to be understood in the sense of Jacod and Shiryaev (2003, p. 349), that is, as weak convergence of $\mathcal{L}(L^n)$ to $\mathcal{L}(L)$ in $\mathscr{P}(\mathbb{D}(\mathbb{R}))$, where $\mathscr{P}(\mathbb{D}(\mathbb{R}))$ denotes the space of all probability measures on the Skorokhod space $\mathbb{D}(\mathbb{R})$ equipped with the Skorokhod topology.

For the classical model where $(S_t)_{t\geq 0}$ is a geometric Brownian motion and $L_t = \sigma B_t + \big(r - \frac{\sigma^2}{2}\big)t$ is a Brownian motion with drift, a very simple approximation scheme using only Bernoulli variables exists: Suppose ξ_{ni}, $1\leq i\leq n$, are iid with $Q\big(\xi_{n1} = \frac{\sigma}{\sqrt{n}}\big) = \frac{1}{2} + \frac{1}{2}\,\frac{r-\frac{\sigma^2}{2}}{\sigma\sqrt{n}} = 1 - Q\big(\xi_{n1} = -\frac{\sigma}{\sqrt{n}}\big)$, then $\sum_{i=1}^n \xi_{ni} \xrightarrow{\mathcal{L}} L_1$ for $n\to\infty$. With slight modifications it is also possible to approximate the process $(L_t)_{0\leq t\leq T}$ on the whole time interval. Take iid Bernoulli variables $\bar{\xi}_{ni}$, $1\leq i\leq k_n = [Tn]$, with $Q\big(\bar{\xi}_{n1} = \frac{\sigma\sqrt{T}}{\sqrt{n}}\big) = \frac{1}{2} + \frac{(r-\frac{\sigma^2}{2})\sqrt{T}}{2\sigma\sqrt{n}} = 1 - Q\big(\bar{\xi}_{n1} = -\frac{\sigma\sqrt{T}}{\sqrt{n}}\big)$, and define $L_t^n := \sum_{i=1}^{[k_n t]} \bar{\xi}_{ni}$, then $(L_t^n)_{0\leq t\leq T} \xrightarrow{\mathcal{L}} (L_t)_{0\leq t\leq T}$ for $n\to\infty$. This approach and its application to option pricing goes back to Cox, Ross, and Rubinstein (1979). In the following we present some possible approximations of GH distributions and Lévy processes.

Remark: If the sample paths $(L_t^n)_{0\leq t\leq T}$ or $(S_t^n)_{0\leq t\leq T}$ are generated in order to derive some option prices thereof, one has to be aware that all simulations have to be performed under a risk neutral martingale measure Q. To emphasize this fact, we used Q in the notation above and also changed the drift of the Brownian motion from μ to r accordingly. Contrary to the Brownian world, in models driven by general Lévy processes the risk neutral measure is typically not unique (in fact, the class of equivalent martingale measures can be very large, as Eberlein and Jacod (1997) have shown), and it is a priori not clear if the driving process L remains in the desired model class under a corresponding measure change. However, here we do not necessarily need to care about this, because Raible (2000, Corollary 2.29) has shown that assuming L is a GH Lévy process under both the real world measure P and the risk neutral measure Q imposes no essential restriction.

For the rest of the section we suppose that $(L_t)_{t\geq 0}$ is a Lévy process induced by a $GH(\lambda, \alpha, \beta, \delta, \mu)$-distribution (including the limits with finite parameters). Our first aim is to define a uan triangular scheme $(X_{ni})_{1\leq i\leq k_n, n\geq 1}$ with $\sum_{i=1}^{k_n} X_{ni} \xrightarrow{\mathcal{L}} L_1$ for $n \to \infty$, where uan means uniformly asymptotically negligible, that is, $\lim_{n\to\infty} \sup_{1\leq i\leq k_n} P(|X_{ni}| > \epsilon) = 0$ for all $\epsilon > 0$. For practical purposes, one may wish to keep the triangular scheme as simple as possible and thus ask which minimal requirements the X_{ni} have to fulfill in any case. If we denote the distribution function of X_{ni} by F_{ni}, necessary conditions for the desired convergence of $\sum_{i=1}^{k_n} X_{ni}$ are

$$\begin{aligned} \sum_{i=1}^{k_n} F_{ni}(y) \underset{n\to\infty}{\longrightarrow} \nu_{GH}\big((-\infty, y]\big) &= \int_{-\infty}^{y} g_{GH}(x)\,\mathrm{d}x, \qquad y < 0, \\ \sum_{i=1}^{k_n} \big(1 - F_{ni}(y)\big) \underset{n\to\infty}{\longrightarrow} \nu_{GH}\big([y, \infty)\big) &= \int_{y}^{\infty} g_{GH}(x)\,\mathrm{d}x, \qquad y > 0. \end{aligned} \tag{1.37}$$

(see for example Loève (1977, p. 323)). The formulas for the Lévy densities derived in the previous section imply that g_{GH} is strictly positive on the whole real line, so we conclude from the above that the joint range $\bigcup_{i=1}^{k_n} X_{ni}(\Omega)$ of the X_{ni} necessarily must be, at least in the limit, a dense subset of $\mathbb{R}$.

Remark: If the X_{ni} are iid and $\sum_{i=1}^{k_n} X_{ni} \xrightarrow{\mathcal{L}} L_1$, then $(L_t)_{0\leq t\leq 1}$ can easily be approximated by $L_t^n := \sum_{i=1}^{[k_n t]} X_{ni}$, $0 \leq t \leq 1$, because in this case we have

$$\phi_{L_t^n}(u) = \phi_{X_{n1}}(u)^{[k_n t]} = \big(\phi_{X_{n1}}^{k_n}(u)\big)^{\frac{[k_n t]}{k_n}},$$

and $\phi_{X_{n1}}^{k_n}(u) \to \phi_{L_1}(u)$ for every $u \in \mathbb{R}$. Moreover, $\frac{[k_n t]}{k_n} \to t$ uniformly on $[0, 1]$, so it follows from the equation above that for arbitrarily fixed u the convergence $\phi_{L_t^n}(u) \to \phi_{L_t}(u) = \phi_{L_1}(u)^t$ is uniform in $t \in [0, 1]$, and Jacod and Shiryaev (2003, Chapter VII, Corollary 4.43) then assures that $(L_t^n)_{0\leq t\leq 1} \xrightarrow{\mathcal{L}} (L_t)_{0\leq t\leq 1}$. The approximation can readily be extended from the time intervall $[0, 1]$ to $[0, T]$ by using iid copies $L^{n,j}$ of L^n to simulate the increments $(L_t - L_j)_{j\leq t\leq j+1}$ where $1 \leq j < T$. Alternatively, if one can find iid random variables $\bar{X}_{ni}$ with $\sum_{i=1}^{k_n'} \bar{X}_{ni} \xrightarrow{\mathcal{L}} L_T$ and defines $L_t^n := \sum_{i=1}^{[k_n' t/T]} \bar{X}_{ni}$, $0 \leq t \leq T$, then a completely analogous reasoning as before yields $(L_t^n)_{0\leq t\leq T} \xrightarrow{\mathcal{L}} (L_t)_{0\leq t\leq T}$.

Equation (1.37) suggests that a convergent triangular scheme with iid random variables $(X_{ni})_{1\leq i\leq k_n, n\geq 1}$ may be obtained from a suitable discretization of the Lévy measure ν_{GH}. This approach was used by Maller, Solomon, and Szimayer (2006) to derive American option prices from the simulated sample paths. They choose sequences $m_n^+, m_n^- \geq 1$ and Δ_n with $m_n^\pm \to \infty$, $\Delta_n \downarrow 0$ and

$\lim_{n\to\infty} m_n^{\pm}\Delta_n > 0$. The random variables X_{ni} are then defined on an equidistant grid by $P(X_{ni} = k\Delta_n) = \frac{1}{k_n}\,\nu_{GH}\big(((k-\frac{1}{2})\Delta_n, (k+\frac{1}{2})\Delta_n]\big)$, where $-m_n^- \leq k \leq m_n^+$, $k \neq 0$, and $P(X_{ni} = 0) = 1 - \sum_{k=-m_n^-,\,k\neq 0}^{m_n^+} P(X_{ni} = k\Delta_n)$. To ensure that the X_{ni} have a proper probability distribution, one has to make the technical assumption $\liminf_{n\to\infty} \sqrt{k_n}\Delta_n > 0$. This essentially amounts to an approximation of the Lévy measure ν_{GH} on $\big(-(m_n^- + \frac{1}{2})\Delta_n, (m_n^+ + \frac{1}{2})\Delta_n\big] \setminus \big(-\frac{\Delta_n}{2}, \frac{\Delta_n}{2}\big]$ with point masses of size $\nu_{GH}\big(((k-\frac{1}{2})\Delta_n, (k+\frac{1}{2})\Delta_n]\big)$ located at $k\Delta_n$ with $-m_n^- \leq k \leq m_n^+$, $k \neq 0$. Because the definition of the X_{ni} is solely based on the truncated Lévy measure and neglects possible non-zero drift terms occuring in the Lévy–Khintchine representation, one further has to add a suitable centering sequence a_n to get the desired convergence, that is, $\sum_{i=1}^{k_n}(X_{ni} - a_n) \xrightarrow{\mathcal{L}} L_1$. If $\mathrm{E}[L_1] < \infty$, a possible choice is $a_n := \frac{\mathrm{E}[L_1]}{k_n} + \mathrm{E}\big[X_{n1}\mathbb{1}_{\{|X_{n1}|\leq 1\}}\big]$, otherwise $\mathrm{E}[L_1]$ in the preceeding equation has to be replaced by the drift term b of the Lévy–Khintchine triplet of L_1.

The results of the previous section allow for an alternative approximation scheme consisting of suitably scaled and shifted Gamma variables. Let us note before that the proof of Proposition 1.26 implies that the characteristic functions of extended generalized Γ-convolutions F arising as normal mean-variance mixtures with mixing distributions $G \in \Gamma_0$ may be represented in a simpler form as in equation (1.26) because in this case we have $b_F = \mu + \beta a_G + \int_{-\infty}^{+\infty} \frac{y}{1+y^2}\,U_F(\mathrm{d}y)$, and the integral term at the end has a finite value, hence it cancels out with the second summand under the integral occuring in (1.26). Applying this to the case of GH distributions we obtain

$$\phi_{GH}(u) = \exp\left[iu\mu - \int_{-\infty}^{+\infty} \ln\left(1 - \frac{iu}{y}\right) U_{GH}(\mathrm{d}y)\right] \tag{1.38}$$

und thus arrive at the following

Proposition 1.31 *Consider a $GH(\lambda,\alpha,\beta,\delta,\mu)$-distribution with corresponding function U_{GH} given by either of the cases considered in Corollary 1.27 and the following assumptions:*

a) For all $n \geq 1$ there exists some $K_n > \alpha - \beta$ and a partition $\alpha - \beta = x_{n1} < x_{n2} < \cdots < x_{nk_n} = K_n$ of $[\alpha - \beta, K_n]$.

b) If $\lambda > 0$, set $X_{n1} := X_{n1}^+ - X_{n1}^- + \mu$ with independent Gamma variables $X_{n1}^+ \sim G(\lambda, \alpha - \beta)$ and $X_{n1}^- \sim G(\lambda, \alpha + \beta)$, otherwise set $X_{n1} \equiv \mu$.
For $2 \leq i \leq k_n$, let $X_{ni} = X_{ni}^+ - X_{ni}^-$ be independent random variables where $X_{ni}^+ \sim G(\lambda_{ni}, \sigma_{ni}^+)$ and $X_{ni}^- \sim G(\lambda_{ni}, \sigma_{ni}^-)$ are independent Gamma variables with $\sigma_{ni}^+ = x_{ni}$, $\sigma_{ni}^- = x_{ni} + 2\beta$ and

$$\lambda_{ni} = U_{GH}(x_{ni}) - U_{GH}(x_{ni-1}) = \delta^2 \int_{\frac{x_{ni-1}^2}{2} + \beta x_{ni-1}}^{\frac{x_{ni}^2}{2} + \beta x_{ni}} g_{|\lambda|}\big(2\delta^2 y - \delta^2(\alpha^2 - \beta^2)\big)\,\mathrm{d}y.$$

If $K_n \uparrow \infty$ and $\sup_{2\leq i\leq k_n} |x_{ni} - x_{ni-1}| \to 0$ for $n \to \infty$, then $\sum_{i=1}^{k_n} X_{ni} \xrightarrow{\mathcal{L}} L_1$ where $L_1 \sim GH(\lambda,\alpha,\beta,\delta,\mu)$.

Proof: The assertion immediately follows from the Continuity Theorem 1.22 and the remarks thereafter (see also the considerations on p. 29). Note that with the above setting U_{GH} is approximated from below on $[\alpha-\beta, K_n]$ and from above on $[-K_n-2\beta, -\alpha-\beta]$ by step functions of the form

$$U_n(x) = \big(\mathbb{1}_{[\alpha-\beta,\infty)}(x) - \mathbb{1}_{(-\infty,-\alpha-\beta]}(x)\big)\cdot\max(0,\lambda) + \\ + \sum_{i=2}^{k_n}\big(U_{GH}(x_{ni}) - U_{GH}(x_{ni-1})\big)\big(\mathbb{1}_{[x_{ni},\infty)}(x) - \mathbb{1}_{(-\infty,-x_{ni}-2\beta]}(x)\big).$$

Since U_{GH} is continuous on $\mathbb{R}\setminus[-\alpha-\beta,\alpha-\beta]$, this is of course by far not the only possibility, but probably one of the simplest. □

The careful reader will also observe that the triangular scheme above is not completely uniformly asymptotically negligible because X_{n1} does in general not fulfill the uan-condition, only the X_{ni} with $i \geq 2$ do so. But since in this case all X_{ni} are already infinitely divisible themselves, the uan-property can be neglected here. What is more important, the X_{ni} are obviously not identically distributed, so the triangular scheme cannot directly be used to simulate paths of a GH Lévy process $(L_t)_{0\leq t\leq T}$ over some finite time intervall $[0,T]$, but it can easily be extended to obtain the desired properties. We first note that L_T can be approximated along exactly the same lines of Proposition 1.31 if one replaces μ by μT and U_{GH} by $U_{GH}T$. This follows immediately from (1.38) and the fact that $\phi_{L_T}(u) = \phi_{L_1}^T$. Therewith we get

Corollary 1.32 *Let $(L_t)_{t\geq 0}$ be a Lévy process with $\mathcal{L}(L_1) = GH(\lambda,\alpha,\beta,\delta,\mu)$ and corresponding function U_{GH}. Suppose that*

a) For all $n \geq 1$ there exists some $K_n > \alpha-\beta$ and a partition $\alpha-\beta = x_{n1} < x_{n2} < \cdots < x_{nm_n} = K_n$ of $[\alpha-\beta, K_n]$.

b) If $\lambda > 0$, set $Y_{n1} := Y_{n1}^+ - Y_{n1}^- + \frac{\mu T}{k_n}$ with independent Gamma variables $Y_{n1}^+ \sim G\big(\frac{\lambda T}{k_n}, \alpha-\beta\big)$ and $Y_{n1}^- \sim G\big(\frac{\lambda T}{k_n}, \alpha+\beta\big)$, otherwise set $Y_{n1} \equiv \frac{\mu T}{k_n}$. For $2 \leq j \leq m_n$, let $Y_{nj} = Y_{nj}^+ - Y_{nj}^-$ be independent random variables where $Y_{nj}^+ \sim G(\lambda_{nj}, \sigma_{nj}^+)$ and $Y_{nj}^- \sim G(\lambda_{nj}, \sigma_{nj}^-)$ are independent and $\sigma_{nj}^+ = x_{nj}$, $\sigma_{nj}^- = x_{nj} + 2\beta$ and $\lambda_{nj} = \frac{T}{k_n}\big(U_{GH}(x_{nj}) - U_{GH}(x_{nj-1})\big)$.

c) Let $(\bar{X}_{ni})_{1\leq i\leq k_n}$ be iid copies of $\bar{X}_n := \sum_{j=1}^{m_n} Y_{nj}$ and $L_t^n := \sum_{i=1}^{[k_n t/T]} \bar{X}_{ni}$, $0 \leq t \leq T$.

If $K_n, k_n \uparrow \infty$ and $\sup_{2\leq j\leq m_n}|x_{nj} - x_{nj-1}| \to 0$ for $n \to \infty$, then we have $(L_t^n)_{0\leq t\leq T} \xrightarrow{\mathcal{L}} (L_t)_{0\leq t\leq T}$.

Proof: The result is an immediate consequence of the convolution property of Gamma distributions (see Proposition 1.11 d) and the remark on p. 12), Proposition 1.31 and the remark on p. 53. □

Remark: In the VG limiting case the approximation scheme simplifies considerably because U_{VG} is constant on $[\alpha-\beta,\infty)$ by Corollary 1.27 b) and thus

$Y_{nj}^+ = Y_{nj}^- \equiv 0$ for $j \geq 2$ (observe that by Proposition 1.9 $\lim_{\lambda\to 0} \mathfrak{L}_{G(\lambda,\sigma)}(u) = \lim_{\lambda\to 0}\big(1+\frac{u}{\sigma}\big)^{-\lambda} = 1 = \mathfrak{L}_{\epsilon_0}(u)$, so we may identify a $G(0,\sigma)$-distribution with the unit mass located in the origin). If we assume that $(Y_{ni}^+)_{1\leq i\leq k_n}$ are iid with $Y_{n1}^+ \sim G\big(\frac{\lambda T}{k_n}, \alpha-\beta\big)$, $(Y_{ni}^-)_{1\leq i\leq k_n}$ are iid with $Y_{n1}^- \sim G\big(\frac{\lambda T}{k_n}, \alpha+\beta\big)$ and $k_n \uparrow \infty$, then the process

$$L_t^n := \sum_{i=1}^{[k_n t/T]} \left(Y_{ni}^+ - Y_{ni}^- + \frac{\mu T}{k_n}\right) = \sum_{i=1}^{[k_n t/T]} Y_{ni}^+ - \sum_{i=1}^{[k_n t/T]} Y_{ni}^- + \frac{[k_n \frac{t}{T}]}{k_n}\mu T$$

converges on $[0,T]$ in law to the VG process generated by $VG(\lambda,\alpha,\beta,\mu)$, and the two sums in the rightmost equation converge in law to the Gamma processes generated by $G(\lambda,\alpha-\beta)$ and $G(\lambda,\alpha+\beta)$, respectively.

Note that under the assumptions of Corollary 1.32 the range of the approximating process $(L_t^n)_{0\leq t\leq T}$ is $\mathbb{R}$, in contrast to the approach of Maller, Solomon, and Szimayer (2006) where by construction L_t^n can only take values on a finite grid. The latter of course eases the calculation of American option prices (for which the corresponding scheme was originally designed), but might be less useful in simulation based pricing of exotic options such as Asian or Lookback options where it is important to reproduce the path properties in a more realistic way. Moreover, the Gamma approximation method might in general be easier to implement than the scheme suggested by Maller et al. It does not require an additional centering sequence, and the discretization of the measure U_{GH} only involves the evaluation of a single integral (see Proposition 1.31), whereas the discretization of the Lévy measure ν_{GH} usually requires the computation of double integrals since the formulas of the Lévy densities g_{GH} already contain an integral term (only some special GH subclasses like NIG or VG admit simpler representations of their Lévy densities).

Having solved the difficulties of calculating the necessary weights and parameters, the remaining task is to simulate the random variables $\bar{X}_{ni}$. Within the framework of Maller et al. this amounts to sampling from a discrete distribution with finite support, whereas under the assumptions of Corollary 1.32 the $\bar{X}_{ni}$ are obtained as sums of finitely many Gamma variables. At first glance the latter method seems to be fairly ineffective compared to the first one since generating many random variates to finally obtain a single one appears much more time-consuming than to sample once from the desired distribution directly. But one should keep in mind that sampling from an arbitrary discrete distribution becomes more complicated and thus slower as the number of elements in the support increases. Gamma variates, however, can be simulated very fast and efficiently by rejection methods for any choice of the parameters λ and σ (see, for example, Devroye (1986, chapter IX.3) or Marsaglia and Tsang (2000) and the references therein). Therefore one could expect that the difference in performance of both approximation methods is not too large and might even decrease as n increases, but we leave a thorough investigation of numerical implementations to future research.

Proposition 1.8 and equation (1.5) imply that every GH Lévy process $(L_t)_{t\geq 0}$ admits a representation of the form $L_t = \mu t + \beta\tau(t) + B_{\tau(t)}$, where $(\tau(t))_{t\geq 0}$ is a GIG process with $\mathcal{L}\big(\tau(1)\big) = GIG\big(\lambda, \delta, \sqrt{\alpha^2 - \beta^2}\big)$ and $(B_t)_{t\geq 0}$ is a standard Brownian motion independent of τ. This provides an alternative simulation method for L using approximations τ_n of the subordinating GIG process τ: Suppose $\bar{X}_{ni}$ are iid random variables with $\sum_{i=1}^{k_n} \bar{X}_{ni} \xrightarrow{\mathcal{L}} \tau(T)$. Define $\tau_n(t) := \sum_{i=1}^{[k_n t/T]} \bar{X}_{ni}$, $0 \leq t \leq T$, and $L_t^n := \mu t + \beta\tau_n(t) + B_{\tau_n(t)}$, then we also have $(L_t^n)_{0\leq t\leq T} \xrightarrow{\mathcal{L}} (L_t)_{0\leq t\leq T}$. This can be proven as follows: Similarly to the proof of Lemma 1.6 (part b) and c)) we see that

$$\phi_{L_t^n}(u) = e^{iu\mu t}\mathfrak{L}_{\bar{X}_{n1}}\big(\tfrac{u^2}{2} - iu\beta\big)^{[k_n t/T]} = e^{iu\mu t}\Big(\mathfrak{L}_{\bar{X}_{n1}}\big(\tfrac{u^2}{2} - iu\beta\big)^{k_n}\Big)^{\frac{[k_n t/T]}{k_n}},$$

and analogously as in the remark on p. 53 we conclude $\phi_{L_t^n}(u) \to \phi_{L_t}(u)$ uniformly on $[0, T]$. Jacod and Shiryaev (2003, Chapter VII, Corollary 4.43) then again yields the desired convergence.

The approximation of GIG processes can be done in a very similar way to that of GH processes. For the sake of completeness, we summarize the result in the following corollary which can be proven along the same lines as before.

Corollary 1.33 *Let $(\tau(t))_{t\geq 0}$ be a Lévy process with $\mathcal{L}\big(\tau(1)\big) = GIG(\lambda, \delta, \gamma)$ and corresponding function U_{GIG}. Suppose that*

a) For all $n \geq 1$ there exists some $K_n > \frac{\gamma^2}{2}$ and a partition $\frac{\gamma^2}{2} = x_{n1} < x_{n2} < \cdots < x_{nm_n} = K_n$ of $\big[\frac{\gamma^2}{2}, K_n\big]$.

b) Let $Y_{n1} \sim G\big(\frac{\lambda T}{k_n}, \frac{\gamma^2}{2}\big)$ if $\lambda > 0$ and $Y_{n1} \equiv 0$ otherwise. For $2 \leq j \leq m_n$, let $Y_{nj} \sim G(\lambda_{nj}, \sigma_{nj})$ be independent Gamma variables with $\sigma_{nj} = x_{nj}$ and

$$\lambda_{nj} = \frac{T}{k_n}\big(U_{GIG}(x_{nj}) - U_{GIG}(x_{nj-1})\big) = \frac{\delta^2 T}{k_n}\int_{x_{nj-1}}^{x_{nj}} g_{|\lambda|}\big(2\delta^2 y - \delta^2\gamma^2\big)\,\mathrm{d}y.$$

c) Let $(\bar{X}_{ni})_{1\leq i\leq k_n}$ be iid copies of $\bar{X}_n := \sum_{j=1}^{m_n} Y_{nj}$ and $\tau_n(t) := \sum_{i=1}^{[k_n t/T]} \bar{X}_{ni}$, $0 \leq t \leq T$.

If $K_n, k_n \uparrow \infty$ and $\sup_{2\leq j\leq m_n} |x_{nj} - x_{nj-1}| \to 0$ for $n \to \infty$, then we have $(\tau_n(t))_{0\leq t\leq T} \xrightarrow{\mathcal{L}} (\tau(t))_{0\leq t\leq T}$.

Remark: For inverse Gaussian distributions there exists a much simpler simulation algorithm developed in Michael, Schucany, and Haas (1976) which only needs a χ_1^2- and a uniformly distributed random variable to generate an IG-distributed random variate thereof. This and the convolution property of inverse Gaussian distributions (see Proposition 1.11 and the remark thereafter) allow for a very fast and easy simulation of inverse Gaussian Lévy processes $(\tau(t))_{t\geq 0}$ with $\mathcal{L}\big(\tau(1)\big) = IG(\delta, \gamma)$: Set $k_n := [Tn]$ and $\tau_n(t) := \sum_{i=1}^{[k_n t]} \bar{X}_{ni}$ where $\bar{X}_{ni} \sim IG\big(\frac{\delta T}{k_n}, \gamma\big)$ are iid, then $(\tau_n(t))_{0\leq t\leq T} \xrightarrow{\mathcal{L}} (\tau(t))_{0\leq t\leq T}$.
This also implies an easy simulation method for NIG Lévy processes $(L_t)_{t\geq 0}$ with $\mathcal{L}(L_1) = NIG(\alpha, \beta, \delta, \mu)$: Since the NIG distributions inherit the convolution property from the inverse Gaussian class (see equation (1.9)), the NIG

process L can be approximated by $L_t^n := \sum_{i=1}^{[k_n t]} \bar{Y}_{ni}$ where k_n is defined as above and $\bar{Y}_{ni}$ are iid with $\bar{Y}_{ni} \sim NIG\big(\alpha, \beta, \frac{\delta T}{k_n}, \frac{\mu T}{k_n}\big)$. Using the mixture representation $\bar{Y}_{ni} \stackrel{d}{=} \frac{\mu T}{k_n} + \beta Z_{ni} + \sqrt{Z_{ni}} W_{ni}$ where $Z_{ni} \sim IG\big(\frac{\delta T}{k_n}, \sqrt{\alpha^2 - \beta^2}\big)$ and $W_{ni} \sim N(0,1)$ is independent of Z_{ni}, the increments $\bar{Y}_{ni}$ can be obtained by generating standard normal and inverse Gaussian random variates with the help of the Michael-Schucany-Haas-algorithm.

Let us return to the problem of simulating GH distributed random variables for a moment. If one uses the triangular scheme defined in Proposition 1.31, the function U_n corresponding to the approximating sum $\sum_{i=1}^{k_n} X_{ni}$ (which is given in explicit form within the proof on p. 55) is constant outside the interval $[-K_n - 2\beta, K_n]$, hence the characteristic function of $\sum_{i=1}^{k_n} X_{ni}$ is given by

$$\phi_{\sum_{i=1}^{k_n} X_{ni}}(u) = \exp\left[iu\mu - \int_{-K_n-2\beta}^{K_n} \ln\left(1 - \frac{iu}{y}\right) U_n(\mathrm{d}y)\right].$$

A comparison with ϕ_{GH} in equation (1.38) shows that this procedure basically leads to a truncation of the tails of the integral in the exponent. If one generates GIG distributed random variables along the same lines, then analogously the right tail of the integral within $\phi_{GIG}(u) = \exp\big[\int_0^\infty \ln\big(1 - \frac{iu}{y}\big) U_{GIG}(\mathrm{d}y)\big]$ will be ignored. The next proposition shows that this may be compensated by adding an independent normal variable to the series $\sum_{i=1}^{k_n} X_{ni}$. It goes back to Bondesson (1982) where the result was mentioned (in a slightly different form), but not strictly proven. We first ensure that all required moments exist: Suppose that X_K is a random variable whose distribution is a generalized Γ-convolution ($\mathcal{L}(X_K) \in \Gamma_0$) with characteristic function

$$\phi_{X_K}(u) = \exp\left[-\int_K^\infty \ln\left(1 - \frac{iu}{y}\right) U(\mathrm{d}y)\right] \tag{1.39}$$

where U fulfills the conditions (1.24), then its mean is given by

$$\begin{aligned} \mathrm{E}[X_K] &= \frac{1}{i}\,\frac{\mathrm{d}\phi_{X_K}(u)}{\mathrm{d}u}\bigg|_{u=0} = \frac{\phi_{X_K}(u)}{i}\int_K^\infty \frac{i}{y - iu}\,U(\mathrm{d}y)\bigg|_{u=0} \\ &= \int_K^\infty \frac{1}{y}\,U(\mathrm{d}y) < \infty \end{aligned} \tag{1.40}$$

if $K \geq 1$ according to (1.24) which also justifies the interchange between differentiation and integration, and for the variance we get

$$\begin{aligned} \sigma_K^2 := \mathrm{Var}[X_K] &= \frac{1}{i}\,\frac{\mathrm{d}}{\mathrm{d}u}\int_K^\infty \frac{1}{y - iu}\,U(\mathrm{d}y)\bigg|_{u=0} \\ &= \frac{1}{i}\int_K^\infty \frac{i}{(y - iu)^2}\,U(\mathrm{d}y)\bigg|_{u=0} = \int_K^\infty \frac{1}{y^2}\,U(\mathrm{d}y). \end{aligned} \tag{1.41}$$

If instead of (1.39) the distribution of X_K is an extended generalized Γ-convolution ($\mathcal{L}(X_K) \in \Gamma$) and has the characteristic function

$$\phi_{X_K}(u) = \exp\left[-\int_{\mathbb{R}\setminus[-K,K]} \left(\ln\left(1 - \frac{iu}{y}\right) + \frac{iuy}{1+y^2}\right) U(\mathrm{d}y)\right] \tag{1.42}$$

with U fulfilling the constraints (1.27), then for $K \geq 1$ we obtain analogously

$$\mathrm{E}[X_K] = \int_{\mathbb{R}\setminus[-K,K]} \frac{1}{y(1+y^2)}\,U(\mathrm{d}y) \quad \text{and} \quad \sigma_K^2 = \int_{\mathbb{R}\setminus[-K,K]} \frac{1}{y^2}\,U(\mathrm{d}y). \quad (1.43)$$

Now we are ready to state the announced

Proposition 1.34 *Suppose X_K has a characteristic function defined by* (1.39) *or* (1.42)*. If $K\sigma_K \to \infty$ for $K \to \infty$, then*

$$\mathcal{L}\left(\frac{X_K - \mathrm{E}[X_K]}{\sigma_K}\right) \xrightarrow{w} N(0,1)$$

where $\mathrm{E}[X_K]$ *and* σ_K^2 *are given by* (1.40) *and* (1.41) *or* (1.43)*, respectively.*

Proof: By the remark on p. 30, every infinitely divisible random variable X which possesses a finite second moment admits a Lévy–Khintchine representation of the form

$$\phi_X(u) = \exp\left(iub_X - \frac{c_X u^2}{2} + \int_{\mathbb{R}} \left(e^{iux} - 1 - iux\right)\nu_X(\mathrm{d}x)\right)$$

from which mean and variance are found to be

$$\mathrm{E}[X] = \frac{1}{i}\frac{\mathrm{d}\phi_X(u)}{\mathrm{d}u}\bigg|_{u=0} = b_X,\ \mathrm{Var}[X] = -\frac{\mathrm{d}^2 \ln\big(\phi_X(u)\big)}{\mathrm{d}u^2}\bigg|_{u=0} = c_X + \int_{\mathbb{R}} x^2\,\nu_X(\mathrm{d}x).$$

If $\phi_{X_K}(u)$ is given by (1.39), then it follows from Proposition 1.20 and equations (1.25) that the Gaussian coefficient in the corresponding Lévy–Khintchine representation vanishes ($c_{X_K} = 0$), and the Lévy measure of X_K is $\nu_{X_K}(\mathrm{d}x) = \mathbb{1}_{(0,\infty)}(x)\frac{1}{x}\int_K^\infty e^{-xy}\,U(\mathrm{d}y)\,\mathrm{d}x$. Thus the characteristic triplet of $X_K^* := \frac{X_K - \mathrm{E}[X_K]}{\sigma_K}$ is

$$b_{X_K^*} = 0, \quad c_{X_K^*} = 0, \quad \nu_{X_K^*}(\mathrm{d}x) = \mathbb{1}_{(0,\infty)}(x)\frac{1}{x}\int_K^\infty e^{-\sigma_K xy}\,U(\mathrm{d}y)\,\mathrm{d}x.$$

Applying Jacod and Shiryaev (2003, Theorem VII.2.14), to prove the weak convergence of $\mathcal{L}(X_K^*)$ to $N(0,1)$ it suffices to show that

i) $\displaystyle\lim_{a\uparrow\infty}\limsup_{K\to\infty}\int_a^\infty x^2\,\nu_{X_K^*}(\mathrm{d}x) = 0.$

ii) $\displaystyle\lim_{K\to\infty}\int_0^\infty g(x)\,\nu_{X_K^*}(\mathrm{d}x) = \int_{-\infty}^{+\infty} g(x)\,\nu_{N(0,1)}(\mathrm{d}x) = 0$

for all continuous, bounded functions $g : \mathbb{R} \to \mathbb{R}$ which are 0 in some neighbourhood of the origin.

(The further conditions $[\beta_1']$ and $[\gamma_1']$ of Theorem VII.2.14 are trivially met because of the standardization of X_K^*.)
Using Fubini's theorem, the expression in i) can be written as

$$\lim_{a\uparrow\infty}\limsup_{K\to\infty}\int_a^\infty x^2\,\frac{1}{x}\int_K^\infty e^{-\sigma_K xy}\,U(\mathrm{d}y)\,\mathrm{d}x = \lim_{a\uparrow\infty}\limsup_{K\to\infty}\int_K^\infty\int_a^\infty xe^{-\sigma_K xy}\,\mathrm{d}x\,U(\mathrm{d}y)$$

$$= \lim_{a\uparrow\infty}\limsup_{K\to\infty}\int_K^\infty \frac{e^{-a\sigma_K y}}{\sigma_K^2 y^2}\,(1 + a\sigma_K y)\,U(\mathrm{d}y) \leq \lim_{a\uparrow\infty}\limsup_{K\to\infty}\int_K^\infty \frac{1}{a\sigma_K^3 y^3}\,U(\mathrm{d}y),$$

because $y \geq K$ and $K\sigma_K \to \infty$ by assumption, so for sufficiently large K we have $e^{-a\sigma_K y} \leq (a\sigma_K y(1+a\sigma_K y))^{-1}$. Continuing the calculation we find

$$\begin{aligned}\lim_{a\uparrow\infty}\limsup_{K\to\infty}\int_a^\infty x^2\,\frac{1}{x}\int_K^\infty e^{-\sigma_K xy}U(\mathrm{d}y)\,\mathrm{d}x &\leq \lim_{a\uparrow\infty}\limsup_{K\to\infty}\int_K^\infty \frac{1}{a\sigma_K^3 y^3}\,U(\mathrm{d}y)\\ \leq \lim_{a\uparrow\infty}\limsup_{K\to\infty}\frac{1}{aK\sigma_K^3}\int_K^\infty\frac{1}{y^2}\,U(\mathrm{d}y) &= \lim_{a\uparrow\infty}\limsup_{K\to\infty}\frac{1}{aK\sigma_K} = 0\end{aligned}$$

as desired.
To verify ii), fix some g and set $M_+ := \sup_{x\in\mathbb{R}} g(x)$, $M_- := \inf_{x\in\mathbb{R}} g(x)$ and $\varepsilon_g := \sup\{\varepsilon > 0 \,|\, g(x) \equiv 0,\ x \in (-\varepsilon,\varepsilon)\}$ (note that the assumptions on g in ii) imply $-\infty < M_- \leq 0 \leq M_+ < \infty$ and $\varepsilon_g > 0$). Then we conclude with similar arguments as before

$$\begin{aligned}\lim_{K\to\infty}\int_0^\infty g(x)\,\frac{1}{x}\int_K^\infty e^{-\sigma_K xy}\,U(\mathrm{d}y)\,\mathrm{d}x &\leq M_+\lim_{K\to\infty}\int_K^\infty\int_{\varepsilon_g}^\infty \frac{e^{-\sigma_K xy}}{x}\,\mathrm{d}x\,U(\mathrm{d}y)\\ \leq M_+\lim_{K\to\infty}\int_K^\infty\int_{\varepsilon_g}^\infty\frac{1}{\sigma_K^3 y^3 x^4}\,\mathrm{d}x\,U(\mathrm{d}y) &= \frac{M_+}{3\varepsilon_g^3}\lim_{K\to\infty}\int_K^\infty\frac{1}{\sigma_K^3 y^3}\,U(\mathrm{d}y) = 0\end{aligned}$$

as well as

$$\lim_{K\to\infty}\int_0^\infty g(x)\,\frac{1}{x}\int_K^\infty e^{-\sigma_K xy}\,U(\mathrm{d}y)\,\mathrm{d}x \geq \frac{M_-}{3\varepsilon_g^3}\lim_{K\to\infty}\int_K^\infty\frac{1}{\sigma_K^3 y^3}\,U(\mathrm{d}y) = 0,$$

thus ii) is also fulfilled.

If ϕ_{X_K} is given by (1.42), then the characteristics of $X_K^* := \frac{X_K - \mathrm{E}[X_K]}{\sigma_K}$ are $b_{X_K^*} = 0$, $c_{X_K^*} = 0$ and

$$\nu_{X_K^*}(\mathrm{d}x) = \left[\frac{\mathbb{1}_{(-\infty,0)}(x)}{|x|}\int_{-\infty}^{-K} e^{-\sigma_K xy}\,U(\mathrm{d}y) + \frac{\mathbb{1}_{(0,\infty)}(x)}{x}\int_K^\infty e^{-\sigma_K xy}\,U(\mathrm{d}y)\right]\mathrm{d}x,$$

and the conditions for normal convergence in this case are

i′) $\displaystyle\lim_{a\uparrow\infty}\limsup_{K\to\infty}\int_{\mathbb{R}\setminus[-a,a]} x^2\,\nu_{X_K^*}(\mathrm{d}x) = 0.$

ii′) $\displaystyle\lim_{K\to\infty}\int_{-\infty}^{+\infty} g(x)\,\nu_{X_K^*}(\mathrm{d}x) = \int_{-\infty}^{+\infty} g(x)\,\nu_{N(0,1)}(\mathrm{d}x) = 0$

for all continuous, bounded functions $g : \mathbb{R} \to \mathbb{R}$ which are 0 in some neighbourhood of the origin.

Either integral occuring in these conditions can obviously be split up into two integrals over $\mathbb{R}_-$ and $\mathbb{R}_+$ which then can separately be shown to converge to zero in exactly the same way as above. This completes the proof. □

By Proposition 1.23 the measure induced by U_{GIG} possesses the Lebesgue density $\delta^2 g_{|\lambda|}\big(2\delta^2 x - \delta^2\gamma^2\big)$ on the set $\big(\frac{\gamma^2}{2},\infty\big)$ for all $GIG(\lambda,\delta,\gamma)$-distributions with $0 < \delta < \infty$ and $0 \leq \gamma < \infty$. Its asymptotic behaviour for $x \to \infty$ is given by $\delta^2 g_{|\lambda|}\big(2\delta^2 x - \delta^2\gamma^2\big) \sim \delta(2\pi^2 x)^{-\frac{1}{2}}$ (see p. 40), hence for sufficiently large K and $\frac{\delta}{\pi\sqrt{2}} > C > 0$ we have

$$\sigma_K^2 = \int_K^\infty \frac{1}{y^2}\, U_{GIG}(\mathrm{d}y) \geq C \int_K^\infty y^{-\frac{5}{2}}\, \mathrm{d}y = \frac{2C}{3}\, K^{-\frac{3}{2}},$$

consequently $\lim_{K\to\infty} K\sigma_K \geq \lim_{K\to\infty}\big(\frac{2C}{3}\big)^{\frac{1}{2}} K^{\frac{1}{4}} = \infty$. Thus if $\sum_{i=1}^{k_n} X_{ni}$ is a sum of Gamma variables defined along the lines of Corollary 1.33 that converges in law to some random variable $X \sim GIG(\lambda, \delta, \gamma)$ with $\delta > 0$, Proposition 1.34 suggests that the convergence can be improved by adding an independent normal variate $X_{nk_n+1} \sim N\big(\mathrm{E}[X_K], \sigma_K^2\big)$ to the series $\sum_{i=1}^{k_n} X_{ni}$. (Here an "improvement of convergence" is to be understood in the sense that the total variation distance between the law of the series and $GIG(\lambda, \delta, \gamma)$ will be reduced by amending the summand X_{nk_n+1}.) At first it may be surprising that adding a Gaussian random variable with range $\mathbb{R}$ could lead to a better approximation of a GIG distributed random variate whose law is concentrated on $\mathbb{R}_+$, but one should observe that by (1.40) and (1.41)

$$\sigma_K^2 = \int_K^\infty \frac{1}{y^2}\, U(\mathrm{d}y) \leq \frac{1}{K}\int_K^\infty \frac{1}{y}\, U(\mathrm{d}y) = \frac{\mathrm{E}[X_K]}{K},$$

hence $\mathrm{E}[X_K] \gg \sigma_K^2$ as K increases, so an $N(\mathrm{E}[X_K], \sigma_K^2)$-distributed random variable will take positive values with high probability.

Moreover, Corollary 1.27 implies that for all $GH(\lambda, \alpha, \beta, \delta, \mu)$-distributions with $\delta > 0$ the measure induced by U_{GH} possesses a Lebesgue density of the form $\mathrm{sign}(x)\delta^2(x+\beta)g_{|\lambda|}\big(\delta^2 x^2 + 2\delta^2\beta x - \delta^2(\alpha^2 - \beta^2)\big)$ on $\mathbb{R}\setminus[-\alpha-\beta, \alpha-\beta]$. On p. 49 we saw that $\delta^2(z+\alpha)g_{|\lambda|}\big(\delta^2(z^2 + 2\alpha z)\big) \sim \frac{\delta}{\pi}$, $z \to \infty$, consequently the density of U_{GH} also converges to $\frac{\delta}{\pi}$ as $x \to \pm\infty$. For sufficiently large K and $\frac{\delta}{\pi} > C > 0$ we thus have

$$\sigma_K^2 = \int_{\mathbb{R}\setminus[-K,K]} \frac{1}{y^2}\, U_{GH}(\mathrm{d}y) \geq C \int_{\mathbb{R}\setminus[-K,K]} \frac{1}{y^2}\, \mathrm{d}y = \frac{2C}{K}$$

and therefore $\lim_{K\to\infty} K\sigma_K \geq \lim_{K\to\infty}\sqrt{2CK} = \infty$. Hence again the conditions of Proposition 1.34 are met, and all conclusions drawn above can also be transferred to the GH distributions.

Remark: If a GH distributed random variable is approximated along the lines of Proposition 1.31, then the characteristic function $\phi_{\bar{X}_K}$ of the "error term" $\bar{X}_K$ is roughly given by

$$\phi_{\bar{X}_K}(u) = \exp\left[-\int_{\mathbb{R}\setminus[-K,K]} \ln\!\left(1 - \frac{iu}{y}\right) U_{GH}(\mathrm{d}y)\right]$$

which slightly differs from (1.42), but it is easily seen that this only has an effect on the mean $\mathrm{E}[\bar{X}_K]$ whereas the variance remains unchanged: $\bar{\sigma}_K^2 := \mathrm{Var}[\bar{X}_K] = \sigma_K^2$, where σ_K^2 is given by (1.43). Thus the assertion of Proposition 1.34 remains valid in this case.

Note that the condition $K\sigma_K \to \infty$ is not fulfilled in the limiting cases with $\delta = 0$, that is, for $GIG(\lambda, 0, \gamma)$- and $VG(\lambda, \alpha, \beta, \mu)$-distributions, because the corresponding functions $U_{GIG(\lambda,0,\gamma)}(x)$ and $U_{VG}(x)$ are constant for $x > \frac{\gamma^2}{2}$ and $x \notin [-\alpha - \beta, \alpha - \beta]$, respectively, hence in both cases $\sigma_K^2 \equiv 0$ if $K > \frac{\gamma^2}{2}$ or $K > \alpha + |\beta|$. But since a VG distributed random variable by Corollary 1.28 in law exactly equals the difference of two Gamma variables, one does not need an additional Gaussian variate for the approximation either, and the relation $GIG(\lambda, 0, \gamma) = G\big(\lambda, \frac{\gamma^2}{2}\big)$ shows that such a variate is also superfluous in the latter case for trivial reasons.

Of course, Proposition 1.34 can be applied to every summand $\bar{X}_{ni}$ occuring in the approximating schemes of GH and GIG Lévy processes defined in Corollaries 1.32 and 1.33 as well. This amounts to the conclusion that for all GH and GIG distributions with parameter $\delta > 0$ the approximation of the corresponding Lévy processes can be improved by adding a suitably scaled Brownian motion with drift to the processes $(L_t^n)_{0\le t\le T}$ and $(\tau_n(t))_{0\le t\le T}$. This may be regarded as a compensation of the small jumps of the processes L and τ that are not covered by L^n and τ_n, because the tail behaviour of the measures induced by U_{GH} and U_{GIG} significantly influences the masses the Lévy measures ν_{GH} and ν_{GIG} concentrate around the origin. (Recall that this is an immediate consequence of the Tauberian Theorem 1.25 and its application in Sections 1.6.4 and 1.6.2.)

These findings coincide with a result of Asmussen and Rosiński (2001) to be explained in the following: Suppose $(X_t)_{t\ge 0}$ is a Lévy process without a Brownian part (this assumption is in general not necessary, but simplifies the exposition and is fulfilled for all GH and GIG Lévy processes we are concerned with here). Then X may be represented as the sum of two independent Lévy processes $X - X^\varepsilon$ and X^ε which are defined by the following decomposition of the characteristic function of X_1:

$$\begin{aligned}
\phi_{X_1}(u) \;&=\; \exp\bigg(iub_X + \int_{\mathbb{R}} \big(e^{iux} - 1 - iux\mathbb{1}_{[-1,1]}(x)\big)\, \nu_X(\mathrm{d}x)\bigg) \\
&=\; \exp\bigg(iub_X + \int_{\mathbb{R}\setminus(-\varepsilon,\varepsilon)} \big(e^{iux} - 1 - iux\mathbb{1}_{[-1,1]}(x)\big)\, \nu_X(\mathrm{d}x)\bigg) \\
&\qquad\cdot \exp\bigg(\int_{(-\varepsilon,\varepsilon)} \big(e^{iux} - 1 - iux\big)\, \nu_X(\mathrm{d}x)\bigg) \\
&=: \phi_{X_1 - X_1^\varepsilon}(u) \cdot \phi_{X_1^\varepsilon}(u)
\end{aligned}$$

The process $X - X^\varepsilon$ is a compound Poisson process with drift which contains all jumps of X whose absolute heights are bigger or equal than ε, whereas X^ε equals the compensated sum of jumps that are smaller than ε. Within the previous framework, one may, roughly speaking, identify $X - X^\varepsilon$ with L^n or τ_n. Now let us concentrate on X^ε. The characteristic function of X_1^ε given above implies

$$\mathrm{E}[X_1^\varepsilon] = 0 \qquad \text{and} \quad \sigma^2(\varepsilon) := \mathrm{Var}[X_1^\varepsilon] = \int_{(-\varepsilon,\varepsilon)} x^2\, \nu_X(\mathrm{d}x).$$

If the variance $\sigma^2(\varepsilon)$ satisfies a certain growth condition, then the scaled process $\sigma(\varepsilon)^{-1}X^\varepsilon$ converges in law to a standard Brownian motion B. More precisely, Asmussen and Rosiński (2001, Theorem 2.1 and Proposition 2.1) showed that

$$\text{If } \lim_{\varepsilon\to 0}\frac{\sigma(\varepsilon)}{\varepsilon} = \infty, \text{ then } \left(\sigma(\varepsilon)^{-1}X_t^\varepsilon\right)_{t\geq 0} \xrightarrow{\mathcal{L}} (B_t)_{t\geq 0} \text{ as } \varepsilon \to 0.$$

Remark: Under an additional assumption on the Lévy measure ν_X, this result can be extended: If $\nu_X(\mathrm{d}x)$ does not have atoms (point masses) in some neighbourhood of the origin, then both assertions above are even equivalent (see Asmussen and Rosiński 2001, Proposition 2.1), that is, $\sigma(\varepsilon)^{-1}X^\varepsilon$ converges in law to B if and only if $\lim_{\varepsilon\to 0}\frac{\sigma(\varepsilon)}{\varepsilon} = \infty$.

From the asymptotic behaviour of the Lévy densities $g_{GIG(\lambda,\delta,\gamma)}$ and $g_{GH(\lambda,\alpha,\beta,\delta,\mu)}$ derived in equations (1.31) and (1.36) we conclude that

$$\sigma_{GIG}(\varepsilon) \sim \sqrt{\frac{\delta}{3}}\left(\frac{2\varepsilon^3}{\pi}\right)^{\frac{1}{4}} \quad \text{and} \quad \sigma_{GH}(\varepsilon) \sim \left(\frac{2\delta\varepsilon}{\pi}\right)^{\frac{1}{2}} \quad \text{for } \varepsilon \to 0,$$

hence we have that $\lim_{\varepsilon\to 0}\frac{\sigma_{GIG}(\varepsilon)}{\varepsilon} = \lim_{\varepsilon\to 0}\frac{\sigma_{GH}(\varepsilon)}{\varepsilon} = \infty$ holds for all GIG and GH distributions with parameter $\delta > 0$, including the limit distributions with $\lambda < 0$. Consequently the small jumps of the corresponding Lévy processes can be approximated by an appropriately scaled Brownian motion.

Again this does not hold true for the Gamma and VG processes, because the asymptotics of their Lévy densities are given by $g_{G(\lambda,\sigma)}(x) \sim \frac{\lambda}{x}$, $x \downarrow 0$, and $g_{VG(\lambda,\alpha,\beta,\mu)}(x) \sim \frac{\lambda}{|x|}$, $x \to 0$, respectively. This implies $\lim_{\varepsilon\downarrow 0}\frac{\sigma_{G(\lambda,\sigma)}(\varepsilon)}{(\varepsilon)} = \frac{\sqrt{\lambda}}{\sqrt{2}}$ and $\lim_{\varepsilon\to 0}\frac{\sigma_{VG(\lambda,\alpha,\beta,\mu)}(\varepsilon)}{(\varepsilon)} = \sqrt{\lambda}$, thus the convergence of the compensated small jumps to a Brownian motion fails according to the remark above. In the VG case, this is also plausible from an intuitive point of view: Recall that the paths of a VG process are of finite variation (see the remark on p. 50), whereas in contrast to this the paths of Brownian motions are almost surely of infinite variation, so any approximation of VG processes that comprehends a Brownian part will lead to sample paths with completely different (and hence wrong) path properties.

It was already mentioned in the remark on p. 8 that a sequence of Lévy processes $(X_t^n)_{t\geq 0}$ converges in law to a Lévy process $(X_t)_{t\geq 0}$ iff $X_1^n \xrightarrow{\mathcal{L}} X_1$. This provides an immediate extension of Proposition 1.34 to

Corollary 1.35 *Let $(X_t^K)_{t\geq 0}$ be a Lévy process defined by $\phi_{X_t^K}(u) = \phi_{X_K}(u)^t$ where $\phi_{X_K}(u)$ is given by (1.39) or (1.42), and set $\sigma_K^2 := \mathrm{Var}[X_1^K]$. If $K\sigma_K \to \infty$ for $K \to \infty$, then*

$$\left(\sigma_K^{-1}(X_t^K - t\mathrm{E}[X_1^K])\right)_{t\geq 0} \xrightarrow{\mathcal{L}} (B_t)_{t\geq 0}$$

where $(B_t)_{t\geq 0}$ denotes a standard Brownian motion.

Intuitively one may assume the correspondence $\varepsilon \leftrightarrow \frac{1}{K}$ such that the above corollary can to some extent be regarded as a reformulation of the Asmussen-Rosiński result for the special case that the underlying Lévy process X is generated by an (extended) generalized Γ-convolution. In fact, under some additional assumptions the conditions $\lim_{K\to\infty} K\sigma_K = \infty$ and $\lim_{\varepsilon\to 0} \frac{\sigma(\varepsilon)}{\varepsilon} = \infty$ can be fulfilled simultaneously and are thus equivalent under these circumstances.

Suppose for example that $(X_t)_{t\geq 0}$ is a Lévy process where $\mathcal{L}(X_1) \in \Gamma_0$, $X_1 = (X_1 - X_K) + X_K$, and ϕ_{X_K} is defined by equation (1.39). If the measure $U(\mathrm{d}y)$ has a Lebesgue density u with the asymptotic behaviour $u(y) \sim cy^{-a}$ for $y \to \infty$ with $0 < a < 1$ and $c > 0$, then the conditions on σ_K and $\sigma(\varepsilon)$ are both met: First, the asymptotics of u imply

$$\sigma_K^2 = \int_K^\infty \frac{1}{y^2}\, U(\mathrm{d}y) \sim c \int_K^\infty y^{-2-a}\, \mathrm{d}y = \frac{c}{1+a}\, K^{-(1+a)}, \quad K \to \infty,$$

hence $\lim_{K\to\infty} K\sigma_K = \lim_{K\to\infty} \frac{\sqrt{c}}{\sqrt{1+a}} K^{\frac{1-a}{2}} = \infty$ because $a < 1$. Moreover, by Proposition 1.20 and equations (1.25) the density of the Lévy measure $\nu_{X_1}(\mathrm{d}x)$ of X_1 is given by $g(x) = \frac{1}{x}\int_0^\infty e^{-xy}\, u(y)\, \mathrm{d}y$, $x > 0$. The assumptions on u allow the application of Theorem 1.25 with $\rho = 1 - a > 0$ and $L(y) \equiv c\Gamma(1-a)$ from which we obtain $g(x) \sim c\Gamma(1-a)x^{-(2-a)}$, $x \downarrow 0$. Consequently

$$\sigma^2(\varepsilon) = \int_0^\varepsilon x^2 g(x)\, \mathrm{d}x \sim c\Gamma(1-a) \int_0^\varepsilon x^a\, \mathrm{d}x = \frac{c\Gamma(1-a)}{1+a}\, \varepsilon^{1+a}, \quad \varepsilon \to 0,$$

and thus $\lim_{\varepsilon\to 0} \frac{\sigma(\varepsilon)}{\varepsilon} = \lim_{\varepsilon\to 0} \big(\frac{c\Gamma(1-a)}{1+a}\big)^{\frac{1}{2}} \varepsilon^{-\frac{1-a}{2}} = \infty$.

If instead $\mathcal{L}(X_1)$ is an extended generalized Γ-convolution ($\mathcal{L}(X_1) \in \Gamma$), then its Lévy measure has a density which is the sum of two Laplace transforms, and analogously as in the proof of Proposition 1.34, every integral under consideration can be split up into two integrals over $\mathbb{R}_-$ and $\mathbb{R}_+$ which can be examined separately as above. Therefore both conditions on σ_K and $\sigma(\varepsilon)$ are also met in this case if the measure $U(\mathrm{d}y)$ possesses a density u which behaves like $u(y) \sim c_1|y|^{-a_1}$, $y \to -\infty$, or $u(y) \sim c_2 y^{-a_2}$, $y \to \infty$, where $-1 < a_1, a_2 < 1$ and $c_1, c_2 > 0$ (note that one of these asymptotic relations is already sufficient because of the possible splitting of the involved integrals).

Remark: The upper bound $a < 1$ resp. $a_1, a_2 < 1$ is required by the constraint $\rho > 0$ in Theorem 1.25, the lower bounds $a > 0$ and $a_1, a_2 > -1$ are enforced by the integrability conditions in (1.24) and (1.27). All GIG and GH distributions with parameter $\delta > 0$ can be regarded as special cases of this more general framework which fulfill the assumptions above with $c = \frac{\delta}{\sqrt{2\pi}}$, $a = \frac{1}{2}$ and $c_1 = c_2 = \frac{\delta}{\pi}$, $a_1 = a_2 = 0$, respectively.

Chapter 2

Multivariate GH distributions and their limits

In the last chapter we have focussed on univariate distributions and processes. However, many tasks and problems arising in financial mathematics are inherently multivariate: consider for example a portfolio of assets or an option whose payoff depends on two or more underlyings, such as swap-, spread-, rainbow- or basket options. To determine a trading strategy which maximizes the expected future value of the portfolio or to estimate potential portfolio losses, as well as to prize the above mentioned options, one essentially needs a multivariate model. In other words, one requires an assumption on the *joint distribution* of all portfolio ingredients respectively stocks on which the option depends. The knowledge of the corresponding univariate marginals is by no means sufficient since they provide no information about the *dependence structure* which considerably influences the risks and returns of the portfolio and the value of the option. Specifying the dependencies between portfolio constituents is also a key ingredient of credit risk models, as we shall see in Chapter 3.

Many higher-dimensional models used in financial mathematics are still based on multivariate normal distributions, despite the fact that empirical investigations strongly reject the hypothesis of multivariate normal distributed asset returns: see for example Affleck-Graves and McDonald (1989), Richardson and Smith (1993) or McNeil, Frey, and Embrechts (2005, Chapter 3, pp. 70–73). Apart from the fact that the marginal log return-distributions deviate significantly from the normal ones, as was already pointed out at the beginning of Chapter 1, a second reason for the rejection of the multivariate normal distribution is its far too simple dependence structure. The components of a multivariate normal distributed random vector are only linearly dependent. This means, their dependencies are already completely characterized by the corresponding covariance matrix, whereas financial data typically exhibits a much more complex dependence structure. In particular, the probability of joint extreme outcomes is severly underestimated by the normal distribution because it also assigns too little weight to the joint tails.

To overcome the deficiencies of the multivariate normal distribution, various alternatives have been proposed in the literature from which we just want to mention the following examples (the list could surely be extended much further): the class of elliptical distributions (Owen and Rabinovitch 1983, Kring, Rachev, Höchstötter, Fabozzi, and Bianchi 2009), (extended) multivariate t distributions (Khan and Zhou 2006, Adcock 2010), multivariate Variance Gamma distributions (Luciano and Schoutens 2006, Semeraro 2008), and multivariate GH distributions (Prause 1999, Chapter 4, Eberlein and Prause 2002, Sections 6 and 7, McNeil, Frey, and Embrechts 2005, Chapter 3.2). In the present chapter the latter are investigated in greater detail, and it is shown how they are related to the other aforementioned distribution classes.

The chapter is organized as follows: Section 2.1 provides some notations and technical preliminaries, especially the concept of multivariate normal mean-variance mixtures. Many of them are natural and straightforward generalizations of the results presented in Section 1.1 of Chapter 1, but are written down here explicitly again for the sake of completeness. Multivariate generalized hyperbolic distributions are introduced in Section 2.2, in which also some basic properties and possible limit distributions are described. The corresponding Lévy–Khintchine representations are derived in Section 2.3, and in the last section we take a closer look at the dependence structure of multivariate generalized hyperbolic distributions. In particular we analyze the tail dependence of GH distributions and provide necessary and sufficient conditions for tail independence.

2.1 Multivariate normal mean-variance mixtures and infinite divisibility

Let us first fix some notations which will be used throughout this chapter: The vectors $u = (u_1, \dots, u_d)^\top$ and $x = (x_1, \dots, x_d)^\top$ are elements of $\mathbb{R}^d$, the superscript $^\top$ stands for the transpose of a vector or matrix. $\langle u, x\rangle = u^\top x = \sum_{i=1}^d u_i x_i$ denotes the scalar product of the vectors u, x and $\|u\| = (u_1^2 + \cdots + u_d^2)^{1/2}$ the Euclidean norm of u. If A is a real-valued $d \times d$-square matrix, then $\det(A)$ denotes the determinant of A. The $d \times d$-identity matrix is labeled I_d. In contrast to u and x, the letters y, s and t are reserved for univariate real variables, that is, we assume $y, s, t \in \mathbb{R}$ or $\mathbb{R}_+$. To properly distinguish between the real number zero and the zero vector, we write $0 \in \mathbb{R}$ and $\mathbf{0} := (0, \dots, 0)^\top \in \mathbb{R}^d$.

Remark: Note that here and in the following $d \geq 2$ indicates the dimension, whereas n is usually used as running index for all kinds of sequences. In particular the notation $N_d(\mu, \Delta)$ will be used for the d-dimensional normal distribution with mean vector μ and covariance matrix Δ.

Having clarified the notation, we now turn to the more sophisticated parts of this section. The definition of multivariate infinite divisibility can be transferred almost literally from Definition 1.1. More precisely we have

Definition 2.1 *A probability measure μ on $(\mathbb{R}^d, \mathcal{B}^d)$ is* infinitely divisible *if for any integer $n \geq 1$ there exists a probability measure μ_n on $(\mathbb{R}^d, \mathcal{B}^d)$ such that μ equals the n-fold convolution of μ_n, that is, $\mu = \mu_n * \cdots * \mu_n =: *_{i=1}^n \mu_n$.*

Similar to the univariate case, the characteristic function $\phi_\mu(u)=\int_{\mathbb{R}^d} e^{i\langle u,x\rangle}\mu(\mathrm{d}x)$ of every infinitely divisible probability measure μ on $(\mathbb{R}^d,\mathcal{B}^d)$ admits a *Lévy–Khintchine representation*

$$\phi_\mu(u) = \exp\left(i\langle u,b\rangle - \frac{1}{2}\langle u,Cu\rangle + \int_{\mathbb{R}^d}\left(e^{i\langle u,x\rangle} - 1 - i\langle u,x\rangle \mathbb{1}_{\{\|x\|\le 1\}}(x)\right)\nu(\mathrm{d}x)\right)$$

where $b\in\mathbb{R}^d$, C is a symmetric positive-semidefinite $d\times d$ matrix and the *Lévy measure* $\nu(\mathrm{d}x)$ on $\mathbb{R}^d$ satisfies $\nu(\{\mathbf{0}\})=0$ as well as $\int_{\mathbb{R}^d}(\|x\|^2\wedge 1)\,\nu(\mathrm{d}x)<\infty$. Again, the triplet (b,C,ν) is unique and completely characterizes μ (Sato 1999, Theorem 8.1).

An $\mathbb{R}^d$-valued stochastic process $(L_t)_{t\ge 0}$ with $L_0=\mathbf{0}$ almost surely that is adapted to some underlying filtration is a *Lévy process* if it has stationary, independent increments and is continuous in probability in the sense of Definition 1.2 in Chapter 1. As before, these properties imply that $\mathcal{L}(L_t)$ is infinitely divisible for all $t\in\mathbb{R}_+$, and the characteristic functions ϕ_{L_t} fulfill $\phi_{L_t}(u)=\phi_{L_1}(u)^t$. Moreover, there is a one-to-one correspondence between multivariate infinitely divisible distributions μ and Lévy processes L via the relation $\phi_{L_t}(u)=\phi_\mu(u)^t$ (Sato 1999, Theorem 7.10). Thus we may occasionally speak of the characteristic triplet of L in the sense that (b,C,ν) is the characteristic triplet of $\mu=\mathcal{L}(L_1)$. Though the class of selfdecomposable distributions is defined in the same way as in the univariate case, we restate its definition for the reader's convenience:

Definition 2.2 *A probability measure μ on $(\mathbb{R}^d,\mathcal{B}^d)$ is called* selfdecomposable *if for every $0<s<1$ there is a probability measure μ_s on $(\mathbb{R}^d,\mathcal{B}^d)$ such that*

$$\phi_\mu(u)=\phi_\mu(su)\phi_{\mu_s}(u).$$

The characterization of multivariate selfdecomposable distributions by means of their Lévy measures is much more involved as the following lemma shows. It can be found in Sato (1999, Theorem 15.10 and Remark 15.12).

Lemma 2.3 *A probability measure μ on $(\mathbb{R}^d,\mathcal{B}^d)$ is selfdecomposable if and only if it is infinitely divisible and its Lévy measure admits the representation*

$$\nu(B)=\int_{\mathcal{S}}\bar\lambda(\mathrm{d}\xi)\int_0^\infty \mathbb{1}_B(r\xi)\,\frac{k_\xi(r)}{r}\,\mathrm{d}r \qquad \forall\, B\in\mathcal{B}^d$$

where $\bar\lambda(\mathrm{d}\xi)$ is a finite measure on the unit sphere $\mathcal{S}:=\{\xi\in\mathbb{R}^d \mid \|\xi\|=1\}$, and $k_\xi(r)$ is a non-negative function that is measurable in $\xi\in\mathcal{S}$ and decreasing in $r>0$. If $\nu(\mathrm{d}x)$ is not the zero measure, then $\bar\lambda(\mathrm{d}\xi)$ and $k_\xi(r)$ can be chosen such that $\bar\lambda(\mathcal{S})=1$ and $\int_0^\infty(r^2\wedge 1)\frac{k_\xi(r)}{r}\,\mathrm{d}r$ is finite and independent of ξ.

Remark: The so-called polar decomposition of the Lévy measure into a spherical and a radial part itself is not a distinguished property of selfdecomposable distributions, but can be done for every Lévy measure $\nu(\mathrm{d}x)$ of an infinitely divisible distribution on $(\mathbb{R}^d,\mathcal{B}^d)$, as Barndorff-Nielsen, Maejima, and Sato (2006, Lemma 2.1) have shown. The characteristic feature of a seldecomposable distribution is that the radial component of its Lévy measure has a Lebesgue density $\frac{k_\xi(r)}{r}$ with decreasing $k_\xi(r)$ (whereas the spherical measure $\bar\lambda(\mathrm{d}\xi)$, which should not be confused with the Lebesgue measure, is not supposed to have a density).

Definition 2.4 *An $\mathbb{R}^d$-valued random variable X is said to have a* multivariate normal mean-variance mixture distribution *if*

$$X \stackrel{d}{=} \mu + Z\beta + \sqrt{Z}AW,$$

where $\mu, \beta \in \mathbb{R}^d$, A is a real-valued $d \times d$-matrix such that $\Delta := AA^\top$ is positive definite, W is a standard normal distributed random vector ($W \sim N_d(\mathbf{0}, I_d)$) and $Z \sim G$ is a real-valued, non-negative random variable independent of W.
Equivalently, a probability measure F on $(\mathbb{R}^d, \mathcal{B}^d)$ is said to be a multivariate normal mean-variance mixture if

$$F(\mathrm{d}x) = \int_{\mathbb{R}_+} N_d(\mu + y\beta, y\Delta)(\mathrm{d}x)\, G(\mathrm{d}y),$$

where the mixing distribution G is a probability measure on $(\mathbb{R}_+, \mathcal{B}_+)$.
We use the short hand notation $F = N_d(\mu + y\beta, y\Delta) \circ G$. If $\mathbb{G}$ is a class of mixing distributions, then $N_d(\mu + y\beta, y\Delta) \circ \mathbb{G} := \{N_d(\mu + y\beta, y\Delta) \circ G \mid G \in \mathbb{G}, \mu \in \mathbb{R}^d\}$.

Remark: Note that one can further assume without loss of generality $|\det(A)| = \det(\Delta) = 1$, since a (positive) multiplicative constant can always be included within the variable Z. More precisely, let $\bar{A} = |\det(A)|^{-1/d} A$, $\bar{\beta} = |\det(A)|^{-2/d}\beta$ and $\bar{Z} = |\det(A)|^{2/d} Z$, then $|\det(\bar{A})| = 1$ and $\mu + Z\beta + \sqrt{Z}AW = \mu + \bar{Z}\bar{\beta} + \sqrt{\bar{Z}}\bar{A}W$. Equivalently, if $\bar{\Delta} = \bar{A}\bar{A}^\top$ and $\bar{G} = \mathcal{L}(\bar{Z})$, then also $\det(\bar{\Delta}) = 1$ and $N_d(\mu + y\beta, y\Delta) \circ G = N_d(\mu + y\bar{\beta}, y\bar{\Delta}) \circ \bar{G}$.

The use of a single univariate mixing variable Z causes dependencies between all entries of X, as we shall see in Section 2.4. A more general approach based on multivariate mixing variables which also allows for independence of all components of X has been proposed in Luciano and Semeraro (2010). But since the multivariate GH distributions to be studied in the subsequent sections are defined along the lines above, we shall not extend the setting further here.

The next lemma summarizes the most important properties of multivariate normal mean-variance mixtures. It is a straightforward generalization of the Lemmas 1.6 and 1.7 in Chapter 1.

Lemma 2.5 *Let $\mathbb{G}$ be a class of probability distributions on $(\mathbb{R}_+, \mathcal{B}_+)$ and $G, G_1, G_2 \in \mathbb{G}$.*

a) If G possesses a moment generating function $M_G(y)$ on some open interval (a, b) with $a < 0 < b$, then $F = N_d(\mu + y\beta, y\Delta) \circ G$ also possesses a moment generating function $M_F(u) = e^{\langle u, \mu \rangle} M_G\big(\frac{\langle u, \Delta u \rangle}{2} + \langle u, \beta \rangle\big)$ that is defined for all $u \in \mathbb{R}^d$ with $a < \frac{\langle u, \Delta u \rangle}{2} + \langle u, \beta \rangle < b$.

b) If $G = G_1 * G_2 \in \mathbb{G}$, then $\big(N_d(\mu_1 + y\beta, y\Delta) \circ G_1\big) * \big(N_d(\mu_2 + y\beta, y\Delta) \circ G_2\big) = N_d(\mu_1 + \mu_2 + y\beta, y\Delta) \circ G \in N_d(\mu + y\beta, y\Delta) \circ \mathbb{G}$.

c) If G is infinitely divisible, then so is $N_d(\mu + y\beta, y\Delta) \circ G$.

d) If G is selfdecomposable, then so is the multivariate normal variance mixture $N_d(\mu, y\Delta) \circ G$.

e) If $(\mu_n)_{n\geq 1}$ and $(\beta_n)_{n\geq 1}$ are convergent sequences of real vectors with finite limits $\mu, \beta \in \mathbb{R}^d$ (that is, $\|\mu\|, \|\beta\| < \infty$), and $(G_n)_{n\geq 1}$ is a sequence of mixing distributions with $G_n \xrightarrow{w} G$, then $N_d(\mu_n + y\beta_n, y\Delta) \circ G_n \xrightarrow{w} N_d(\mu + y\beta, y\Delta) \circ G$.

Proof: Because the moment generating function and the characteristic function of a multivariate normal distribution $N_d(\mu, \Delta)$ are given by $M_{N_d(\mu,\Delta)}(u) = e^{\frac{\langle u,\Delta u\rangle}{2} + \langle u,\mu\rangle}$ and $\phi_{N_d(\mu,\Delta)}(u) = M_{N_d(\mu,\Delta)}(iu)$, respectively, parts a)–c) can be verified completely analogously to the proof of Lemma 1.6. From this one especially obtains that the characteristic function of $F = N_d(\mu+y\beta, y\Delta)\circ G$ admits the representation $\phi_F(u) = e^{i\langle u,\mu\rangle}\mathfrak{L}_G\big(\frac{\langle u,\Delta u\rangle}{2} - i\langle u, \beta\rangle\big)$. Using this and the fact that the definition of selfdecomposability is the same for uni- and multivariate distributions, it is easily seen that the proof of Lemma 1.6 d) can also be transferred almost literally to the present case.
Setting $F_n := N_d(\mu_n + y\beta_n, y\Delta) \circ G_n$ and $F := N_d(\mu + y\beta, y\Delta) \circ G$, to prove e) it suffices to show that for an arbitrarily fixed $u \in \mathbb{R}^d$ we have

$$\phi_{F_n}(u) = e^{i\langle u,\mu_n\rangle}\mathfrak{L}_{G_n}\Big(\frac{\langle u,\Delta u\rangle}{2} - i\langle u, \beta_n\rangle\Big) \underset{n\to\infty}{\longrightarrow} e^{i\langle u,\mu\rangle}\mathfrak{L}_G\Big(\frac{\langle u,\Delta u\rangle}{2} - i\langle u, \beta\rangle\Big) = \phi_F(u)$$

but this follows with exactly the same reasoning as in the proof of Lemma 1.7. (It would even be possible to slightly generalize the assertion by additionally assuming that there exists a sequence $(\Delta^n)_{n\geq 1}$ of covariance matrices which converges element-wise to Δ, that is, $\max_{1\leq i,j\leq d} |\Delta^n_{ij} - \Delta_{ij}| \to 0$ if $n \to \infty$.) $\square$

Remark: Observe that in contrast to the univariate case the selfdecomposability of the mixing distribution only transfers to the corresponding normal variance mixtures, but in general not to normal mean-variance mixtures with $\beta \neq \mathbf{0}$. As an example we shall see in Section 2.3 that multivariate VG distributions are selfdecomposable if and only if $\beta = \mathbf{0}$.

The next lemma is the multivariate analogon of Proposition 1.8 in Chapter 1. We state it here for further reference in Section 2.3.

Lemma 2.6 *Let $F = N_d(\mu + y\beta, y\Delta) \circ G$ be a multivariate normal mean-variance mixture with infinitely divisible mixing distribution G and $(X_t)_{t\geq 0}$, $(\tau(t))_{t\geq 0}$ be two Lévy processes with $\mathcal{L}(X_1) = F$ and $\mathcal{L}(\tau(1)) = G$.
Set $(\bar{B}_t)_{t\geq 0} := (AB_t)_{t\geq 0}$ where $(B_t)_{t\geq 0}$ is a d-dimensional standard Brownian motion independent of $(\tau(t))_{t\geq 0}$, and A is a $d \times d$-matrix fulfilling $AA^\top = \Delta$. Then $(Y_t)_{t\geq 0}$, defined by*

$$Y_t := \mu t + \beta\tau(t) + \bar{B}_{\tau(t)},$$

is a Lévy process that is identical in law to $(X_t)_{t\geq 0}$.

Proof: The independence of $(B_t)_{t\geq 0}$, $(\tau(t))_{t\geq 0}$ implies that $(\beta\tau(t) + \bar{B}_{\tau(t)})_{t\geq 0}$ and hence $(Y_t)_{t\geq 0}$ are Lévy processes according to Sato (1999, Theorem 30.1) (see Theorem 2.13 in this thesis). The characteristic function of Y_1 is given by

$$\begin{aligned}\phi_{Y_1}(u) &= \mathrm{E}\big[e^{i\langle u,Y_1\rangle}\big] = e^{i\langle u,\mu\rangle}\,\mathrm{E}\big[e^{i\langle u,\tau(1)\beta\rangle}\mathrm{E}[e^{i\langle u,\bar{B}_{\tau(1)}\rangle}|\tau(1)]\big]\\ &= e^{i\langle u,\mu\rangle}\,\mathrm{E}\Big[e^{-\big(\frac{\langle u,\Delta u\rangle}{2} - i\langle u,\beta\rangle\big)\tau(1)}\Big] = e^{i\langle u,\mu\rangle}\,\mathfrak{L}_G\big(\tfrac{\langle u,\Delta u\rangle}{2} - i\langle u, \beta\rangle\big),\end{aligned}$$

because $\mathcal{L}(\tau(1)) = G$. From the proof of Lemma 2.5 we know that $\phi_{Y_1}(u) = \phi_F(u) = \phi_{X_1}(u)$, so $(X_t)_{t\geq 0}$ and $(Y_t)_{t\geq 0}$ are two Lévy processes with $\mathcal{L}(X_1) = \mathcal{L}(Y_1)$, and the assertion then follows from Sato (1999, Theorem 7.10 (iii)). $\square$

In the last part of this section we want to highlight the relationship between multivariate normal mean-variance mixtures and elliptical distributions. From a financial point of view, the latter are of some interest because they have the nice property that within this class the Value-at-Risk (VaR) is a coherent risk measure in the sense of Artzner, Delbaen, Eber, and Heath (1999) (this has been shown in Embrechts, McNeil, and Straumann (2002, Theorem 1), see also McNeil, Frey, and Embrechts (2005, Theorem 6.8)). However, in the following we only report some basic facts of elliptical distributions, for a more comprehensive overview we refer to Cambanis, Huang, and Simons (1981) and the book of Fang, Kotz, and Ng (1990). Our presentation here is inspired by McNeil, Frey, and Embrechts (2005, Section 3.3).

Since elliptical distributions emerge as a natural generalization of the class of spherical distributions, we first introduce the latter.

Definition 2.7 *An $\mathbb{R}^d$-valued random vector X has a* spherical distribution *if there exists a function $\psi : \mathbb{R}_+ \to \mathbb{R}$ such that the characteristic function $\phi_X(u) = \mathrm{E}\big[e^{i\langle u,X\rangle}\big]$ of X admits the representation*

$$\phi_X(u) = \psi\big(\langle u,u\rangle\big) \qquad \forall\, u \in \mathbb{R}^d.$$

The function $\psi(t)$ uniquely determining $\phi_X(u)$ is called characteristic generator *of the spherical distribution $\mathcal{L}(X)$, and the notation $X \sim S_d\big(\psi(t)\big)$ will be used.*

Remark: The probably most popular example of a spherical distribution is the multivariate standard normal distribution $N_d(\mathbf{0}, I_d)$. Its characteristic function is given by $\phi_{N_d(\mathbf{0},I_d)}(u) = e^{-\langle u,u\rangle/2}$, and from the above definition it immediately follows that $N_d(\mathbf{0}, I_d) = S_d\big(e^{-t/2}\big)$. Moreover, every normal variance mixture $F = N_d(\mathbf{0}, yI_d) \circ G$ is spherical because from the proof of Lemma 2.5 we know that $\phi_F(u) = \mathfrak{L}_G\big(\frac{\langle u,u\rangle}{2}\big)$, hence $N_d(\mathbf{0}, yI_d) \circ G = S_d\big(\mathfrak{L}_G\big(\frac{t}{2}\big)\big)$.

Not only the characteristic functions, but even the densities (if existent) of spherical distributions have a very special form as the following considerations show: Suppose $\mathcal{L}(X)$ is spherical and possesses a bounded Lebesgue density $f(x)$, then by the inversion formula we have

$$f(x) = \frac{1}{(2\pi)^d}\int_{\mathbb{R}^d} e^{-i\langle u,x\rangle}\phi_X(u)\,\mathrm{d}u = \frac{1}{(2\pi)^d}\int_{\mathbb{R}^d} e^{-i\langle u,x\rangle}\psi\big(\langle u,u\rangle\big)\,\mathrm{d}u.$$

Now if O is an orthogonal $d\times d$-matrix (that is, $OO^\top = O^\top O = I_d$), then $\langle O^\top u, O^\top u\rangle = \langle u,u\rangle$ and

$$\begin{aligned} f(Ox) &= \frac{1}{(2\pi)^d}\int_{\mathbb{R}^d} e^{-i\langle u,Ox\rangle}\psi\big(\langle u,u\rangle\big)\,\mathrm{d}u \\ &= \frac{1}{(2\pi)^d}\int_{\mathbb{R}^d} e^{-i\langle O^\top u,x\rangle}\psi\big(\langle O^\top u,O^\top u\rangle\big)\,\mathrm{d}u \\ &= \frac{1}{(2\pi)^d}\int_{\mathbb{R}^d} e^{-i\langle \bar u,x\rangle}\psi\big(\langle \bar u,\bar u\rangle\big)\,\mathrm{d}\bar u = f(x), \end{aligned}$$

where in the last line we made the coordinate transform $\bar{u} = O^\top u$ and used the fact that $|\det(O)| = |\det(O^\top)| = 1$. Since the equation $f(x) = f(Ox)$ holds for every orthogonal matrix O, $f(x)$ must necessarily be of the form

$$f(x) = h\big(\langle x, x\rangle\big) \qquad \text{for some function } h : \mathbb{R} \to \mathbb{R}_+ .$$

Consequently the level sets $\{x \in \mathbb{R}^d \,|\, f(x) = c\}$, $c > 0$, of f are exactly the $d-1$-dimensional hyperspheres $\{x \in \mathbb{R}^d \,|\, \langle x, x\rangle = \bar{c}\}$, $\bar{c} > 0$, which gave the class of spherical distributions its name. Now we turn to the more general case of elliptical distributions which are defined by

Definition 2.8 *An $\mathbb{R}^d$-valued random vector X has an* elliptical distribution *if there exists a function $\psi : \mathbb{R}_+ \to \mathbb{R}$, a symmetric, positive semidefinite $d \times d$-matrix Σ and some $\mu \in \mathbb{R}^d$ such that the characteristic function $\phi_X(u)$ of X admits the representation*

$$\phi_X(u) = e^{i\langle u,\mu\rangle}\psi\big(\langle u, \Sigma u\rangle\big) \qquad \forall\, u \in \mathbb{R}^d.$$

The elliptical distribution $\mathcal{L}(X)$ then is denoted by $E_d\big(\mu, \Sigma, \psi(t)\big)$.

Remark: Clearly $S_d\big(\psi(t)\big) = E_d\big(\mathbf{0}, I_d, \psi(t)\big)$, but in contrast to the spherical subclass the representation $E_d\big(\mu, \Sigma, \psi(t)\big)$ of elliptical distributions is not unique. Only μ is uniquely determined, but Σ and $\psi(t)$ are not: Setting $\bar{\Sigma} := c\Sigma$ and $\bar{\psi}(t) := \psi(\frac{t}{c})$ for some arbitrary $c > 0$ yields $E_d\big(\mu, \Sigma, \psi(t)\big) = E_d\big(\mu, \bar{\Sigma}, \bar{\psi}(t)\big)$.

A useful alternative characterization of elliptical distributions is provided by

Corollary 2.9 $X \sim E_d\big(\mu, \Sigma, \psi(t)\big)$ *if and only if*

$$X \stackrel{d}{=} \mu + AY$$

where $Y \sim S_d\big(\psi(t)\big)$ and A is a $d \times d$-matrix fulfilling $AA^\top = \Sigma$.

Proof: If $X \stackrel{d}{=} \mu + AY$ with $Y \sim S_d\big(\psi(t)\big)$ and $AA^\top = \Sigma$, then using Definition 2.7 we obtain

$$\phi_X(u) = e^{i\langle u,\mu\rangle}\phi_Y\big(A^\top u\big) = e^{i\langle u,\mu\rangle}\psi\big(\langle A^\top u, A^\top u\rangle\big) = e^{i\langle u,\mu\rangle}\psi\big(\langle u, \Sigma u\rangle\big),$$

hence $X \sim E_d\big(\mu, \Sigma, \psi(t)\big)$ by Definition 2.8. On the other hand, for every symmetric, positive semidefinite $d \times d$-matrix Σ it is always possible to find a $d \times d$-matrix A such that $AA^\top = \Sigma$. Thus we can also go through the above chain of equations from the right to the left which proves the "only if"-part. $\square$

Remark: Every spherically distributed random vector $Y \sim S_d\big(\psi(t)\big)$ has the alternative representation $Y \stackrel{d}{=} RS$, where R is an $\mathbb{R}_+$-valued random variable and S is a random vector which is independent of R and uniformly distributed on the unit sphere $\mathcal{S} := \{\xi \in \mathbb{R}^d \,|\, \|\xi\| = 1\}$ (see McNeil, Frey, and Embrechts 2005, Theorem 3.22). Hence $X \sim E_d\big(\mu, \Sigma, \psi(t)\big)$ if and only if $X \stackrel{d}{=} \mu + RAS$ with A as defined in the above corollary.

This allows us to draw some conclusions on the density $f(x)$ of an elliptically distributed random vector X which exists if $\mathcal{L}(Y)$ possesses a Lebesgue density and the matrix Σ is positive definite. In this case we have $Y = A^{-1}(X - \mu)$, and since $\mathcal{L}(Y)$ is spherical, its density must be of the form $h\big(\langle x,x\rangle\big)$ as shown above. Because $(A^{-1})^\top A^{-1} = \Sigma^{-1}$, the density of X is thus given by

$$f(x) = \frac{1}{\sqrt{\det(\Sigma)}}\, h\big(\langle x-\mu, \Sigma^{-1}(x-\mu)\rangle\big)$$

whose level sets obviously are the ellipsoids $\{x \in \mathbb{R}^d \mid \langle x-\mu, \Sigma^{-1}(x-\mu)\rangle = \bar{c}\}$, $\bar{c} > 0$. Therefore this class of distributions is sometimes also called "elliptically contoured distributions". The last corollary of this section shows their relation to the class of multivariate normal mean-variance mixtures.

Corollary 2.10 *A normal mean-variance mixture $F = N_d(\mu + y\beta, y\Delta) \circ G$ is an elliptical distribution if and only if $\beta = \mathbf{0}$, that is, if and only if it is a normal variance mixture.*

Proof: The proof of Lemma 2.5 implies that the characteristic function of F is given by $\phi_F(u) = e^{i\langle u,\mu\rangle}\mathfrak{L}_G\big(\frac{\langle u,\Delta u\rangle}{2} - i\langle u,\beta\rangle\big)$ which evidently has the representation $e^{i\langle u,\mu\rangle}\psi\big(\langle u,\Sigma u\rangle\big)$ required by Definition 2.8 with $\Sigma = \Delta$ and $\psi(t) = \mathfrak{L}_G\big(\frac{t}{2}\big)$ if and only if $\beta = \mathbf{0}$. □

2.2 Multivariate generalized hyperbolic distributions

Multivariate generalized hyperbolic distributions have already been introduced as a natural generalization of the univariate case at the end of the seminal paper Barndorff-Nielsen (1977) and were investigated further in Blæsild (1981) and Blæsild and Jensen (1981). They are defined as normal mean-variance mixtures with GIG mixing distributions in the following way:

$$GH_d(\lambda,\alpha,\beta,\delta,\mu,\Delta) := N_d(\mu + y\Delta\beta, y\Delta) \circ GIG\big(\lambda, \delta, \sqrt{\alpha^2 - \langle\beta,\Delta\beta\rangle}\big), \quad (2.1)$$

where it is usually assumed without loss of generality (see the remark on p. 68) that $\det(\Delta) = 1$ which we shall also do in the following if not stated otherwise. Due to the parameter restrictions of GIG distributions (see p. 8), the other GH parameters have to fulfill the constraints

$$\lambda \in \mathbb{R},\ \ \alpha,\delta \in \mathbb{R}_+,\ \ \beta,\mu \in \mathbb{R}^d \quad \text{and} \quad \begin{array}{ll} \delta \geq 0,\ 0 \leq \sqrt{\langle\beta,\Delta\beta\rangle} < \alpha, & \text{if } \lambda > 0, \\ \delta > 0,\ 0 \leq \sqrt{\langle\beta,\Delta\beta\rangle} < \alpha, & \text{if } \lambda = 0, \\ \delta > 0,\ 0 \leq \sqrt{\langle\beta,\Delta\beta\rangle} \leq \alpha, & \text{if } \lambda < 0. \end{array}$$

The meaning and influence of the parameters is essentially the same as in the univariate case (see p. 13). Again, parametrizations with $\delta = 0$, $\alpha = 0$ or $\sqrt{\langle\beta,\Delta\beta\rangle} = \alpha$ have to be understood as limiting cases.

Remark: Note that the above definition of multivariate GH distributions as normal mean-variance mixtures of the form $N_d(\mu + y\Delta\beta, y\Delta) \circ G$ is of course equivalent to the representation $N_d(\mu + y\tilde{\beta}, y\Delta) \circ G$ used in the previous section because the $d \times d$-matrix Δ is always regular by assumption. The modification of the mean term just simplifies some formulas as we shall see below.

For notational consistency with Chapter 1, the term $GH_d(\lambda,\alpha,\beta,\delta,\mu,\Delta)$ will be reserved for multivariate GH distributions with $\beta,\mu \in \mathbb{R}^d$, whereas $GH(\lambda,\alpha,\beta,\delta,\mu)$ denotes a univariate GH distribution with $\beta,\mu \in \mathbb{R}$ as before.

If $\delta > 0$ and $\sqrt{\langle \beta, \Delta\beta \rangle} < \alpha$, then the density of $GH_d(\lambda, \alpha, \beta, \delta, \mu, \Delta)$ can be obtained from (2.1) with the help of equations (1.2) and (A.1) as follows:

$$
\begin{aligned}
d_{GH_d(\lambda,\alpha,\beta,\delta,\mu,\Delta)}(x) &= \int_0^\infty d_{N_d(\mu+y\Delta\beta,y\Delta)}(x)\, d_{GIG\left(\lambda,\delta,\sqrt{\alpha^2-\langle\beta,\Delta\beta\rangle}\right)}(y)\,\mathrm{d}y \\
&\underset{(1.2)}{=} \int_0^\infty (2\pi y)^{-\frac{d}{2}}\, e^{-\frac{1}{2}\langle x-\mu-y\Delta\beta,(y\Delta)^{-1}(x-\mu-y\Delta\beta)\rangle}\, \frac{(\alpha^2-\langle\beta,\Delta\beta\rangle)^{\frac{\lambda}{2}}}{\delta^\lambda\, 2K_\lambda(\delta\sqrt{\alpha^2-\langle\beta,\Delta\beta\rangle})}\, y^{\lambda-1} \\
&\qquad\qquad \cdot e^{-\frac{1}{2}\left(\frac{\delta^2}{y}+(\alpha^2-\langle\beta,\Delta\beta\rangle)y\right)}\,\mathrm{d}y \\
&= \frac{(\alpha^2-\langle\beta,\Delta\beta\rangle)^{\frac{\lambda}{2}}}{(2\pi)^{\frac{d}{2}}\delta^\lambda\, 2K_\lambda(\delta\sqrt{\alpha^2-\langle\beta,\Delta\beta\rangle})}\, e^{\langle\beta,x-\mu\rangle} \\
&\qquad\qquad \cdot \int_0^\infty y^{\lambda-\frac{d}{2}-1}\, e^{-\frac{1}{2}\left(\frac{1}{y}(\langle x-\mu,\Delta^{-1}(x-\mu)\rangle+\delta^2)+\alpha^2 y\right)}\,\mathrm{d}y \\
&= \frac{(\alpha^2-\langle\beta,\Delta\beta\rangle)^{\frac{\lambda}{2}}}{(2\pi)^{\frac{d}{2}}\delta^\lambda\, 2K_\lambda(\delta\sqrt{\alpha^2-\langle\beta,\Delta\beta\rangle})}\, e^{\langle\beta,x-\mu\rangle} \left(\frac{\sqrt{\langle x-\mu,\Delta^{-1}(x-\mu)\rangle+\delta^2}}{\alpha}\right)^{\lambda-\frac{d}{2}} \\
&\qquad\qquad \cdot \int_0^\infty \bar{y}^{\lambda-\frac{d}{2}-1}\, e^{-\frac{1}{2}\left(\frac{1}{\bar{y}}+\bar{y}\right)\alpha\sqrt{\langle x-\mu,\Delta^{-1}(x-\mu)\rangle+\delta^2}}\,\mathrm{d}\bar{y} \\
&\underset{(A.1)}{=} \frac{(\alpha^2-\langle\beta,\Delta\beta\rangle)^{\frac{\lambda}{2}}}{(2\pi)^{\frac{d}{2}}\alpha^{\lambda-\frac{d}{2}}\delta^\lambda\, K_\lambda(\delta\sqrt{\alpha^2-\langle\beta,\Delta\beta\rangle})}\, \left(\langle x-\mu,\Delta^{-1}(x-\mu)\rangle+\delta^2\right)^{(\lambda-\frac{d}{2})/2} \\
&\qquad\qquad \cdot K_{\lambda-\frac{d}{2}}\left(\alpha\sqrt{\langle x-\mu,\Delta^{-1}(x-\mu)\rangle+\delta^2}\right) e^{\langle\beta,x-\mu\rangle}
\end{aligned} \tag{2.2}
$$

where in the last but one equation the substitution $\bar{y} = \frac{\alpha}{\sqrt{\langle x-\mu,\Delta^{-1}(x-\mu)\rangle+\delta^2}}y$ was made. The univariate density given in equations (1.6) and (1.7) is immediately obtained from (2.2) by setting $d = \Delta = 1$.

Remark: If the $d \times d$-matrix Δ is replaced by a matrix $\bar{\Delta}$ of the same dimensions with $\det(\bar{\Delta}) \neq 1$, then the normal density $d_{N_d(\mu+y\bar{\Delta}\beta,y\bar{\Delta})}(x)$ has an additional factor $\det(\bar{\Delta})^{-1/2}$ which will be incorporated in the norming constant of $d_{GH_d(\lambda,\alpha,\beta,\delta,\mu,\bar{\Delta})}(x)$ as the above calculation shows. Suppose $\bar{\Delta} = c^{1/d}\Delta$ for some $c > 0$, then $\det(\bar{\Delta}) = c$, and if we also replace $\lambda, \alpha, \beta, \delta, \mu$ by the barred parameters

$$\bar{\lambda} := \lambda, \quad \bar{\alpha} := c^{\frac{1}{2d}}\alpha, \quad \bar{\beta} := \beta, \quad \bar{\delta} := c^{-\frac{1}{2d}}\delta, \quad \bar{\mu} := \mu,$$

then it is easily seen from (2.2) that the densities of $GH_d(\lambda, \alpha, \beta, \delta, \mu, \Delta)$ and $GH_d(\bar{\lambda}, \bar{\alpha}, \bar{\beta}, \bar{\delta}, \bar{\mu}, \bar{\Delta})$ and thus both distributions coincide. Note that these considerations also remain true for all subsequently defined limit distributions. This again shows that the assumption $\det(\Delta) = 1$ is not an essential restriction. The barred parameters will be used later at some points in Section 2.4 to indicate that $\det(\bar{\Delta}) = 1$ is not assumed there.

If multivariate GH distributions would have been defined as a mixture of the form $N_d(\mu + y\beta, y\Delta) \circ GIG\left(\lambda, \delta, \sqrt{\alpha^2 - \langle\beta, \Delta\beta\rangle}\right)$ (see the remark on the previous page), then the last factor of the density (2.2) would be $e^{\langle\Delta^{-1}\beta,x-\mu\rangle}$ instead of $e^{\langle\beta,x-\mu\rangle}$, and $\bar{\beta}$ would have to be defined by $\bar{\beta} = c^{1/d}\beta$.

For some special choices of λ the density formula simplifies considerably using the representation (A.7) of the Bessel function $K_{\frac{1}{2}}(x)$. With $\lambda = -\frac{1}{2}$ one obtains the *multivariate normal inverse Gaussian distribution* $NIG_d(\alpha, \beta, \delta, \mu, \Delta)$ possessing the density

$$d_{NIG_d(\alpha,\beta,\delta,\mu,\Delta)}(x) = \sqrt{\frac{2}{\pi}}\,\frac{\delta\alpha^{\frac{d+1}{2}} e^{\delta\sqrt{\alpha^2-\langle\beta,\Delta\beta\rangle}}}{(2\pi)^{\frac{d}{2}}}\big(\langle x-\mu, \Delta^{-1}(x-\mu)\rangle + \delta^2\big)^{-\frac{d+1}{4}} \cdot K_{\frac{d+1}{2}}\big(\alpha\sqrt{\langle x-\mu, \Delta^{-1}(x-\mu)\rangle + \delta^2}\big)\, e^{\langle\beta, x-\mu\rangle},$$

and $\lambda = \frac{d+1}{2}$ yields the *d-dimensional hyperbolic distribution* $HYP_d(\alpha, \beta, \delta, \mu, \Delta)$ with density

$$d_{HYP_d(\alpha,\beta,\delta,\mu,\Delta)}(x) = \frac{(2\pi)^{-\frac{d-1}{2}}\,(\alpha^2 - \langle\beta, \Delta\beta\rangle)^{\frac{d+1}{4}}}{2\alpha\delta^{\frac{d+1}{2}} K_{\frac{d+1}{2}}(\delta\sqrt{\alpha^2-\langle\beta,\Delta\beta\rangle})}\, e^{-\alpha\sqrt{\langle x-\mu,\Delta^{-1}(x-\mu)\rangle+\delta^2}} \cdot e^{\langle\beta, x-\mu\rangle}.$$

The proof of Lemma 2.5, Proposition 1.9 and (2.1) imply that the characteristic function of $GH_d(\lambda, \alpha, \beta, \delta, \mu, \Delta)$ is given by

$$\begin{aligned}\phi_{GH_d(\lambda,\alpha,\beta,\delta,\mu,\Delta)}(u) &= e^{i\langle u,\mu\rangle}\mathfrak{L}_{GIG\left(\lambda,\delta,\sqrt{\alpha^2-\langle\beta,\Delta\beta\rangle}\right)}\Big(\tfrac{\langle u,\Delta u\rangle}{2} - i\langle u, \Delta\beta\rangle\Big) \\ &= e^{i\langle u,\mu\rangle}\left(\frac{\alpha^2-\langle\beta,\Delta\beta\rangle}{\alpha^2 - \langle\beta+iu, \Delta(\beta+iu)\rangle}\right)^{\frac{\lambda}{2}} \frac{K_\lambda\big(\delta\sqrt{\alpha^2-\langle\beta+iu,\Delta(\beta+iu)\rangle}\big)}{K_\lambda\big(\delta\sqrt{\alpha^2-\langle\beta,\Delta\beta\rangle}\big)}\end{aligned} \tag{2.3}$$

Since all GIG distributions and thus by Lemma 2.5 c) also all multivariate GH distributions (including the limits mentioned below) are infinitely divisible, the above characteristic function alternatively admits a Lévy–Khintchine representation which will be derived in Section 2.3.

Let us briefly mention possible weak limits of multivariate GH distributions here. If $\lambda > 0$ and $\delta \to 0$, then by (2.1), (1.3) and Lemma 2.5 e) we have convergence to a *multivariate Variance-Gamma distribution*:

$$GH_d(\lambda, \alpha, \beta, \delta, \mu, \Delta) \xrightarrow{w} N_d(\mu + y\Delta\beta, y\Delta) \circ G\big(\lambda, \tfrac{\alpha^2-\langle\beta,\Delta\beta\rangle}{2}\big) = VG_d(\lambda, \alpha, \beta, \mu, \Delta).$$

The corresponding density can easily be derived from equation (2.2) observing that $\delta^\lambda K_\lambda(\delta\sqrt{\alpha^2-\langle\beta,\Delta\beta\rangle}) \to 2^{\lambda-1}\Gamma(\lambda)(\alpha^2-\langle\beta,\Delta\beta\rangle)^{-\lambda/2}$ for $\delta \to 0$ by (A.8) which yields

$$d_{VG_d(\lambda,\alpha,\beta,\mu,\Delta)}(x) = \frac{(\alpha^2-\langle\beta,\Delta\beta\rangle)^\lambda}{(2\pi)^{\frac{d}{2}}\alpha^{\lambda-\frac{d}{2}}2^{\lambda-1}\Gamma(\lambda)}\big(\langle x-\mu, \Delta^{-1}(x-\mu)\rangle\big)^{(\lambda-\frac{d}{2})/2} \cdot K_{\lambda-\frac{d}{2}}\big(\alpha\sqrt{\langle x-\mu,\Delta^{-1}(x-\mu)\rangle}\big)\, e^{\langle\beta,x-\mu\rangle}, \tag{2.4}$$

and the characteristic function is obtained by

$$\begin{aligned}\phi_{VG_d(\lambda,\alpha,\beta,\mu,\Delta)}(u) &= e^{i\langle u,\mu\rangle}\mathfrak{L}_{G\left(\lambda,\frac{\alpha^2-\langle\beta,\Delta\beta\rangle}{2}\right)}\Big(\tfrac{\langle u,\Delta u\rangle}{2} - i\langle u,\Delta\beta\rangle\Big) \\ &= e^{i\langle u,\mu\rangle}\left(\frac{\alpha^2-\langle\beta,\Delta\beta\rangle}{\alpha^2-\langle\beta+iu,\Delta(\beta+iu)\rangle}\right)^\lambda.\end{aligned} \tag{2.5}$$

For $\lambda < 0$ and $\alpha \to 0$ as well as $\beta \to \mathbf{0}$ we arrive at the multivariate scaled and shifted t *distribution* with $f = -2\lambda$ degrees of freedom:

$$GH_d(\lambda, \alpha, \beta, \delta, \mu, \Delta) \xrightarrow{w} N_d(\mu, y\Delta) \circ iG\big(\lambda, \tfrac{\delta^2}{2}\big) = t_d(\lambda, \delta, \mu, \Delta).$$

Its density can be calculated as follows:

$$\begin{aligned} d_{t_d(\lambda,\delta,\mu,\Delta)}(x) &= \int_0^\infty d_{N_d(\mu,y\Delta)}(x)\, d_{iG(\lambda,\frac{\delta^2}{2})}(y)\,\mathrm{d}y \\ &\underset{(1.4)}{=} \int_0^\infty (2\pi y)^{-\frac{d}{2}}\, e^{-\frac{1}{2}\langle x-\mu,(y\Delta)^{-1}(x-\mu)\rangle} \left(\frac{2}{\delta^2}\right)^\lambda \frac{y^{\lambda-1}}{\Gamma(-\lambda)}\, e^{-\frac{\delta^2}{2y}}\,\mathrm{d}y \\ &= \left(\frac{2}{\delta^2}\right)^\lambda \frac{(2\pi)^{-\frac{d}{2}}}{\Gamma(-\lambda)} \int_0^\infty y^{\lambda-\frac{d}{2}-1}\, e^{-\frac{1}{2}\left(\frac{1}{y}(\langle x-\mu,\Delta^{-1}(x-\mu)\rangle+\delta^2)\right)}\,\mathrm{d}y \\ &= \frac{\Gamma\big(-\lambda+\frac{d}{2}\big)}{(\delta^2\pi)^{\frac{d}{2}}\,\Gamma(-\lambda)} \left(1+\frac{\langle x-\mu,\Delta^{-1}(x-\mu)\rangle}{\delta^2}\right)^{\lambda-\frac{d}{2}}, \end{aligned} \tag{2.6}$$

where the last line follows from the fact that the integrand in the last but one equation is equal to the density of $iG\big(\lambda - \frac{d}{2}, \sqrt{\langle x-\mu, \Delta^{-1}(x-\mu)\rangle + \delta^2}\big)$ without the corresponding norming constant $\big(2/(\langle x-\mu, \Delta^{-1}(x-\mu)\rangle+\delta^2)\big)^{\lambda-d/2}\cdot$ $\Gamma(-\lambda+\frac{d}{2})^{-1}$. Hence the value of the integral must equal the inverse of the latter. The characteristic function is given by

$$\phi_{t_d(\lambda,\delta,\mu,\Delta)}(u) = e^{i\langle u,\mu\rangle}\,\mathfrak{L}_{iG\left(\lambda,\frac{\delta^2}{2}\right)}\Big(\tfrac{\langle u,\Delta u\rangle}{2}\Big) = e^{i\langle u,\mu\rangle}\left(\frac{2}{\delta}\right)^\lambda \frac{2K_\lambda\big(\delta\sqrt{\langle u,\Delta u\rangle}\big)}{\Gamma(-\lambda)\,(\langle u,\Delta u\rangle)^{\frac{\lambda}{2}}}. \tag{2.7}$$

If $\lambda < 0$, but $\langle \beta, \Delta\beta\rangle \to \alpha^2$, then we have weak convergence to the normal mean-variance mixture

$$GH_d(\lambda, \alpha, \beta, \delta, \mu, \Delta) \xrightarrow{w} N_d(\mu + y\Delta\beta, y\Delta) \circ iG\big(\lambda, \tfrac{\delta^2}{2}\big).$$

Combining the arguments leading to (2.2) and (2.6), its density is seen to be

$$\begin{aligned} d_{GH_d(\lambda,\sqrt{\langle\beta,\Delta\beta\rangle},\beta,\delta,\mu,\Delta)}(x) = {} & \frac{2^{\lambda+1-\frac{d}{2}}\,\delta^{-2\lambda}}{\pi^{\frac{d}{2}}\Gamma(-\lambda)\alpha^{\lambda-\frac{d}{2}}} \big(\langle x-\mu,\Delta^{-1}(x-\mu)\rangle+\delta^2\big)^{(\lambda-\frac{d}{2})/2} \\ & \cdot K_{\lambda-\frac{d}{2}}\big(\alpha\sqrt{\langle x-\mu,\Delta^{-1}(x-\mu)\rangle+\delta^2}\big)\, e^{\langle\beta,x-\mu\rangle} \end{aligned} \tag{2.8}$$

where $\alpha = \sqrt{\langle\beta, \Delta\beta\rangle}$, and the corresponding characteristic function is obtained similarly as in the t limiting case to be

$$\phi_{GH_d(\lambda,\sqrt{\langle\beta,\Delta\beta\rangle},\beta,\delta,\mu,\Delta)}(u) = \left(\frac{2}{\delta}\right)^\lambda \frac{2K_\lambda\big(\delta\sqrt{\langle u,\Delta u\rangle - 2i\langle u,\Delta\beta\rangle}\big)}{\Gamma(-\lambda)\,\big(\langle u,\Delta u\rangle - 2i\langle u,\Delta\beta\rangle\big)^{\frac{\lambda}{2}}}. \tag{2.9}$$

Last but not least also the multivariate normal distribution emerges as a weak limit if $\alpha \to \infty$, $\delta \to \infty$ and $\frac{\delta}{\alpha} \to \sigma^2 < \infty$. Analogously as in the univariate case, these assumptions entail, together with Corollary 1.10 and Lemma 2.5 e), that

$$GH_d(\lambda, \alpha, \beta, \delta, \mu, \Delta) \xrightarrow{w} N_d(\mu + y\Delta\beta, y\Delta) \circ \epsilon_{\sigma^2} = N_d(\mu + \sigma^2\Delta\beta, \sigma^2\Delta).$$

The most important properties of multivariate GH distributions are summarized in the following theorem which goes back to Blæsild (1981, Theorem 1), see also Blæsild and Jensen (1981, p. 49f). It shows that this distribution class is closed under forming marginals, conditioning and affine transformations.

Theorem 2.11 *Suppose $X \sim GH_d(\lambda, \alpha, \beta, \delta, \mu, \Delta)$. Let $(X_1, X_2)^\top$ be a partition of X where X_1 has the dimension r and X_2 the dimension $k = d - r$, and let $(\beta_1, \beta_2)^\top$ and $(\mu_1, \mu_2)^\top$ be similar partitions of β and μ. Furthermore, let*

$$\Delta = \begin{pmatrix} \Delta_{11} & \Delta_{12} \\ \Delta_{21} & \Delta_{22} \end{pmatrix}$$

be a partition of Δ such that Δ_{11} is an $r \times r$-matrix. Then the following holds:

a) $X_1 \sim GH_r(\lambda^*, \alpha^*, \beta^*, \delta^*, \mu^*, \Delta^*)$ with starred parameters given by $\lambda^* = \lambda$, $\alpha^* = \det(\Delta_{11})^{-\frac{1}{2r}} \sqrt{\alpha^2 - \langle \beta_2, (\Delta_{22} - \Delta_{21}\Delta_{11}^{-1}\Delta_{12})\beta_2 \rangle}$, $\beta^* = \beta_1 + \Delta_{11}^{-1}\Delta_{12}\beta_2$, $\delta^* = \det(\Delta_{11})^{\frac{1}{2r}}\delta$, $\mu^* = \mu_1$ and $\Delta^* = \det(\Delta_{11})^{-\frac{1}{r}}\Delta_{11}$.

b) The conditional distribution of X_2 given $X_1 = x_1$ is $GH_k(\tilde{\lambda}, \tilde{\alpha}, \tilde{\beta}, \tilde{\delta}, \tilde{\mu}, \tilde{\Delta})$ with tilded parameters given by $\tilde{\lambda} = \lambda - \frac{r}{2}$, $\tilde{\alpha} = \det(\Delta_{11})^{\frac{1}{2k}}\alpha$, $\tilde{\beta} = \beta_2$, $\tilde{\delta} = \det(\Delta_{11})^{-\frac{1}{2k}} \sqrt{\delta^2 + \langle x_1 - \mu_1, \Delta_{11}^{-1}(x_1 - \mu_1) \rangle}$, $\tilde{\mu} = \mu_2 + \Delta_{21}\Delta_{11}^{-1}(x_1 - \mu_1)$ and $\tilde{\Delta} = \det(\Delta_{11})^{\frac{1}{k}}(\Delta_{22} - \Delta_{21}\Delta_{11}^{-1}\Delta_{12})$.

c) Suppose $Y = BX + b$ where B is a regular $d \times d$-matrix and $b \in \mathbb{R}^d$, then $Y \sim GH_d(\hat{\lambda}, \hat{\alpha}, \hat{\beta}, \hat{\delta}, \hat{\mu}, \hat{\Delta})$ where $\hat{\lambda} = \lambda$, $\hat{\alpha} = |\det(B)|^{-\frac{1}{d}}\alpha$, $\hat{\beta} = (B^{-1})^\top\beta$, $\hat{\delta} = |\det(B)|^{\frac{1}{d}}\delta$, $\hat{\mu} = B\mu + b$ and $\hat{\Delta} = |\det(B)|^{-\frac{2}{d}} B\Delta B^\top$.

Remark: An important fact we want to stress here is that the above theorem remains also valid for all multivariate GH limit distributions considered before. Thus one can in particular conclude from part b) that the limiting subclass of VG distributions itself is, in contrast to the t limit distributions, not closed under conditioning. This holds because the parameter $\tilde{\delta}$ of the conditional distribution in general is greater than zero, and the parameter $\tilde{\lambda} = \lambda - \frac{r}{2}$ may become negative if the subdimension r is sufficiently large.

Moreover, all margins of $t_d(\lambda, \delta, \mu, \Delta)$ are again t distributed $t_r(\lambda, \delta^*, \mu^*, \Delta^*)$ because if the joint distribution has the parameters $\alpha = 0$ and $\beta = \mathbf{0}$, part a) of the theorem implies that $\alpha^* = 0$ and $\beta^* = \mathbf{0}$ for every marginal distribution. Similarly, all margins of $VG_d(\lambda, \alpha, \beta, \mu, \Delta)$ are again VG distributions because if $\delta = 0$, then also $\delta^* = 0$. In addition it can be shown that all margins of $GH_d\big(\lambda, \sqrt{\langle \beta, \Delta\beta \rangle}, \beta, \delta, \mu, \Delta\big)$-distributions are of the same limiting type as their joint distribution, too.

Further note that a d-dimensional hyperbolic distribution does not have univariate hyperbolic margins because part a) of the theorem states that the parameter λ^* of the marginal distributions is $\lambda^* = \lambda = \frac{d+1}{2} \neq 1$ for $d \geq 2$.

Since every linear mapping $\tilde{X} = aX + b$ with $a \in \mathbb{R} \setminus \{0\}$, $b \in \mathbb{R}^d$, of a GH distributed random vector $X \sim GH_d(\lambda, \alpha, \beta, \delta, \mu, \Delta)$ can alternatively be written in the form $\tilde{X} = BX + b$ where $B = aI_d$ and hence $\det(B) = a^d$, it follows from Theorem 2.11 c) that $\tilde{X} \sim GH_d(\lambda, |a|^{-1}\alpha, a^{-1}\beta, |a|\delta, a\mu + b, \Delta)$. Thus the parameters

$$\zeta := \delta\sqrt{\alpha^2 - \langle\beta, \Delta\beta\rangle} \qquad \text{and} \qquad \Pi := \frac{\Delta\beta}{\sqrt{\alpha^2 - \langle\beta, \Delta\beta\rangle}} \tag{2.10}$$

which may replace α and β are scale- and location-invariant. The definition and notation of the invariant parameters of multivariate GH distributions is, however, by no means unique in the literature.

The next corollary deals with the special case of linear transforms $\langle h, X\rangle$, $h \in \mathbb{R}^d$, of GH distributed random vectors X. The corresponding distributions can, of course, also be derived with the help of parts a) and c) of Theorem 2.11, but we shall give a short alternative proof below based on characteristic functions.

Corollary 2.12 *Let $X \sim GH_d(\lambda, \alpha, \beta, \delta, \mu, \Delta)$ and $Y := \langle h, X\rangle$ with $h \in \mathbb{R}^d$, $h \neq \mathbf{0}$, then $Y \sim GH(\hat{\lambda}, \hat{\alpha}, \hat{\beta}, \hat{\delta}, \hat{\mu})$ where $\hat{\lambda} = \lambda$, $\hat{\alpha} = \sqrt{\frac{\alpha^2 - \langle\beta,\Delta\beta\rangle}{\langle h,\Delta h\rangle} + \left(\frac{\langle h,\Delta\beta\rangle}{\langle h,\Delta h\rangle}\right)^2}$, $\hat{\beta} = \frac{\langle h,\Delta\beta\rangle}{\langle h,\Delta h\rangle}$, $\hat{\delta} = \delta\sqrt{\langle h, \Delta h\rangle}$ and $\hat{\mu} = \langle h, \mu\rangle$.*

Proof: We have $\phi_Y(y) = \mathrm{E}[e^{iy\langle h,X\rangle}] = \phi_X(yh)$ and thus by (2.3)

$$\begin{aligned}
&\phi_Y(y) = \\
&= e^{i\langle yh,\mu\rangle}\left(\frac{\alpha^2 - \langle\beta, \Delta\beta\rangle}{\alpha^2 - \langle\beta + iyh, \Delta(\beta + iyh)\rangle}\right)^{\frac{\lambda}{2}} \frac{K_\lambda\big(\delta\sqrt{\alpha^2 - \langle\beta + iyh, \Delta(\beta + iyh)\rangle}\big)}{K_\lambda\big(\delta\sqrt{\alpha^2 - \langle\beta, \Delta\beta\rangle}\big)} \\
&= e^{iy\langle h,\mu\rangle}\left(\frac{\langle h, \Delta h\rangle\left(\frac{\alpha^2-\langle\beta,\Delta\beta\rangle}{\langle h,\Delta h\rangle} + \left(\frac{\langle h,\Delta\beta\rangle}{\langle h,\Delta h\rangle}\right)^2 - \left(\frac{\langle h,\Delta\beta\rangle}{\langle h,\Delta h\rangle}\right)^2\right)}{\langle h, \Delta h\rangle\left(\frac{\alpha^2-\langle\beta,\Delta\beta\rangle}{\langle h,\Delta h\rangle} + \left(\frac{\langle h,\Delta\beta\rangle}{\langle h,\Delta h\rangle}\right)^2 - \left(\frac{\langle h,\Delta\beta\rangle}{\langle h,\Delta h\rangle} + iy\right)^2\right)}\right)^{\frac{\lambda}{2}} \\
&\quad\cdot \frac{K_\lambda\left(\delta\sqrt{\langle h, \Delta h\rangle}\sqrt{\frac{\alpha^2-\langle\beta,\Delta\beta\rangle}{\langle h,\Delta h\rangle} + \left(\frac{\langle h,\Delta\beta\rangle}{\langle h,\Delta h\rangle}\right)^2 - \left(\frac{\langle h,\Delta\beta\rangle}{\langle h,\Delta h\rangle} + iy\right)^2}\right)}{K_\lambda\left(\delta\sqrt{\langle h, \Delta h\rangle}\sqrt{\frac{\alpha^2-\langle\beta,\Delta\beta\rangle}{\langle h,\Delta h\rangle} + \left(\frac{\langle h,\Delta\beta\rangle}{\langle h,\Delta h\rangle}\right)^2 - \left(\frac{\langle h,\Delta\beta\rangle}{\langle h,\Delta h\rangle}\right)^2}\right)} \\
&= e^{iy\hat{\mu}}\left(\frac{\hat{\alpha}^2 - \hat{\beta}^2}{\hat{\alpha}^2 - (\hat{\beta} + iy)^2}\right)^{\frac{\lambda}{2}} \frac{K_\lambda\left(\hat{\delta}\sqrt{\hat{\alpha}^2 - (\hat{\beta} + iy)^2}\right)}{K_\lambda\left(\hat{\delta}\sqrt{\hat{\alpha}^2 - \hat{\beta}^2}\right)} = \phi_{GH(\hat{\lambda},\hat{\alpha},\hat{\beta},\hat{\delta},\hat{\mu})}(y).
\end{aligned}$$

Analogous calculations using the characteristic functions given in (2.5), (2.7), and (2.9) show that the assertion remains valid in all corresponding limiting cases. □

Let us finally take a closer look at the moments of multivariate GH distributions. By Definition 2.4, every random variable X possessing a multivariate normal mean-variance mixture distribution admits the stochastic representation $X \stackrel{d}{=} \mu + Z\beta + \sqrt{Z}AW$ with independent random variables Z and $W \sim N_d(\mathbf{0}, I_d)$. The standardization of W and its independence from Z imply that

$$\begin{aligned} \mathrm{E}(X) &= \mu + \mathrm{E}(Z)\beta \\ \mathrm{Cov}(X) &= \mathrm{E}\big[(X - \mathrm{E}(X))(X - \mathrm{E}(X))^\top\big] = \mathrm{E}(Z)\Delta + \mathrm{Var}(Z)\beta\beta^\top \end{aligned} \tag{2.11}$$

with $\Delta = AA^\top$, provided that $\mathrm{E}(|Z|), \mathrm{Var}(Z) < \infty$. If $X \sim GH_d(\lambda, \alpha, \beta, \delta, \mu, \Delta)$, then by (2.1) $X \stackrel{d}{=} \mu + Z\Delta\beta + \sqrt{Z}AW$ and $Z \sim GIG\big(\lambda, \delta, \sqrt{\alpha^2 - \langle\beta, \Delta\beta\rangle}\big)$. Using the moment formulas on p. 11 we get explicit expressions for $\mathrm{E}(Z)$ and $\mathrm{Var}(Z)$ which can be inserted into the general equations above to finally obtain

$$\begin{aligned} \mathrm{E}[GH_d(\lambda, \alpha, \beta, \delta, \mu, \Delta)] &= \mu + \frac{\delta^2}{\zeta}\frac{K_{\lambda+1}(\zeta)}{K_\lambda(\zeta)}\beta, \\ \mathrm{Cov}[GH_d(\lambda, \alpha, \beta, \delta, \mu, \Delta)] &= \frac{\delta^2}{\zeta}\frac{K_{\lambda+1}(\zeta)}{K_\lambda(\zeta)}\Delta + \frac{\delta^4}{\zeta^2}\left(\frac{K_{\lambda+2}(\zeta)}{K_\lambda(\zeta)} - \frac{K^2_{\lambda+1}(\zeta)}{K^2_\lambda(\zeta)}\right)\beta\beta^\top \end{aligned} \tag{2.12}$$

with $\zeta = \delta\sqrt{\alpha^2 - \langle\beta, \Delta\beta\rangle}$ as defined in (2.10). In case of the Variance-Gamma limits we have

$$\begin{aligned} \mathrm{E}[VG_d(\lambda, \alpha, \beta, \mu, \Delta)] &= \mu + \frac{2\lambda}{\alpha^2 - \langle\beta, \Delta\beta\rangle}\beta, \\ \mathrm{Cov}[VG_d(\lambda, \alpha, \beta, \mu, \Delta)] &= \frac{2\lambda}{\alpha^2 - \langle\beta, \Delta\beta\rangle}\Delta + \frac{4\lambda}{(\alpha^2 - \langle\beta, \Delta\beta\rangle)^2}\beta\beta^\top. \end{aligned} \tag{2.13}$$

Observe that by Lemma 2.5 both multivariate GH and VG distributions possess moment generating functions and hence finite moments of arbitrary order because the mixing GIG and Gamma distributions do have this property. This is no longer true for the limit distributions with $\lambda < 0$ because the corresponding inverse Gamma mixing distributions only have finite moments up to order $r < -\lambda$. By Theorem 2.11 a), the marginal distributions of $t_d(\lambda, \delta, \mu, \Delta)$ are given by $t(\lambda, \sqrt{\Delta_{ii}}\delta, \mu_i)$, $1 \le i \le d$ (recall that $\alpha = 0$ and $\beta = \mathbf{0}$ in this case), and from their tail behaviour (see p. 22/23) one can easily conclude that mean vector and covariance matrix of the t limit distributions are well defined and finite only if $\lambda < -\frac{1}{2}$ resp. $\lambda < -1$. If these constraints are fulfilled, then

$$\mathrm{E}[t_d(\lambda, \delta, \mu, \Delta)] = \mu \qquad \text{and} \qquad \mathrm{Cov}[t_d(\lambda, \delta, \mu, \Delta)] = \frac{\delta^2}{-2\lambda - 2}\Delta\,. \tag{2.14}$$

In the other limiting case where $\langle\beta, \Delta\beta\rangle = \alpha^2 > 0$ equations (2.11) state that necessary and sufficient conditions for the existence of a mean vector and covariance matrix of the limit distributions are that the inverse Gamma mixing distributions have finite means and variances which holds true if and only if $\lambda < -1$ and $\lambda < -2$, respectively. If λ is appropriately small, then

$$\begin{aligned} \mathrm{E}\big[GH_d\big(\lambda, \sqrt{\langle\beta, \Delta\beta\rangle}, \beta, \delta, \mu, \Delta\big)\big] &= \mu + \frac{\delta^2}{-2\lambda - 2}\beta, \\ \mathrm{Cov}\big[GH_d\big(\lambda, \sqrt{\langle\beta, \Delta\beta\rangle}, \beta, \delta, \mu, \Delta\big)\big] &= \frac{\delta^2}{-2\lambda - 2}\Delta + \frac{\delta^4}{4(\lambda+1)^2(-\lambda-2)}\beta\beta^\top. \end{aligned} \tag{2.15}$$

2.3 Lévy–Khintchine representations of multivariate GH distributions

It was already mentioned on page 74 that also all higher-dimensional GH distributions are infinitely divisible because they are defined as multivariate normal mean-variance mixtures with an infinitely divisible mixing distribution. In the present section we derive the corresponding Lévy–Khintchine representations of their characteristic functions which allows us in particular to characterize the subclass of multivariate VG distributions that are in addition selfdecomposable. It will turn out that, in contrast to the univariate case, in general only symmetric VG distributions with $\beta = \mathbf{0}$ do have this property.

The key result for the subsequent analysis is the following theorem from Sato (1999, Theorem 30.1) which we cite here (with a slightly adapted notation) for further reference.

Theorem 2.13 (Subordinated Lévy processes) *Let $(Z_t)_{t\geq 0}$ be an increasing Lévy process on $\mathbb{R}$ with Lévy measure $\rho(\mathrm{d}y)$, drift b_0 and $P^{Z_s} =: G_s$, that is,*

$$\mathrm{E}\big[e^{-vZ_s}\big] = \mathfrak{L}_{G_s}(v) = \int_0^\infty e^{-vy}\, G_s(\mathrm{d}y) = e^{s\Psi(-v)}, \qquad v \geq 0,$$

where for any complex z with $Re(z) \leq 0$

$$\Psi(z) = b_0 z + \int_0^\infty \big(e^{zy} - 1\big)\, \rho(\mathrm{d}y) \quad \textit{with} \;\; b_0 \geq 0 \;\; \textit{and} \quad \int_0^\infty (1 \wedge y)\, \rho(\mathrm{d}y) < \infty.$$

Let further $(X_t)_{t\geq 0}$ be a Lévy process on $\mathbb{R}^d$ with characteristic triplet (b, C, ν) and $P^{X_s} =: \Phi_s$. Suppose $(X_t)_{t\geq 0}$ and $(Z_t)_{t\geq 0}$ are independent and define

$$Y_t(\omega) := X_{Z_t(\omega)}(\omega), \qquad t \geq 0.$$

Then $(Y_t)_{t\geq 0}$ is a Lévy process on $\mathbb{R}^d$ with

$$P[Y_t \in B] = \int_0^\infty \Phi_s(B)\, G_t(\mathrm{d}s), \qquad B \in \mathcal{B}^d,$$

$$\phi_{Y_t}(u) = \mathrm{E}\big[e^{i\langle u, Y_t\rangle}\big] = e^{t\Psi(\log(\phi_{X_1}(u)))},$$

and the characteristic triplet $(\tilde{b}, \tilde{C}, \tilde{\nu})$ of Y is given by

$$\tilde{b} = b_0 b + \int_0^\infty \int_{\{x\in\mathbb{R}^d \,|\, \|x\|\leq 1\}} x\, \Phi_s(\mathrm{d}x)\, \rho(\mathrm{d}s), \tag{2.16}$$

$$\tilde{C} = b_0 C, \tag{2.17}$$

$$\tilde{\nu}(B) = b_0 \nu(B) + \int_0^\infty \Phi_s(B)\, \rho(\mathrm{d}s), \qquad B \in \mathcal{B}\big(\mathbb{R}^d \setminus \{\mathbf{0}\}\big). \tag{2.18}$$

Remark: The characteristic triplets (b, C, ν) and $(\tilde{b}, \tilde{C}, \tilde{\nu})$ in the above theorem correspond to the Lévy–Khintchine representation

$$\phi_\mu(u) = \exp\left(i\langle u, b\rangle - \frac{1}{2}\langle u, Cu\rangle + \int_{\mathbb{R}^d} \big(e^{i\langle u,x\rangle} - 1 - i\langle u, x\rangle \mathbb{1}_{\{\|x\|\leq 1\}}(x)\big)\nu(\mathrm{d}x)\right)$$

with $\mu = P^{X_1}$ or $\mu = P^{Y_1}$. But analogously as in the one-dimensional case one has the equivalence $\int_{\mathbb{R}^d} \|x\|^r \, \mu(\mathrm{d}x) < \infty \Leftrightarrow \int_{\{x \,|\, \|x\|>1\}} \|x\|^r \, \nu(\mathrm{d}x) < \infty$ (Sato 1999, Theorem 25.3 and Proposition 25.4), so if all marginal distributions of μ possess finite first moments, the truncation function within the integral term of the Lévy–Khintchine formula can again be omitted, implying that at the same time the drift vector b is modified to

$$\bar{b} = b + \int_{\{x \,|\, \|x\|>1\}} x \, \nu(\mathrm{d}x) = \mathrm{E}[\mu]$$

where the integral has to be understood componentwise (see also Sato 1999, Example 25.12).

Before applying Theorem 2.13 to the special case of multivariate GH distributions resp. Lévy processes, we first have to calculate the required alternative representation of the Laplace transforms of GIG distributions. This is done in the following

Lemma 2.14 *The Laplace transforms of $GIG(\lambda,\delta,\gamma)$-distributions, including the weak limits with parameters $\delta = 0$ and $\gamma = 0$, can be written in the form*

$$\mathfrak{L}_{GIG(\lambda,\delta,\gamma)}(v) = \exp\left(\int_0^\infty \left(e^{-vy} - 1\right) g_{GIG(\lambda,\delta,\gamma)}(y)\, \mathrm{d}y\right)$$

where $g_{GIG(\lambda,\delta,\gamma)}(y)$ denotes the density of the corresponding Lévy measure.

Proof: It has beeen shown in Chapter 1.6, Proposition 1.23, that all GIG distributions and the weak Gamma and inverse Gamma limits are generalized Gamma convolutions with generating pairs $(0, U_{GIG(\lambda,\delta,\gamma)})$. By Definition 1.19, their characteristic functions thus admit the representation

$$\phi_{GIG(\lambda,\delta,\gamma)}(v) = \exp\left[-\int_0^\infty \ln\left(1 - \frac{iv}{y}\right) U_{GIG(\lambda,\delta,\gamma)}(\mathrm{d}y)\right].$$

Because $\mathfrak{L}_{GIG(\lambda,\delta,\gamma)}(v) = \phi_{GIG(\lambda,\delta,\gamma)}(iv)$ (see the remark on p. 10), we have

$$\begin{aligned}
&\ln\bigl(\mathfrak{L}_{GIG(\lambda,\delta,\gamma)}(v)\bigr) = -\int_0^\infty \ln\left(1 + \frac{v}{y}\right) U_{GIG(\lambda,\delta,\gamma)}(\mathrm{d}y) \\
&= -\int_0^\infty \bigl(\ln(v+y) - \ln(y)\bigr)\, U_{GIG(\lambda,\delta,\gamma)}(\mathrm{d}y) = -\int_0^\infty \int_y^{v+y} \frac{1}{t}\, \mathrm{d}t\, U_{GIG(\lambda,\delta,\gamma)}(\mathrm{d}y) \\
&= -\int_0^\infty \int_y^{v+y} \int_0^\infty e^{-st}\, \mathrm{d}s\, \mathrm{d}t\, U_{GIG(\lambda,\delta,\gamma)}(\mathrm{d}y) \\
&= -\int_0^\infty \int_0^\infty \int_y^{v+y} e^{-st}\, \mathrm{d}t\, U_{GIG(\lambda,\delta,\gamma)}(\mathrm{d}y)\, \mathrm{d}s \\
&= -\int_0^\infty \int_0^\infty \frac{e^{-sy} - e^{-s(v+y)}}{s}\, U_{GIG(\lambda,\delta,\gamma)}(\mathrm{d}y)\, \mathrm{d}s \\
&= \int_0^\infty \bigl(e^{-vs} - 1\bigr) \frac{1}{s} \int_0^\infty e^{-sy}\, U_{GIG(\lambda,\delta,\gamma)}(\mathrm{d}y)\, \mathrm{d}s = \int_0^\infty \bigl(e^{-vs} - 1\bigr)\, g_{GIG(\lambda,\delta,\gamma)}(s)\, \mathrm{d}s
\end{aligned}$$

where the last equation follows from the proof of Proposition 1.20 (see also the proof of Proposition 1.24). Hence the desired representation of the Laplace transforms is valid for all GIG distributions with finite parameters $\delta, \gamma \geq 0$. $\square$

With the help of the preceeding prerequisites we are now ready to derive the announced Lévy–Khintchine representations of multivariate GH distributions and its weak limits.

Proposition 2.15 *The characteristic functions of $GH_d(\lambda, \alpha, \beta, \delta, \mu, \Delta)$-distributions can be represented as follows:*

a) If $\delta > 0$ and $\langle \beta, \Delta\beta \rangle < \alpha^2$, then

$$\phi_{GH_d(\lambda,\alpha,\beta,\delta,\mu,\Delta)}(u) = \\ = \exp\left[i\langle u, \mathrm{E}[GH_d]\rangle + \int_{\mathbb{R}^d} \big(e^{i\langle u,x\rangle} - 1 - i\langle u, x\rangle\big)\, g_{GH_d(\lambda,\alpha,\beta,\delta,\mu,\Delta)}(x)\,\mathrm{d}x \right]$$

where $\mathrm{E}[GH_d]$ is given in (2.12), and the density of the Lévy measure is

$$g_{GH_d(\lambda,\alpha,\beta,\delta,\mu,\Delta)}(x) = \\ = \frac{2e^{\langle x,\beta\rangle}}{\big(2\pi\sqrt{\langle x, \Delta^{-1}x\rangle}\big)^{\frac{d}{2}}} \left[\int_0^\infty \frac{(2y+\alpha^2)^{\frac{d}{4}} K_{\frac{d}{2}}\big(\sqrt{(2y+\alpha^2)\langle x, \Delta^{-1}x\rangle}\big)}{\pi^2 y \big[J^2_{|\lambda|}(\delta\sqrt{2y}) + Y^2_{|\lambda|}(\delta\sqrt{2y})\big]}\,\mathrm{d}y \right. \\ \left. + \max(0,\lambda)\,\alpha^{\frac{d}{2}} K_{\frac{d}{2}}\big(\alpha\sqrt{\langle x, \Delta^{-1}x\rangle}\big) \right]. \tag{2.19}$$

b) If $\lambda > 0$ and $\delta = 0$ (Variance-Gamma limit), we have

$$\phi_{VG_d(\lambda,\alpha,\beta,\mu,\Delta)}(u) = \\ = \exp\left[i\langle u, \mathrm{E}[VG_d]\rangle + \int_{\mathbb{R}^d} \big(e^{i\langle u,x\rangle} - 1 - i\langle u, x\rangle\big)\, g_{VG_d(\lambda,\alpha,\beta,\mu,\Delta)}(x)\,\mathrm{d}x \right]$$

where $\mathrm{E}[VG_d]$ is given in (2.13), and the Lévy density is

$$g_{VG_d(\lambda,\alpha,\beta,\mu,\Delta)}(x) = \frac{2\alpha^{\frac{d}{2}}\lambda e^{\langle x,\beta\rangle}}{\big(2\pi\sqrt{\langle x, \Delta^{-1}x\rangle}\big)^{\frac{d}{2}}}\, K_{\frac{d}{2}}\big(\alpha\sqrt{\langle x, \Delta^{-1}x\rangle}\big). \tag{2.20}$$

c) If $\lambda < 0$ and $\alpha = 0$, $\beta = \mathbf{0}$ (t limit), then

$$\phi_{t_d(\lambda,\delta,\mu,\Delta)}(u) = \\ = \exp\left[i\langle u, \mu\rangle + \int_{\mathbb{R}^d} \big(e^{i\langle u,x\rangle} - 1 - i\langle u, x\rangle \mathbb{1}_{\{\|x\|\leq 1\}}(x)\big)\, g_{t_d(\lambda,\delta,\mu,\Delta)}(x)\,\mathrm{d}x \right],$$

and the corresponding Lévy density is

$$g_{t_d(\lambda,\delta,\mu,\Delta)}(x) = \\ = \frac{2}{\big(2\pi\sqrt{\langle x, \Delta^{-1}x\rangle}\big)^{\frac{d}{2}}} \int_0^\infty \frac{(2y)^{\frac{d}{4}} K_{\frac{d}{2}}\big(\sqrt{(2y)\langle x, \Delta^{-1}x\rangle}\big)}{\pi^2 y \big[J^2_{|\lambda|}(\delta\sqrt{2y}) + Y^2_{|\lambda|}(\delta\sqrt{2y})\big]}\,\mathrm{d}y. \tag{2.21}$$

If $\lambda < -\frac{1}{2}$, then the truncation function within the integral term of the characteristic function can be omitted without further changes.

d) In the case $\lambda < 0$ and $\langle\beta, \Delta\beta\rangle = \alpha^2$ we have

$$\phi_{GH_d(\lambda,\sqrt{\langle\beta,\Delta\beta\rangle},\beta,\delta,\mu,\Delta)}(u) = \\ = \exp\left[i\langle u, b\rangle + \int_{\mathbb{R}^d} \left(e^{i\langle u,x\rangle} - 1 - i\langle u, x\rangle \mathbb{1}_{\{\|x\|\le 1\}}(x)\right) g(x)\,\mathrm{d}x\right],$$

where the drift vector is given by

$$b = \int_0^\infty \int_{\{x\in\mathbb{R}^d \,|\, \|x\|\le 1\}} x\, d_{N_d((s\Delta)\beta, s\Delta)}(x)\, g_{iG(\lambda,\frac{\delta^2}{2})}(s)\,\mathrm{d}x\,\mathrm{d}s,$$

and the Lévy density $g(x) = g_{GH_d(\lambda,\sqrt{\langle\beta,\Delta\beta\rangle},\beta,\delta,\mu,\Delta)}(x)$ is obtained from (2.19) by inserting $\alpha = \sqrt{\langle\beta,\Delta\beta\rangle}$. If $\lambda < -1$, one can omit the truncation function within the integral term of the characteristic function and replace the above expression for b by the mean $\mathrm{E}\big[GH_d\big(\lambda, \sqrt{\langle\beta,\Delta\beta\rangle}, \beta, \delta, \mu, \Delta\big)\big]$ from (2.15).

PROOF: a) Suppose $\mu = \mathbf{0}$ for the moment, then the defining equality (2.1), the infinite divisibility of GIG distributions and Lemma 2.6 imply that the Lévy process $(Y_t)_{t\ge 0}$ induced by $GH_d(\lambda,\alpha,\beta,\delta,\mathbf{0},\Delta)$ admits the stochastic representation $Y_t = \hat{B}_{\tau(t)}$ where $\hat{B}_t = \Delta\beta t + AB_t$ is a d-dimensional Brownian motion with drift $\Delta\beta$ and covariance matrix $\mathrm{Cov}(\hat{B}_1) = AA^\top = \Delta$, and $(\tau(t))_{t\ge 0}$ is an increasing Lévy process with $\mathcal{L}(\tau(1)) = GIG(\lambda, \delta, \sqrt{\alpha^2 - \langle\beta,\Delta\beta\rangle})$. Hence $(Y_t)_{t\ge 0}$ is a subordinated Lévy process of the form $Y_t = X_{Z_t}$ assumed in Theorem 2.13 with $X_t = \hat{B}_t$ and $Z_t = \tau(t)$, and the characteristic triplet $(\tilde{b}, \tilde{C}, \tilde{\nu})$ of $\mathcal{L}(Y_1) = GH_d(\lambda,\alpha,\beta,\delta,\mathbf{0},\Delta)$ can thus be obtained from equations (2.16)–(2.18). The parameter assumptions $\delta > 0$ and $\langle\beta,\Delta\beta\rangle < \alpha^2$ ensure a finite mean, so by the remark on p. 79f we can choose $\tilde{b} = \mathrm{E}[GH_d(\lambda,\alpha,\beta,\delta,\mathbf{0},\Delta)]$ and omit the truncation function within the integral term of the Lévy–Khintchine formula. If $\mu \ne \mathbf{0}$, then the proof of Lemma 2.5 as well as equation (2.3) show that this only leads to an additional summand $i\langle u, \mu\rangle$ in the exponent of the characteristic function which by (2.12) is directly included in $i\langle u, \tilde{b}\rangle$ if $\tilde{b} = \mathrm{E}[GH_d(\lambda,\alpha,\beta,\delta,\mu,\Delta)]$ and obviously does not affect the components $\tilde{C}$ and $\tilde{\nu}$ of the characteristic triplet.

Lemma 2.14 shows that the Laplace transforms of all GIG distributions with finite parameters have an exponent $\Psi(z) = \log\big(\mathfrak{L}_{GIG}(-z)\big)$ with drift parameter $b_0 = 0$, therefore by (2.17) the Gaussian part $\frac{1}{2}\langle u, \tilde{C}u\rangle$ of the Lévy–Khintchine representation vanishes for all $GH_d(\lambda,\alpha,\beta,\delta,\mu,\Delta)$-distributions and their weak limits considered in parts b)–d). Moreover, in the present case we have $\Phi_s = \mathcal{L}(X_s) = \mathcal{L}(\hat{B}_s) = N_d(s(\Delta\beta), s\Delta)$ and $\rho(\mathrm{d}s) = g_{GIG}(s)\,\mathrm{d}s$, so equation (2.18) and Fubini's theorem imply that the Lévy measures of multivariate GH distributions and their weak limits are given by

$$\tilde{\nu}(\mathrm{d}x) = \int_0^\infty \Phi_s(\mathrm{d}x)\,\rho(\mathrm{d}s) = \int_0^\infty d_{N_d(s(\Delta\beta),s\Delta)}(x)\, g_{GIG}(s)\,\mathrm{d}s\ \mathrm{d}x$$

and hence always possess a Lebesgue density. If $\delta > 0$ and $\langle\beta,\Delta\beta\rangle < \alpha^2$, this density is obtained with the help of Proposition 1.24 a) as follows:

$$g_{GH_d(\lambda,\alpha,\beta,\delta,\mu,\Delta)}(x) = \int_0^\infty d_{N_d(s(\Delta\beta),s\Delta)}(x)\, g_{GIG(\lambda,\delta,\sqrt{\alpha^2-\langle\beta,\Delta\beta\rangle})}(s)\,\mathrm{d}s$$

$$= \int_0^\infty \frac{e^{-\frac{1}{2}\langle x-(s\Delta)\beta,(s\Delta)^{-1}(x-(s\Delta)\beta)\rangle}}{(2\pi s)^{\frac{d}{2}}}\,\frac{e^{-\frac{s}{2}(\alpha^2-\langle\beta,\Delta\beta\rangle)}}{s}$$

$$\cdot\left[\int_0^\infty \frac{e^{-sy}}{\pi^2 y\big[J_{|\lambda|}^2(\delta\sqrt{2y})+Y_{|\lambda|}^2(\delta\sqrt{2y})\big]}\,\mathrm{d}y + \max(0,\lambda)\right]\mathrm{d}s$$

$$= e^{\langle x,\beta\rangle}\int_0^\infty \frac{e^{-\frac{1}{2}(\langle x,(s\Delta)^{-1}x\rangle+s\alpha^2)}}{(2\pi s)^{\frac{d}{2}}\, s} \cdot\left[\int_0^\infty \frac{e^{-sy}}{\pi^2 y\big[J_{|\lambda|}^2(\delta\sqrt{2y})+Y_{|\lambda|}^2(\delta\sqrt{2y})\big]}\,\mathrm{d}y + \max(0,\lambda)\right]\mathrm{d}s \tag{2.22}$$

$$= \frac{e^{\langle x,\beta\rangle}}{(2\pi)^{\frac{d}{2}}}\left[\int_0^\infty \frac{\int_0^\infty s^{-\frac{d}{2}-1}e^{-\frac{1}{2}\left(\frac{\langle x,\Delta^{-1}x\rangle}{s}+s(2y+\alpha^2)\right)}\,\mathrm{d}s}{\pi^2 y\big[J_{|\lambda|}^2(\delta\sqrt{2y})+Y_{|\lambda|}^2(\delta\sqrt{2y})\big]}\,\mathrm{d}y + \max(0,\lambda)\int_0^\infty s^{-\frac{d}{2}-1}e^{-\frac{1}{2}\left(\frac{\langle x,\Delta^{-1}x\rangle}{s}+s\alpha^2\right)}\,\mathrm{d}s\right] \tag{2.23}$$

Observing that $s^{-\frac{d}{2}-1}e^{-\frac{1}{2}\left(\frac{\langle x,\Delta^{-1}x\rangle}{s}+s(2y+\alpha^2)\right)}$ is the non-normalized density of a $GIG\big(-\frac{d}{2},\sqrt{\langle x,\Delta^{-1}x\rangle},\sqrt{2y+\alpha^2}\big)$-distribution, the value of the integral over this function from 0 to ∞ must be equal to the inverse of the corresponding norming constant. A comparison with (1.2) yields

$$\int_0^\infty s^{-\frac{d}{2}-1}e^{-\frac{1}{2}\left(\frac{\langle x,\Delta^{-1}x\rangle}{s}+s(2y+\alpha^2)\right)}\,\mathrm{d}s = 2\left(\frac{2y+\alpha^2}{\langle x,\Delta^{-1}x\rangle}\right)^{\frac{d}{4}} K_{\frac{d}{2}}\big(\sqrt{(2y+\alpha^2)\langle x,\Delta^{-1}x\rangle}\big).$$

Similarly, $s^{-\frac{d}{2}-1}e^{-\frac{1}{2}\left(\frac{\langle x,\Delta^{-1}x\rangle}{s}+s\alpha^2\right)}$ is the density of a $GIG\big(-\frac{d}{2},\sqrt{\langle x,\Delta^{-1}x\rangle},\alpha\big)$-distribution without norming constant and thus

$$\int_0^\infty s^{-\frac{d}{2}-1}e^{-\frac{1}{2}\left(\frac{\langle x,\Delta^{-1}x\rangle}{s}+s\alpha^2\right)}\,\mathrm{d}s = 2\alpha^{\frac{d}{2}}\big(\langle x,\Delta^{-1}x\rangle\big)^{-\frac{d}{4}} K_{\frac{d}{2}}\big(\alpha\sqrt{\langle x,\Delta^{-1}x\rangle}\big).$$

Inserting both expressions into (2.23) we obtain (2.19).

b) Because all VG distributions have finite means, it follows from the same reasoning as in part a) that the drift vector within the Lévy–Khintchine representation is given by $\mathrm{E}[VG_d(\lambda,\alpha,\beta,\mu,\Delta)]$, the Gaussian part vanishes, and the density of the Lévy measure can be calculated using Proposition 1.24 b) as follows:

$$g_{VG_d(\lambda,\alpha,\beta,\mu,\Delta)}(x) = \int_0^\infty d_{N_d(s(\Delta\beta),s\Delta)}(x)\, g_{G\left(\lambda,\frac{\alpha^2-\langle\beta,\Delta\beta\rangle}{2}\right)}(s)\,\mathrm{d}s$$

$$= \int_0^\infty \frac{e^{-\frac{1}{2}\langle x-(s\Delta)\beta,(s\Delta)^{-1}(x-(s\Delta)\beta)\rangle}}{(2\pi s)^{\frac{d}{2}}}\,\frac{\lambda e^{-\frac{1}{2}(\alpha^2-\langle\beta,\Delta\beta\rangle)}}{s}\,\mathrm{d}s$$

$$= \frac{\lambda e^{\langle x,\beta\rangle}}{(2\pi)^{\frac{d}{2}}} \int_0^\infty s^{-\frac{d}{2}-1} e^{-\frac{1}{2}(\langle x,(s\Delta)^{-1}x\rangle + s\alpha^2)}\,\mathrm{d}s$$
$$= \frac{2\alpha^{\frac{d}{2}}\lambda e^{\langle x,\beta\rangle}}{\big(2\pi\sqrt{\langle x,\Delta^{-1}x\rangle}\big)^{\frac{d}{2}}} K_{\frac{d}{2}}\big(\alpha\sqrt{\langle x,\Delta^{-1}x\rangle}\big)$$

where the last equation has already been shown in the proof of part a).

c) Since the t limit-distributions have a finite mean only if $\lambda < -\frac{1}{2}$, we have to calculate the drift vector $\tilde{b}$ according to equation (2.16) of Theorem 2.13. To apply it correctly, we again assume $\mu = \mathbf{0}$ first. From the normal variance mixture representation of $t_d(\lambda,\delta,\mathbf{0},\Delta)$ (see p. 75) we then conclude that $\Phi_s = N_d(\mathbf{0}, s\Delta)$ and $\rho(\mathrm{d}s) = g_{iG(\lambda,\frac{\delta^2}{2})}(s)\,\mathrm{d}s$ in this case. By Lemma 2.14, $b_0 = 0$ holds. Therefore we obtain

$$\tilde{b} = \int_0^\infty \int_{\{x\in\mathbb{R}^d \,|\, \|x\|\le 1\}} x\, d_{N_d(\mathbf{0},s\Delta)}(x)\,\mathrm{d}x\ \, g_{iG(\lambda,\frac{\delta^2}{2})}(s)\,\mathrm{d}s = \mathbf{0},$$

because the value of the inner integral is $\mathbf{0}$ by symmetry. If $\mu \neq \mathbf{0}$, then the drift vector becomes $\tilde{b} = \mathbf{0} + \mu = \mu$ as pointed out in the proof of part a). The drift vector remains unchanged if the truncation function within the integral term of the Lévy–Khintchine representation is removed for sufficiently small λ, because $\mathrm{E}[t_d(\lambda,\delta,\mu,\Delta)] = \mu$ by (2.14).

Moreover, Proposition 1.24 shows that the Lévy density $g_{iG(\lambda,\frac{\delta^2}{2})}(s)$ can be obtained from $g_{GIG(\lambda,\delta,\gamma)}(s)$ by setting $\gamma = 0$ within the latter, hence the density $g_{t_d(\lambda,\delta,\mu,\Delta)}(x)$ of the Lévy measure of $t_d(\lambda,\delta,\mu,\Delta)$ is immediately obtained from the corresponding calculation in the proof of part a) by inserting $\alpha = 0$, $\beta = \mathbf{0}$ and observing that $\lambda < 0$ here. Alternatively, these insertions can be made directly within the formula for $g_{GH_d(\lambda,\alpha,\beta,\delta,\mu,\Delta)}(x)$, which yields (2.21).

d) The assertions follow along almost the same lines as before, but the drift coefficient $\tilde{b}$ can hardly be calculated more explicitly because in this limit case we always have $\beta \neq \mathbf{0}$. Consequently the density of $\Phi_s = N_d(s(\Delta\beta), s\Delta)$ is never symmetric around the origin. $\square$

Note that the Lévy densities (2.19)and (2.21) can be simplified considerably if $\lambda = \pm\frac{1}{2}$. Using equations (A.12) and assuming $\delta > 0$, the inner integral in (2.22) becomes

$$\int_0^\infty \frac{e^{-sy}}{\pi^2 y\big[J^2_{\frac{1}{2}}(\delta\sqrt{2y}) + Y^2_{\frac{1}{2}}(\delta\sqrt{2y})\big]}\,\mathrm{d}y = \frac{\delta}{\pi}\int_0^\infty \frac{e^{-sy}}{\sqrt{2y}}\,\mathrm{d}y$$
$$\underset{t:=\sqrt{y}}{=} \frac{2\delta}{\sqrt{\pi}}\int_0^\infty \frac{1}{\sqrt{2\pi}}\,e^{-st^2}\,\mathrm{d}t = \frac{\delta}{\sqrt{2\pi s}}\int_{-\infty}^\infty \frac{1}{\sqrt{2\pi(2s)^{-1}}}\,e^{-st^2}\,\mathrm{d}t = \frac{\delta}{\sqrt{2\pi s}},$$

because the last integrand equals the density of a $N\big(0,\frac{1}{2s}\big)$-distribution. Inserting this expression into (2.22) and continuing the calculation, we obtain

$$
\begin{aligned}
g_{GH_d(|\lambda|=\frac{1}{2},\alpha,\beta,\delta,\mu,\Delta)}(x) &= e^{\langle x,\beta\rangle}\int_0^\infty \frac{e^{-\frac{1}{2}(\langle x,(s\Delta)^{-1}x\rangle+s\alpha^2)}}{(2\pi s)^{\frac{d}{2}}\,s}\left[\frac{\delta}{\sqrt{2\pi s}}+\max(0,\lambda)\right]\mathrm{d}s \\
&= \frac{e^{\langle x,\beta\rangle}}{(2\pi)^{\frac{d}{2}}}\left[\int_0^\infty \frac{\delta}{\sqrt{2\pi}}\,s^{-\frac{d+1}{2}-1}e^{-\frac{1}{2}\left(\frac{\langle x,\Delta^{-1}x\rangle}{s}+s\alpha^2\right)}\,\mathrm{d}s\right. \\
&\qquad\left.+\max(0,\lambda)\int_0^\infty s^{-\frac{d}{2}-1}e^{-\frac{1}{2}\left(\frac{\langle x,\Delta^{-1}x\rangle}{s}+s\alpha^2\right)}\,\mathrm{d}s\right] \\
&= \frac{2\alpha^{\frac{d}{2}}\,e^{\langle x,\beta\rangle}}{\left(2\pi\sqrt{\langle x,\Delta^{-1}x\rangle}\right)^{\frac{d}{2}}}\left[\frac{\delta\sqrt{\alpha}}{\sqrt{2\pi}\langle x,\Delta^{-1}x\rangle^{\frac{1}{4}}}\,K_{\frac{d+1}{2}}\left(\alpha\sqrt{\langle x,\Delta^{-1}x\rangle}\right)\right. \\
&\qquad\left.+\max(0,\lambda)K_{\frac{d}{2}}\left(\alpha\sqrt{\langle x,\Delta^{-1}x\rangle}\right)\right].
\end{aligned}
\tag{2.24}
$$

The last line follows from the same arguments used in the proof of Proposition 2.15 a).

Remark: If $d=\Delta=1$, then $\sqrt{\langle x,\Delta^{-1}x\rangle}$ becomes $|x|$, and together with the representation of the Bessel function $K_{\frac{1}{2}}(x)$ (see equation (A.7)), the Lévy density (2.19) in Proposition 2.15 a) simplifies to

$$
\begin{aligned}
&g_{GH(\lambda,\alpha,\beta,\delta,\mu)}(x) = \\
&\quad= \frac{e^{\beta x}}{|x|}\left[\int_0^\infty \frac{e^{-|x|\sqrt{2y+\alpha^2}}}{\pi^2 y\left[J^2_{|\lambda|}\left(\delta\sqrt{2y}\right)+Y^2_{|\lambda|}\left(\delta\sqrt{2y}\right)\right]}\,\mathrm{d}y+\max(0,\lambda)\,e^{-\alpha|x|}\right],
\end{aligned}
$$

which exactly coincides with the representation derived in Chapter 1, Proposition 1.29 a). This also holds true for the limit distributions considered in parts b)–d) of both propositions, showing that the formulas (2.19)–(2.21) are in fact the natural generalizations of the univariate Lévy densities. Inserting $d=\Delta=1$ and $\lambda=-\frac{1}{2}$ in (2.24) yields in the same way

$$
g_{NIG(\alpha,\beta,\delta,\mu)}(x) = \frac{\delta\alpha}{\pi|x|}\,e^{\beta x}K_1\left(\alpha|x|\right).
$$

This agrees with our calculations on page 51.

All symmetric $GH_d(\lambda,\alpha,\mathbf{0},\delta,\mu,\Delta)$-distributions and their weak limits considered above (in particular all $t_d(\lambda,\delta,\mu,\Delta)$-distributions) are selfdecomposable. This is an immediate consequence of the fact that all GH distributions (including the weak limits) with $\beta=\mathbf{0}$ are normal variance mixtures with selfdecomposable GIG mixing distributions (see pp. 72–75) and thus are selfdecomposable themselves by Lemma 2.5 d). However, this property in general does not transfer to the skewed distributions as the subsequent proposition shows.

Proposition 2.16 *A multivariate $VG_d(\lambda,\alpha,\beta,\mu,\Delta)$-distribution is selfdecomposable if and only if $\beta=\mathbf{0}$.*

Proof: Because of the preceding remarks it only remains to show the "only if"-part for which we use some ideas from Takano (1989). We first derive the polar decomposition of the corresponding Lévy densities, that is, we move from the Cartesian coordinates x to polar coordinates $r := \|x\|$ and $\xi := \frac{x}{\|x\|}$. Inserting the new variables into the density formula (2.20) and observing that, according to the transformation rule of integration $\mathrm{d}x = r^{d-1}\,\mathrm{d}r\,\mathrm{d}\xi$, we obtain

$$\nu_{VG_d(\lambda,\alpha,\beta,\mu,\Delta)}(\mathrm{d}r\,\mathrm{d}\xi) = \frac{2\lambda\, r^{\frac{d}{2}-1}\alpha^{\frac{d}{2}} e^{r\langle\xi,\beta\rangle}}{\big(2\pi\sqrt{\langle\xi,\Delta^{-1}\xi\rangle}\big)^{\frac{d}{2}}}\, K_{\frac{d}{2}}\big(\alpha r\sqrt{\langle\xi,\Delta^{-1}\xi\rangle}\big)\,\mathrm{d}r\,\mathrm{d}\xi$$

Setting $\bar{\lambda}(\mathrm{d}\xi) := 2\lambda\big(2\pi\langle\xi,\Delta^{-1}\xi\rangle\big)^{-\frac{d}{2}}\,\mathrm{d}\xi$ and

$$k_\xi(r) := e^{r\langle\xi,\beta\rangle}\big(\alpha r\sqrt{\langle\xi,\Delta^{-1}\xi\rangle}\big)^{\frac{d}{2}} K_{\frac{d}{2}}\big(\alpha r\sqrt{\langle\xi,\Delta^{-1}\xi\rangle}\big),$$

we get the representation $\nu_{VG_d(\lambda,\alpha,\beta,\mu,\Delta)}(\mathrm{d}r\,\mathrm{d}\xi) = \frac{k_\xi(r)}{r}\,\mathrm{d}r\,\bar{\lambda}(\mathrm{d}\xi)$. By Lemma 2.3, the distribution $VG_d(\lambda,\alpha,\beta,\mu,\Delta)$ is selfdecomposable if and only if the function $k_\xi(r)$ is decreasing in r for every (or, to be more precisely, $\bar{\lambda}$-almost all) ξ. But below we will show that for every $\beta \neq \mathbf{0}$ there exist some ξ_β and a neighbourhood $B(\xi_\beta) \subset \mathcal{S} = \{\xi \in \mathbb{R}^d \mid \|\xi\| = 1\}$ of ξ_β with $\bar{\lambda}(B(\xi_\beta)) > 0$ such that $k_{\bar{\xi}}(r)$ has a strictly positive derivative near the origin for every $\bar{\xi} \in B(\xi_\beta)$. Hence a VG distribution with $\beta \neq \mathbf{0}$ cannot be selfdecomposable.

Let us write $k_\xi(r) = e^{r\langle\xi,\beta\rangle} L_{\frac{d}{2}}\big(\alpha r\sqrt{\langle\xi,\Delta^{-1}\xi\rangle}\big)$ with $L_{\frac{d}{2}}(y) := y^{\frac{d}{2}} K_{\frac{d}{2}}(y)$. Equations (A.8) and (A.9) from Appendix A imply that for $d \geq 2$ we have

$$\lim_{y\to 0} L_{\frac{d}{2}}(y) = 2^{\frac{d-2}{2}}\,\Gamma\big(\tfrac{d}{2}\big) \quad\text{and}\quad \lim_{y\to 0} y L_{\frac{d-2}{2}}(y) = 0,$$

and from equations (A.4) and (A.3) we conclude

$$\begin{aligned} L'_{\frac{d}{2}}(y) &= \frac{d}{2}\, y^{\frac{d-2}{2}}\, K_{\frac{d}{2}}(y) + y^{\frac{d}{2}} K'_{\frac{d}{2}}(y) \\ &= \frac{d}{2}\, y^{\frac{d-2}{2}}\, K_{\frac{d}{2}}(y) - y^{\frac{d-2}{2}}\left(\frac{d}{2}\, K_{\frac{d}{2}}(y) + y K_{\frac{d-2}{2}}(y)\right) = -y L_{\frac{d-2}{2}}(y). \end{aligned} \tag{2.25}$$

Consequently the derivative of $k_\xi(r)$ with respect to r is

$$k'_\xi(r) = e^{r\langle\xi,\beta\rangle}\Big[\langle\xi,\beta\rangle\, L_{\frac{d}{2}}\big(\alpha r\sqrt{\langle\xi,\Delta^{-1}\xi\rangle}\big) - r\alpha^2\langle\xi,\Delta^{-1}\xi\rangle\, L_{\frac{d-2}{2}}\big(\alpha r\sqrt{\langle\xi,\Delta^{-1}\xi\rangle}\big)\Big]$$

and thus $\lim_{r\to 0} k'_\xi(r) = \langle\xi,\beta\rangle\, 2^{\frac{d-2}{2}}\,\Gamma\big(\frac{d}{2}\big)$. Now if $\beta \neq \mathbf{0}$, we obviously can find some ξ_β and a neighbourhood $B(\xi_\beta) \subset \mathcal{S}$ of ξ_β with $\bar{\lambda}(B(\xi_\beta)) > 0$ such that $\langle\bar{\xi},\beta\rangle > 0$ for all $\bar{\xi} \in B(\xi_\beta)$ (note that by the above definition $\bar{\lambda}(\mathrm{d}\xi)$ is equivalent to the Lebesgue measure restricted on the sphere $\mathcal{S}$), and from the continuity of $k'_\xi(r)$ it follows that $k'_{\bar{\xi}}(r)$ is strictly positive on some open intervall $(0, r_0)$ for all $\bar{\xi} \in B(\xi_\beta)$. This proves the assertion. □

Unfortunately, the previous proof cannot be transferred to other GH distributions than the VG subclass as will be explained below. By Proposition 2.15

a) and d), the corresponding Lévy measures are, using the polar coordinates r and ξ defined above and the relation $\mathrm{d}x = r^{d-1}\,\mathrm{d}r\,\mathrm{d}\xi$, given by

$$\begin{aligned}
&\nu_{GH_d(\lambda,\alpha,\beta,\delta,\mu,\Delta)}(\mathrm{d}r\,\mathrm{d}\xi) = \\
&\quad = \frac{2e^{r\langle\xi,\beta\rangle}}{r\big(2\pi\sqrt{\langle\xi,\Delta^{-1}\xi\rangle}\big)^{\frac{d}{2}}}\left[\int_0^\infty \frac{r^{\frac{d}{2}}(2y+\alpha^2)^{\frac{d}{4}}K_{\frac{d}{2}}\big(r\sqrt{(2y+\alpha^2)\langle\xi,\Delta^{-1}\xi\rangle}\big)}{\pi^2 y\big[J^2_{|\lambda|}(\delta\sqrt{2y})+Y^2_{|\lambda|}(\delta\sqrt{2y})\big]}\,\mathrm{d}y\right. \\
&\qquad\qquad \left. + \max(0,\lambda)\,(\alpha r)^{\frac{d}{2}}K_{\frac{d}{2}}\big(\alpha r\sqrt{\langle\xi,\Delta^{-1}\xi\rangle}\big)\right]\mathrm{d}r\,\mathrm{d}\xi\,.
\end{aligned}$$

Setting $\bar\lambda(\mathrm{d}\xi) := 2\big(2\pi\langle\xi,\Delta^{-1}\xi\rangle\big)^{-\frac{d}{2}}\,\mathrm{d}\xi$ and

$$\begin{aligned}
k_\xi(r) &:= e^{r\langle\xi,\beta\rangle}\left[\int_0^\infty \frac{\big(r\sqrt{(2y+\alpha^2)\langle\xi,\Delta^{-1}\xi\rangle}\big)^{\frac{d}{2}}K_{\frac{d}{2}}\big(r\sqrt{(2y+\alpha^2)\langle\xi,\Delta^{-1}\xi\rangle}\big)}{\pi^2 y\big[J^2_{|\lambda|}(\delta\sqrt{2y})+Y^2_{|\lambda|}(\delta\sqrt{2y})\big]}\,\mathrm{d}y\right. \\
&\qquad\left. + \max(0,\lambda)\,\big(\alpha r\sqrt{\langle\xi,\Delta^{-1}\xi\rangle}\big)^{\frac{d}{2}}K_{\frac{d}{2}}\big(\alpha r\sqrt{\langle\xi,\Delta^{-1}\xi\rangle}\big)\right] \\
&= e^{r\langle\xi,\beta\rangle}\left[\int_0^\infty \frac{L_{\frac{d}{2}}\big(r\sqrt{(2y+\alpha^2)\langle\xi,\Delta^{-1}\xi\rangle}\big)}{\pi^2 y\big[J^2_{|\lambda|}(\delta\sqrt{2y})+Y^2_{|\lambda|}(\delta\sqrt{2y})\big]}\,\mathrm{d}y\right. \\
&\qquad\left. + \max(0,\lambda)L_{\frac{d}{2}}\big(\alpha r\sqrt{\langle\xi,\Delta^{-1}\xi\rangle}\big)\right]
\end{aligned} \tag{2.26}$$

with $L_{\frac{d}{2}}(y) := y^{\frac{d}{2}}K_{\frac{d}{2}}(y)$ as before yields $\nu_{GH_d(\lambda,\alpha,\beta,\delta,\mu,\Delta)}(\mathrm{d}r\,\mathrm{d}\xi) = \frac{k_\xi(r)}{r}\,\mathrm{d}r\,\bar\lambda(\mathrm{d}\xi)$. Differentiating $k_\xi(r)$ with respect to r we get, using (2.25),

$$\begin{aligned}
k'_\xi(r) &= e^{r\langle\xi,\beta\rangle}\left[\left(\langle\xi,\beta\rangle\int_0^\infty \frac{L_{\frac{d}{2}}\big(r\sqrt{(2y+\alpha^2)\langle\xi,\Delta^{-1}\xi\rangle}\big)}{\pi^2 y\big[J^2_{|\lambda|}(\delta\sqrt{2y})+Y^2_{|\lambda|}(\delta\sqrt{2y})\big]}\,\mathrm{d}y\right.\right. \\
&\qquad\left. -r\int_0^\infty \frac{(2y+\alpha^2)\langle\xi,\Delta^{-1}\xi\rangle\,L_{\frac{d-2}{2}}\big(r\sqrt{(2y+\alpha^2)\langle\xi,\Delta^{-1}\xi\rangle}\big)}{\pi^2 y\big[J^2_{|\lambda|}(\delta\sqrt{2y})+Y^2_{|\lambda|}(\delta\sqrt{2y})\big]}\,\mathrm{d}y\right) \\
&\qquad + \max(0,\lambda)\Big(\langle\xi,\beta\rangle\,L_{\frac{d}{2}}\big(\alpha r\sqrt{\langle\xi,\Delta^{-1}\xi\rangle}\big) \\
&\qquad\qquad\left. -r\alpha^2\langle\xi,\Delta^{-1}\xi\rangle\,L_{\frac{d-2}{2}}\big(\alpha r\sqrt{\langle\xi,\Delta^{-1}\xi\rangle}\big)\Big)\right],
\end{aligned} \tag{2.27}$$

where the interchange between differentiation and integration in the second integral term can be justified as follows: Since every Lévy measure $\nu(\mathrm{d}x)$ on $\mathbb{R}^d$ fulfills $\nu\big(\{x\in\mathbb{R}^d \,|\, \|x\|\geq\epsilon\}\big) < \infty$ for all $\epsilon > 0$ and arbitrary dimension d, the function $k_\xi(r)$ and hence especially the integral in (2.26) must be finite for all $r > 0$, $\xi\in\mathcal{S}$, and $d\in\mathbb{N}$. By equation (A.5) we have the inequality $L_{\frac{d+2}{2}}(y) = y^{\frac{d+2}{2}}K_{\frac{d+2}{2}}(y) > y^{\frac{d+2}{2}}K_{\frac{d-2}{2}}(y) = y^2 L_{\frac{d-2}{2}}(y)$ for $y > 0$, which provides an integrable majorant for the derivative. But here the problem arises that the behaviour of $k'_\xi(r)$ near the origin can hardly be determined. Since the denominator within both integrals asymptotically behaves like $\delta(2\pi^2 y)^{-\frac{1}{2}}$ for $y \to \infty$ (see pp. 36 and 40) and $L_{\frac{d}{2}}(y)$ tends to a constant for $y \to 0$, the first integral in (2.27) diverges to infinity if $r \to 0$. The second integral does so as well, but it is not clear if this may be compensated by the preceding factor r such that the product remains bounded if r tends to zero. But even if this would not be the case, the overall difference of the two integral terms might converge to a finite limit. All in all, it seems not feasible to deduce from (2.27) whether the right hand side is ≤ 0 for all $r > 0$ or if there exists some open intervall $(r_1, r_2) \subset \mathbb{R}_+$ on which $k'_\xi(r)$ is strictly positive. A complete characterization of all selfdecomposable GH distributions will probably require some different approaches and techniques and is therefore left for future research.

2.4 On the dependence structure of multivariate GH distributions

Correlation is probably the most established dependence measure due to its simplicity and its predominant role within the normal world where it characterizes dependencies almost completely. This follows from the fact that the components W_i, $1 \leq i \leq d$, of a standard normal distributed random vector $W \sim N_d(\mathbf{0}, I_d)$ are independent from each other (the joint density is just the product of the marginal ones in this case) and the stochastic representation

$$X \sim N_d(\mu, \Delta) \quad \Longleftrightarrow \quad X \overset{d}{=} \mu + AW \;\; \text{where } W \sim N_d(\mathbf{0}, I_d) \text{ and } AA^\top = \Delta.$$

Since X in distribution is nothing but a linear transform of a random vector W with independent (normal distributed) entries, the components of X can, roughly speaking, exhibit at most linear dependencies, and exactly these are specified and quantified by the pairwise correlations. However, things completely change if we depart from normality and consider normal variance mixtures instead. Suppose

$$X \sim N_d(\mu, y\Delta) \circ G, \;\; \text{that is,} \;\; X \overset{d}{=} \mu + \sqrt{Z} AW$$

where $\mathcal{L}(Z) = G$, $W \sim N_d(\mathbf{0}, I_d)$ and $AA^\top = \Delta$ according to Definition 2.4. As we already remarked on p. 68, the mixing variable Z causes dependencies between the components of X, but these are typically not captured by correlation as the following lemma shows. It is a slightly more general version of McNeil, Frey, and Embrechts (2005, Lemma 3.5) which we adopt here since—in our opinion—the result is as simple as illustrative.

Lemma 2.17 *Suppose that $X \overset{d}{=} \mu + \sqrt{Z}AW$ has a normal variance mixture distribution where $\mathrm{E}(Z) < \infty$ and $\Delta = AA^\top$ is a $d \times d$-diagonal matrix such that $\mathrm{Cov}(X_i, X_j) = 0$, $1 \leq i, j \leq d$, $i \neq j$, by (2.11). Then the X_i, $1 \leq i \leq d$, are independent if and only if Z is almost surely constant, that is, if and only if X is multivariate normal distributed.*

Proof: Because Δ is diagonal (and positive definite by Definition 2.4), we can assume without loss of generality that also the matrix A is diagonal and

$A_{ii} = \sqrt{\Delta_{ii}}$, $1 \le i \le d$. The independence of Z and W and Jensen's inequality then imply

$$\begin{aligned}\mathrm{E}\left(\prod_{i=1}^{d}|X_i-\mu_i|\right) &= \mathrm{E}\left((\sqrt{Z})^d\prod_{i=1}^{d}|\sqrt{\Delta_{ii}}W_i|\right) = \mathrm{E}\big((\sqrt{Z})^d\big)\prod_{i=1}^{d}\mathrm{E}\big(|\sqrt{\Delta_{ii}}W_i|\big)\\ &\ge \mathrm{E}\big(\sqrt{Z}\big)^d\prod_{i=1}^{d}\mathrm{E}\big(|\sqrt{\Delta_{ii}}W_i|\big) = \prod_{i=1}^{d}\mathrm{E}\big(|X_i-\mu_i|\big).\end{aligned}$$

Since the function $f(x) = x^d$ is strictly convex on $\mathbb{R}_+$ for $d \ge 2$, equality throughout holds if and only if Z is constant almost surely. □

Remark: The above result can even be extended: If $X \stackrel{d}{=} \mu + Z\beta + \sqrt{Z}AW$ has a normal mean-variance mixture distribution with $0 < \mathrm{Var}(Z) < \infty$, $\Delta = AA^\top$ is a $d \times d$-diagonal matrix and $\mathrm{Cov}(X_i, X_j) = 0$ for some $1 \le i \ne j \le d$, then X_i and X_j are not independent either. This can be seen as follows: Since Δ is diagonal and $\mathrm{Var}(Z) > 0$, by (2.11) $\mathrm{Cov}(X_i, X_j) = 0$ implies that $(\beta\beta^\top)_{ij} = 0$. This means, either $\beta_i = 0$ or $\beta_j = 0$ (or both, but then we would be within the setting of Lemma 2.17 again). Suppose $\beta_i \ne 0$ and $\beta_j = 0$, then we calculate using similar arguments as above

$$\begin{aligned}\mathrm{E}\big((X_i-\mu_i)\,|X_j-\mu_j|\big) &= \mathrm{E}\big((\beta_i Z + \sqrt{Z}\sqrt{\Delta_{ii}}W_i)\,|\sqrt{Z}\sqrt{\Delta_{jj}}W_j|\big)\\ &= \mathrm{E}\big((\beta_i Z^{\frac{3}{2}} + Z\sqrt{\Delta_{ii}}W_i)\big)\mathrm{E}\big(|\sqrt{\Delta_{jj}}W_j|\big) = \beta_i\,\mathrm{E}\big(Z^{\frac{3}{2}}\big)\,\mathrm{E}\big(|\sqrt{\Delta_{jj}}W_j|\big)\\ &> \beta_i\,\mathrm{E}(Z)^{\frac{3}{2}}\,\mathrm{E}\big(|\sqrt{\Delta_{jj}}W_j|\big) = \mathrm{E}(\beta_i Z)\,\mathrm{E}(Z)^{\frac{1}{2}}\,\mathrm{E}\big(|\sqrt{\Delta_{jj}}W_j|\big)\\ &> \mathrm{E}(\beta_i Z)\,\mathrm{E}\big(Z^{\frac{1}{2}}\big)\mathrm{E}\big(|\sqrt{\Delta_{jj}}W_j|\big) = \mathrm{E}(X_i-\mu_i)\,\mathrm{E}\big(|X_j-\mu_j|\big),\end{aligned}$$

and the inequalities are strict because $f(x) = x^{\frac{3}{2}}$ and $g(x) = \sqrt{x}$ are strictly convex resp. concave and $\mathcal{L}(Z)$ is non-degenerate by assumption.

Thus in general zero correlation within multivariate normal mean-variance mixture models must not be interpreted as independence. In particular the components X_i of a GH distributed random vector $X \sim GH_d(\lambda, \alpha, \beta, \delta, \mu, \Delta)$ can never be independent because Theorem 2.11 b) states that the conditional distribution $\mathcal{L}\big(X_i\,|\,(X_1, \dots, X_{i-1}, X_{i+1}, \dots, X_d)^\top = \bar{x}\big) = GH(\tilde{\lambda}, \tilde{\alpha}, \tilde{\beta}, \tilde{\delta}, \tilde{\mu})$ always depends on the vector $\bar{x}$ (at least the parameter $\tilde{\delta}$ does so) for every $1 \le i \le d$. Moreover, it should be observed that for generalized hyperbolic distributed random variables the maximal attainable absolute correlation is usually strictly smaller than one: the Cauchy–Schwarz inequality states that $|\mathrm{Corr}(X_1, X_2)| = 1$ can occur if and only if $X_2 = aX_1 + b$ almost surely for some $a, b \in \mathbb{R}$ and $a \ne 0$, but if $X_1 \sim GH(\lambda_1, \alpha_1, \beta_1, \delta_1, \mu_1)$ and $X_2 \sim GH(\lambda_2, \alpha_2, \beta_2, \delta_2, \mu_2)$, the required linear relationship imposes some conditions on the GH parameters. Recall that $aX_1 + b \sim GH(\lambda, \frac{\alpha}{|a|}, \frac{\beta}{a}, \delta|a|, a\mu + b)$ by Theorem 2.11 c). Thus using the scale- and location-invariant parameters $\zeta_i = \delta_i(\alpha_i^2 - \beta_i^2)^{\frac{1}{2}}$ and $\rho_i = \frac{\beta_i}{\alpha_i}$, $i = 1, 2$, we conclude that $X_2 = aX_1 + b$ can hold only if $\zeta_1 = \zeta_2$, $|\rho_1| = |\rho_2|$ and $\lambda_1 = \lambda_2$.

Having seen that correlation is in general not the tool to precisely describe and measure dependencies in multivariate models, one may ask if there exists a more powerful notion for this purpose. The answer is positive and provided by

Definition 2.18 *A d-dimensional copula C is a distribution function on $[0,1]^d$ with standard uniform marginal distributions, that is, $C : [0,1]^d \to [0,1]$ has the following properties:*

a) $C(u) = C(u_1, \ldots, u_d)$ *is increasing in each argument u_i,*

b) $C(1, \ldots, 1, u_i, 1, \ldots, 1) = u_i$ *for all $1 \le i \le d$ and $u_i \in [0,1]$,*

c) For all $(a_1, \ldots, a_d)^\top$, $(b_1, \ldots, b_d)^\top \in [0,1]^d$ with $a_i \le b_i$, $1 \le i \le d$, we have

$$\sum_{i_1=1}^{2} \cdots \sum_{i_d=1}^{2} (-1)^{i_1+\cdots+i_d} C(u_{1i_1}, \ldots, u_{di_d}) \ge 0,$$

where $u_{j1} = a_j$ and $u_{j2} = b_j$ for all $1 \le j \le d$.

Remark: Properties a) and b) immediately follow from the definition of $C(u)$ as a distribution function with identically uniformly distributed marginals on $[0,1]$, c) essentially is a reformulation of the fact that if $U = (U_1, \ldots, U_d)^\top$ is a random vector possessing the distribution function $C(u)$, then necessarily $P(a_1 \le U_1 \le b_1, \ldots, a_d \le U_d \le b_d) \ge 0$. It can also be shown that these properties are sufficient, that is, every function $C : [0,1]^d \to [0,1]$ fulfilling a), b) and c) is a copula. Clearly, the k-dimensional margins of a copula C are also copulas for every $2 \le k < d$.

We do not intend to give a detailed survey on the fairly large theory of copulas in the following, but simply want to mention some basic facts and results related to the topic of tail-dependence we shall be concerned with later on. For further reading, we refer to the literature hereafter (which of course is a subjective and incomplete choice): A concise and readable introduction to copulas can be found in McNeil, Frey, and Embrechts (2005, Chapter 5), a comprehensive overview is provided by the monograph of Nelsen (1999), and the application of copulas to finance is discussed in Cherubini, Luciano, and Vecchiato (2004).

The central role of copulas in the study of multivariate distributions is highlighted by the following fundamental result which goes back to Sklar (1959). It not only shows that copulas are inherent in every multivariate distribution, but also that the latter can be constructed by plugging the desired marginal distributions into a suitably chosen copula.

Theorem 2.19 (Sklar's Theorem) *Let F be a d-dimensional distribution function with margins $F_1, \ldots, F_d$. Then there exists a copula $C : [0,1]^d \to [0,1]$ such that for all $x = (x_1, \ldots, x_d)^\top \in [-\infty, \infty]^d$*

$$F(x_1, \ldots, x_d) = C\big(F_1(x_1), \ldots, F_d(x_d)\big). \tag{2.28}$$

If $F_1, \ldots, F_d$ are all continuous, then C is unique, otherwise C is uniquely determined on $F_1(\mathbb{R}) \times \cdots \times F_d(\mathbb{R})$ where $F_i(\mathbb{R})$ denotes the range of F_i.

Conversely, if $C : [0,1]^d \to [0,1]$ is a copula and $F_1, \dots, F_d$ are univariate distribution functions, then the function $F(x)$ defined by (2.28) *is a multivariate distribution function with margins $F_1, \dots, F_d$.*

Remark: In most books, Sklar's Theorem is only proven in the special case of continuous margins F_i (see, for example, McNeil, Frey, and Embrechts 2005, p. 187), the proof of the general case is usually referred to the original articles or, if given explicitly, written down in a rather lengthy and technical way (see Nelsen 1999, p. 16ff). A short and elegant proof of the general case which is based on the distributional transform can be found in Rüschendorf (2009).

If all marginal distribution functions F_i of F are continuous and their *generalized inverses* F_i^{-1} are defined by $F_i^{-1}(u_i) := \inf\{y \,|\, F_i(y) \geq u_i\}$ (with the usual convention $\inf \emptyset = \infty$), then $F_i\big(F_i^{-1}(u_i)\big) = u_i$. Thus it immediately follows from (2.28) by inserting $x_i = F_i^{-1}(u_i)$, $u_i \in [0,1]$, $1 \leq i \leq d$, that in this case the unique copula $C_F(u)$ contained in F is given by

$$C_F(u_1, \dots, u_d) = F\big(F_1^{-1}(u_1), \dots, F_d^{-1}(u_d)\big). \tag{2.29}$$

The computation of this so-called *implied copula* $C_F(u)$ is in general numerically demanding if the distribution function $F(x)$ is not known explicitly. Suppose for example that only the density $f(x)$ of F can be expressed in closed form, then already the determination of a single value $F(x_0)$ requires to evaluate a d-dimensional integral which especially for greater dimensions d can hardly be done sufficiently precise in reasonable time. But for multivariate normal mean-variance mixtures it is sometimes possible to significantly reduce the numerical complexity: Suppose that $F = N_d(\mu + y\beta, y\Delta) \circ G$ with known margins F_i possessing Lebesgue densities f_i as above, and let O be an orthogonal $d \times d$-matrix such that $O\Delta O^\top$ is diagonal, then

$$\begin{aligned} C_F(u_1, \dots, u_d) &= F\big(F_1^{-1}(u_1), \dots, F_d^{-1}(u_d)\big) \\ &= \int_{-\infty}^{F_d^{-1}(u_d)} \cdots \int_{-\infty}^{F_1^{-1}(u_1)} \int_0^\infty d_{N_d(\mu+y\beta,y\Delta)}(x_1, \dots, x_d)\, G(\mathrm{d}y)\ \mathrm{d}x_1 \dots \mathrm{d}x_d \\ &= \int_0^\infty \prod_{i=1}^d \Phi_{((O(\mu+y\beta))_i, (O\Delta O^\top)_{ii})} \Big(\big(O(F_1^{-1}(u_1), \dots, F_d^{-1}(u_d))^\top\big)_i \Big)\, G(\mathrm{d}y), \end{aligned}$$

where $\Phi_{(\mu,\sigma^2)}$ denotes the (univariate) distribution function of $N(\mu, \sigma^2)$. The last expression can be evaluated much easier on a computer since it only requires the calculation of one-dimensional integrals (possibly more than one because the values $F_i^{-1}(u_i)$ of the marginal quantile functions may only be obtained by integrating the corresponding densities $f_i(x_i)$ numerically).

If in addition to the marginal distributions F_i also F itself possesses a Lebesgue density $f(x)$, a further simplification can be achieved by using the *(implied) copula density* $c_F(u)$ which is defined by

$$c_F(u_1, \dots, u_d) := \frac{\partial C_F(u_1, \dots, u_d)}{\partial u_1 \dots \partial u_d} = \frac{f\big(F_1^{-1}(u_1), \dots, F_d^{-1}(u_d)\big)}{f_1\big(F_1^{-1}(u_1)\big) \cdots f_d\big(F_d^{-1}(u_d)\big)}, \tag{2.30}$$

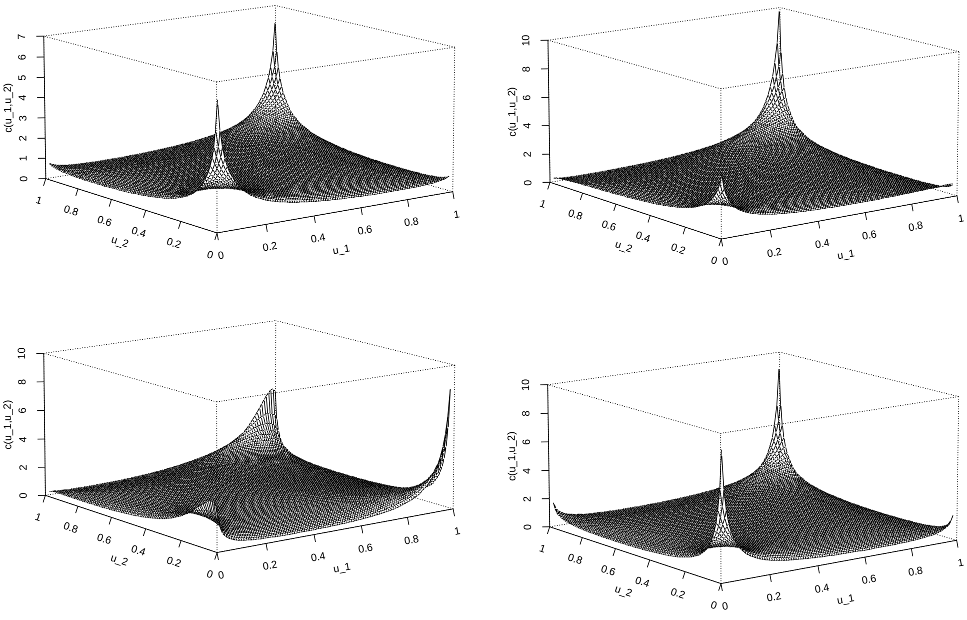


Figure 2.1: Densities of implied copulas of bivariate GH distributions and their limits. The underlying distributions are as follows:
top left: symmetric $NIG_2(10, \mathbf{0}, 0.2, \mathbf{0}, \bar{\Delta})$,
top right: skewed $NIG_2(10, \left(\begin{smallmatrix}4\\1\end{smallmatrix}\right), 0.2, \mathbf{0}, \bar{\Delta})$,
bottom left: skewed $NIG_2(4, \left(\begin{smallmatrix}3\\-2\end{smallmatrix}\right), 0.2, \mathbf{0}, \bar{\Delta})$, bottom right: $t_2(-2, 2, \mathbf{0}, \bar{\Delta})$.
For all distributions $\bar{\Delta} = \left(\begin{smallmatrix}1 & \rho\\ \rho & 1\end{smallmatrix}\right)$ with $\rho = 0.3$.

where the last equation immediately follows from (2.29). Combining (2.30) and Theorem 2.11 a) allows to calculate the copula densities $c_{GH_d(\lambda,\alpha,\beta,\delta,\mu,\Delta)}(u)$ of all multivariate GH distributions including the aforementioned limits. Some results for the bivariate case are visualized in Figure 2.1 above. Note that the choice of $\rho = 0.3$ implies $\det(\bar{\Delta}) = 1 - \rho^2 < 1$, so the parameters of the t- and NIG distributions are the barred ones $(\bar{\lambda}, \bar{\alpha}, \bar{\beta}, \bar{\delta}, \bar{\mu})$ defined in the remark on page 73. If $\bar{\beta} = \beta = \mathbf{0}$, then by equations (2.12)–(2.14) $\bar{\Delta}$ equals the correlation matrix of the related distribution.

Apart from being inherent in every multivariate distribution, the importance of copulas relies on the fact that they encode the dependencies between the margins F_i of F. Many popular dependence measures like, for example, Kendall's tau, Spearman's rho, or the Gini coefficient can be expressed and calculated solely in terms of the associated copulas (see McNeil, Frey, and Embrechts 2005, Proposition 5.29, and Nelsen 1999, Corollary 5.1.13). Thus the assertion of Sklar's Theorem might alternatively be stated in the following way: Every multivariate distribution can be split up into two parts, the marginal distributions and the dependence structure. The next proposition shows that copulas and hence all dependence measures that can be derived from them are invariant under strictly increasing transformations of the margins.

Proposition 2.20 *Suppose that* $(X_1, \dots, X_d)^\top$ *is a random vector with joint distribution function* F*, continuous margins* F_i*,* $1 \le i \le d$*, and implied copula* C_F *given by* (2.29)*. Let* $T_1, \dots, T_d$ *be strictly increasing functions and* G *be the joint distribution function of* $\big(T_1(X_1), \dots, T_d(X_d)\big)^\top$*. Then the implied copulas of* F *and* G *coincide, that is,* $C_F = C_G$*.*

PROOF: See McNeil, Frey, and Embrechts (2005, Proposition 5.6). □

REMARK: From the above proposition it especially follows that the correlation of two random variables does not depend on the inherent copula of their joint distribution alone because correlation is invariant under (strictly) increasing linear transformations only, but not under arbitrary increasing mappings. Correlation is also linked to the marginal distributions since it requires them to possess finite second moments to be well defined, whereas by Sklar's Theorem a copula of the joint distribution always exists without imposing any conditions on the margins.

We now turn to the dependence measure we shall be concerned with for the rest of the present section, the coefficients of tail dependence, which are formally defined by

Definition 2.21 *Let* F *be the joint distribution function of the bivariate random vector* $(X_1, X_2)^\top$ *and* F_1, F_2 *be the marginal distribution functions of* X_1 *and* X_2*, then the* coefficient of upper tail dependence *of* F *resp.* X_1 *and* X_2 *is*

$$\lambda_u := \lambda_u(F) = \lambda_u(X_1, X_2) = \lim_{q \uparrow 1} P\big(X_2 > F_2^{-1}(q) \,|\, X_1 > F_1^{-1}(q)\big),$$

provided a limit $\lambda_u \in [0,1]$ *exists. If* $0 < \lambda_u \le 1$*, then* F *resp.* X_1 *and* X_2 *are said to be* upper tail dependent*; if* $\lambda_u = 0$*, they are called* upper tail independent *or asymptotically independent in the upper tail. Similarly, the* coefficient of lower tail dependence *is*

$$\lambda_l := \lambda_l(F) = \lambda_l(X_1, X_2) = \lim_{q \downarrow 0} P\big(X_2 \le F_2^{-1}(q) \,|\, X_1 \le F_1^{-1}(q)\big),$$

again provided a limit $\lambda_l \in [0,1]$ *exists. If* $\lambda_u = \lambda_l = 0$*, then* F *resp.* X_1 *and* X_2 *are* tail independent*.*

REMARK: If the distribution functions F_1 and F_2 are not continuous and strictly increasing, F_1^{-1} and F_2^{-1} in the previous definition again have to be understood as generalized inverses as defined on page 91.

The larger (or less) q, the more rare is the event $\{X_i > F_i^{-1}(q)\}$ (respectively $\{X_i \le F_i^{-1}(q)\}$). Thus the coefficients of tail dependence are nothing but the limits of the conditional probabilities that the second random variable takes extremal values given the first one also does so. In other words, they may be regarded as the probabilities of joint extremal outcomes of X_1 and X_2. This concept also is of some importance in finance: Suppose for example that X_1 and X_2 represent two risky assets. If their joint distribution is lower tail dependent, the possibility that both of them suffer severe losses at the same time cannot be neglected. In portfolio credit risk models, X_1 and X_2 may be the state variables of two different firms or credit instruments, and the coefficient of lower tail dependence can then be interpreted as the probability of a joint default. Tail dependence is a copula property, which is illustrated by the subsequent

Proposition 2.22 *Let $(X_1, X_2)^\top$ be a bivariate random vector with joint distribution function F, continuous margins F_1, F_2, and implied copula C_F as defined in (2.29). Then the following holds:*

a) The coefficients of lower and upper tail dependence can be calculated by

$$\lambda_l = \lim_{q \downarrow 0} \frac{C_F(q,q)}{q} \qquad \text{and} \qquad \lambda_u = \lim_{q \uparrow 1} \frac{1 - 2q + C_F(q,q)}{1-q}.$$

b) If in addition F_1, F_2 are strictly increasing, λ_l and λ_u can be obtained by

$$\lambda_l = \lim_{q \downarrow 0} P\big(X_2 \le F_2^{-1}(q) \,|\, X_1 = F_1^{-1}(q)\big) + \lim_{q \downarrow 0} P\big(X_1 \le F_1^{-1}(q) \,|\, X_2 = F_2^{-1}(q)\big),$$
$$\lambda_u = \lim_{q \uparrow 1} P\big(X_2 > F_2^{-1}(q) \,|\, X_1 = F_1^{-1}(q)\big) + \lim_{q \uparrow 1} P\big(X_1 > F_1^{-1}(q) \,|\, X_2 = F_2^{-1}(q)\big).$$

Proof: The assertion of part a) of the proposition can be found in many textbooks on copulas and dependence, and part b) essentially follows from the ideas of McNeil, Frey, and Embrechts (2005, pp. 197 and 210). However, due to its importance we provide a detailed proof here for the sake of completeness.

a) By Definition 2.21 we have

$$\begin{aligned}\lambda_l &= \lim_{q \downarrow 0} P\big(X_2 \le F_2^{-1}(q) \,|\, X_1 \le F_1^{-1}(q)\big) = \lim_{q \downarrow 0} \frac{P\big(X_1 \le F_1^{-1}(q), X_2 \le F_2^{-1}(q)\big)}{P\big(X_1 \le F_1^{-1}(q)\big)} \\ &= \lim_{q \downarrow 0} \frac{F\big(F_1^{-1}(q), F_2^{-1}(q)\big)}{F_1\big(F_1^{-1}(q)\big)} \underset{(2.29)}{=} \lim_{q \downarrow 0} \frac{C_F(q,q)}{q},\end{aligned}$$

and similarly we obtain

$$\begin{aligned}\lambda_u &= \lim_{q \uparrow 1} P\big(X_2 > F_2^{-1}(q) \,|\, X_1 > F_1^{-1}(q)\big) = \lim_{q \uparrow 1} \frac{P\big(X_1 > F_1^{-1}(q), X_2 > F_2^{-1}(q)\big)}{P\big(X_1 > F_1^{-1}(q)\big)} \\ &= \lim_{q \uparrow 1} \frac{1 - P\big(\{X_1 \le F_1^{-1}(q)\} \cup \{X_2 \le F_2^{-1}(q)\}\big)}{1 - P\big(X_1 \le F_1^{-1}(q)\big)} \\ &= \lim_{q \uparrow 1} \frac{1 - F_1\big(F_1^{-1}(q)\big) - F_2\big(F_2^{-1}(q)\big) + F\big(F_1^{-1}(q), F_2^{-1}(q)\big)}{1 - F_1\big(F_1^{-1}(q)\big)} \\ &= \lim_{q \uparrow 1} \frac{1 - 2q + C_F(q,q)}{1-q}.\end{aligned}$$

(Note that the continuity of F_i is required to apply equation (2.29) and for $F_i\big(F_i^{-1}(q)\big) = q$ to hold).

b) Since F_1 and F_2 are continuous, the random variables $U_i := F_i(X_i)$, $i = 1, 2$, are both uniformly distributed on $(0,1)$. By Definition 2.18, the joint distribution function of $(U_1, U_2)^\top$ thus is a copula itself which by Proposition 2.20

equals C_F because F_1 and F_2 are strictly increasing by assumption, that is, $P(U_1 \le u_1, U_2 \le u_2) = C_F(u_1, u_2)$. Moreover, for every bivariate copula C we have $0 \le \frac{\partial C(u_1,u_2)}{\partial u_i} \le 1$, $i = 1, 2$, and the partial derivatives exist for Lebesgue-almost all $(u_1, u_2) \in [0, 1]^2$ (Nelsen 1999, Theorem 2.2.7). Together we obtain

$$\begin{aligned}\frac{\partial C}{\partial u_1}(u_1, u_2) &= \lim_{\epsilon \to 0} \frac{C_F(u_1 + \epsilon, u_2) - C_F(u_1, u_2)}{\epsilon} \\ &= \lim_{\epsilon \to 0} \frac{P(u_1 \le U_1 \le u_1 + \epsilon, U_2 \le u_2)}{P(u_1 \le U_1 \le u_1 + \epsilon)} = P(U_2 \le u_2 \,|\, U_1 = u_1)\end{aligned}$$

and analogously $\frac{\partial C(u_1,u_2)}{\partial u_2} = P(U_1 \le u_1 \,|\, U_2 = u_2)$. The continuity and strict monotonicity of the F_i further implies $\{U_i \le u_i\} = \{X_i \le F_i^{-1}(u_i)\}$ for all $u_i \in (0, 1)$ and hence $P(U_i \le u_i \,|\, U_j = u_j) = P\big(X_i \le F_i^{-1}(u_i) \,|\, X_j = F_j^{-1}(u_j)\big)$ for $i, j \in \{1, 2\}$, $i \ne j$. Applying L 'Hôpital's rule to the expressions derived in part a) we finally get

$$\begin{aligned}\lambda_l &= \lim_{q \downarrow 0} \frac{C_F(q, q)}{q} = \lim_{q \downarrow 0} \frac{\mathrm{d}C(q, q)}{\mathrm{d}q} = \lim_{q \downarrow 0} \left(\frac{\partial C}{\partial u_1}(q, q) + \frac{\partial C}{\partial u_2}(q, q) \right) \\ &= \lim_{q \downarrow 0} \big(P(U_2 \le q \,|\, U_1 = q) + P(U_1 \le q \,|\, U_2 = q) \big) \\ &= \lim_{q \downarrow 0} \big(P\big(X_2 \le F_2^{-1}(q) \,|\, X_1 = F_1^{-1}(q)\big) + P\big(X_1 \le F_1^{-1}(q) \,|\, X_2 = F_2^{-1}(q)\big) \big)\end{aligned}$$

and, with the same reasoning,

$$\begin{aligned}\lambda_u &= \lim_{q \uparrow 1} \frac{1 - 2q + C_F(q, q)}{1 - q} = \lim_{q \uparrow 1} \left(2 - \frac{\partial C}{\partial u_1}(q, q) - \frac{\partial C}{\partial u_2}(q, q) \right) \\ &= \lim_{q \uparrow 1} \big((1 - P(U_2 \le q \,|\, U_1 = q) + 1 - P(U_1 \le q \,|\, U_2 = q) \big) \\ &= \lim_{q \uparrow 1} \big(P\big(X_2 > F_2^{-1}(q) \,|\, X_1 = F_1^{-1}(q)\big) + P\big(X_1 > F_1^{-1}(q) \,|\, X_2 = F_2^{-1}(q)\big) \big) .\end{aligned}$$

□

With the help of these preliminaries we are now able to give a complete answer to the question which members of the multivariate GH family show tail dependence and which do not. To our knowledge, only symmetric GH distributions have been considered in this regard in the literature so far. By equation (2.1) and Corollary 2.10, every multidimensional GH distribution with parameter $\beta = \mathbf{0}$ belongs to the class of elliptical distributions, thus the tail independence of $GH_d(\lambda, \alpha, \mathbf{0}, \delta, \mu, \Delta)$ (apart from the t limit case with $\alpha = 0$) can be deduced from the more general result below of Hult and Lindskog (2002, Theorem 4.3). It uses the representation $X \stackrel{d}{=} \mu + RAS$ of an elliptically distributed random vector X which was introduced in Corollary 2.9 and the remark thereafter.

Theorem 2.23 *Let* $X \stackrel{d}{=} \mu + RAS \sim E_d\big(\mu, \Sigma, \psi(t)\big)$ *be an elliptically distributed random vector with* $\Sigma_{ii} > 0$, $1 \le i \le d$, *and* $|\rho_{ij}| := |\Sigma_{ij}/\sqrt{\Sigma_{ii}\Sigma_{jj}}| < 1$ *for all* $i \ne j$. *Then the following statements are equivalent:*

a) *The distribution function* F_R *of* R *is regularly varying with exponent* $p < 0$*, that is,* $F_R \in \mathscr{R}_p$ *(see Definition 1.15).*

b) $(X_i, X_j)^\top$ *is tail dependent for all* $i \neq j$*.*

Moreover, if $F_R \in \mathscr{R}_p$ *with* $p < 0$*, then for all* $i \neq j$

$$\lambda_u(X_i, X_j) = \lambda_l(X_i, X_j) = \frac{\int_{(\pi/2 - \arcsin(\rho_{ij}))/2}^{\pi/2} \cos^{|p|}(t)\,\mathrm{d}t}{\int_0^{\pi/2} \cos^{|p|}(t)\,\mathrm{d}t}.$$

If $X \sim N_d(\mu, y\Delta) \circ G$ has a normal variance mixture distribution which is elliptical by Corollary 2.10, then X admits the two stochastic representations $\mu + \sqrt{Z}AW \stackrel{d}{=} X \stackrel{d}{=} \mu + RAS$ where the vector μ and the $d \times d$-matrix A on the left and right hand side coincide. This equation suggests that the tail behaviour of the distribution F_R of R is mainly influenced by the distribution G of Z and vice versa. Indeed, one can show that F_R is regularly varying with exponent $2p < 0$ ($F_R \in \mathscr{R}_{2p}$) if and only if $G \in \mathscr{R}_p$ (see McNeil, Frey, and Embrechts 2005, pp. 92 and 295f).

Suppose now $X \sim GH_d(\lambda, \alpha, \mathbf{0}, \delta, \mu, \Delta)$ (excluding the t limiting case for a moment), then by equation (1.2) the density of the corresponding mixing distribution $GIG(\lambda, \delta, \alpha)$ has a semi-heavy right tail in the sense of Definition 1.12 with constants $a_2 = \lambda - 1$, $b_2 = \frac{\alpha^2}{2}$ and $c_2 = \frac{(\alpha/\delta)^\lambda}{2K_\lambda(\delta\alpha)}$. (In case of the VG limit, the density of the mixing Gamma distribution $G\big(\lambda, \frac{\alpha^2}{2}\big)$ also has a semi-heavy right tail with the same constants a_2 and b_2, but $c_2 = \frac{(\alpha^2/2)^\lambda}{\Gamma(\lambda)}$.) By Proposition 1.13 and Definition 1.14, the distribution functions of $GIG(\lambda, \delta, \alpha)$ and $G\big(\lambda, \frac{\alpha^2}{2}\big)$ both have an exponential right tail with rate b_2. In view of Definition 1.15 and the subsequent remark, distribution functions with exponential right tails can be regarded as regularly varying with exponent $-\infty$. Consequently, for the distribution function F_R of R in the representation $X \stackrel{d}{=} \mu + RAS$ we have $F_R \in \mathscr{R}_{-\infty}$ as well. Applying Theorem 2.23 yields

$$\lambda_u(X_i, X_j) = \lambda_l(X_i, X_j) = \lim_{p \to -\infty} \frac{\int_{(\pi/2 - \arcsin(\rho_{ij}))/2}^{\pi/2} \cos^{|p|}(t)\,\mathrm{d}t}{\int_0^{\pi/2} \cos^{|p|}(t)\,\mathrm{d}t} = 0,$$

showing the tail independence of all symmetric $GH_d(\lambda, \alpha, \mathbf{0}, \delta, \mu, \Delta)$-distributions with parameter $\alpha > 0$.

Remark: The convergence of the ratio of the two integrals can be justified as follows: Since $h : (-\frac{\pi}{2}, \frac{\pi}{2}) \to \mathbb{R}_-$ with $h(x) = \log(\cos(x))$ has an absolute maximum at $x_0 = 0$ and $h''(x) = -\cos^{-2}(x)$, an application of Laplace's method shows that for all $0 < b \leq \frac{\pi}{2}$

$$\int_0^b \cos^{|p|}(t)\,\mathrm{d}t = \int_0^b e^{|p|h(t)}\,\mathrm{d}t \sim \sqrt{\frac{\pi}{-2|p|h''(0)}}\, e^{|p|h(0)} = \sqrt{\frac{\pi}{2|p|}}, \quad |p| \to \infty,$$

consequently

$$\lim_{p \to -\infty} \frac{\int_{(\pi/2 - \arcsin(\rho_{ij}))/2}^{\pi/2} \cos^{|p|}(t)\,\mathrm{d}t}{\int_0^{\pi/2} \cos^{|p|}(t)\,\mathrm{d}t} = 1 - \lim_{p \to -\infty} \frac{\int_0^{(\pi/2 - \arcsin(\rho_{ij}))/2} \cos^{|p|}(t)\,\mathrm{d}t}{\int_0^{\pi/2} \cos^{|p|}(t)\,\mathrm{d}t} = 0.$$

In the t limiting case, however, we have $X \overset{d}{=} \mu + \sqrt{Z}AW \sim t_d(\lambda, \delta, \mu, \Delta)$ with $Z \sim iG\big(\lambda, \frac{\delta^2}{2}\big)$, and from equation (1.4) it is easily seen that the density $d_{iG(\lambda,\delta^2/2)}$ is regularly varying with exponent $\lambda - 1$. Hence $G = F_Z \in \mathscr{R}_\lambda$ and thus, as pointed out above, $F_R \in \mathscr{R}_{2\lambda}$, so we conclude from Theorem 2.23 that $\lambda_u(X_i, X_j) = \lambda_l(X_i, X_j) > 0$ for all t distributions $t_d(\lambda, \delta, \mu, \Delta)$. The coefficients are quantified more accurately in Proposition 2.24 below.

This main result of the present section shows that the dependence behaviour can change dramatically if we move from symmetric to skewed GH distributions with parameter $\beta \neq \mathbf{0}$: in addition to tail independence also complete dependence can occur, that is, both of the coefficients λ_l and λ_u may be equal to one. More precisely we have

Proposition 2.24 *Let $X \sim GH_2(\lambda, \alpha, \beta, \delta, \mu, \Delta)$ and define $\rho := \frac{\Delta_{12}}{\sqrt{\Delta_{11}\Delta_{22}}}$ as well as $\bar{\beta}_i := \sqrt{\Delta_{ii}}\beta_i$ for $i = 1, 2$. Then the following holds:*

a) If $0 \leq \sqrt{\langle \beta, \Delta\beta \rangle} < \alpha$, then the GH distribution (including possible VG limits) is tail independent if $-1 < \rho \leq 0$. If $0 < \rho < 1$, then

$$\lambda_l(X_1, X_2) = \lambda_u(X_1, X_2) = \begin{cases} 0, \; c_*, c_*^{-1} > \rho, \\ 1, \; \min(c_*, c_*^{-1}) < \rho, \end{cases}$$

where $c_* := \frac{\sqrt{\alpha^2 - (1-\rho^2)\bar{\beta}_2^2} + \bar{\beta}_1 + \rho\bar{\beta}_2}{\sqrt{\alpha^2 - (1-\rho^2)\bar{\beta}_1^2} + \bar{\beta}_2 + \rho\bar{\beta}_1}$.

b) If $\lambda < 0$ and $\alpha = 0$, then $X \sim t_2(\lambda, \delta, \mu, \Delta)$ and

$$\lambda_u(X_1, X_2) = \lambda_l(X_1, X_2) = 2\,F_{t(\lambda - \frac{1}{2}, \sqrt{-2\lambda+1}, 0)}\left(-\sqrt{\frac{(-2\lambda+1)(1-\rho)}{1+\rho}}\right),$$

where $F_{t(\lambda-\frac{1}{2},\sqrt{-2\lambda+1},0)}$ denotes the distribution function of the univariate Student's t-distribution $t\big(\lambda - \frac{1}{2}, \sqrt{-2\lambda+1}, 0\big)$ with $f = -2\lambda + 1$ degrees of freedom.

c) Let $\lambda < 0$ and $0 < \sqrt{\langle \beta, \Delta\beta \rangle} = \alpha$. If $(\bar{\beta}_1 + \rho\bar{\beta}_2)(\bar{\beta}_2 + \rho\bar{\beta}_1) < 0$, then

$$\lambda_u(X_1, X_2) = \lambda_l(X_1, X_2) = \begin{cases} 0, \; \rho < 0, \\ 1, \; \rho > 0. \end{cases}$$

If $(\bar{\beta}_1 + \rho\bar{\beta}_2)(\bar{\beta}_2 + \rho\bar{\beta}_1) > 0$, then

$$\lambda_u(X_1, X_2) = \lambda_l(X_1, X_2) = \begin{cases} 0, \; c_*, c_*^{-1} > \rho, \\ 1, \; \min(c_*, c_*^{-1}) < \rho, \end{cases} \quad \text{where } c_* := \frac{\bar{\beta}_1 + \rho\bar{\beta}_2}{\bar{\beta}_2 + \rho\bar{\beta}_1}.$$

For the proof, we need the following lemma which is a slightly modified version of Banachewicz and van der Vaart (2008, Lemma 3.1):

Lemma 2.25 *Suppose $F : \mathbb{R} \to [0, 1]$ is a continuous and strictly increasing distribution function.*

a) If $F(y) \sim c_1|y|^{-a_1}$ as $y \to -\infty$ and $1 - F(y) \sim c_2 y^{-a_2}$ as $y \to \infty$ for some $a_1, a_2, c_1, c_2 > 0$, then $F^{-1}(u) \sim -\left(\frac{c_1}{u}\right)^{\frac{1}{a_1}}$ and $F^{-1}(1-u) \sim \left(\frac{c_2}{u}\right)^{\frac{1}{a_2}}$ for $u \downarrow 0$.

b) If instead $F(y) \sim c_1|y|^{a_1}e^{-b_1|y|}$ as $y \to -\infty$ and $1 - F(y) \sim c_2 y^{a_2} e^{-b_2 y}$ as $y \to \infty$ for some $a_1, a_2 \in \mathbb{R}$ and $b_1, b_2, c_1, c_2 > 0$, then $F^{-1}(u) \sim \frac{\log(u)}{b_1}$ and $F^{-1}(1-u) \sim -\frac{\log(u)}{b_2}$ for $u \downarrow 0$.

Proof: a) If $1 - F(y) \sim c_2 y^{-a_2}$ as $y \to \infty$, then for any $r > 0$

$$\lim_{u \downarrow 0} \frac{1 - F\big(r\left(\frac{c_2}{u}\right)^{\frac{1}{a_2}}\big)}{u} = r^{-a_2}.$$

For $r < 1$, the right hand side of the above equation is greater than one, so we conclude that in this case $1 - F\big(r\left(\frac{c_2}{u}\right)^{\frac{1}{a_2}}\big) > u$ for sufficiently small u and hence $F^{-1}(1-u) > r\left(\frac{c_2}{u}\right)^{\frac{1}{a_2}}$ (note that the assumptions on F imply $F^{-1}\big(F(y)\big) = y$ for all $y \in \mathbb{R}$). If $r > 1$, then we similarly obtain $1 - F\big(r\left(\frac{c_2}{u}\right)^{\frac{1}{a_2}}\big) < u$ and thus $F^{-1}(1-u) < r\left(\frac{c_2}{u}\right)^{\frac{1}{a_2}}$ for sufficiently small u. This proves the assertion for $F^{-1}(1-u)$, and the asymptotic behaviour of $F^{-1}(u)$ for $u \downarrow 0$ can be shown analogously.

b) If $1 - F(y) \sim c_2 y^{a_2} e^{-b_2 y}$ as $y \to \infty$, then we have

$$\lim_{u \downarrow 0} \frac{1 - F\big(-\frac{r\log(u)}{b_2}\big)}{u} = \lim_{u \downarrow 0} c_2 \left(-\frac{r\log(u)}{b_2}\right)^{a_2} u^{r-1} = \begin{cases} \infty, & r < 1, \\ 0, & r > 1. \end{cases}$$

With the same reasoning as before we conclude $F^{-1}(1-u) \sim -\frac{\log(u)}{b_2}$ for $u \downarrow 0$, and the corresponding result for $F^{-1}(u)$ is easily obtained along the same lines. □

Proof of Proposition 2.24: Propositions 2.22 and 2.20 state that tail dependence is a copula property and therefore invariant under strictly increasing transformations of X_1 and X_2. But if $X \sim GH_2(\lambda, \alpha, \beta, \delta, \mu, \Delta)$, the linear transformation $Y = \left(\begin{smallmatrix} 1/\sqrt{\Delta_{11}} & 0 \\ 0 & 1/\sqrt{\Delta_{22}} \end{smallmatrix}\right)(X - \mu)$ obviously is strictly increasing in each component, and Theorem 2.11 c) implies that $Y \sim GH_2(\bar{\lambda}, \bar{\alpha}, \bar{\beta}, \bar{\delta}, \mathbf{0}, \bar{\Delta})$ with $\bar{\lambda} = \lambda$, $\bar{\alpha} = \alpha$, $\bar{\beta} = \left(\begin{smallmatrix} \sqrt{\Delta_{11}} & 0 \\ 0 & \sqrt{\Delta_{22}} \end{smallmatrix}\right)\beta$, $\bar{\delta} = \delta$, $\bar{\Delta} = \left(\begin{smallmatrix} 1 & \rho \\ \rho & 1 \end{smallmatrix}\right)$ and $\rho := \Delta_{12}/\sqrt{\Delta_{11}\Delta_{22}}$. Note that we here use the barred parameters defined in the remark on p. 73 because in general $\det(\bar{\Delta}) = 1 - \rho^2 < 1$. As already pointed out in the remarks on pages 73 and 76, these considerations remain also valid for all GH limit distributions. Hence we can and will always assume $X \sim GH_2(\lambda, \alpha, \bar{\beta}, \delta, \mathbf{0}, \bar{\Delta})$ in the following. The fact that Δ is supposed to be positive definite with $\det(\Delta) = 1$ by definition implies the inequality $0 < \frac{1}{\Delta_{11}\Delta_{22}} = \frac{\Delta_{11}\Delta_{22} - \Delta_{12}^2}{\Delta_{11}\Delta_{22}} = 1 - \rho^2$, thus $|\rho| < 1$.

a) If $X \sim GH_2(\lambda, \alpha, \bar{\beta}, \delta, \mathbf{0}, \bar{\Delta})$ and $0 \le \sqrt{\langle \beta, \Delta\beta \rangle} < \alpha$, then by Theorem 2.11 a) the marginal distributions are $X_1 \sim GH\big(\lambda, (\alpha^2 - (1-\rho^2)\bar{\beta}_2^2)^{1/2}, \bar{\beta}_1 + \rho\bar{\beta}_2, \delta, 0\big)$

and $X_2 \sim GH\big(\lambda, (\alpha^2 - (1-\rho^2)\bar{\beta}_1^2)^{1/2}, \bar{\beta}_2 + \rho\bar{\beta}_1, \delta, 0\big)$. To simplify notations we set $\hat{\alpha}_1 := (\alpha^2 - (1-\rho^2)\bar{\beta}_2^2)^{1/2}$, $\hat{\beta}_1 := \bar{\beta}_1 + \rho\bar{\beta}_2$, and $\hat{\alpha}_2 := (\alpha^2 - (1-\rho^2)\bar{\beta}_1^2)^{1/2}$, $\hat{\beta}_2 := \bar{\beta}_2 + \rho\bar{\beta}_1$, then we obtain $\hat{\alpha}_1^2 - \hat{\beta}_1^2 = \hat{\alpha}_2^2 - \hat{\beta}_2^2 = \alpha^2 - \langle \beta, \Delta\beta\rangle > 0$. Thus the densities of $\mathcal{L}(X_1)$ and $\mathcal{L}(X_2)$ both have semi-heavy tails (see Definition 1.12 and the remark thereafter), and Proposition 1.13 (or equivalently Corollary 1.17) implies that the corresponding distribution functions F_1 and F_2 fulfill the assumptions of Lemma 2.25 b) with $b_1 = \hat{\alpha}_i + \hat{\beta}_i$ and $b_2 = \hat{\alpha}_i - \hat{\beta}_i$, $i = 1, 2$. From this we conclude that $F_1^{-1}(q) \sim c_l F_2^{-1}(q)$ for $q \downarrow 0$ as well as $F_1^{-1}(q) \sim c_u F_2^{-1}(q)$ for $q \uparrow 1$ where $c_l := \frac{\hat{\alpha}_2 + \hat{\beta}_2}{\hat{\alpha}_1 + \hat{\beta}_1} > 0$ and $c_u := \frac{\hat{\alpha}_2 - \hat{\beta}_2}{\hat{\alpha}_1 - \hat{\beta}_1} > 0$. Note that $c_l c_u = \frac{\hat{\alpha}_2^2 - \hat{\beta}_2^2}{\hat{\alpha}_1^2 - \hat{\beta}_1^2} = 1$ and thus $c_u = c_l^{-1}$. All this also holds in the VG limit case with $\delta = 0$ because Theorem 2.11 a) still applies there and the univariate VG marginal densities have semi-heavy tails, too (see page 22).

By Theorem 2.11 b), the conditional distribution of X_i given $X_j = x_j$ (where here and in the following $i, j \in \{1, 2\}$ and $i \neq j$) is given by $P(X_i \,|\, X_j = x_j) = GH\big(\lambda - \frac{1}{2}, \alpha(1-\rho^2)^{-1/2}, \bar{\beta}_i, \sqrt{\delta^2 + x_j^2}\sqrt{1-\rho^2}, \rho x_j\big)$, and part c) of the same theorem then yields

$$
\begin{aligned}
&P\left(\frac{X_i - \rho x_j}{\sqrt{\delta^2 + x_j^2}\sqrt{1-\rho^2}} \,\middle|\, X_j = x_j\right) \\
&\qquad\qquad = GH\Big(\lambda - \tfrac{1}{2}, \alpha\sqrt{\delta^2 + x_j^2}, \bar{\beta}_i\sqrt{\delta^2 + x_j^2}\sqrt{1-\rho^2}, 1, 0\Big) \\
&\qquad\qquad =: GH^*_{i|j}\big(\lambda - \tfrac{1}{2}, \alpha, \bar{\beta}_i, \delta, \rho, x_j\big).
\end{aligned}
$$

Again, this also remains true in the VG limit case (see the remark on p. 76). Let $F^q_{i|j}$ denote the distribution function of $GH^*_{i|j}\big(\lambda - \frac{1}{2}, \alpha, \bar{\beta}_i, \delta, \rho, F_j^{-1}(q)\big)$ and set

$$
h_{i|j}(q) := (1-\rho^2)^{-\frac{1}{2}} \frac{F_i^{-1}(q) - \rho F_j^{-1}(q)}{\sqrt{\delta^2 + \big(F_j^{-1}(q)\big)^2}} \qquad \text{for } q \in (0,1),
$$

then we have

$$
\begin{aligned}
\lim_{q\downarrow 0} P\big(X_i \le F_i^{-1}(q) \,\big|\, X_j = F_j^{-1}(q)\big) &= \lim_{q\downarrow 0} F^q_{i|j}\big(h_{i|j}(q)\big), \\
\lim_{q\uparrow 1} P\big(X_i > F_i^{-1}(q) \,\big|\, X_j = F_j^{-1}(q)\big) &= \lim_{q\uparrow 1} 1 - F^q_{i|j}\big(h_{i|j}(q)\big).
\end{aligned}
$$

Moreover, if $\alpha > |\beta| \ge 0$, then $GH(\lambda, r\alpha, r\beta, \delta, \mu) \xrightarrow{w} \epsilon_\mu$ for $r \to \infty$ because

$$
\begin{aligned}
&\lim_{r\to\infty} \phi_{GH(\lambda, r\alpha, r\beta, \delta, \mu)}(u) = \\
&\quad = \lim_{r\to\infty} e^{iu\mu} \left(\frac{(r\alpha)^2 - (r\beta)^2}{(r\alpha)^2 - (r\beta + iu)^2}\right)^{\frac{\lambda}{2}} \frac{K_\lambda\big(\delta\sqrt{(r\alpha)^2 - (r\beta + iu)^2}\big)}{K_\lambda\big(\delta\sqrt{(r\alpha)^2 - (r\beta)^2}\big)} \\
&\quad = \lim_{r\to\infty} e^{iu\mu} \left(\frac{\alpha^2 - \beta^2}{\alpha^2 - (\beta + \frac{iu}{r})^2}\right)^{\frac{\lambda}{2}} \frac{K_\lambda\big(r\delta\sqrt{\alpha^2 - (\beta + \frac{iu}{r})^2}\big)}{K_\lambda\big(r\delta\sqrt{\alpha^2 - \beta^2}\big)} = e^{iu\mu}
\end{aligned}
$$

which implies that $GH^*_{i|j}\big(\lambda-\frac{1}{2},\alpha,\bar\beta_i,\delta,\rho,F_j^{-1}(q)\big)$ converges weakly to the degenerate distribution ϵ_0 if $q\downarrow 0$ or $q\uparrow 1$. From the asymptotic relations of the quantile functions $F_1^{-1}(q)$ and $F_2^{-1}(q)$ we further obtain

$$\lim_{q\downarrow 0} h_{i|j}(q)=(1-\rho^2)^{-\frac{1}{2}}\big(\rho-c_l^{j-i}\big),\quad\text{and}\quad \lim_{q\uparrow 1} h_{i|j}(q)=(1-\rho^2)^{-\frac{1}{2}}\big(c_l^{i-j}-\rho\big)$$

(remember $c_u=c_l^{-1}$), consequently

$$\lim_{q\downarrow 0} P\big(X_i\le F_i^{-1}(q)\,\big|\,X_j=F_j^{-1}(q)\big)=F_{\epsilon_0}\left(\frac{\rho-c_l^{j-i}}{\sqrt{1-\rho^2}}\right)=\begin{cases}0, & c_l^{j-i}>\rho,\\ 1, & c_l^{j-i}<\rho,\end{cases}$$

as well as

$$\lim_{q\uparrow 1} P\big(X_i> F_i^{-1}(q)\,\big|\,X_j=F_j^{-1}(q)\big)=1-F_{\epsilon_0}\left(\frac{c_l^{i-j}-\rho}{\sqrt{1-\rho^2}}\right)=\begin{cases}0, & c_l^{i-j}>\rho,\\ 1, & c_l^{i-j}<\rho,\end{cases}$$

and Proposition 2.22 b) finally implies that $\lambda_l(X_1,X_2)=\lambda_u(X_1,X_2)=0$ if and only if $c_l,c_l^{-1}>\rho$. Since $c_l>0$, the conditions are trivially met if $\rho\le 0$. If $0<\rho<1$, then at most one of the quantities c_l and c_l^{-1} can be smaller than ρ (note that the convergence to a well-defined limit cannot be assured if $c_l^{j-i}=\rho>0$, therefore we exclude these possibilities in our considerations). This completes the proof of part a).

c) Because Theorem 2.11 a) still applies if $X\sim GH_2(\lambda,\alpha,\bar\beta,\delta,\mathbf{0},\bar\Delta)$, $\lambda<0$, and $0<\sqrt{\langle\beta,\Delta\beta\rangle}=\alpha$, we have, using the notations from above, that $X_i\sim GH(\lambda,\hat\alpha_i,\hat\beta_i,\delta,0)$, $i=1,2$. However, in this case $\hat\alpha_i^2-\hat\beta_i^2=\alpha^2-\sqrt{\langle\beta,\Delta\beta\rangle}=0$, hence both marginal distributions are univariate GH limit distributions with $\lambda<0$ and $\hat\alpha_i=|\hat\beta_i|$. If $\hat\beta_i>0$, we conclude from equations (1.20), (A.10) and Proposition 1.13 that the tail behaviour of the distribution function is given by $F_i(y)\sim c_{i1}|y|^{\lambda-1}e^{-2\hat\alpha_i|y|}$ for $y\to-\infty$ and $1-F_i(y)\sim c_{i2}|y|^{\lambda}$ as $y\to\infty$ where

$$c_{i1}=\frac{2^{\lambda-1}}{\hat\alpha_i^{\lambda+1}\delta^{2\lambda}\Gamma(|\lambda|)}\qquad\text{and}\qquad c_{i2}=\frac{2^{\lambda}}{|\lambda|\,\hat\alpha_i^{\lambda}\delta^{2\lambda}\Gamma(|\lambda|)}.$$

Lemma 2.25 now states that $F_i^{-1}(q)\sim\frac{\log(q)}{2\hat\alpha_i}$ for $q\downarrow 0$ and $F_i^{-1}(q)\sim\big(\frac{c_{i2}}{1-q}\big)^{\frac{1}{|\lambda|}}$ for $q\uparrow 1$. If $\hat\beta_i<0$, then we analogously obtain $F_i^{-1}(q)\sim-\big(\frac{c_{i2}}{q}\big)^{\frac{1}{|\lambda|}}$ as $q\downarrow 0$ and $F_i^{-1}(q)\sim-\frac{\log(1-q)}{2\hat\alpha_i}$ as $q\uparrow 1$. Because the case $\hat\beta_i=\bar\beta_i+\rho\bar\beta_j=0$ is ruled out by assumption, the equality $0=\hat\alpha_i^2-\hat\beta_i^2=\alpha^2-(1-\rho^2)\bar\beta_j^2-(\bar\beta_i+\rho\bar\beta_j)^2$ implies that $\alpha>\sqrt{1-\rho^2}|\bar\beta_i|$. Thus we can proceed along the same lines as in the proof of part a) and get

$$\begin{aligned}\lim_{q\downarrow 0} P\big(X_i\le F_i^{-1}(q)\,\big|\,X_j=F_j^{-1}(q)\big)&=F_{\epsilon_0}\Big(\lim_{q\downarrow 0}h_{i|j}(q)\Big),\\ \lim_{q\uparrow 1} P\big(X_i> F_i^{-1}(q)\,\big|\,X_j=F_j^{-1}(q)\big)&=1-F_{\epsilon_0}\Big(\lim_{q\uparrow 1}h_{i|j}(q)\Big)\end{aligned}$$

if we again exclude the cases where $h_{i|j}(q)\to 0$ for the same reasons as above.

Suppose $\hat\beta_1, \hat\beta_2 > 0$, then $F_1^{-1}(q) \sim c_l F_2^{-1}(q)$ with $c_l = \frac{\hat\alpha_2}{\hat\alpha_1} = \frac{\hat\beta_2}{\hat\beta_1} > 0$ as $q \downarrow 0$ and $F_1^{-1}(q) \sim c_u F_2^{-1}(q)$ with $c_u = \big(\frac{c_{12}}{c_{22}}\big)^{1/|\lambda|} = \big(\frac{\hat\alpha_2^\lambda}{\hat\alpha_1^\lambda}\big)^{1/|\lambda|} = \frac{\hat\beta_1}{\hat\beta_2} = c_l^{-1}$ for $q \uparrow 1$. Consequently we again have

$$\lim_{q\downarrow 0} h_{i|j}(q) = (1-\rho^2)^{-\frac{1}{2}}\big(\rho - c_l^{j-i}\big), \quad \text{and} \quad \lim_{q\uparrow 1} h_{i|j}(q) = (1-\rho^2)^{-\frac{1}{2}}\big(c_l^{i-j} - \rho\big)$$

and conclude, analogously as before, that $\lambda_l(X_1, X_2) = \lambda_u(X_1, X_2) = 0$ if and only if $c_l, c_l^{-1} > \rho$. If $\hat\beta_1, \hat\beta_2 < 0$, the tail behaviour of the quantile functions is just exchanged ($c_l \rightsquigarrow c_l^{-1}$ and $c_u = c_l^{-1} \rightsquigarrow c_l$), hence the assertion remains also valid in this case.

Finally, let $\hat\beta_1 > 0$ and $\hat\beta_2 < 0$, then $F_1^{-1}(q) \sim \frac{\log(q)}{2\hat\alpha_1}$ and $F_2^{-1}(q) \sim -\big(\frac{c_{22}}{q}\big)^{\frac{1}{|\lambda|}}$ as $q \downarrow 0$, thus $\lim_{q\downarrow 0} \frac{F_1^{-1}(q)}{F_2^{-1}(q)} = 0$ and

$$\lim_{q\downarrow 0} h_{i|j}(q) = \begin{cases} (1-\rho^2)^{-\frac{1}{2}}\rho, & i-j=-1, \\ -\infty, & i-j=1, \end{cases}$$

hence $\lambda_l(X_1, X_2) = 0$ if and only if $\rho < 0$. Further $F_1^{-1}(q) \sim \big(\frac{c_{12}}{1-q}\big)^{\frac{1}{|\lambda|}}$ and $F_2^{-1}(q) \sim -\frac{\log(1-q)}{2\hat\alpha_1}$ for $q \uparrow 1$, consequently $\lim_{q\uparrow 1} \frac{F_2^{-1}(q)}{F_1^{-1}(q)} = 0$ and

$$\lim_{q\uparrow 1} h_{i|j}(q) = \begin{cases} -(1-\rho^2)^{-\frac{1}{2}}\rho, & i-j=1, \\ \infty, & i-j=-1, \end{cases}$$

which implies that also $\lambda_u(X_1, X_2) = 0$ if and only if $\rho < 0$. Trivially, all conclusions remain true if $\hat\beta_1 < 0$ and $\hat\beta_2 > 0$.

b) The proof of this part goes back to Embrechts, McNeil, and Straumann (2002), see also McNeil, Frey, and Embrechts (2005, p. 211). If $\lambda < 0$ and $\alpha = 0$, we can assume $X \sim GH_2(\lambda, 0, \mathbf{0}, \delta, \mathbf{0}, \bar\Delta) = t_2(\lambda, \delta, \mathbf{0}, \bar\Delta)$, and the marginal distributions are given by $\mathcal{L}(X_1) = \mathcal{L}(X_2) = GH(\lambda, 0, 0, \delta, 0) = t(\lambda, \delta, 0)$ according to Theorem 2.11 a), hence we have $F_1^{-1}(q) = F_2^{-1}(q)$ for all $q \in (0,1)$ in this case. By Theorem 2.11 b), the conditional distributions also coincide, that is, $P(X_2 \,|\, X_1 = x) = P(X_1 \,|\, X_2 = x) = t\big(\lambda, \sqrt{\delta^2 + x^2}\sqrt{1-\rho^2}, \rho x\big)$, and part c) of the same theorem implies

$$\begin{aligned} P\left(\frac{\sqrt{-2\lambda+1}}{\sqrt{1-\rho^2}}\, \frac{X_2 - \rho x}{\sqrt{\delta^2 + x^2}} \,\middle|\, X_1 = x\right) &= P\left(\frac{\sqrt{-2\lambda+1}}{\sqrt{1-\rho^2}}\, \frac{X_1 - \rho x}{\sqrt{\delta^2 + x^2}} \,\middle|\, X_2 = x\right) \\ &= t\big(\lambda - \tfrac{1}{2}, \sqrt{-2\lambda+1}, 0\big). \end{aligned}$$

(Note that, in principle, the additional scaling factor $\sqrt{-2\lambda+1}$ is not necessary, but leads to the relation $\delta^2 = -2\lambda + 1 = -2(\lambda - \frac{1}{2})$ of the parameters of the conditional distribution which therewith becomes a classical Student's t-distributiom with $f = -2\lambda + 1$ degrees of freedom.) If we set

$$h(q) := \frac{\sqrt{-2\lambda+1}}{\sqrt{1-\rho^2}}\, \frac{F_2^{-1}(q) - \rho F_1^{-1}(q)}{\sqrt{\delta^2 + (F_1^{-1}(q))^2}} \quad \text{for } q \in (0,1),$$

we get, using that $F_1^{-1}(q) = F_2^{-1}(q)$,

$$\lim_{q\downarrow 0} h(q) = -\frac{\sqrt{-2\lambda+1}\,(1-\rho)}{\sqrt{1-\rho^2}} = -\sqrt{\frac{(-2\lambda+1)(1-\rho)}{1+\rho}} = -\lim_{q\uparrow 1} h(q),$$

consequently

$$\begin{aligned}
&\lim_{q\downarrow 0} P\big(X_2 \le F_2^{-1}(q) \,\big|\, X_1 = F_1^{-1}(q)\big) = \lim_{q\downarrow 0} P\big(X_1 \le F_1^{-1}(q) \,\big|\, X_2 = F_2^{-1}(q)\big) \\
&= \lim_{q\downarrow 0} F_{t(\lambda-\frac{1}{2},\sqrt{-2\lambda+1},0)}\big(h(q)\big) = F_{t(\lambda-\frac{1}{2},\sqrt{-2\lambda+1},0)}\left(-\sqrt{\frac{(-2\lambda+1)(1-\rho)}{1+\rho}}\right)
\end{aligned}$$

and

$$\begin{aligned}
\lim_{q\uparrow 1} P\big(X_2 > F_2^{-1}(q) \,\big|\, X_1 = F_1^{-1}(q)\big) &= \lim_{q\uparrow 1} P\big(X_1 > F_1^{-1}(q) \,\big|\, X_2 = F_2^{-1}(q)\big) \\
&= \lim_{q\uparrow 1} 1 - F_{t(\lambda-\frac{1}{2},\sqrt{-2\lambda+1},0)}\big(h(q)\big) \\
&= 1 - F_{t(\lambda-\frac{1}{2},\sqrt{-2\lambda+1},0)}\left(\sqrt{\frac{(-2\lambda+1)(1-\rho)}{1+\rho}}\right).
\end{aligned}$$

The symmetry relation $F_{t(\lambda-1/2,\sqrt{-2\lambda+1},0)}(-x) = 1 - F_{t(\lambda-1/2,\sqrt{-2\lambda+1},0)}(x)$ and Proposition 2.22 b) now yield the desired result. $\square$

The conditions $c_* > \rho$ and $c_*^{-1} > \rho$ in Proposition 2.24 a) are trivially fulfilled if $\bar\beta_1 = \bar\beta_2$, because then $c_* = c_*^{-1} = 1$. This in particular includes the case $\beta = \mathbf{0}$ which provides an alternative proof for the tail independence of symmetric GH distributions (apart from the t limit case). In general, however, a checking of the condition might seem to be a little bit cumbersome. The following corollary provides a simpler criterion for tail independence of GH distributions.

Corollary 2.26 *Suppose that $X \sim GH_2(\lambda,\alpha,\beta,\delta,\mu,\Delta)$ and $\rho := \frac{\Delta_{12}}{\sqrt{\Delta_{11}\Delta_{22}}} > 0$. Then $\lambda_l(X_1,X_2) = \lambda_u(X_1,X_2) = 0$ if either $\sqrt{\langle\beta,\Delta\beta\rangle} < \alpha$ and $\beta_1\beta_2 \ge 0$ or $0 < \sqrt{\langle\beta,\Delta\beta\rangle} = \alpha$ and $\beta_1\beta_2 > 0$.*

Proof: According to Proposition 2.24 a) and c), we just have to show that the conditions $\beta_1\beta_2 \ge 0$ resp. > 0 imply $c_*, c_*^{-1} > \rho$. Assume $\sqrt{\langle\beta,\Delta\beta\rangle} < \alpha$ first. If both $\beta_1,\beta_2 \ge 0$, then so are $\bar\beta_1 = \sqrt{\Delta_{11}}\beta_1$ and $\bar\beta_2 = \sqrt{\Delta_{22}}\beta_2$. Since $\rho > 0$, we see from the inequality $0 < \alpha^2 - \langle\beta,\Delta\beta\rangle = \alpha^2 - \bar\beta_1^2 - 2\rho\bar\beta_1\bar\beta_2 - \bar\beta_2^2$ that $\bar\beta_i < \alpha$, $i = 1,2$. Therewith we obtain

$$c_* = \frac{\sqrt{\alpha^2-(1-\rho^2)\bar\beta_2^2} + \bar\beta_1 + \rho\bar\beta_2}{\sqrt{\alpha^2-(1-\rho^2)\bar\beta_1^2} + \bar\beta_2 + \rho\bar\beta_1} > \frac{\sqrt{\alpha^2-(1-\rho^2)\alpha^2} + \rho\bar\beta_1 + \rho\bar\beta_2}{\alpha + \bar\beta_1 + \bar\beta_2} = \rho,$$

and an analogous estimate shows that also $c_*^{-1} > \rho$. If $\beta_1 \le 0$ and $\beta_2 \le 0$, we use the fact that c_*^{-1} may alternatively be represented by $c_*^{-1} = \frac{\sqrt{\alpha^2-(1-\rho^2)\bar\beta_2^2}-\bar\beta_1-\rho\bar\beta_2}{\sqrt{\alpha^2-(1-\rho^2)\bar\beta_1^2}-\bar\beta_2-\rho\bar\beta_1}$ and similarly conclude that $c_*, c_*^{-1} > \rho$.

Now, let $0 < \sqrt{\langle \beta, \Delta\beta \rangle} = \alpha$ and note that the condition $\beta_1\beta_2 > 0$ implies $(\bar\beta_1 + \rho\bar\beta_2)(\bar\beta_2 + \rho\bar\beta_1) > 0$. If both $\beta_1, \beta_2 > 0$, then $c_* = \frac{\bar\beta_1 + \rho\bar\beta_2}{\bar\beta_2 + \rho\bar\beta_1} > \frac{\rho\bar\beta_1 + \rho\bar\beta_2}{\bar\beta_2 + \bar\beta_1} = \rho$, and $c_*^{-1} > \rho$ follows analogously. If $\beta_1, \beta_2 < 0$, the same result is obtained by using the representation $c_* = \frac{-\bar\beta_1 - \rho\bar\beta_2}{-\bar\beta_2 - \rho\bar\beta_1}$. □

Remark: An immediate consequence of the preceding corollary is that complete dependence ($\lambda_l(X_1, X_2) = \lambda_u(X_1, X_2) = 1$) within bivariate GH distributions can only occur if the parameters β_1 and β_2 have opposite signs, and one might conjecture that the conditions $c_*, c_*^{-1} > \rho$ are also always fulfilled in these cases such that a two-dimensional GH distribution would be tail independent for almost any choice of parameters. However, this is not true, and it is fairly easy to construct counterexamples: Take $\alpha = 4$, $\bar\beta_1 = 3$, $\bar\beta_2 = -2$, and $\rho = 0.3$, then $\alpha^2 - \langle \beta, \Delta\beta \rangle = \alpha^2 - \bar\beta_1^2 - 2\rho\bar\beta_1\bar\beta_2 - \bar\beta_2^2 = 6.6$ and

$$c_*^{-1} = \frac{\sqrt{\alpha^2 - (1-\rho^2)\bar\beta_1^2} + \bar\beta_2 + \rho\bar\beta_1}{\sqrt{\alpha^2 - (1-\rho^2)\bar\beta_2^2} + \bar\beta_1 + \rho\bar\beta_2} \approx 0.286 < \rho.$$

The corresponding copula density is shown in Figure 2.1. In view of Proposition 2.24, the densities displayed there represent all possible tail dependencies of GH distributions: $NIG_2(10, \mathbf{0}, 0.2, \mathbf{0}, \bar\Delta)$ and $NIG_2(10, \binom{4}{1}, 0.2, \mathbf{0}, \bar\Delta)$ are tail independent, $NIG_2(4, \binom{3}{-2}, 0.2, \mathbf{0}, \bar\Delta)$ is completely dependent, and the t distribution $t_2(-2, 2, \mathbf{0}, \bar\Delta)$ lies in between.

The fact that for GH distributions the coefficients of tail dependence can only take the most extreme values 0 and 1 may surely be surprising at first glance, but this phenomenon can also be observed in other distribution classes (making it possibly less astonishing). For example, Banachewicz and van der Vaart (2008) found a similar behaviour for the upper tail dependence coefficient $\lambda_u(X_1, X_2)$ of a skewed grouped t distribution. An alternative derivation and discussion of their results can also be found in Fung and Seneta (2010).

Thus the dependence structure of multivariate GH distributions is fairly strict in some sense since it neither allows independent components nor non-trivial values of the tail dependence coefficients. A possible way to relax these restrictions is to consider affine mappings of random vectors with independent GH distributed components: If $Y \overset{d}{=} AX + \mu$, where $\mu \in \mathbb{R}^d$, A is a lower triangular $d \times d$-matrix, and $X = (X_1, \dots, X_d)^\top$ with independent $X_i \sim GH(\lambda_i, \alpha_i, \beta_i, 1, 0)$, $1 \le i \le d$, then Y is said to have a *multivariate affine GH distribution*. Dependent on the choice of A, $\mathcal{L}(Y)$ can either possess independent margins or show upper and lower tail dependence. Schmidt, Hrycej, and Stützle (2006) provide a thorough discussion of this model.

Chapter 3

Applications to credit portfolio modeling and CDO pricing

Credit risk represents by far the biggest risk in the activities of a traditional bank. In particular, during recession periods financial institutions loose enormous amounts as a consequence of bad loans and default events. Traditionally, the risk arising from a loan contract could not be transferred and remained in the books of the lending institution until maturity. This has changed completely since the introduction of credit derivatives such as credit default swaps (CDSs) and collateralized debt obligations (CDOs) roughly fifteen years ago. The volume in trading these products at the exchanges and directly between individual parties (OTC) has increased enormously. This success is due to the fact that credit derivatives allow the transfer of credit risk to a larger community of investors. The risk profile of a bank can now be shaped according to specified limits, and concentrations of risk caused by geographic and industry sector factors can be reduced.

However, credit derivatives are complex products, and a sound risk-management methodology based on appropriate quantitative models is needed to judge and control the risks involved in a portfolio of such instruments. Quantitative approaches are particularly important in order to understand the risks involved in portfolio products such as CDOs. Here we need mathematical models which allow to derive the statistical distribution of portfolio losses. This distribution is influenced by the default probabilities of the individual instruments in the portfolio, and, more importantly, by the *joint behaviour* of the components of the portfolio. Therefore the probabilistic dependence structure of default events has to be modeled appropriately.

In the present chapter, this will be achieved by an extension of the factor model approach which goes back to Vasiček (1987, 1991). He assumed the factors to be normal distributed which still is the industry standard up to now. We shall replace the normal distribution by the much more flexible class of GH distributions which have been introduced and extensively studied in the first chapter of this thesis. As will be shown, this approach leads to a substantial improvement of performance in the pricing of synthetic CDO tranches. The main results of this chapter can also be found in Eberlein, Frey, and v. Hammerstein (2008).

3.1 CDOs: Basic concepts and modeling approaches

A *collateralized debt obligation* (CDO) is a structured product based on an underlying portfolio of reference entities subject to credit risk, such as corporate bonds, mortgages, loans, or credit derivatives. Although several types of CDOs are traded in the market which mainly differ in the content of the portfolio and the cash flows between counterparties, the basic structure is the same. The originator (usually a bank) sells the assets of the portfolio to a so-called *special purpose vehicle* (SPV), a company which is set up only for the purpose of carrying out the securitization and the necessary transactions. The SPV does not need capital itself, instead it issues notes to finance the acquisition of the assets. Each note belongs to a certain loss piece or *tranche* after the portfolio has been divided into a number of them. Consequently, the portfolio is no longer regarded as an asset pool but as a collateral pool. The tranches have different seniorities: The first loss piece or *equity tranche* has the lowest, followed by *junior mezzanine*, *mezzanine*, *senior* and finally *super-senior* tranches. The interest payments the SPV has to make to the buyer of a CDO tranche are financed from the cash flow generated by the collateral pool. Therefore the performance and the default risk of the portfolio is taken over by the investors. Since all liabilities of the SPV as a tranche seller are funded by proceeds from the portfolio, CDOs can be regarded as a subclass of so-called asset-backed securities. If the assets consist mainly of bonds resp. loans, the CDO is also called collateralized bond obligation (CBO) resp. collateralized loan obligation (CLO). For a *synthetic CDO* which we shall discuss in greater detail below, the portfolio contains only credit default swaps. The motivation to build a CDO is given by economic reasons:

- By selling the assets to the SPV, the originator removes them from his balance sheet, and therefore he is able to reduce his regulatory capital. The capital which is set free can then be used for new business opportunities.
- The proceeds from the sale of the CDO tranches are typically higher than the initial value of the asset portfolio because the risk-return profile of the tranches is more attractive for investors. This is both the result from and the reason for slicing the portfolio into tranches and the implicit collation and rebalancing hereby. Arbitrage CDOs are mainly set up to exploit this difference.

In general, CDO contracts can be quite sophisticated because there are no regulations for the compilation of the reference portfolio and its tranching or the payments to be made between the parties. The originator and the SPV can design the contract in a taylormade way, depending on the purposes they want to achieve. To avoid unnecessary complications, we concentrate in the following on synthetic CDOs which are based on a portfolio of credit default swaps.

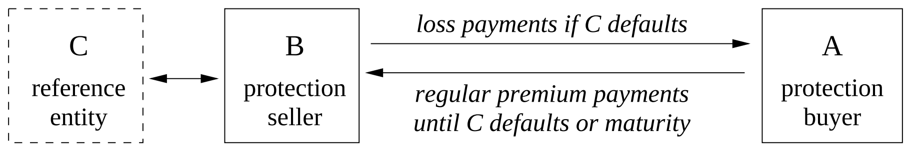


Figure 3.1: Basic structure of a CDS

3.1.1 Structure and payoffs of CDSs and synthetic CDOs

As mentioned before, the reference portfolio of a synthetic CDO consists entirely of *credit default swaps* (CDSs). These are insurance contracts protecting from losses caused by default of defaultable assets. The protection buyer A periodically pays a fixed premium to the protection seller B until a prespecified credit event occurs or the contract terminates. In turn, B makes a payment to A that covers his losses if the credit event has happened during the lifetime of the contract. Since there are many possibilities to specify the default event as well as the default payment, different types of CDSs are traded in the market, depending on the terms the counterparties have agreed on. The basic structure is shown in Figure 3.1. Throughout this chapter we will make the following assumptions: The reference entity of the CDS is a defaultable bond with nominal value L, and the credit event is the default of the bond issuer. If default has happened, B pays $(1-R)L$ to A where R denotes the *recovery rate*. On the other side, A quarterly pays a fixed premium of $0.25 r_{CDS} L$ where r_{CDS} is the annualized fair CDS rate. To determine this rate explicitly, we fix some notation:

r is the riskless interest rate, assumed to be constant over the lifetime $[0,T]$ of the CDS,

$u(t)$ is the discounted value of all premiums paid up to time t when the annualized premium is standardized to 1,

$G_1(t)$ is the distribution function of the default time T_1 with corresponding density $g_1(t)$ (its existence will be justified by the assumptions in subsequent sections).

The expected value of the discounted premiums (*premium leg*) can then be written as

$$PL(r_{CDS}) = r_{CDS}\, L \int_0^T u(t) g_1(t)\,\mathrm{d}t + r_{CDS}\, L\, u(T)\big(1 - G_1(T)\big)\,.$$

The expected discounted default payment (*default leg*) is given by

$$D = (1-R)L \int_0^T g_1(t)\, e^{-rt}\,\mathrm{d}t\,.$$

The no-arbitrage condition $PL(r_{CDS}) = D$ then implies

$$r_{CDS} = \frac{(1-R)\int_0^T g_1(t)\, e^{-rt}\,\mathrm{d}t}{\int_0^T u(t) g_1(t)\,\mathrm{d}t + u(T)\big(1-G_1(T)\big)} = \frac{D}{PL(1)}\,. \tag{3.1}$$

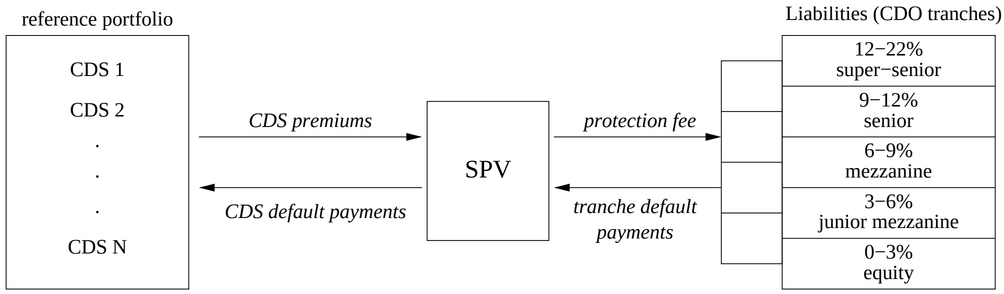


Figure 3.2: Schematic representation of the payments in a synthetic CDO. The choice of the attachment points corresponds to DJ iTraxx Europe standard tranches.

To explain the structure and the cash flows of a synthetic CDO, assume that its reference portfolio consists of N different CDSs with the same notional value L. We divide this portfolio in subsequent tranches. Each tranche covers a certain range of the percentage losses of the total portfolio value NL defined by lower and upper *attachment points* $0 \leq K_l, K_u \leq 1$. The buyer of a tranche compensates as protection seller for all losses that exceed the amount of $K_l NL$ up to a maximum of $K_u NL$. On the other hand, the SPV as protection buyer has to make quarterly payments of $0.25 r_c V_t$, where V_t is the notional value of the tranche at payment date t. Note that V_t starts with $NL(K_u - K_l)$ and is reduced by every default that hits the tranche. r_c is the fair tranche rate. See also Figure 3.2.

In recent years a new and simplified way of buying and selling CDO tranches has become very popular, the trading of *single index tranches*. For this purpose standardized portfolios and tranches are defined. Two counterparties can agree to buy and sell protection on an individual tranche and exchange the cash flows shown in the right half of Figure 3.2. The underlying CDS portfolio, however, is never physically created, it is merely a reference portfolio from which the cash flows are derived. So the left hand side of Figure 3.2 vanishes in this case, and the SPV is replaced by the protection buyer. The portfolios for the two most traded indices, the Dow Jones CDX NA IG and the Dow Jones iTraxx Europe, are composed of 125 investment grade US and European firms, respectively. The index itself is nothing but the weighted credit default swap spread of the reference portfolio. In Sections 3.1.2 and 3.2, we shall derive the corresponding default probabilities. We will use market quotes for different iTraxx tranches and maturities to calibrate our models later in Section 3.2.2.

In the following we denote the attachment points by $0 = K_0 < K_1 < \cdots < K_m \leq 1$ such that the lower and upper attachment points of tranche i are K_{i-1} and K_i, respectively. Suppose, for example, that $(1-R)j = K_{i-1}N$ and $(1-R)k = K_i N$ for some $j < k,\ j, k \in \mathbb{N}$. Then the protection seller B of tranche i pays $(1-R)L$ if the $(j+1)^{st}$ reference entity in the portfolio defaults. For each of the following possible $k - j - 1$ defaults, the protection buyer receives the same amount from B. After the k^{th} default occurred, the outstanding notional of the tranche is zero and the contract terminates. However, the losses will usually not match the attachment points. In general, some of them are divided up between subsequent tranches: If $\frac{(j-1)(1-R)}{N} < K_i < \frac{j(1-R)}{N}$ for some $j \in \mathbb{N}$, then tranche i bears a loss of $NL\big(K_i - \frac{(j-1)(1-R)}{N}\big)$ (and is exhausted thereafter) if the j^{th} default occurs. The overshoot is absorbed by the following tranche whose outstanding notional is reduced by $NL\big(\frac{j(1-R)}{N} - K_i\big)$. We use the following notation:

K_{i-1}, K_i are the lower/upper attachment points of tranche i,

Z_t is the relative amount of CDSs which have defaulted up to time t, expressed as a fraction of the total number N,

$L_t^i = \min[(1-R)Z_t, K_i] - \min[(1-R)Z_t, K_{i-1}]$ is the loss of tranche i up to time t, expressed as a fraction of the total notional value NL,

r_i is the fair spread rate of tranche i,

$0 = t_0 < \cdots < t_n$ are the payment dates of protection buyer and seller,

$\beta(t_0, t_k)$ is the discount factor for time t_k.

Remark: Under the assumption of a constant riskless interest rate r we would have $\beta(t_0, t_k) = e^{-rt_k}$. Since this assumption is too restrictive one uses zero coupon bond prices for discounting instead. Therefore $\beta(t_0, t_k)$ will denote the price of a zero coupon bond with maturity t_k at time t_0.

The assumption that all CDSs have the same notional value may seem somewhat artificial, but it is fulfilled for CDOs on standardized portfolios like the Dow Jones CDX or the iTraxx Europe.

With the above notation, the premium leg as well as the default leg of tranche i can be expressed as

$$\begin{aligned} PL_i(r_i) &= \sum_{k=1}^{n} (t_k - t_{k-1}) \beta(t_0, t_k)\, r_i\, \mathrm{E}\big[\big(K_i - K_{i-1} - L_{t_k}^i\big) NL\big], \\ D_i &= \sum_{k=1}^{n} \beta(t_0, t_k) \mathrm{E}\big[\big(L_{t_k}^i - L_{t_{k-1}}^i\big) NL\big], \end{aligned} \tag{3.2}$$

where $\mathrm{E}[\cdot]$ denotes expectation. For the fair spread rate one obtains

$$r_i = \frac{\sum_{k=1}^{n} \beta(t_0, t_k)\big(\mathrm{E}\big[L_{t_k}^i\big] - \mathrm{E}\big[L_{t_{k-1}}^i\big]\big)}{\sum_{k=1}^{n} (t_k - t_{k-1}) \beta(t_0, t_k)\big(K_i - K_{i-1} - \mathrm{E}\big[L_{t_k}^i\big]\big)}\,. \tag{3.3}$$

Remark: To get arbitrage-free prices, all expectations above have to be taken under a risk neutral probability measure, which is assumed implicitly. One should be aware that risk neutral probabilities cannot be estimated from historical default data.

Since payment dates and attachment points are specified in the CDO contract and discount factors can be obtained from the market, the remaining task is to develop a realistic portfolio model from which the risk neutral distribution of Z_t can be derived, that is, we need to model the joint distribution of the default times $T_1, \ldots, T_N$ of the reference entities.

3.1.2 Factor models with normal distributions

To construct this joint distribution, the first step is to define the marginal distributions $Q_i(t) = P(T_i \leq t)$. The standard approach, which was proposed in Li (2000), is to assume that the default times T_i are exponential distributed, that is, $Q_i(t) = 1 - e^{-\lambda_i t}$. The *default intensities* λ_i can be estimated from the clean spreads $\frac{r^i_{CDS}}{1-R}$ where r^i_{CDS} is the fair CDS spread of firm i which can be derived using the formula (3.1). In fact, the relationship $\lambda_i \approx \frac{r^i_{CDS}}{1-R}$ is obtained directly from (3.1) by inserting the default density $g_1(t) = \lambda_i e^{-\lambda_i t}$ (see McNeil, Frey, and Embrechts 2005, Chapter 9.3.3).

As mentioned before, the CDX and iTraxx indices quote an average CDS spread for the whole portfolio in basis points (100 bp = 1%), therefore the market convention is to set

$$\lambda_i \equiv \lambda_a = \frac{s_a}{(1-R)10000}, \tag{3.4}$$

where s_a is the average CDX or iTraxx spread in basis points. This implies that all firms in the portfolio have the same default probability. One can criticize this assumption from a theoretical point of view, but it simplifies and fastens the calculation of the loss distribution considerably as we will see below. Since λ_a is obtained from data of derivative markets, it can be considered as a risk neutral parameter, and therefore the $Q_i(t)$ can be regarded as risk neutral probability distributions as well.

The second step to obtain the joint distribution of the default times is to impose a suitable coupling between the marginals. Since all firms are subject to the same economic environment and many of them are linked by direct business relations, the assumption of independence of defaults between different firms obviously is not realistic. The empirically observed occurrence of disproportionally many defaults in certain time periods also contradicts the independence assumption. Therefore the main task in credit portfolio modeling is to implement a realistic dependence structure which generates loss distributions that are consistent with market observations. The following approach goes back to Vasiček (1987) and was motivated by the model of Merton (1974).

For each CDS in the CDO portfolio, we define a random variable X_i as follows:

$$X_i := \sqrt{\rho}\, M + \sqrt{1-\rho}\, Z_i, \qquad 0 \leq \rho < 1, \quad i = 1, \ldots, N, \tag{3.5}$$

where $M, Z_1, \ldots, Z_N$ are independent and standard normal distributed. Obviously $X_i \sim N(0,1)$ and $\mathrm{Corr}(X_i, X_j) = \rho$, $i \neq j$. X_i can be interpreted as state variable for the firm that issued the bond which CDS number i secures. The state is driven by two factors: the *systematic factor* M represents the macroeconomic environment to which all firms are exposed, whereas the *idiosyncratic factor* Z_i incorporates firm specific strengths or weaknesses.

To model the individual defaults, we define time-dependent thresholds by

$$d_i(t) := \Phi^{-1}\big(Q_i(t)\big)$$

where $\Phi^{-1}(x)$ denotes the inverse of the standard normal distribution function resp. the quantile function of $N(0,1)$. Observe that the $d_i(t)$ are increasing because so are Φ^{-1} and Q_i. Therefore we can define each default time T_i as the first point in time at which the corresponding variable X_i is smaller than the threshold $d_i(t)$, that is,

$$T_i := \inf\{t \geq 0 \,|\, X_i \leq d_i(t)\}, \qquad i = 1, \ldots, N. \tag{3.6}$$

This also ensures that the T_i have the desired distribution, because

$$P(T_i \leq t) = P\big(X_i \leq \Phi^{-1}\big(Q_i(t)\big)\big) = P\big(\Phi(X_i) \leq Q_i(t)\big) = Q_i(t),$$

where the last equation follows from the fact that the random variable $\Phi(X_i)$ is uniformly distributed on the interval $[0,1]$. Moreover, the leftmost equation shows that $T_i \stackrel{d}{=} Q_i^{-1}\big(\Phi(X_i)\big)$, so the default times inherit the dependence structure of the X_i. Since the latter are not observable, but serve only as auxiliary variables to construct dependencies, such models are also termed "latent variable" models. Note that by (3.4) we have $Q_i(t) \equiv Q(t)$ and thus $d_i(t) \equiv d(t)$, therefore we omit the index i in the following.

Remark: Instead of inducing dependence by latent variables that are linked by the factor equation (3.5), one can also define the dependence structure of the default times more directly by inserting the marginal distribution functions into an appropriately chosen copula (see Sklar's Theorem 2.19 and the subsequent discussion in Chapter 2.4). We do not discuss this approach here further, but give some references at the end of Section 3.1.3.

To derive the loss distribution, let A_k^t be the event that exactly k defaults have happened up to time t. From equations (3.6) and (3.5) we get

$$P(T_i < t \,|\, M) = P(X_i < d(t) \,|\, M) = \Phi\left(\frac{d(t) - \sqrt{\rho}M}{\sqrt{1-\rho}}\right).$$

Since the X_i are independent conditional on M, the conditional probability $P(A_k^t \,|\, M)$ equals the probability of a binomial distribution with parameters N and $p = P(T_i < t \,|\, M)$:

$$P(A_k^t \,|\, M) = \binom{N}{k} \Phi\left(\frac{d(t) - \sqrt{\rho}M}{\sqrt{1-\rho}}\right)^k \left(1 - \Phi\left(\frac{d(t) - \sqrt{\rho}M}{\sqrt{1-\rho}}\right)\right)^{N-k}.$$

The probability that at time t the relative number of defaults Z_t does not exceed

q is

$$
\begin{aligned}
F_{Z_t}(q) &= \sum_{k=0}^{[Nq]} P(A_k^t) \\
&= \int_{-\infty}^{\infty} \sum_{k=0}^{[Nq]} \binom{N}{k} \Phi\left(\frac{d(t)-\sqrt{\rho}y}{\sqrt{1-\rho}}\right)^k \left(1-\Phi\left(\frac{d(t)-\sqrt{\rho}y}{\sqrt{1-\rho}}\right)\right)^{N-k} F_M(\mathrm{d}y)\,.
\end{aligned}
$$

If the portfolio is very large, one can simplify F_{Z_t} further using the following approximation which was introduced in Vasiček (1991) and is known as *large homogeneous portfolio (LHP) approximation.* Let $p_t(M) := \Phi\Big(\frac{d(t)-\sqrt{\rho}M}{\sqrt{1-\rho}}\Big)$ and G_{p_t} be the corresponding distribution function, then we can rewrite F_{Z_t} in the following way:

$$
F_{Z_t}(q) = \int_0^1 \sum_{k=0}^{[Nq]} \binom{N}{k} s^k (1-s)^{N-k}\, G_{p_t}(\mathrm{d}s). \tag{3.7}
$$

Applying the LHP approximation means that we have to determine the behaviour of the integrand for $N \to \infty$. For this purpose, suppose that Y_i are independent and identically distributed Bernoulli variables with $P(Y_i = 1) = s = 1 - P(Y_i = 0)$. Then the strong law of large numbers states that $\bar{Y}_N = \frac{1}{N}\sum_{i=1}^N Y_i \to s$ almost surely which implies convergence of the distribution functions $F_{\bar{Y}_N}(x) \to \mathbb{1}_{[0,x]}(s)$ pointwise on $\mathbb{R} \setminus \{s\}$. For all $q \neq s$ we thus have

$$
\sum_{k=0}^{[Nq]} \binom{N}{k} s^k (1-s)^{N-k} = P\left(\sum_{i=1}^N Y_i \le Nq\right) = P\big(\bar{Y}_N \le q\big) \underset{N\to\infty}{\longrightarrow} \mathbb{1}_{[0,q]}(s).
$$

Since the sum on the left hand side is bounded by 1, we can apply the dominated convergence theorem to (3.7) and obtain

$$
\begin{aligned}
F_{Z_t}(q) &\approx \int_0^1 \mathbb{1}_{[0,q]}(s)\,\mathrm{d}G_{p_t}(s) = G_{p_t}(q) = P\left(-\frac{\sqrt{1-\rho}\,\Phi^{-1}(q)-d(t)}{\sqrt{\rho}} \le M\right) \\
&= \Phi\left(\frac{\sqrt{1-\rho}\,\Phi^{-1}(q)-d(t)}{\sqrt{\rho}}\right)
\end{aligned} \tag{3.8}
$$

where in the last equation the symmetry relation $1-\Phi(x) = \Phi(-x)$ has been used. This distribution is, together with the above assumptions, the current market standard for the calculation of CDO spreads according to equation (3.3). Since the relative portfolio loss up to time t is given by $(1-R)Z_t$, the expectations $\mathrm{E}\big[L_{t_k}^i\big]$ within (3.3) can be written as follows:

$$
\mathrm{E}\big[L_{t_k}^i\big] = \int_{\frac{K_{i-1}}{1-R}\wedge 1}^{\frac{K_i}{1-R}\wedge 1} (1-R)\big(q - \tfrac{K_{i-1}}{1-R}\big) F_{Z_{t_k}}(\mathrm{d}q) + (K_i - K_{i-1})\big[1 - F_{Z_{t_k}}\big(\tfrac{K_i}{1-R}\wedge 1\big)\big]. \tag{3.9}
$$

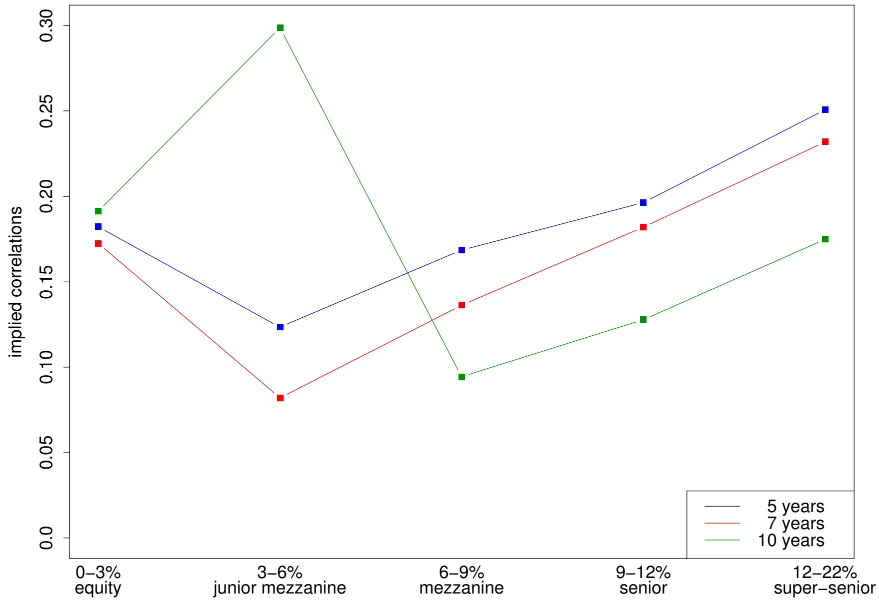


Figure 3.3: Implied correlations calculated from the prices of DJ iTraxx Europe standard tranches at November 13, 2006, for different maturities T.

3.1.3 Deficiencies and extensions

The pricing formula obtained from (3.3), (3.8), and (3.9) contains one unknown quantity: the correlation parameter ρ. This parameter has to be estimated before one can calculate the fair rate of a CDO tranche. A priori it is not clear which data and which estimation procedure one could use to get ρ. In the Merton approach, defaults are driven by the evolution of the asset value of a firm. Consequently, the dependence between defaults is derived from the dependence between asset values. The latter cannot be observed directly, therefore some practitioners have used equity correlations which can be estimated from stock price data. A more direct and plausible alternative would be to infer correlations from historical default data, but since default typically is a rare event, this would require data sets over very long time periods which are usually not available.

With the development of a liquid market for single index tranches in the last years, a new source of correlation information has arisen: the *implied correlations* from index tranche prices. Similar to the determination of implied volatilities from option prices by inverting the Black–Scholes formula, one can invert the above pricing formula and solve numerically for the correlation parameter ρ which reproduces the quoted market price. This also provides a method to examine if the model and its assumptions are appropriate. If this is the case, the correlations derived from market prices of different tranches of the same index should coincide. However, in reality one observes a so-called *correlation smile*: the implied correlations of the equity and (super-)senior tranches are typically much higher than those of the mezzanine tranches. An example of this stylized feature is shown in Figure 3.3.

However, one should observe that in general implied correlations are uniquely determined for the equity tranche only. For higher tranches, it can happen that there exist two different solutions, or even none at all, which yield the observed market spread. To circumvent this problem, market practitioners have developed the concept of *base correlations*. To explain this, let us take a look back to equation (3.2): If payment dates, discount factors, and the total notional value are given and fixed, the premium leg and the default leg of each tranche can be regarded as functions that depend on the correlation, the market spread, and the corresponding attachment points, that is, $PL_i = PL_i(\rho, r_i, K_{i-1}, K_i)$ and $D_i = D_i(\rho, K_{i-1}, K_i)$. Note that the dependence on ρ stems from the expectations $\mathrm{E}\big[L^i_{t_k}\big]$, see also equations (3.8) and (3.9). The no-arbitrage condition can then be written in the following form: $0 = PL_i(\rho, r_i, K_{i-1}, K_i) - D_i(\rho, K_{i-1}, K_i)$. Inserting the attachment points and the market spread, this becomes a defining equation for ρ, and the implied correlation $\hat{\rho}_i$ is nothing but a root of it.

The central idea in the definition of base correlations $\tilde{\rho}_i$ is that investing in a tranche having the attachment points K_{i-1}, K_i is equivalent to being short in a tranche with attachment points $0, K_{i-1}$ and being long in a tranche with attachment points $0, K_i$. Thus we may reformulate the no-arbitrage condition as follows:

$$\begin{aligned} 0 &= PL_i(\rho, r_i, K_{i-1}, K_i) - D_i(\rho, K_{i-1}, K_i) \\ &= \big[PL(\rho, r_i, 0, K_i) - D(\rho, 0, K_i)\big] - \big[PL(\tilde{\rho}_{i-1}, r_i, 0, K_{i-1}) - D(\tilde{\rho}_{i-1}, 0, K_{i-1})\big] \end{aligned}$$

Inserting attachment points and market spreads into the second equation, it can recursively be solved to get the values $\tilde{\rho}_i$. For $i = 1$, the second term in square brackets vanishes since $K_0 = 0$, and the equation becomes $0 = PL(\rho, r_1, 0, K_1) - D(\rho, 0, K_1)$ which coincides with the defining one for $\hat{\rho}_1$. Hence the solution is $\tilde{\rho}_1 = \hat{\rho}_1$, that is, in case of the equity tranche base correlation and implied correlation are the same. $\tilde{\rho}_2$ then is obtained as the root of

$$0 = \big[PL(\rho, r_2, 0, K_2) - D(\rho, 0, K_2)\big] - \big[PL(\tilde{\rho}_1, r_2, 0, K_1) - D(\tilde{\rho}_1, 0, K_1)\big],$$

and $\tilde{\rho}_i$, $i \geq 3$, can be calculated consecutively along the same lines. The advantage of this approach is that equations of the type $c_i = PL(\rho, r_i, 0, K_i) - D(\rho, 0, K_i)$ have exactly one solution (similar to the implied correlation of an equity tranche), thus the base correlations $\tilde{\rho}_i$ are uniquely determined. Therefore prices of CDO tranches can alternatively be expressed in terms of base correlations which in fact many market participants do. Figure 3.4 shows the base correlations corresponding to the iTraxx quotes of November 13, 2006. (These are recomputed as described above because our dataset only contains the different spreads r_i in basis points.) Despite their advantages in practice, one should be aware that base correlations do not remedy any weakness of the model. The model imperfections here express themselves by the increase of $\tilde{\rho}_i$ from tranche to tranche. In a perfect model, $\tilde{\rho}_i$ should be constant for all i (and hence unnecessary).

The correlation smile as well as the inconstancy of the base correlations indicate that the classical model is not flexible enough to generate realistic dependence structures. This is only partly due to the simplifications made by

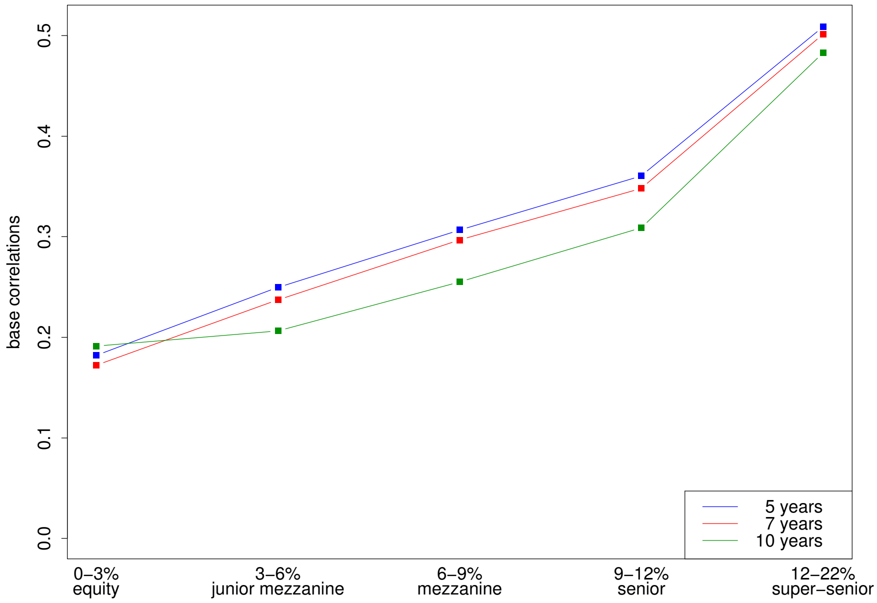


Figure 3.4: Base correlations calculated from the prices of DJ iTraxx Europe standard tranches at November 13, 2006, for different maturities T.

using the LHP approach. The deeper reason for this phenomenon lies in the fact that the model with normal factors strongly underestimates the probabilities of joint defaults. This has led to severe mispricings and inadequate risk forecasts in the past. The problem became evident in the so-called *correlation crisis* in May 2005: the factor model based on normal distributions was unable to follow the movement of market quotes occuring in reaction to the downgrading of Ford and General Motors to non-investment grade.

A number of different approaches for dealing with this problem have been investigated. A rather intuitive extension to remedy the deficiencies of the normal factor model which we shall exploit in Section 3.2 is to allow for factor distributions which are much more flexible than the standard normal ones. Different factor distributions do not only change the shape of F_{Z_t}, but also have a great influence on the copula implicitly contained in the joint distribution of the latent variables. In fact, the replacement of the normal distribution leads to a fundamental modification of the dependence structure which becomes much more complex and can even exhibit tail dependence. The first paper in which alternative factor distributions are used is Hull and White (2004) where both factors are assumed to follow a Student's t-distribution with 5 degrees of freedom. In Kalemanova, Schmid, and Werner (2007), normal inverse Gaussian distributions are applied for pricing synthetic CDOs, and in Albrecher, Ladoucette, and Schoutens (2007) several models based on Gamma, inverse Gaussian, Variance-Gamma, normal inverse Gaussian and Meixner distributions are presented. In the last paper the systematic and idiosyncratic factors are represented by the values of a suitably scaled and shifted Lévy process at times ρ and $1 - \rho$.

Another way to extend the classical model is to implement *stochastic corre-*

lations and *random factor loadings*. In the first approach, which was developed in Gregory and Laurent (2004), the constant correlation parameter ρ in (3.5) is replaced by a random variable taking values in $[0,1]$. The cumulative default distribution can then be derived similarly as before, but one has to condition on both, the systematic factor and the correlation variable. The concept of random factor loadings was first published in Andersen and Sidenius (2005). There the X_i are defined by $X_i := m_i(M) + \sigma_i(M)Z_i$ with some deterministic functions m_i and σ_i. In the simplest case, $X_i = m + (l\mathbb{1}_{\{M<e\}} + h\mathbb{1}_{\{M\geq e\}})M + \nu Z_i$ where $l, h, e \in \mathbb{R}$ are additional parameters and m, ν are constants chosen such that $\mathrm{E}[X_i] = 0$ and $\mathrm{Var}[X_i] = 1$. Further information and numerical details for the calibration of such models to market data can be found in Burtschell, Gregory, and Laurent (2007).

As already mentioned in the remark on p. 111, other approaches use copula models to define the dependencies between the default times T_i. The first papers where copulas were used in credit risk models are Li (2000) and Schönbucher and Schubert (2001). A more recent approach based on Archimedean copulas can be found in Berrada, Dupuis, Jacquier, Papageorgiou, and Rémillard (2006). The pricing performance of models with Clayton and Marshall–Olkin copulas was investigated and compared with some other popular approaches in Burtschell, Gregory, and Laurent (2005). There the prices calculated from the Clayton copula model showed a slightly better fit to the market quotes, but they were still relatively close to those generated by the Gaussian model. The Marshall–Olkin copulas performed worse, since the deviations from market prices were greater than those of other models considered.

3.2 Calibration with GH distributions

As outlined above, we want to overcome the deficiencies of the standard model by using more advanced and flexible distributions. The implementation of alternative factor distributions not only provides additional parameters for a more precise calibration, but also has a significant impact on the dependence structure of the default times as the following considerations show. Recall that the general factor model is given by

$$X_i := \sqrt{\rho}\, M + \sqrt{1-\rho}\, Z_i, \qquad 0 \leq \rho < 1, \quad i = 1, \dots, N, \tag{3.10}$$

where $M, Z_1, \dots, Z_N$ are assumed to be independent and, in addition, the Z_i are identically distributed (hence so are the X_i). The corresponding distribution functions are denoted by F_M, F_Z, F_X and are supposed to be continuous and strictly increasing on $\mathbb{R}$. Analogously to (3.6), the default times T_i are defined by

$$T_i := \inf\bigl\{t \geq 0 \,\big|\, X_i \leq F_X^{-1}\bigl(1 - e^{-\lambda_a t}\bigr)\bigr\}, \qquad i = 1, \dots, N,$$

with λ_a from equation (3.4), thus $P(T_i \leq t) = 1 - e^{-\lambda_a t} =: Q(t)$. Since F_X is continuous and strictly increasing, the random variables $F_X(X_i)$ are uniformly distributed on $(0,1)$. Further, let G_X and G_U denote the distribution functions of $(X_1, \dots, X_N)^\top$ and $U := (F_X(X_1), \dots, F_X(X_N))^\top$, respectively, then G_U is a copula according to Definition 2.18 which coincides with C_{G_X} by Proposition 2.20. Consequently

$$\begin{aligned} F(t_1,\dots,t_N) &:= P(T_1 \le t_1,\dots,T_N \le t_N) \\ &= P\big(F_X(X_1) \le Q(t_1),\dots,F_X(X_N) \le Q(t_N)\big) \\ &= C_{G_X}\big(Q(t_1),\dots,Q(t_N)\big), \end{aligned}$$

and hence $C_F = C_{G_X}$, that is, the implied copulas of the joint distributions of the X_i and the default times T_i are the same. (This can also be regarded as a more rigorous mathematical formulation of the assertion that the T_i inherit the dependence structure of the X_i, see p. 111.) Moreover, the conditional independence of the X_i given M implies

$$\begin{aligned} C_F(u_1,\dots,u_N) &= C_{G_X}(u_1,\dots,u_N) = G_X\big(F_X^{-1}(u_1),\dots,F_X^{-1}(u_N)\big) \\ &= \mathrm{E}\big[P\big(X_1 \le F_X^{-1}(u_1),\dots,X_N \le F_X^{-1}(u_N) \,|\, M\big)\big] \\ &= \int_{\mathbb{R}} \prod_{i=1}^{N} F_Z\left(\frac{F_X^{-1}(u_i) - \sqrt{\rho}\,y}{\sqrt{1-\rho}}\right) F_M(\mathrm{d}y) \end{aligned}$$

which shows the direct and predominant influence of the factor distributions F_M and F_Z on the dependencies of the default times. In particular, the distributions F and G_X are tail dependent ($\lambda_u(T_i,T_j) = \lambda_u(X_i,X_j) > 0$, $1 \le i \ne j \le N$) if and only if the systematic factor M is heavy tailed, that is, $F_M \in \mathscr{R}_p$ for some $-\infty < p < 0$ (see Definition 1.15 in Chapter 1.3). This was proven by Malevergne and Sornette (2004).

Remark: By the above equations, the factor model may also be regarded as a special case of the more general approach to couple the individual default times by a suitably chosen copula. Because of its particular structure, C_F is sometimes called *factor copula.*

We therefore suppose, in addition to the aforementioned assumptions, that the factor distributions are given by $M \sim GH(\lambda_M,\alpha_M,\beta_M,\delta_M,\mu_M)$ and $Z_i \sim GH(\lambda_Z,\alpha_Z,\beta_Z,\delta_Z,\mu_Z)$ for all $1 \le i \le N$. Applying the LHP approximation and letting $N \to \infty$, the cumulative default distribution F_{Z_t} can be derived analogously as described at the end of Section 3.1.2. One obtains

$$F_{Z_t}(q) \approx 1 - F_M\left(\frac{F_X^{-1}(Q(t)) - \sqrt{1-\rho}\,F_Z^{-1}(q)}{\sqrt{\rho}}\right). \tag{3.11}$$

Note that this expression cannot be simplified further as in equation (3.8) because the distribution of M is in general not symmetric (symmetry only holds if $\beta_M = 0$). For the calibration of the model to the iTraxx data later on, we restrict ourselves to some specific subclasses and limiting cases of the generalized hyperbolic class. We shall use normal inverse Gaussian and hyperbolic distributions, whose characteristic functions are

$$\phi_{NIG(\alpha,\beta,\delta,\mu)}(u) = e^{iu\mu}\, e^{\delta(\sqrt{\alpha^2-\beta^2}-\sqrt{\alpha^2-(\beta+iu)^2})} \tag{3.12}$$

and

$$\phi_{HYP(\alpha,\beta,\delta,\mu)}(u) = e^{iu\mu}\left(\frac{\alpha^2-\beta^2}{\alpha^2-(\beta+iu)^2}\right)^{\frac{1}{2}} \frac{K_1\big(\delta\sqrt{\alpha^2-(\beta+iu)^2}\big)}{K_1\big(\delta\sqrt{\alpha^2-\beta^2}\big)}. \tag{3.13}$$

The corresponding probability densities can be found in Chapter 1.3 on page 14. Further, we shall apply Variance-Gamma and t distributions which have the characteristic funtions

$$\phi_{VG(\lambda,\alpha,\beta,\mu)}(u) = e^{iu\mu}\left(\frac{\alpha^2-\beta^2}{\alpha^2-(\beta+iu)^2}\right)^{\lambda}, \tag{3.14}$$

$$\phi_{t(\lambda,\delta,\mu)}(u) = e^{iu\mu}\left(\frac{2}{\delta}\right)^{\lambda}\frac{2K_\lambda(\delta|u|)}{\Gamma(-\lambda)|u|^\lambda}. \tag{3.15}$$

The appropriate densities are given in Chapter 1.4.1 on pages 21 and 22.

Remark: As shown in Chapters 1.3 and 1.4.1, the densities and distribution functions of almost all GH distributions possess exponentially decreasing tails, only the t- and skew Student t limit distributions (see equation (1.20)) have a power tail. According to the already alluded results of Malevergne and Sornette (2004), the joint distributions F and G_X of the default times and the latent variables X_i therefore show (upper) tail dependence if and only if the systematic factor M is t- or skew Student t-distributed.

Moreover, by equation (1.21) in Chapter 1.4.2 we have $GH(\lambda,\alpha,\beta,\delta,\mu) \xrightarrow{w} N(\mu+\beta\sigma^2,\sigma^2)$ if $\alpha,\delta\to\infty$ and $\frac{\delta}{\alpha}\to\sigma^2$, thus the normal factor model is included as a limit in our setting.

3.2.1 Factor scaling and calculation of quantiles

To preserve the role of ρ as a correlation parameter, we have to standardize the factor distributions such that they have zero mean and unit variance. In the general case of GH distributions we fix shape, skewness and tail behaviour by specifying α,β,λ, and then calculate $\bar\delta$ and $\bar\mu$ that scale and shift the density appropriately. For this purpose we first solve the equation

$$1 = \mathrm{Var}[GH(\lambda,\alpha,\beta,\delta,\mu)] = \frac{\delta^2}{\zeta}\frac{K_{\lambda+1}(\zeta)}{K_\lambda(\zeta)} + \beta^2\frac{\delta^4}{\zeta^2}\left(\frac{K_{\lambda+2}(\zeta)}{K_\lambda(\zeta)} - \frac{K^2_{\lambda+1}(\zeta)}{K^2_\lambda(\zeta)}\right)$$

with $\zeta := \delta\sqrt{\alpha^2-\beta^2}$ numerically to obtain $\bar\delta$, and then determine $\bar\mu$ that fulfills

$$0 = \mathrm{E}[GH(\lambda,\alpha,\beta,\bar\delta,\bar\mu)] = \bar\mu + \frac{\beta\bar\delta^2}{\bar\zeta}\frac{K_{\lambda+1}(\bar\zeta)}{K_\lambda(\bar\zeta)}, \quad \bar\zeta := \bar\delta\sqrt{\alpha^2-\beta^2}.$$

Since the Bessel functions $K_{n+1/2}$, $n\ge 0$, can be expressed explicitly in closed forms (see equation (A.6) in Appendix A), the calculations simplify considerably for the NIG subclass. There we have

$$\mathrm{Var}[NIG(\alpha,\beta,\delta,\mu)] = \frac{\delta\alpha^2}{(\alpha^2-\beta^2)^{\frac{3}{2}}}, \quad \mathrm{E}[NIG(\alpha,\beta,\delta,\mu)] = \mu + \frac{\beta\delta}{\sqrt{\alpha^2-\beta^2}},$$

so the distribution can be standardized by choosing $\bar{\delta} = \frac{(\alpha^2-\beta^2)^{\frac{3}{2}}}{\alpha^2}$ and $\bar{\mu} = -\frac{\beta(\alpha^2-\beta^2)}{\alpha^2}$. In the VG limiting case the variance is given by

$$\mathrm{Var}[VG(\lambda,\alpha,\beta,\mu)] = \frac{2\lambda}{\alpha^2-\beta^2} + \frac{4\lambda\beta^2}{(\alpha^2-\beta^2)^2} =: \sigma_{VG}^2,$$

so it is tempting to use λ as a scaling parameter, but this would change the tail behaviour which we want to keep fixed. Recalling that, by Corollary 1.28, a VG distributed random variable $X \sim VG(\lambda,\alpha,\beta,\mu)$ admits the stochastic representation $X \stackrel{d}{=} X_1 - X_2 + \mu$ with $X_1 \sim G(\lambda,\alpha-\beta)$ and $X_2 \sim G(\lambda,\alpha+\beta)$, the correct scaling that preserves the shape is $\bar{\alpha} = \sigma_{VG}\,\alpha$, $\bar{\beta} = \sigma_{VG}\,\beta$. Then $\bar{\mu}$ has to fulfill

$$0 = \mathrm{E}[VG(\lambda,\bar{\alpha},\bar{\beta},\bar{\mu})] = \bar{\mu} + \frac{2\lambda\bar{\beta}}{\bar{\alpha}^2-\bar{\beta}^2}.$$

The second moment of a t distribution $t(\lambda,\delta,\mu)$ exists only if $\lambda < -1$ (confer page 23). With this constraint, mean and variance are given by

$$\mathrm{Var}[t(\lambda,\delta,\mu)] = \frac{\delta^2}{-2\lambda-2} \qquad \text{and} \quad \mathrm{E}[t(\lambda,\delta,\mu)] = \mu,$$

therefore one has to choose $\bar{\delta} = \sqrt{-2\lambda-2}$ and $\bar{\mu} = 0$.

We thus have a minimum number of three free parameters in our generalized factor model, namely λ_M, λ_Z, and ρ, if both M and Z_i are t-distributed, up to a maximum number of seven $(\lambda_M,\alpha_M,\beta_M,\lambda_Z,\alpha_Z,\beta_Z,\rho)$ if both factors are GH or VG distributed. If we restrict the distributions of M and Z_i to certain GH subclasses by fixing λ_M and λ_Z, five free parameters are remaining.

After the standardization of the factor distributions, the remaining problem is to compute the quantiles $F_X^{-1}(Q(t))$ which enter the default distribution F_{Z_t} according to equation (3.11). Since the class of GH distributions is in general not closed under convolutions as was pointed out in Chapter 1.3, the distribution function F_X is not known explicitly. Therefore the central task in the implementation of the model is to develop a fast and stable algorithm for the numerical calculation of the quantiles $F_X^{-1}(q)$, because simulation techniques have to be ruled out from the very beginning for two reasons: The default probabilities $Q(t)$ are very small, so one would have to generate a very large data set to get reasonably accurate quantile estimates, and the simulation would have to be restarted whenever at least one model parameter has been modified. Since the pricing formula is evaluated thousands of times with different parameters during the calibration procedure, this would be too time-consuming. Further, the routine used to calibrate the models tries to find an extremal point by searching the direction of the steepest ascend within the parameter space in each optimization step. This can only be done successfully if the model prices depend exclusively on the parameters and not additionally on random effects. In the latter case the optimizer may behave erratically and possibly will never reach an extremum.

Therefore we compute the quantiles via Fourier inversion. Let $\hat{P}_X$, $\hat{P}_M$, and $\hat{P}_Z$ denote the characteristic functions of X_i, M, and Z_i, then by equation (3.10) and the independence of the factors we have $\hat{P}_X(t) = \hat{P}_M(\sqrt{\rho}\,t)\hat{P}_Z(\sqrt{1-\rho}\,t)$. The inversion formula yields

$$F_X(y) = \lim_{a\to-\infty}\lim_{c\to\infty}\frac{1}{2\pi}\int_{-c}^{c}\frac{e^{-iat}-e^{-iyt}}{it}\,\hat{P}_M(\sqrt{\rho}\,t)\hat{P}_Z(\sqrt{1-\rho}\,t)\,\mathrm{d}t.$$

Thus we can approximate the distribution function F_X by choosing sufficiently small a and large c and evaluating the above integral numerically. The desired quantiles $F_X^{-1}(Q(t))$ are then derived by Newton's method. The accuracy can be adjusted by modifying a and c accordingly, which might also depend on the parameters of the factor distributions to improve the results. The characteristic functions we used for our calibrations are given explicitly in (3.12)–(3.15).

In contrast to this approach, there exist at least two special settings in which the quantiles $F_X^{-1}(Q(t))$ can be calculated directly. The first one relies on the convolution property of the NIG subclass,

$$NIG(\alpha,\beta,\delta_1,\mu_1) * NIG(\alpha,\beta,\delta_2,\mu_2) = NIG(\alpha,\beta,\delta_1+\delta_2,\mu_1+\mu_2),$$

and the fact that if $Y \sim NIG(\alpha,\beta,\delta,\mu)$, then $aY \sim NIG\big(\frac{\alpha}{|a|},\frac{\beta}{a},\delta|a|,\mu a\big)$ (see pp. 13–14). Thus if both M and Z_i follow an NIG distribution and the distribution parameters of the latter are defined by $\alpha_Z := \frac{\alpha_M\sqrt{1-\rho}}{\sqrt{\rho}}$ and $\beta_Z = \frac{\beta_M\sqrt{1-\rho}}{\sqrt{\rho}}$, then equation (3.10) implies that $X_i \sim NIG\big(\frac{\alpha_M}{\sqrt{\rho}},\frac{\beta_M}{\sqrt{\rho}},\frac{\bar{\delta}_M}{\sqrt{\rho}},\frac{\bar{\mu}_M}{\sqrt{\rho}}\big)$. Here $\bar{\delta}_M$ and $\bar{\mu}_M$ are the parameters of the standardized distribution of M as described before.

In the VG limiting case, the behaviour of the parameters α,β and μ is the same under scaling, and the corresponding convolution property is

$$VG(\lambda_1,\alpha,\beta,\mu_1) * VG(\lambda_2,\alpha,\beta,\mu_2) = VG(\lambda_1+\lambda_2,\alpha,\beta,\mu_1+\mu_2).$$

Consequently, if both factors are VG distributed and the free parameters of the idiosyncratic factor are chosen as $\lambda_Z = \frac{\lambda_M(1-\rho)}{\rho}$, $\alpha_Z = \alpha_M$, $\beta_Z = \beta_M$, then $X_i \sim VG\big(\frac{\lambda_M}{\rho},\frac{\bar{\alpha}_M}{\sqrt{\rho}},\frac{\bar{\beta}_M}{\sqrt{\rho}},\frac{\bar{\mu}_M}{\sqrt{\rho}}\big)$. The remaining two convolution formulas in (1.9) may be exploited similarly to obtain the distribution of X_i in closed form.

This stability under convolutions, together with the appropriate parameter choices for the idiosyncratic factor, was used in Kalemanova, Schmid, and Werner (2007) and all models considered in Albrecher, Ladoucette, and Schoutens (2007). We do not use this approach here because it reduces the number of free parameters and therefore the flexibility of the factor model. Moreover, in such a setting the distribution of the idiosyncratic factor is uniquely determined by the systematic factor, which contradicts the intuitive idea behind the factor model and lacks an economic interpretation.

3.2.2 Calibration results for the DJ iTraxx Europe

We calibrate our generalized factor model with market quotes of DJ iTraxx Europe standard tranches. As mentioned before, the iTraxx Europe index is based on a reference portfolio of 125 European investment grade firms and quotes its average credit spread which can be used to estimate the default intensity of all constituents according to equation (3.4). The diversification of the portfolio always remains the same. It contains CDSs of 10 firms from automotive industry, 30 consumers, 20 energy firms, 20 industrials, 20 TMTs (technology, media and telecommunication companies) and 25 financials. In each sector, the firms with the highest liquidity and volume of trade with respect to their defaultable assets (bonds and CDSs) are selected. The iTraxx portfolio is reviewed and updated quarterly. Not only companies that have defaulted in between are replaced by new ones, but also those which no longer fulfill the liquidity and trading demands. Of course, the recomposition affects future deals only. Once two counterparties have agreed to buy and sell protection on a certain iTraxx tranche, the current portfolio is kept fixed for them in order to determine the corresponding cash flows described in Section 3.1.1. The names and attachment points of the five iTraxx standard tranches are given in Figures 3.2, 3.3, and 3.4. For each of them, four contracts with different maturities (3, 5, 7 and 10 years) are available.

The settlement date of the sixth iTraxx series was December 20, 2006, so the 5, 7, and 10 year contracts mature on December 20, 2011, 2013, and 2016, respectively. We consider the market prices of the latter on all standard tranches at November 13, 2006. For the mezzanine and senior tranches, these are equal to the annualized fair spreads r_i which can be obtained from equation (3.3) and are also termed *running spreads*. However, the market convention for pricing the equity tranche is somewhat different: In this case the protection buyer has to pay a certain percentage s_1 of the notional value K_1NL as an *up-front fee* at the starting time t_0 of the contract and a fixed spread of 500 bp on the outstanding notional at $t_1,\dots,t_n$. Therefore the premium leg for the equity tranche is given by

$$PL_1(s_1) = s_1K_1NL + 0.05\sum_{k=1}^{n}(t_k - t_{k-1})\beta(t_0,t_k)\mathrm{E}\big[\big(K_1 - L^1_{t_k}\big)NL\big],$$

and the no-arbitrage condition $PL_1(s_1) = D_1$ then implies

$$s_1 = \frac{\sum_{k=1}^{n}\beta(t_0,t_k)\Big(\mathrm{E}\big[L^1_{t_k}\big] - \mathrm{E}\big[L^1_{t_{k-1}}\big] - 0.05(t_k - t_{k-1})\big(K_1 - \mathrm{E}\big[L^1_{t_k}\big]\big)\Big)}{K_1}. \quad (3.16)$$

Since the running spread is set to a constant of 500 bp, the varying market price quoted for the equity tranche is the percentage s_1 defining the magnitude of the up-front fee.

We calibrate our generalized factor model by least squares optimization, that is, we first specify to which subclass of the GH family the distributions F_M and F_Z belong, and then determine the correlation and distribution parameters numerically which minimize the sum of the squared differences between model and market prices over all tranches. Although our algorithm for computing the quantiles $F_X^{-1}(Q(t))$ allows us to combine factor distributions of different GH subclasses, we restrict both factors to the same subclass for simplicity reasons. Therefore in the following table and figures the expression VG, for example, denotes a factor model where M and the Z_i are Variance-Gamma distributed. The model prices are calculated from equations (3.3) and (3.16), using the

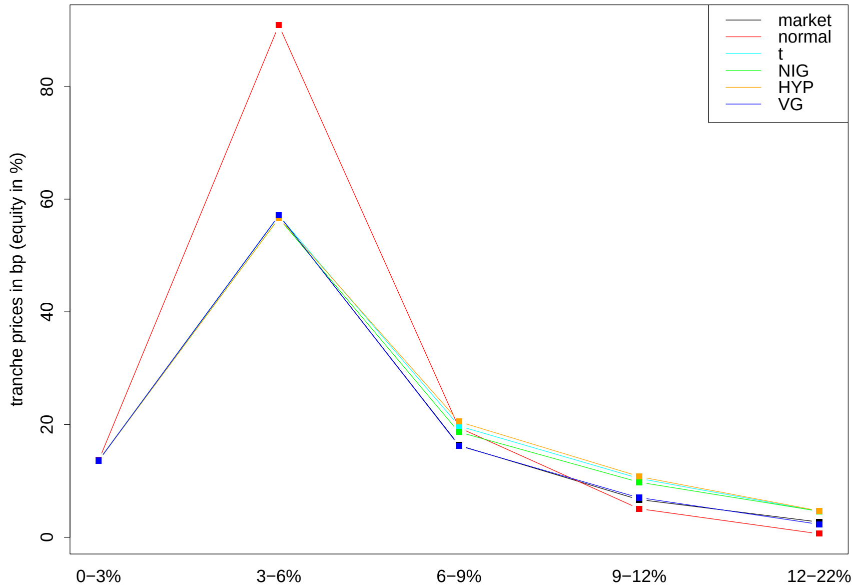


Figure 3.5: Comparison of calibrated model prices and market prices of the 5 year iTraxx contracts.

cumulative default distribution (3.11) resp. (3.8) for the normal factor model which serves as a benchmark. The recovery rate R which has a great influence on the expected losses $\mathrm{E}[L^i_{t_k}]$ according to equation (3.9) is always set to 40%; this is the common market assumption for the iTraxx portfolio.

One should observe that the prices of the equity tranches are usually given in percent, whereas the spreads of all other tranches are quoted in basis points. In order to use the same units for all tranches in the objective function to be minimized, the equity prices are transformed into basis points within the optimization algorithm. Thus they are much higher than the mezzanine and senior spreads and therefore react to parameter changes in a more sensitive way, which amounts to an increased weighting of the equity tranche in the calibration procedure. This is also desirable from an economical point of view since the costs for mispricing the equity tranche are typically greater than for all other tranches.

Remark: For the same reason, the normal factor model is usually calibrated by determining the implied correlation of the equity tranche first and then using this to calculate the fair spreads of the other tranches. This ensures that at least the equity price is matched perfectly. To provide a better comparison with our model, we give up this convention and also use least squares estimation in this case. Therefore the fit of the equity tranche is sometimes less accurate, but the distance between model and market prices is smaller for the higher tranches instead.

Our calibration results are summarized in Table 3.1. The normal benchmark model performs worst in all cases, which can also be seen from Figures 3.5 and 3.6. The performance of the t model is comparable with the NIG and HYP

Tranches	0–3 %	3–6 %	6–9 %	9–12 %	12–22 %	estimated parameters
	iTraxx 5Y S6 $(s_a = 24.88\,\text{bp})$					
Market	**13.60 %**	**57.16 bp**	**16.31 bp**	**6.65 bp**	**2.67 bp**	
Normal	13.64 %	90.93 bp	19.42 bp	5.03 bp	0.60 bp	$\rho = 0.181$
t	13.60 %	57.12 bp	19.75 bp	10.42 bp	4.59 bp	$\lambda_M = -1.982$, $\lambda_Z = -65.317$, $\rho = 0.133$
NIG	13.60 %	56.67 bp	18.66 bp	9.74 bp	4.60 bp	$\alpha_M = 5.683$, $\alpha_Z = 2.934$, $\beta_M = -0.174$, $\beta_Z = -2.599$, $\rho = 0.616$
HYP	13.60 %	56.67 bp	20.51 bp	10.76 bp	4.65 bp	$\alpha_M = 2.773$, $\alpha_Z = 2.320$, $\beta_M = -1.510$, $\beta_Z = -1.280$, $\rho = 0.290$
VG	13.60 %	57.16 bp	16.21 bp	7.00 bp	2.25 bp	$\lambda_M = 1.565$, $\lambda_Z = 2.118$, $\alpha_M = 4.112$, $\alpha_Z = 6.355$, $\beta_M = 1.415$, $\beta_Z = -2.177$, $\rho = 0.444$
	iTraxx 7Y S6 $(s_a = 33.38\,\text{bp})$					
Market	**28.71 %**	**140.27 bp**	**41.64 bp**	**21.05 bp**	**7.43 bp**	
Normal	28.75 %	205.39 bp	58.04 bp	18.54 bp	2.75 bp	$\rho = 0.172$
t	28.71 %	139.45 bp	47.44 bp	24.94 bp	11.16 bp	$\lambda_M = -1.633$, $\lambda_Z = -65.209$, $\rho = 0.174$
NIG	28.71 %	138.27 bp	48.27 bp	25.16 bp	11.48 bp	$\alpha_M = 4.346$, $\alpha_Z = 2.537$, $\beta_M = -0.037$, $\beta_Z = -2.171$, $\rho = 0.541$
HYP	28.71 %	138.60 bp	51.21 bp	26.84 bp	11.45 bp	$\alpha_M = 3.561$, $\alpha_Z = 2.227$, $\beta_M = -2.084$, $\beta_Z = -1.181$, $\rho = 0.326$
VG	28.71 %	140.15 bp	42.72 bp	19.77 bp	7.04 bp	$\lambda_M = 1.061$, $\lambda_Z = 1.842$, $\alpha_M = 3.696$, $\alpha_Z = 7.821$, $\beta_M = 1.320$, $\beta_Z = -1.582$, $\rho = 0.415$
	iTraxx 10Y S6 $(s_a = 43.38\,\text{bp})$					
Market	**42.67 %**	**360.34 bp**	**105.08 bp**	**43.33 bp**	**13.52 bp**	
Normal	42.69 %	387.27 bp	157.51 bp	70.08 bp	16.34 bp	$\rho = 0.191$
t	42.69 %	342.74 bp	130.40 bp	63.56 bp	23.39 bp	$\lambda_M = -2.195$, $\lambda_Z = -65.072$, $\rho = 0.212$
NIG	42.67 %	358.94 bp	111.56 bp	56.02 bp	22.63 bp	$\alpha_M = 0.824$, $\alpha_Z = 11.156$, $\beta_M = 0.734$, $\beta_Z = 10.647$, $\rho = 0.275$
HYP	42.67 %	356.55 bp	104.59 bp	33.61 bp	5.96 bp	$\alpha_M = 2.613$, $\alpha_Z = 1.700$, $\beta_M = 0.897$, $\beta_Z = -0.025$, $\rho = 0.181$
VG	42.67 %	358.79 bp	107.92 bp	41.06 bp	10.97 bp	$\lambda_M = 1.422$, $\lambda_Z = 2.438$, $\alpha_M = 11.352$, $\alpha_Z = 4.210$, $\beta_M = 4.620$, $\beta_Z = -2.711$, $\rho = 0.421$

Table 3.1: Market prices of the 5, 7, and 10 year iTraxx contracts at November 13, 2006, and calibrated model prices.

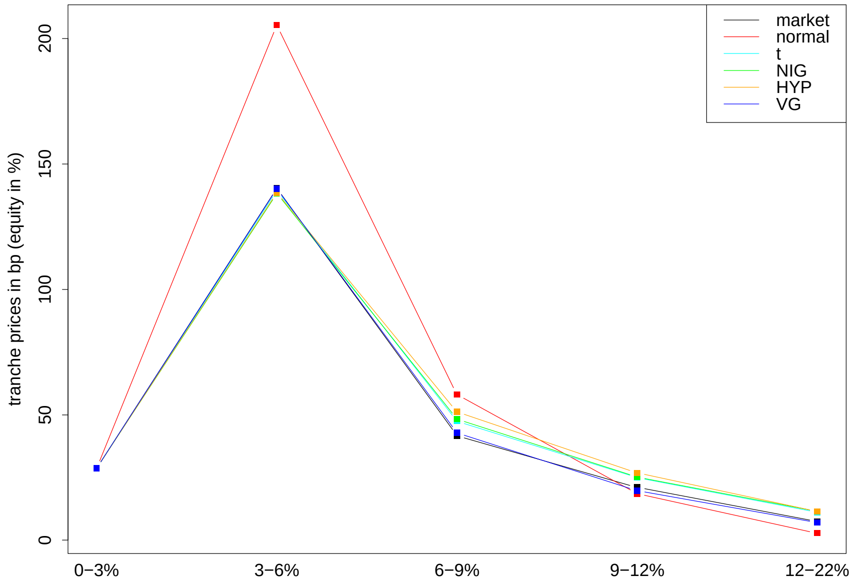


Figure 3.6: Comparison of calibrated model prices and market prices of the 7 year iTraxx contracts.

models for the 5 and 7 year iTraxx contracts, but worse for the 10 year contracts. The goodness of fit of the NIG and HYP models is similar. The absolute pricing errors of the NIG model are slightly smaller for the shorter maturities, but greater than those of the HYP model for the 10 year maturity. The VG model always provides the best fit. Since the t-model is the only one exhibiting tail dependence (confer the remark on page 118) but does not outperform the NIG, HYP and VG models, one may conclude that this property is negligible in the presence of more flexible factor distributions. This may also be confirmed by the fact that all estimated GH parameters β_M and β_Z are different from zero, which means the factor distributions are skewed. Furthermore, the parameter ρ is usually higher in the GH factor models than in the normal benchmark model. This indicates that correlation is still of some importance, but has a different impact on the pricing formula because of the more complex dependence structure.

The VG model even has the potential to fit the market prices of all tranches and maturities simultaneously with high accuracy, which we shall show below. However, before that we want to point out that the calibration over different maturities requires some additional care to avoid inconsistencies when calculating the default probabilities. As can be seen from Figure 3.7, the average iTraxx spreads s_a are increasing in maturity, and by equation (3.4) so do the default intensities λ_a. This means that the estimated default probabilities $Q(t) = 1 - e^{-\lambda_a t}$ of a CDO with a longer lifetime are always greater than those of a CDO with a shorter maturity. While this can be neglected when concentrating on just one maturity, this fact has to be taken into account when considering iTraxx CDOs of different maturities together. Since the underlying portfolio is the same, the default probabilities should coincide during the common lifetime.

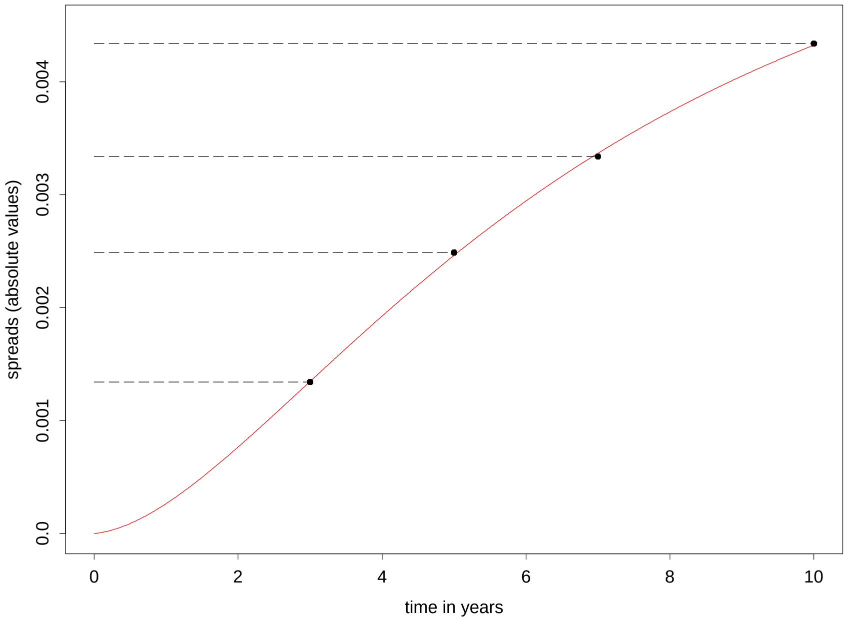


Figure 3.7: Constant iTraxx spreads of November 13, 2006, and fitted Nelson–Siegel curve $\hat{r}_{NS}$ with parameters $\hat{\beta}_0 = 0.0072, \hat{\beta}_1 = -0.0072, \hat{\beta}_2 = -0.0069, \hat{\tau}_1 = 2.0950$.

To avoid these problems, we now assume that the average spreads $s_a = s(t)$ are time-dependent and follow a Nelson–Siegel curve. This parametric family of functions has been introduced in Nelson and Siegel (1987) and has become very popular in interest rate theory for the modeling of yield curves where the task is the following: Let $\beta(0, t_k)$ denote today's price of a zero coupon bond with maturity t_k as before, then one has to find a function f (*instantaneous forward rates*) such that the model prices $\beta(0, t_k) = \exp\bigl(-\int_0^{t_k} f(t)\,\mathrm{d}t\bigr)$ approximate the market prices reasonably well for all maturities t_k. Since instantaneous forward rates cannot be observed directly in the market, one often uses an equivalent expression in terms of *spot rates*: $\beta(0, t_k) = \exp(-r(t_k)\,t_k)$, where the spot rate is given by $r(t_k) = \frac{1}{t_k}\int_0^{t_k} f(t)\,\mathrm{d}t$. Nelson and Siegel suggested to model the forward rates by

$$f_{NS(\beta_0,\beta_1,\beta_2,\tau_1)}(t) = \beta_0 + \beta_1 e^{-\frac{t}{\tau_1}} + \beta_2 \frac{t}{\tau_1}\, e^{-\frac{t}{\tau_1}}.$$

The corresponding spot rates are given by

$$r_{NS(\beta_0,\beta_1,\beta_2,\tau_1)}(t) = \beta_0 + (\beta_1 + \beta_2)\frac{\tau_1}{t}\bigl(1 - e^{-\frac{t}{\tau_1}}\bigr) - \beta_2 e^{-\frac{t}{\tau_1}}. \tag{3.17}$$

In order to obtain time-consistent default probabilities resp. intensities, we replace s_a in equation (3.4) by a Nelson–Siegel spot rate curve (3.17) that has been fitted to the four quoted average iTraxx spreads, that is,

$$\lambda_a = \lambda(t) = \frac{\hat{r}_{NS}(t)}{(1-R)10000}, \tag{3.18}$$

Tranches	Market	VG	Market	VG	Market	VG
	5Y		7Y		10Y	
0–3 %	13.60 %	13.60 %	28.71 %	28.72 %	42.67 %	42.67 %
3–6 %	57.16 bp	53.30 bp	140.27 bp	132.27 bp	360.34 bp	357.60 bp
6–9 %	16.31 bp	17.19 bp	41.64 bp	41.83 bp	105.08 bp	111.17 bp
9–12 %	6.65 bp	8.23 bp	21.05 bp	19.90 bp	43.33 bp	52.00 bp
12–22 %	2.67 bp	3.05 bp	7.43 bp	7.34 bp	13.52 bp	18.97 bp

Table 3.2: Results of the VG model calibration simultaneously over all maturities. The estimated parameters are as follows: $\lambda_M = 0.920$, $\alpha_M = 5.553$, $\beta_M = 1.157$, $\lambda_Z = 2.080$, $\alpha_Z = 2.306$, $\beta_Z = -0.753$, $\rho = 0.321$.

and $Q(t) := 1 - e^{-\lambda(t)\,t}$. The Nelson–Siegel curve estimated from the iTraxx spreads of November 13, 2006, is shown in Figure 3.7. At first glance the differences between constant and time-varying spreads seem to be fairly large, but one should observe that these are the absolute values which have already been divided by 10000 and therefore range from 0 to 0.004338, so the differences in the default probabilities are almost negligible.

Under the additional assumption (3.18), we have calibrated a model with VG distributed factors to the tranche prices of all maturities simultaneously. The results are summarized in Table 3.2 and visualized in Figure 3.8. The fit is excellent. The maximal absolute pricing error is less than 9 bp, and for the 5 and 7 year maturities the errors are, apart from the junior mezzanine tranches,

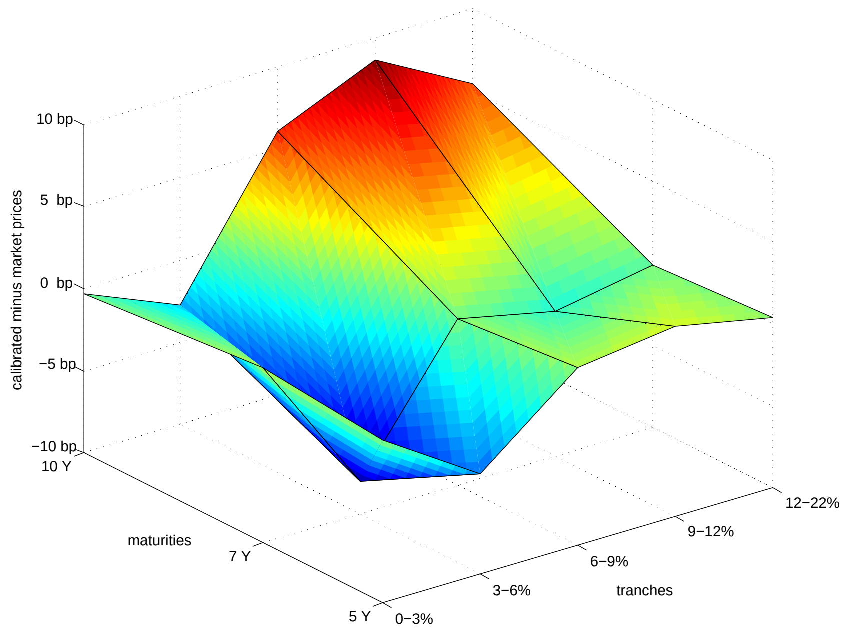


Figure 3.8: Graphical representation of the differences between model and market prices obtained from the simultaneous VG calibration.

almost as small as in the previous calibrations. The junior mezzanine tranche is underpriced for all maturities, but it is difficult to say whether this is caused by model or by market imperfections. Nevertheless the overall pricing performance of the extended VG model is comparable or better than the performance of the models considered in Albrecher, Ladoucette, and Schoutens (2007), Burtschell, Gregory, and Laurent (2005), and Kalemanova, Schmid, and Werner (2007), although the latter were only calibrated to tranche quotes of a single maturity.

Also note that this model admits a flat correlation structure not only over all tranches, but also over different maturities: all model prices in Table 3.2 were calculated using the same parameter ρ. Thus the correlation smiles shown in Figure 3.3 which in some sense question the factor equation (3.5) resp. (3.10) are completely eliminated, and there is no need for a somewhat artificial base correlation framework. Therefore the intuitive idea of the factor approach is preserved, but one should keep in mind that in the case of GH distributed factors the dependence structure of the joint distribution of the X_i is more complex and cannot be described by correlation alone.

3.3 Summary and outlook

In this chapter we presented a detailed description of synthetic CDOs, a widespread instrument in portfolio credit risk management, and the normal factor model used to price them. Though the latter has been established as a market standard, it is completely unable to capture and reproduce the quoted prices, which can be seen from the implied correlation smile as well as from the poor calibration results. The main reasons for this unsatisfactory behaviour are the lack of additional parameters and a too simple dependence structure between individual default times. We have shown how these deficiencies can be remedied by implementing more flexible and advanced factor distributions. Extended models using generalized hyperbolic distributions provide an excellent fit to quoted market spreads, but remain analytically and numerically tractable yet. This in particular means that the cumulative default distribution F_{Z_t} derived from the model assumptions can be computed reasonably fast and accurate.

A consistent modeling of F_{Z_t} is not only essential for a correct pricing of credit portfolio tranches, but also for an adequate rating and determination of tranche sizes as the following example clarifies. Suppose there is a senior tranche of a large homogeneous credit portfolio with lower attachment point K_l and upper attachment point $K_u = 1$. Depending on the risk aversion of potential investors, the tranche has to get a certain rating. Let us assume for simplicity that it will be assigned to a certain rating class i if the probability that the notional value of the tranche is reduced by defaults during the lifetime of the contract is smaller than the corresponding rating probability pd_i. Using the notation of the previous sections, this means that the senior tranche will get the rating i if and only if $F_{Z_T}(K_l) \geq 1 - pd_i$. Thus the desired rating can always be achieved by choosing K_l accordingly, but this choice crucially depends on the shape of F_{Z_T} and hence on the underlying factor model.

Assuming zero recovery ($R = 0$) and an individual default probability

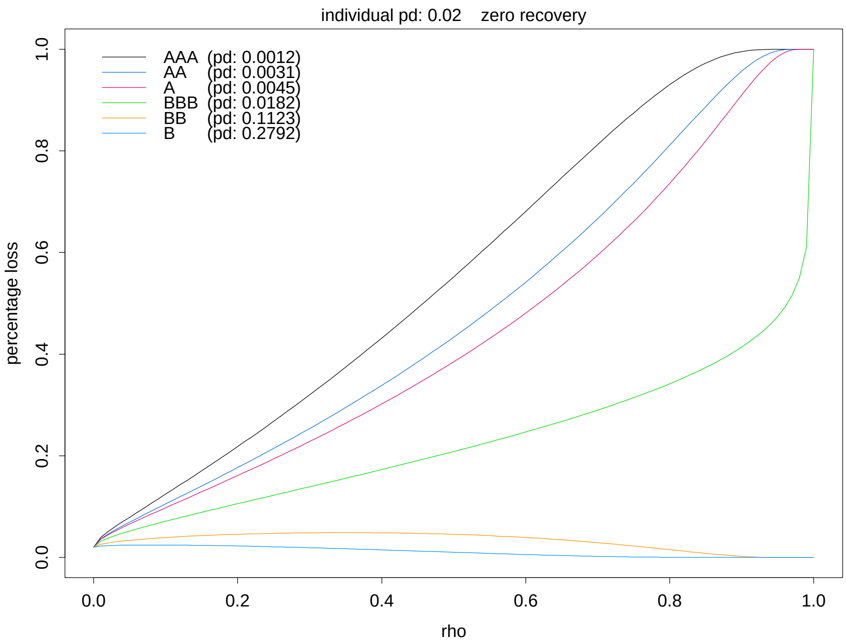
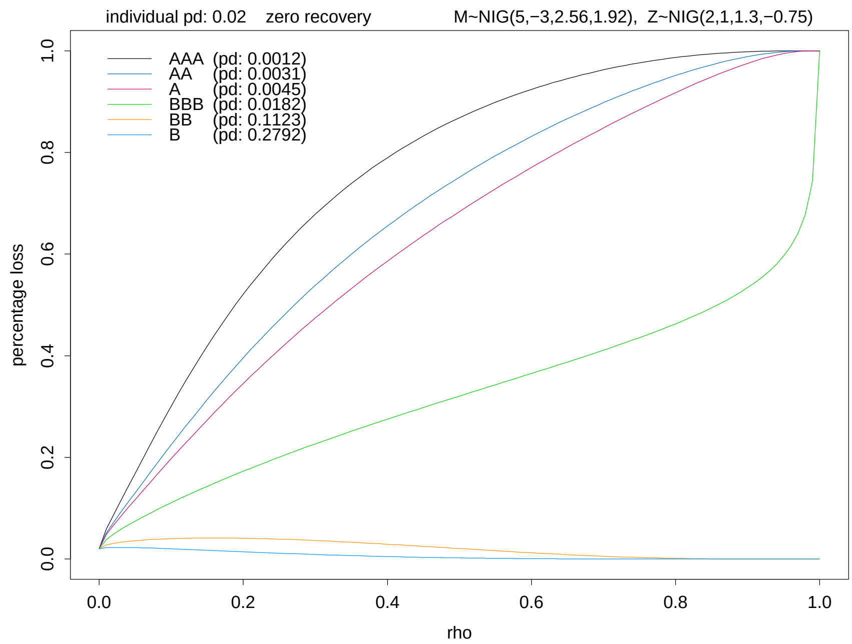


Figure 3.9: Dependence of optimal lower attachment points $K_l^* = F_{Z_T}^{-1}(1-p_i)$ for different rating classes on ρ in the normal factor model $(M, Z_i \sim N(0,1))$.

$Q(T) = 0.02$, Figure 3.9 shows the dependence of the lower attachment points $K_l^* = F_{Z_T}^{-1}(1 - pd_i)$ on the correlation parameter ρ in the normal factor model where F_{Z_T} is given by equation (3.8). The probabilities pd_i associated to the different rating classes given there are just chosen for illustration purposes and do not correspond to real market values. Note that in this model ρ is the only parameter that can influence K_l^*. However, things change significantly if we move to the extended model and allow for alternative factor distributions such that F_{Z_T} is given by (3.11). Two examples are visualized in Figure 3.10. In comparison the the normal factor model, the differences in the shapes of the curves $K_l^*(\rho)$ are obvious. This again shows the great impact of the factor distributions on the dependence structure and the loss distribution of the portfolio.

In the present state, our model only incorporates constant default intensities λ_a resp. deterministic intensity functions $\lambda(t)$ (see equations (3.4) and (3.18)). A topic for future research is the extension to dynamic intensity models which can also capture the movements of tranche spreads over several trading days.



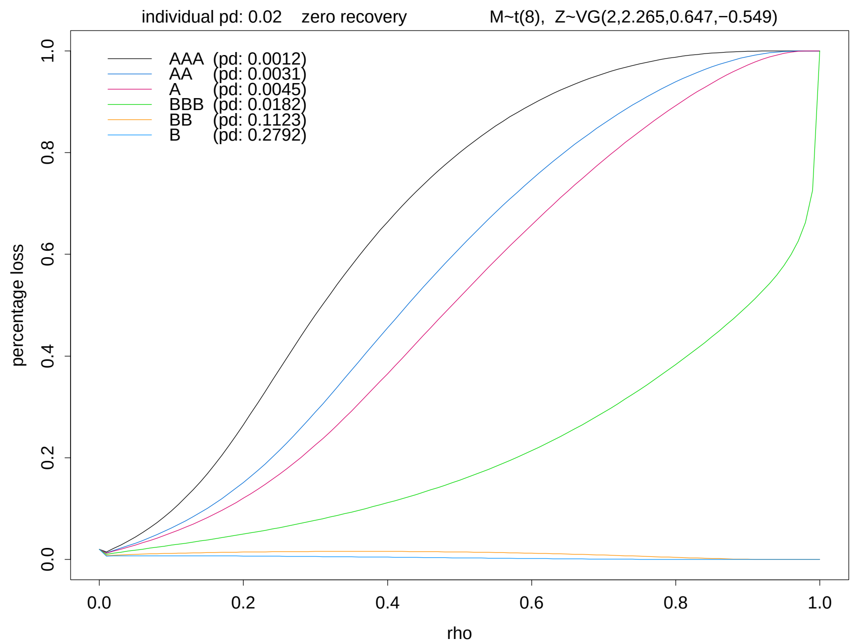


Figure 3.10: Dependence of optimal lower attachment points $K_l^* = F_{Z_T}^{-1}(1-p_i)$ for different rating classes on ρ in the extended factor model $(M \sim GH(\lambda_M, \alpha_M, \beta_M, \bar{\delta}_M, \bar{\mu}_M),\ Z_i \sim GH(\lambda_Z, \alpha_Z, \beta_Z, \bar{\delta}_Z, \bar{\mu}_Z))$.

Appendix A

Bessel functions

We summarize some properties of Bessel functions used within this thesis.

The modified Bessel functions of third kind $K_\lambda(z)$ are solutions of the differential equation

$$z^2 \frac{d^2 f}{dz^2} + z \frac{df}{dz} - (z^2 + \lambda^2) f = 0.$$

They are regular functions of z throughout the complex z-plane cut along the negative real axis, and for fixed $z \neq 0$, $K_\lambda(z)$ is an entire function of λ. $K_\lambda(z)$ tends to zero for all λ as $|z| \to \infty$ in the sector $|\arg(z)| < \frac{\pi}{2}$. Moreover, $K_\lambda(z)$ is real and positive if $z = x \in \mathbb{R}$ and $x > 0$ (Abramowitz and Stegun 1968, p. 374).

Integral representation

$$K_\lambda(x) = \frac{1}{2} \int_0^\infty y^{\lambda-1}\, e^{-\frac{x}{2}(y+y^{-1})}\, \mathrm{d}y, \quad x > 0. \tag{A.1}$$

Reference: Watson (1952, p. 182, formula (8))

Basic properties

$$K_\lambda(x) = K_{-\lambda}(x), \tag{A.2}$$

$$K_{\lambda+1}(x) = \frac{2\lambda}{x} K_\lambda(x) + K_{\lambda-1}(x), \tag{A.3}$$

$$-2K'_\lambda(x) = K_{\lambda-1}(x) + K_{\lambda+1}(x), \tag{A.4}$$

$$K_{\lambda+\epsilon}(x) - K_\lambda(x) > 0 \quad \text{for all } \lambda \geq 0 \text{ and } \epsilon > 0. \tag{A.5}$$

References: Watson (1952, p. 79, formulas (8), (1) and (2)) for (A.2)–(A.4), the last inequality is mentioned in Lorch (1967, p. 2) and goes back to Soni (1965).

Series representation for $\boldsymbol{\lambda = n + \frac{1}{2}}$, $\boldsymbol{n \in \mathbb{N}_0}$

$$K_{n+\frac{1}{2}}(x) = \sqrt{\frac{\pi}{2x}}\, e^{-x} \left[1 + \sum_{i=1}^{n} \frac{(n+i)!}{(n-i)!i!} (2x)^{-i} \right], \tag{A.6}$$

$$\Rightarrow \quad K_{-\frac{1}{2}}(x) = K_{\frac{1}{2}}(x) = \sqrt{\frac{\pi}{2x}}\, e^{-x}. \tag{A.7}$$

Reference: Watson (1952, p. 80, formula (12)).

Asymptotic behaviour

$$K_\lambda(x) \sim \frac{1}{2}\,\Gamma(|\lambda|)\left(\frac{x}{2}\right)^{-|\lambda|}, \qquad x \downarrow 0,\ \lambda \neq 0, \tag{A.8}$$

$$K_0(x) \sim -\ln(x), \qquad x \downarrow 0, \tag{A.9}$$

$$K_\lambda(x) \sim \sqrt{\frac{\pi}{2x}}\, e^{-x}, \qquad x \to \infty, \tag{A.10}$$

References: Abramowitz and Stegun (1968, fomulas 9.6.8, 9.6.9 and 9.7.2).

The Bessel functions of first and second kind, $J_\lambda(z)$ and $Y_\lambda(z)$, respectively, are solutions of the differential equation

$$z^2\,\frac{d^2 f}{dz^2} + z\,\frac{df}{dz} + (z^2 - \lambda^2) f = 0.$$

They are also regular functions of z throughout the complex z-plane cut along the negative real axis, and for fixed $z \neq 0$, $J_\lambda(z)$ and $Y_\lambda(z)$ are entire functions of λ. Moreover, both functions are real valued if $z = x \in \mathbb{R}$ and $x > 0$ (Abramowitz and Stegun 1968, p. 358).

Relations between J_λ and Y_λ

$$Y_\lambda(x) = \frac{J_\lambda(x)\cos(\lambda\pi) - J_{-\lambda}(x)}{\sin(\lambda\pi)} \tag{A.11}$$

Reference: Watson (1952, p. 64, formula (1)).

Representations for $|\lambda| = \frac{1}{2}$

$$\begin{aligned} J_{\frac{1}{2}}(x) &= \sqrt{\frac{2}{\pi x}}\sin(x), & J_{-\frac{1}{2}}(x) &= \sqrt{\frac{2}{\pi x}}\cos(x), \\ Y_{\frac{1}{2}}(x) &= -\sqrt{\frac{2}{\pi x}}\cos(x), & Y_{-\frac{1}{2}}(x) &= -\sqrt{\frac{2}{\pi x}}\sin(x) \end{aligned} \tag{A.12}$$

References: The representations of $J_{\frac{1}{2}}$ and $J_{-\frac{1}{2}}$ can be found in Watson (1952, p. 54, formula (3) and p. 55, formula (6)), those of $Y_{\frac{1}{2}}$ and $Y_{-\frac{1}{2}}$ then immediately follow from (A.11).

Asymptotic behaviour for $x \downarrow 0$

$$J_\lambda(x) \sim \left(\frac{x}{2}\right)^{\lambda}\left(\Gamma(\lambda+1)\right)^{-1}, \qquad \lambda \geq 0, \tag{A.13}$$

$$Y_\lambda(x) \sim -\frac{\Gamma(\lambda)}{\pi}\left(\frac{x}{2}\right)^{-\lambda}, \qquad \lambda > 0, \tag{A.14}$$

$$Y_0(x) \sim \frac{2}{\pi}\,\ln(x). \tag{A.15}$$

References: Abramowitz and Stegun (1968, fomulas 9.1.7, 9.1.9 and 9.1.8).

Asymptotic expansions for $x \to \infty$

$$
\begin{aligned}
J_\lambda(x) \sim \sqrt{\frac{2}{\pi x}} \Bigg[&\cos\left(x - \frac{\lambda\pi}{2} - \frac{\pi}{4}\right) \sum_{m \geq 0} (-1)^m \frac{(\lambda, 2m)}{(2x)^{2m}} \\
&- \sin\left(x - \frac{\lambda\pi}{2} - \frac{\pi}{4}\right) \sum_{m \geq 0} (-1)^m \frac{(\lambda, 2m+1)}{(2x)^{2m+1}} \Bigg],
\end{aligned} \tag{A.16}
$$

$$
\begin{aligned}
Y_\lambda(x) \sim \sqrt{\frac{2}{\pi x}} \Bigg[&\sin\left(x - \frac{\lambda\pi}{2} - \frac{\pi}{4}\right) \sum_{m \geq 0} (-1)^m \frac{(\lambda, 2m)}{(2x)^{2m}} \\
&+ \cos\left(x - \frac{\lambda\pi}{2} - \frac{\pi}{4}\right) \sum_{m \geq 0} (-1)^m \frac{(\lambda, 2m+1)}{(2x)^{2m+1}} \Bigg],
\end{aligned} \tag{A.17}
$$

$$
\text{where } (\lambda, m) = \frac{(4\lambda^2 - 1^2)(4\lambda^2 - 3^2)\ldots(4\lambda^2 - (2m-1)^2)}{2^{2m}\, m!}\,.
$$

References: Watson (1952, p. 199).

References

Aas, K. and I. H. Haff (2006). The generalized hyperbolic skew Student's t-distribution. *Journal of Financial Econometrics 4*(2), 275–309.

Abramowitz, M. and I. A. Stegun (1968). *Handbook of Mathematical Functions* (5th ed.). New York: Dover Publications.

Adcock, C. J. (2010). Asset pricing and portfolio selection based on the multivariate extended skew-Student-t distribution. *Annals of Operations Research 176*(1), 221–234.

Affleck-Graves, J. and B. McDonald (1989). Nonnormalities and tests of asset pricing theories. *Journal of Finance 44*(4), 889–908.

Albrecher, H., S. A. Ladoucette, and W. Schoutens (2007). A generic one-factor Lévy model for pricing synthetic CDOs. In M. C. Fu, R. A. Jarrow, J.-Y. J. Yen, and R. J. Elliott (Eds.), *Advances in Mathematical Finance*, Applied and Numerical Harmonic Analysis, pp. 259–277. Boston: Birkhäuser.

Andersen, L. and J. Sidenius (2005). Extensions to the Gaussian copula: random recovery and random factor loadings. *Journal of Credit Risk 1*(1), 29–70.

Artzner, P., F. Delbaen, J.-M. Eber, and D. Heath (1999). Coherent measures of risk. *Mathematical Finance 9*(3), 203–228.

Asmussen, S. and J. Rosiński (2001). Approximations of small jumps of Lévy processes with a view towards simulation. *Journal of Applied Probability 38*, 482–493.

Banachewicz, K. and A. van der Vaart (2008). Tail dependence of skewed grouped t-distributions. *Statistics & Probability Letters 78*, 2388–2399.

Barndorff-Nielsen, O. E. (1977). Exponentially decreasing distributions for the logarithm of particle size. *Proceedings of the Royal Society London A 353*, 401–419.

Barndorff-Nielsen, O. E. (1998). Processes of normal inverse Gaussian type. *Finance and Stochastics 2*, 41–68.

Barndorff-Nielsen, O. E. and C. Halgreen (1977). Infinite divisibility of the hyperbolic and generalized inverse Gaussian distributions. *Zeitschrift für Wahrscheinlichkeitstheorie und verwandte Gebiete 38*, 309–312.

Barndorff-Nielsen, O. E., M. Maejima, and K.-I. Sato (2006). Some classes of multivariate infinitely divisible distributions admitting stochastic integral representations. *Bernoulli 12*(1), 1–33.

Barndorff-Nielsen, O. E. and K. Prause (2001). Apparent scaling. *Finance and Stochastics 5*, 103–113.

Berrada, T., D. Dupuis, E. Jacquier, N. Papageorgiou, and B. Rémillard (2006). Credit migration and basket derivatives pricing with copulas. *Journal of Computational Finance 10*(1), 43–68.

Bingham, N. H., C. M. Goldie, and E. Omey (2006). Regularly varying probability densities. *Publications de l'Institut Mathématique Beograd 80(94)*, 47–57.

Bingham, N. H., C. M. Goldie, and J. L. Teugels (1987). *Regular Variation*, Volume 27 of *Encyclopedia of Mathematics and Its Applications.* Cambridge: Cambridge University Press.

Blæsild, P. (1981). The two-dimensional hyperbolic distribution and related distributions, with an application to Johannsen's bean data. *Biometrika 68*, 251–263.

Blæsild, P. and J. L. Jensen (1981). Multivariate distributions of hyperbolic type. In C. Taillie, G. P. Patil, and B. Baldessari (Eds.), *Statistical Distributions in Scientific Work*, Volume 4, pp. 45–66. Dordrecht: Reidel.

Blattberg, R. C. and N. J. Gonedes (1974). A comparison of the stable and Student distributions as statistical models for stock prices. *Journal of Business 47*, 244–280.

Bondesson, L. (1982). On simulation from infinitely divisible distributions. *Advances in Applied Probability 14*, 855–869.

Bondesson, L. (1992). *Generalized Gamma Convolutions and Related Classes of Distributions and Densities.* Lecture Notes in Statistics 76. Berlin: Springer.

Burtschell, X., J. Gregory, and J.-P. Laurent (2005). A comparative analysis of CDO pricing models. working paper, BNP Paribas (updated version: 2008).

Burtschell, X., J. Gregory, and J.-P. Laurent (2007). Beyond the Gaussian copula: stochastic and local correlation. *Journal of Credit Risk 3*(1), 31–62.

Cambanis, S., S. Huang, and G. Simons (1981). On the theory of elliptically contoured distributions. *Journal of Multivariate Analysis 11*, 368–385.

Carr, P., H. Geman, D. B. Madan, and M. Yor (2002). The fine structure of asset returns: an empirical investigation. *Journal of Business 75*, 305–332.

Carr, P., H. Geman, D. B. Madan, and M. Yor (2007). Self-decomposability and option pricing. *Mathematical Finance 17*, 31–57.

Cherubini, U., E. Luciano, and W. Vecchiato (2004). *Copula Methods in Finance*. Chichester: Wiley.

Cline, D. B. H. (1986). Convolution tails, product tails and domains of attraction. *Probability Theory and Related Fields 72*, 529–557.

Cont, R. and P. Tankov (2004). *Financial Modelling with Jump Processes*. Boca Raton: Chapman & Hall / CRC Press.

Cox, J. C., S. A. Ross, and M. Rubinstein (1979). Option pricing: A simplified approach. *Journal of Financial Economics 7*, 229–264.

Devroye, L. (1986). *Non-Uniform Random Variate Generation*. New York: Springer.

Doetsch, G. (1950). *Handbuch der Laplace-Transformation*, Volume I. Basel: Birkhäuser.

Eberlein, E. (2001). Application of generalized hyperbolic Lévy motions to finance. In O. E. Barndorff-Nielsen, T. Mikosch, and S. Resnick (Eds.), *Lévy Processes: Theory and Applications*, pp. 319–336. Boston: Birkhäuser.

Eberlein, E., R. Frey, and E. A. v. Hammerstein (2008). Advanced credit portfolio modeling and CDO pricing. In W. Jäger and H.-J. Krebs (Eds.), *Mathematics - Key Technology for the Future: Joint Projects between Universities and Industry 2004–2007*, pp. 253–280. Berlin: Springer.

Eberlein, E. and J. Jacod (1997). On the range of options prices. *Finance and Stochastics 1*, 131–140.

Eberlein, E. and U. Keller (1995). Hyperbolic distributions in finance. *Bernoulli 1*, 281–299.

Eberlein, E., U. Keller, and K. Prause (1998). New insights into smile, mispricing, and value at risk: the hyperbolic model. *Journal of Business 71*, 371–405.

Eberlein, E. and W. Kluge (2006). Exact pricing formulae for caps and swaptions in a Lévy term structure model. *Journal of Computational Finance 9*, 99–125.

Eberlein, E. and W. Kluge (2007). Calibration of Lévy term structure models. In M. C. Fu, R. A. Jarrow, J.-Y. J. Yen, and R. J. Elliott (Eds.), *Advances in Mathematical Finance*, Applied and Numerical Harmonic Analysis, pp. 147–172. Boston: Birkhäuser.

Eberlein, E. and N. Koval (2006). A cross-currency Lévy market model. *Quantitative Finance 6*, 465–480.

Eberlein, E. and F. Özkan (2003). Time consistency of Lévy models. *Quantitative Finance 3*, 40–50.

Eberlein, E. and K. Prause (2002). The generalized hyperbolic model: financial derivatives and risk measures. In H. Geman, D. Madan, S. Pliska, and T. Vorst (Eds.), *Mathematical Finance – Bachelier Congress 2000*, pp. 245–267. Berlin: Springer.

Eberlein, E. and S. Raible (1999). Term structure models driven by general Lévy processes. *Mathematical Finance 9*, 31–54.

Eberlein, E. and E. A. v. Hammerstein (2004). Generalized hyperbolic and inverse Gaussian distributions: limiting cases and approximation of processes. In R. C. Dalang, M. Dozzi, and F. Russo (Eds.), *Seminar on Stochastic Analysis, Random Fields and Applications IV*, Volume 58 of *Progress in Probability*, pp. 221–264. Basel: Birkhäuser.

Embrechts, P. and C. M. Goldie (1980). On closure and factorization properties of subexponential distributions. *Journal of the Australian Mathematical Society A 29*, 243–256.

Embrechts, P., A. J. McNeil, and D. Straumann (2002). Correlation and dependence in risk management: properties and pitfalls. In M. Dempster (Ed.), *Risk Management: Value at Risk and Beyond*, pp. 176–223. Cambridge: Cambridge University Press.

Fang, K.-T., S. Kotz, and K. W. Ng (1990). *Symmetric Multivariate and Related Distributions*, Volume 36 of *Monographs on statistics and applied probability*. London: Chapman & Hall.

Feller, W. (1971). *An Introduction to Probability Theory and Its Applications II* (2^{nd} ed.). New York: Wiley.

Fung, T. and E. Seneta (2010). Tail dependence for two skew t distributions. *Statistics & Probability Letters 80*, 784–791.

Geman, H., D. B. Madan, and M. Yor (2001). Asset prices are Brownian motion: only in business time. In M. Avellaneda (Ed.), *Quantitative Analysis in Financial Markets*, Volume II, pp. 103–146. Singapore: World Scientific.

Good, I. J. (1953). On the population frequencies of species and the estimation of population parameters. *Biometrika 40*, 237–264.

Gregory, J. and J.-P. Laurent (2004). In the core of correlation. *Risk October*, 87–91.

Grosswald, E. (1976). The Student-t-distribution of any degree of freedom is infinitely divisible. *Zeitschrift für Wahrscheinlichkeitstheorie und verwandte Gebiete 36*, 103–109.

Halgreen, C. (1979). Self-decomposability of the generalized inverse Gaussian and hyperbolic distributions. *Zeitschrift für Wahrscheinlichkeitstheorie und verwandte Gebiete 47*, 13–17.

Hull, J. C. and A. White (2004). Valuation of a CDO and an n^{th} to default CDS without Monte Carlo simulation. *Journal of Derivatives 12*, 8–23.

Hult, H. and F. Lindskog (2002). Multivariate extremes, aggregation and dependence in elliptical distributions. *Advances in Applied Probability 34*(3), 587–608.

Jacod, J. and A. N. Shiryaev (2003). *Limit Theorems for Stochastic Processes* (2^{nd} ed.). Berlin: Springer.

Jänich, K. (1996). *Funktionentheorie. Eine Einführung* (4^{th} ed.). Berlin: Springer.

Johnson, N. L., S. Kotz, and N. Balakrishnan (1995). *Continuous Univariate Distributions* (2^{nd} ed.), Volume 2. New York: Wiley.

Jørgensen, B. (1982). *Statistical Properties of the Generalized Inverse Gaussian distribution.* Lecture Notes in Statistics 9. New York: Springer.

Kalemanova, A., B. Schmid, and R. Werner (2007). The normal inverse Gaussian distribution for synthetic CDO pricing. *Journal of Derivatives 14*(3), 80–94.

Khan, R. and G. Zhou (2006). Modeling Non-Normality Using Multivariate t: Implications for Asset Pricing. working paper, University of Toronto.

Kring, S., S. T. Rachev, M. Höchstötter, F. J. Fabozzi, and M. L. Bianchi (2009). Multi-tail generalized elliptical distributions for asset returns. *Econometrics Journal 12*, 272–291.

Küchler, U. and S. Tappe (2008). Bilateral gamma distributions and processes in financial mathematics. *Stochastic Processes and their Applications 118*, 261–283.

Li, D. X. (2000). On default correlation: a copula function approach. *Journal of Fixed Income 9*, 43–54.

Loève, M. (1977). *Probability Theory I* (4^{th} ed.). New York: Springer.

Lorch, L. (1967). Inequalities for some Whittaker functions. *Archivum Mathematicum 3*(1), 1–9.

Luciano, E. and W. Schoutens (2006). A multivariate jump-driven financial asset model. *Quantitative Finance 6*(5), 385–402.

Luciano, E. and P. Semeraro (2010). A generalized normal mean-variance mixture for return processes in finance. *International Journal of Theoretical and Applied Finance 13*(3), 415–440.

Madan, D. B., P. Carr, and E. C. Chang (1998). The variance gamma process and option pricing. *European Finance Review 2*, 79–105.

Madan, D. B. and E. Seneta (1990). The variance gamma (V.G.) model for share market returns. *Journal of Business 63*, 511–524.

Malevergne, Y. and D. Sornette (2004). How to account for extreme co-movements between individual stocks and the market. *Journal of Risk 6*(3), 71–116.

Maller, R. A., D. H. Solomon, and A. Szimayer (2006). A multinomial approximation for American option prices in Lévy process models. *Mathematical Finance 16*(4), 613–633.

Mandelbrot, B. B. (1963). The variation of certain speculative prices. *Journal of Business 36*, 394–419.

Marsaglia, G. and W. W. Tsang (2000). A simple method for generating gamma variables. *ACM Transactions on Mathematical Software 26*(3), 363–372.

McNeil, A., R. Frey, and P. Embrechts (2005). *Quantitative Risk Management.* Princeton: Princeton University Press.

Merton, R. C. (1974). On the pricing of corporate debt: the risk structure of interest rates. *Journal of Finance 29*, 449–470.

Michael, J. R., W. R. Schucany, and R. W. Haas (1976). Generating random variates using transformations with multiple roots. *The American Statistican 30*(2), 88–90.

Nelsen, R. B. (1999). *An Introduction to Copulas.* Lecture Notes in Statistics 139. New York: Springer.

Nelson, C. R. and A. F. Siegel (1987). Parsimonious modeling of yield curves. *Journal of Business 60*, 473–489.

Owen, J. and R. Rabinovitch (1983). On the class of elliptical distributions and their applications to the theory of portfolio choice. *Journal of Finance 38*(3), 745–752.

Prause, K. (1999). The generalized hyperbolic model: estimation, financial derivatives, and risk measures. Ph.D. thesis, University of Freiburg.

Raible, S. (2000). Lévy processes in finance: theory, numerics and empirical facts. Ph.D. thesis, University of Freiburg.

Richardson, M. and T. Smith (1993). A test for multivariate normality in stock returns. *Journal of Business 66*(2), 295–321.

Rüschendorf, L. (2009). On the distributional transform, Sklar's theorem, and the empirical copula process. *Journal of Statistical Planning and Inference 139*(11), 3921–3927.

Sato, K.-I. (1999). *Lévy Processes and Infinitely Divisible Distributions.* Cambridge: Cambridge University Press.

Sato, K.-I. (2001). Subordination and self-decomposability. *Statistics & Probability Letters 54*, 317–324.

Schmidt, R., T. Hrycej, and E. Stützle (2006). Multivariate distribution models with generalized hyperbolic margins. *Computational Statistics & Data Analysis 50*, 2065–2096.

Schönbucher, P. J. and D. Schubert (2001). Copula-dependent default risk in intensity models. working paper, University of Bonn.

Schoutens, W. (2003). *Lévy Processes in Finance: Pricing Financial Derivatives.* Chichester: Wiley.

Schoutens, W. (2006). Exotic options under Lévy models: An overview. *Journal of Computational and Applied Mathematics 189*, 526–538.

Semeraro, P. (2008). A multivariate variance gamma model for financial applications. *International Journal of Theoretical and Applied Finance 11*(1), 1–18.

Sichel, H. (1973). Statistical valuation of diamondiferous deposits. *Journal of the Southern African Institute of Mining and Metallurgy 73*, 235–243.

Sichel, H. (1974). On a distribution representing sentence-length in written prose. *Journal of the Royal Statistical Society A 137*, 25–34.

Sklar, A. (1959). Fonctions de répartition à n dimensions et leurs marges. *Publications de l'Institut de Statistique de l'Université de Paris 8*, 229–231.

Soni, R. P. (1965). On an inequality for modified Bessel functions. *Journal of Mathematics and Physics 44*, 406–407.

Takano, K. (1989). On mixtures of the normal distribution by the generalized gamma convolutions. *Bulletin of the Faculty of Science, Ibaraki University, Series A 21*, 29–41. Correction and addendum in Vol. 22 (1990), 49–52.

Thorin, O. (1977a). On the infinite divisibility of the Pareto distribution. *Scandinavian Actuarial Journal*, 31–40.

Thorin, O. (1977b). On the infinite divisibility of the lognormal distribution. *Scandinavian Actuarial Journal*, 121–148.

Thorin, O. (1978). An extension of the notion of a generalized Γ-convolution. *Scandinavian Actuarial Journal*, 141–149.

Vasiček, O. A. (1987). Probability of loss on loan portfolio. Mimeo, KMV Corporation.

Vasiček, O. A. (1991). Limiting loan loss probability distribution. Mimeo, KMV Corporation.

Watson, G. N. (1952). *A Treatise on the Theory of Bessel Functions* (reprinted 2nd ed.). Cambridge: Cambridge University Press.

Wolfe, S. J. (1971). On moments of infinitely divisible distribution functions. *Annals of Mathematical Statistics 42*, 2036–2043.