

# **ghyp: A package on generalized hyperbolic distributions**

Wolfgang Breymann\*      David Lüthi<sup>†</sup>

March 30, 2011

In this document the generalized hyperbolic (GH) distribution is explained to the extent which is required to understand the internals of the R package **ghyp**. Essentially, the density and moment generating functions of the GH distribution and its special cases are provided together with some important properties of the GH family. In addition the algorithm underlying the fitting procedure for the multivariate GH distribution is explained. Then we shed light on how to switch between different parametrizations of the GH distribution. In the appendix we describe the generalized inverse Gaussian distribution and give some useful facts regarding the modified Bessel function of the third kind. Finally, we write on the design on the package **ghyp** and give some code chunks which indicate how the software can be used.

---

\*Institute of Data Analysis and Process Design, e-mail: wolfgang.breymann@zhaw.ch

<sup>†</sup>SIM Global Advantage Investment Management, e-mail: luethid@gmail.com

## Contents

<b>1. Introduction</b>	<b>3</b>
<b>2. Definition</b>	<b>3</b>
2.1. Expected value and variance . . . . .	4
2.2. Density . . . . .	4
2.3. Moment generating function . . . . .	4
2.4. Linear transformations . . . . .	5
<b>3. Special cases of the generalized hyperbolic distribution</b>	<b>5</b>
<b>4. Parametrization</b>	<b>5</b>
4.1. $(\lambda, \chi, \psi, \mu, \Sigma, \gamma)$ -Parametrization . . . . .	6
4.2. $(\lambda, \bar{\alpha}, \mu, \Sigma, \gamma)$ -Parametrization . . . . .	6
4.3. $(\lambda, \alpha, \mu, \Delta, \delta, \beta)$ -Parametrization . . . . .	7
4.4. Switching between different parametrizations . . . . .	7
4.5. Location and scale invariant parametrizations . . . . .	8
4.6. Numerical implementation . . . . .	8
<b>5. Fitting generalized hyperbolic distributions to data</b>	<b>8</b>
5.1. EM-Scheme . . . . .	9
5.2. MCECM estimation . . . . .	9
<b>6. Applications, comments and references</b>	<b>10</b>
<b>A. Shape of the univariate generalized hyperbolic distribution</b>	<b>12</b>
<b>B. Modified Bessel function of the third kind</b>	<b>13</b>
<b>C. Generalized Inverse Gaussian distribution</b>	<b>15</b>
C.1. Gamma distribution . . . . .	15
C.2. Inverse gamma distribution . . . . .	16
<b>D. Densities of the special cases of the GH distribution</b>	<b>17</b>
D.1. Student-t distribution . . . . .	17
D.2. Variance gamma distribution . . . . .	18
<b>E. Conditional density of the mixing variable <math>W</math></b>	<b>18</b>
E.1. Generalized hyperbolic, hyperbolic and NIG distribution . . . . .	18
E.2. Student-t distribution . . . . .	19
E.3. Variance gamma distribution . . . . .	19
<b>F. Distribution objects</b>	<b>19</b>
<b>G. Constructors</b>	<b>20</b>

<b>H. Examples</b>	<b>20</b>
H.1. Initializing distribution objects . . . . .	20
H.2. Object-oriented approach . . . . .	21
H.3. Fitting generalized hyperbolic distributions to data . . . . .	22

## 1. Introduction

The origin of this package goes back to the first authors' years at RiskLab, when he worked together with Alexander McNeil to develop an algorithm for fitting multivariate generalized hyperbolic (GH) distributions. Accordingly, the functionality of this package partly overlaps McNeil's library QRMlib (McNeil, 2005). However, there are quite some differences in the implementation. From the user's point of view, an important difference may be that one can choose between different parametrizations. In addition, with the present library it is possible to fit multivariate as well as univariate generalized hyperbolic distributions in a unified framework. Furthermore, the package **ghyp** provides a much richer set of functionality, including many plotting resources and graphical diagnostic tools as well as various utilities as the likelihood-ratio test, Akaike information based model selection, and linear transformations for instance. In general, we put emphasis on financial application and provide several risk and performance measurement functionalities and a portfolio optimization routine. The finance related facilities are not described in this document, but in the reference manual of the package **ghyp**. This document is primarily intended to give the necessary information and formulas to deal with univariate and multivariate generalized hyperbolic distributions. It focuses on (i) different parametrizations, (ii) fitting multivariate GH distributions, and (iii) special cases of the GH distribution.

Please note that this document is by far not complete. A nice reference for univariate GH distributions and the generalized inverse Gaussian (GIG) distribution is Prause (1999) and Paoletta (2007). An instructive guide for multivariate GH distributions is McNeil, Frey, and Embrechts (2005), where one can find also some empirical work and applications of the GH distribution in finance.

## 2. Definition

Facts about generalized hyperbolic (GH) distributions are cited according to McNeil, Frey, and Embrechts (2005) chapter 3.2.

The random vector  $\mathbf{X}$  is said to have a multivariate GH distribution if

$$\mathbf{X} \stackrel{d}{=} \mu + W\gamma + \sqrt{W}AZ \quad (2.1)$$

where

- (i)  $\mathbf{Z} \sim N_k(\mathbf{0}, I_k)$
- (ii)  $A \in \mathbb{R}^{d \times k}$
- (iii)  $\mu, \gamma \in \mathbb{R}^d$
- (iv)  $W \geq 0$  is a scalar-valued random variable which is independent of  $\mathbf{Z}$  and has a Generalized Inverse Gaussian distribution, written  $GIG(\lambda, \chi, \psi)$  (cf. appendix C).

Note that there are at least five alternative definitions leading to different parametrizations. In section 4 we will present the parametrizations which are implemented in the package **ghyp**. The above definition, though intuitive, acquires an identifiability problem which will be described in section 4.1. The parameters of a GH distribution given by the above definition admit the following interpretation:

- $\lambda, \chi, \psi$  determine the shape of the distribution. That is, how much weight is assigned to the tails and to the center. In general, the larger those parameters the closer the distribution is to the normal distribution.
- $\mu$  is the location parameter.
- $\Sigma = AA'$  is the dispersion-matrix.
- $\gamma$  is the skewness parameter. If  $\gamma = 0$ , then the distribution is symmetric around  $\mu$ .

Observe that the conditional distribution of  $\mathbf{X}|W = w$  is normal,

$$\mathbf{X}|W = w \sim \mathcal{N}_d(\mu + w\gamma, w\Sigma), \quad (2.2)$$

where  $\Sigma = AA'$ .

### 2.1. Expected value and variance

The expected value and the variance are given by

$$\mathbb{E}[\mathbf{X}] = \mu + \mathbb{E}[W]\gamma \quad (2.3)$$

$$\begin{aligned} \text{var}(\mathbf{X}) &= \mathbb{E}[\text{cov}(\mathbf{X}|W)] + \text{cov}(\mathbb{E}[\mathbf{X}|W]) \\ &= \text{var}(W)\gamma\gamma' + \mathbb{E}[W]\Sigma. \end{aligned} \quad (2.4)$$

### 2.2. Density

Since the conditional distribution of  $\mathbf{X}$  given  $W$  is Gaussian with mean  $\mu + W\gamma$  and variance  $W\Sigma$  the GH density can be found by mixing  $\mathbf{X}|W$  with respect to  $W$ .

$$\begin{aligned} f_{\mathbf{X}}(\mathbf{x}) &= \int_0^\infty f_{\mathbf{X}|W}(\mathbf{x}|w) f_W(w) dw \\ &= \int_0^\infty \frac{e^{(\mathbf{x}-\mu)'\Sigma^{-1}\gamma}}{(2\pi)^{\frac{d}{2}}|\Sigma|^{\frac{1}{2}}w^{\frac{d}{2}}} \exp\left\{-\frac{Q(\mathbf{x})}{2w} - \frac{\gamma'\Sigma^{-1}\gamma}{2/w}\right\} f_W(w) dw \\ &= \frac{(\sqrt{\psi/\chi})^\lambda (\psi + \gamma'\Sigma^{-1}\gamma)^{\frac{d}{2}-\lambda}}{(2\pi)^{\frac{d}{2}}|\Sigma|^{\frac{1}{2}}K_\lambda(\sqrt{\chi\psi})} \times \frac{K_{\lambda-\frac{d}{2}}(\sqrt{(\chi + Q(\mathbf{x}))(\psi + \gamma'\Sigma^{-1}\gamma)}) e^{(\mathbf{x}-\mu)'\Sigma^{-1}\gamma}}{(\sqrt{(\chi + Q(\mathbf{x}))(\psi + \gamma'\Sigma^{-1}\gamma)})^{\frac{d}{2}-\lambda}}, \end{aligned} \quad (2.5)$$

where the relation (B.2) for the modified Bessel function of the third kind  $K_\lambda(\cdot)$  is used (cf. section B) and  $Q(\mathbf{x})$  denotes the mahalanobis distance  $Q(\mathbf{x}) = (\mathbf{x} - \mu)'\Sigma^{-1}(\mathbf{x} - \mu)$ . The domain of variation of the parameters  $\lambda, \chi$  and  $\psi$  is given in section 4 and appendix C.

### 2.3. Moment generating function

An appealing property of normal mixtures is that the moment generating function is easily calculated once the moment generating function of the mixture is known. Based on equation (C.2) we obtain the moment generating function of a GH distributed random variable  $X$  as

$$\begin{aligned} \mathbb{M}_{GH}(t) &= \mathbb{E}[\mathbb{E}[\exp\{\mathbf{t}'\mathbf{X}\} | W]] = e^{\mathbf{t}'\mu} \mathbb{E}[\exp\{W(\mathbf{t}'\gamma + 1/2 \mathbf{t}'\Sigma\mathbf{t})\}] \\ &= e^{\mathbf{t}'\mu} \left( \frac{\psi}{\psi - 2\mathbf{t}'\gamma - \mathbf{t}'\Sigma\mathbf{t}} \right)^{\lambda/2} \frac{K_\lambda(\sqrt{\psi(\chi - 2\mathbf{t}'\gamma - \mathbf{t}'\Sigma\mathbf{t})})}{K_\lambda(\sqrt{\chi\psi})}, \quad \chi \geq 2\mathbf{t}'\gamma + \mathbf{t}'\Sigma\mathbf{t}. \end{aligned}$$

For moment generating functions of the special cases of the GH distribution we refer to Prause (1999) and Paoletta (2007).

## 2.4. Linear transformations

The GH class is closed under linear transformations:

**Proposition 1.** *If  $\mathbf{X} \sim GH_d(\lambda, \chi, \psi, \mu, \Sigma, \gamma)$  and  $\mathbf{Y} = B\mathbf{X} + \mathbf{b}$ , where  $B \in \mathbb{R}^{k \times d}$  and  $\mathbf{b} \in \mathbb{R}^k$ , then  $Y \sim GH_k(\lambda, \chi, \psi, B\mu + \mathbf{b}, B\Sigma B', B\gamma)$ .*

*Proof.* The characteristic function of  $\mathbf{X}$  is

$$\phi_X(\mathbf{t}) = \mathbb{E} \left[ \mathbb{E} \left[ e^{i\mathbf{t}'\mathbf{X}} | W \right] \right] = \mathbb{E} \left[ e^{i\mathbf{t}'(\mu + W\gamma) - 1/2 W \mathbf{t}' \Sigma \mathbf{t}} \right] = e^{i\mathbf{t}'\mu} \hat{H}(1/2 \mathbf{t}' \Sigma \mathbf{t} - i\mathbf{t}'\gamma),$$

where  $\hat{H}(\theta) = \int_0^\infty e^{-\theta v} dH(v)$  denotes the Laplace-Stieltjes transform of the distribution function  $H$  of  $W$ . Let  $\mathbf{Y} = B\mathbf{X} + \mathbf{b}$ . The characteristic function is then

$$\begin{aligned} \phi_Y(t) &= \mathbb{E} \left[ \mathbb{E} \left[ e^{i\mathbf{t}'(B\mathbf{X} + \mathbf{b})} | W \right] \right] = \mathbb{E} \left[ e^{i\mathbf{t}'(B(\mu + W\gamma) + \mathbf{b}) - 1/2 W \mathbf{t}' B \Sigma B' \mathbf{t}} \right] \\ &= e^{i\mathbf{t}'(B\mu + \mathbf{b})} \hat{H}(1/2 \mathbf{t}' B \Sigma B' \mathbf{t} - i\mathbf{t}' B \gamma). \end{aligned}$$

Therefore,  $B\mathbf{X} + \mathbf{b} \sim GH_k(\lambda, \chi, \psi, B\mu + \mathbf{b}, B\Sigma B', B\gamma)$ . □

## 3. Special cases of the generalized hyperbolic distribution

The GH distribution contains several special cases known under special names.

- If  $\lambda = \frac{d+1}{2}$  the name generalized is dropped and we have a multivariate *hyperbolic (hyp)* distribution. The univariate margins are still GH distributed. Inversely, when  $\lambda = 1$  we get a multivariate GH distribution with hyperbolic margins.
- If  $\lambda = -\frac{1}{2}$  the distribution is called *Normal Inverse Gaussian (NIG)*.
- If  $\chi = 0$  and  $\lambda > 0$  one gets a limiting case which is known amongst others as *Variance Gamma (VG)* distribution.
- If  $\psi = 0$  and  $\lambda < 0$  the *generalized hyperbolic Student-t* distribution is obtained (called simply *Student-t* in what follows).

Further information about the special cases and the necessary formulas to fit these distributions to multivariate data can be found in the appendixes C and D. The parameter constraints for the special cases in different parametrizations are described in the following section.

## 4. Parametrization

There are several alternative parametrizations for the GH distribution. In the R package **ghyp** the user can choose between three of them. There exist further parametrizations which are not implemented and not mentioned here. For these parametrizations we refer to Prause (1999) and Paoletta (2007).

Appendix H.1 explains how to use the parametrizations implemented in the package **ghyp**. Table 4.1 describes the parameter ranges for each parametrization and each special case. Clearly, the dispersion matrices  $\Sigma$  and  $\Delta$  have to fulfill the usual conditions for covariance matrices, i.e., symmetry and positive definiteness as well as full rank. Appendix H.1 also gives a table where the constructor functions for each combination of distribution and parametrization are listed.

	$(\lambda, \chi, \psi, \mu, \Sigma, \gamma)$ -Parametrization					
	$\lambda$	$\chi$	$\psi$	$\mu$	$\Sigma$	$\gamma$
ghyp	$\lambda \in \mathbb{R}$	$\chi > 0$	$\psi > 0$	$\mu \in \mathbb{R}^d$	$\Sigma \in \mathbb{R}^\Sigma$	$\gamma \in \mathbb{R}^d$
hyp	$\lambda = \frac{d+1}{2}$	$\chi > 0$	$\psi > 0$	$\mu \in \mathbb{R}^d$	$\Sigma \in \mathbb{R}^\Sigma$	$\gamma \in \mathbb{R}^d$
NIG	$\lambda = -\frac{1}{2}$	$\chi > 0$	$\psi > 0$	$\mu \in \mathbb{R}^d$	$\Sigma \in \mathbb{R}^\Sigma$	$\gamma \in \mathbb{R}^d$
t	$\lambda < 0$	$\chi > 0$	$\psi = 0$	$\mu \in \mathbb{R}^d$	$\Sigma \in \mathbb{R}^\Sigma$	$\gamma \in \mathbb{R}^d$
VG	$\lambda > 0$	$\chi = 0$	$\psi > 0$	$\mu \in \mathbb{R}^d$	$\Sigma \in \mathbb{R}^\Sigma$	$\gamma \in \mathbb{R}^d$

	$(\lambda, \bar{\alpha}, \mu, \Sigma, \gamma)$ -Parametrization				
	$\lambda$	$\bar{\alpha}$	$\mu$	$\Sigma$	$\gamma$
ghyp	$\lambda \in \mathbb{R}$	$\bar{\alpha} > 0$	$\mu \in \mathbb{R}^d$	$\Sigma \in \mathbb{R}^\Sigma$	$\gamma \in \mathbb{R}^d$
hyp	$\lambda = \frac{d+1}{2}$	$\bar{\alpha} > 0$	$\mu \in \mathbb{R}^d$	$\Sigma \in \mathbb{R}^\Sigma$	$\gamma \in \mathbb{R}^d$
NIG	$\lambda = \frac{1}{2}$	$\bar{\alpha} > 0$	$\mu \in \mathbb{R}^d$	$\Sigma \in \mathbb{R}^\Sigma$	$\gamma \in \mathbb{R}^d$
t	$\lambda = -\frac{\nu}{2} < -1$	$\bar{\alpha} = 0$	$\mu \in \mathbb{R}^d$	$\Sigma \in \mathbb{R}^\Sigma$	$\gamma \in \mathbb{R}^d$
VG	$\lambda > 0$	$\bar{\alpha} = 0$	$\mu \in \mathbb{R}^d$	$\Sigma \in \mathbb{R}^\Sigma$	$\gamma \in \mathbb{R}^d$

	$(\lambda, \alpha, \mu, \Delta, \delta, \beta)$ -Parametrization					
	$\lambda$	$\alpha$	$\delta$	$\mu$	$\Delta$	$\beta$
ghyp	$\lambda \in \mathbb{R}$	$\alpha > 0$	$\delta > 0$	$\mu \in \mathbb{R}^d$	$\Delta \in \mathbb{R}^\Delta$	$\beta \in \{\mathbf{x} \in \mathbb{R}^d : \alpha^2 - \mathbf{x}'\Delta\mathbf{x} > 0\}$
hyp	$\lambda = \frac{d+1}{2}$	$\alpha > 0$	$\delta > 0$	$\mu \in \mathbb{R}^d$	$\Delta \in \mathbb{R}^\Delta$	$\beta \in \{\mathbf{x} \in \mathbb{R}^d : \alpha^2 - \mathbf{x}'\Delta\mathbf{x} > 0\}$
NIG	$\lambda = -\frac{1}{2}$	$\alpha > 0$	$\delta > 0$	$\mu \in \mathbb{R}^d$	$\Delta \in \mathbb{R}^\Delta$	$\beta \in \{\mathbf{x} \in \mathbb{R}^d : \alpha^2 - \mathbf{x}'\Delta\mathbf{x} > 0\}$
t	$\lambda < 0$	$\alpha = \sqrt{\beta'\Delta\beta}$	$\delta > 0$	$\mu \in \mathbb{R}^d$	$\Delta \in \mathbb{R}^\Delta$	$\beta \in \mathbb{R}^d$
VG	$\lambda > 0$	$\alpha > 0$	$\delta = 0$	$\mu \in \mathbb{R}^d$	$\Delta \in \mathbb{R}^\Delta$	$\beta \in \{\mathbf{x} \in \mathbb{R}^d : \alpha^2 - \mathbf{x}'\Delta\mathbf{x} > 0\}$

**Table 4.1:** The domain of variation for the parameters of the GH distribution and some of its special cases for different parametrizations. We denote the set of all feasible covariance matrices in  $\mathbb{R}^{d \times d}$  with  $\mathbb{R}^\Sigma$ . Furthermore, let  $\mathbb{R}^\Delta = \{A \in \mathbb{R}^\Sigma : |A| = 1\}$ .

#### 4.1. $(\lambda, \chi, \psi, \mu, \Sigma, \gamma)$ -Parametrization

The  $(\lambda, \chi, \psi, \mu, \Sigma, \gamma)$ -parametrization is obtained as the normal mean-variance mixture distribution when  $W \sim GIG(\lambda, \chi, \psi)$ . This parametrization has a drawback of an identification problem. Indeed, the distributions  $\text{GH}_d(\lambda, \chi, \psi, \mu, \Sigma, \gamma)$  and  $\text{GH}_d(\lambda, \chi/k, k\psi, \mu, k\Sigma, k\gamma)$  are identical for any  $k > 0$ . Therefore, an identifying problem occurs when we start to fit the parameters of a GH distribution to data. This problem may be solved by introducing a suitable constraint. One possibility is to require the determinant of the dispersion matrix  $\Sigma$  to be 1.

#### 4.2. $(\lambda, \bar{\alpha}, \mu, \Sigma, \gamma)$ -Parametrization

There is a more elegant way to eliminate the degree of freedom than constraining the determinant of the dispersion matrix  $\Sigma$  to 1. We simply require the expected value of the generalized inverse Gaussian distributed mixing variable  $W$  to be 1 (cf. C). This makes the interpretation of the skewness parameters  $\gamma$  easier and in addition, the fitting procedure becomes faster (cf. 5.1).

Note that the expected value of a GIG distributed random variable exists as long as  $\psi > 0$ . If  $\psi \rightarrow 0$  the GIG distribution approaches the Inverse Gamma distribution (the mixing distribution belonging to a Student-t distribution) for which the expectation only exists if  $\gamma < -1$  (cf. appendix C and D.1).

We define

$$\mathbb{E}[W] = \sqrt{\frac{\chi}{\psi}} \frac{K_{\lambda+1}(\sqrt{\chi\psi})}{K_{\lambda}(\sqrt{\chi\psi})} = 1. \quad (4.1)$$

and set

$$\bar{\alpha} = \sqrt{\chi\psi}. \quad (4.2)$$

It follows that

$$\psi = \bar{\alpha} \frac{K_{\lambda+1}(\bar{\alpha})}{K_{\lambda}(\bar{\alpha})} \text{ and } \chi = \frac{\bar{\alpha}^2}{\psi} = \bar{\alpha} \frac{K_{\lambda}(\bar{\alpha})}{K_{\lambda+1}(\bar{\alpha})}. \quad (4.3)$$

Note that whenever  $\lambda = -0.5$  (NIG distribution) we have that  $\psi = \bar{\alpha}$  and  $\chi = \bar{\alpha}$ . This is because of the symmetry of the Bessel function (cf. equation B.5).

The drawback of the  $(\lambda, \bar{\alpha}, \mu, \Sigma, \gamma)$ -parametrization is that it does not exist in the case  $\bar{\alpha} = 0$  and  $\lambda \in [-1, 0]$ , which corresponds to a Student-t distribution with non-existing variance. Note that the  $(\lambda, \bar{\alpha}, \mu, \Sigma, \gamma)$ -parametrization yields to a slightly different parametrization for the special case of a Student-t distribution (cf. section D.1 for details). The limit of the equations (4.3) as  $\bar{\alpha} \downarrow 0$  can be found in (D.3) and (D.9).

### 4.3. $(\lambda, \alpha, \mu, \Delta, \delta, \beta)$ -Parametrization

When the GH distribution was introduced in Barndorff-Nielsen (1977), the following parametrization for the multivariate case was used:

$$f_{\mathbf{X}}(\mathbf{x}) = \frac{(\alpha^2 - \boldsymbol{\beta}'\Delta\boldsymbol{\beta})^{\lambda/2}}{(2\pi)^{\frac{d}{2}} \sqrt{|\Delta|} \alpha^{\lambda - \frac{d}{2}} \delta^{\lambda} K_{\lambda}(\delta \sqrt{\alpha^2 - \boldsymbol{\beta}'\Delta\boldsymbol{\beta}})} \times \frac{K_{\lambda - \frac{d}{2}}(\alpha \sqrt{\delta^2 + (\mathbf{x} - \mu)' \Delta^{-1} (\mathbf{x} - \mu)}) e^{\boldsymbol{\beta}'(\mathbf{x} - \mu)}}{(\sqrt{\delta^2 + (\mathbf{x} - \mu)' \Delta^{-1} (\mathbf{x} - \mu)})^{\frac{d}{2} - \lambda}}. \quad (4.4)$$

Similar to the  $(\lambda, \chi, \psi, \mu, \Sigma, \gamma)$  parametrization, there is an identifying problem which can be solved by constraining the determinant of  $\Delta$  to 1. The mixture representation belonging to this parametrization is

$$X|W = w \sim \mathcal{N}_d(\mu + w\boldsymbol{\beta}\Delta, w\Delta) \quad (4.5)$$

$$W \sim GIG(\lambda, \delta^2, \alpha^2 - \boldsymbol{\beta}'\Delta\boldsymbol{\beta}). \quad (4.6)$$

In the univariate case the above expression reduces to

$$f_{\mathbf{X}}(\mathbf{x}) = \frac{(\alpha^2 - \beta^2)^{\lambda/2}}{\sqrt{2\pi} \alpha^{\lambda - \frac{1}{2}} \delta^{\lambda} K_{\lambda}(\delta \sqrt{\alpha^2 - \beta^2})} \times \frac{K_{\lambda - \frac{1}{2}}(\alpha \sqrt{\delta^2 + (x - \mu)^2})}{(\sqrt{\delta^2 + (x - \mu)^2})^{\frac{1}{2} - \lambda}} e^{\beta(x - \mu)}, \quad (4.7)$$

which is the most widely used parametrization of the GH distribution in literature.

### 4.4. Switching between different parametrizations

The following formulas can be used to switch between the  $(\lambda, \bar{\alpha}, \mu, \Sigma, \gamma)$ ,  $(\lambda, \chi, \psi, \mu, \Sigma, \gamma)$ , and the  $(\lambda, \alpha, \mu, \Delta, \delta, \beta)$ -parametrization. The parameters  $\lambda$  and  $\mu$  remain the same, regardless of the parametrization.

The way to obtain the  $(\lambda, \alpha, \mu, \Delta, \delta, \beta)$ -parametrization from the  $(\lambda, \bar{\alpha}, \mu, \Sigma, \gamma)$ -parametrization leads over the  $(\lambda, \chi, \psi, \mu, \Sigma, \gamma)$ -parametrization:

$$(\lambda, \bar{\alpha}, \mu, \Sigma, \gamma) \quad \Leftrightarrow \quad (\lambda, \chi, \psi, \mu, \Sigma, \gamma) \quad \Leftrightarrow \quad (\lambda, \alpha, \mu, \Delta, \delta, \beta)$$



$(\lambda, \bar{\alpha}, \mu, \Sigma, \gamma) \rightarrow (\lambda, \chi, \psi, \mu, \Sigma, \gamma)$ : Use the relations in (4.3) to obtain  $\chi$  and  $\psi$ . The parameters  $\Sigma$  and  $\gamma$  remain the same.

$(\lambda, \chi, \psi, \mu, \Sigma, \gamma) \rightarrow (\lambda, \bar{\alpha}, \mu, \Sigma, \gamma)$ : Set  $k = \sqrt{\frac{\chi}{\psi}} \frac{K_{\lambda+1}(\sqrt{\chi\psi})}{K_{\lambda}(\sqrt{\chi\psi})}$ .

$$\bar{\alpha} = \sqrt{\chi\psi}, \quad \Sigma \equiv k\Sigma, \quad \gamma \equiv k\gamma \quad (4.8)$$

$(\lambda, \chi, \psi, \mu, \Sigma, \gamma) \rightarrow (\lambda, \alpha, \mu, \Delta, \delta, \beta)$ :

$$\begin{aligned} \Delta &= |\Sigma|^{-\frac{1}{d}} \Sigma, \quad \beta = \Sigma^{-1} \gamma \\ \delta &= \sqrt{\chi |\Sigma|^{\frac{1}{d}}}, \quad \alpha = \sqrt{|\Sigma|^{-\frac{1}{d}} (\psi + \gamma' \Sigma^{-1} \gamma)} \end{aligned} \quad (4.9)$$

$(\lambda, \alpha, \mu, \Delta, \delta, \beta) \rightarrow (\lambda, \chi, \psi, \mu, \Sigma, \gamma)$ :

$$\Sigma = \Delta, \quad \gamma = \Delta\beta, \quad \chi = \delta^2, \quad \psi = \alpha^2 - \beta' \Delta \beta. \quad (4.10)$$

#### 4.5. Location and scale invariant parametrizations

All parametrizations mentioned above can be modified to location and scale invariant parametrizations. For a  $d$ -variate GH random variable  $\mathbf{X}$  there may occur problems if it is linearly transformed ( $\mathbf{Y} = B\mathbf{X} + \mathbf{b}$ ) while  $B$  affects the dimensionality (i.e.  $B \notin \mathbb{R}^{d \times d}$ ). Thus, let's consider the univariate case for a moment.

2.4 Let  $\sigma^2 = \Sigma$ . For the  $(\lambda, \chi, \psi, \mu, \Sigma, \gamma)$  and  $(\lambda, \bar{\alpha}, \mu, \Sigma, \gamma)$  parametrization one can define  $\bar{\gamma} = \sigma\gamma$ . The density reads

$$f_X(x) = \frac{(\sqrt{\psi/\chi})^\lambda (\psi + \bar{\gamma}^2)^{\frac{1}{2}-\lambda}}{\sqrt{2\pi}\sigma K_\lambda(\sqrt{\chi\psi})} \times \frac{K_{\lambda-\frac{1}{2}}(\sqrt{(\chi + q(x)^2)(\psi + \bar{\gamma}^2)}) e^{q(x)\bar{\gamma}}}{(\sqrt{(\chi + q(x)^2)(\psi + \bar{\gamma}^2)})^{\frac{1}{2}-\lambda}}, \quad (4.11)$$

where  $q(x) = (x - \mu)/\sigma$ .

In case of the  $(\lambda, \alpha, \mu, \Delta, \delta, \beta)$  parametrization one can define  $\bar{\alpha} = \alpha\delta$  and  $\bar{\beta} = \beta\delta$  (see the appendix of Prause (1999)) to get the invariant density

$$f_X(x) = \frac{(\bar{\alpha}^2 - \bar{\beta}^2)^{\lambda/2}}{\sqrt{2\pi} \bar{\alpha}^{\lambda-\frac{1}{2}} \delta K_\lambda(\sqrt{\bar{\alpha}^2 - \bar{\beta}^2})} \times \frac{K_{\lambda-\frac{1}{2}}(\bar{\alpha}\sqrt{1 + q(x)^2})}{(\sqrt{1 + q(x)^2})^{\frac{1}{2}-\lambda}} e^{\beta q(x)}. \quad (4.12)$$

#### 4.6. Numerical implementation

Internally, the package **ghyp** uses the  $(\lambda, \chi, \psi, \mu, \Sigma, \gamma)$ -parametrization. However, fitting is done in the  $(\lambda, \bar{\alpha}, \mu, \Sigma, \gamma)$ -parametrization since this parametrization does not necessitate additional constraints to eliminate the redundant degree of freedom. Consequently, what cannot be represented by the  $(\lambda, \bar{\alpha}, \mu, \Sigma, \gamma)$ -parametrization cannot be fitted (cf. section 4.2).

### 5. Fitting generalized hyperbolic distributions to data

Numerical optimizers can be used to fit univariate GH distributions to data by means of maximum likelihood estimation. Multivariate GH distributions can be fitted with expectation-maximization (EM) type algorithms (see Dempster et al. (1977) and Meng and Rubin (1993)).

### 5.1. EM-Scheme

Assume we have iid data  $\mathbf{x}_1, \dots, \mathbf{x}_n$  and parameters represented by  $\Theta = (\lambda, \bar{\alpha}, \mu, \Sigma, \gamma)$ . The problem is to maximize

$$\ln L(\Theta; \mathbf{x}_1, \dots, \mathbf{x}_n) = \sum_{i=1}^n \ln f_{\mathbf{X}}(\mathbf{x}_i; \Theta). \quad (5.1)$$

This problem is not easy to solve due to the number of parameters and necessity of maximizing over covariance matrices. We can proceed by introducing an augmented likelihood function

$$\ln \tilde{L}(\Theta; \mathbf{x}_1, \dots, \mathbf{x}_n, w_1, \dots, w_n) = \sum_{i=1}^n \ln f_{\mathbf{X}|W}(\mathbf{x}_i | w_i; \mu, \Sigma, \gamma) + \sum_{i=1}^n \ln f_W(w_i; \lambda, \bar{\alpha}) \quad (5.2)$$

and spend the effort on the estimation of the latent mixing variables  $w_i$  coming from the mixture representation (2.2). This is where the EM algorithm comes into play.

**E-step:** Calculate the conditional expectation of the likelihood function (5.2) given the data  $\mathbf{x}_1, \dots, \mathbf{x}_n$  and the current estimates of parameters  $\Theta^{[k]}$ . This results in the objective function

$$Q(\Theta; \Theta^{[k]}) = \mathbb{E} \left[ \ln \tilde{L}(\Theta; \mathbf{x}_1, \dots, \mathbf{x}_n, w_1, \dots, w_n) | \mathbf{x}_1, \dots, \mathbf{x}_n; \Theta^{[k]} \right]. \quad (5.3)$$

**M-step:** Maximize the objective function with respect to  $\Theta$  to obtain the next set of estimates  $\Theta^{[k+1]}$ .

Alternating between these steps yields to the maximum likelihood estimation of the parameter set  $\Theta$ .

In practice, performing the E-Step means maximizing the second summand of (5.2) numerically. The log density of the GIG distribution (cf. C.1) is

$$\ln f_W(w) = \frac{\lambda}{2} \ln(\psi/\chi) - \ln(2K_\lambda \sqrt{\chi\psi}) + (\lambda - 1) \ln w - \frac{\chi}{2} \frac{1}{w} - \frac{\psi}{2} w. \quad (5.4)$$

When using the  $(\lambda, \bar{\alpha})$ -parametrization this problem is of dimension two instead of three as it is in the  $(\lambda, \chi, \psi)$ -parametrization. As a consequence the performance increases.

Since the  $w_i$ 's are latent one has to replace  $w$ ,  $1/w$  and  $\ln w$  with the respective expected values in order to maximize the log likelihood function. Let

$$\eta_i^{[k]} := \mathbb{E} [w_i | \mathbf{x}_i; \Theta^{[k]}], \quad \delta_i^{[k]} := \mathbb{E} [w_i^{-1} | \mathbf{x}_i; \Theta^{[k]}], \quad \xi_i^{[k]} := \mathbb{E} [\ln w_i | \mathbf{x}_i; \Theta^{[k]}]. \quad (5.5)$$

We have to find the conditional density of  $w_i$  given  $\mathbf{x}_i$  to calculate these quantities (cf. (E.1)).

### 5.2. MCECM estimation

In the R implementation a modified EM scheme is used, which is called multi-cycle, expectation, conditional estimation (MCECM) algorithm (Meng and Rubin, 1993; McNeil, Frey, and Embrechts, 2005). The different steps of the MCECM algorithm are sketched as follows:

- (1) Select reasonable starting values for  $\Theta^{[k]}$ . For example  $\lambda = 1$ ,  $\bar{\alpha} = 1$ ,  $\mu$  is set to the sample mean,  $\Sigma$  to the sample covariance matrix and  $\gamma$  to a zero skewness vector.
- (2) Calculate  $\chi^{[k]}$  and  $\psi^{[k]}$  as a function of  $\bar{\alpha}^{[k]}$  using (4.3).
- (3) Use (5.5), (C.3) and (E.1) to calculate the weights  $\eta_i^{[k]}$  and  $\delta_i^{[k]}$ . Average the weights to get

$$\bar{\eta}^{[k]} = \frac{1}{n} \sum_{i=1}^n \eta_i^{[k]} \text{ and } \bar{\delta}^{[k]} = \frac{1}{n} \sum_{i=1}^n \delta_i^{[k]}. \quad (5.6)$$

- (4) If an asymmetric model is to be fitted set  $\gamma$  to  $\mathbf{0}$ , else set

$$\gamma^{[k+1]} = \frac{1}{n} \frac{\sum_{i=1}^n \delta_i^{[k]} (\bar{\mathbf{x}} - \mathbf{x}_i)}{\bar{\eta}^{[k]} \bar{\delta}^{[k]} - 1}. \quad (5.7)$$

- (5) Update  $\mu$  and  $\Sigma$ :

$$\mu^{[k+1]} = \frac{1}{n} \frac{\sum_{i=1}^n \delta_i^{[k]} \mathbf{x}_i - \gamma^{[k+1]}}{\bar{\delta}^{[k]}} \quad (5.8)$$

$$\Sigma^{[k+1]} = \frac{1}{n} \sum_{i=1}^n \delta_i^{[k]} (\mathbf{x}_i - \mu^{[k+1]})(\mathbf{x}_i - \mu^{[k+1]})' - \bar{\eta}^{[k]} \gamma^{[k+1]} \gamma^{[k+1]}' \quad (5.9)$$

- (6) Set  $\Theta^{[k,2]} = (\lambda^{[k]}, \bar{\alpha}^{[k]}, \mu^{[k+1]}, \Sigma^{[k+1]}, \gamma^{[k+1]})$  and calculate weights  $\eta_i^{[k,2]}$ ,  $\delta_i^{[k,2]}$  and  $\xi_i^{[k,2]}$  using (5.5), (C.4) and (C.3).
- (7) Maximize the second summand of (5.2) with density (5.4) with respect to  $\lambda$ ,  $\chi$  and  $\psi$  to complete the calculation of  $\Theta^{[k,2]}$  and go back to step (2). Note that the objective function must calculate  $\chi$  and  $\psi$  in dependence of  $\lambda$  and  $\bar{\alpha}$  using relation (4.3).

## 6. Applications, comments and references

Even though the GH distribution was initially invented to study the distribution of the logarithm of particle sizes, we will focus on applications of the GH distribution family in finance and risk measurement.

We have seen above that the GH distribution is very flexible in the sense that it nests several other distributions such as the Student-t (cf. D.1). To give some references and applications of the GH distribution let us first summarize some of its important properties. Beside of the above mentioned flexibility, three major facts led to the popularity of GH distribution family in finance:

- (1) The GH distribution features both *fat tails* and *skewness*. These properties account for some of the frequently reported stylized facts of financial returns but also of financial return volatility (hyperbolic GARCH effects).
- (2) The GH family is naturally extended to multivariate distributions<sup>1</sup>. A multivariate GH distribution does exhibit some kind of *non-linear dependence*, for example *tail-dependence*. This reflects the fact that extremes mostly occur for several risk-drivers

<sup>1</sup>The extension to multivariate distributions is natural because of the mixing structure (see eq. (2.2)).

simultaneously in financial markets. This property is of fundamental importance for risk-management, and can influence for instance portfolio weights.

- (3) The GH distribution is infinitely divisible (cf. Barndorff-Nielsen and Halgreen (1977)). This is a necessary and sufficient condition to build *Lévy processes*. Lévy processes are widespread in finance because of their time-continuity and their ability to model jumps.

Based on these properties one can classify the applications of the GH distributions into the fields *empirical modelling*, *risks and dependence modelling*, *Lévy processes & derivatives*, and *portfolio selection*. In the following, we try to assign papers to each of the classes of applications mentioned above. Rather than giving summaries and conclusions for each paper, we simply cite them and refer the interested reader to the articles and the references therein. Note that some articles deal with special cases of the GH distribution only.

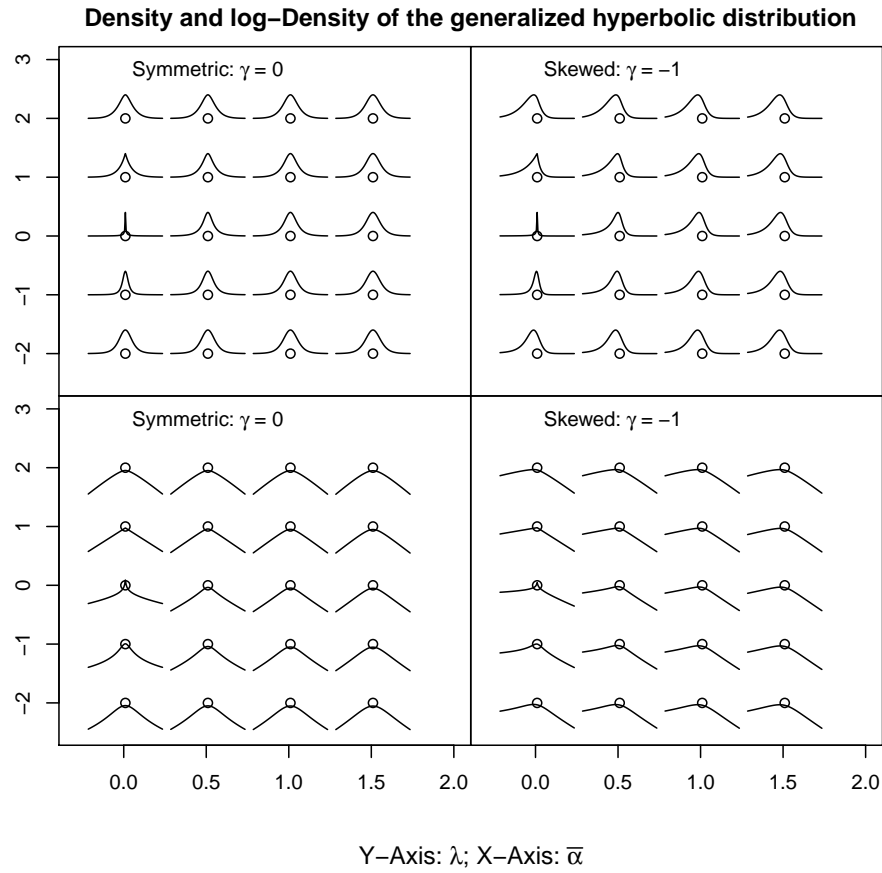
**Empirical modelling:** Eberlein and Keller (1995); Barndorff-Nielsen and Prause (2001); Forsberg and Bollerslev (2002); Davidson (2004); Fergusson and Platen (2006).

**Risk and dependence modelling:** Eberlein et al. (1998); Breymann et al. (2003); McNeil et al. (2005); Chen et al. (2005); Kassberger and Kiesel (2006).

**Lévy processes & derivatives:** Barndorff-Nielsen (1997a,b); Bibby and Renssen (1997); Madan et al. (1998); Raible (2000); Cont and Tankov (2003); Prause (1999).

**Portfolio selection:** Kassberger (2007).

### A. Shape of the univariate generalized hyperbolic distribution



**Figure A.1:** The shape of the univariate generalized hyperbolic density drawn with different shape parameters  $(\lambda, \bar{\alpha})$ . The location and scale parameter  $\mu$  and  $\sigma$  are set to 0 and 1, respectively. The skewness parameter  $\gamma$  is 0 in the left column and  $-1$  in the right column of the graphics array.

## B. Modified Bessel function of the third kind

The modified Bessel function of the third kind appears in the GH as well as in the GIG density (2.5, C.1). This function has the integral representation

$$K_\lambda(x) = \frac{1}{2} \int_0^\infty w^{\lambda-1} \exp \left\{ -\frac{1}{2}x(w + w^{-1}) \right\} dw, \quad x > 0. \quad (\text{B.1})$$

The substitution  $w = x\sqrt{\chi/\psi}$  can be used to obtain the following relation, which facilitates to bring the GH density (2.5) into a closed-form expression.

$$\int_0^\infty w^{\lambda-1} \exp \left\{ -\frac{1}{2} \left( \frac{\chi}{w} + w\psi \right) \right\} dw = 2 \left( \frac{\chi}{\psi} \right)^{\frac{\lambda}{2}} K_\lambda(\sqrt{\chi\psi}) \quad (\text{B.2})$$

When calculating the densities of the special cases of the GH density we can use the asymptotic relations for small arguments  $x$ .

$$K_\lambda(x) \sim \Gamma(\lambda) 2^{\lambda-1} x^{-\lambda} \quad \text{as } x \downarrow 0 \quad \text{and } \lambda > 0 \quad (\text{B.3})$$

and

$$K_\lambda(x) \sim \Gamma(-\lambda) 2^{-\lambda-1} x^\lambda \quad \text{as } x \downarrow 0 \quad \text{and } \lambda < 0. \quad (\text{B.4})$$

(B.4) follows from (B.3) and the observation that the Bessel function is symmetric with respect to the index  $\lambda$ :

$$K_\lambda(x) = K_{-\lambda}(x) \quad (\text{B.5})$$

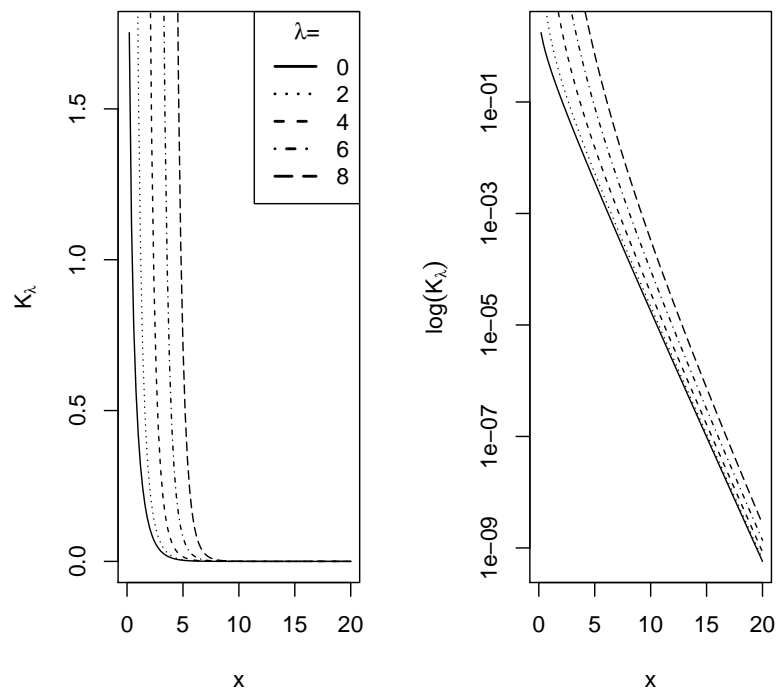
An asymptotic relation for large arguments  $x$  is given by

$$K_\lambda(x) \sim \sqrt{\frac{\pi}{2}} x^{-\frac{1}{2}} e^{-x} \quad \text{as } x \rightarrow \infty. \quad (\text{B.6})$$

Also, for  $\lambda = 0.5$  the Bessel function can be stated explicitly with

$$K_{0.5}(x) = K_{-0.5}(x) = \sqrt{\frac{\pi}{2x}} e^{-x}, \quad x > 0. \quad (\text{B.7})$$

We refer to Abramowitz and Stegun (1972) for further information on Bessel functions.



**Figure B.1:** The modified Bessel function of the third kind drawn with different indices  $\lambda$ .

## C. Generalized Inverse Gaussian distribution

The density of a Generalized Inverse Gaussian (GIG) distribution is given as

$$f_{GIG}(w) = \left(\frac{\psi}{\chi}\right)^{\frac{\lambda}{2}} \frac{w^{\lambda-1}}{2K_{\lambda}(\sqrt{\chi\psi})} \exp\left\{-\frac{1}{2}\left(\frac{\chi}{w} + \psi w\right)\right\}, \quad (\text{C.1})$$

with parameters satisfying

$$\begin{aligned} \chi > 0, \psi &\geq 0, & \lambda < 0 \\ \chi > 0, \psi &> 0, & \lambda = 0 \\ \chi \geq 0, \psi &> 0, & \lambda > 0. \end{aligned}$$

The GIG density nests the inverse Gaussian (IG) density ( $\lambda = -0.5$ ) and contains many special cases as limits, among others the Gamma ( $\Gamma$ ) and Inverse Gamma ( $\Gamma$ ) densities (cf. C.1 and C.2 below).

The moment generating function of the GIG distribution (cf. Paolella (2007)) is determined by

$$\mathbb{M}_{GIG}(t) = \left(\frac{\psi}{\psi - 2t}\right)^{\lambda/2} \frac{K_{\lambda}(\sqrt{\chi(\psi - 2t)})}{K_{\lambda}(\sqrt{\chi\psi})}. \quad (\text{C.2})$$

The  $n$ -th moment of a GIG distributed random variable  $X$  can be found to be

$$\mathbb{E}[X^n] = \left(\frac{\chi}{\psi}\right)^{\frac{n}{2}} \frac{K_{\lambda+n}(\sqrt{\chi\psi})}{K_{\lambda}(\sqrt{\chi\psi})}. \quad (\text{C.3})$$

Furthermore,

$$\mathbb{E}[\ln X] = \frac{d\mathbb{E}[X^{\alpha}]}{d\alpha} \Big|_{\alpha=0}. \quad (\text{C.4})$$

Note that numerical calculations of  $\mathbb{E}[\ln X]$  may be performed with the integral representation as well. In the R package `ghyp` the derivative construction is implemented.

### C.1. Gamma distribution

When  $\chi = 0$  and  $\lambda > 0$  the GIG distribution reduces to the Gamma ( $\Gamma$ ) distribution:

$$\lim_{\chi \downarrow 0} f_{GIG}(x|\lambda, \chi, \psi) = f_{\Gamma}(x|\lambda, 1/2\psi), \quad x \in \mathbb{R}, \lambda > 0, \chi > 0 \quad (\text{C.5})$$

The density of  $X \sim \Gamma(\alpha, \beta)$  is

$$f_{\Gamma}(w) = \frac{\beta^{\alpha}}{\Gamma(\alpha)} w^{\alpha-1} \exp\{-\beta w\}. \quad (\text{C.6})$$

The expected value and the variance are

$$\mathbb{E}[X] = \frac{\beta}{\alpha} \quad \text{and} \quad \text{var}(X) = \frac{\alpha}{\beta^2}, \quad (\text{C.7})$$

respectively. The expected value of the logarithm is  $\mathbb{E}[\ln X] = \psi(\alpha) - \ln(\beta)$  where  $\psi(\cdot)$  is the digamma function. We will see that this value is not needed to fit a multivariate variance gamma distribution (cf. E.3).



## C.2. Inverse gamma distribution

When  $\psi = 0$  and  $\lambda < 0$  the GIG distribution reduces to the Inverse Gamma (IG) distribution:

$$\lim_{\psi \downarrow 0} f_{GIG}(x|\lambda, \chi, \psi) = f_{IG}(x|\lambda, 1/2\chi), \quad x \in \mathbb{R}, \lambda < 0, \psi > 0 \quad (C.8)$$

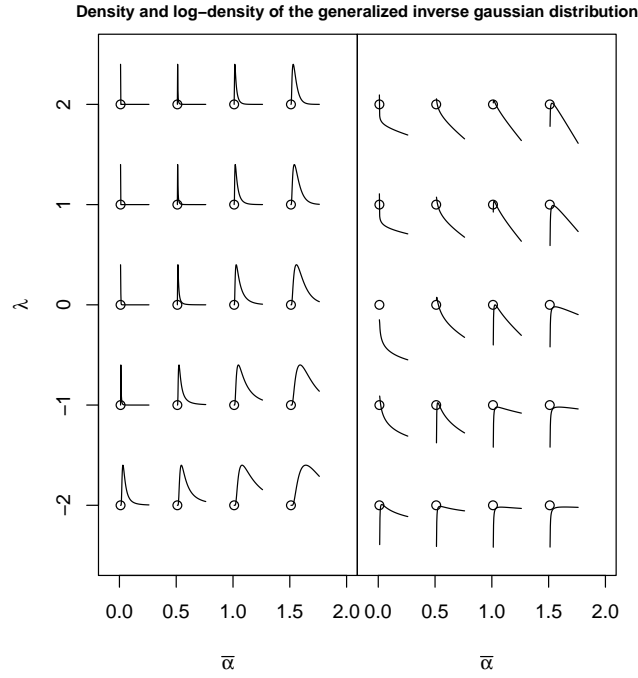
The density of  $X \sim \text{IG}(\alpha, \beta)$  is

$$f_X(w) = \frac{\beta^\alpha}{\Gamma(\alpha)} w^{-\alpha-1} \exp\left\{-\frac{\beta}{w}\right\}. \quad (C.9)$$

The expected value and the variance are

$$\mathbb{E}[X] = \frac{\beta}{\alpha - 1} \quad \text{and} \quad \text{var}(X) = \frac{\beta^2}{(\alpha - 1)^2(\alpha - 2)}, \quad (C.10)$$

and exist provided that  $\alpha > 1$  and  $\alpha > 2$ , respectively. The expected value of the logarithm is  $\mathbb{E}[\ln X] = \ln(\beta) - \psi(\alpha)$ . This value is required in order to fit a symmetric multivariate Student-t distribution by means of the MCECM algorithm (cf. E.2).



**Figure C.1:** The density and the log-density of the generalized inverse gaussian distribution drawn with different shape parameters  $(\lambda, \bar{\alpha})$ . See (4.3) for the transformation from  $\bar{\alpha}$  to  $(\chi, \psi)$ .

## D. Densities of the special cases of the GH distribution

As mentioned in section 3 the GH distribution contains several special and limiting cases. In what follows the densities of the limiting cases are derived. In the case of a hyperbolic or normal NIG we simply fix the parameter  $\lambda$  either to  $(d+1)/2$  or  $-0.5$ .

### D.1. Student-t distribution

With relation (B.4) it can be easily shown that when  $\psi \rightarrow 0$  and  $\lambda < 0$  the density of a GH distribution results in

$$f_{\mathbf{X}}(\mathbf{x}) = \frac{\chi^{-\lambda}(\gamma'\Sigma^{-1}\gamma)^{\frac{d}{2}-\lambda}}{(2\pi)^{\frac{d}{2}}|\Sigma|^{\frac{1}{2}}\Gamma(-\lambda)2^{-\lambda-1}} \times \frac{K_{\lambda-\frac{d}{2}}(\sqrt{(\chi+\mathbf{Q}(\mathbf{x}))\gamma'\Sigma^{-1}\gamma})e^{(\mathbf{x}-\mu)'\Sigma^{-1}\gamma}}{(\sqrt{(\chi+\mathbf{Q}(\mathbf{x}))\gamma'\Sigma^{-1}\gamma})^{\frac{d}{2}-\lambda}}. \quad (\text{D.1})$$

As  $\gamma \rightarrow 0$  we obtain again with relation (B.4) the symmetric multivariate Student-t density

$$f_{\mathbf{X}}(\mathbf{x}) = \frac{\chi^{-\lambda}\Gamma(-\lambda+\frac{d}{2})}{\pi^{\frac{d}{2}}|\Sigma|^{\frac{1}{2}}\Gamma(-\lambda)} \times (\chi+\mathbf{Q}(\mathbf{x}))^{\lambda-\frac{d}{2}}. \quad (\text{D.2})$$

We switch to the Student-t parametrization and set the degree of freedom  $\nu = -2\lambda$ <sup>2</sup>. Because  $\psi = 0$  the transformation of  $\bar{\alpha}$  to  $\chi$  and  $\psi$  (cf. 4.3) reduces to

$$\chi = \bar{\alpha} \frac{K_{\lambda}(\bar{\alpha})}{K_{\lambda+1}(\bar{\alpha})} \xrightarrow{\bar{\alpha} \rightarrow 0} 2(-\lambda-1) = \nu-2. \quad (\text{D.3})$$

Plugging in the values for  $\lambda$  and  $\nu$ , the densities take the form

$$f_{\mathbf{X}}(\mathbf{x}) = \frac{(\nu-2)^{\frac{\nu}{2}}(\gamma'\Sigma^{-1}\gamma)^{\frac{\nu+d}{2}}}{(2\pi)^{\frac{d}{2}}|\Sigma|^{\frac{1}{2}}\Gamma(\frac{\nu}{2})2^{\frac{\nu}{2}-1}} \times \frac{K_{\frac{\nu+d}{2}}(\sqrt{(\nu-2+\mathbf{Q}(\mathbf{x}))\gamma'\Sigma^{-1}\gamma})e^{(\mathbf{x}-\mu)'\Sigma^{-1}\gamma}}{(\sqrt{(\nu-2+\mathbf{Q}(\mathbf{x}))\gamma'\Sigma^{-1}\gamma})^{\frac{\nu+d}{2}}} \quad (\text{D.4})$$

and for the symmetric case as  $\gamma \rightarrow 0$

$$f_{\mathbf{X}}(\mathbf{x}) = \frac{(\nu-2)^{\frac{\nu}{2}}\Gamma(\frac{\nu+d}{2})}{\pi^{\frac{d}{2}}|\Sigma|^{\frac{1}{2}}\Gamma(\frac{\nu}{2})(\nu-2+\mathbf{Q}(\mathbf{x}))^{\frac{\nu+d}{2}}}. \quad (\text{D.5})$$

It is important to note that the variance does not exist in the symmetric case for  $\nu \leq 2$  while for the asymmetric case the variance does not exist for  $\nu \leq 4$ . This is because the variance of an asymmetric GH distribution involves the variance of the mixing distribution. In case of a Student-t distribution, the mixing variable  $w$  is inverse gamma distributed and has finite variance only if  $\beta > 2$  which corresponds to  $\lambda < -2$ , i.e.  $\nu > 4$  (cf. C.10). Alternatively, in the univariate case, this can be seen by the fact that the Student-t density has regularly varying tails. For  $x \rightarrow \infty$ , one obtains

$$f_X(x) = L(x)x^{-\nu-1}, \quad \text{for } \gamma = 0 \quad (\text{D.6})$$

$$f_X(x) = L(x)x^{-\frac{\nu}{2}-1}, \quad \text{for } \gamma > 0, \quad (\text{D.7})$$

where  $L(x)$  denotes a slowly varying function at  $\infty$ . The asymptotic relation for the modified Bessel function of the third kind (B.6) was applied to (D.4) to arrive at (D.7).

<sup>2</sup>Note that the  $(\lambda, \bar{\alpha}, \mu, \Sigma, \gamma)$  parametrization yields to a slightly different Student-t parametrization: In this package the parameter  $\Sigma$  denotes the variance in the multivariate case and the standard deviation in the univariate case. Thus, set  $\sigma = \sqrt{\nu/(\nu-2)}$  in the univariate case to get the same results as with the standard R implementation of the Student-t distribution.

## D.2. Variance gamma distribution

Relation (B.3) can be used again to show that for  $\chi \rightarrow 0$  and  $\lambda > 0$  the density of the GH distribution results in

$$f_{\mathbf{X}}(\mathbf{x}) = \frac{\psi^\lambda (\psi + \gamma' \Sigma^{-1} \gamma)^{\frac{d}{2} - \lambda}}{(2\pi)^{\frac{d}{2}} |\Sigma|^{\frac{1}{2}} \Gamma(\lambda) 2^{\lambda-1}} \times \frac{K_{\lambda - \frac{d}{2}}(\sqrt{Q(\mathbf{x})}(\psi + \gamma' \Sigma^{-1} \gamma)) e^{(\mathbf{x} - \mu)' \Sigma^{-1} \gamma}}{(\sqrt{Q(\mathbf{x})}(\psi + \gamma' \Sigma^{-1} \gamma))^{\frac{d}{2} - \lambda}}. \quad (\text{D.8})$$

In the case of a variance gamma distribution the transformation of  $\bar{\alpha}$  to  $\chi$  and  $\psi$  (cf. 4.3) reduces to

$$\psi = \bar{\alpha} \frac{K_{\lambda+1}(\bar{\alpha})}{K_{\lambda}(\bar{\alpha})} = 2\lambda \quad (\text{D.9})$$

Thus, the density is

$$f_{\mathbf{X}}(\mathbf{x}) = \frac{2\lambda^\lambda (2\lambda + \gamma' \Sigma^{-1} \gamma)^{\frac{d}{2} - \lambda}}{(2\pi)^{\frac{d}{2}} |\Sigma|^{\frac{1}{2}} \Gamma(\lambda)} \times \frac{K_{\lambda - \frac{d}{2}}(\sqrt{Q(\mathbf{x})}(2\lambda + \gamma' \Sigma^{-1} \gamma)) e^{(\mathbf{x} - \mu)' \Sigma^{-1} \gamma}}{(\sqrt{Q(\mathbf{x})}(2\lambda + \gamma' \Sigma^{-1} \gamma))^{\frac{d}{2} - \lambda}}. \quad (\text{D.10})$$

A limiting case arises when  $Q(\mathbf{x}) \rightarrow 0$ , that is when  $\mathbf{x} - \mu \rightarrow 0$ . As long as  $\lambda - \frac{d}{2} > 0$  relation (B.3) can be used to verify that the density reduces to

$$f_{\mathbf{X}}(\mathbf{x}) = \frac{\psi^\lambda (\psi + \gamma' \Sigma^{-1} \gamma)^{\frac{d}{2} - \lambda} \Gamma(\lambda - \frac{d}{2})}{2^d \pi^{\frac{d}{2}} |\Sigma|^{\frac{1}{2}} \Gamma(\lambda)}. \quad (\text{D.11})$$

By replacing  $\psi$  with  $2\lambda$  the limiting density is obtained in the  $(\lambda, \bar{\alpha}, \mu, \Sigma, \gamma)$ -parametrization.<sup>3</sup>

For  $\lambda - \frac{d}{2} \leq 0$  the density diverges.<sup>4</sup>

## E. Conditional density of the mixing variable $W$

Performing the E-Step of the MCECM algorithm requires the calculation of the conditional expectation of  $w_i$  given  $\mathbf{x}_i$ . In this section the conditional density is derived.

### E.1. Generalized hyperbolic, hyperbolic and NIG distribution

The mixing term  $w$  is GIG distributed. By using (2.5) and (C.1) the density of  $w_i$  given  $\mathbf{x}_i$  can be calculated to be again the GIG density with parameters  $(\lambda - \frac{d}{2}, Q(\mathbf{x}) + \chi, \psi + \gamma' \Sigma^{-1} \gamma)$ .

$$\begin{aligned} f_{w|\mathbf{x}}(w) &= \frac{f_{\mathbf{X},W}(\mathbf{x}, w)}{f_{\mathbf{X}}(\mathbf{x})} \\ &= \frac{f_{\mathbf{X}|W}(\mathbf{x}) f_{GIG}(w)}{\int_0^\infty f_{\mathbf{X}|W}(\mathbf{x}) f_{GIG}(w) dw} \\ &= \left( \frac{\gamma' \Sigma^{-1} \gamma + \psi}{Q(\mathbf{x}) + \chi} \right)^{0.5(\lambda - \frac{d}{2})} \times \\ &\quad \frac{w^{\lambda - \frac{d}{2} - 1} \exp \left\{ -\frac{1}{2} \left( \frac{Q(\mathbf{x}) + \chi}{w} + w (\gamma' \Sigma^{-1} \gamma + \psi) \right) \right\}}{2 K_{\lambda - \frac{d}{2}}(\sqrt{(Q(\mathbf{x}) + \chi) (\gamma' \Sigma^{-1} \gamma + \psi)})} \end{aligned} \quad (\text{E.1})$$

<sup>3</sup>The numeric implementation in R uses spline interpolation for the case where  $\lambda - \frac{d}{2} > 0$  and  $Q(\mathbf{x}) < \epsilon$ .

<sup>4</sup>The current workaround in R simply sets observations where  $Q(\mathbf{x}) < \epsilon$  to  $\epsilon$  when  $\lambda - \frac{d}{2} \leq 0$ .

## E.2. Student-t distribution

The mixing term  $w$  is  $\Pi$  distributed. Again, the conditional density of  $w_i$  given  $\mathbf{x}_i$  results in the GIG density. The equations (2.5) and (C.9) were used. The parameters of the GIG density are  $(\lambda - \frac{d}{2}, Q(\mathbf{x}) + \chi, \gamma' \Sigma^{-1} \gamma)$ . When  $\gamma$  becomes 0 the conditional density reduces to the  $\Pi$  density with parameters  $(\frac{d}{2} - \lambda, \frac{Q(\mathbf{x}) + \chi}{2})$ .

$$\begin{aligned}
 f_{w|\mathbf{x}}(w) &= \frac{f_{\mathbf{X},W}(\mathbf{x}, w)}{f_{\mathbf{X}}(\mathbf{x})} \\
 &= \frac{f_{\mathbf{X}|W}(\mathbf{x}) f_{\Pi}(w)}{\int_0^\infty f_{\mathbf{X}|W}(\mathbf{x}) f_{\Pi}(w) dw} \\
 &= \left( \frac{\gamma' \Sigma^{-1} \gamma}{Q(\mathbf{x}) + \chi} \right)^{0.5(\lambda - \frac{d}{2})} \times \frac{w^{\lambda - \frac{d}{2} - 1} \exp \left\{ -\frac{1}{2} \left( \frac{Q(\mathbf{x}) + \chi}{w} + w \gamma' \Sigma^{-1} \gamma \right) \right\}}{2 K_{\lambda - \frac{d}{2}}(\sqrt{(Q(\mathbf{x}) + \chi) \gamma' \Sigma^{-1} \gamma})} \quad (\text{E.2})
 \end{aligned}$$

## E.3. Variance gamma distribution

The mixing term  $w$  is  $\Gamma$  distributed. By using (2.5) and (C.6) the density of  $w_i$  given  $\mathbf{x}_i$  can be calculated to be again the GIG density with parameters  $(\lambda - \frac{d}{2}, Q(\mathbf{x}), \psi + \gamma' \Sigma^{-1} \gamma)$ .

$$\begin{aligned}
 f_{w|\mathbf{x}}(w) &= \frac{f_{\mathbf{X},W}(\mathbf{x}, w)}{f_{\mathbf{X}}(\mathbf{x})} \\
 &= \frac{f_{\mathbf{X}|W}(\mathbf{x}) f_{\Gamma}(w)}{\int_0^\infty f_{\mathbf{X}|W}(\mathbf{x}) f_{\Gamma}(w) dw} \\
 &= \left( \frac{\gamma' \Sigma^{-1} \gamma + \psi}{Q(\mathbf{x})} \right)^{0.5(\lambda - \frac{d}{2})} \times \quad (\text{E.3})
 \end{aligned}$$

$$\frac{w^{\lambda - \frac{d}{2} - 1} \exp \left\{ -\frac{1}{2} \left( \frac{Q(\mathbf{x})}{w} + w (\gamma' \Sigma^{-1} \gamma + \psi) \right) \right\}}{2 K_{\lambda - \frac{d}{2}}(\sqrt{Q(\mathbf{x}) (\gamma' \Sigma^{-1} \gamma + \psi)})} \quad (\text{E.4})$$

## F. Distribution objects

In the package **ghyp** we follow an object-oriented programming approach and introduce distribution objects. There are mainly four reasons for that:

1. Unlike most distributions the GH distribution involves at least 5 parameters which have to fulfill some consistency requirements. Consistency checks can be performed once for all when an object is initialized. In addition, constructors for different parametrizations can be added easily and do not necessitate a change of all the function headers.
2. Once initialized the common functions belonging to a distribution can be called conveniently by passing the distribution object. A repeated input of the parameters is avoided.
3. Distribution objects returned from fitting procedures can be directly passed to, e.g., the density function since fitted distribution objects add information to the distribution object and consequently inherit from the class of the distribution object.

4. Generic method dispatching can be used to provide a uniform interface to, e.g., calculate the expected value `mean(distribution.object)`. Additionally, one can take advantage of generic programming since R provides virtual classes and some forms of polymorphism.

See appendix H for several examples and H.2 for particular examples concerning the object-oriented approach.

## G. Constructors

Before giving examples on how GH distribution objects can be initialized, let us list the constructor functions for different distributions and parametrizations. Note that the constructors below are used to initialize both, univariate and multivariate distributions.

Distribution	Parametrization		
	$(\lambda, \chi, \psi, \mu, \Sigma, \gamma)$	$(\lambda, \bar{\alpha}, \mu, \Sigma, \gamma)$	$(\lambda, \alpha, \mu, \Delta, \delta, \beta)$
GH	<code>ghyp(...)</code>	<code>ghyp(..., alpha.bar=x)</code>	<code>ghyp.ad(...)</code>
hyp	<code>hyp(...)</code>	<code>hyp(..., alpha.bar=x)</code>	<code>hyp.ad(...)</code>
NIG	<code>NIG(...)</code>	<code>NIG(..., alpha.bar=x)</code>	<code>NIG.ad(...)</code>
Student-t	<code>student.t(..., chi=x)</code>	<code>student.t(...)</code>	<code>student.t.ad(...)</code>
VG	<code>VG(..., psi=x)</code>	<code>VG(...)</code>	<code>VG.ad(...)</code>

Apparently, the same constructor functions are used for the  $(\lambda, \chi, \psi, \mu, \Sigma, \gamma)$  and the  $(\lambda, \bar{\alpha}, \mu, \Sigma, \gamma)$  parametrization.

In case of the GH, hyp, and NIG distribution, the idea is that as long as the parameter `alpha.bar` is not submitted the  $(\lambda, \chi, \psi, \mu, \Sigma, \gamma)$  parametrization is used. Conversely, if `alpha.bar` is submitted, the  $(\lambda, \bar{\alpha}, \mu, \Sigma, \gamma)$  parametrization will be used and the parameters  $\chi$  and  $\psi$  are neglected. Thus, typing either `ghyp()`, `hyp()`, or `NIG()` initializes an distribution object in the  $(\lambda, \chi, \psi, \mu, \Sigma, \gamma)$  parametrization.

This is different for the Student-t and the VG distribution. Per default, these distributions are initialized in the  $(\lambda, \bar{\alpha}, \mu, \Sigma, \gamma)$  parametrization. In case of the Student-t distribution, the  $(\lambda, \chi, \psi, \mu, \Sigma, \gamma)$  parametrization is choosen as soon as a  $\chi$  (`chi`) is submitted which does not satisfy eq. D.3 In case of the VG distribution, the  $(\lambda, \chi, \psi, \mu, \Sigma, \gamma)$  parametrization is choosen as soon as a  $\psi$  (`psi`) is submitted which does not satisfy eq. D.9

To get the standard Student-t parametrization (i.e. the R-core implementation) use `student.t(nu = nu, chi = nu)`.

## H. Examples

This section provides examples of distribution objects and the object-oriented approach as well as fitting to data and portfolio optimization.

### H.1. Initializing distribution objects

This example shows how GH distribution objects can be initialized by either using the  $(\lambda, \chi, \psi, \mu, \Sigma, \gamma)$ , the  $(\lambda, \bar{\alpha}, \mu, \Sigma, \gamma)$  or the  $(\lambda, \alpha, \mu, \Delta, \delta, \beta)$ -parametrization.

```
> library(ghyp)
> ghyp(lambda = -2, alpha.bar = 0.5, mu = 10, sigma = 5, gamma = 1)
```

Asymmetric Generalized Hyperbolic Distribution:

Parameters:

lambda	alpha.bar	mu	sigma	gamma
-2.0	0.5	10.0	5.0	1.0

```
> ghyp.ad(lambda = -2, alpha = 1, beta = 0.5, mu = 10, delta = 1)
```

Asymmetric Generalized Hyperbolic Distribution:

Parameters:

lambda	alpha	delta	beta	mu
-2.0	1.0	1.0	0.5	10.0

```
> ghyp(lambda = -2, chi = 5, psi = 0.1, mu = 10:11, sigma = diag(5:6),
+       gamma = -1:0)
```

Asymmetric Generalized Hyperbolic Distribution:

Parameters:

lambda	chi	psi
-2.0	5.0	0.1

mu:

```
[1] 10 11
```

sigma:

	[,1]	[,2]
[1,]	5	0
[2,]	0	6

gamma:

```
[1] -1 0
```

## H.2. Object-oriented approach

First of all a GH distribution object is initialized and a consistency check takes place. The second command shows how the initialized distribution object is passed to the density function. Then a Student-t distribution is fitted to the daily log-returns of the Novartis stock. The fitted distribution object is passed to the quantile function. Since the fitted distribution object inherits from the distribution object this constitutes no problem.

The generic methods *hist*, *pairs*, *coef*, *plot*, *lines*, *transform*, *[,]*, *mean* and *vcov* are defined for distribution objects inheriting from the class *ghyp*.

The generic methods *logLik*, *AIC* and *summary* are added for the class *mle.ghyp*, inheriting from the class *ghyp*.

```
## Consistency check when initializing a GH distribution object.
## Valid input:
univariate.ghyp.object <- student.t(nu = 3.5, mu = 10,
                                     sigma = 5, gamma = 1)

## Passing a distribution object to the density function
dghyp(10:14,univariate.ghyp.object)

## Passing a fitted distribution object to the quantile function
fitted.ghyp.object <- fit.tuv(smi.stocks["Novartis"], silent = T)
qghyp(c(0.01,0.05),fitted.ghyp.object)

## Generic method dispatching: the histogram method
hist(fitted.ghyp.object, legend.cex = 0.7)

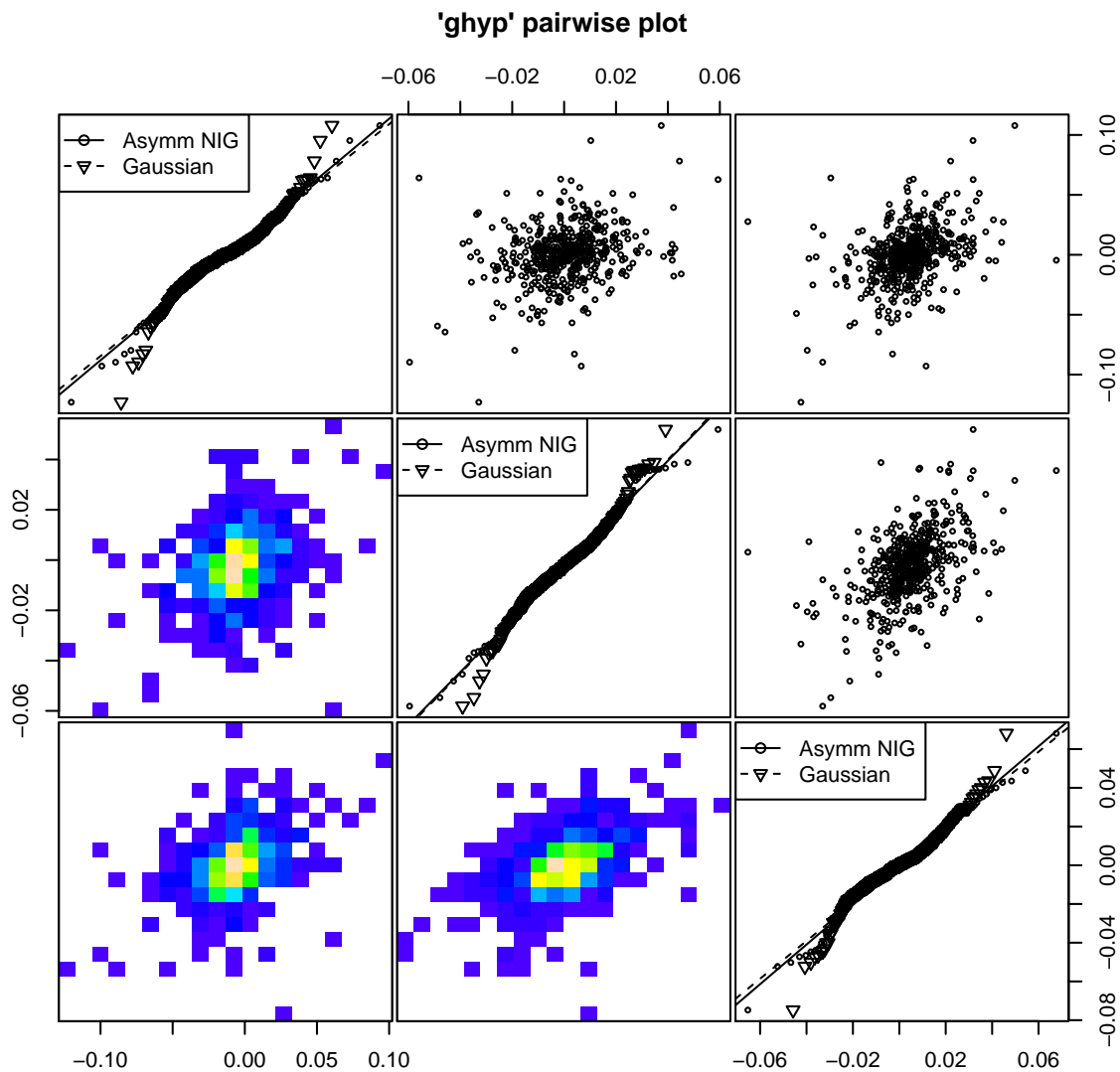
## Generic programming:
mean(fitted.ghyp.object)      ## fitted.ghyp.object is of class
                             ## "mle.ghyp" which extends "ghyp"

vcov(univariate.ghyp.object)
```

### H.3. Fitting generalized hyperbolic distributions to data

A multivariate NIG distribution is fitted to the daily returns of three swiss blue chips: Credit Suisse, Nestle and Novartis. A *pairs* plot is drawn in order to do some graphical diagnostics of the accuracy of the fit.

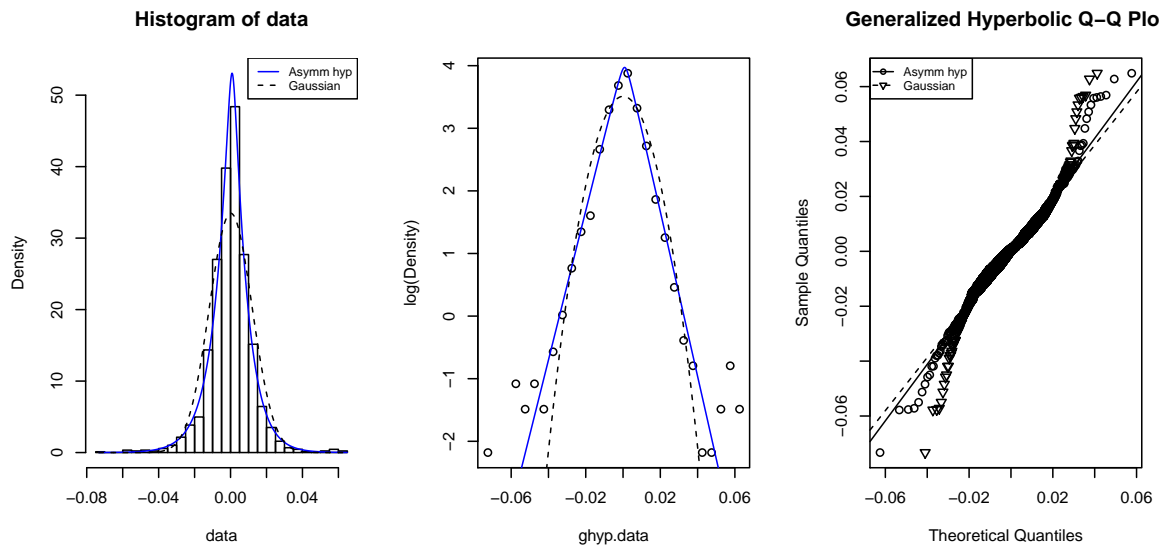
```
> data(smi.stocks)
> fitted.returns.mv <- fit.NIGmv(data = smi.stocks[1:500, c("CS",
+   "Nestle", "Novartis")], silent = TRUE)
> pairs(fitted.returns.mv, cex = 0.5, nbins = 20)
```



In the following part daily log-returns of the SMI are fitted to the hyperbolic distribution. Again, some graphical verification is done to check the accuracy of the fit.

```
> fitted.smi.returns <- fit.hypuv(data = smi.stocks[, c("SMI")],
+   silent = TRUE)
> par(mfrow = c(1, 3))
> hist(fitted.smi.returns, ghyp.col = "blue", legend.cex = 0.7)
> hist(fitted.smi.returns, log.hist = T, nclass = 30, plot.legend = F,
+   ghyp.col = "blue")
> qqghyp(fitted.smi.returns, plot.legend = T, legend.cex = 0.7)
```





## References

- Milton Abramowitz and Irene A. Stegun. *Handbook of mathematical functions*. Dover books on mathematics. Dover, New York, NY, 9. dover printing edition, 1972. ISBN 0486612724.
- O. E. Barndorff-Nielsen. Exponentially decreasing distributions for the logarithm of particle size. *Proceedings of the Royal Society of London. Series A, Mathematical and Physical Sciences*, 353(1674):401–419, 1977.
- O. E. Barndorff-Nielsen. Normal inverse Gaussian distributions and stochastic volatility modelling. *Scandinavian Journal of Statistics*, 24(1):1–13, 1997a.
- O. E. Barndorff-Nielsen. Processes of normal inverse gaussian type. *Finance and Stochastics*, 2(1):41–68, 1997b.
- O. E. Barndorff-Nielsen and O. Halgreen. Infinite divisibility of the hyperbolic and generalized inverse gaussian distribution. *Zeitschrift für Wahrscheinlichkeitstheorie und verwandte Gebiete*, 38(4):309–311, 1977.
- Ole E. Barndorff-Nielsen and Karsten Prause. Apparent scaling. *Finance and Stochastics*, 5(1):103–113, 2001.
- B. M. Bibby and M. Sørensen. A hyperbolic diffusion model for stock prices. *Finance & Stochastics* 2, pages 25–41, 1997.
- Wolfgang Breyermann, Alexandra Dias, and Paul Embrechts. Dependence Structures for Multivariate High-Frequency Data in Finance. *Quantitative Finance*, 3(1):1–14, 2003.
- Ying Chen, Wolfgang Härdle, and Seok-Oh Jeong. *Nonparametric Risk Management with Generalized Hyperbolic Distributions*, volume 1063 of *Preprint / Weierstraß-Institut für Angewandte Analysis und Stochastik*. WIAS, Berlin, 2005.

- Rama Cont and Peter Tankov. *Financial Modelling with Jump Processes*. Chapman & Hall/CRC Financial Mathematics Series, 2003.
- James Davidson. Moment and memory properties of linear conditional heteroscedasticity models, and a new model. *Journal of Business & Economic Statistics*, 22(1):16–29, January 2004.
- A. P. Dempster, N. M. Laird, and D. B. Rubin. Maximum likelihood from incomplete data via the EM algorithm. *Journal of the Royal Statistical Society*, 39(1):1–38, 1977.
- E. Eberlein and U. Keller. Hyperbolic distributions in finance. *Bernoulli* 1, pages 281–299, 1995.
- E. Eberlein, U. Keller, and K. Prause. New insights into smile, mispricing and value at risk measures. *Journal of Business* 71, pages 371–405, 1998.
- Kevin Fergusson and Eckhard Platen. On the distributional characterization of log-returns of a world stock index. *Applied Mathematical Finance*, 13(1):19–38, March 2006.
- Lars Forsberg and Tim Bollerslev. Bridging the gap between the distribution of realized (ECU) volatility and arch modelling (of the EURO): the GARCH-NIG model. 17:535–548, 2002.
- Stefan Kassberger. Efficient portfolio optimization in the multivariate generalized hyperbolic framework. *SSRN eLibrary*, 2007.
- Stefan Kassberger and Rüdiger Kiesel. A fully parametric approach to return modelling and risk management of hedge funds. *Financial markets and portfolio management*, 20(4): 472–491, 2006.
- Dilip B. Madan, Peter Carr, and Eric C. Chang. The variance gamma process and option pricing. *European Finance Review*, 2:79–105, 1998.
- Alexander J. McNeil. Free S-Plus and R library designed to accompany the book quantitative risk management: Concepts, techniques and tools. <http://cran.r-project.org/web/packages/QRMLib/index.html>, 2005.
- Alexander J. McNeil, Rüdiger Frey, and Paul Embrechts. *Quantitative risk management: Concepts, techniques and tools*. Princeton University Press, Princeton, 2005. ISBN 978–0–691–12255–7.
- X.-L. Meng and D.-B. Rubin. Maximum likelihood estimation via the ECM algorithm: A general framework. *Biometrika*, 80(2):267–278, 1993.
- Marc Paoletta. *Intermediate probability: A computational approach*. Wiley, Chichester, 2007. ISBN 978–0–470–02637–3.
- Karsten Prause. *The Generalized Hyperbolic Model: Estimation Financial Derivatives and Risk Measures*. PhD thesis, Universität Freiburg i. Br., Freiburg i. Br., 1999.
- Sebastian Raible. *Lévy processes in finance: Theory, numerics, and empirical facts*. PhD thesis, Universität Freiburg i. Br., 2000.