Efficient and robust portfolio optimization in the multivariate Generalized Hyperbolic framework

MARTIN HELLMICH† and STEFAN KASSBERGER*

†Frankfurt School of Finance and Management, Sonnemannstraße 9–11, 60314 Frankfurt, Germany
‡Institute of Mathematical Finance, Ulm University, Helmholtzstraße 18, 89069 Ulm, Germany

(Received 14 October 2007; in final form 18 May 2009)

In this paper, we apply the multivariate Generalized Hyperbolic (mGH) distribution to portfolio modeling, using Conditional Value at Risk (CVaR) as a risk measure. Exploiting the fact that portfolios whose constituents follow an mGH distribution are univariate GH distributed, we prove some results relating to measurement and decomposition of portfolio risk, and show how to efficiently tackle portfolio optimization. Moreover, we develop a robust portfolio optimization approach in the mGH framework, using Worst Case Conditional Value at Risk (WCVaR) as risk measure.

Keywords: Portfolio optimization; Robust optimization; Asset allocation; Risk management; Multivariate Generalized Hyperbolic distribution; Conditional Value at Risk; Worst Case Conditional Value at Risk

JEL Classification: C16; C52; C61; C63; G32

1. Introduction

Modern portfolio optimization is a far cry from the classical mean–variance approach pioneered by Markowitz (1952). The departure from the time-honored Markowitz framework has been spurred by two intimately related insights. First, the use of a Gaussian distribution to describe the returns of financial assets will inevitably lead to what can at best be called a rough approximation to reality. Second, variance can be an inadequate risk measure if a more flexible, non-Gaussian return distribution is adopted.

It has become a generally accepted fact, supported by numerous empirical studies, that empirical asset return distributions are non-normal. In fact, they are almost always found to exhibit skewness (i.e. asymmetry) and excess kurtosis, which renders the normal (Gaussian) distribution an inadequate model (Prause 1999, Raible 2000, Schoutens 2003, Cont and Tankov 2004). Thus, realistic modelling calls for alternative probability distributions. In recent years, several viable alternatives to the Gaussian distribution, capable of capturing commonly observed empirical features, have been proposed for use in financial modelling. For example, Madan and Seneta (1990) suggest the Variance Gamma distribution, Eberlein and Keller (1995) and Bingham and Kiesel (2001) advocate the use of the Hyperbolic distribution, Barndorff-Nielsen (1997) proposes the Normal Inverse Gaussian distribution, Eberlein (2001) applies the Generalized Hyperbolic distribution, and Aas and Haff (2006) find that the Generalized Hyperbolic Skew Student’s t distribution matches empirical data very well.

While these studies document the superior capabilities of the Generalized Hyperbolic class and its subclasses when it comes to realistically describing univariate financial data, recent empirical studies conducted in a multivariate setting make a convincing case for the multivariate Generalized Hyperbolic (mGH) distribution and its subclasses as a model for multivariate financial data as well. For instance, McNeil et al. (2005) calibrate the mGH model and its subclasses to both multivariate stock- and multivariate exchange-rate returns. In a likelihood-ratio test against the general mGH model, the Gaussian model is always rejected. Aas et al. (2005) and Kassberger and Kiesel (2006) employ the multivariate NIG (Normal Inverse Gaussian) distribution successfully for risk management purposes. The latter study demonstrates that the NIG distribution provides a much better...
fit to the empirical distribution of hedge fund returns than the normal distribution. The Gaussian distribution is found to seriously understate the probability of tail events, while the heavier tails of the mGH class seem to describe actual tail behavior well. Tail-related risk measures such as Value at Risk (VaR) and Conditional Value at Risk (CVaR) are shown to be severely misleading when calculated on the basis of the Gaussian distribution. This problem is found to carry over into the portfolio context.

All the aforementioned distributions have two important features in common. First, they can all be considered as marginal distributions of (multivariate) Lévy processes (i.e. processes with independent and stationary, but not necessarily Gaussian increments). Second, they all belong to the class of (multivariate) Generalized Hyperbolic distributions, which encompasses the Gaussian distribution as a limiting case. Therefore, the mGH class offers a natural generalization of the multivariate Gaussian class.

The departure from the normal distribution and the adoption of more realistic distributions, however, call for adequate risk measures and computational tools. Portfolio optimization using non-Gaussian distributions should not be performed in a mean–variance framework, because in the non-Gaussian case it is inappropriate to describe the riskiness of a financial asset solely by the variance of its returns (thereby ignoring higher moments). In recent years, CVaR, also known as Expected Shortfall or Tail-VaR, has been embraced by academics and practitioners alike as a tractable and theoretically well-founded alternative to classical risk measures such as VaR or variance.

In addition to being based on realistic distributional assumptions and an informative risk measure, an alternative portfolio optimization approach should be computationally feasible—even for problems involving a large number of assets—in order to be applicable to real-world situations. Moreover, it should be amenable to a robust formulation of the portfolio optimization problem. Robust formulations are based on the insight that optimal portfolios can be remarkably sensitive to only slight variations in the input parameters, which are often fraught with estimation error. The combined effect of the uncertainties in the parameters can render the result of a portfolio optimization procedure highly unreliable. To counteract this phenomenon, robust approaches rely on uncertainty sets that contain the ‘true’ parameters for a specific confidence level, instead of point estimates of the parameters, thereby taking parameter uncertainty into account. For a survey of robust optimization, the interested reader is referred to BERTSIMAS et al. (2008).

A modern portfolio optimization approach is thus characterized by the following desirable features: allowance for realistic return distributions, use of a realistic risk measure, computational tractability, and admissibility of a tractable robust formulation. We contribute to the literature by proposing a portfolio optimization approach that incorporates all of these features. Our approach is based on the mGH distribution, relies on CVaR, leads to a convex optimization problem, and allows for a robust formulation that can be solved just as efficiently as the original problem.

The remainder of this paper is structured as follows. Section 2 introduces CVaR as an alternative risk measure and gives an overview of several standard forms of the portfolio optimization problem. In section 3, the multivariate Generalized Hyperbolic class of distributions is introduced. In addition, results relating to the determination of the CVaR for mGH portfolios are established, and a decomposition formula for the CVaR of a portfolio is presented and proved. These results, while interesting in their own right for risk management purposes, form the foundation for an efficient formulation of the portfolio optimization problem in the mGH framework, which is the subject of section 4. Furthermore, section 4 introduces a robust formulation of the portfolio optimization problem, which relies on Worst Case Conditional Value at Risk (WCVaR) as a risk measure. It is shown that the robust portfolio optimization problem can be solved as efficiently as the original problem. Section 5 is devoted to a numerical study in which the methodologies developed in the paper are applied to empirical data. Section 6 sums up the main insights and concludes.

2. Risk measures, performance measures, and portfolio optimization

2.1. Coherent measures of risk

Since a non-Gaussian distribution cannot be characterized solely in terms of its means and its covariance matrix, a deviation from the multivariate Gaussian paradigm of portfolio optimization has to be supported by the adoption of alternative risk measures, such as Value at Risk or CVaR. In their seminal paper, ARTZNER et al. (1999) specify a number of desirable properties a risk measure should have and introduce the notion of a coherent risk measure (see also Malevergne and SORNETTE (2006)). In the following definition, $L_1$ and $L_2$ can be interpreted as random losses. A risk measure $\phi$ that maps a random loss to a real number is said to be coherent if it satisfies the following axioms.

- $(A1)$ Translation invariance: $\phi(L_1 + l) = \phi(L_1) + l$, for all random losses $L_1$ and all $l \in \mathbb{R}$.
- $(A2)$ Subadditivity: $\phi(L_1 + L_2) \leq \phi(L_1) + \phi(L_2)$, for all random losses $L_1$, $L_2$.
- $(A3)$ Positive homogeneity: $\phi(\lambda L_1) = \lambda \phi(L_1)$, for all random losses $L_1$ and all $\lambda > 0$.
- $(A4)$ Monotonicity: $\phi(L_1) \leq \phi(L_2)$, for all random losses $L_1$, $L_2$ with $L_1 \leq L_2$ almost surely.

It is worth noting that subadditivity and positive homogeneity imply convexity, whereas the converse generally does not hold.

Value at Risk (VaR) has become an industry standard for measuring financial risks. VaR has derived much of its popularity from the fact that it gives a handy and easy-to-understand representation of potential losses.
If $X$ is the random return associated with an asset, then $L = -X$ is the relative loss, and the VaR at level $\beta \in (0, 1)$, denoted by $VaR_\beta(L)$, is defined as $VaR_\beta(L) = \inf \{ l \in \mathbb{R} : \mathbb{P}(L > l) \leq 1 - \beta \} = \inf \{ l \in \mathbb{R} : F_L(l) \geq \beta \}$. Hence, $VaR_\beta(L)$ is the smallest relative loss level whose probability of being exceeded is at most $1 - \beta$. For continuous, strictly increasing loss distribution functions (which we will assume throughout the paper), VaR can be more simply expressed as the $\beta$-quantile of the loss distribution function $F_L$:

$$VaR_\beta(L) = F_L^{-1}(\beta).$$

Of course, VaR can also be defined in terms of absolute losses. However, as we are going to model returns rather than VaR, as $CVaR = \frac{1}{\beta}Cov(L)$, we have:

A straightforward consequence of this definition is the relation $CVaR_\beta(L) \geq VaR_\beta(L)$. CVaR is more informative than VaR, as $CVaR_\beta(L)$ takes the loss distribution beyond the point $VaR_\beta(L)$ into account and thus also measures the severity of losses that exceed $VaR_\beta(L)$. VaR, in contrast, ignores losses beyond $VaR_\beta(L)$ and thus discards information implicit in the loss distribution. CVaR is well suited as a risk measure in the context of portfolio optimization, for reasons that will be elaborated on in what follows.

2.2. Portfolio optimization using CVaR

Portfolio optimization problems appear in various guises. The following result, which is proved by Krokhmal et al. (2002), establishes a link among three of the most common formulations. Let $\phi : \mathcal{A} \rightarrow \mathbb{R}$ be a convex risk measure, and let $R : \mathcal{A} \rightarrow \mathbb{R}$ be a concave reward function, both defined on the convex set $\mathcal{A} \subseteq \mathbb{R}^d$. Let $x \in \mathcal{A}$ be a vector of portfolio weights, i.e. assume that $\sum_{i=1}^d x_i = 1$. Then the following three optimization problems lead to the same efficient frontiers when varying the parameters $\lambda$, $\rho$, and $\omega$, respectively:

$$(P1) \quad \min_{x \in \mathcal{A}} \phi(x) - \lambda R(x),$$

subject to $\lambda \geq 0$,

$$(P2) \quad \min_{x \in \mathcal{A}} \phi(x),$$

subject to $R(x) \geq \rho$,

$$(P3) \quad \max_{x \in \mathcal{A}} R(x),$$

subject to $\phi(x) \leq \omega$.

In other words, a portfolio that is efficient for one of these three problem formulations will also be efficient for the other two. In our subsequent considerations, we will identify $R(x)$ with the expected portfolio return, which is a linear (and thus concave) function of portfolio weights, and $\phi(x)$ with the portfolio CVaR, which is convex in the portfolio weights.

While the above formulations involving the minimization of a linear functional of risk and reward are very common in the literature, other formulations of the portfolio optimization problem entail the maximization of a reward–risk ratio. For instance, the use of Return-on-Risk-Capital (RORC for short), defined as $R(x)/\phi(x)$, is motivated by Fischer and Roehrl (2005). Rachev et al. (2007) provide an overview of various other reward–risk ratios.

3. Beyond Gaussian mean–variance optimization: Using the mGH distribution for portfolio modelling

3.1. Modelling multivariate returns with the mGH distribution

As already pointed out, there is compelling empirical evidence that returns of financial assets are not Gaussian. As a consequence, a more realistic model is called for. Because of its great generality and relatively high numerical tractability, the mGH distribution is an ideal candidate.

3.1.1. The mGH distribution as a normal mean–variance mixture. A random variable $W \in \mathbb{R}_+$ is said to have a Generalized Inverse Gaussian (GIG) distribution with parameters $\lambda$, $\chi$, and $\psi$, denoted by $W \sim GIG(\lambda, \chi, \psi)$, if its density is given by

$$f_{GIG}(y; \lambda, \chi, \psi) = \begin{cases} \frac{\lambda^{\chi/2}}{2K_\chi(\sqrt{\lambda})} y^{\chi-1} \exp\left(-\frac{\chi y^{-1} + \psi y}{2}\right), & y > 0, \\ 0, & y \leq 0, \end{cases}$$

where, for $x > 0$, $K_\chi(x)$ is the modified Bessel function of the third kind with index $\lambda$:

$$K_\chi(x) = \frac{1}{2} \int_0^\infty y^{\chi-1} \exp\left(-\frac{x(y + y^{-1})}{2}\right) dy.$$
The parameters in (1) are assumed to satisfy $\chi > 0$ and $\psi \geq 0$ if $\lambda < 0$; $\chi > 0$ and $\psi > 0$ if $\lambda = 0$; and $\chi \geq 0$ and $\psi > 0$ if $\lambda > 0$. The expected value of $Y$ can be expressed as:

$$\mathbb{E}(W) = \frac{\sqrt{\chi/\psi} K_{1+(}\sqrt{\psi/\chi})}{K_0(\sqrt{\chi/\psi})}. \quad (2)$$

The class of mGH distributions can now be introduced as the class of normal mean–variance mixtures with a GIG-distributed mixing variable. A random variable $X = (X_1, \ldots, X_d)$ is said to follow a $d$-dimensional mGH distribution with parameters $\lambda, \chi, \psi, \mu, \gamma, \Sigma$, denoted by $X \sim mGH(\lambda, \chi, \psi, \mu, \gamma, \Sigma)$, if

$$X \overset{d}{=} \mu + W\gamma + \sqrt{W}AZ,$$

where $\mu, \gamma \in \mathbb{R}^d$ are deterministic, $Z \sim N_0(0, I_d)$ follows a $k$-dimensional normal distribution, $W \sim GIG(\lambda, \chi, \psi)$ is a positive, scalar random variable independent of $Z$, $A \in \mathbb{R}^{k \times d}$ denotes a $d \times k$ matrix, and $\Sigma = AA'$. We find that $X | W = w \sim N_0(\mu + w\gamma, w\Sigma)$, i.e. the conditional distribution of $X$ given $W$ is normal, which explains the name normal mean–variance mixture. The mixing variable $W$ can be thought of as a stochastic volatility factor. From the above definition, it follows directly that $\mathbb{E}(X) = \mu + \mathbb{E}(W)\gamma$ and $\text{Cov}(X) = \mathbb{E}(W)\Sigma + \text{Var}(W)\gamma\gamma'$. It is interesting to note that the absence of correlation of the components of $X$ implies independence if and only if $W$ is almost surely constant, i.e. if $X$ is multivariate normal. For $\gamma = 0$, the class of normal variance mixture distributions is obtained. These distributions fall into the class of elliptical distributions, which will be formally introduced later.

For non-singular $\Sigma$, it can be shown that the following representation for the density $f_{mGH}$ of a $d$-dimensional $GH_d(\lambda, \chi, \psi, \mu, \gamma, \Sigma)$ distributed random variable holds:

$$f_{mGH}(y; \lambda, \chi, \psi, \mu, \gamma, \Sigma) = c \frac{K_{d-\frac{d}{2}}(\sqrt{\chi+(y-\mu)\Sigma^{-1}(y-\mu)})(\psi + \gamma\Sigma^{-1}\gamma')}{\exp((y-\mu)\Sigma^{-1}(y-\mu)) \sqrt{\chi+(y-\mu)\Sigma^{-1}(y-\mu)} \sqrt{(\psi + \gamma\Sigma^{-1}\gamma')}}^{d/2-\lambda},$$

with $c$ a normalizing constant,

$$c = \frac{(\chi/\psi)^{\lambda/2} \sqrt{(\psi + \gamma\Sigma^{-1}\gamma')}}{2\pi^{d/2} \sqrt{\det(\Sigma^{1/2} K_0(\sqrt{\chi/\psi}))}}, \quad (3)$$

where $|\cdot|$ denotes the determinant. Observe that, for every $\lambda > 0$, the distributions $GH_d(\lambda, \chi, \psi, \mu, \gamma, \Sigma)$ and $GH_d(\lambda, \chi/\alpha, a\psi, \mu, a\gamma, a\Sigma)$ coincide, since, for all $y \in \mathbb{R}$,

$$f_{GH}(y; \lambda, \chi, \psi, \mu, \gamma, \Sigma) = f_{GH}(y; \lambda, \chi/\alpha, a\psi, \mu, a\gamma, a\Sigma). \quad (5)$$

This non-uniqueness gives rise to an identifiability problem when trying to calibrate the parameters. However, this problem can be addressed in several ways, for example by requiring the determinant of $\Sigma$ to assume a pre-specified value, or by fixing the value of either $\chi$ or $\psi$.

### 3.1.2. Subclasses of the mGH class

The mGH class of distributions is very general and accommodates many subclasses that have become popular in financial modelling. The purpose of this section is to provide a brief survey of some of these subclasses. Compare also McNeil et al. (2005), who provide a discussion of the tail behavior of these classes.

#### Hyperbolic distributions

For $\lambda = \frac{1}{2}(d+1)$, one arrives at the $d$-dimensional Hyperbolic distribution. However, the univariate marginals of a $d$-dimensional Hyperbolic distribution with $d \geq 2$ are not univariate Hyperbolic distributions. See Eberlein and Keller (1995) for an application of univariate Hyperbolic distributions to financial modelling.

#### Normal Inverse Gaussian (NIG) distributions

For $\lambda = -\frac{1}{2}$, one obtains the class of NIG distributions, which has become widely applied to financial data (see, e.g., Aas et al. (2005) and Kassberger and Kiesel (2006)). The tails of NIG distributions are slightly heavier than those of the Hyperbolic class.

#### Variance Gamma (VG) distributions

For $\lambda > 0$ and $\chi = 0$, one obtains a limiting case which is known as the Variance Gamma class. See Madan and Seneta (1990) for an application of univariate VG distributions to equity return modelling.

#### Skew Student's t distributions

For $\lambda = -\frac{1}{2} \chi$ and $\psi = 0$, another limiting case is obtained, which is often called the Skew Student’s t distribution. The interesting aspect of this distribution is that, in contrast to the aforementioned ones, it is able to account for heavy-tailedness. See Aas and Haff (2006) for an application in a univariate setting.

#### Elliptically symmetric mGH (symGH) distributions

For $\gamma = 0$, one obtains the subclass of elliptically symmetric mGH distributions, which are henceforth called symmetric mGH or symGH distributions. Compared with the general mGH distribution, the density of a symGH distribution simplifies considerably:

$$f_{symGH}(y; \lambda, \chi, \psi, \mu, \Sigma) = \frac{(\chi/\psi)^{\lambda/2} \sqrt{(\chi+(y-\mu)\Sigma^{-1}(y-\mu))\psi}}{(2\pi)^{d/2} \sqrt{\det(\Sigma^{1/2} K_0(\sqrt{\chi/\psi}))} \sqrt{(\chi+(y-\mu)\Sigma^{-1}(y-\mu))\psi}^{d/2-\lambda}} \cdot K_{d-\frac{d}{2}}(\sqrt{\chi+(y-\mu)\Sigma^{-1}(y-\mu)}\psi)^{d/2-\lambda}. \quad (6)$$

The symGH class belongs to the class of elliptical distributions.

### 3.2. Distributional properties, CVaR, and portfolio risk decomposition of mGH portfolios

In this section, we take advantage of the analytical tractability of the mGH class to state the distribution of a portfolio whose constituents follow an mGH distribution. This result will be used to derive analytical expressions for the portfolio’s CVaR and for the risk contribution of a single asset to overall portfolio risk.
3.2.1. Distribution of portfolio returns and CVaR. From the parametrization of the mGH class, it can be easily inferred that it is closed under linear transformations. More precisely, let \( X \sim GH_d(\lambda, \chi, \psi, \mu, \gamma, \Sigma) \) and \( Y = BX + b \), where \( B \in \mathbb{R}^{p \times d} \) and \( b \in \mathbb{R}^p \). Then
\[
Y = BX + b = B\mu + b + WB\chi + \sqrt{W}BAZ
\sim GH_d(\lambda, \chi, \psi, B\mu + b, By, B\Sigma B').
\]
Thus, linear transformations of mGH random variables leave the distribution of the GIG mixing variable unchanged. In particular, it follows that every component \( X_i \) of \( X \) is governed by a univariate GH distribution: \( X_i \sim GH_1(0, 1, \psi, \mu_i, \gamma, \Sigma_i) \). Furthermore, for \( x = (x_1, \ldots, x_d)' \in \mathbb{R}^d \),
\[
x'X = \sum_{i=1}^d x_i X_i \sim GH_d(\lambda, \chi, \psi, \mu, \gamma, \Sigma).
\]
If, in addition, the \( x_i \) are required to sum to unity (i.e. that \( \sum_i x_i = 1 \), where \( 1 = (1, \ldots, 1) \in \mathbb{R}^d \)), and thus can be interpreted as the weights of the individual assets in a portfolio, we can conclude that if the returns of the constituents of a portfolio follow an mGH distribution, then the return of the portfolio is univariate GH distributed.

Now, we derive the univariate GH density of portfolio returns, which will turn out to be considerably simpler than its multivariate counterpart. Using \( f \), we can represent a univariate GH density of the form \( f_{GH_1}(y; \lambda, \chi, \psi, \mu, \gamma, \Sigma) \) as \( f_{GH_1}(y; \lambda, \chi, \psi, \mu, \gamma, \Sigma) \). This shows that, in the univariate case, without loss of generality, the dispersion parameter \( \Sigma \in \mathbb{R}_+ \), can be assumed to be 1, and the density thus simplifies (compare (3) and (4))
\[
f_{GH_1}(y; \lambda, \chi, \psi, \mu, \gamma) = (\chi\psi)^{-1/2} \psi y \exp((y - \mu)') \left( (\chi + (y - \mu)') \exp((y - \mu)') \right)^{-1/2 - \lambda}.
\]
Now assume that the returns of \( d \) assets \( X = (X_1, \ldots, X_d)' \) are distributed according to \( GH_d(\lambda, \chi, \psi, \mu, \gamma, \Sigma) \). Then the return \( x'X \) of a portfolio with weights \( x \) with \( \sum_i x_i = 1 \) follows a \( GH_d(\lambda, \chi, \psi, \mu, \gamma, \Sigma) \) distribution. Denoting by \( f_{GH_d}(y) \) the density of the portfolio return \( x'X \) evaluated at \( y \), CVaR can be computed as follows:
\[
CVaR_\beta(-x'X) = \mathbb{E}[-x'X | -x'X \geq CVaR_\beta(-x'X)]
= -\mathbb{E}[x'X | x'X \leq CVaR_\beta(x'X)]
= \frac{1}{1 - \beta} \int_{-\infty}^{\text{CVaR}_\beta} y \cdot f_{GH_d}(y) \, dy.
\]
4. CVaR-based portfolio optimization in the mGH framework

4.1. The general case

First, we study a portfolio optimization problem of the class (P2), which is representative of the class of problems (P1) through (P3). We choose (P2) because of its similarity to the classical Markowitz problem, which involves minimizing risk (as measured by portfolio variance) under a minimum constraint for the expected return.

Consider the portfolio optimization problem (P2):

\[
\begin{align*}
\text{min}_x & \quad CVaR_\beta (-x'X), \\
\text{subject to} & \quad x \in \mathcal{X} = \{x \in \mathbb{R}^d_+: v'x \geq \rho, \Gamma x = 1\},
\end{align*}
\]

where \( X \sim GH_\mu(\lambda, \chi, \psi, \mu, \gamma, \Sigma) \), \( v = (v_1, \ldots, v_d) \in \mathbb{R}^d \), and \( v_i = \mathbb{E}(X_i) \) is the expected return of asset \( i \). Hence, the objective is to minimize CVaR under the condition that the expected portfolio return is at least \( \rho \).

Since both the objective function and the constraints in (P2) are convex, this problem falls into the category of convex optimization problems, which makes the sophisticated machinery of convex optimization available. In particular, convex optimization problems do not suffer from the existence of local minima which are not at the same time global minima. However, directly evaluating the objective function via formula (7) might be undesirable from a numerical point of view, since it entails use of a numerical root-finding procedure. This problem can be circumvented by applying the insights of Rockafellar and Uryasev (2000, 2002), who introduce an auxiliary function

\[
F_\beta(x, \alpha) = \frac{1}{\beta} - \frac{1}{\beta} \int_{\mathbb{R}^d} \left[-x'y - \alpha\right]^+ p(y) \, dy,
\]

where \( \alpha \) is a real number, \(-x'y\) denotes the loss associated with the vector of portfolio weights \( x \in \mathbb{R}^d \) and the return vector \( y \in \mathbb{R}^d \), and \( p: \mathbb{R}^d \mapsto \mathbb{R}_+ \) is the d-dimensional probability density function of the asset returns. Rockafellar and Uryasev demonstrate that \( F_\beta(x, \alpha) \) is convex with respect to \((x, \alpha)\), and that, given \( x \), CVaR can be calculated by minimizing \( F_\beta(x, \alpha) \) with respect to \( \alpha \):

\[
CVaR_\beta (-x'X) = \min_{\alpha \in \mathbb{R}} F_\beta(x, \alpha).
\]

If we postulate that the asset returns follow a d-dimensional mGH distribution with density function \( f_{GH}(y; \lambda, \chi, \psi, \mu, \gamma, \Sigma) \), then \( F_\beta(x, \alpha) \) can be considerably simplified; in particular, d-dimensional integration can be avoided. Denoting by \( f_{GH}(z; \lambda, \chi, \psi, \mu, \gamma, \Sigma) \) the (univariate) density of the portfolio return \( x'X \), we obtain

\[
\begin{align*}
F_\beta(x, \alpha) &= \alpha + \frac{1}{\beta} \int_{\mathbb{R}^d} \left(-z - \alpha\right)_+ \, dz \\
&= \alpha + \frac{1}{\beta} \int_{-\infty}^{-\alpha} \left(-z - \alpha\right) \, dz \\
&= \alpha + \frac{1}{\beta} \int_{-\infty}^{-\alpha} \left(-z - \alpha\right) \, dz \\
&= \frac{1}{\beta} \int_{-\infty}^{-\alpha} \left(-z - \alpha\right) \, dz.
\end{align*}
\]

Hence, the original problem can be recast as the convex program

\[
\begin{align*}
\text{min}_{x, \alpha} & \quad F_\beta(x, \alpha), \\
\text{subject to} & \quad x \in \mathcal{X} = \{x \in \mathbb{R}^d_+: v'x \geq \rho, \Gamma x = 1\},
\end{align*}
\]

which, in contrast to the original problem (P2), does not require numerical root-finding.

For problems involving the maximization of RORC, i.e. problems of the form

\[
\begin{align*}
\text{max}_x & \quad \frac{x'v}{CVaR_\beta (-x'X)}, \\
\text{subject to} & \quad x \in \mathcal{X} = \{x \in \mathbb{R}^d_+: v'x \geq \rho, \Gamma x = 1\},
\end{align*}
\]

the objective function will in general be non-convex. For this type of problem, we can take advantage of the fact that a solution to (P4) will lie on the efficient frontier induced by the respective problem (P2"); evidently, portfolios that are inefficient in the sense of (P2") cannot be solutions to (P4). Thus, in order to find a solution to (P4), one can calculate the efficient frontier of (P2")—or, more precisely, efficient portfolios for several values of \( \rho \), to approximate the efficient frontier—and then simply evaluate the objective function of (P4) for these portfolios. Thus, we are able to reduce the non-convex optimization problem (P4) to a fixed number (depending on the accuracy required) of convex optimization problems that are efficiently solvable. This is highly advantageous, since for non-convex portfolio optimization problems, one typically has to fall back on heuristic optimization procedures. Fischer and Roehrl (2005), for instance, advocate using swarm-intelligence methods for RORC optimization, which are computationally expensive and cannot guarantee that a global optimum is attained.

4.2. The elliptical case

In this section, we recall the notion of elliptical (also called elliptically symmetric or elliptically contoured) distributions, and demonstrate how to exploit their structural properties in the context of portfolio optimization. Elliptical distributions have been applied to financial modelling since the seminal paper by Owen and Rabinovitch (1983), and they remain popular to this day (see, e.g., Bingham and Kiesel 2002, Landsman and Valdez 2003, Hamada and Valdez 2008).

We confine our considerations to absolutely continuous multivariate distributions with mean vector \( \mu \) and positive definite dispersion matrix \( \Sigma \), as these are the ones that are relevant from a practical standpoint. A distribution of this type is said to be elliptical if its density \( f: \mathbb{R}^d \mapsto \mathbb{R} \) has the form

\[
f(x) = \frac{g((x - \mu)'\Sigma^{-1}(x - \mu))}{|\Sigma|^{1/2}}.
\]
where g: \mathbb{R} \rightarrow \mathbb{R} is a scalar function termed the 
-density generator. Thus, the density of an elliptical distribution is 
a function of the quadratic form (x - \mu) \Sigma^{-1} (x - \mu), and 
its level sets are elliptically symmetric in \mathbb{R}^{d}, which 
extplains the name. Inspecting formula (6), one recognizes 
the symmetric mGH distribution to be a member of the 
elliptic family. More generally, normal variance 
mixtures (and thus, in particular, the multivariate normal 
distribution) can be shown to be elliptical.

Elliptical distributions have several nice properties, 
which facilitate their application to practical problems. 
For example, linear combinations of the components 
of elliptical random vectors remain elliptical and have 
the same characteristic generator. In particular, the univariate 
marginal distributions of elliptical distributions are also 
elliptical and inherit the generator of the parent distribution. 
The essential property for portfolio-optimization purposes, however, was formulated by Embrechts et al. (2002): Suppose that 
X follows a d-dimensional elliptical distribution. Let \eta_j = \mathbb{E}(X_j), and let \phi be a positive homogeneous, 
translation-invariant risk measure. Define the subset of portfolios having expected return \rho as 
\mathcal{X} := \{x \in \mathbb{R}^d : \phi(x) \leq \rho, 1^T x = 1\}.

Then 
\text{argmin}_{x \in \mathcal{X}} \phi(-x^T X) = \text{argmin}_{x \in \mathcal{X}} \var(X^T x).

Thus, for an elliptical portfolio distribution, instead of 
solving a portfolio optimization problem with a 
positive homogeneous, translation-invariant risk measure as 
objective function under the condition that a given expected return is attained, one can solve the corresponding 
problem with the variance as objective function. Hence, the 
optimization problem reduces to a simple Markowitz-type optimization. Furthermore, it is important to note that the optimal allocation will be independent of the risk measure used. Of course, the values of the objective functions will differ in general, but the set of solutions will not. The above insight applies to risk measures such as VaR and CVaR, since both are positive homogeneous and translation invariant.

Now let \mathcal{X} be an elliptical mGH distributed random variable, i.e. \mathcal{X} \sim \text{GH}(\lambda, \chi, \psi, \mu, 0, \Sigma), and let \mathcal{X} \in \mathbb{R}^d (with \Sigma = 1) be the portfolio composition. Then \mathcal{X} \sim \text{GH}(\lambda, \chi, \psi, \mu, 0, \Sigma), and thus \mathcal{X}^T is also elliptical. Moreover, \mathbb{E}(\mathcal{X}^T) = x^T \mu \text{ and } \var(\mathcal{X}^T) = \mathbb{E}(W)^T \Sigma x, \text{ where } W \sim \text{mGH}(\lambda, \chi, \psi). \text{ Therefore, a portfolio optimization problem of the above type can be recast as the quadratic program }

\begin{equation}
\begin{aligned}
\min & \ x^T \Sigma x, \\
\text{subject to} & \ x \in \mathcal{X} = \{x \in \mathbb{R}^d : \phi(x) \leq \rho, 1^T x = 1\}.
\end{aligned}
\end{equation}

The advantage this formulation offers over the general (non-elliptical) mGH case is twofold. First, the simplicity and additional structure of the objective function can be exploited to solve quadratic programs more efficiently than general convex programs using dedicated quadratic optimization algorithms. Second, once an optimal solution \(x^*\) of the above problem has been found, it has a universal character: Not only does it solve the Markowitz-type variance minimization problem above, but it also minimizes portfolio VaR and portfolio CVaR for all levels \(\beta\). Thus, if the aim is to minimize CVaR for different levels \(\beta\), one needs to solve the optimization problem only once. This is in contrast to the situation in the general (non-elliptical) case, where optimization with respect to different risk measures or different levels \(\beta\) typically leads to different optimal allocations. Once the quadratic program has been solved, one may wish to calculate the CVaR corresponding to the optimal solution \(x^*\), which can be done using formula (7).

4.3. Robust portfolio optimization using Worst Case CVaR

The applicability of a portfolio optimization approach to real-world problems not only hinges on its numerical tractability but also crucially depends on its robustness: Small changes in input data should have only a minor impact on optimization results. This insight has spurred interest in robust portfolio optimization approaches (see, e.g., El Ghaoui et al. 2003, Goldfarb and Iyengar 2003, Halldorsson and Tütüncü 2003, or Zhu and Fukushima 2009). The central idea of robust portfolio optimization is to use uncertainty sets for the unknown parameters instead of only point estimates, and to compute portfolios whose worst-case performance (meaning the performance under the least favorable parameters in the uncertainty set) is optimal. In this section, we develop a robust optimization approach in the mGH framework, using Worst Case Conditional Value at Risk (WCVaR) as a risk measure. The resulting robust optimization problem will be demonstrated to be as tractable as its ‘classical’ (non-robust) CVaR-based counterpart examined above.

Let \(\mathcal{P}\) be a class of multivariate asset return distributions, let \(\mathcal{X}^p\) be a random vector of asset returns with distribution \(P \in \mathcal{P}\), and let \(x \in \mathcal{X}\) be a vector of portfolio weights. The WCVaR of a portfolio with weights \(x\) at level \(\beta \in (0, 1)\) is defined as

\[ WCVaR^\beta(x) = \sup_{P \in \mathcal{P}} \text{VaR}_\beta(-x^T \mathcal{X}^p). \]

As demonstrated by Zhu and Fukushima (2009), WCVaR inherits subadditivity, positive homogeneity, monotonicity, and translation invariance from CVaR—and therefore, just as CVaR, is a coherent risk measure. Moreover, the authors show that WCVaR is convex in \(x\).

The robust counterpart of the classical portfolio optimization problem involves minimizing WCVaR on a non-empty, compact, convex set of portfolio weights \(\mathcal{X} \subset \mathbb{R}^d_{++} \) as

\[ \min_{x \in \mathcal{X}} WCVaR^\beta(x). \]

The solution \(x^*\) of (R1) will then be the allocation with optimal worst-case properties.
4.3.1. Robust optimization within the mGH class. Our objective in this subsection is to state a robust optimization problem in the mGH class, making use of its specific characteristics. First, a parametric family \( P \) of distributions needs to be specified. The parameter space can be chosen in several ways; see Bertsimas et al. (2008) for an overview of different approaches. We consider separable uncertainty sets, which have been used extensively in the literature on robust portfolio optimization (e.g. by Halldórsson and Tütüncü 2003, Tütüncü and Koenig 2004, and Kim and Boyd 2007).

Assume that \( P \triangleq \{ GH_d(\lambda, \chi, \psi, \mu, \gamma, \Sigma) : (\mu, \gamma, \Sigma) \in \mathcal{M} \} \) is a family of mGH distributions with \( \lambda, \chi, \psi \) fixed. \((\mu, \gamma, \Sigma)\) is assumed to be an element of a separable polyhedral uncertainty set \( \mathcal{M} \triangleq I_\mu \times I_\gamma \times I_\Sigma \) with

\[
I_\mu = \{ \mu \in \mathbb{R}^d : \mu_L \leq \mu \leq \mu_U \},
I_\gamma = \{ \gamma \in \mathbb{R} : \gamma_L \leq \gamma \leq \gamma_U \},
I_\Sigma = \{ \Sigma \in \mathbb{R}^{d \times d} : \Sigma_L \leq \Sigma \leq \Sigma_U, \Sigma \text{ pos. definite} \}
\]

compact intervals. All inequalities in the set definitions are to be understood component-wise. Although the parameters governing the mixing distribution are kept fixed, our approach is flexible enough to incorporate uncertainty not only in the means and covariances of the return distribution but also in skewness and kurtosis, since \( \gamma \), which influences the latter two moments, is incorporated into the uncertainty set.

Since \( \mathcal{M} \) is compact, the supremum is attained:

\[
WCVaR_{\alpha}^D(x) = \sup_{P \in \mathcal{P}} CVaR_{\alpha}(-x'X^P)
\]

Lemma 4.1: Let \( X \subset \mathbb{R}_+^d \) be a convex set.

(a) \( F_P(x, \alpha, \lambda, \chi, \psi, \mu, \gamma, \Sigma) \) is component-wise monotonically decreasing in \( \mu \) and \( \gamma \) and component-wise monotonically increasing in \( \Sigma \). In particular, for any \((x, \alpha) \in X \times \mathbb{R}_+^d \),

\[
\max_{(\mu, \gamma, \Sigma) \in \mathcal{M}} F_P(x, \alpha, \lambda, \chi, \psi, \mu, \gamma, \Sigma) = F_P(x, \alpha, \lambda, \chi, \psi, \mu_L, \gamma_L, \Sigma_U).
\]

(b) \( F_P(x, \alpha, \lambda, \chi, \psi, \mu, \gamma, \Sigma) \) is convex in \((x, \alpha) \) on \( X \times \mathbb{R}_+^d \).

Proof: See appendix B. \( \square \)

Lemma 4.1 is key to the proof of proposition 4.2.

Proposition 4.2: The following relations hold:

(a) \( WCVaR_{\alpha}^D(x) = \min_{w \in \mathbb{R}_+} F_P(x, \alpha, \lambda, \chi, \psi, \mu_L, \gamma_L, \Sigma_U), \)

with \( F_P(x, \alpha, \lambda, \chi, \psi, \mu_L, \gamma_L, \Sigma_U) \) convex in \( \alpha \) on \( \mathbb{R}_+ \);

(b) \( \min_{x \in X} WCVaR_{\alpha}^D(x) = \min_{(x, \alpha) \in X \times \mathbb{R}_+} F_P(x, \alpha, \lambda, \chi, \psi, \mu_L, \gamma_L, \Sigma_U), \) (R1”)

with \( F_P(x, \alpha, \lambda, \chi, \psi, \mu_L, \gamma_L, \Sigma_U) \) convex in \((x, \alpha) \) on \( X \times \mathbb{R}_+ \).

Proof: See appendix C. \( \square \)

The above proposition demonstrates that the WCVaR for any portfolio \( x \in X \) is attained for the parameters \( \mu_L, \gamma_L \) and \( \Sigma_U \). Furthermore, it shows that the original robust optimization problem can be massively simplified and cast as a convex program (R1”), which is as efficiently solvable as the corresponding classical optimization problem.

4.3.2. Towards an efficient formulation of the robust problem. In the following, we demonstrate how (R1”) can be substantially simplified by exploiting the specific structure of the mGH class. To this end, we collect some essential properties of \( F_P \) in the following lemma.

4.3.3. Introducing a minimum return constraint. This subsection examines robust portfolio optimization using WCVaR under a minimum constraint for the worst-case expected portfolio return. Let \( P, X^P \), and \( X \) be defined as above. The worst-case expected return for a portfolio \( x \in X \) is

\[
\min_{x \in X} \mathbb{E}(x'X^P) = \min_{(x, \alpha) \in X \times \mathbb{R}_+} \sum_{i=1}^d x_i \mu_i + \gamma \mathbb{E}(W)
\]

with \( W \sim GIG(\chi, \psi, \lambda) \) and \( \mathbb{E}(W) \) as in formula (2). Apparently, the worst-case expected portfolio return is assumed for the parameter vectors \( \mu_L \) and \( \gamma_L \), which also appear in the parameter set for which the WCVaR is attained. If we use the notation \( v_{WCV} = \mu_L + \gamma_L \mathbb{E}(W) \) for the worst-case expected return vector, then the robust optimization problem under a minimum return constraint for the worst-case expected return can be stated as the convex program

\[
\begin{array}{ll}
\min_{(x, \alpha)} & F_P(x, \alpha, \lambda, \chi, \psi, \mu_L, \gamma_L, \Sigma_U), \\
\text{subject to} & x \in X = \{ x \in \mathbb{R}^d_+ : x'1 = 1, v_{WCV} x \geq 0 \}, \\
& \alpha \in \mathbb{R}_+.
\end{array}
\]
5. Numerical results

In this section, we present a numerical example, based on empirical data, in which the theory developed above is applied. We consider four indices: two stock price indices, namely the Dow Jones Eurostoxx 50 and the S&P 500, a bond index, namely the iBoxx Euro, and a commodity index, namely the S&P GSCI. We calibrate an mGH model to 210 weekly returns of these indices observed between September 2004 and September 2008. Using the EM algorithm (see, e.g., McNeil et al. 2005) for calibration, we obtained the following parameter estimates for the joint return distribution, where the order of the elements in the following vectors and matrices corresponds to the order in the above enumeration:

\[ \lambda = 1.725, \quad \gamma = 2.714, \quad \psi = 6.766, \]
\[ \mu = 10^{-3} \begin{pmatrix} 12.465 \\ 6.167 \\ 1.265 \\ 8.751 \end{pmatrix}, \quad \Sigma = 10^{-4} \begin{pmatrix} 3.822 & 2.712 & -0.202 & -0.771 \\ 2.712 & 3.096 & -0.203 & -0.911 \\ -0.202 & -0.203 & 0.186 & -0.005 \\ -0.771 & -0.911 & -0.005 & 11.101 \end{pmatrix}, \]

Let \( X \) follow an mGH distribution with the above parameters. Then

\[ \mathbb{E}(X) = 10^{-3} \begin{pmatrix} 1.451 \\ 1.034 \\ 0.474 \\ 2.187 \end{pmatrix}, \]
\[ \text{Cov}(X) = 10^{-4} \begin{pmatrix} 4.072 & 2.829 & -0.184 & -0.622 \\ 2.829 & 3.150 & -0.195 & -0.841 \\ -0.184 & -0.195 & 0.187 & 0.006 \\ -0.622 & -0.841 & 0.006 & 11.105 \end{pmatrix}, \]
\[ \text{Corr}(X) = \begin{pmatrix} 1 & 0.777 & -0.234 & -0.086 \\ 0.777 & 1 & -0.275 & -0.144 \\ -0.234 & -0.275 & 1 & -0.014 \\ -0.086 & -0.144 & -0.014 & 1 \end{pmatrix}, \]
\[ \text{Skew}(X) = \begin{pmatrix} -0.335 \\ -0.177 \\ -0.113 \\ -0.122 \end{pmatrix}, \quad \text{ExcKurt}(X) = \begin{pmatrix} 0.738 \\ 0.653 \\ 0.632 \\ 0.634 \end{pmatrix}. \]

Apparently, the GSCI commodity index features the highest expected return (0.2187 percentage points per week), while the iBoxx bond index has the lowest expected return. The volatility of the iBoxx is by far the lowest, the stock price indices exhibit a similar level of volatility, and the GSCI can be seen to be substantially more volatile than the other indices. The stock price indices are strongly positively correlated, but slightly negatively correlated to both the bond and the commodity indices, whereas the latter two are almost uncorrelated. All indices exhibit only moderate negative skewness and excess kurtosis when observed on a weekly basis. However, as Konikov and Madan (2002) point out, the skewness of the marginal distribution of a Lévy process decreases as the reciprocal of the square root of time, whereas its excess kurtosis decreases as the reciprocal of time. Bearing in mind that the mGH distribution is infinitely divisible and therefore can be regarded as the marginal distribution of a multivariate Lévy process, we can conclude that for daily returns, both negative skewness and (especially) excess kurtosis would be considerably more pronounced (scale the weekly figures by \( \sqrt{5} \) and 5, respectively).

Based on these parameters and \( \beta = 0.95 \), we perform a mean-CVaR optimization under minimum return constraints, i.e. we solve (P2”). Figure 1 presents the mean-CVaR efficient frontier (top left graph), the compositions of the efficient portfolios (top right graph), and the CVaR contributions of the individual assets in the efficient portfolios (bottom graph). The weekly CVaR ranges from 0.79 percentage points for the minimum-CVaR portfolio (which at the same time has the minimum expected return among all efficient portfolios) to 7.11 percentage points for the portfolio with maximum expected return. The minimum-CVaR portfolio is made up mainly of a position in the iBoxx, while encompassing only small positions in the S&P 500 and the GSCI. The maximum-return portfolio consists solely of a position in the GSCI, the index with maximum expected return.

The graph at the bottom of figure 1 displays the CVaR contributions of the individual assets given the portfolio compositions shown in the top right graph. The upper boundary of the colored area corresponds exactly to the efficient frontier shown in the top left graph, the only difference being that, in the top left graph, expected return corresponds to the vertical axis and CVaR to the horizontal axis. An interesting phenomenon can be witnessed when relating portfolio compositions to CVaR contributions: Although the weights of the individual assets change linearly when moving towards higher returns, their risk contributions do not. This effect becomes particularly evident for portfolios that consist only of the Eurostoxx 50 and the GSCI: A linear decrease in the weight of the Eurostoxx 50 induces a superlinear decrease in its risk contribution. This observation can be explained by the relatively higher diversification benefits—and thus the relatively lower risk contributions—associated with smaller positions.

Now we tackle the corresponding robust portfolio optimization problem (R1”). First of all, the uncertainty set needs to be specified. In the spirit of the approach taken by Kim and Boyd (2007), we assume that the uncertainty sets arise from the base-case parameters.
Thus, the worst-case returns will have lower means, more pronounced negative skewness (recall that all components of \( \gamma \) are negative in our example, and thus \( \gamma_L < \gamma \)), and higher variances and covariances, leading to lower expected returns and higher risk of efficient portfolios. The deteriorated risk–return profile becomes evident when comparing the robust efficient frontier (i.e. the worst-case optimal portfolios) displayed in the top left graph of figure 2 with the efficient frontier in the base case.

Comparing the compositions of base-case and worst-case efficient portfolios (top right graphs of figures 1 and 2, respectively), one recognizes that the weight of the iBoxx position (the least risky investment) has increased throughout the full spectrum of expected returns, while the weight of the Eurostoxx 50 is essentially zero throughout. Looking more closely at the CVaR contributions, we note that even for portfolios where the nominal weight of the GSCI is significantly smaller than that of the iBoxx, the GSCI’s risk contribution can be the higher of the two, on account of its far higher volatility. As in the baseline scenario, we witness linearly decreasing portfolio weights giving rise to superlinearly decreasing CVaR contributions.

By proposition 4.2, the worst-case scenario is fully determined by the interval bounds \( \mu_L, \gamma_L \), and \( \Sigma_U \), which can be calculated to be

\[
I_{\mu} = \{\tilde{\mu} \in \mathbb{R}^d : \mu_L \leq \tilde{\mu} \leq \mu_U\} = \{\tilde{\mu} \in \mathbb{R}^d : |\mu_i - \tilde{\mu}_i| \leq 0.1 \mu_i\},
\]

\[
I_{\gamma} = \{\tilde{\gamma} \in \mathbb{R}^d : \gamma_L \leq \tilde{\gamma} \leq \gamma_U\} = \{\tilde{\gamma} \in \mathbb{R}^d : |\gamma_i - \tilde{\gamma}_i| \leq 0.1 \gamma_i\},
\]

\[
I_{\Sigma} = \{\tilde{\Sigma} \in \mathbb{R}^{d \times d} : \Sigma_L \leq \tilde{\Sigma} \leq \Sigma_U, \tilde{\Sigma} \text{ pos. definite}\}
\]

\[
= \{\tilde{\Sigma} = (\tilde{\sigma}_j) \in \mathbb{R}^{d \times d} : |\sigma_{ij} - \tilde{\sigma}_{ij}| \leq 0.1 \sigma_{ij}, \tilde{\Sigma} \text{ pos. definite}\}.
\]

By proposition 4.2, the worst-case scenario is fully determined by the interval bounds \( \mu_L, \gamma_L \), and \( \Sigma_U \), which can be calculated to be

\[
\mu_L = 0.9 \mu, \quad \gamma_L = 1.1 \gamma,
\]

\[
\Sigma_U = 10^{-4} \begin{pmatrix}
4.204 & 2.983 & -0.182 & -0.694 \\
2.983 & 3.405 & -0.183 & -0.819 \\
-0.182 & -0.183 & 0.204 & -0.004 \\
-0.694 & -0.819 & -0.004 & 12.117
\end{pmatrix},
\]

where the positive definiteness of \( \Sigma_U \) is easily verified. Thus, the worst-case returns will have lower means, more pronounced negative skewness (recall that all components of \( \gamma \) are negative in our example, and thus \( \gamma_L < \gamma \)), and higher variances and covariances, leading to lower expected returns and higher risk of efficient portfolios.

Comparing the compositions of base-case and worst-case efficient portfolios (top right graphs of figures 1 and 2, respectively), one recognizes that the weight of the iBoxx position (the least risky investment) has increased throughout the full spectrum of expected returns, while the weight of the Eurostoxx 50 is essentially zero throughout. Looking more closely at the CVaR contributions, we note that even for portfolios where the nominal weight of the GSCI is significantly smaller than that of the iBoxx, the GSCI’s risk contribution can be the higher of the two, on account of its far higher volatility. As in the baseline scenario, we witness linearly decreasing portfolio weights giving rise to superlinearly decreasing CVaR contributions.

Figure 1. Efficient frontier, portfolio composition, and CVaR contributions in the base case.
Finally, we compare the performance of both classical and robust portfolios in the baseline and worst-case scenarios (see figure 3). The blue and red curves represent the efficient frontiers of figures 1 and 2, respectively. The black curve reflects the performance of robust efficient portfolios under the parameters of the base case, while the green curve represents the performance of classical efficient portfolios in the worst case. As evidenced by the green curve, classical efficient portfolios perform quite badly should the worst case obtain, and some even lead to negative expected returns. Moreover, because of the lack of monotonicity in the green curve, some allocations are severely inefficient, leading to lower expected returns while at the same time being riskier than other feasible portfolios. In contrast, robust portfolios (red curve) perform substantially better than the classical ones in the worst case while being only slightly worse in the base case. Furthermore, except for very low-risk portfolios, robust portfolios can be seen to exhibit a monotonic relation of risk and return in the base case. It is worth noting that, since the classical and robust minimum-CVaR portfolios are not identical, both the blue and black curves and the red and green curves start from slightly different (though visually almost indistinguishable) points, respectively. Overall, one notes that robust portfolios perform reasonably well in both scenarios, while classical portfolios exhibit pronounced sensitivity to the scenario that actually obtains and thus might lead to severely inefficient allocations.
6. Conclusion

In this paper, we develop a tractable yet flexible approach to portfolio risk management and portfolio optimization based on the mGH distribution. As the normal distribution is a limiting case of the mGH class, the approach presented in this paper can be considered a natural generalization of the Markowitz approach. Exploiting the fact that portfolios whose constituents follow an mGH distribution are univariate GH distributed, we provide analytical formulas for portfolio CVaR and the contributions of individual assets thereto. Then we demonstrate how to efficiently compute optimal portfolios in the mGH framework. Using WC VaR as a risk measure, we formulate a robust optimization approach within the mGH framework, which is shown to lead to optimization problems that can be solved as efficiently as their classical counterparts. Finally, we apply our insights to a numerical example, highlighting the advantages of robust portfolios.

Acknowledgements

SK gratefully acknowledges financial support from the German Research Council (DFG) and helpful discussions with Thomas Liebmann.

References


Appendix A: Proof proposition 3.1

Proof: Define a matrix

\[
B = \begin{pmatrix}
0 & \ldots & 0 & x_1 & 0 & \ldots & 0
\end{pmatrix} \in \mathbb{R}^{2 \times d}.
\]

Then

\[
\begin{pmatrix}
x_1 & X_1 \\
x'X
\end{pmatrix} = BX
\]

\[
\sim GH \left( \lambda, \chi, \psi, \begin{pmatrix} x_1 \mu_i \end{pmatrix}, \begin{pmatrix} x_i \sigma_i \\
x_i \sum_j x_j \sigma_{ij} \\
x_i \sum_k \sum_l x_k x_l \sigma_{ijkl}
\end{pmatrix}, \left( \begin{pmatrix} x_1 \gamma_1 \\
x_i \sum_j x_j \sigma_{ij} \\
x_i \sum_k \sum_l x_k x_l \sigma_{ijkl}
\end{pmatrix} \right) \right).
\]

Denote by \( f_{GH} \) the density function of the above distribution. We find that

\[
x'X \sim GH \left( \lambda, \chi, \psi, X, x', \gamma \right),
\]

which can be used to infer that \( Var_{R_1}(x'X) = GH^{-1}(1 - \beta) \). Now, using the definition of conditional expectation,

\[
\begin{align*}
\mathbb{E}[-x_i & | -x' X \geq Var_{R}(x'X)] \\
& = -\mathbb{E}[x_i | x' X \leq Var_{R_1}(x'X)] \\
& = -\frac{1}{1 - \beta} \mathbb{E}[x_i | x' X \leq GH^{-1}(1 - \beta)] \\
& = -\frac{1}{1 - \beta} \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} y_1 \cdot f_{GH}(y_1, y_2) dy_2 dy_1.
\end{align*}
\]

\[\square\]

Appendix B: Proof of lemma 4.1

Proof:

(a) We show component-wise monotonicity of \( F_\beta(x, \alpha; \lambda, \chi, \psi, \mu, \gamma, \Sigma) \) in \( \mu, \gamma \) and \( \Sigma \). We have

\[
F_\beta(x, \alpha; \lambda, \chi, \psi, \mu, \gamma, \Sigma) = \alpha - \frac{1}{1 - \beta} \int_{-\infty}^{\infty} (z + \alpha) f_{GH}(z; \lambda, \chi, \psi, x', \gamma', x' \Sigma x) dz
\]

\[
= \alpha - \frac{1}{1 - \beta} \mathbb{E}(\{Z + \alpha\} \cdot \{Z \leq -\alpha\})
\]

where \( Z \sim GH \left( \lambda, \chi, \psi, X, x', \gamma', x' \Sigma x \right) \). Letting \( Y \sim N(0, 1) \) and \( W \sim G \left( \lambda, \chi, \psi \right) \) be independent random variables, and \( f_Y \) and \( f_{GH} \) the corresponding densities, and using the definition of an mGH random variable as a normal mean–variance mixture, we find that \( Z = d X' \mu + W' x' \gamma + \sqrt{W x' \Sigma x} Y \), and that

\[
\begin{align*}
\mathbb{E}(Z + \alpha) \cdot \mathbb{E}(Z \leq -\alpha)
& = -\mathbb{E}(\{x' \mu + W' x' \gamma + \sqrt{W x' \Sigma x} Y + \alpha\}) \\
& \cdot \mathbb{E}(\{x' \mu + W' x' \gamma + \sqrt{W x' \Sigma x} Y \leq -\alpha\})
\end{align*}
\]

\[
= \int_{0}^{\infty} \neg \mathbb{E}(x' \mu + W' x' \gamma + \sqrt{W x' \Sigma x} Y + \alpha) \\
\cdot \mathbb{I}(x' \mu + W' x' \gamma + \sqrt{W x' \Sigma x} Y \leq -\alpha) \mid W = w \\
\cdot f_{GH}(w; \lambda, \chi, \psi) dw
\]

\[
= \int_{0}^{\infty} \left( \int_{-\infty}^{\infty} - (x' \mu + W' x' \gamma + \sqrt{W x' \Sigma x} Y + \alpha) f_Y(y) dy \right) \\
\cdot f_{GH}(w; \lambda, \chi, \psi) dw
\]

\[
= \int_{0}^{\infty} \left( \int_{0}^{\infty} (x' \mu + W' x' \gamma + \sqrt{W x' \Sigma x} Y + \alpha) f_Y(y) dy \right) \\
\cdot f_{GH}(w; \lambda, \chi, \psi) dw,
\]

with an upper integration bound

\[
g(w) = \frac{-\alpha + x' \mu + W' x' \gamma}{\sqrt{W x' \Sigma x}}.
\]

Define the auxiliary function \( h: \mathbb{R}_+ \times \mathbb{R}^2 \times \mathbb{R}_+ \mapsto \mathbb{R} \)

\[
h(w, \mu, \gamma, \sigma) = \int_{-\infty}^{\infty} (\mu w + \gamma w^2 + \sigma w^3) f_Y(y) dy.
\]

Taking partial derivatives of \( h \) and setting \( z = (\alpha + \mu + \gamma w)/\sqrt{\sigma w} \) yields

\[
\frac{d h(z, \mu, \sigma)}{d \mu} = \Phi(z) - 1 < 0,
\]

\[
\frac{d h(z, \mu, \sigma)}{d \gamma} = w \Phi(z) - w < 0,
\]

\[
\frac{d h(z, \mu, \sigma)}{d \sigma} = \frac{w^2}{2 \sqrt{2 \pi}} \exp \left( -\frac{z^2}{2} \right) > 0,
\]

where \( w > 0 \) and \( \Phi(z) \) is the cumulative density function of the normal distribution. Using \( h, F_\beta \) can be rewritten in the form

\[
F_\beta(x, \alpha; \lambda, \chi, \psi, \mu, \gamma, \Sigma) = \alpha + \frac{1}{1 - \beta} \int_{0}^{\infty} h(w, x', \gamma', x' \Sigma x) f_{GH}(w; \lambda, \chi, \psi) dw.
\]

As \( x \in X \subseteq \mathbb{R}^d \)—and, in particular, is non-negative—it follows that

\[
\frac{d F_\beta(x, \alpha; \lambda, \chi, \psi, \mu, \gamma, \Sigma)}{d \mu} \leq 0,
\]

\[
\frac{d F_\beta(x, \alpha; \lambda, \chi, \psi, \mu, \gamma, \Sigma)}{d \gamma} \leq 0,
\]

\[
\frac{d F_\beta(x, \alpha; \lambda, \chi, \psi, \mu, \gamma, \Sigma)}{d \sigma} \geq 0,
\]

where \( \Sigma = (\sigma_{ij}) \). The remaining statement in part (a) of lemma 4.1 follows directly from this and the definition of \( M \).
Appendix C: Proof of proposition 4.2

Proof:

(a) Recall that

\[ WCVaR^\beta(x) = \max_{(\mu, \gamma, \Sigma) \in \Sigma} \min_{\alpha} \left( F_\beta(x; \alpha; \lambda, \chi, \psi, \mu, \gamma, \Sigma) \right) \]

Using the minimax inequality and part (a) of lemma 4.1, we find that

\[ \max_{(\mu, \gamma, \Sigma) \in \Sigma} \min_{\alpha} \left( F_\beta(x; \alpha; \lambda, \chi, \psi, \mu, \gamma, \Sigma) \right) \leq \min_{\alpha} \max_{(\mu, \gamma, \Sigma) \in \Sigma} \left( F_\beta(x; \alpha; \lambda, \chi, \psi, \mu, \gamma, \Sigma) \right) \]

\[ = \min_{\alpha} \left( F_\beta(x; \alpha; \lambda, \chi, \psi, \mu, \gamma, \Sigma) \right) \]

Convexity was proved in part (b) of lemma 4.1.

(b) By virtue of part (a),

\[ \min_{x \in X} WCVaR^\beta(x) = \min_{x \in X} \left( \min_{\alpha} \left( F_\beta(x; \alpha; \lambda, \chi, \psi, \mu, \gamma, \Sigma) \right) \right) \]

\[ = \min_{(x, \alpha) \in X \times \alpha} \left( F_\beta(x; \alpha; \lambda, \chi, \psi, \mu, \gamma, \Sigma) \right). \]

The statement on convexity follows from part (b) of lemma 4.1.

On the other hand,

\[ \max_{(\mu, \gamma, \Sigma) \in \Sigma} \min_{\alpha} \left( F_\beta(x; \alpha; \lambda, \chi, \psi, \mu, \gamma, \Sigma) \right) \geq \min_{\alpha} \min_{\alpha} \left( F_\beta(x; \alpha; \lambda, \chi, \psi, \mu, \gamma, \Sigma) \right) \]

so that equality follows:

\[ \max_{(\mu, \gamma, \Sigma) \in \Sigma} \min_{\alpha} \left( F_\beta(x; \alpha; \lambda, \chi, \psi, \mu, \gamma, \Sigma) \right) = \min_{\alpha} \left( F_\beta(x; \alpha; \lambda, \chi, \psi, \mu, \gamma, \Sigma) \right). \]