Optimal Portfolio Selection Based on Value-at-Risk and Expected Shortfall Under Generalized Hyperbolic Distribution

B. A. Surya *
School of Business and Management
Bandung Institute of Technology

R. Kurniawan
New Frontier Solutions Pte. Ltd
Jakarta - Indonesia

February 10, 2012

Abstract

This paper discusses optimal portfolio selection problems under value-at-risk and expected shortfall as the risk measures. We employ multivariate generalized hyperbolic distribution as the joint distribution for the risk factors of underlying portfolio assets, which include stocks, currencies and bonds. Working under this distribution, we find the optimal portfolio strategy.

Keywords: Multivariate generalized hyperbolic distribution; Value-at-risk; Expected shortfall; Portfolio optimization

1 Introduction

It’s been well known in years that financial data are often not normally distributed. They exhibit properties that normally distributed data don’t have. For example, it has been observed that empirical return distributions almost always exhibit excess kurtosis and heavy tail (Cont, [11]). Mandelbrot [24] also concluded that the logarithm of relative price changes on financial and commodity markets exhibit a heavy-tailed distribution. More recently, Madan and Seneta [23] proposed a Lévy process with Variance Gamma distributed increments to model log price processes.

*Corresponding email address: {budhi.surya}@sbm-itb.ac.id.
Variance Gamma itself is a special case of Generalized Hyperbolic (GHYP) distribution which was originally introduced by Barndorff-Nielsen [4]. Other subclasses of the GH distribution also proved to provide an excellent fit to empirically observed increments of financial log price processes, in particular, log return distributions, such as the Hyperbolic distribution (Eberlein and Keller [12]), the Normal Inverse Gaussian Distribution (Barndorff-Nielsen [5]), the Generalized Hyperbolic skew Student t distribution (Aas and Haff [1]), and most notably, the Student t and normal distributions themselves. These give enough reasons for the popularity of Generalized Hyperbolic family distributions: they provide a good fit to financial return data and also consistent with continuous-time models where the logarithm of asset price processes follow Lévy processes.

Since it is clear that Gaussian distribution does not provide a good fit to return data anymore and that volatility cannot perform well under non-normal framework, alternative risk measures must be considered. They are also needed to account for the extreme movements of assets return reflected by the heavy-tailed distributions. A risk measure called Value-at-Risk (VaR) can satisfy this need. Instead of measuring return deviations from its mean, it determines the point of relative loss level that is exceeded at a certain degree. Consequently, when suitably adjusted, it can measure the behavior of negative return distributions at a point far away from the expected return. Hence, it is able to take into account the extreme movements of assets return. Furthermore, it gives an easy representation of potential losses since it is none other than the quantile of loss distribution, when the distribution is continuous. This is the aspect that has mainly contributed to its popularity. However, it has a serious drawback. Although it is coherent for elliptical distributions, when applied to nonelliptical distributions it can lead to a centralized portfolio (Artzner et. al. [3]), which is against the portfolio diversity principle. In the same case, it is also a generally nonconvex function of portfolio weights, hence making portfolio optimization an expensive computational problem.

The more recent and popularly used risk measure is the expected shortfall (ES). It was made popular by Artzner et al. [3] in response to VaR’s drawback. Unlike VaR, it always leads to a diversity portfolio. It is also a coherent risk measure. Furthermore, this measure takes into account the behaviour of return distributions at and beyond a certain point. So, like VaR, it shows the tail behaviour of the distributions, but, at a much wider scope. Ultimately, these attributes make it a seemingly better risk measure than its classic counterpart.

The remainder of this paper is organized as follows. In Section 2 we briefly discuss general property of GHYP distribution. More details of GHYP calibration using EM algorithm is discussed in the Appendix. Section 3 briefly discusses value-at-risk and expected shortfall. The asset structures of portfolio is elaborated in more details in Section 4. Section 5 discusses the profit and loss (P&L) distribution in terms of the multivariate generalized hyperbolic distribution. Portfolio optimization problems are formulated in Section 6. Section 7 discusses numerical examples on real data of the theoretical framework discussed in the above sections. Section 8 concludes this paper.
2 Generalized Hyperbolic Distribution

Before presenting the Generalized Hyperbolic (GH) distribution, it is essential to present the underlying distribution that builds it.

**Definition 2.1 Generalized Inverse Gaussian distribution (GIG).** The random variable \( W \) is said to have a Generalized Inverse Gaussian (GIG) distribution, written by \( W \sim N^\sim(\lambda, \chi, \psi) \), if its probability density function is

\[
h(w; \lambda, \chi, \psi) = \frac{\chi^{-\lambda}}{2K_\lambda(\sqrt{\chi\psi})} w^{\lambda-1} \exp\left(-\frac{1}{2} \left(\chi w^{-1} + \psi w\right)\right), \quad w, \chi, \psi > 0, \quad \lambda \in \mathbb{R}
\]

(2.1)

where \( K_\lambda(\sqrt{\chi\psi}) \) is a modified bessel function of the second kind with index \( \lambda \).

The Generalized Hyperbolic distribution belongs to the normal mean variance mixture distribution class defined by the following definition.

**Definition 2.2 Multivariate Normal Mean Variance Mixture Distribution (MNMVM).** A random variable \( X \in \mathbb{R}^d \) is MNMVM distributed if it has the following representation

\[
X = \mu + W\gamma + \sqrt{W}AZ,
\]

(2.2)

where \( \mu, \gamma \in \mathbb{R}^d \) and \( A \in \mathbb{R}^{d \times k} \) are the distribution parameters, \( Z \sim N_k(0, I_k) \) is a standard multivariate normal random variable, and \( W \) is a nonnegative mixing random variable. Also note that the representation requires \( \Sigma := AA^t \) to be positive definite. In univariate version of the model, \( \Sigma \) is replaced by \( \sigma^2 \).

This new class of distribution is proposed by McNeil et. al. [27]. The most important aspect of this distribution is that it provides the very first representation for the Generalized Hyperbolic distribution that has the benefit of preserving its parameters under linear transformation. Another representation for the Generalized Hyperbolic distribution was previously proposed by Barndorff-Nielsen [6], but it does not show this special property. This representation also proves to be very useful in the calibrating process. The Generalized Hyperbolic distribution can now be defined from representation (2.2).

**Definition 2.3 Generalized Hyperbolic Distribution (GH).** A random variable \( X \in \mathbb{R}^d \) is said to be a GH distributed random variable, denoted by

\[
X \sim GH(\lambda, \chi, \psi, \mu, \Sigma, \gamma)
\]

(2.3)

iff it has the representation (2.2) with \( W \sim N^\sim(\lambda, \chi, \psi) \) is a scalar GIG distributed random variable. Additionally, \( X \) is called symmetric iff \( \gamma = 0 \).
This representation is consistent with the definition of GH distribution first proposed by Barndorff-Nielsen [4] with pdf

\[
f(x) = c K_\lambda \left( \sqrt{\chi + (x - \mu)'\Sigma^{-1}(x - \mu)} \right) \left( \psi + \gamma'\Sigma^{-1}\gamma \right)^{\frac{d}{2} - \lambda} e^{(x - \mu)'\Sigma^{-1}\gamma},
\]

\[\chi, \psi > 0, \quad \lambda \in \mathbb{R}\]

with the normalizing constant

\[
c = \left( \frac{\sqrt{\chi\psi}}{2\pi} \right)^{\frac{d}{2}} |\Sigma|^{\frac{1}{2}} K_\lambda \left( \sqrt{\chi\psi} \right).
\]

Representation (2.2) gives a significant contribution in the linearity property of a generalized hyperbolic distribution. The following theorem plays central role in solving optimal portfolio selection problems, discussed in more details in Section 6.

**Theorem 2.4** If \(X \in \mathbb{R}^d\) is a Generalized Hyperbolic random variable, i.e, \(X \sim GH(\lambda, \chi, \psi, \mu, \Sigma, \gamma)\), then

\[
BX + b \sim GH(\lambda, \chi, \psi, B\mu + b, B\Sigma B', B\gamma),
\]

where \(B \in \mathbb{R}^{l \times d}\) and \(b \in \mathbb{R}^d\).

# 3 Risk Measures

This section briefly discusses general property of risk measures associated with uncertain loss in a portfolio due to the volatility of financial market. Hence, a measurement of risk must take account the randomness of the loss and is used to determine capital reserve to anticipate future loss. In this paper we employ value-at-risk and expected shortfall as risk measure, which we shall now discuss.

## 3.1 Value-at-Risk

**Definition 3.1 Value at Risk (VaR).** For a given \(\beta \in (0, 1)\), the VaR at level \(\beta\) for a portfolio loss \(L \in \mathcal{M}\) is defined by

\[
\text{VaR}_\beta(L) := \inf \{l \in \mathbb{R} : \mathbb{P}(L > l) \leq 1 - \beta\}.
\]

For a continuous and strictly positive loss density functions, its distribution function is continuous and strictly increasing. In this case, VaR can be simply expressed as the \(\beta\)-quantile of the loss distribution function \(F_L\):

\[
\text{VaR}_\beta(L) = F_L^{-1}(\beta).
\]
Artzner et al. [3] point out that VaR can lack subadditivity when applied to non-elliptical distributions, in particular the asymmetric GH distributions, which amounts to ignoring the benefit of portfolio diversification. Moreover, when applied in portfolio optimization problem, VaR is generally a non-convex function of portfolio weights. This non-convexity may lead to multiple local extrema. This renders portfolio optimization as computationally expensive problem.

Although it is generally not coherent, in the case of symmetric GH distributions and $\beta \geq 0.5$, VaR at level $\beta$ is subadditive.

**Theorem 3.2** Let $X$ be a $d$-dimensional symmetric Generalized Hyperbolic random variable with $X \sim GH(\lambda, \chi, \psi, \mu, \Sigma, 0)$. Then, for every $0.5 \leq \beta < 1$, $\text{VaR}_\beta : \mathcal{M} \to \mathbb{R}$ is coherent.

See McNeil et. al. [27] for proof.

### 3.2 Expected Shortfall

**Definition 3.3 Expected Shortfall (ES).** For a given $\beta \in (0, 1)$, ES is defined as the expectation of loss $L \in \mathcal{M}$ conditional on loss being at least $\text{VaR}_\beta(L)$:

$$\text{ES}_\beta(L) := \mathbb{E}[L | L > \text{VaR}_\beta(L)].$$  \hfill (3.3)

Acerbi and Tasche [2] first proved that ES is a coherent risk measure, hence retaining convexity property.

**Theorem 3.4** Let $X$ be a $d$-dimensional random variable. Then, for every $\beta \in (0, 1)$, $\text{ES}_\beta : \mathcal{M} \to \mathbb{R}$ is a coherent risk measure.

The following lemma, also from Acerbi and Tasche [2], can be useful to calculate the empirical Expected Shortfall.

**Lemma 3.1 (Acerbi and Tasche, 2002)** Let $L_1, \ldots, L_n$ be i.i.d. random variables with distribution equal to the distribution of some random variable $L$ and $L_{1:n}, \ldots, L_{n:n}$ be the corresponding order statistics such that $L_{1:n} \geq \ldots \geq L_{n:n}$. Then, for every $\beta \in (0, 1)$, with probability 1,

$$\lim_{n \to \infty} \frac{\sum_{i=1}^{\lfloor n(1-\beta) \rfloor} L_{i:n}}{\lfloor n(1-\beta) \rfloor} = \text{ES}_\beta(L)$$  \hfill (3.4)

where $\lfloor \cdot \rfloor$ is a floor function.

### 4 Portfolio Structure

The most vital aspect in portfolio optimization problem is the modeling of the portfolio risk. Since the risk comes from the portfolio loss value over some periods, it is important to define the the loss function of a portfolio. First, set to be used
for future reference. Denote by \( V(s) \) the portfolio value at time \( s \). For a given time horizon \( \Delta \), the loss of the portfolio over the period \([s, s + \Delta]\) is defined by

\[
L_{[s, s+\Delta]} := -(V(s + \Delta) - V(s)).
\] (4.1)

In establishing the portfolio theory, \( \Delta \) is assumed to be a fixed constant. In this case, it is more convenient to use the following definition

\[
L_{t+1} := L_{[t\Delta, (t+1)\Delta]} = -(V_{t+1} - V_t).
\] (4.2)

Here, time series notation is adopted where \( V_t := V(t\Delta) \). Any random variables with \( t \) as the subscript are assumed to be defined in similar way from here on.

In the context of risk management where the calendar time \( s \) is measured in years, if daily losses are being considered, \( \Delta \) can be set as \( \Delta = 1/250 \). In this case, \( V_t \) and \( V_{t+1} \) represent the portfolio value on days \( t \) and \( t + 1 \), respectively, and, \( L_{t+1} \) is the loss from day \( t \) to day \( t + 1 \). From here on equation (4.2) will be used to define the portfolio loss with \( t \) measured in the time horizon specified by \( \Delta \).

In standard practice, \( V_t \) is modelled as a function of time \( t \) and a \( d \)-dimensional random vector \( Z_t = (Z_{t,1}, \ldots, Z_{t,d})' \) of risk factors. In this work, they are assumed to follow some discrete stochastic process. Hence, \( V_t \) has representation

\[
V_t = f(t, Z_t)
\] (4.3)

for some measurable function \( f : \mathbb{R}_+ \times \mathbb{R}^d \rightarrow \mathbb{R} \). The choice of \( f \) depends on the assets contained in the considered portfolio, while the risk factors are usually chosen to be the logarithmic price of financial assets, yields or logarithmic exchange rates. In this paper, the risk factors are chosen to take one of these forms since the distribution models of their time increments have been empirically known as have been made clear in the introduction section.

Define the increment process \((X_t)\) by \( X_t := Z_t - Z_{t-1} \). Using the mapping (4.3) the portfolio loss can be written as

\[
L_{t+1} = -(f(t + 1, Z_t + X_{t+1}) - f(t, Z_t)).
\] (4.4)

If \( f \) is differentiable, a first-order approximation \( L_{t+1}^\Delta \) of (4.4) can be considered,

\[
L_{t+1}^\Delta := -(f_t(t, Z_t) + \sum_{i=1}^d f_{Z,t,i}(t, Z_t) X_{t+1,i}),
\] (4.5)

where the subscripts to \( f \) denote partial derivatives. The first-order approximation gives a convenient computation of loss since it represents it as the linear combination of risk-factor changes. However, it is best to be used only when the risk-factor changes are likely to be small (i.e. if the risk is measured in small time horizon) and when the portfolio value is almost linear in the risk factors (i.e. if the function \( f \) has small second derivatives). The former can be shown when analyzing the counterpart of function \( f \) in the case when time is measured in longer time horizon, say, in years.
4.1 Stock Portfolio

Consider a fixed portfolio of \(d_s\) stocks and denote by \(\lambda_i^s\) the number of shares of stock \(i\) in the portfolio at time \(t\). Denote the stock \(i\) price process by \((S_{t,i})\). Since the log price process of financial assets have been modelled quite satisfyingly, the risk factor is assumed to be \(Z_{t,i} := \ln S_{t,i}\). The risk factor change then assumes the form of stock’s log return, i.e. \(X_{t+1,i} = \ln \left( \frac{S_{t+1,i}}{S_{t,i}} \right)\). Then,

\[
V_t = \sum_{i=1}^{d_s} \lambda_i^s \exp(Z_{t,i})
\]

and

\[
L_{t+1} = -\sum_{i=1}^{d_s} \lambda_i^s S_{t,i} \left( \exp(X_{t+1,i}) - 1 \right).
\]

Using first order approximation as in equation (4.5) on (4.7), the loss function can be linearized as

\[
L^\Delta_{t+1} = -\sum_{i=1}^{d_s} \lambda_i^s S_{t,i} X_{t+1,i} = -V_t \sum_{i=1}^{d_s} w_{t,i}^s \ln \left( \frac{S_{t+1,i}}{S_{t,i}} \right),
\]

where \(w_{t,i}^s := \lambda_i^s S_{t,i}/V_t\), the stock portfolio weight. Equation (4.8) gives the linearized risk mapping for stock portfolio. The error from such linearization is small as long as the stock log return is generally small.

4.2 Zero Coupon Bond Portfolio

Definition 4.1 (Zero Coupon Bond Portfolio) A zero coupon bond with maturity date \(T\) and face value \(K\), also known as a zero, is a contract which promises its holder a payment of amount \(K\) to be paid at date \(T\). Denote its price at time \(t\), where \(0 \leq t \leq T\) by \(p_z(t,T)\).

Definition 4.2 (Continuously Compounded Yield) A continuously compounded yield at time \(t\) with \(0 \leq t \leq T\), denoted by \(y_z(t,T)\) for a zero coupon bond with maturity \(T\) and face value \(K\) is defined as the factor \(y\) that satisfies

\[
p_z(t,T) = Ke^{-y(T-t)},
\]

or equivalently as

\[
y_z(t,T) := -\left( 1/(T-t) \right) \ln \left( \frac{p_z(t,T)}{K} \right).
\]

Since \(y_z(t,T)\) takes the form of log price of a financial asset, in this case, a zero coupon bond, it is natural to assume it as the risk factor in zero coupon bond portfolio.
A fixed coupon bond is a contract that provides payments at time \( C \), and face value \( K \), for \( i = 1, 2, \ldots, d \). Let the current time \( t \) be such that \( t < \min_{1 \leq i \leq d} T_i \). Denote by \( \lambda^c_i \) the number of bonds with maturity \( T_i \) in the portfolio. Let \( Z_{t,i} := K(T_i - t) \) be the risk factor. Hence, \( X_{t+1,i} = y_c((t+1)\Delta, T_i) - y_c(t\Delta, T_i) \), the increment of the risk factors. The value of the portfolio at time \( t \) is then

\[
V_t = \sum_{i=1}^{d_c} \lambda^c_i p_c(t\Delta, T_i) = \sum_{i=1}^{d_c} \lambda^c_i K \exp(-(T_i - t\Delta)y_c(t\Delta, T_i)). \tag{4.11}
\]

Using first order approximation in similar fashion as with the stock portfolio, the linearized loss can be obtained as

\[
L_{t+1}^\Delta = -\sum_{i=1}^{d_c} \lambda^c_i p_c(t\Delta, T_i) (y_c(t\Delta, T_i) \Delta - (T_i - t\Delta) X_{t+1,i})
\]

\[
= -V_t \sum_{i=1}^{d_c} w^c_{t,i} (y_c(t\Delta, T_i) \Delta - (T_i - t\Delta) X_{t+1,i}), \tag{4.12}
\]

where \( w^c_{t,i} := \lambda^c_i p_c(t, T_i)/V_t \) denotes the weight of bond portfolio.

### 4.3 Fixed Coupon Bond Portfolio

**Definition 4.3 (Fixed Coupon Bond)** A fixed coupon bond is a contract that guarantees its holder a sequence of deterministic payments \( C_1, C_2, \ldots, C_n \), called the coupons, at time \( T_1, T_2, \ldots, T_n \) which is arranged in ascending order. For simplicity, let \( C_n \) includes its face value. Denote its price at time \( t \) as \( p_c(t, T) \) where \( T = T_n \).

**Definition 4.4 (Yield to Maturity)** A yield to maturity at \( t < T_1 \), denoted by \( y_c(t, T) \) for a fixed coupon bond with payments \( C_1, C_2, \ldots, C_n \) at time \( T_1, T_2, \ldots, T_n \), where \( T = T_n \), the maturity date, is defined as the value \( y \) which satisfies

\[
p_c(t, T) = \sum_{i=1}^{n} C_i e^{-y(T_i-t)} . \tag{4.13}
\]

Hence, yield to maturity is defined implicitly and must be solved by a numerical root finding method.

Now, consider a fixed portfolio with \( d^c \) fixed coupon bonds, where the \( i \)-th bond provides payments \( C^{(i)}_1, C^{(i)}_2, \ldots, C^{(i)}_n \) at \( T^{(i)}_1, \ldots, T^{(i)}_n \) for \( i = 1, 2, \ldots, d^c \). Let the current time \( t \) be such that \( t < \min_{1 \leq i \leq d^c} T^{(i)}_1 \). Denote by \( \lambda^c_i \) the number of \( i \)-th bond in the portfolio. Let \( Z_{t,i} := y_c(t\Delta, T^{(i)}) \) be the risk factor. Hence, \( X_{t+1,i} = y_c((t+1)\Delta, T^{(i)}) - y_c(t\Delta, T^{(i)}) \), the increment of the risk factors. The value of the portfolio at time \( t \) is then

\[
V_t = \sum_{i=1}^{d_c} \lambda^c_i p_c(t\Delta, T^{(i)}) = \sum_{i=1}^{d_c} \lambda^c_i \left( \sum_{j=1}^{n_i} C^{(i)}_j e^{-(T^{(i)}_j - t\Delta) y_c(t\Delta, T^{(i)})} \right). \tag{4.14}
\]
Using first order approximation in similar fashion as with the stock portfolio, the linearized loss can be obtained as

\[ L_{t+1}^\Delta = - \sum_{i=1}^{d_e} \lambda_e^i p_e(t, T^{(i)}) \left( y_e \left(t, T^{(i)}\right) \Delta - D_i X_{t+1,i} \right) \]

\[ = - V_t \sum_{i=1}^{d_e} w_{t,i}^e \left( y_e \left(t, T^{(i)}\right) \Delta - D_i X_{t+1,i} \right), \quad (4.15) \]

where \( w_{t,i}^e := \lambda_e^i p_e(t, T^{(i)}) / V_t \) denotes the weight of bond portfolio and

\[ D_i := \sum_{j=1}^{n_i} C_j^{(i)} e^{-\left(T_j^{(i)} - t\right)} g_e(t, T^{(i)}) \left(T_j^{(i)} - t\Delta\right) \]

\[ p_e(t, T^{(i)}) \quad (4.16) \]

denotes the \( i \)-th bond’s duration, a measure of the sensitivity of the bond price with respect to yield changes, since \( D_i = \frac{\partial p_e(t, T^{(i)})}{\partial y_e(t, T^{(i)})} / p_e(t, T^{(i)})\).

### 4.4 Currency Portfolio

The risk mapping for stock and currency portfolios are similar in nature. To see this, consider a currency portfolio with \( d_e \) number of foreign currencies and denote by \( \lambda_e^i \) the value of currency \( i \) in the corresponding currency denomination, in the portfolio at time \( t \). Denote currency \( i \) exchange rate process by \((e_{t,i})\) in foreign/domestic value. The risk factor is assumed to be \( Z_{t,i} := \ln e_{t,i} \) with reasons similar to the ones in stock portfolio case. So, the portfolio value at time \( t \) is

\[ V_t = \sum_{i=1}^{d_e} \lambda_e^i \exp(Z_{t,i}) \quad (4.17) \]

Hence, the currency portfolio is similar to the stock portfolio. With some adjustments, equations (4.17)-(4.8) can be used to derive the risk mapping for linearized currency portfolio loss, which is

\[ L_{t+1}^\Delta = - V_t \sum_{i=1}^{d_e} w_{t,i}^e \ln \left( \frac{e_{t+1,i}}{e_{t,i}} \right), \quad (4.18) \]

where \( w_{t,i}^e := \lambda_e^i e_{t,i} / V_t \), the currency portfolio weight.

It is also appropriate to consider a portfolio consisting assets valued in foreign currency since not all assets used in this research are denominated in IDR. For this reason, consider a fixed portfolio consisting of \( d \) noncurrency assets, each valued in foreign currency, i.e., let asset \( i \) be valued in currency \( i \). Let \( p(t, Z_{t,i}) \) be the price of asset \( i \) which depends on risk factor \( Z_{t,i} \), \( \lambda_i \) be the amount of asset \( i \) and \( e_{t,i} \) be the exchange rate of currency \( i \). Then, this portfolio value at time \( t \) in IDR is

\[ V_t = \sum_{i=1}^{d} \lambda_i p(t, Z_{t,i}) e_{t,i} = \sum_{i=1}^{d} \lambda_i p(t, Z_{t,i}) \exp(\ln e_{t,i}). \quad (4.19) \]
Hence, each asset of the portfolio contains two risk factors: its intrinsic risk factor and the log exchange rate. It follows that the linearized portfolio loss can be obtained by equation (4.5) as

$$L_{t+1}^\Delta = - \sum_{i=1}^{d_s} \lambda_i \left( p_{Z_{t,i}}(t, Z_{t,i}) e_{t,i} X_{t+1,i}^s + p(t, Z_{t,i}) e_{t,i} \ln \left( \frac{e_{t+1,i}}{e_{t,i}} \right) \right)$$

(4.20)

$$= -V_t \sum_{i=1}^{d_s} w_{t,i} \left( \frac{p_{Z_{t,i}}(t, Z_{t,i})}{p(t, Z_{t,i})} X_{t+1,i}^s + X_{t+1,i}^e \right),$$

(4.21)

where $w_{t,i} := \lambda_i p(t, Z_{t,i}) e_{t,i} / V_t$ is the asset’s weight, while $X_{t+1,i}^s$ and $X_{t+1,i}^e$ are the risk factor increments for the asset and the foreign currency, respectively. Note that the weight for assets valued in foreign currency differs slightly from the weight of other domestic assets. This weight can be regarded as the usual weight multiplied by the exchange rate of the currency of which it is denominated. This interpretation is consistent with the conversion process of its foreign value to its domestic value.

5 Profit and Loss Distribution

Since the risk mapping for each of the portfolio’s assets have been obtained, it is time to determine the distribution of the portfolio loss distribution by employing the linearity property of Generalized Hyperbolic distribution presented by theorem 2.4 in the previous chapter. Due to this linearity property, to make future calculation tractable, only the linearized portfolio loss function is considered.

Using the notation from the preceding sections, consider a fixed portfolio containing $d_s$ stocks, $d_e$ currencies, $d_z$ zeros and $d_c$ fixed coupon bonds. First, assume that all of the assets are denominated in the domestic currency. Using the formula of linearized loss from preceding section, the linearized loss of this portfolio can be obtained as

$$L_{t+1}^\Delta = -V_t \left( \sum_{i=1}^{d_s} w_{t,i}^s X_{t+1,i}^s + \sum_{i=1}^{d_z} w_{t,i}^z (y_z (t \Delta, T_i) \Delta - (T_i - t \Delta) X_{t+1,i}^z) + \sum_{i=1}^{d_c} w_{t,i}^c (y_c (t \Delta, T^{(i)}) \Delta - D_i X_{t+1,i}^c) + \sum_{i=1}^{d_e} w_{t,i}^e X_{t+1,i}^e \right),$$

(5.1)

where $X_{t+1,i}^s := \ln \left( \frac{S_{t+1,i}}{S_{t,i}} \right)$, $X_{t+1,i}^e := \ln \left( \frac{e_{t+1,i}}{e_{t,i}} \right)$, $X_{t+1,i}^z := (y_z(t + 1, T_i) - y_z(t, T_i))$ and $X_{t+1,i}^c := (y_c((t + 1) \Delta, T^{(i)}) - y_c(t \Delta, T^{(i)}))$ are the risk factors for stocks, currencies, zeros and fixed coupon bonds, respectively. Next, as has been assumed, let
\( \mathbf{X} := (X^s_{t+1,1}, \ldots, X^s_{t+1,d_s}, X^s_{t+1,1}, \ldots, X^s_{t+1,1}, \ldots, X^e_{t+1,1}, \ldots, X^e_{t+1,1}) \sim \text{GH}(\lambda, \chi, \psi, \mu, \Sigma, \gamma). \) Note that \( \mathbf{X} \in \mathbb{R}^d, \) where \( d := d_s + d_e + d_z + d_c. \) By (5.1), it is clear that

\[
L_{t+1}^\Delta = -V_t \mathbf{w}' (\mathbf{b} + \mathbf{B} \mathbf{X})
\]

where \( \mathbf{w} \) is the weight vector corresponding to each portfolio assets (arranged in similar fashion as \( \mathbf{X} \)), and \( \mathbf{b} \in \mathbb{R}^d \) and \( \mathbf{B} \in \mathbb{R}^{d \times d} \) are constant vector and diagonal matrix, respectively, such that

\[
b_i = \begin{cases} 
y_z(t \Delta, T_i) \Delta, & i = d_s + 1, \ldots, d_s + d_z 
y_c(t \Delta, T_i) \Delta, & i = d_s + d_z + 1, \ldots, d_s + d_z + d_c 
0, & \text{otherwise}
\end{cases}
\]

and

\[
B_{ii} = \begin{cases} 
-(T_i - t \Delta), & i = d_s + 1, \ldots, d_s + d_z 
-D_i, & i = d_s + d_z + 1, \ldots, d_s + d_z + d_c 
0, & i = d - d_e + 1, \ldots, d \text{ and if currency } i \text{ is not held}
1, & \text{otherwise}
\end{cases}
\]

Now, consider the same case, with the exception that some of the noncurrency assets are denominated in foreign currency. By the arguments within section 4.4, equation (5.2) can still be used to represent the linearized loss of this portfolio with the exception that each components of the weight vector that corresponds with assets valued in foreign currency must be multiplied by the foreign exchange rate at time \( t \), and that the diagonal matrix \( \mathbf{B} \) has to be modified into a sparse matrix with same diagonal entries as before, but with nondiagonal entries

\[
B_{ij} = \begin{cases} 
1, & \text{if asset } i \text{ is valued in currency } j 
0, & \text{otherwise}
\end{cases}
\]

where \( i = 1, \ldots, d - d_e \) and \( j = d - d_e + 1, \ldots, d. \)

By evoking theorem 2.4, the preceding representation can be simplified into

\[
L_{t+1}^\Delta = -V_t \mathbf{w}' \tilde{\mathbf{X}},
\]

where \( \tilde{\mathbf{X}} \sim \text{GH}(\lambda, \chi, \psi, \mathbf{b} + \mathbf{B} \mathbf{\mu}, \mathbf{B} \Sigma \mathbf{B}, \mathbf{B} \gamma). \) Also note that by the same theorem, \( L_{t+1}^\Delta \sim \text{GH}(\lambda, \chi, \psi, -V_t \mathbf{w}' (\mathbf{b} + \mathbf{B} \mathbf{\mu}), (V_t)^2 \mathbf{w}' \mathbf{B} \Sigma \mathbf{B}' \mathbf{w}, -V_t \mathbf{w}' \mathbf{B} \gamma). \) This is one of the advantages of modeling risk factor increments with Generalized Hyperbolic distribution, as the linearized portfolio is also Generalized Hyperbolic distributed. Note also that for optimization purposes, equation (5.6) can be used to represent portfolio losses. This concludes that a portfolio loss function can be represented by its weight vector and the risk factor increment vector through equation (5.6).
6 Portfolio Optimization

As is argued by the linearity property of Generalized Hyperbolic distribution, portfolio loss function is approximated by its linearized counterpart, and so is the portfolio profit function. So, from here on, the notation $L$ will be used as the linearized portfolio function, replacing the role of $L^\Delta$. To set up the portfolio optimization problem, let

$$X \sim GH(\lambda, \chi, \psi, \mu, \Sigma, \gamma),$$

(6.1)

a $d$-dimensional Generalized Hyperbolic random variable. This $X$ can be regarded as the variable $\tilde{X}$ in equation (5.6) back in section 5 by adjusting the parameters of $X$ since this equation can be used to represent portfolio losses. Next, let $M := \{ L : L = l + \lambda'X, \lambda \in \mathbb{R}^d, l \in \mathbb{R} \}$ be the set of portfolio losses and $\phi : M \to \mathbb{R}$ be the coherent risk measure. Note that by the discussion from the previous section, the distribution of $L$ itself is a function of portfolio weights, hence, so are its expectation and its risk measure.

The followings are the optimization versions which will be discussed.

**Definition 6.1 (Markowitz Optimization Problem)**

$$\min \varsigma(w)$$

subject to $R(w) = \lambda$ and $w'1 = 1$, \hspace{1cm} (P1)

where $\lambda$ is the target expected return and $w$ is the portfolio weights.

$$\varsigma(w) := \phi(L(w))$$

(6.2)

$$R(w) := \mathbb{E}[-L(w)]$$

(6.3)

By representation (5.6), the portfolio loss $L$ has a one-to-one relationship with $w$, the portfolio weights. Hence, it follows, that, if $\phi(L)$ is convex on $M$, then $\varsigma(w)$ is convex on $\mathbb{R}^d$.

**Definition 6.2 (RORC Optimization Problem)**

$$\max \frac{R(w)}{\varsigma(w)}$$

subject to $w'1 = 1$. \hspace{1cm} (P2)

The objective function in this problem can be nonconvex, but, it can easily be seen that the solution of this problem lies on the efficient frontier of the problem (P1). In cases where the maximum of the RORC cannot be found analytically, this fact can be used to find the maximum of the ratio by selecting the point on the efficient frontier that yields the highest value of the ratio. This way, the nonconvex problem is reduced to a fixed number of convex optimization problems which are efficiently solvable.
6.1 Optimization in Symmetric Case

In the symmetric case, \( \gamma \) in equation (6.1) has to be set to 0, and so,

\[
E[w'X] = w'E[X] = w'\mu. \tag{6.4}
\]

Hence, the set of feasible weights for the Markowitz optimization problem can be defined as \( S := \{ w : w'1 = 1, w' \mu = \mu_p \} \). The problem will first be discussed.

Assume \( \beta > 0.5 \). Since \( w'X \sim \text{GH}(\lambda, \chi, \psi, w' \mu, w' \Sigma w, 0) \), then \( \tilde{X} := (\tilde{w}'X + w'\mu)/\sqrt{w'\Sigma w} \sim \text{GH}(\lambda, \chi, \psi, 0, 1, 0) \) doesn’t depend on \( w \). Hence, note that

\[
\text{VaR}_\beta(L) = -\mu_p + q_1 \sqrt{w' \Sigma w} \tag{6.5}
\]

\[
\text{ES}_\beta(L) = -\mu_p + q_2 \sqrt{w' \Sigma w} \tag{6.6}
\]

where \( q_1 := \text{VaR}_\beta(\tilde{X}) \) and \( q_2 := \text{ES}_\beta(\tilde{X}) \) don’t depend on \( w \). Since \( E[\tilde{X}] = 0, \beta > 0.5 \) and \( \tilde{X} \) is symmetric, it is clear that \( q_1 \) and \( q_2 \) are nonnegative constants. This implies that

\[
\text{argmin}_w \text{VaR}(L) = \text{argmin}_w \text{ES}(L) = \text{argmin}_w w' \Sigma w. \tag{6.7}
\]

**Proposition 6.3 (Equality of Markowitz-Optimal Weights)** In the framework of symmetric Generalized Hyperbolic, the optimal portfolio composition obtained from Markowitz optimization problem using Value-at-Risk or Expected Shortfall at confidence level \( \beta \geq 0.5 \) is equal to that of using volatility as risk measure.

See McNeil et. al. [27] for proof.

So, optimizing using VaR and ES is equivalent to using volatility. Hence, by standard Lagrangian method, the optimal weights can be obtained as

\[
w = K_1 + K_2 \mu_p, \tag{6.8}
\]

where \( K_1 \) and \( K_2 \) are constant vectors

\[
K_1 = \frac{1}{D}[A_1 \Sigma^{-1}1 - A_2 \Sigma^{-1}\mu] \tag{6.9}
\]

\[
K_2 = \frac{1}{D}[A_3 \Sigma^{-1}\mu - A_2 \Sigma^{-1}1], \tag{6.10}
\]

with \( A_1 := \mu' \Sigma^{-1} \mu, A_2 := \mu' \Sigma^{-1}1, A_3 := 1' \Sigma^{-1}1 \) and \( D := A_1 A_3 - A_2^2 \).

Equations (6.5) and (6.6) can be further used to find \( \mu_p \) which yields the minimum value of VaR and ES. Let \( q \) be either \( q_1 \) or \( q_2 \). Since VaR and ES can be regarded as a function of \( \mu_p \), define a function \( f \) with

\[
f(\mu_p) := -\mu_p + q \sqrt{w' \Sigma w}. \tag{6.11}
\]

Substituting \( w \) from (6.8) to (6.11) yields

\[
f(\mu_p) = -\mu_p + q \sqrt{K_2' \Sigma K_2 \mu_p^2 + 2K_1' \Sigma K_2 \mu_p + K_1' \Sigma K_1}. \tag{6.12}
\]
By differentiation, it can be obtained that

\begin{align*}
  f'(\mu_p) &= -1 + q \frac{K_2'\Sigma K_2\mu_p + K_1'\Sigma K_1}{\sqrt{K_2'\Sigma K_2\mu_p^2 + 2K_1'\Sigma K_1\mu_p + K_1'\Sigma K_1}}, \quad (6.13) \\
  f''(\mu_p) &= q \frac{(K_2'\Sigma K_2)(K_1'\Sigma K_1) - (K_1'\Sigma K_2)^2}{(K_2'\Sigma K_2\mu_p^2 + 2K_1'\Sigma K_1\mu_p + K_1'\Sigma K_1)^{3/2}}. \quad (6.14)
\end{align*}

By Schwarz inequality, \((K_2'\Sigma K_2)(K_1'\Sigma K_1) - (K_1'\Sigma K_2)^2 \geq 0\). It follows then that \(f''\) is nonnegative and so \(f\) is convex with respect to \(\mu_p\). Observe that if \(f'' = 0\), the minimum of \(f\) doesn’t exist. Assume it is positive. To see another condition for the existence of the minimum, observe that

\[
\lim_{\mu_p \to \pm\infty} f'(\mu_p) = -1 \pm q\sqrt{K_2'\Sigma K_2}. \quad (6.15)
\]

By the monotonicity of \(f'\), it follows that the minimum of \(f\) exists if

\[
\lim_{\mu_p \to -\infty} f'(\mu_p) = -1 + q\sqrt{K_2'\Sigma K_2} > 0. \quad (6.16)
\]

Assume the minimum exists. Define \(J_1 := K_2'\Sigma K_2\), \(J_2 := K_1'\Sigma K_2\) and \(J_3 := K_1'\Sigma K_1\) to simplify matters. Since \(f\) is convex, the minimum can be found by solving the root of \(f'\). Observe that the root-finding problem of \(f'\) can be transformed into the problem of solving

\[
J_1(1 - q^2J_1)\mu_p^2 + 2J_2(1 - q^2J_1)\mu_p + (J_3 - q^2J_2^2) = 0. \quad (6.17)
\]

Equation (6.17) yields

\[
\mu_1 = \frac{-J_2(1 - q^2J_1) - \sqrt{D}}{2J_1(1 - q^2J_1)} \quad \text{and} \quad \mu_2 = \frac{-J_2(1 - q^2J_1) + \sqrt{D}}{2J_1(1 - q^2J_1)}, \quad (6.18)
\]

where \(D := 4(-1 + q^2J_1)(J_1J_3 - J_2^2)\). If \(D \neq 0\), then, only one between \(\mu_1\) and \(\mu_2\) can be the solution. To see which one, observe that

\[
f'(\mu_p) = \begin{cases} 
-1 - q \sqrt{\frac{(J_1\mu_p + J_2)^2}{J_1\mu_p^2 + 2J_2\mu_p + J_3}} & \text{if } \mu_p \leq -\frac{J_2}{J_1} \\
-1 + q \sqrt{\frac{(J_1\mu_p + J_2)^2}{J_1\mu_p^2 + 2J_2\mu_p + J_3}} & \text{if } \mu_p > -\frac{J_2}{J_1}.
\end{cases} \quad (6.19)
\]

Hence, it is necessary for \(\mu_p\) to be greater than \(-\frac{J_2}{J_1}\) to have \(f'(\mu_p) = 0\). Since it is clear that \(\mu_1 \leq -\frac{J_2}{J_1}\), the root is \(\mu_2\). This result can be summed up as

**Proposition 6.4 (Global Minimum of VaR and ES)** In symmetric GH framework of portfolio optimization, the minimum of VaR and ES at confidence level \(\beta \geq 0.5\) of portfolio loss exists iff

\[
(K_2'\Sigma K_2)(K_1'\Sigma K_1) - (K_1'\Sigma K_2)^2 > 0 \quad (6.20)
\]
and
\[ -1 + q\sqrt{K_2'\Sigma K_2} > 0. \] (6.21)

If it exists, the minimum is achieved at
\[ \mu_p = -\frac{2(K_1'\Sigma K_2)(1 - q^2(K_2'\Sigma K_2)) + \sqrt{D}}{2(K_2'\Sigma K_2)(1 - q^2(K_2'\Sigma K_2))}, \] (6.22)
where \( D = 4(-1 + q^2(K_2'\Sigma K_2))((K_2'\Sigma K_2)(K_1'\Sigma K_1) - (K_1'\Sigma K_2)^2) \). Note that \( q = q_1 \) in the context of VaR and \( q_2 \) otherwise.

Next, to find the optimal weights of the RORC optimization problem, the following proposition can be used.

**Proposition 6.5 (RORC in Symetric Framework)** Let \( \phi \) be Value-at-Risk or Expected Shortfall at confidence level \( \beta \geq 0.5 \) at the framework of symmetric Generalized Hyperbolic. Let \( \text{RORC}_\sigma \) and \( \text{RORC}_\phi \) be the RORC for return-volatility and either return-VaR or return-ES optimization problems, respectively, with same set of weight constraints. Then, if \( \mu^* = \arg\max \text{RORC}_\sigma(\mu) \) and the global minimum value of \( \phi \) is positive, \( \mu^* = \arg\max \text{RORC}_\phi(\mu) \).

**Proof** Let \( \phi \) be either the minimum of \( \text{VaR}_\beta \) or \( \text{ES}_\beta \) given an expected return \( \mu_p \). Using equations (6.5) and (6.6), then \( \phi \) can be expressed as a function of \( \mu_p \) only, i.e.,
\[ \phi(\mu_p) = -\mu_p + q\sigma(\mu_p), \] (6.23)
where \( \sigma(\mu_p) \) is the minimum volatility corresponding to \( \mu_p \). Let \( \mu^* = \arg\max \text{RORC}_\sigma(\mu) \), then for every \( \tilde{\mu} \in \mathbb{R} \),
\[ \frac{\tilde{\mu}}{\sigma(\tilde{\mu})} \leq \frac{\mu^*}{\sigma(\mu^*)} \] (6.24)
\[ \Leftrightarrow \tilde{\mu} \sigma(\mu^*) \leq \mu^* \sigma(\tilde{\mu}). \] (6.25)

Next, note that
\[ \frac{\tilde{\mu}}{\phi(\tilde{\mu})} - \frac{\mu^*}{\phi(\mu^*)} = \frac{\mu^* \phi(\mu^*) - \mu^* \phi(\tilde{\mu})}{\phi(\mu^*)} \] (6.26)
\[ = \frac{q(\mu^* \sigma(\mu^*) - \mu^* \sigma(\tilde{\mu}))}{\phi(\mu^*)} \leq 0, \] (6.27)
where the last equality is due to equation (6.23), while the inequality is due to equation (6.25) and the positivity of \( q \) and \( \phi \). The proof is complete.

Hence, optimal returns which yield the minimum RORC for optimizations using either VaR or ES are equal to those using volatility. It is therefore easier to compute the minimum RORC using volatility as risk measure. In the following discussions, the existence conditions of minimum RORC will be derived using this property.
Note that given targeted expected return, \( \mu_p \), the portfolio loss volatility can be expressed as a function of \( \mu_p \)

\[
\sigma(L) = \sqrt{K w' \Sigma w} = \sqrt{K \left( K_1' \Sigma K_1 - \frac{(K_1' \Sigma K_2)^2}{K_2' \Sigma K_2} \right)}, \tag{6.28}
\]

where \( K := \mathbb{E}[W] \), the expectation of the underlying distribution of GH and depends only on the inner parameters of GH: \( \chi \), \( \lambda \) and \( \psi \), and therefore is a constant. Hence, the RORC using volatility as risk measure can be expressed as

\[
\text{RORC}(\mu_p) = \frac{\mu_p}{\sigma(\mu_p)}. \tag{6.29}
\]

The first derivative of RORC will first be analyzed. It can be obtained that

\[
\text{RORC}'(\mu_p) = \frac{\sigma - \mu_p (\sigma^2)'}{\sigma^2} = \frac{2\sigma^2 - \mu_p (\sigma^2)'}{2\sigma^3} = \frac{K, 2K_1' \Sigma K_2 \mu_p + K_1' \Sigma K_1}{\sigma^3}. \tag{6.30}
\]

Set \( \mu_p^* := -\frac{K_1' \Sigma K_1}{2K_1' \Sigma K_2} \). From the third equation of (6.30), there can be three cases. First, if \( K_1' \Sigma K_2 < 0 \), then the RORC function will be strictly increasing over \( (-\infty, \mu_p^*) \) and strictly decreasing over \( (\mu_p^*, \infty) \). In this case, the RORC will then be maximized at \( \mu_p = \mu_p^* \). Second, if \( K_1' \Sigma K_2 \geq 0 \), then the RORC will increase to its asymptote

\[
\lim_{\mu_p \to \infty} \frac{\mu_p}{\sigma} = \frac{1}{\sqrt{K, K_1' \Sigma K_2}}. \tag{6.31}
\]

as \( \mu_p \) tends to \( \infty \), where its asymptote is the maximum value. In the case where \( K_1' \Sigma K_2 > 0 \), this asymptote is its maximum value since

\[
\lim_{\mu_p \to -\infty} \frac{\mu_p}{\sigma} = -\frac{1}{\sqrt{K, K_1' \Sigma K_2}}. \tag{6.32}
\]

The preceding results therein lead to the following proposition.

**Proposition 6.6 (Conditions of the Existence of Minimum RORC)**

In the symmetric GH framework, maximum RORC using either VaR or ES at confidence level \( \beta \geq 0.5 \) exists iff \( K_1' \Sigma K_2 < 0 \). If it exists, the maximum is achieved at the expected return on the level

\[
\mu_p^* := -\frac{K_1' \Sigma K_1}{2K_1' \Sigma K_2}. \tag{6.33}
\]
6.2 Optimization in Asymmetric Case

Since VaR is not necessarily a coherent risk measure in the asymmetric GH framework, only optimizations using ES will be discussed. In this case, $\text{ES}_\beta(L) = \text{ES}_\beta(-w'X)$ cannot be expressed in basic functions such as in equation (6.6). It is not a linear transformation of volatility, since the factor $q_2$ in the equation contains the term $w'\gamma$, i.e.

$$q_2 = \text{ES}_\beta(Y), \quad Y \sim GH \left( \lambda, \chi, \psi, 0, 1, \frac{w'\gamma}{w'\Sigma w} \right). \tag{6.34}$$

But, for numerical computations, it can be expressed in integral form as follows

$$\text{ES}_\beta(-w'X) = \mathbb{E}[ -w'X \mid -w'X > \text{VaR}_\beta(-w'X) ] \tag{6.35}$$

$$= -\mathbb{E}[w'X \mid w'X \leq \text{VaR}_{1-\beta}(w'X)] \tag{6.36}$$

$$= - \left( \frac{1}{1 - \beta} \int_{-\infty}^{F_Y^{-1}(1-\beta)} y f_Y(y) dy \right), \tag{6.37}$$

where $Y = w'X \sim GH (\lambda, \chi, \psi, w'\mu, w'\Sigma w, w'\gamma)$, and $F_Y$ and $f_Y$ denote the cdf and pdf of $Y$, respectively. The computation of $F_Y^{-1}(1 - \beta)$ can be done by numerical root finding method, while the integral in (6.37) is done by numerical integration.

7 Numerical Results

For simulation purposes, the following assets are used: 4 stocks, 3 foreign currencies, 2 Indonesian government-issued zero coupon bonds and 1 Indonesian government-issued international fixed coupon bond. The stocks are chosen international blue chip companies which are

1. ASII (Astra International), denominated in IDR
2. BACH (Bank of China), denominated in CNY
3. INTC (Intel Corporation), denominated in USD
4. RDS (Royal Dutch Shell), denominated in EUR

The foreign currencies are chosen to represent the three most developed regions: U.S.A, Europe and China.

1. CNY (Chinese Yuan)
2. EUR (Euro)
3. USD (US dollar)

The zero coupon bonds are government issued of series
1. ZC3 with specifications
   - ISIN: IDB000000309
   - BB Number: EH0300275
   - Amount Issued: IDR 1,500,000.00
   - Amount Outstanding: IDR 1,249,000.00
   - Par Amount: IDR 1,000,000
   - Maturity: 11/20/2012
   - Denomination: IDR

2. ZC5 with specifications
   - ISIN: IDB000000507
   - BB Number: EH1785003
   - Amount Issued: IDR 3,150,000.00
   - Amount Outstanding: IDR 1,263,000.00
   - Par Amount: IDR 1,000,000
   - Maturity: 2/20/2013
   - Denomination: IDR

while the fixed coupon bond is of series

1. INDO38 with specifications
   - Par Amount: USD 1,000,000
   - Coupon: 7.75% p.a.
   - Coupon payment date: 17 January & 17 July
   - Maturity: 11/17/2038
   - Denomination: USD

Additionally, the following terms are used on the portfolio weights: unconstrained and constrained. An asset weight is called unconstrained if asset shortings are permitted, and constrained if no shorts are allowed.

7.1 Calibration Results

As has been explained before, the source of uncertainty of a portfolio comes from its risk factors. The increments of those risk factors are assumed to be Generalized Hyperbolic distributed. For stocks and currencies, the increments are their log returns, respectively. While for zero coupon bonds and fixed coupon bonds, they are the continuously compounded yields and yield-to-maturity, respectively. For
bonds, the calculations of time to coupon maturities are done under actual/365 rule which assumes that the number of days in every year is simplified to be 365, while calculations of period less than 1 year is done on actual account.

The data obtained for all of the assets are their trading prices, so, the stocks and currencies log returns can be calculated directly. For bonds, the prices are quoted as the percentage of their par values. Yields of zero coupon bonds can be calculated using equation (4.10). Their trading prices can be used as the value of $p_z(t,T)$. Yield to maturity of a fixed coupon bond can be calculated by solving equation (4.13) for $y$ by root finding methods. Since both sides of the equation can be divided by the bond’s par value, the trading price can be inputted directly to variable $p_c(t,T)$.

Note that for INDO38, there are 57 coupon payments with the last coincide with the redeem payment of the bond, i.e., $C_i = 3.875$ for $i = 1, \ldots, 56$ and $C_{57} = 103.875$.

The calibration results are presented for the percentage risk factors increments of the assets used. They will be presented for both the symmetric and asymmetric models of univariate and multivariate Generalized Hyperbolic distributions. When discussing about the multivariate Generalized Hyperbolic calibration of the data, the following order is used for the order of the elements of GH random vector:

1. ZC3 yield increment
2. ZC5 yield increment
3. INDO38 yield increment
4. ASII log return
5. BACH log return
6. INTC log return
7. RDS log return
8. CNY log return
9. EUR log return
10. USD log return

We denote by $X_s$ and $X$ the symmetric and asymmetric GH random vectors for the daily returns, respectively, where $X_s \sim GH(\lambda_s, \chi_s, \psi_s, \mu_s, \Sigma_s, \gamma_s)$ and $X \sim GH(\lambda, \chi, \psi, \mu, \Sigma, \gamma)$.

The calibration is performed by employing the EM algorithm detailed in the Appendix. Again, note that the calibration can only be done with fixed $\lambda$. For nonnegative $\lambda$, $\chi$-algorithm is used, while for negative $\lambda$, $\psi$-algorithm is used. For each $\lambda$, the calibration is terminated when the likelihood increments between the current and previous iteration is less than a specified value. More specifically, denote the observed Generalized Hyperbolic random vector by $X = (X_1, \ldots, X_n)$ and
Generalized Hyperbolic estimated parameters at iteration \( k \) by 
\[
\hat{\theta}^k = (\lambda^{(k)}, \chi^{(k)}, \psi^{(k)}, 
\mu^{(k)}, \Sigma^{(k)}, \gamma^{(k)})
\]
Given positive \( \epsilon \), the calibration process then terminates at iteration 
\( k + 1 \) if
\[
L(\hat{\theta}^{k+1}; X) - L(\hat{\theta}^k; X) < \epsilon.
\] (7.1)

The EM algorithm is then combined with a one dimensional optimization method
to find the value of \( \lambda \) that yields the highest likelihood value. MATLAB 'fminbnd'
function is used for purpose, where the search region for \( \lambda \) is from \(-10 \) to \(10 \).

Tables 1-2 show the calibrated parameters of the symmetric and asymmetric
univariate Generalized Hyperbolic distribution for all of the assets used.

<table>
<thead>
<tr>
<th></th>
<th>( \mu )</th>
<th>( \sigma^2 )</th>
<th>( \chi )</th>
<th>( \psi )</th>
<th>( \lambda )</th>
</tr>
</thead>
<tbody>
<tr>
<td>ZC3</td>
<td>0.0115</td>
<td>0.1130</td>
<td>0.0068</td>
<td>0.1178</td>
<td>-0.2031</td>
</tr>
<tr>
<td>ZC5</td>
<td>0.0074</td>
<td>0.1617</td>
<td>0.0031</td>
<td>0.1571</td>
<td>-0.1139</td>
</tr>
<tr>
<td>INDO38</td>
<td>0.0005</td>
<td>0.1046</td>
<td>0.0009</td>
<td>0.0735</td>
<td>-0.1823</td>
</tr>
<tr>
<td>ASII</td>
<td>0.0240</td>
<td>11.5590</td>
<td>0.2071</td>
<td>0.6014</td>
<td>-0.1384</td>
</tr>
<tr>
<td>BACH</td>
<td>-0.0498</td>
<td>3.6587</td>
<td>1.0949</td>
<td>0.3008</td>
<td>-1.1128</td>
</tr>
<tr>
<td>INTC</td>
<td>-0.0427</td>
<td>5.1474</td>
<td>0.1715</td>
<td>2.0345</td>
<td>0.8028</td>
</tr>
<tr>
<td>RDS</td>
<td>0.0759</td>
<td>4.0809</td>
<td>0.8588</td>
<td>0.0680</td>
<td>-1.2296</td>
</tr>
<tr>
<td>CNY</td>
<td>0.0176</td>
<td>0.5832</td>
<td>0.1424</td>
<td>0.0949</td>
<td>-0.5731</td>
</tr>
<tr>
<td>EUR</td>
<td>-0.0330</td>
<td>0.8128</td>
<td>1.8219</td>
<td>0.0000</td>
<td>-1.9109</td>
</tr>
<tr>
<td>USD</td>
<td>0.0074</td>
<td>0.6621</td>
<td>0.1063</td>
<td>0.1019</td>
<td>-0.5073</td>
</tr>
</tbody>
</table>

Table 1: Parameters of calibrated univariate symmetric Generalized Hyperbolic
distribution for some portfolio assets.

<table>
<thead>
<tr>
<th></th>
<th>( \mu )</th>
<th>( \gamma )</th>
<th>( \sigma^2 )</th>
<th>( \chi )</th>
<th>( \psi )</th>
<th>( \lambda )</th>
</tr>
</thead>
<tbody>
<tr>
<td>ZC3</td>
<td>0.0120</td>
<td>-0.0173</td>
<td>0.1227</td>
<td>0.0057</td>
<td>0.1193</td>
<td>-0.1909</td>
</tr>
<tr>
<td>ZC5</td>
<td>0.0075</td>
<td>-0.0041</td>
<td>0.1615</td>
<td>0.0031</td>
<td>0.1572</td>
<td>-0.1137</td>
</tr>
<tr>
<td>INDO38</td>
<td>0.0005</td>
<td>0.0077</td>
<td>0.1037</td>
<td>0.0009</td>
<td>0.0740</td>
<td>-0.1818</td>
</tr>
<tr>
<td>ASII</td>
<td>-0.0173</td>
<td>0.0003</td>
<td>11.5450</td>
<td>0.2087</td>
<td>0.6000</td>
<td>-0.1413</td>
</tr>
<tr>
<td>BACH</td>
<td>-0.0543</td>
<td>0.0065</td>
<td>3.6590</td>
<td>1.0956</td>
<td>0.3008</td>
<td>-1.1133</td>
</tr>
<tr>
<td>INTC</td>
<td>0.0464</td>
<td>-0.1231</td>
<td>5.1327</td>
<td>0.1136</td>
<td>2.1686</td>
<td>0.9358</td>
</tr>
<tr>
<td>RDS</td>
<td>0.1388</td>
<td>-0.1176</td>
<td>4.0292</td>
<td>0.8633</td>
<td>0.0773</td>
<td>-1.2176</td>
</tr>
<tr>
<td>CNY</td>
<td>0.0125</td>
<td>0.0240</td>
<td>0.5761</td>
<td>0.1419</td>
<td>0.1000</td>
<td>-0.5638</td>
</tr>
<tr>
<td>EUR</td>
<td>-0.0364</td>
<td>0.0052</td>
<td>0.8130</td>
<td>1.8223</td>
<td>0.0000</td>
<td>-1.9111</td>
</tr>
<tr>
<td>USD</td>
<td>0.0096</td>
<td>-0.0119</td>
<td>0.6618</td>
<td>0.1074</td>
<td>0.1008</td>
<td>-0.5107</td>
</tr>
</tbody>
</table>

Table 2: Parameters of calibrated univariate asymmetric Generalized Hyperbolic
distribution for some portfolio assets.

The followings are the parameters of the calibrated symmetric multivariate GH
distribution:
\[ \lambda = -1.1245 \]
\[ \chi = 0.3631 \]
\[ \psi = 0.0003 \]

\[ \mu = \begin{pmatrix}
0.0040 \\
0.0059 \\
-0.0024 \\
0.1523 \\
0.0002 \\
-0.0306 \\
0.0996 \\
-0.0069 \\
-0.0143 \\
-0.0150
\end{pmatrix} \]

\[ \Sigma = \begin{pmatrix}
0.1490 & 0.0544 & 0.0004 & -0.1129 & 0.0373 & -0.0261 & -0.0753 & 0.0155 & -0.0285 & 0.0138 \\
0.1790 & 0.0048 & -0.1648 & 0.0794 & -0.0162 & 0.0423 & 0.0259 & -0.0174 & 0.0279 \\
0.0455 & -0.2220 & -0.0495 & -0.0422 & -0.0732 & 0.0325 & -0.0118 & 0.0036 \\
30.8840 & 1.9736 & 1.8663 & 2.5462 & -1.2549 & 0.0044 & -1.1895 \\
12.4030 & 0.0826 & 0.9799 & -0.4477 & 0.1270 & -0.4281 \\
16.7750 & 4.5935 & -0.1699 & 0.3917 & -0.1978 \\
9.4950 & -0.3300 & 0.3854 & -0.0251 \\
2.2710 & 0.3456 & 1.0010 \\
\end{pmatrix} \]

and also some other values related to \( X_s \)

\[ \mathbb{E}[X_s] = \begin{pmatrix}
0.0040 \\
0.0059 \\
-0.0024 \\
0.1523 \\
0.0002 \\
-0.0306 \\
0.0996 \\
-0.0069 \\
-0.0143 \\
-0.0150
\end{pmatrix} \]

\[ \text{Cov}(X_s) = \begin{pmatrix}
0.1490 & 0.0544 & 0.0004 & -0.1129 & 0.0373 & -0.0261 & -0.0753 & 0.0155 & -0.0285 & 0.0138 \\
0.1790 & 0.0048 & -0.1648 & 0.0794 & -0.0162 & 0.0423 & 0.0259 & -0.0174 & 0.0279 \\
0.0455 & -0.2220 & -0.0495 & -0.0422 & -0.0732 & 0.0325 & -0.0118 & 0.0036 \\
30.8840 & 1.9736 & 1.8663 & 2.5462 & -1.2549 & 0.0044 & -1.1895 \\
12.4030 & 0.0826 & 0.9799 & -0.4477 & 0.1270 & -0.4281 \\
16.7750 & 4.5935 & -0.1699 & 0.3917 & -0.1978 \\
9.4950 & -0.3300 & 0.3854 & -0.0251 \\
2.2710 & 0.3456 & 1.0010 \\
\end{pmatrix} \]
and also some other values related to $X$.

The followings are the parameters of the calibrated asymmetric multivariate GH distribution:

$$\begin{align*}
\lambda &= -1.1259 \\
\chi &= 0.3766 \\
\psi &= 0.0005 \\
\mu &= \begin{pmatrix}
0.0052 \\
0.0064 \\
-0.0044 \\
0.1675 \\
0.0099 \\
-0.0214 \\
0.1160 \\
-0.0158 \\
-0.0112 \\
-0.0180
\end{pmatrix} \\
\gamma &= \begin{pmatrix}
-0.0069 \\
-0.0025 \\
0.0107 \\
-0.0807 \\
-0.0491 \\
-0.0471 \\
-0.0807 \\
0.0446 \\
-0.0168 \\
0.0135
\end{pmatrix}
\end{align*}$$

$$\Sigma = \begin{pmatrix}
0.1440 & 0.0526 & 0.0005 & -0.1101 & 0.0356 & -0.0255 & -0.0734 & 0.0153 & -0.0277 & 0.0133 \\
0.1732 & 0.0047 & -0.1596 & 0.0775 & -0.0159 & 0.0406 & 0.0252 & -0.0167 & 0.0272 \\
0.0440 & -0.2135 & -0.0470 & -0.0399 & -0.0697 & 0.0309 & -0.0111 & 0.0353 \\
29.8220 & 1.8966 & 1.7985 & 2.4522 & -1.1873 & 0.0075 & -1.1468 \\
11.9650 & 0.0800 & 0.9397 & -0.4301 & 0.1211 & -0.4127 \\
16.1970 & 4.4408 & -0.1615 & 0.3785 & -0.1906 \\
9.1630 & -0.3155 & 0.3128 & -0.4100 \\
0.9198 & 0.3684 & 0.8932 \\
2.1903 & 0.3343 & 0.9673
\end{pmatrix}$$
\[
\mathbb{E}[\mathbf{X}] = 10^{-2} \times \begin{pmatrix}
-0.0016 \\
0.0039 \\
0.0063 \\
0.0868 \\
-0.0391 \\
-0.0685 \\
0.0353 \\
0.0288 \\
-0.0280 \\
-0.0046
\end{pmatrix}
\]

\[
\text{Cov}(\mathbf{X}) = \begin{pmatrix}
0.1553 & 0.0567 & -0.0173 & 0.0236 & 0.1169 & 0.0526 & 0.0603 & -0.0585 & 0.0002 & -0.0090 \\
0.1747 & -0.0018 & -0.1106 & 0.1073 & 0.0127 & 0.0896 & -0.0019 & -0.0065 & 0.0190 \\
0.0717 & -0.4225 & -0.1742 & -0.1620 & -0.2788 & 0.1463 & -0.0547 & 0.0702 \\
31.3940 & 2.8534 & 2.7170 & 4.0249 & -2.0557 & 0.3353 & -1.4092 \\
12.5470 & 0.6387 & 1.8964 & -0.9584 & 0.3205 & -0.5723 \\
16.7340 & 5.3592 & -0.6687 & 0.5700 & -0.3438 \\
10.7360 & -1.1839 & 0.6405 & -0.6724 \\
1.3993 & 0.1874 & 1.0381 \\
2.2586 & 0.2797 \\
1.0110
\end{pmatrix}
\]

\[
\text{Corr}(\mathbf{X}) = \begin{pmatrix}
1.0000 & 0.3444 & -0.1635 & 0.0107 & 0.0837 & 0.0326 & 0.0467 & -0.1255 & 0.0003 & -0.0227 \\
1.0000 & -0.0160 & -0.0472 & 0.0724 & 0.0074 & 0.0654 & -0.0038 & -0.0103 & 0.0452 \\
1.0000 & -0.2915 & -0.1836 & -0.1479 & -0.3176 & 0.4617 & -0.1358 & 0.2606 \\
1.0000 & 0.1438 & 0.1185 & 0.2192 & -0.3102 & 0.0398 & -0.2501 \\
1.0000 & 0.0441 & 0.1634 & -0.2287 & 0.0602 & -0.1607 \\
1.0000 & 0.3999 & -0.1382 & 0.0927 & -0.0836 \\
1.0000 & -0.3055 & 0.1301 & -0.2041 \\
1.0000 & 0.1054 & 0.8728 \\
1.0000 & 0.1851 \\
1.0000
\end{pmatrix}
\]

Overall, results of the calibrated asymmetric model are similar to those of the symmetric one. Stocks still yield the highest returns amongst other asset classes, with the highest being that of ASII log return. Correlations between assets of the same class are still positive. The difference is that the correlations between zero coupon bonds and INDO38 are negative now, while they are mostly positive with stocks and are mostly positive now. Their correlations with CNY are still negative, but are of varying signs with EUR and USD. Stocks are still negatively correlated with CNY and USD, while being positively correlated with EUR.

### 7.2 Goodness of Fit

In this section, the goodness of fit of calibrated parameters of Generalized Hyperbolic on the data used are analyzed. To gain some confidence that Generalized Hyperbolic provides a good fit, some univariate examples of the data will be analyzed first.
Figures 1 shows the comparison between the QQplot of normal distribution and the asymmetric Generalized Hyperbolic distribution. As was previously discussed, the normal distribution does not provide a good fit since its QQplot is not linear. They form the ‘inverted S’ shape that show that the actual distribution of the data show higher kurtosis than normal and has heavier tail. It seems that the calibrated Generalized Hyperbolic distribution shows a significant improvement over normal distribution. In all cases, the QQplot of the GH distribution is almost linear, showing a good fit to the empirical distribution.

![Figure 1: QQplot comparison between normal and asymmetric GH distribution of daily a) ZC5 yield increment; b) INDO38 yield increment; c) ASII log return; d) CNY log return](image)

To look better at how the kurtosis and skewness of the theoretical distributions match the ones of the actual distributions, a comparison between the histogram of the empirical distribution and the pdf of the theoretical distributions are used as are shown by figures 2-5. The figures include the comparisons between normal distribution, symmetric and asymmetric Generalized Hyperbolic distributions, as well as the marginal distributions of calibrated symmetric and asymmetric Generalized Hyperbolic distributions. The empirical log pdf is generated by using Epanechnikov smoothing kernel (see [20]).

It can be seen in all cases that GH distribution gives the better fit over normal distribution. One of the main reason is that because it can adapt its shape to have the excess kurtosis feature that all of the assets exhibit. Moreover, as is shown by the log pdf plots, the distribution’s tail heaviness can match that of the empirical distribution, unlike the normal one.

Between the univariate and the marginal distributions, it can also be seen that the univariate gives the better fit. The main reason is because in the marginal case,
the shape parameters are fixed as the result of the interdependence structure between the assets, unlike in the univariate case. This lessens the freedom of individual parameter calibration to fit univariate empirical data.

Comparisons between the empirical and theoretical kurtosis and skewness, as well as the log likelihood values of the theoretical distributions can be seen at table 3. The likelihood values are the highest in the univariate cases, as is predicted. For most assets the empirical skewness is significantly different than zero, so the symmetric model is somewhat doubtful. The fitted kurtosis of the univariate models also deviate very much from the empirical kurtosis for some assets, although not as much as in the marginal cases. Although the deviations are large, it does not
necessarily condemn the GH model as providing a poor fit. Further test has to be done by a distance test, in this case, the Kolmogorov-Smirnov statistical test.

7.3 Numerical Results of Markowitz Optimization Problem

Figures 6-8 show the comparison of the efficient frontiers between different distributions for each of the risk measures. In all cases, the risk measure for the normal distribution is lesser than that of symmetric GH distribution. In the case of volatility, the volatility from normal distribution is also smaller than that of asymmetric GH distribution. The same cannot be said for the Expected Shortfall case. Initially,
<table>
<thead>
<tr>
<th>Asset</th>
<th>Empirical</th>
<th>Normal</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>$s$</td>
<td>$k$</td>
</tr>
<tr>
<td>ZC3</td>
<td>0.1176</td>
<td>16.4520</td>
</tr>
<tr>
<td>ZC5</td>
<td>3.0542</td>
<td>45.8650</td>
</tr>
<tr>
<td>INDO38</td>
<td>1.9973</td>
<td>40.3670</td>
</tr>
<tr>
<td>ASII</td>
<td>-0.0528</td>
<td>9.0945</td>
</tr>
<tr>
<td>BACH</td>
<td>0.2904</td>
<td>7.3296</td>
</tr>
<tr>
<td>INTC</td>
<td>0.1575</td>
<td>5.1840</td>
</tr>
<tr>
<td>RDS</td>
<td>0.3179</td>
<td>10.2140</td>
</tr>
<tr>
<td>CNY</td>
<td>7.7082</td>
<td>131.9600</td>
</tr>
<tr>
<td>EUR</td>
<td>2.1573</td>
<td>36.2930</td>
</tr>
<tr>
<td>USD</td>
<td>3.7037</td>
<td>83.3790</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Asset</th>
<th>Symmetric GH</th>
<th>Asymmetric GH</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>$s$ $k$ $\log L$</td>
<td>$s$ $k$ $\log L$</td>
</tr>
<tr>
<td>ZC3</td>
<td>40.7560 266.5800 -1.8266</td>
<td>39.1420 266.9800</td>
</tr>
<tr>
<td>ZC5</td>
<td>33.9100 166.9800 -0.3127</td>
<td>33.8370 167.0100</td>
</tr>
<tr>
<td>INDO38</td>
<td>66.7540 582.8000 1.5014</td>
<td>65.2660 582.9600</td>
</tr>
<tr>
<td>ASII</td>
<td>9.6297 -1481.7000 0.1760</td>
<td>9.5947 -1481.6000</td>
</tr>
<tr>
<td>BACH</td>
<td>8.6667 -1174.6000 0.0194</td>
<td>8.6655 -1174.6000</td>
</tr>
<tr>
<td>INTC</td>
<td>5.5683 -1295.3000 -0.1362</td>
<td>5.4644 -1295.1000</td>
</tr>
<tr>
<td>RDS</td>
<td>17.6340 -1167.7000 -0.8004</td>
<td>16.4360 -1167.2000</td>
</tr>
<tr>
<td>CNY</td>
<td>31.4910 -460.8500 0.8655</td>
<td>30.0330 -460.6200</td>
</tr>
<tr>
<td>EUR</td>
<td>62.1430 -722.5800 2.8576</td>
<td>499.7000 -722.5800</td>
</tr>
<tr>
<td>USD</td>
<td>32.1470 -478.3800 -0.4276</td>
<td>32.2290 -478.3300</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Asset</th>
<th>Marginal Symmetric GH</th>
<th>Marginal Asymmetric GH</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>$s$ $k$ $\log L$</td>
<td>$s$ $k$ $\log L$</td>
</tr>
<tr>
<td>ZC3</td>
<td>1086.8000 110.2100 -16.0320</td>
<td>856.1400 110.2000</td>
</tr>
<tr>
<td>ZC5</td>
<td>1086.8000 23.2110 -4.4976</td>
<td>744.6300 23.0750</td>
</tr>
<tr>
<td>INDO38</td>
<td>1086.8000 223.3800 70.7120</td>
<td>1076.5000 222.8400</td>
</tr>
<tr>
<td>ASII</td>
<td>1086.8000 -3004.4000 -12.3760</td>
<td>819.3300 -3015.3000</td>
</tr>
<tr>
<td>BACH</td>
<td>1086.8000 -2229.2000 -11.7700</td>
<td>812.2700 -2231.4000</td>
</tr>
<tr>
<td>INTC</td>
<td>1086.8000 -2588.7000 -9.3410</td>
<td>787.2300 -2591.3000</td>
</tr>
<tr>
<td>RDS</td>
<td>1086.8000 -2210.9000 -27.5430</td>
<td>957.3000 -2215.0000</td>
</tr>
<tr>
<td>CNY</td>
<td>1086.8000 -586.2800 61.9930</td>
<td>1077.3000 -587.6500</td>
</tr>
<tr>
<td>USD</td>
<td>1086.8000 -609.8200 11.2600</td>
<td>807.4000 -610.2900</td>
</tr>
</tbody>
</table>

Table 3: Comparison between empirical and calibrated univariate distributions of percentage assets returns (times 100), showing log-likelihood values ($\log L$), skewness and kurtosis.
<table>
<thead>
<tr>
<th>Asset</th>
<th>Normal KS</th>
<th>Symmetric GH KS</th>
<th>Asymmetric GH KS</th>
<th>Normal p-value</th>
<th>Symmetric GH p-value</th>
<th>Asymmetric GH p-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>ZC3</td>
<td>0.3383</td>
<td>0.051796</td>
<td>0.080783</td>
<td>0.043042</td>
<td>0.70182</td>
<td></td>
</tr>
<tr>
<td>ZC5</td>
<td>0.31493</td>
<td>0.030318</td>
<td>0.64157</td>
<td>0.028854</td>
<td>0.14052</td>
<td></td>
</tr>
<tr>
<td>INDO38</td>
<td>0.37469</td>
<td>0.043183</td>
<td>0.21466</td>
<td>0.047114</td>
<td>0.98393</td>
<td></td>
</tr>
<tr>
<td>ASII</td>
<td>0.18378</td>
<td>0.018722</td>
<td>0.98477</td>
<td>0.018817</td>
<td>0.14692</td>
<td></td>
</tr>
<tr>
<td>BACH</td>
<td>0.11802</td>
<td>0.046611</td>
<td>0.046717</td>
<td>0.021985</td>
<td>0.9707</td>
<td></td>
</tr>
<tr>
<td>INTD</td>
<td>0.15973</td>
<td>0.022724</td>
<td>0.92802</td>
<td>0.019985</td>
<td>0.9852</td>
<td></td>
</tr>
<tr>
<td>RDS</td>
<td>0.10445</td>
<td>0.023106</td>
<td>0.90688</td>
<td>0.018673</td>
<td>0.77144</td>
<td></td>
</tr>
<tr>
<td>CNY</td>
<td>0.19423</td>
<td>0.023102</td>
<td>0.907</td>
<td>0.027115</td>
<td>0.14052</td>
<td></td>
</tr>
<tr>
<td>EUR</td>
<td>0.097319</td>
<td>0.020691</td>
<td>0.95997</td>
<td>0.020442</td>
<td>0.964</td>
<td></td>
</tr>
<tr>
<td>USD</td>
<td>0.19096</td>
<td>0.023673</td>
<td>0.89089</td>
<td>0.025499</td>
<td>0.8316</td>
<td></td>
</tr>
</tbody>
</table>

Table 4: Comparison between empirical and calibrated univariate models of assets returns, showing Kolmogorov-Smirnov statistics (KS) and p-values.

<table>
<thead>
<tr>
<th>Asset</th>
<th>Marginal Normal KS</th>
<th>Symmetric GH KS</th>
<th>Marginal Asymmetric GH KS</th>
<th>Marginal Normal p-value</th>
<th>Symmetric GH p-value</th>
<th>Marginal Asymmetric GH p-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>ZC3</td>
<td>0.17342</td>
<td>0.17153</td>
<td>1.04E-15</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>ZC5</td>
<td>0.1518</td>
<td>0.15189</td>
<td>2.07E-12</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>INDO38</td>
<td>0.15755</td>
<td>0.16279</td>
<td>3.42E-14</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>ASII</td>
<td>0.061469</td>
<td>0.05985</td>
<td>0.027558</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>BACH</td>
<td>0.062909</td>
<td>0.06158</td>
<td>0.021436</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>INTD</td>
<td>0.038392</td>
<td>0.037</td>
<td>0.38606</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>RDS</td>
<td>0.04057</td>
<td>0.037675</td>
<td>0.36393</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>CNY</td>
<td>0.09</td>
<td>0.085171</td>
<td>0.0003499</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>EUR</td>
<td>0.037214</td>
<td>0.035756</td>
<td>0.42899</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>USD</td>
<td>0.096168</td>
<td>0.095194</td>
<td>3.92E-05</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Table 5: Comparison between log likelihood values for normal and asymmetric univariate models of assets returns, showing Likelihood Ratio test statistics (LR) and p-values.

28
Table 6: Comparison between log likelihood values for symmetric and asymmetric univariate models of assets returns, showing Likelihood Ratio test statistics (LR) and p-values.

<table>
<thead>
<tr>
<th>Asset</th>
<th>Asymmetric GH</th>
<th>Symmetric GH</th>
<th>LR</th>
<th>p-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>ZC3</td>
<td>266.98</td>
<td>266.58</td>
<td>0.80241</td>
<td>0.37037</td>
</tr>
<tr>
<td>ZC5</td>
<td>167.01</td>
<td>166.98</td>
<td>0.058991</td>
<td>0.8081</td>
</tr>
<tr>
<td>INDO38</td>
<td>582.96</td>
<td>582.8</td>
<td>0.33459</td>
<td>0.56297</td>
</tr>
<tr>
<td>ASII</td>
<td>-1481.6</td>
<td>-1481.7</td>
<td>0.22602</td>
<td>0.63497</td>
</tr>
<tr>
<td>BBCA</td>
<td>-1174.6</td>
<td>-1174.6</td>
<td>0.0024875</td>
<td>0.96022</td>
</tr>
<tr>
<td>BACH</td>
<td>-1295.1</td>
<td>-1295.3</td>
<td>0.48384</td>
<td>0.48669</td>
</tr>
<tr>
<td>RDSA</td>
<td>-1167.2</td>
<td>-1167.7</td>
<td>0.92701</td>
<td>0.33564</td>
</tr>
<tr>
<td>CNY</td>
<td>-460.62</td>
<td>-460.85</td>
<td>0.46393</td>
<td>0.49579</td>
</tr>
<tr>
<td>EUR</td>
<td>-722.58</td>
<td>-722.58</td>
<td>0.0077574</td>
<td>0.92982</td>
</tr>
<tr>
<td>USD</td>
<td>-478.33</td>
<td>-478.38</td>
<td>0.10278</td>
<td>0.74852</td>
</tr>
</tbody>
</table>

Figure 6: Portfolio efficient frontiers with unconstrained weights using volatility for asymmetric, symmetric GH and normal case

the frontier using normal distribution has the minimum risk. But, the risk of asymmetric GH frontier is slow to increase and hence at some point, around 0.9 percent expected return, it is surpassed by the risk of normal frontier. Overall, the efficient frontier of asymmetric GH has the steepest slope.

Figures 9 and 10 show the comparison between return-volatility frontiers using different optimal weights: weights obtained from return-volatility, return-VaR an return-ES portfolio optimizations. In the symmetric case, the difference between these frontiers is negligible, proving the consistency of this result with that of the fact that optimal weights from these types of optimizations are equal, as has been shown by equation (6.7). In the asymmetric case, the frontiers clearly differ since the optimal weights are not necessarily equal. The difference is significant in the case of unconstrained weights, ranging from 0.6 percent to 1.6 percent of volatility level. However, if no shorts are permitted, the difference becomes very insignificant for
expected returns more than 0.038 percent, even the maximum is only 0.02 percent of volatility level.

7.4 Numerical Results of RORC Optimization Problem

In the symmetric case, since the global minimum of each risk measure is positive, proposition 6.1 can be used to find the RORCs for all cases more practically. One only has to compute the expected return and volatility levels that give the best RORC, then using equations (6.5) and (6.6), the corresponding Value-at-Risk and Expected Shortfall can be computed. Combined with the expected return, they produce the maximum RORCs. Table 7 give the maximum RORCs in each case as well as the corresponding values of each risk measure and the expected return, while table 8 gives the composition of portfolio assets which yields the maximum RORC in both the constrained and unconstrained cases. Since volatility gives the
Figure 9: Comparison of return-volatility portfolio efficient frontiers using optimal weights from different risk measures at 95% confidence level in symmetric case

Figure 10: Comparison of return-volatility portfolio efficient frontiers using optimal weights from different risk measures at 95% confidence level in asymmetric case

smallest value compared to Value-at-Risk and Expected Shortfall, it’s clear that the RORC for volatility yields the smallest value, while the RORC for Expected Shortfall yields the largest value.

Also note that for the unconstrained cases, the condition for the existence of maximum RORC is not satisfied in every case. The stated RORC in the unconstrained cases is the asymptotic value. For the constrained cases, however, the maximum exists. The composition used to achieve the maximum RORC in the constrained case is shown by table 8. Apparently, only three assets are used: INDO38, ASII and RDSA. This may indicate that the other assets give bad performance and have to be replaced to gain a better portfolio.

The RORC and the composition of portfolio weights for volatility and Expected Shortfall are shown by tables 9-12. For return-volatility optimization, the maximum RORC for the unconstrained case does not exist. The same cannot be be said in the case of Expected Shortfall since we do not have the tool to determine its existence for certain in the unconstrained case. The best that can be done is to analyze the
<table>
<thead>
<tr>
<th>Expected Return</th>
<th>Volatility</th>
<th>VaR_{0.95}</th>
<th>ES_{0.95}</th>
</tr>
</thead>
<tbody>
<tr>
<td>Constrained</td>
<td>0.1100</td>
<td>3.1936</td>
<td>3.3666</td>
</tr>
<tr>
<td>Unconstrained</td>
<td>∞</td>
<td>∞</td>
<td>∞</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>RORC_σ</th>
<th>RORC_{VaR}</th>
<th>RORC_{ES}</th>
</tr>
</thead>
<tbody>
<tr>
<td>Constrained</td>
<td>3.4444</td>
<td>3.2674</td>
</tr>
<tr>
<td>Unconstrained</td>
<td>5.3025</td>
<td>5.1204</td>
</tr>
</tbody>
</table>

Table 7: Maximum percentage RORCs and corresponding percentage expected returns and risks of optimization in symmetric case.

<table>
<thead>
<tr>
<th>ZC03</th>
<th>ZC05</th>
<th>INDO38</th>
<th>ASII</th>
<th>BACH</th>
<th>INTC</th>
<th>RDSA</th>
</tr>
</thead>
<tbody>
<tr>
<td>Constrained</td>
<td>0</td>
<td>0</td>
<td>4.1367</td>
<td>40.9615</td>
<td>0</td>
<td>0</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>CNY</th>
<th>EUR</th>
<th>USD</th>
</tr>
</thead>
<tbody>
<tr>
<td>Constrained</td>
<td>0</td>
<td>0</td>
</tr>
</tbody>
</table>

Table 8: Percentage weights of portfolio assets at maximum RORC in symmetric case.

<table>
<thead>
<tr>
<th>Expected Return</th>
<th>Volatility</th>
<th>RORC</th>
</tr>
</thead>
<tbody>
<tr>
<td>Constrained</td>
<td>0.035</td>
<td>1.0117</td>
</tr>
<tr>
<td>Unconstrained</td>
<td>∞</td>
<td>∞</td>
</tr>
</tbody>
</table>

Table 9: Maximum percentage RORCs and corresponding percentage expected return and volatility of return-volatility optimization in asymmetric case.

<table>
<thead>
<tr>
<th>ZC03</th>
<th>ZC05</th>
<th>INDO38</th>
<th>ASII</th>
<th>BACH</th>
<th>INTC</th>
<th>RDSA</th>
</tr>
</thead>
<tbody>
<tr>
<td>Constrained</td>
<td>2.4702</td>
<td>0</td>
<td>0</td>
<td>12.9822</td>
<td>0</td>
<td>0</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>CNY</th>
<th>EUR</th>
<th>USD</th>
</tr>
</thead>
<tbody>
<tr>
<td>Constrained</td>
<td>81.3363</td>
<td>0</td>
</tr>
</tbody>
</table>

Table 10: Percentage weights of portfolio assets at maximum RORC of return-volatility optimization in asymmetric case.

<table>
<thead>
<tr>
<th>Expected Return</th>
<th>ES_{0.95}</th>
<th>RORC</th>
</tr>
</thead>
<tbody>
<tr>
<td>Constrained</td>
<td>0.0290</td>
<td>1.4507</td>
</tr>
<tr>
<td>Unconstrained</td>
<td>0.0860</td>
<td>1.3766</td>
</tr>
</tbody>
</table>

Table 11: Maximum percentage RORCs and corresponding percentage expected return and Expected Shortfall of return-Expected Shortfall optimization in asymmetric case. Note: Valuation in unconstrained case is done for percentage expected return in (0,0.086]
Table 12: Percentage weights of portfolio assets at maximum RORC of return-Expected Shortfall at 95% confidence level optimization in asymmetric case.

<table>
<thead>
<tr>
<th></th>
<th>ZC03</th>
<th>ZC05</th>
<th>INDO38</th>
<th>ASII</th>
<th>BACH</th>
<th>INTC</th>
<th>RDSA</th>
</tr>
</thead>
<tbody>
<tr>
<td>Constrained</td>
<td>17.2417</td>
<td>0</td>
<td>0</td>
<td>8.0013</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Unconstrained</td>
<td>48.6934</td>
<td>13.5626</td>
<td>-6.9148</td>
<td>4.0255</td>
<td>0.0524</td>
<td>-2.7210</td>
<td>1.5085</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th></th>
<th>CNY</th>
<th>EUR</th>
<th>USD</th>
</tr>
</thead>
<tbody>
<tr>
<td>Constrained</td>
<td>74.7571</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Unconstrained</td>
<td>219.6778</td>
<td>-9.5916</td>
<td>-168.2929</td>
</tr>
</tbody>
</table>

Table 12: Percentage weights of portfolio assets at maximum RORC of return-Expected Shortfall at 95% confidence level optimization in asymmetric case.

trend of the RROC plot against the expected return as is shown by figure 11. By figure 11.d, the trend is increasing with no decreasing sign. So, it can be guessed that the RORC for Expected Shortfall in the unconstrained case does not have a maximum. Note that the RORC value in the unconstrained section in table 11 is the maximum RORC over the interval (0, 0.0860) of percentage expected return.

Figure 11 also shows that in all cases the RORC achieved by Expected Shortfall is the lowest, followed by VaR, where the highest is achieved using volatility as risk measure. The results are consistent with fact that $VaR_\beta \leq ES_\beta$, for any given $\beta$, although volatility is not necessarily less than both $VaR$ and $ES$.

Figure 11: RORC for varying expected returns: a) Symmetric case with constrained weight; b) Symmetric case with unconstrained weight; c) Asymmetric case with constrained weight; d) Asymmetric case with unconstrained weight

Finally, table 13 gives a benchmark for RORC values according to JKSE composite index return. The volatility of the index return is measured by sample standard deviation, the Value-at-Risk by sample quantile and the Expected Shortfall by the statistic given by Lemma 3.1. According to this benchmark, every portfolio acquired
Table 13: Maximum percentage RORCs of JKSE composite index at 95% confidence level.

<table>
<thead>
<tr>
<th>RORC_σ</th>
<th>RORC_{VaR}</th>
<th>RORC_ES</th>
</tr>
</thead>
<tbody>
<tr>
<td>2.2192</td>
<td>1.6179</td>
<td>0.7728</td>
</tr>
</tbody>
</table>

in this paper using both symmetric and asymmetric Generalized Hyperbolic distribution give better performance in both the constrained and unconstrained cases. This shows that the optimization method used in this paper is more superior than the one used by JKSE.

8 Concluding Remarks

In this paper, the author develops a way to solve portfolio optimization problems when Value-at-Risk and Expected Shortfall are used as risk measures, and when assets return distribution is modelled by Generalized Hyperbolic distribution. In the framework of symmetric Generalized Hyperbolic, the author managed to obtain analytical solutions to Markowitz and RORC portfolio optimization problems using Value-at-Risk and Expected Shortfall. The solutions can be obtained with the help of the linearity property of Generalized Hyperbolic. Moreover, by the linearity of Generalized Hyperbolic distribution, the optimal weights obtained in Markowitz optimization model using Value-at-Risk and Expected Shortfall have been found to be equal to those which are yielded from using volatility as risk measure. In short, the problem is reduced to classical Markowitz optimization problem.

In the asymmetric framework, the optimization problems using Expected Shortfall are solved numerically. Originally, evaluating Expected Shortfall function means evaluating high dimensional integral. But, with the linearity property of the Generalized Hyperbolic distribution, this almost intractable problem can be greatly simplified into a one dimensional integral problem. The Markowitz-optimal composition obtained from using Expected Shortfall have also been found to be not necessarily equal to those obtained from using volatility. But, when no shortings are allowed, the compositions are very similar, especially for high expected returns. The nonconvexity of the RORC problem is able to be avoided by transforming it into a number of convex Markowitz optimization problems. The number depends on the desired accuracy.

In many cases, especially when asset shortings are allowed, optimal weights cannot be achieved for the RORC optimization version. The reason is that the existence condition for the maximum RORC is not satisfied (its supremum however can be achieved). However, when Expected Shortfall is used in the asymmetric framework, the author does not have an existence condition of the maximum RORC. Hence, the author resorts to a plot of RORC against portfolio expected return. The
result is that the trend of the plot is forever increasing to an asymptotic value. Moreover, the plot behaves similarly with the plot obtained when volatility is used as the risk measure, which in fact, is produced from the case where optimal weights cannot be achieved. These findings provide the author some confidence that the maximum RORC in the original case is unachievable and that the RORC tends to some asymptotic value.

An important use of the RORC optimization model is that it serves as a tool to compare the performance of portfolios used in this paper with a benchmark portfolio. In this paper, JKSE portfolio is used as the benchmark. The result is that the portfolios produced using both symmetric and asymmetric models of Generalized Hyperbolic distribution perform better than JKSE portfolio, with performance measured by RORC. They are better at any type of risk measures used in this paper. This concludes that the optimization method used in this paper is more superior than the one used by JKSE.

A Calibration of GHYP Distribution

A.1 EM Algorithm

EM algorithm is a tool to estimate unknown parameters of a distribution. This estimate is based on the maximum likelihood method.

**Definition A.1 Likelihood Function** Suppose $X = (X_1, \ldots, X_n)$ is a vector of $n$ independent and identically distributed (i.i.d.) random variables $X_i \in \mathbb{R}^d$, called random samples, with pdf $f$. Denote the parameter space of the distribution by $\Omega$. Define the likelihood function as the joint density of the random samples, denoted by

$$L(\theta; X) := \prod_{i=1}^{n} f(X_i; \theta), \quad \theta \in \Omega$$

**Theorem A.2** Suppose $X = (X_1, \ldots, X_n)$ is a vector of $n$ independent and identically distributed (i.i.d.) random variables $X_i \in \mathbb{R}^d$ with pdf $f$ and parameter space $\Omega$. Let $\theta_0$ be the true parameter of this distribution. If the pdf has common support for all $\theta \in \Omega$, then

$$\lim_{n \to \infty} \mathbb{P}(L(\theta_0; X) \geq L(\theta; X)) = 1, \text{ for all } \theta \neq \theta_0$$

For proof, see [17].

Based on theorem, an estimate of the true distribution parameter can obtained as

$$\hat{\theta} = \arg \max \theta L(\theta; X).$$

This estimate is called the maximum likelihood estimator (MLE).

In some cases, the maxima of the likelihood function is intractable due to the existence of unobserved latent random variables. In our case, this unobserved variable can be regarded as the mixing random variable $W$ in (2.2), while $X$ is the observed
variable. Using this unobserved variable, together with the observed variable, the EM algorithm is created to estimate the MLE.

**Algorithm A.3 EM Algorithm.** Denote the observed random variables by \( X = (X_1, \ldots, X_n) \) and the unobserved random variables by \( Y = (Y_1, \ldots, Y_n) \), where \( X_i \in \mathbb{R}^d \) and \( Y_i \in \mathbb{R}^{\tilde{d}} \). Assume that the \( X_i \)s are i.i.d. and so are the \( Y_i \)s. Furthermore, let \( X \) be independent from \( Y \). Let

\[
\tilde{L} := \tilde{L}(\theta; X, Y),
\]

the joint pdf of \( X \) and \( Y \), where \( \theta \) is the parameter of \( X \) distribution. The EM algorithm is as follows

1. Give an initial estimate \( \hat{\theta}^{(0)} \) of the true distribution parameter.

2. **Expectation Step.** Compute

\[
Q(\theta; \hat{\theta}^{(0)}) := \mathbb{E} \left[ \log \tilde{L}(\theta; X, Y) \mid X, \hat{\theta}^{(0)} \right]
\]

3. **Maximization Step.** Compute

\[
\hat{\theta}^{(1)} = \arg \max_{\theta} Q(\theta; \hat{\theta}^{(0)}),
\]

the estimate of the first iteration.

4. Obtain the estimate of the \( m + 1 \)-th iteration by executing step 2 and step 3 using the \( m \)-th estimate \( \hat{\theta}^{(m)} \).

**Theorem A.4** The sequence of estimates \( \hat{\theta}^{(m)} \), defined by algorithm A.3, satisfies

\[
L(\hat{\theta}^{(m+1)}; X) \geq L(\hat{\theta}^{(m)}; X).
\]

For proof, see [17].

The theorem does not guarantee that the EM estimates will converge to the MLE, but, it increases the likelihood function on every executed iteration. This is the basis for estimating the MLE.

### A.2 Calibration Using EM

To use the EM algorithm to calibrate GH distribution, the observed and unobserved variables must be identified. From representation (2.2), the random variable \( X \) can be regarded as the observed variable, while \( W \) as the unobserved variable. The next step is to formulate \( Q(\theta; \hat{\theta}^{(k)}) \), where \( \hat{\theta}^{(k)} = (\lambda^{(k)}, \chi^{(k)}, \psi^{(k)}, \mu^{(k)}, \Sigma^{(k)}, \gamma^{(k)}) \).

Let \( X = (X_1, \ldots, X_n) \) be a vector of \( n \) independent and identically distributed (i.i.d.) GH random variables \( X_i \in \mathbb{R}^d \), and \( W = (W_1, \ldots, W_n) \) be a vector of \( n \) independent and identically distributed nonnegative random variables, and that \( X \) is independent form \( W \). Suppose the \( k \)-th estimate \( \hat{\theta}^{(k)} \) has been obtained, then
the function \( \log \tilde{L} \) needs to be calculated. Since from representation (2.2) we have 
\[
X_i|W_i \sim N(\mu + W_i\gamma, W_i\Sigma),
\]
\[
\log \tilde{L}(\theta; X, Y) = \sum_{i=1}^{n} \log f_{X_i|W_i}(x_i|w_i; \mu, \Sigma, \gamma) + \sum_{i=1}^{n} \log h_{W_i}(w_i; \lambda, \chi, \psi). \tag{A.8}
\]

Let 
\[
L_1 := \sum_{i=1}^{n} \log f_{X_i|W_i}(x_i|w_i; \mu, \Sigma, \gamma), \quad \text{and} \tag{A.9}
\]
\[
L_2 := \sum_{i=1}^{n} \log h_{W_i}(w_i; \lambda, \chi, \psi). \tag{A.10}
\]
Then, to get the next estimate, we have to maximize 
\[
Q(\theta; \hat{\theta}) = \mathbb{E}[L_1|X, \hat{\theta}^{(k)}] + \mathbb{E}[L_2|X, \hat{\theta}^{(k)}]. \tag{A.11}
\]

From (A.8) and (A.11), it is evident that we can maximize each term of (A.11) separately. This is because, from (A.8), 
\( L_1 \) does not contain \( \lambda, \chi, \psi \) and 
\( L_2 \) does not contain \( \mu, \Sigma, \gamma \).

From the formula of the pdf of multivariate normal distribution 
\[
N(\mu, \Sigma) \text{ of dimension } d 
\]
\[
f(x) = \frac{1}{(2\pi)^{d/2}|\Sigma|^{1/2}} e^{\frac{1}{2} (x - \mu)' \Sigma^{-1} (x - \mu)}, \tag{A.12}
\]
we have 
\[
f_{X_i|W_i}(x_i|w_i; \mu, \Sigma, \gamma) = \frac{1}{(2\pi)^{d/2}|w_i\Sigma|^{1/2}} \times
\]
\[
\exp \left\{ -\frac{1}{2} (x_i - (\mu + w_i\gamma))' \frac{1}{w_i} \Sigma^{-1} (x_i - (\mu + w_i\gamma)) \right\}, \tag{A.13}
\]
\[
= \frac{1}{(2\pi)^{d/2} w_i^{1/2} |\Sigma|^{1/2}} e^{(x_i - \mu)' \Sigma^{-1} \gamma} e^{-\frac{1}{2} \gamma' w_i^{-1} \gamma} e^{-\frac{1}{2} \gamma' \Sigma^{-1} \gamma}. \tag{A.14}
\]
Hence, 
\[
L_1 = -\frac{n}{2} \log |\Sigma| - \frac{d}{2} \sum_{i=1}^{n} \log w_i + \sum_{i=1}^{n} (x_i - \mu)' \Sigma^{-1} \gamma 
- \frac{1}{2} \sum_{i=1}^{n} \frac{1}{w_i} \rho_{x_i} - \frac{1}{2} \gamma' \Sigma^{-1} \gamma \sum_{i=1}^{n} w_i. \tag{A.15}
\]
If \( W \sim GIG(\lambda, \chi, \psi) \), \( L_2 \) takes the form 
\[
L_2 = (\lambda - 1) \sum_{i=1}^{n} \log w_i - \frac{\chi}{2} \sum_{i=1}^{n} w_i^{-1} - \frac{\psi}{2} \sum_{i=1}^{n} w_i - \frac{n\lambda}{2} \log \chi 
+ \frac{n\lambda}{2} \log \psi - n \log \left( 2K_{\lambda} \left( \sqrt{\lambda\psi} \right) \right). \tag{A.16}
\]
In (A.15)-(A.16), the factor $w_i$ comes in the form of $w_i, w_i^{-1}$, and log $w_i$. Hence, to compute $E[L_2|X, \hat{\theta}^{(k)}]$, it is necessary to compute

$$\eta_i^{(k)} := E[W_i|X, \hat{\theta}^{(k)}], \quad \delta_i^{(k)} := E[W_i^{-1}|X, \hat{\theta}^{(k)}], \quad \xi_i^{(k)} := E[\log W_i|X, \hat{\theta}^{(k)}]. \quad (A.17)$$

Also define

$$\bar{\eta}_i^{(k)} = \frac{1}{n} \sum_{i=1}^{n} \eta_i^{(k)}, \quad \bar{\delta}_i^{(k)} = \frac{1}{n} \sum_{i=1}^{n} \delta_i^{(k)}, \quad \bar{\xi}_i^{(k)} = \frac{1}{n} \sum_{i=1}^{n} \xi_i^{(k)} \quad (A.18)$$

To compute them, the distribution of $W_i|X_i; \hat{\theta}^{(k)}$ needs to be determined. It can be determined using its pdf with formula

$$f_{W_i|X_i}(w_i|x_i; \hat{\theta}^{(k)}) = \frac{f_{X_i|W_i}(x_i|w_i; \hat{\theta}^{(k)}) h_{W_i}(w_i; \hat{\theta}^{(k)})}{f_{X_i}(x_i; \hat{\theta}^{(k)})} \quad (A.19)$$

By substituting (2.1), (2.4), and (A.14) into (A.19), it can be checked cross checked with formula (2.1) that the resulting pdf is of GIG distribution $N^\sim(\lambda^{(k)} - \frac{d}{2}, \rho_{x_i^{(k)}} + \chi^{(k)}, \psi^{(k)} + \gamma^{(k)} \Sigma^{-1(k)} \gamma^{(k)})$. Using (??) and (??), we have

$$\begin{align*}
\delta_i^{(k)} &= \left( \frac{\rho_{x_i^{(k)}} + \chi^{(k)}}{\psi^{(k)} + \gamma^{(k)} \Sigma^{-1(k)} \gamma^{(k)}} \right)^{-\frac{1}{2}} \\
&\quad \times \left( K_{\lambda^{(k)} - \frac{d}{2} - 1} \left( \sqrt{\frac{1}{K_{\lambda^{(k)} - \frac{d}{2}} \left( \sqrt{\rho_{x_i^{(k)}} + \chi^{(k)}} \left( \psi^{(k)} + \gamma^{(k)} \Sigma^{-1(k)} \gamma^{(k)} \right) \right)} \right) \right) \quad (A.20)
\end{align*}$$

$$\begin{align*}
\eta_i^{(k)} &= \left( \frac{\rho_{x_i^{(k)}} + \chi^{(k)}}{\psi^{(k)} + \gamma^{(k)} \Sigma^{-1(k)} \gamma^{(k)}} \right)^{\frac{1}{2}} \\
&\quad \times \left( K_{\lambda^{(k)} - \frac{d}{2} + 1} \left( \sqrt{\frac{1}{K_{\lambda^{(k)} - \frac{d}{2}} \left( \sqrt{\rho_{x_i^{(k)}} + \chi^{(k)}} \left( \psi^{(k)} + \gamma^{(k)} \Sigma^{-1(k)} \gamma^{(k)} \right) \right)} \right) \right) \quad (A.21)
\end{align*}$$

$$\begin{align*}
\xi_i^{(k)} &= \frac{1}{2} \log \left( \frac{\rho_{x_i^{(k)}} + \chi^{(k)}}{\psi^{(k)} + \gamma^{(k)} \Sigma^{-1(k)} \gamma^{(k)}} \right) + \\
&\quad \frac{\partial K_{\lambda^{(k)} - \frac{d}{2} + a} \left( \sqrt{\frac{1}{K_{\lambda^{(k)} - \frac{d}{2}} \left( \sqrt{\rho_{x_i^{(k)}} + \chi^{(k)}} \left( \psi^{(k)} + \gamma^{(k)} \Sigma^{-1(k)} \gamma^{(k)} \right) \right)} \right)}{\partial \alpha} \bigg|_{\alpha=0} \\
&\quad \frac{1}{K_{\lambda^{(k)} - \frac{d}{2}} \left( \sqrt{\frac{1}{K_{\lambda^{(k)} - \frac{d}{2}} \left( \sqrt{\rho_{x_i^{(k)}} + \chi^{(k)}} \left( \psi^{(k)} + \gamma^{(k)} \Sigma^{-1(k)} \gamma^{(k)} \right) \right)} \right) \right) \quad (A.22)
\end{align*}$$

So, to obtain $E[L_1|X, \hat{\theta}^{(k)}]$ and $E[L_2|X, \hat{\theta}^{(k)}]$, we only have to substitute (A.20) to (A.22) into (A.15) and (A.16). Now, we can proceed to the maximization process.
Since $L_1$ does not depend on the distribution of $W$ as explained before, we only need to consider the values of the parameters $\mu, \Sigma, \gamma$ to maximize $\mathbb{E}[L_1|X, \hat{\theta}^{(k)}]$. Following the standard routine optimization as Hu does (see [18]), set

\[
\frac{\partial \mathbb{E}[L_1|X, \hat{\theta}^{(k)}]}{\partial \mu} = 0 \quad (A.23)
\]

\[
\frac{\partial \mathbb{E}[L_1|X, \hat{\theta}^{(k)}]}{\partial \gamma} = 0 \quad (A.24)
\]

\[
\frac{\partial \mathbb{E}[L_1|X, \hat{\theta}^{(k)}]}{\partial \Sigma} = 0 \quad (A.25)
\]

We can find the next estimate $(\mu^{(k+1)}, \Sigma^{(k+1)}, \gamma^{(k+1)})$ by solving the above system. The solution to the system can be estimated by first solving (A.23) to find $\mu^{(k+1)}$, and then substituting it to (A.24) to find $\gamma^{(k+1)}$. Finally, $\Sigma^{(k+1)}$ can be obtained by substituting $(\mu^{(k+1)}, \gamma^{(k+1)})$ into (A.25). The process is as follows.

Using (A.15), we have

\[
\frac{\partial \mathbb{E}[L_1|X, \hat{\theta}^{(k)}]}{\partial \mu} = -n\gamma'\Sigma^{-1} + \left( \sum_{i=1}^{n} \delta_i^{(k)} (x_i - \mu)' \right) \Sigma^{-1} \quad (A.26)
\]

\[
\frac{\partial \mathbb{E}[L_1|X, \hat{\theta}^{(k)}]}{\partial \gamma} = \Sigma^{-1} \left( \sum_{i=1}^{n} (x_i - \mu) - \sum_{i=1}^{n} \eta_i^{(k)} \gamma \right) \quad (A.27)
\]

\[
\frac{\partial \mathbb{E}[L_1|X, \hat{\theta}^{(k)}]}{\partial \Sigma} = \Sigma^{-1} \left( -\frac{n}{2} - \sum_{i=1}^{n} (x_i - \mu)' \Sigma^{-1} + \frac{1}{2} \sum_{i=1}^{n} \delta_i^{(k)} (x_i - \mu) (x_i - \mu)' \times \right.
\]

\[\left. \Sigma^{-1} + \sum_{i=1}^{n} \eta_i^{(k)} \gamma \gamma' \Sigma^{-1} \right) \quad (A.28)
\]

By setting equation (A.26) to 0, we can get the next estimate for $\mu$ as

\[
\mu^{(k+1)} = \frac{n^{-1} \sum_{i=1}^{n} \delta_i^{(k)} x_i - \gamma}{\delta^{(k)}} \quad (A.29)
\]

By setting equation (A.27) to 0 and then substituting $\mu$ with formula (A.29), we get the estimate

\[
\gamma^{(k+1)} = \frac{n^{-1} \sum_{i=1}^{n} \delta_i^{(k)} (\bar{x} - x_i)}{\delta^{(k)} \eta^{(k)} - 1} \quad (A.30)
\]

From (A.30), since $\gamma^{(k+1)}$ does not depend on $\mu^{(k+1)}$, $\gamma$ in (A.29) can be changed into $\gamma^{(k+1)}$ to get

\[
\mu^{(k+1)} = \frac{n^{-1} \sum_{i=1}^{n} \delta_i^{(k)} x_i - \gamma^{(k+1)}}{\delta^{(k)}} \quad (A.31)
\]
By setting equation (A.28) to 0 and then substituting \( \mu \) and \( \gamma \) with formula (A.31) and (A.30), we get the estimate

\[
\Sigma^{(k+1)} = \frac{1}{n} \sum_{i=1}^{n} \delta_i^{(k)} (x_i - \mu^{(k+1)}) (x_i - \mu^{(k+1)})' - \bar{\eta}^{(k)} \gamma^{(k+1)} \gamma^{(k+1)}'.
\] (A.32)

The maximization of \( \mathbb{E}[L_2 | X, \hat{\theta}^{(k)}] \) depends on the parameter \((\lambda, \chi, \psi)\). In calibrating GH distribution, it is intractable to do so because of \( \lambda \). In order to make it tractable, we have to fix \( \lambda \). As Hu [18] stated, no has ever calibrated \( \lambda \) exactly, so what we can do is executing the algorithm for several choices of our \( \lambda \) to get the values of the other parameters. After we have done so, we evaluate the likelihood function over those sets of values, and then we make the conjecture for the value of \( \lambda \) that results in the highest likelihood value. So, from now on, we calibrate GH distribution only for fixed \( \lambda \). Note that \( \lambda \) from other distributions can still be calibrated.

Define

\[
\phi := \sqrt{\chi \psi}.
\] (A.33)

To overcome the identifiability problem when calibrating GH distribution, one of the free parameters must be fixed. To make it simple, either \( \chi \) or \( \psi \) can be fixed. By substituting (A.20)-(A.21) into (A.16) we obtain

\[
\mathbb{E}[L_2 | X, \hat{\theta}^{(k)}] = (\lambda - 1) \sum_{i=1}^{n} \xi_i^{(k)} - \frac{\chi}{2} \sum_{i=1}^{n} \delta_i^{(k)} - \frac{\psi}{2} \sum_{i=1}^{n} \eta_i^{(k)} - \frac{n \lambda}{2} \log \chi
\]

\[
+ \frac{n \lambda}{2} \log \psi - n \log (2 K_{\lambda} (\phi)).
\] (A.34)

If we fix \( \chi \), to maximize (A.34) we need to set

\[
\frac{\partial \mathbb{E}[L_2 | X, \hat{\theta}^{(k)}]}{\partial \psi} = 0.
\] (A.35)

Using the formula

\[
\frac{d \log K_{\lambda} (x)}{dx} = \frac{\lambda}{x} - \frac{K_{\lambda+1} (x)}{K_{\lambda} (x)}
\] (A.36)

, we have

\[
\frac{\partial \mathbb{E}[L_2 | X, \hat{\theta}^{(k)}]}{\partial \psi} = -\frac{n}{2} \left( \bar{\eta}^{(k)} - \sqrt{\chi} K_{\lambda+1} (\phi) \right).
\] (A.37)

Then, we can get \( \phi^{(k+1)} \) by solving

\[
\phi \bar{\eta}^{(k)} K_{\lambda} (\phi) - K_{\lambda+1} (\phi) \chi = 0.
\] (A.38)

Hence, we can get \( \psi^{(k+1)} \) by

\[
\psi^{(k+1)} = \frac{\phi^{(k+1)}}{\chi}.
\] (A.39)
This method where $\chi$ is fixed will be called as $\chi$-algorithm.

If we fix $\psi$, to maximize (A.34) we need to set

$$\frac{\partial E[L_2|X, \hat{\theta}^{(k)}]}{\partial \chi} = 0.$$  
(A.40)

Using the formula

$$\frac{d \log K_\lambda(x)}{dx} = -\frac{\lambda}{x} - \frac{K_{\lambda-1}(x)}{K_\lambda(x)},$$  
(A.41)

we have

$$\frac{\partial E[L_2|X, \hat{\theta}^{(k)}]}{\partial \chi} = -\frac{n}{2} \left( \frac{\delta^{(k)}}{\hat{\chi}} + \sqrt{\frac{\psi}{\chi}} \frac{K_{\lambda-1}(\phi)}{K_\lambda(\phi)} \right).$$  
(A.42)

Then, we can get $\phi^{(k+1)}$ by solving

$$\phi \delta^{(k)} K_\lambda(\phi) - K_{\lambda-1}(\phi) \psi = 0.$$  
(A.43)

Hence, we can get $\chi^{(k+1)}$ by

$$\chi^{(k+1)} = \frac{\phi^{(k+1)^2}}{\psi}.$$  
(A.44)

This method where $\psi$ is fixed will be called as $\psi$-algorithm.

References


