



---

*Research article*

## **Portfolio selection based on asymmetric Laplace distribution, coherent risk measure, and expectation-maximization estimation**

Yue Shi<sup>1</sup>, Chi Tim Ng<sup>2,\*</sup> and Ka-Fai Cedric Yiu<sup>3</sup>

<sup>1</sup> School of Mathematics and Systems Science, Beihang University, Beijing, China

<sup>2</sup> Department of Statistics, Chonnam National University, 500-757, Gwangju, Korea

<sup>3</sup> Department of Applied Mathematics, Hong Kong Polytechnic University, Kowloon, Hong Kong

\* **Correspondence:** Email: easterlyng@gmail.com; Tel: +821049663177.

**Abstract:** In this paper, portfolio selection problem is studied under Asymmetric Laplace Distribution (ALD) framework. Asymmetric Laplace distribution is able to capture tail-heaviness, skewness, and leptokurtosis observed in empirical financial data that cannot be explained by traditional Gaussian distribution. Under Asymmetric Laplace distribution framework, portfolio selection methods based on different risk measures are discussed. Moreover, we derived the Expectation-Maximization (EM) procedure for parameter estimation of Asymmetric Laplace distribution. Performance of the proposed method is illustrated via extensive simulation studies. Two real data examples are complemented to confirm that the Asymmetric Laplace distribution based portfolio selection models are efficient.

**Keywords:** portfolio selection; Asymmetric Laplace distributions; EM procedure; tail-heaviness; Skewness

**JEL codes:** G11

---

### **1. Introduction**

Portfolio selection aims at either maximizing the return or minimizing the risk. In 1952, Markowitz (1952) suggests to select the portfolio by minimizing the standard deviation at a given expected return under the assumption that asset returns are normally distributed. This means that standard deviation is chosen as the risk measure. Markowitz's work laid down the cornerstone for modern portfolio selection theory framework.

Risk measures and probability distributions are two important constituents of the portfolio selection theory. Traditional Markowitz's model (Markowitz, 1952) is established based on normality assumption and standard deviation is chosen as the risk measure.

One disadvantage of taking standard deviation (StD) as a risk measure is that the loss in the extreme cases tends to be underestimated. To overcome such a difficulty, the idea of Value at Risk (VaR) is also widely used in practice. Artzner et al. (1999) suggests that desirable risk measure should be “coherent”. However, VaR does not fulfill the subadditivity condition as required by the definition of “coherence”. Yiu (2004) proposed an optimal portfolio selection under Value-at-Risk. On the other hand, Expected Shortfall (ES) is coherent as a popular risk measure for portfolio selection that aims at averaging the tail uncertainties.

It is well-known that financial data cannot be described satisfactorily by normal distribution. The normality assumption is restrictive and is generally violated due to financial market uncertainties and managers’ risk aversion. As Behr and Pter (2009) pointed out, alternatives for multivariate normal distribution are necessary for portfolio selection. A desirable alternative model should be able to explain tail heaviness, skewness, and excess kurtosis. Various heavy tailed distributions have been applied to portfolio selection problems. Among these, Mandelbrot (1997) concluded that the daily rate of return of stock price data exhibit heavy tailed distributions; Hu and Kercheval (2010) apply multivariate skewed  $t$  and student  $t$  distribution for efficient frontier analysis; Generalized hyperbolic distribution is extensively studied in (Behr and Pter, 2009; Eberlein, 2001; Hellmich and Kassberger, 2011; Hu and Kercheval, 2007; Surya and Kurniawan, 2014; Socgnia and Wilcox, 2014), with special cases including hyperbolic distribution (Bingham and Kiesel, 2001; Eberlein and Keller, 1995), Variance Gamma distribution (Seneta, 2004), Normal Inverse Gaussian distribution (Barndor-Nielsen, 1995), etc.

Recently, Asymmetric Laplace distribution has received various attention in the literature, to name a few, (Ayebo and Kozubowski, 2003; Kollo and Srivastava, 2005; Kozubowski and Podgrski, 1999; Kozubowski and Podgrski, 2001; Punathumparambath, 2012). Compared to Normal distribution, the Asymmetric Laplace distribution describes asymmetry, steep peak, and tail heaviness better. Portfolio selection models are extensively studied under Asymmetric Laplace framework. Zhu (2007), Kozubowski and Podgrski (2001) apply Asymmetric Laplace distribution to financial data. By assuming that the asset data is generated from autoregressive moving average (ARMA) time series models with Asymmetric Laplace noise, Zhu (2007) establish the asymptotic inference theory under very mild conditions and present methods of computing conditional Value at Risk (CVaR). Zhao et al. (2015) further propose a so-called mean-CVaR-skewness portfolio selection strategy under Asymmetric Laplace distribution, this model can be further transformed to quadratic programming problem with explicit solutions.

In this paper, we extended Hu’s work (Hu, 2010) to Asymmetric Laplace framework. We first derived the equivalence of mean-VaR/ES/Std-skewness-kurtosis models, and show that these models can be reduced to quadratic programming problem. Since Zhao et al. (2015) utilized moment estimation for parameter estimation of Asymmetric Laplace distribution which less efficient compare to maximum likelihood estimation. Taken into consideration of the normal mean-variance mixture of Asymmetric Laplace distribution, followed by Expectation-Maximization algorithm for multivariate Laplace distribution in Arslan (2010), we derived the EM algorithm for Asymmetric Laplace Distributions that outperforms moment estimation in Zhao et al. (2015). The advantage of the proposed EM algorithm is to alleviate the complicated calculation of Bessel function. This improves many existing methods of estimating Asymmetric Laplace distributions, for example, Hrlimann (2013), Kollo and Srivastava (2005) and Visk (2009). Extensive simulation studies and efficient

frontier analysis are complemented to confirm that our algorithm performs better than moment estimation for parameter estimation.

The rest of the article is organized as follows. In Section 2, properties of Asymmetric Laplace distributions and coherent risk measures are summarized. In Section 3, portfolio selection under Asymmetric Laplace framework are derived, complement with Expectation-Maximization (EM) algorithm for parameter estimation of Asymmetric Laplace distributions. In Section 4, simulation studies are provided to show the efficiency of the Expectation-Maximization procedure. Section 5 presents real data analysis of Asymmetric Laplace Distributions based portfolio selection models, followed by conclusive remarks in Section 6.

## 2. Preliminary knowledge

### 2.1. Asymmetric Laplace distribution

Kotz et al. (2001) proposed the Asymmetric Laplace Distribution with density function

$$f(\mathbf{x}) = \frac{2e^{\mathbf{x}'\boldsymbol{\Sigma}^{-1}\boldsymbol{\mu}}}{(2\pi)^{n/2}|\boldsymbol{\Sigma}|^{1/2}} \left( \frac{\mathbf{x}'\boldsymbol{\Sigma}^{-1}\mathbf{x}}{2 + \boldsymbol{\mu}'\boldsymbol{\Sigma}^{-1}\boldsymbol{\mu}} \right)^{v/2} K_v(\sqrt{(2 + \boldsymbol{\mu}'\boldsymbol{\Sigma}^{-1}\boldsymbol{\mu})(\mathbf{x}'\boldsymbol{\Sigma}^{-1}\mathbf{x})}), \quad (2.1)$$

denoted as  $\mathbf{X} \sim AL_n(\boldsymbol{\mu}, \boldsymbol{\Sigma})$ . Here,  $n$  is the dimension of random vector  $\mathbf{X}$ ,  $v = (2 - n)/2$  and  $K_v(u)$  is the modified Bessel function of the third kind with the following two popular representations:

$$K_v(u) = \frac{1}{2} \left( \frac{u}{2} \right)^v \int_0^\infty t^{-v-1} \exp\left\{-t - \frac{u^2}{4t}\right\} dt, \quad u > 0, \quad (2.2)$$

$$K_v(u) = \frac{(u/2)^v \Gamma(1/2)}{\Gamma(v + 1/2)} \int_1^\infty e^{-ut} (t^2 - 1)^{v-1/2} dt, \quad u > 0, v \geq -1/2. \quad (2.3)$$

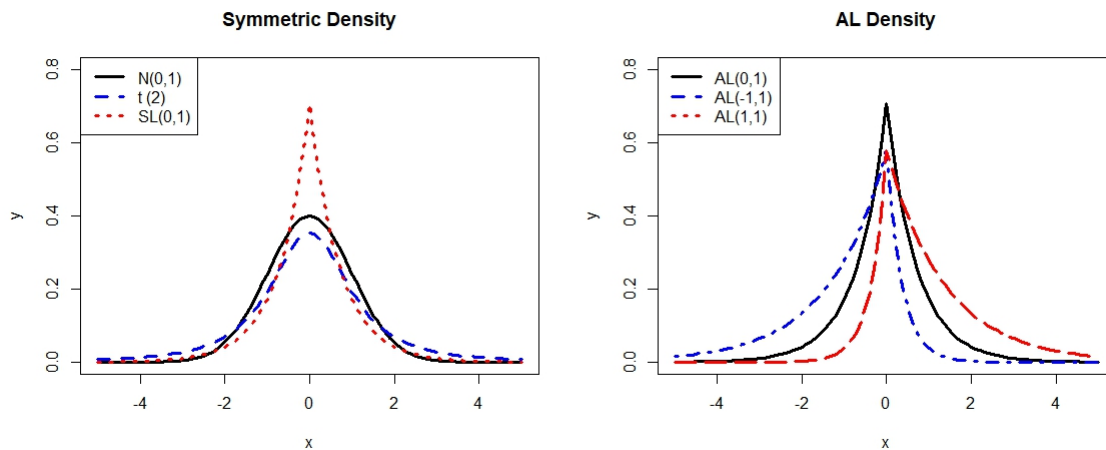
When  $\boldsymbol{\mu} = \mathbf{0}_n$ , we can obtain Symmetric Laplace distribution  $SL(\boldsymbol{\Sigma})$  with density

$$f(\mathbf{x}) = 2(2\pi)^{-n/2} |\boldsymbol{\Sigma}|^{-1/2} (\mathbf{x}'\boldsymbol{\Sigma}^{-1}\mathbf{x}/2)^{v/2} K_v(\sqrt{2\mathbf{x}'\boldsymbol{\Sigma}^{-1}\mathbf{x}}).$$

When  $n = 1$ , we have  $\boldsymbol{\Sigma} = \sigma_{11} = \sigma$ . In such cases, (2.1) becomes the univariate Laplace distribution  $AL_1(\mu, \sigma)$  distribution with parameters  $\mu$  and  $\sigma$ . The corresponding density function is

$$f(x) = \frac{1}{\gamma} \exp\left\{-\frac{|x|}{\sigma^2} [\gamma - \mu \cdot \text{sign}(x)]\right\} \quad \text{with} \quad \gamma = \sqrt{\mu^2 + 2\sigma^2}. \quad (2.4)$$

The symmetric case ( $\mu = 0$ ) leads to the univariate Laplace distribution  $SL_1(0, \sigma)$ .



**Figure 1.** Univariate densities.

Figure 1 displays plot of symmetric densities and AL densities. Symmetric densities including standard normal distribution, student  $t$  distribution with 2 degrees of freedom, and univariate symmetric Laplace distribution, denoted as  $N(0, 1)$ ,  $t(2)$ ,  $SL_1(0, 1)$ . The student  $t$  distribution possesses heavier tail than normal distribution, whereas  $SL_1(0, 1)$  distribution imposes greater peakedness and heavier tail than normal case. As for plots of AL densities, when  $\mu > 0$ , the density skews to the right. On the other hand, when  $\mu < 0$ , the density skews to the left.

Important results of univariate and multivariate asymptotic Laplace distributions that will be used later on are presented below.

**Proposition 2.1.** (See Kotz, 2001)

- (1). If  $\mathbf{X} = (X_1, \dots, X_n)$  follows multivariate Asymmetric Laplace distribution, i.e.,  $\mathbf{X} \sim AL_n(\boldsymbol{\mu}, \boldsymbol{\Sigma})$ ,  $n$  is the number of securities. The linear combination  $\mathbf{w}'\mathbf{X} = w_1X_1 + \dots + w_nX_n$  follows univariate Asymmetric Laplace distribution, i.e.  $\mathbf{w}'\mathbf{X} \sim AL_1(\mu, \sigma)$ , with  $\mu = \mathbf{w}'\boldsymbol{\mu}$ ,  $\sigma = \sqrt{\mathbf{w}'\boldsymbol{\Sigma}\mathbf{w}}$ ,  $\mathbf{w} = (w_1, \dots, w_n)'$ .
- (2). Assume that univariate random variable  $Y \sim AL_1(\mu, \sigma)$ . To measure the asymmetry and peakedness of the distribution, define the skewness ( $\text{Skew}[Y]$ ) and kurtosis ( $\text{Kurt}[Y]$ ) as the third and fourth standardized moment of a random variable  $Y$ . Then,

$$\text{Skew}[Y] = \frac{\mathbb{E}(Y - \mathbb{E}Y)^3}{[\mathbb{E}(Y - \mathbb{E}Y)^2]^{3/2}} = \frac{2\mu^3 + 3\mu\sigma^2}{(\mu^2 + \sigma^2)^{3/2}},$$

$$\text{Kurt}[Y] = \frac{\mathbb{E}(Y - \mathbb{E}Y)^4}{[\text{Var}(Y)]^2} = \frac{9\mu^4 + 6\sigma^4 + 18\mu^2\sigma^2}{(\mu^2 + \sigma^2)^2}.$$

- (3). Let  $\mathbf{X} = (X_1, X_2, \dots, X_n) \sim AL_n(\boldsymbol{\mu}, \boldsymbol{\Sigma})$ . Then the first and second order moments of  $\mathbf{X}$  are

$$\mathbb{E}(\mathbf{X}) = \boldsymbol{\mu} \quad \text{and} \quad \text{Cov}(\mathbf{X}) = \boldsymbol{\Sigma} + \boldsymbol{\mu}'\boldsymbol{\mu}.$$

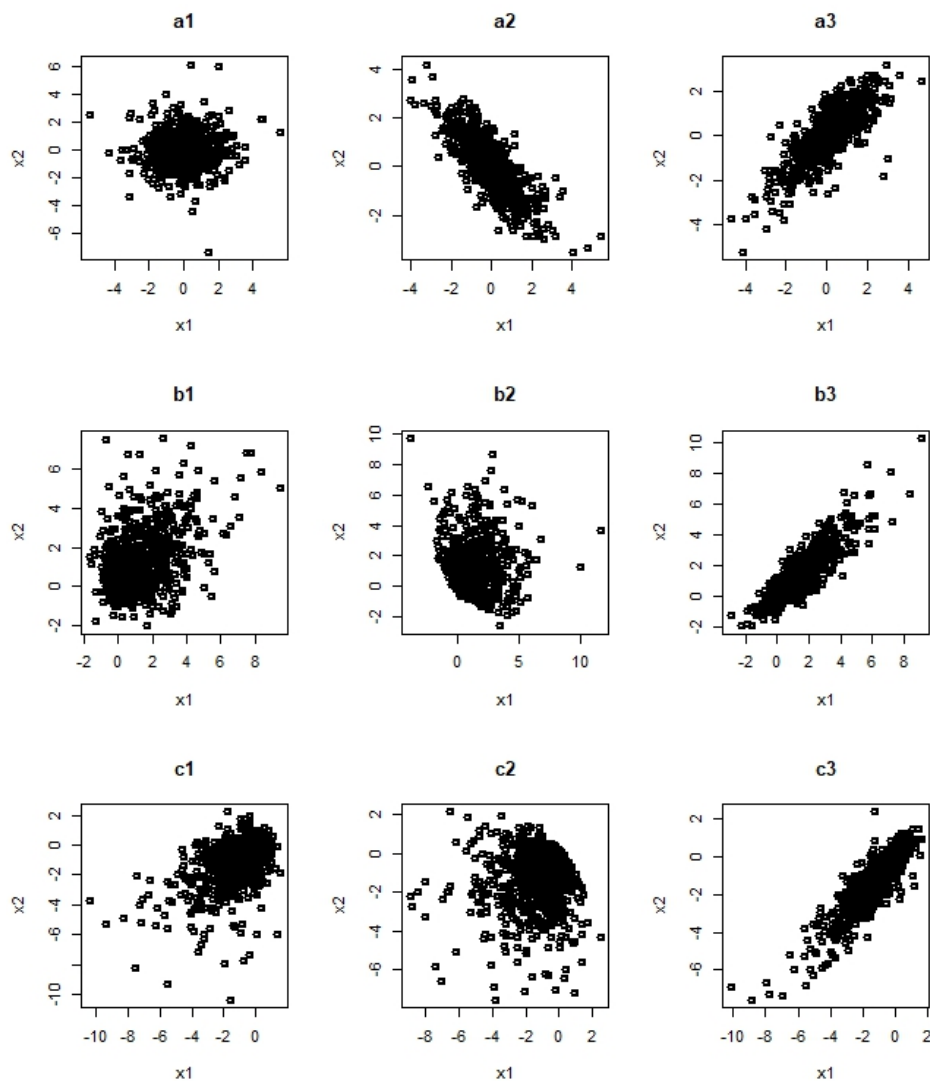
- (4). The Asymmetric Laplace distribution can be represented as a mixture of normal vector and a standard exponential variable, i.e.,  $\mathbf{X} \sim AL_n(\boldsymbol{\mu}, \boldsymbol{\Sigma})$  can be represented as

$$\mathbf{X} = \boldsymbol{\mu}Z + Z^{1/2}\mathbf{Y},$$

where  $Y \sim N_n(\mathbf{0}, \Sigma)$ ,  $Z \sim \text{Exp}(1)$ . This indicates that we can simulate multivariate Asymmetric Laplace random vector  $X \sim AL_n(\mu, \Sigma)$  as follows:

1. Generate a multivariate normal variable  $Y \sim N_n(\mathbf{0}, \Sigma)$ ;
2. Generate a standard exponential variable  $Z \sim \text{Exp}(1)$ ;
3. Construct Asymmetric Laplace random vector as  $X = \mu Z + Z^{1/2}Y$ .

Figure 2 displays several realizations of bivariate Asymmetric Laplace distribution with different levels of asymmetry and peakedness.



**Figure 2.** Bivariate Asymmetric Laplace data with  $\mu$  cases: (a1,a2,a3):  $\mu = (0, 0)$ ; (b1,b2,b3):  $\mu = (1, 1)$ ; (c1,c2,c3):  $\mu = (-1, -1)$ . Covariance matrix  $\Sigma$  cases. (a1,b1,c1):  $\sigma_{11} = \sigma_{22} = 1, \sigma_{12} = \sigma_{21} = 0$ ; (a2,b2,c2):  $\sigma_{11} = \sigma_{22} = 1, \sigma_{12} = \sigma_{21} = 0.8$ ; (a3,b3,c3):  $\sigma_{11} = \sigma_{22} = 1, \sigma_{12} = \sigma_{21} = -0.8$ .

## 2.2. Risk measures

Since mean and covariance matrix cannot be used to characterize non-Gaussian distribution, alternative risk measures are necessary for portfolio selection problems. Artzner et al. (1999) suggests that a desirable risk measure should be defined fulfilling certain properties and such a risk measure is said to be coherent.

A risk measure  $\phi$  that maps a random variable to a real number is coherent if it satisfies the following conditions:

- 1). Translation invariance:  $\phi(l + h) = \phi(l) + h$ , for all random losses  $l$  and all  $h \in \mathbb{R}$ ;
- 2). Subadditivity:  $\phi(l + h) \leq \phi(l) + \phi(h)$ , for all random losses  $l, h$ ;
- 3). Positive homogeneity:  $\phi(\lambda l) = \lambda\phi(l)$  for all random losses  $l$  and all  $\lambda > 0$ ;
- 4). Monotonicity:  $\phi(l_1) \leq \phi(l_2)$  for all random losses  $l_1, l_2$  with  $l_1 \leq l_2$  almost surely.

Standard deviation is not coherent in general excepting the Gaussian cases. VaR is coherent when the underlying distribution is elliptically distributed. Expected Shortfall, or the so-called conditional value at risk (CVaR) is a coherent risk measure since it always satisfies subadditivity, monotonicity, positive homogeneity, and convexity. For any fixed  $\alpha \in (0, 1)$ ,  $\alpha$ -VaR is the  $\alpha$ -quantile loss while  $\alpha$ -ES is the average of all  $\beta$ -VaR for  $\beta \in (\alpha, 1)$ . Both VaR and CVaR measure the potential maximal loss. VaR and ES can be written as

$$\text{VaR}_\alpha = F^{-1}(\alpha) \quad \text{and} \quad \text{ES}_\alpha = E[L|L \leq -\text{VaR}_\alpha] = -\frac{1}{\alpha} \int_{-\infty}^{-\text{VaR}_\alpha} \text{VaR}_\beta d\beta,$$

where  $F(\cdot)$  is the cumulative distribution function of loss  $L$  and  $\text{ES}_\alpha$  is the expected loss above  $\text{VaR}_\alpha$ . Thus, the estimation process are

$$\int_{-\infty}^{-\text{VaR}_\alpha} f_X(x) dx = \alpha \quad \text{and} \quad \text{ES}_\alpha = -\frac{1}{\alpha} \int_{-\infty}^{-\text{VaR}_\alpha} x f(x) dx. \quad (2.5)$$

Under normality assumption,  $\text{VaR}_\alpha$  and  $\text{ES}_\alpha$  are

$$\begin{aligned} \text{VaR}_\alpha &= \mu + \sigma \Phi^{-1}(1 - \alpha), \\ \text{ES}_\alpha &= \mu + \sigma \frac{\psi(\Phi^{-1}(1 - \alpha))}{\alpha}. \end{aligned}$$

where  $\psi(\cdot)$  as the normal density distribution, and  $\Phi^{-1}(\cdot)$  is the quantile distribution. As shown in Hu et al. (2010), portfolio selected by minimizing standard deviation,  $\text{VaR}_\alpha$ , and  $\text{ES}_\alpha$  are the equivalent under elliptical distribution assumption.

It is well-documented that asset securities are not normally distributed. As an alternative to Gaussian distribution, Asymmetric Laplace distribution exhibits tail-heaviness, skewness, and peakedness.

## 3. Portfolio selection under ALD framework

Let  $\mathbf{X} = (X_1, X_2, \dots, X_n) \sim AL_n(\boldsymbol{\mu}_\varphi, \boldsymbol{\Sigma}_\varphi)$  be the return vectors of  $n$  securities, and  $\mathbf{w} = (w_1, w_2, \dots, w_n)'$  is the allocation weight vector. Then, the portfolio is defined as

$$\mathcal{P}(\mathbf{w}) = \mathbf{w}'\mathbf{X} = \sum_{i=1}^n w_i X_i.$$

According to Proposition 2.1 (2),  $\mathcal{P}(\mathbf{w}) \sim AL_1(\mu, \sigma)$  with  $\mu = \mathbf{w}'\boldsymbol{\mu}$ ,  $\sigma = \sqrt{\mathbf{w}'\boldsymbol{\Sigma}\mathbf{w}}$ .

From Theorem 3.1–3.2 below, in order to select a portfolio under Asymmetric Laplace distribution, it suffices to obtain the unknown parameters  $\boldsymbol{\mu}_{\mathcal{P}}$  and  $\boldsymbol{\Sigma}_{\mathcal{P}}$ . Thus portfolio selection models under Asymmetric Laplace distribution lead to parameter estimation for  $AL_n(\boldsymbol{\mu}_{\mathcal{P}}, \boldsymbol{\Sigma}_{\mathcal{P}})$ . Zhao et al. (2015) proposed the multi-objective portfolio selection model under Asymmetric Laplace framework and derived the simplified model that can be reformulated as quadratic programming problem. However, to estimate the unknown parameters, the authors adopt a moment estimation method that is less efficient compared to maximum likelihood method. Since Asymmetric Laplace distribution can be represented as a mixture of exponential distribution and multivariate normal distribution, we derived the Expectation-Maximization algorithm for parameter estimation of Asymmetric Laplace distribution. The algorithm for estimating these unknown parameters is discussed in Section 3.2.

### 3.1. Portfolio selection theorems

**Theorem 3.1.** Let  $\mathbf{X} = (X_1, \dots, X_n) \sim AL_n(\boldsymbol{\mu}_{\mathcal{P}}, \boldsymbol{\Sigma}_{\mathcal{P}})$  be a  $n$ -dimensional random vector that follow multivariate Asymmetric Laplace distribution, each element  $(X_i, i = 1, 2, \dots, n)$  represent a stock. Let  $\mathbf{w}$  be the weight vector and  $\mathcal{P}(\mathbf{w}) = \mathbf{w}'\mathbf{X} = \sum_{i=1}^n w_i X_i$  be the portfolio. Then, under Asymmetric Laplace framework, risk measures of StD,  $\text{VaR}_{\alpha}$ , and  $\text{ES}_{\alpha}$  at  $\alpha \in (0, 1)$  level formulated as

$$\begin{aligned} \text{Standard Deviation:} \quad \text{StD}(\mathcal{P}(\mathbf{w})) &= \frac{\sigma}{\sqrt{2}}; \\ \text{Value at Risk:} \quad \text{VaR}_{\alpha}(\mathcal{P}(\mathbf{w})) &= -\frac{\sigma^2}{\gamma + \mu} \ln \frac{\alpha\gamma(\gamma + \mu)}{\sigma^2}; \\ \text{Expected Shortfall:} \quad \text{ES}_{\alpha}(\mathcal{P}(\mathbf{w})) &= \frac{\sigma^2}{\gamma + \mu} - \frac{\sigma^2}{\gamma + \mu} \ln \frac{\alpha\gamma(\gamma + \mu)}{\sigma^2}. \end{aligned}$$

Here,  $\mu = \mathbf{w}'\boldsymbol{\mu}_{\mathcal{P}} = \text{mean}(\mathcal{P}(\mathbf{w}))$ ,  $\sigma = \sqrt{\mathbf{w}'\boldsymbol{\Sigma}_{\mathcal{P}}\mathbf{w}} = \text{std}(\mathcal{P}(\mathbf{w}))$  and  $\gamma = \sqrt{\mu^2 + 2\sigma^2}$ .

**Proof.** Let  $\boldsymbol{\mu}_{\mathcal{P}} = (\mu_1, \dots, \mu_n)$  be the mean return vector of the securities  $(X_1, \dots, X_n)$  and  $\boldsymbol{\Sigma}_{\mathcal{P}} = (\sigma_{\mathcal{P}})_{i,j=1}^n$  be the scale matrix of  $(X_1, \dots, X_n)$ . Denote the allocation vector by  $\mathbf{w} = (w_1, \dots, w_n)'$ . Then, the portfolio  $\mathcal{P}(\mathbf{w}) = \sum_{i=1}^n w_i X_i$  follows univariate Asymmetric Laplace distribution with

$$\mathcal{P}(\mathbf{w}) = \sum_{i=1}^n w_i X_i \sim AL_1(\mu, \sigma) \quad \text{with} \quad \mu = \sum_{i=1}^n \mu_i w_i, \sigma = \left( \sum_{i=1}^n \sum_{j=1}^n \sigma_{\mathcal{P}_{ij}} w_i w_j \right)^{1/2}.$$

If  $\boldsymbol{\mu} = \mathbf{0}_n$ , the univariate symmetric Asymmetric Laplace distribution becomes  $AL_1(0, \sigma)$  with density

$$g(x) = \frac{1}{\gamma} \exp\left\{-\frac{|x|}{\sigma^2}\gamma\right\} \quad \text{with} \quad \gamma = \sqrt{2}\sigma.$$

Thus, standard deviation (StD) of portfolio  $\mathcal{P}(\mathbf{w}) = \mathbf{w}'\mathbf{X}$  is

$$\text{StD}(\mathcal{P}(\mathbf{w})) = \int_{-\infty}^{+\infty} \frac{1}{\gamma} |x| \exp\left\{-\frac{\gamma}{\sigma^2}|x|\right\} dx = 2 \int_0^{+\infty} \frac{x}{\gamma} \exp\left\{-\frac{\gamma}{\sigma^2}x\right\} dx = \frac{2\sigma^4}{\gamma^3} = \frac{\sigma}{\sqrt{2}}.$$

According to the definition of  $\text{VaR}_{\alpha}$  and  $\text{ES}_{\alpha}$  as defined in (2.5) and univariate Asymmetric Laplace density (2.4), we have

$$-\int_{-\infty}^{-\text{VaR}_{\alpha}} \frac{1}{\gamma} \exp\left\{-\frac{|x|}{\sigma^2}[\gamma - \mu \cdot \text{sgn}(x)]\right\} dx = \alpha,$$

$$\frac{\sigma^2}{\gamma(\gamma + \mu)} \exp\left\{-\frac{\gamma + \mu}{\sigma^2} \text{VaR}_\alpha\right\} = \alpha.$$

Thus,  $\text{VaR}_\alpha$  and  $\text{ES}_\alpha$  are

$$\begin{aligned} \text{VaR}_\alpha(\mathcal{P}(\mathbf{w})) &= -\frac{\sigma^2}{\sqrt{\mu^2 + 2\sigma^2} + \mu} \ln \frac{\alpha(\mu^2 + 2\sigma^2 + \mu\sqrt{\mu^2 + 2\sigma^2})}{\sigma^2} = -\frac{\sigma^2}{\gamma + \mu} \ln \frac{\alpha\gamma(\gamma + \mu)}{\sigma^2}; \\ \text{ES}_\alpha(\mathcal{P}(\mathbf{w})) &= -\frac{1}{\alpha} \int_{-\infty}^{-\text{VaR}_\alpha} x f_X(x) dx = -\frac{1}{\alpha} \int_{-\infty}^{-\text{VaR}_\alpha} x \frac{1}{\gamma} \exp\left\{-\frac{|x|}{\sigma^2} [\gamma - \mu \cdot \text{sgn}(x)]\right\} dx \\ &= \frac{\sigma^2}{\mu + \sqrt{\mu^2 + 2\sigma^2}} - \frac{\sigma^2}{\mu + \sqrt{\mu^2 + 2\sigma^2}} \ln \left\{2\alpha + \frac{\alpha(\mu^2 + \mu\sqrt{\mu^2 + 2\sigma^2})}{\sigma^2}\right\} \\ &= \frac{\sigma^2}{\gamma + \mu} - \frac{\sigma^2}{\gamma + \mu} \ln \frac{\alpha\gamma(\gamma + \mu)}{\sigma^2}. \end{aligned}$$

□

Then we have the following theorem.

**Theorem 3.2.** Let  $X \sim AL_n(\boldsymbol{\mu}_\mathcal{P}, \boldsymbol{\Sigma}_\mathcal{P})$ . Then, portfolio  $\mathcal{P}(\mathbf{w}) = \mathbf{w}'X$  with following models based on  $\text{ES}_\alpha$ ,  $\text{VaR}_\alpha$ , and  $\text{StD}$  (as defined in Theorem 3.1)

$$\begin{aligned} &\min_{\mathbf{w}} \text{ES}_\alpha(\mathcal{P}(\mathbf{w})) \text{ or } \min_{\mathbf{w}} \text{VaR}_\alpha(\mathcal{P}(\mathbf{w})) \text{ or } \min_{\mathbf{w}} \text{StD}(\mathcal{P}(\mathbf{w})) \\ &\max_{\mathbf{w}} \text{Skew}[\mathcal{P}(\mathbf{w})] = \frac{2\mu^3 + 3\mu\sigma^2}{(\mu^2 + \sigma^2)^{3/2}} \\ &\max_{\mathbf{w}} \text{Kurt}[\mathcal{P}(\mathbf{w})] = \frac{9\mu^4 + 6\sigma^4 + 18\mu^2\sigma^2}{(\mu^2 + \sigma^2)^2} \\ &\text{s.t. } \mathbf{w}'\boldsymbol{\mu} = r_0, \mathbf{w}'\mathbf{1} = 1. \end{aligned}$$

are equivalent. Here,  $\mu = \mathbf{w}'\boldsymbol{\mu}_\mathcal{P} = \text{mean}[\mathcal{P}(\mathbf{w})]$ ,  $\sigma = \sqrt{\mathbf{w}'\boldsymbol{\Sigma}_\mathcal{P}\mathbf{w}} = \text{std}[\mathcal{P}(\mathbf{w})]$ ,  $\mathbf{w} = (w_1, w_2, \dots, w_n)'$ .

**Proof.** Let  $g(\mu, \sigma) = \frac{\sigma^2}{\mu + \sqrt{\mu^2 + 2\sigma^2}}$ . Then,  $\text{ES}_\alpha[\mathcal{P}(\mathbf{w})]$  and  $\text{VaR}_\alpha[\mathcal{P}(\mathbf{w})]$  are

$$\begin{aligned} \text{VaR}_\alpha[\mathcal{P}(\mathbf{w})] &= -g(\mu, \sigma) \ln\left(2\alpha + \frac{\alpha\mu}{g(\mu, \sigma)}\right) = -g(\mu, \sigma) \left[\ln \alpha + \ln\left(2 + \frac{\mu}{g(\mu, \sigma)}\right)\right], \\ \text{ES}_\alpha[\mathcal{P}(\mathbf{w})] &= g(\mu, \sigma) - g(\mu, \sigma) \ln\left(2\alpha + \frac{\alpha\mu}{g(\mu, \sigma)}\right) = (1 - \ln \alpha)g(\mu, \sigma) - g(\mu, \sigma) \ln\left(2 + \frac{\mu}{g(\mu, \sigma)}\right). \end{aligned}$$

Differentiating the above expressions with respect to  $\sigma$ , we have

$$\begin{aligned} \frac{\partial \text{VaR}_\alpha[\mathcal{P}(\mathbf{w})]}{\partial \sigma} &= \frac{\partial g(\mu, \sigma)}{\partial \sigma} \left[ -\ln \alpha - \ln\left(2 + \frac{\mu}{g(\mu, \sigma)}\right) + \frac{\frac{\mu}{g(\mu, \sigma)}}{2 + \frac{\mu}{g(\mu, \sigma)}} \right] > 0, \\ \frac{\partial \text{ES}_\alpha[\mathcal{P}(\mathbf{w})]}{\partial \sigma} &= \frac{\partial g(\mu, \sigma)}{\partial \sigma} \left[ 1 - \ln \alpha - \ln\left(2 + \frac{\mu}{g(\mu, \sigma)}\right) + \frac{\frac{\mu}{g(\mu, \sigma)}}{2 + \frac{\mu}{g(\mu, \sigma)}} \right] > 0, \end{aligned}$$



where

$$\frac{\partial g(\mu, \sigma)}{\partial \sigma} = \frac{\partial \left[ \frac{\sigma^2}{\mu + \sqrt{\mu^2 + 2\sigma^2}} \right]}{\partial \sigma} = 2\sigma \frac{\mu + \frac{\mu^2 + \sigma^2}{\sqrt{\mu^2 + 2\sigma^2}}}{(\mu^2 + \sqrt{\mu^2 + 2\sigma^2})^2} > 0.$$

The derivative of skewness measure with respect to  $\sigma$  is

$$\frac{\partial \text{Skew}[\mathcal{P}(\mathbf{w})]}{\partial \sigma} = \frac{\partial \left[ \frac{2\mu^3 + 3\mu\sigma^2}{(\mu^2 + \sigma^2)^{3/2}} \right]}{\partial \sigma} = \frac{-3\mu\sigma^3}{(\mu^2 + \sigma^2)^{5/2}} < 0.$$

The derivative of kurtosis measure with respect to  $\sigma$  is

$$\frac{\partial \text{Kurt}[\mathcal{P}(\mathbf{w})]}{\partial \sigma} = \frac{\frac{9\mu^4 + 6\sigma^4 + 18\mu^2\sigma^2}{(\mu^2 + \sigma^2)^2}}{\partial \sigma} = \frac{-12\mu^4\sigma^3 - 12\mu^2\sigma^5}{(\mu^2 + \sigma^2)^4} < 0.$$

The monotonicity of  $\text{VaR}_\alpha[\mathcal{P}(\mathbf{w})]$ ,  $\text{ES}_\alpha[\mathcal{P}(\mathbf{w})]$ ,  $\text{Skew}[\mathcal{P}(\mathbf{w})]$ , and  $\text{Kurt}[\mathcal{P}(\mathbf{w})]$  with respect to  $\sigma$  indicate that the portfolio selection problems based on these risk measures are equivalent. This means that minimizing  $\text{VaR}_\alpha[\mathcal{P}(\mathbf{w})]$ ,  $\text{ES}_\alpha[\mathcal{P}(\mathbf{w})]$ ,  $\text{StD}[\mathcal{P}(\mathbf{w})]$  are equivalent to minimizing  $\mathbf{w}'\Sigma\varphi\mathbf{w}$ .  $\square$

### 3.2. Parameter estimation of Asymmetric Laplace distribution

Assume  $\mathbf{X} = (X_1, X_2, \dots, X_n) \sim \text{AL}_n(\boldsymbol{\mu}, \boldsymbol{\Sigma})$ . Let  $\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_T \in \mathbb{R}^n$  be the  $T$  observations. We aim at fitting a multivariate Asymmetric Laplace distribution  $\mathcal{AL}_n(\boldsymbol{\mu}, \boldsymbol{\Sigma})$  with unknown parameters  $\boldsymbol{\mu}, \boldsymbol{\Sigma}$ .

Hrilmann (2013), Kollo and Srivastava (2005), Visk (2009) consider moment matching methods that is less efficient than maximum likelihood estimation. Kotz et al. (2002) and Kotz et al. (2001) presented the maximum likelihood estimators for parameter estimation of Asymmetric Laplace distributions. However, maximum likelihood estimation require computation of complicated Bessel function. Thus we derived the expectation-maximization algorithm for parameter estimation of Asymmetric Laplace distribution.

#### 3.2.1. Moment estimation

As Zhao et al. (2015) pointed out, according to Proposition 2.1 (3), Asymmetric Laplace distribution can be estimated via moment method (Moment-AL) with

$$\hat{\boldsymbol{\mu}} = \bar{\mathbf{x}} \quad \text{and} \quad \hat{\boldsymbol{\Sigma}} = \text{cov}(\mathbf{X}) - \hat{\boldsymbol{\mu}}'\hat{\boldsymbol{\mu}},$$

where  $\bar{\mathbf{x}} = \frac{1}{n} \sum_{i=1}^n \mathbf{x}_i$ ,  $\text{Cov}(\mathbf{X}) = \sum_{i=1}^n (\mathbf{x}_i - \bar{\mathbf{x}})^T (\mathbf{x}_i - \bar{\mathbf{x}})$ .

#### 3.2.2. Maximum likelihood estimation

Consider sample points  $\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_n$  and density function of Asymmetric Laplace distribution as defined in (2.1). Taken logarithm with respect to likelihood function, the log-likelihood is

$$\begin{aligned} \ell(\boldsymbol{\mu}, \boldsymbol{\Sigma}) &= \ln L(\boldsymbol{\mu}, \boldsymbol{\Sigma}) = \sum_{i=1}^T \ln f(\mathbf{x}_i; \boldsymbol{\mu}, \boldsymbol{\Sigma}) \\ &= \sum_{i=1}^T \mathbf{x}_i' \boldsymbol{\Sigma}^{-1} \boldsymbol{\mu} + T \ln 2 - \frac{Tn}{2} \ln(2\pi) - \frac{T}{2} \ln(|\boldsymbol{\Sigma}|) + \frac{\nu}{2} \sum_{i=1}^T \ln(\mathbf{x}_i' \boldsymbol{\Sigma}^{-1} \mathbf{x}_i) - \end{aligned}$$

$$\frac{\nu T}{2} \ln(2 + \boldsymbol{\mu}' \boldsymbol{\Sigma}^{-1} \boldsymbol{\mu}) + \sum_{t=1}^T \ln K_{\nu} \left\{ \sqrt{(2 + \boldsymbol{\mu}' \boldsymbol{\Sigma}^{-1} \boldsymbol{\mu})(\mathbf{x}_t' \boldsymbol{\Sigma}^{-1} \mathbf{x}_t)} \right\}.$$

Generally, we can directly maximize the log-likelihood function  $\ell(\boldsymbol{\mu}, \boldsymbol{\Sigma})$  with respect to parameters  $\boldsymbol{\mu}, \boldsymbol{\Sigma}$  and thus obtain the maximum likelihood estimator. Unfortunately, the density function involves modified Bessel function of the third kind with density (2.2), (2.3) that are too complex and complicated for numerical maximization. However, Gaussian-Exponential mixture representation of the Asymmetric Laplace distribution allows us to employ the expectation-maximization algorithm without involving modified Bessel functions.

### 3.2.3. Expectation-maximization algorithm

Then we derive the Expectation-Maximization algorithm for parameter estimation of multivariate Asymmetric Laplace Distribution (mALD), we follow the EM derivation for Multivariate Skew Laplace distribution in Arslan (2010).

Let  $\mathbf{X} = (X_1, X_2, \dots, X_n)$  be Asymmetric Laplace distributed random vector. Proposition 2.1 suggests that  $\mathbf{X}$  can be generated from a latent random variable  $Z = z$  through multivariate Gaussian distribution with  $z\boldsymbol{\mu}, z\boldsymbol{\Sigma}$ , i.e.  $X|Z = z \sim N_n(z\boldsymbol{\mu}, z\boldsymbol{\Sigma}), Z \sim \text{Exp}(1)$  with density

$$\begin{aligned} f_{X|Z}(\mathbf{x}, z) &= \frac{1}{(2\pi)^{n/2} |z\boldsymbol{\Sigma}|^{1/2}} \exp \left\{ -\frac{1}{2} (\mathbf{x} - z\boldsymbol{\mu})' (z\boldsymbol{\Sigma})^{-1} (\mathbf{x} - z\boldsymbol{\mu}) \right\}, \\ f_Z(z) &= e^{-z} 1_{\{z \geq 0\}}. \end{aligned}$$

Thus the joint density function of  $\mathbf{X}$  and  $Z$  is

$$f_{X,Z}(\mathbf{x}, z) = f_{X|Z}(\mathbf{x}, z) f_Z(z) = \frac{1}{(2\pi)^{\frac{n}{2}} z^{\frac{n}{2}} |\boldsymbol{\Sigma}|^{\frac{1}{2}}} \exp \left\{ -\frac{1}{2z} \mathbf{x}' \boldsymbol{\Sigma}^{-1} \mathbf{x} + \mathbf{x}' \boldsymbol{\Sigma}^{-1} \boldsymbol{\mu} - \frac{z}{2} \boldsymbol{\mu}' \boldsymbol{\Sigma}^{-1} \boldsymbol{\mu} - z 1_{\{z \geq 0\}} \right\}.$$

Suppose that there are  $T$  observations  $\mathbf{X}_1, \dots, \mathbf{X}_T$  generated from the latent random variables  $z_1, z_2, \dots, z_T$  respectively. The complete data is defined as  $\{(\mathbf{x}_t, z_t)\}, t = 1, 2, \dots, T$ . In the EM algorithm,  $\mathbf{x}_t$  and  $z_t$  are the observed and missing data respectively. The log-likelihood up to an additive constant can be written as

$$\tilde{L}(\boldsymbol{\mu}, \boldsymbol{\Sigma}) = \sum_{t=1}^T \ln f_{X,Z}(\mathbf{x}_t, z_t) = -\frac{T}{2} \ln |\boldsymbol{\Sigma}| - \frac{1}{2} \sum_{t=1}^T \frac{1}{z_t} \mathbf{x}_t' \boldsymbol{\Sigma}^{-1} \mathbf{x}_t + \sum_{t=1}^T \mathbf{x}_t' \boldsymbol{\Sigma}^{-1} \boldsymbol{\mu} - \frac{1}{2} \boldsymbol{\mu}' \boldsymbol{\Sigma}^{-1} \boldsymbol{\mu} \sum_{t=1}^T z_t - \sum_{t=1}^T z_t 1_{\{z_t \geq 0\}}.$$

Note that the last term of the above equation does not contain any unknown parameters and thus is negligible. Then, the E-step becomes

$$E(\tilde{L}(\boldsymbol{\mu}, \boldsymbol{\Sigma}) | \mathbf{x}_t, \hat{\boldsymbol{\mu}}, \hat{\boldsymbol{\Sigma}}) \propto -\frac{T}{2} \ln |\boldsymbol{\Sigma}| + \sum_{t=1}^T \mathbf{x}_t' \boldsymbol{\Sigma}^{-1} \boldsymbol{\mu} - \frac{1}{2} \sum_{t=1}^T E(z_t^{-1} | \mathbf{x}_t, \hat{\boldsymbol{\mu}}, \hat{\boldsymbol{\Sigma}}) \mathbf{x}_t' \boldsymbol{\Sigma}^{-1} \mathbf{x}_t - \frac{1}{2} \boldsymbol{\mu}' \boldsymbol{\Sigma}^{-1} \boldsymbol{\mu} \sum_{t=1}^T E(z_t | \mathbf{x}_t, \hat{\boldsymbol{\mu}}, \hat{\boldsymbol{\Sigma}}).$$

where  $E(z_t | \mathbf{x}_t, \hat{\boldsymbol{\mu}}, \hat{\boldsymbol{\Sigma}})$  and  $E(z_t^{-1} | \mathbf{x}_t, \hat{\boldsymbol{\mu}}, \hat{\boldsymbol{\Sigma}})$  are the conditional expectations of  $z_t$  and  $z_t^{-1}$  given  $\mathbf{x}_t$  and the current estimates  $\hat{\boldsymbol{\mu}}, \hat{\boldsymbol{\Sigma}}$ .

To evaluate conditional expectations  $E(z_t^{-1}|\mathbf{x}_t, \hat{\boldsymbol{\mu}}, \hat{\boldsymbol{\Sigma}})$  and  $E(z_t|\mathbf{x}_t, \hat{\boldsymbol{\mu}}, \hat{\boldsymbol{\Sigma}})$ , we need the conditional density of  $Z$  given  $\mathbf{X}$ ,  $f_{Z|\mathbf{X}}$ . After some straightforward algebra, the conditional distribution of  $Z$  given  $\mathbf{X}$  is an inverse Gaussian distribution with density function

$$\begin{aligned} f_{Z|\mathbf{X}}(z|\mathbf{x}, \boldsymbol{\mu}, \boldsymbol{\Sigma}) &= \frac{f_{\mathbf{X},Z}(\mathbf{x}, z)}{f_{\mathbf{X}}(\mathbf{x})} = \frac{\frac{1}{(2\pi)^{\frac{n}{2}} z^{\frac{n}{2}} |\boldsymbol{\Sigma}|^{\frac{1}{2}}} \exp\left\{-\frac{1}{2z} \mathbf{x}' \boldsymbol{\Sigma}^{-1} \mathbf{x} + \mathbf{x}' \boldsymbol{\Sigma}^{-1} \boldsymbol{\mu} - \frac{z}{2} \boldsymbol{\mu}' \boldsymbol{\Sigma}^{-1} \boldsymbol{\mu} - z \mathbf{1}_{\{z \geq 0\}}\right\}}{\frac{2e^{\mathbf{x}' \boldsymbol{\Sigma}^{-1} \boldsymbol{\mu}}}{(2\pi)^{n/2} |\boldsymbol{\Sigma}|^{1/2}} \left(\frac{\mathbf{x}' \boldsymbol{\Sigma}^{-1} \mathbf{x}}{2 + \boldsymbol{\mu}' \boldsymbol{\Sigma}^{-1} \boldsymbol{\mu}}\right)^{n/2} K_n\left(\sqrt{(2 + \boldsymbol{\mu}' \boldsymbol{\Sigma}^{-1} \boldsymbol{\mu})(\mathbf{x}' \boldsymbol{\Sigma}^{-1} \mathbf{x})}\right)} \\ &= \left(\frac{2 + \boldsymbol{\mu}' \boldsymbol{\Sigma}^{-1} \boldsymbol{\mu}}{\mathbf{x}' \boldsymbol{\Sigma}^{-1} \mathbf{x}}\right)^{n/2} z^{-n/2} \frac{\exp\left\{-\frac{1}{2}\left[z^{-1} \mathbf{x}' \boldsymbol{\Sigma}^{-1} \mathbf{x} + z(\boldsymbol{\mu}' \boldsymbol{\Sigma}^{-1} \boldsymbol{\mu} + z \mathbf{1}_{\{z \geq 0\}})\right]\right\}}{2K_n\left(\sqrt{(2 + \boldsymbol{\mu}' \boldsymbol{\Sigma}^{-1} \boldsymbol{\mu})(\mathbf{x}' \boldsymbol{\Sigma}^{-1} \mathbf{x})}\right)}. \end{aligned} \quad (3.1)$$

**Lemma 3.1.** (GIG (Stacy, 1962)) A random variable  $X$  follows Generalized Inverse Gaussian distribution (denoted as  $X \sim N^-(\lambda, \chi, \psi)$ ) if its density function could be represented as

$$f(x) = \frac{\chi^{-\lambda} (\sqrt{\chi\psi})^\lambda}{2K_\lambda(\sqrt{\chi\psi})} x^{\lambda-1} \exp\left\{-\frac{1}{2}(\chi x^{-1} + \psi x)\right\}, \quad x > 0.$$

where  $K_\lambda$  denotes the third kind modified Bessel function, and the parameters satisfy

$$\begin{cases} \chi > 0, \psi \geq 0, & \text{if } \lambda < 0; \\ \chi > 0, \psi > 0, & \text{if } \lambda = 0; \\ \chi \geq 0, \psi > 0, & \text{if } \lambda > 0. \end{cases}$$

After some algebraic manipulations, it is easy to show that  $Z|\mathbf{X}$  follows Generalized Inverse Gaussian distribution:

$$Z|\mathbf{X} \sim N^-\left(\frac{2-n}{2}, \mathbf{x}' \boldsymbol{\Sigma}^{-1} \mathbf{x}, 2 + \boldsymbol{\mu}' \boldsymbol{\Sigma}^{-1} \boldsymbol{\mu}\right).$$

If  $\chi > 0, \psi > 0$ , the moments could be calculated through the following formulas:

$$\begin{aligned} E(X^\alpha) &= \left(\frac{\chi}{\psi}\right)^{\alpha/2} \frac{K_{\lambda+\alpha}(\sqrt{\chi\psi})}{K_\lambda(\sqrt{\chi\psi})}, \quad \alpha \in \mathbb{R}, \\ E(\ln X) &= \left.\frac{dE(X^\alpha)}{d\alpha}\right|_{\alpha=0}. \end{aligned}$$

Denote  $\chi = \mathbf{x}' \boldsymbol{\Sigma}^{-1} \mathbf{x}, \psi = 2 + \boldsymbol{\mu}' \boldsymbol{\Sigma}^{-1} \boldsymbol{\mu}$ . Then,  $Z|\mathbf{X} \sim N^-(\frac{2-n}{2}, \chi, \psi)$ . From the conditional density function of (3.1), we can obtain the conditional expectations with the following moment properties:

$$\begin{aligned} a_t &= E(z_t|\mathbf{x}_t, \hat{\boldsymbol{\mu}}, \hat{\boldsymbol{\Sigma}}) = \sqrt{\frac{\chi_t}{\psi} \frac{K_{\frac{n}{2}-2}(\sqrt{\chi_t\psi})}{K_{\frac{n}{2}-1}(\sqrt{\chi_t\psi})}}, \quad t = 1, 2, \dots, T; \\ b_t &= E(z_t^{-1}|\mathbf{x}_t, \hat{\boldsymbol{\mu}}, \hat{\boldsymbol{\Sigma}}) = \sqrt{\frac{\psi}{\chi_t} \frac{K_{\frac{n}{2}}(\sqrt{\chi_t\psi})}{K_{\frac{n}{2}-1}(\sqrt{\chi_t\psi})}}, \quad t = 1, 2, \dots, T. \end{aligned}$$

where  $\chi_t = \mathbf{x}_t' \boldsymbol{\Sigma}^{-1} \mathbf{x}_t, R(\lambda) = \frac{K_{\lambda+1}(x)}{K_\lambda(x)}$  is strictly decreasing in  $x$  with  $\lim_{x \rightarrow \infty} R_\lambda(x) = 1$  and  $\lim_{x \rightarrow 0^+} R_\lambda(x) = \infty$ . Thus,  $a_t > 0, b_t > 0, t = 1, 2, \dots, T$ .

Finally, if the conditional expectation  $E(z_t|\mathbf{x}_t, \hat{\boldsymbol{\mu}}, \hat{\boldsymbol{\Sigma}})$  and  $E(z_t^{-1}|\mathbf{x}_t, \hat{\boldsymbol{\mu}}, \hat{\boldsymbol{\Sigma}})$  are replaced by  $a_t$  and  $b_t$  respectively, the objective function becomes

$$Q(\boldsymbol{\mu}, \boldsymbol{\Sigma}|\mathbf{x}_t, \hat{\boldsymbol{\mu}}, \hat{\boldsymbol{\Sigma}}) = -\frac{T}{2} \ln |\boldsymbol{\Sigma}| + \sum_{t=1}^T \mathbf{x}_t' \boldsymbol{\Sigma}^{-1} \boldsymbol{\mu} - \frac{1}{2} \sum_{t=1}^T b_t \mathbf{x}_t' \boldsymbol{\Sigma}^{-1} \mathbf{x}_t - \frac{1}{2} \boldsymbol{\mu}' \boldsymbol{\Sigma}^{-1} \boldsymbol{\mu} \sum_{t=1}^T a_t. \quad (3.2)$$

Denote  $S = \Sigma^{-1}$ . The objective function (3.2) becomes

$$Q(\mu, \Sigma | x_t, \hat{\mu}, \hat{S}) = \frac{T}{2} \ln |S| + \sum_{t=1}^T x'_t S \mu - \frac{1}{2} \sum_{t=1}^T b_t x'_t S x_t - \frac{1}{2} \mu' S \mu \sum_{t=1}^T a_t. \quad (3.3)$$

Taking derivative of objective function (3.3) with respect to  $\mu, S$ , we obtain

$$\begin{aligned} \frac{\partial Q}{\partial \mu} &= \sum_{t=1}^T x'_t S - \sum_{t=1}^T a_t \mu' S = 0, \\ \frac{\partial Q}{\partial S} &= \frac{T}{2} S^{-1} - \frac{1}{2} \sum_{t=1}^T b_t x'_t x_t + \sum_{t=1}^T x'_t \mu - \frac{1}{2} \sum_{t=1}^T a_t \mu' \mu = 0. \end{aligned}$$

Substituting  $S$  by  $\Sigma$  and setting these derivatives to zero yield

$$\begin{aligned} \sum_{t=1}^T x'_t \Sigma^{-1} - \sum_{t=1}^T a_t \mu' \Sigma^{-1} &= \mathbf{0}, \\ \frac{T}{2} \Sigma - \frac{1}{2} \sum_{t=1}^T b_t x'_t x_t + \sum_{t=1}^T x'_t \mu - \frac{1}{2} \sum_{t=1}^T a_t \mu' \mu &= \mathbf{0}. \end{aligned}$$

Thus, maximization of the objective function  $Q(\mu, \Sigma | x_t, \hat{\mu}, \hat{\Sigma})$  can be achieved by the following iterative updating formulas:

$$\hat{\mu} = \frac{\bar{x}}{\bar{a}}; \quad \hat{\Sigma} = \frac{\overline{b_t x'_t x_t}}{\bar{a}} - \frac{\bar{x}' \bar{x}}{\bar{a}}.$$

where  $\bar{a}, \bar{b}$  stand for the average of  $\{a_t\}_{t=1}^T$  and  $\{b_t\}_{t=1}^T$  respectively and  $\bar{x}$  is the average of  $\{x_t\}_{t=1}^T$ . In what follows, we present the iterative reweighted Expectation-Maximization algorithm for parameter estimation of Asymmetric Laplace distribution.

---

**Algorithm 1** Iterative reweighting algorithm
 

---

1. Set iteration number  $k = 1$  and select initial estimates for parameters  $\boldsymbol{\mu}^{(0)}, \boldsymbol{\Sigma}^{(0)}$ .
2. (E-Step) At  $k$ -th iteration with current estimates  $\boldsymbol{\mu}^{(k)}, \boldsymbol{\Sigma}^{(k)}$ , define the corresponding log-likelihood as

$$l^{(k)} = \log \sum_{t=1}^T f(\mathbf{x}_t | \boldsymbol{\mu}^{(k)}, \boldsymbol{\Sigma}^{(k)}), \quad k = 1, 2, \dots$$

With notations  $\chi_t = \mathbf{x}_t' \boldsymbol{\Sigma}^{-1} \mathbf{x}_t$ ,  $\psi = 2 + \boldsymbol{\mu}' \boldsymbol{\Sigma}^{-1} \boldsymbol{\mu}$ , we can obtain iterative weights

$$a_t = E(z_t | \mathbf{x}_t, \hat{\boldsymbol{\mu}}, \hat{\boldsymbol{\Sigma}}) = \sqrt{\frac{\chi_t}{\psi} \frac{K_{2-\frac{n}{2}}(\sqrt{\chi_t \psi})}{K_{1-\frac{n}{2}}(\sqrt{\chi_t \psi})}}, \quad t = 1, 2, \dots, T;$$

$$b_t = E(z_t^{-1} | \mathbf{x}_t, \hat{\boldsymbol{\mu}}, \hat{\boldsymbol{\Sigma}}) = \sqrt{\frac{\psi}{\chi_t} \frac{K_{-\frac{n}{2}}(\sqrt{\chi_t \psi})}{K_{1-\frac{n}{2}}(\sqrt{\chi_t \psi})}}, \quad t = 1, 2, \dots, T.$$

3. (M-Step) Employ the following iteration formulas to calculate the new estimates  $\boldsymbol{\mu}^{(k+1)}, \boldsymbol{\Sigma}^{(k+1)}$  at  $(k + 1)$ -th iteration:

$$\boldsymbol{\mu}^{(k+1)} = \frac{\bar{\mathbf{x}}}{\bar{a}}, \quad \boldsymbol{\Sigma}^{(k+1)} = \overline{b_t \mathbf{x}_t' \mathbf{x}_t} - \frac{\bar{\mathbf{x}}' \bar{\mathbf{x}}}{\bar{a}}. \quad (3.4)$$

The log-likelihood at  $(k + 1)$ -th iteration becomes

$$l^{(k+1)} = \log \sum_{t=1}^T f(\mathbf{x}_t | \boldsymbol{\mu}^{(k+1)}, \boldsymbol{\Sigma}^{(k+1)}).$$

4. Repeat these iteration steps until convergence with criterion  $l^{(k+1)} - l^{(k)} < \varepsilon$ , where  $\varepsilon > 0$  is a small number that control the convergence precision, for convenience, we take  $\varepsilon = 1e^{-16}$ .
- 

#### 4. Simulation studies

To evaluate the performance of portfolio selection models and parameter estimation methods in Section 3, we generate 100 datasets from Gaussian distribution and Asymmetric Laplace distribution respectively. Each dataset consists of  $T = 200$  observations with the following parameter settings: Case (1):  $n = 3$ ,  $\boldsymbol{\mu} = (0.03, 0.06, 0.09)$ ; Case (2):  $n = 5$ ,  $\boldsymbol{\mu} = (0.01, 0.02, 0.06, 0.08, 0.09)$ ; Case (3):  $n = 10$ ,  $\boldsymbol{\mu} = (0.01, 0.02, 0.03, \dots, 0.10)$ . For each case, we set  $\boldsymbol{\Sigma} = \text{diag}(\boldsymbol{\mu}/10)$ . All the simulation studies are carried out on a PC with Intel Core i7 3.6 GHz processor under R platform.

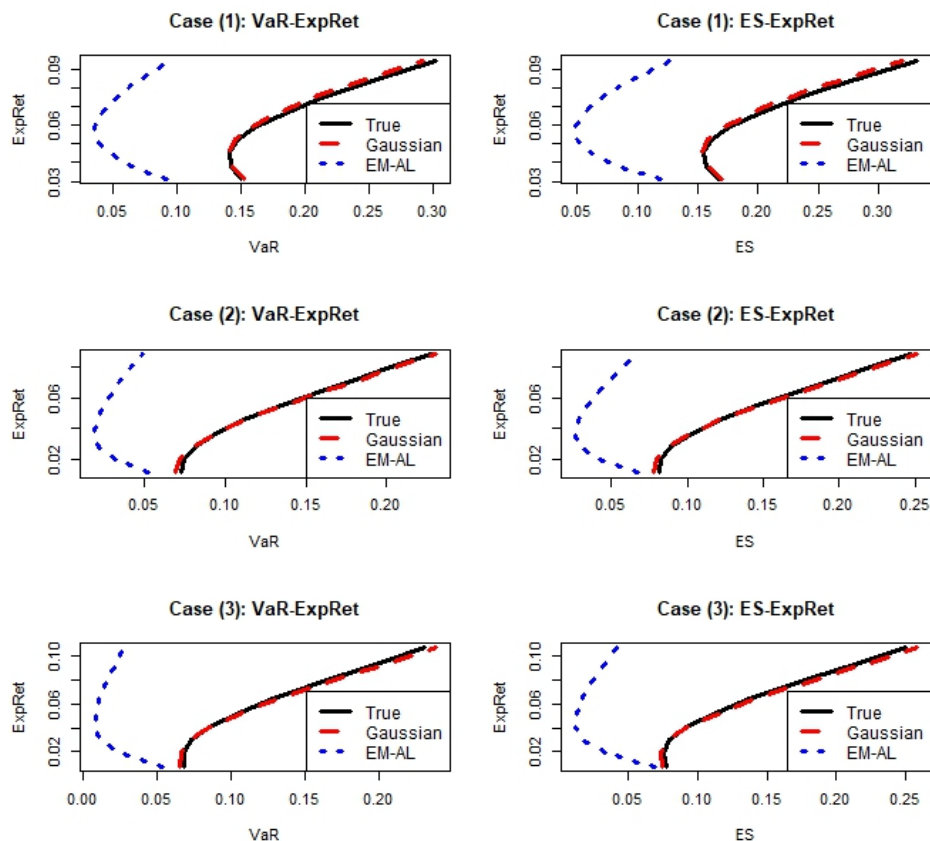
Each dataset are estimated under both multivariate Gaussian and Asymmetric Laplace distribution. ALD (EM-AL) is estimated using the EM algorithm described in Section 3.2. We evaluate the estimation performance using Bias measure, defined as  $\text{Bias} = \|\hat{\boldsymbol{\mu}} - \boldsymbol{\mu}\|_1 + \|\hat{\boldsymbol{\Sigma}} - \boldsymbol{\Sigma}\|_1$ . The mean log-likelihood and mean bias of the simulated 200 datasets are reported in Table 1.

Table 1 indicate that if the model is correctly specified, the estimation performance is always the best in terms of bias. If the data is generated from Gaussian distribution, the estimation from Gaussian model is the best, so does Asymmetric Laplace distribution. If data is generated from Gaussian distribution, then the estimation log-likelihood of Gaussian model is larger than Asymmetric Laplace

distribution, this is true for Asymmetric Laplace data as well.

**Table 1.** Model fitting results of Gaussian data and Asymmetric Laplace data using Gauss Model and EM-AL Model.

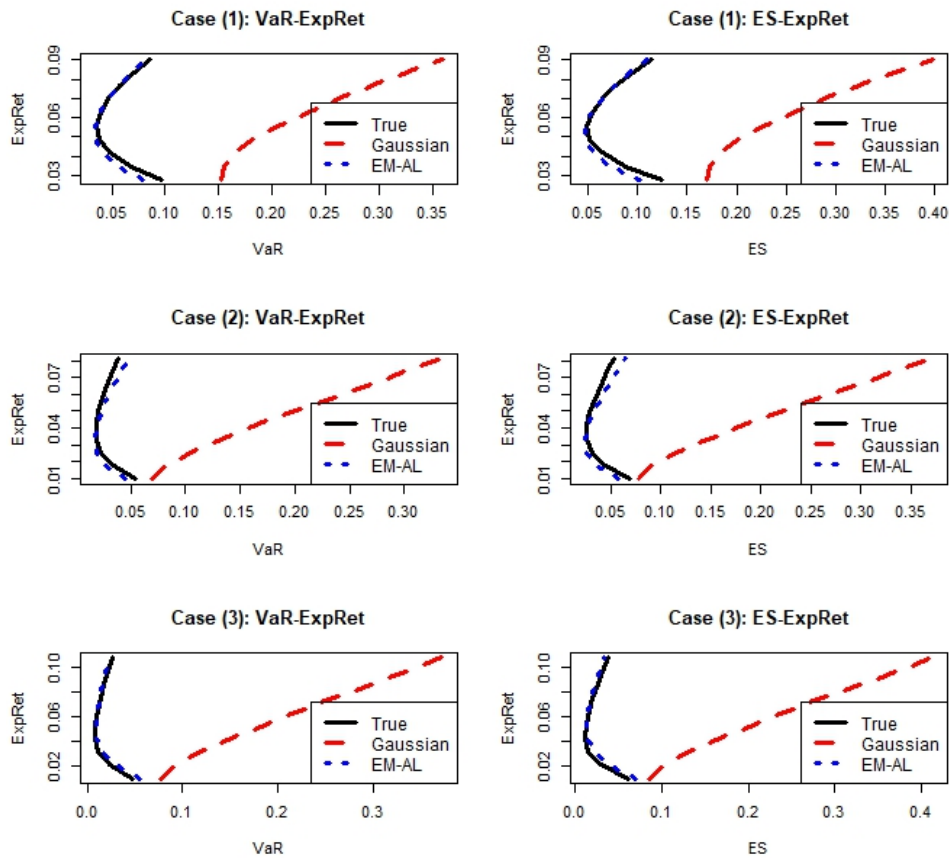
Gaussian Data					
	Log-Likelihood		Bias		
	Gauss	EM-AL	Gauss	Moment-AL	EM-AL
Case (1)	709.6159	641.2472	0.0153	0.0451	0.0183
Case (2)	1363.0231	1260.3676	0.0280	0.0886	0.0319
Case (3)	2593.1151	2432.2192	0.0672	0.3309	0.0766
Asymmetric Laplace Data					
	Log-Likelihood		Bias		
	Gauss	EM-AL	Gauss	Moment-AL	EM-AL
Case (1)	619.1798	735.0847	0.0542	0.0252	0.0226
Case (2)	1250.7110	1463.9149	0.0870	0.0376	0.0302
Case (3)	2432.3401	2919.9878	0.3750	0.1304	0.0832



**Figure 3.** Efficient frontiers of simulated Gaussian data of case (1)–(3) using Gaussian and EM-AL model.

Figure 3 show that for Gaussian data, since Gaussian data fit the model better, efficient frontiers

under Gaussian data are more close to Gaussian models; Figure 4 indicate that for generated Asymmetric Laplace data, efficient frontiers nearly equivalent to true Asymmetric Laplace data. Figure 3–4 suggest that we can first modeling data using Gaussian and Asymmetric Laplace distribution, and use the fitted log-likelihood to determine the distribution, then we evaluate the performance with the corresponding efficient frontier analysis.



**Figure 4.** Efficient frontiers of simulated Asymmetric Laplace data of case (1)–(3) using Gaussian and EM-AL model.

## 5. Real data analysis

In this section, we apply our proposed methodology to two real financial datasets, Hang Seng Index and Nasdaq Index, both datasets are downloaded from Bloomberg, with daily data range from January 4, 2011, to December 29, 2017. The variable of interest is the rate of returns multiplied by the annualized ratio  $\sqrt{252}$ , formulated as

$$\text{LogRet}(t) = \sqrt{252} \{ \log(\text{price}[t+1]) - \log(\text{price}[t]) \}, \quad t = 1, 2, \dots, 1721.$$

These two datasets are analyzed through efficient frontier analysis under ALD framework on R platform. Theorem 3.1–3.2 indicate that portfolio selection models under ALD framework can be reduced to the following quadratic programming problem:

$$\min_w \sigma^2 = \mathbf{w}'\Sigma\mathbf{w} \quad \text{s.t.} \quad \mathbf{w}'\boldsymbol{\mu} = r_0, \mathbf{w}'\mathbf{1} = 1.$$

with explicit solution (see Lai and Xing, 2008) as follows:

$$\hat{\mathbf{w}} = \frac{D - r_0 B}{AD - B^2} \Sigma^{-1} \mathbf{1} + \frac{r_0 A - B}{AD - B^2} \Sigma^{-1} \boldsymbol{\mu}. \quad (5.1)$$

Here,  $A = \mathbf{1}'\Sigma^{-1}\mathbf{1}$ ,  $B = \mathbf{1}'\Sigma^{-1}\boldsymbol{\mu}$  and  $D = \boldsymbol{\mu}'\Sigma^{-1}\boldsymbol{\mu}$ .

### 5.1. Example 1: Hang seng index

In the first example, we construct a portfolio consisting of 8 Hang Seng indexes: HK1, HK175, HK2007, HK2318, HK4, HK6, HK66. The summary descriptive statistics are reported in Table 2. It is clear that the all stock returns exhibit larger skewness and kurtosis. The median of these stocks are close to zero, the log-likelihood of Asymmetric Laplace distribution is larger than gaussian distribution, indicating that Asymmetric Laplace distribution would be a good fit than gaussian distribution.

**Table 2.** Hang Seng data statistics.

Descriptive Statistics							
	StD	Mean	Median	Skewness	Kurtosis	Jarq.Test	Jarq.Prob
HK1	19.7819	0.2370	0.0000	3.9601	71.8163	375244.7434	0.0000
HK175	3.7101	0.2183	0.0000	1.1510	28.6192	59264.8528	0.0000
HK2007	2.1968	0.1115	0.0000	0.5928	16.5623	19825.4392	0.0000
HK2318	12.6007	0.3431	0.0000	0.6174	6.2371	2908.6947	0.0000
HK4	5.8690	0.0409	0.0000	0.4894	7.6362	4263.7750	0.0000
HK6	13.0712	0.1517	0.0000	-1.1430	20.3112	30037.3385	0.0000
HK66	5.8708	0.1545	0.0000	-1.7941	19.9392	29510.3684	0.0000
			Gaussian			EM-AL	
Log-likelihood			-39707.45			-37865.83	
Parameter Estimation							
	HK1	HK175	HK2007	HK2318	HK4	HK6	HK66
$\mu$	0.2370	0.2183	0.1115	0.3431	0.0409	0.1517	0.1545
$\Sigma$	HK1	HK175	HK2007	HK2318	HK4	HK6	HK66
HK1	406.4426	12.8089	11.4547	130.0934	68.3857	94.4069	60.3087
HK175	12.8089	7.3656	1.6330	12.0356	4.7467	4.8725	3.6070
HK2007	11.4547	1.6330	3.8506	9.9261	4.6830	3.9055	2.4395
HK2318	130.0934	12.0356	9.9261	174.0423	42.1879	48.2877	35.0499
HK4	68.3857	4.7467	4.6830	42.1879	44.6336	26.5314	17.6232
HK6	94.4069	4.8725	3.9055	48.2877	26.5314	205.1471	32.5570
HK66	60.3087	3.6070	2.4395	35.0499	17.6232	32.5570	41.6153

Then we fit the data to Asymmetric Laplace distribution through EM algorithm as described in Section 3.2. Parameter estimation results are displayed in Table 2, we construct portfolios under Asymmetric Laplace framework at different levels of expected return. Consider increasing target expected return values

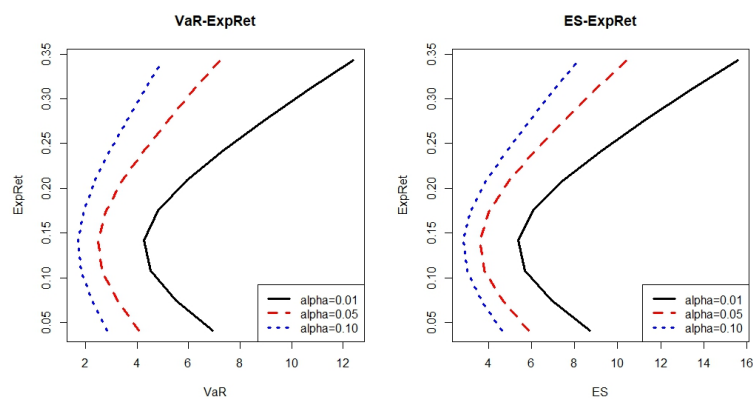
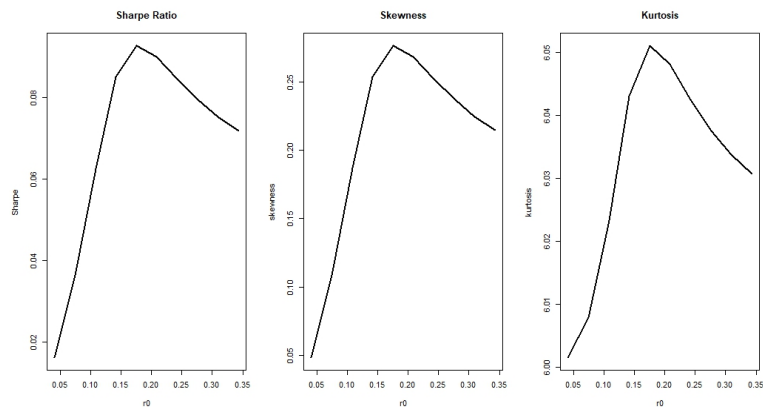
$$r_0 = 0.040, 0.0745, 0.1081, 0.1417, 0.1753, 0.2088, 0.2424, 0.2760, 0.3096, 0.3431.$$

Portfolio selection results are summarized in Table 3, with kurtosis, skewness, sharpe ratio, VaR and ES results at  $\alpha = 0.01, 0.05, 0.10$  levels.



**Table 3.** Efficient frontier results of Hang Seng data

	$r$	$\mu$	$\sigma$	Skew	Kurt	Sharpe	VaR <sub>0.01</sub>	ES <sub>0.01</sub>	VaR <sub>0.05</sub>	ES <sub>0.05</sub>	VaR <sub>0.10</sub>	ES <sub>0.10</sub>
1	0.0409	0.0409	2.5461	0.0482	6.0016	0.0161	6.9429	8.7229	4.0782	5.8582	2.8444	4.6244
2	0.0745	0.0745	2.0565	0.1086	6.0079	0.0362	5.5080	6.9254	3.2268	4.6442	2.2444	3.6618
3	0.1081	0.1081	1.7299	0.1869	6.0233	0.0625	4.5256	5.6960	2.6420	3.8124	1.8308	3.0011
4	0.1417	0.1417	1.6650	0.2537	6.0430	0.0851	4.2684	5.3771	2.4841	3.5928	1.7156	2.8243
5	0.1753	0.1753	1.8891	0.2763	6.0510	0.0928	4.8095	6.0606	2.7960	4.0471	1.9288	3.1799
6	0.2088	0.2088	2.3199	0.2682	6.0480	0.0900	5.9208	7.4601	3.4434	4.9827	2.3764	3.9157
7	0.2424	0.2424	2.8656	0.2523	6.0425	0.0846	7.3493	9.2579	4.2774	6.1861	2.9544	4.8631
8	0.2760	0.2760	3.4724	0.2372	6.0375	0.0795	8.9467	11.2679	5.2108	7.5321	3.6018	5.9231
9	0.3096	0.3096	4.1134	0.2247	6.0337	0.0753	10.6386	13.3966	6.1999	8.9578	4.2882	7.0462
10	0.3431	0.3431	4.7749	0.2147	6.0307	0.0719	12.3871	15.5962	7.2222	10.4313	4.9978	8.2069

**Figure 5.** ALD Efficient Frontier of Hang Seng Data.**Figure 6.** Skewness, Kurtosis and Sharpe Ratio Tendency of Hang Seng Data.

The efficient frontier tendencies are displayed in Figure 5. It is suggested that aggressive investors should impose higher confidence levels and conservative investors may choose smaller confidence levels. Figure 6 depicts the kurtosis, skewness, and sharpe ratio tendency of portfolio selection models. Results show that Sharpe Ratio, Skewness and Kurtosis increase fast and decreases slowly down as the target expected returns increases.

## 5.2. Example 2: Nasdaq index

In the second example, we consider Nasdaq index, including CTRP, MNST, NFLX, NTES, NVDA, TTWO, and report the descriptive statistics in Table 4. All the indexes exhibit significant skewness and kurtosis. Jarq.Test results indicate that this dataset deviates from normality significantly. We fit the log returns data to Gaussian and Asymmetric Laplace distributions. Since Asymmetric Laplace model achieve higher log-likelihood results compared to Gaussian model, we choose EM-AL model for data fitting. Parameter estimation results are displayed in Table 4.

**Table 4.** Nasdaq data statistics.

Descriptive Statistics							
	StD	Mean	Median	Skewness	Kurtosis	Jarq.Test	Jarq.Prob
CTRP	12.5853	0.2100	0.0000	1.5907	16.5100	20786.4126	0.0000
MNST	10.0679	0.4904	0.1587	1.9632	25.7988	50065.6119	0.0000
NFLX	32.9311	1.5015	-0.0079	1.0314	20.0251	29796.6074	0.0000
NTES	58.0910	2.7817	1.2700	1.2457	26.0418	50315.0307	0.0000
NVDA	25.2360	1.6024	0.3175	1.7413	40.3981	120864.0386	0.0000
TTWO	12.2826	0.8786	0.1587	2.1392	52.3926	203127.9368	0.0000
			Gaussian			EM-AL	
Log-Likelihood			-46400.25			-42798.03	
Parameter Estimation							
		CTRP	MNST	NFLX	NTES	NVDA	TTWO
$\mu$		0.2100	0.4904	1.5015	2.7817	1.6024	0.8786
$\Sigma$		CTRP	MNST	NFLX	NTES	NVDA	TTWO
CTRP	187.1236	25.4281	129.9104	246.6268	46.1539	40.0369	
MNST	25.4281	116.7107	51.2572	82.0634	26.2438	23.5405	
NFLX	129.9104	51.2572	900.1261	327.6980	122.5361	83.8650	
NTES	246.6268	82.0634	327.6980	2380.8894	213.1414	121.1493	
NVDA	46.1539	26.2438	122.5361	213.1414	259.7019	53.8530	
TTWO	40.0369	23.5405	83.8650	121.1493	53.8530	111.0020	

Then we consider increasing target expected returns

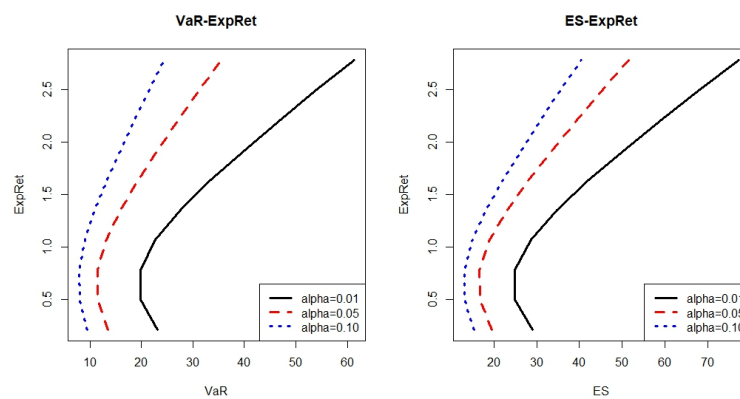
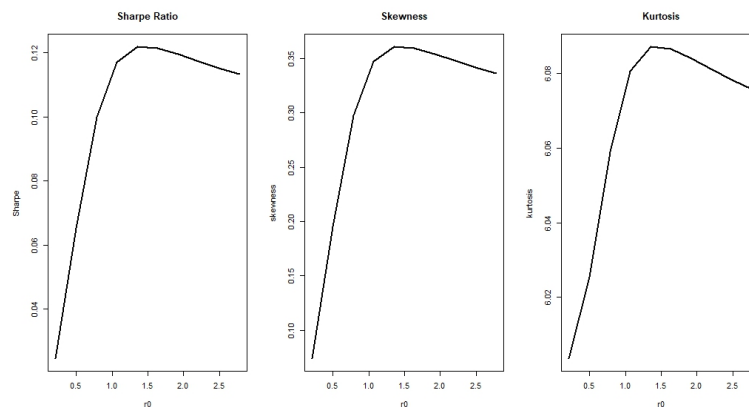
$$r_0 = 0.2100, 0.4958, 0.7815, 1.0673, 1.3530, 1.6388, 1.9245, 2.2102, 2.4960, 2.7817.$$

Results of skewness, kurtosis, sharpe ratio and VaR, ES results are summarized in Table 5. Figure 7 displays the efficient frontiers at confidence level  $\alpha = 0.01, 0.05, 0.10$ . These results show that the portfolio capture higher risk at higher  $\alpha$  levels. Figure 8 displays the skewness, kurtosis, and Sharpe Ratio Tendency. The optimal portfolios can be obtained from (5.1) with the corresponding VaR, ES, skewness, kurtosis and Sharpe Ratio.

Table 5 and Figure 7 suggests that as  $r_0$  increases, all ES ( $ES_{0.01}, ES_{0.05}, ES_{0.10}$ ) increases, indicating that higher return is derived from higher risk. It is interesting that under the ALD assumption, as  $r_0$  increases, Sharpe ratio and skewness first decreases then increases accordingly. As  $\alpha$  increases, VaR and ES measures decreases. Thus, conservative investors can choose larger  $\alpha$  levels and aggressive investors would select smaller  $\alpha$  levels.

**Table 5.** Efficient frontier analysis of Nasdaq data.

	$r$	$\mu$	$\sigma$	Skew	Kurt	Sharpe	VaR <sub>0.01</sub>	ES <sub>0.01</sub>	VaR <sub>0.05</sub>	ES <sub>0.05</sub>	VaR <sub>0.10</sub>	ES <sub>0.10</sub>
1	0.2100	0.2100	8.5237	0.0739	6.0036	0.0246	23.0672	28.9903	13.5344	19.4575	9.4288	15.3519
2	0.4958	0.4958	7.5957	0.1951	6.0254	0.0653	19.8220	24.9509	11.5675	16.6963	8.0125	13.1413
3	0.7815	0.7815	7.8219	0.2973	6.0590	0.0999	19.7857	24.9397	11.4907	16.6447	7.9183	13.0723
4	1.0673	1.0673	9.1167	0.3472	6.0806	0.1171	22.7065	28.6414	13.1547	19.0896	9.0409	14.9758
5	1.3530	1.3530	11.1127	0.3608	6.0870	0.1218	27.5608	34.7713	15.9560	23.1665	10.9581	18.1686
6	1.6388	1.6388	13.5025	0.3597	6.0865	0.1214	33.4993	42.2627	19.3952	28.1586	13.3208	22.0843
7	1.9245	1.9245	16.1117	0.3541	6.0838	0.1194	40.0422	50.5132	23.1898	33.6608	15.9318	26.4028
8	2.2102	2.2102	18.8494	0.3478	6.0808	0.1173	46.9392	59.2083	27.1927	39.4619	18.6883	30.9575
9	2.4960	2.4960	21.6671	0.3418	6.0781	0.1152	54.0563	68.1800	31.3251	45.4488	21.5353	35.6590
10	2.7817	2.7817	24.5371	0.3365	6.0757	0.1134	61.3178	77.3329	35.5424	51.5576	24.4416	40.4567

**Figure 7.** ALD efficient frontier of Hang Seng data.**Figure 8.** Skewness, kurtosis and Sharpe Ratio tendency of Nasdaq data.

## 6. Conclusion and prospects

In this paper, we derive several equivalent portfolio selection models under ALD framework, these models can be transformed to quadratic programming problem with explicit solutions. The Expectation-Maximization algorithm for parameter estimation of Asymmetric Laplace distribution is obtained and outperforms moment estimation.

There are several advantages of Asymmetric Laplace distribution based models. First, the equivalence of risk measures such as VaR, ES and StD at maximization of skewness and minimization of kurtosis facilitate portfolio selection models significantly. Second, confidence levels of these models offer investors various portfolio selection choices. Conservative investors can choose larger  $\alpha$  levels and aggressive investors can select smaller  $\alpha$  levels. Finally, ALD model is able to explain skewness and kurtosis in financial data. Therefore, the Asymmetric Laplace distribution can be widely applied to handle real financial datasets.

We may further extend the Asymmetric Laplace based portfolio selection model to cases of mixture Asymmetric Laplace distributions. Another direction is to combine clustering techniques (see Dias et al., 2015; Iorio et al., 2018) with Asymmetric Laplace distribution for portfolio selection of time series models.

### Acknowledgments

Chi Tim, Ng's work is supported by the 2016 Chonnam National University Research Program grant (No. 2016–2762).

### Conflict of interest

The authors declare no conflict of interest.

### References

- Arslan O (2010) An alternative multivariate skew Laplace distribution: properties and estimation. *Stat Pap* 51: 865–887.
- Artzner P, Delbaen F, Eber JM, et al. (1999) Coherent measures of risk. *Math financ* 9: 203–228.
- Ayebo A, Kozubowski TJ (2003) An asymmetric generalization of Gaussian and Laplace laws. *J Probab Stat Science* 1: 187–210.
- Barndorff-Nielsen OE (1995) Normal inverse Gaussian distributions and the modeling of stock returns. 24: 1–13.
- Bingham NH, Kiesel R (2001) Modelling asset returns with hyperbolic distributions. *In Return Distributions in Finance*, 1–20.
- Behr A, Ptter U (2009) Alternatives to the normal model of stock returns: Gaussian mixture, generalised logF and generalised hyperbolic models. *Ann Financ* 5: 49–68.
- Dias JG, Vermunt JK, Ramos S (2015) Clustering financial time series: New insights from an extended hidden Markov model. *European J Oper Res* 243: 852–864.
- Eberlein E (2001) Application of generalized hyperbolic Lvy motions to finance. *In Lvy processes*, Birkhuser, Boston, MA, 319–336.
- Eberlein E, Keller U (1995) Hyperbolic distributions in finance. *Bernoulli* 1: 281–299.

- Hellmich M, Kassberger S (2011) Efficient and robust portfolio optimization in the multivariate generalized hyperbolic framework. *Quant Financ* 11: 1503–1516.
- Hu W, Kercheval A (2007) Risk management with generalized hyperbolic distributions. *In Proceedings of the Fourth IASTED International Conference on Financial Engineering and Applications*, ACTA Press, 19–24.
- Hu W, Kercheval AN (2010) Portfolio optimization for student t and skewed t returns. *Quant Financ* 10: 91–105.
- Hrlimann W (2013) A moment method for the multivariate asymmetric Laplace distribution. *Stat Probabil Lett* 83: 1247–1253.
- Iorio C, Frasso G, D'Ambrosio A, et al. (2018) A P-spline based clustering approach for portfolio selection. *Expert Syst Appl* 95: 88–103.
- Kollo T, Srivastava MS (2005) Estimation and testing of parameters in multivariate Laplace distribution. *Commun Stat-Theor M* 33: 2363–2387.
- Kotz S, Kozubowski TJ, Podgrski K (2001) Asymmetric multivariate Laplace distribution. *In The Laplace Distribution and Generalizations*, Birkhuser, Boston, MA, 239–272.
- Kotz S, Kozubowski TJ, Podgrski K (2002) Maximum likelihood estimation of asymmetric Laplace parameters. *Ann I Stat Math* 54: 816–826.
- Kotz S, Kozubowski T, Podgorski K (2012) The Laplace distribution and generalizations: a revisit with applications to communications, economics, engineering, and finance. Springer Science & Business Media.
- Kozubowski TJ, Podgrski K (1999) A class of asymmetric distributions. *Actuarial Research Clearing House* 1: 113–134.
- Kozubowski TJ, Podgrski K (2001) Asymmetric Laplace laws and modeling financial data. *Math Comput Model* 34: 1003–1021.
- Lai TL, Xing H (2008) Statistical models and methods for financial markets. New York: Springer.
- Mandelbrot BB (1997) The variation of certain speculative prices. *In Fractals and scaling in finance*, Springer, New York, NY, 371–418.
- Markowitz H (1952) Portfolio selection. *The journal of finance* 7: 77–91.
- Punathumparambath B (2012) The multivariate asymmetric slash Laplace distribution and its applications. *Statistica* 72: 235.
- Seneta E (2004) Fitting the variance-gamma model to financial data. *J Appl Probab* 41: 177–187.
- Socgnia VK, Wilcox D (2014) A comparison of generalized hyperbolic distribution models for equity returns. *J Appl Math*.
- Stacy EW (1962) A generalization of the gamma distribution. *Ann math stat* 1187–1192.

- 
- Surya BA, Kurniawan R (2014) Optimal portfolio selection based on expected shortfall under generalized hyperbolic distribution. *Asia-Pacific Financ Mark* 21: 193–236.
- Visk H (2009) On the parameter estimation of the asymmetric multivariate Laplace distribution. *Commun StatTheor M* 38: 461–470.
- Yiu KFC (2004) Optimal portfolios under a value-at-risk constraint. *J Economic Dyn Control* 28: 1317–1334.
- Zhao S, Lu Q, Han L, et al. (2015) A mean-CVaR-skewness portfolio optimization model based on asymmetric Laplace distribution. *Ann Oper Res* 226: 727–739.
- Zhu Y (2007) Application of Asymmetric Laplace Laws in Financial Risk Measures and Time Series Analysis. *Doctoral dissertation*, University of Florida.



AIMS Press

©2018 the Author(s), licensee AIMS Press. This is an open access article distributed under the terms of the Creative Commons Attribution License (<http://creativecommons.org/licenses/by/4.0>)