An asymmetric generalization of Gaussian and Laplace laws

Abraham Ayebo and Tomasz J Kozubowski

Department of Mathematics, University of Nevada at Reno, Reno, NV 89557

Abstract: We study a class of skew continuous distributions on the real line that arises from symmetric exponential power laws by incorporating inverse scale factors into the positive and negative orthants. Skew and symmetric Laplace and normal laws are included in this class as special cases. We present main properties of skew exponential power laws, derive maximum likelihood estimators of their parameters, and discuss their applications in finance.

Keywords: Asymmetric Laplace law, double exponential distribution, exponential power distribution, foreign currency exchange rate, generalized error distribution, generalized exponential distribution, maximum likelihood estimation, skew normal distribution

AMS 2000 Subject Classification: Primary 62E10; Secondary 62F10, 62F12, 62P05, 94A17

1. Introduction

Consider a positive exponential power distribution with density

\[ h(x) = \frac{\alpha}{\Gamma\left(\frac{1}{\alpha}\right)} \left\{ \begin{array}{ll}
\exp\left(-\frac{1}{\sigma^\alpha}x^\alpha\right) & \text{for } x \geq 0, \\
0 & \text{for } x < 0,
\end{array} \right. \quad (1.1) \]

where \( \alpha > 0 \) is the shape parameter and \( \sigma > 0 \) is the scale parameter (see, e.g., Bartoszewicz (1985) and references therein). A symmetrization of (1.1) leads to the exponential power distribution (generalized error distribution) on \( \mathbb{R} \) with p.d.f.

\[ g(x) = \frac{\alpha}{2\sigma\Gamma(1/\alpha)} e^{-|x|^{\alpha}/\sigma^\alpha}, \quad x \in \mathbb{R}. \quad (1.2) \]

This distribution was introduced by Subbotin (1923) and popularized by Box and Tiao (1962, 1964, 1973), who used it in robustness studies (see also Tiao and Lund (1970), Swamy and Mehta (1977), West (1984), and more recent Osiewalski and Steel (1993)).

Since exponential power laws including their special cases of normal ($\alpha = 2$) and Laplace ($\alpha = 1$) distributions are symmetric, they are not appropriate for modeling data with asymmetric empirical distributions. However, various practical applications require models for unimodal but skew data. In this work we propose a skew exponential power model, derive its properties, and investigate its applications in stochastic modeling.

We follow Fernandez and Steel (1998) (see also Fernandez et al. (1995)), who introduced skewness into a symmetric distribution by incorporating inverse scale factors into the negative and positive orthants. More precisely, if $g$ is a symmetric p.d.f. on $\mathbb{R}$, then for any $\kappa > 0$ we obtain a skew density

\begin{equation}
(1.3) \quad f(x) = \frac{2\kappa}{1 + \kappa^2} \begin{cases} 
 g(x\kappa) & \text{for } x \geq 0, \\
 g\left(\frac{x}{\kappa}\right) & \text{for } x < 0.
\end{cases}
\end{equation}

Normal density generates the class of skew normal distribution (see Tiao and Lund (1970) and Mudholkar and Hutson (2000) and references therein). The Laplace density

\begin{equation}
(1.4) \quad g(x) = \frac{1}{2\sigma} e^{-|x|/\sigma}, \ x \in \mathbb{R},
\end{equation}

leads to the class of skew Laplace distributions, valuable for stochastic modeling in a variety of fields, including finance, economics, and the sciences (see Kotz et al. (2001)).

We consider the class of skew distributions obtained via (1.3) when $g$ is a symmetric exponential power density (1.2). Similarly to the symmetric EP laws, the resulting skew
EP distributions exhibit variety of tail behaviors including light tailed near-normal laws ($\alpha \approx 2$) and heavier tailed near-Laplace distributions ($\alpha \approx 1$). Additionally, by introducing asymmetry we obtain a flexible class with potential applications in stochastic modeling and robustness studies.

Our paper is organized as follows. Definitions and basic properties are presented in Section 2. Section 3 is devoted to estimation of the parameters. In Section 4 we consider applications of EP models to financial data. Technical results and proofs are collected in Section 5.

2. Definition and properties

Following the procedure of Fernandez and Steel (1998) described above and introducing an additional location parameter, we define an asymmetric exponential power distribution.

**Definition 2.1.** A random variable $Y$ is said to have an asymmetric exponential power (EP) distribution if there exist parameters $\alpha > 0$, $\theta \in \mathbb{R}$, $\sigma > 0$, and $\kappa > 0$ such that the density function of $Y$ has the form

$$f(x) = \frac{\alpha}{\sigma \Gamma\left(\frac{1}{\alpha}\right)} \frac{\kappa}{1 + \kappa^2} \exp\left( -\frac{\kappa^\alpha}{\sigma^\alpha}[(x-\theta)^+]^\alpha - \frac{1}{\sigma^\alpha \kappa^\alpha}[(x-\theta)^-]^\alpha \right),$$

where

$$u^+ = \begin{cases} u & \text{if } u \geq 0, \\ 0 & \text{if } u < 0 \end{cases} \quad \text{and} \quad u^- = \begin{cases} -u & \text{if } u \leq 0, \\ 0 & \text{if } u > 0. \end{cases}$$

We denote the distribution of $Y$ by $\mathcal{EP}_\alpha(\theta, \sigma, \kappa)$ and write $Y \sim \mathcal{EP}_\alpha(\theta, \sigma, \kappa)$.

Parameters $\theta$ and $\sigma$ correspond to location and scale, respectively, while $\kappa$ controls skewness, and $\alpha$ is the shape parameter. For $\kappa = 1$, the distribution is symmetric about $\theta$. The distribution function of $Y \sim \mathcal{EP}_\alpha(\theta, \sigma, \kappa)$ is given by

$$F(y) = Pr(Y \leq y) = \begin{cases} \frac{\kappa^2}{1 + \kappa^2} \frac{1}{\Gamma\left(\frac{1}{\alpha}\right)} \Gamma\left(\frac{1}{\alpha}, \left[ \frac{\theta-y}{\sigma \kappa} \right]^\alpha \right) & \text{for } y < \theta, \\ 1 - \frac{\kappa^2}{1 + \kappa^2} \frac{1}{\Gamma\left(\frac{1}{\alpha}\right)} \Gamma\left(\frac{1}{\alpha}, \frac{\kappa^\alpha}{\sigma^\alpha} (y-\theta)^\alpha \right) & \text{for } y \geq \theta, \end{cases}$$

where

$$\Gamma(a, x) = \int_x^\infty t^{a-1} \exp(-t) dt, \quad a > 0, \quad x > 0,$$
is the incomplete gamma function.

2.1. Special Cases. If \( \kappa = 1 \) we obtain the symmetric exponential power distribution (1.2). In case \( \kappa \neq 1 \), letting \( \alpha = 1 \) leads to the skew Laplace distribution with density

\[
f(x) = \frac{1}{\sigma} \frac{\kappa}{1 + \kappa^2} \begin{cases} 
\exp \left( -\frac{\kappa}{\sigma} |x - \theta| \right) & \text{for } x \geq \theta, \\
\exp \left( -\frac{1}{\sigma \kappa} |x - \theta| \right) & \text{for } x < \theta,
\end{cases}
\]

see, e.g., Kotz et al. (2001). For \( \alpha = 2 \), we obtain the skew normal distribution studied recently by Mudholkar and Hutson (2000).

2.2. Moments and related parameters. Moments corresponding to exponential power random variables are straightforward to compute using Lemma 1 in Section 5.

Proposition 2.2. If \( Y \sim \mathcal{E}P_{\alpha}(\theta, \sigma, \kappa) \), then

\[
E(Y - \theta)^n = \sigma^n \frac{\kappa}{\Gamma(\frac{1}{\alpha})(1 + \kappa^2)} \Gamma \left( \frac{n+1}{\alpha} \right) \left( (-1)^n \kappa^{n+1} + \frac{1}{\kappa^{n+1}} \right), \quad n = 0, 1, 2, ...
\]

\[
E(|Y - \theta|^\eta) = \sigma^\eta \frac{\kappa}{\Gamma(\frac{1}{\alpha})(1 + \kappa^2)} \Gamma \left( \frac{\eta+1}{\alpha} \right) \left( \kappa^{\eta+1} + \frac{1}{\kappa^{\eta+1}} \right), \quad \eta > -1.
\]

\[
E((Y - \theta)^+)^\eta = \sigma^\eta \frac{\Gamma(\frac{n+1}{\alpha})}{\Gamma(\frac{1}{\alpha})^2(1 + \kappa^2)^\eta}, \quad \eta > -1.
\]

\[
E((Y - \theta)^-)^\eta = \sigma^\eta \frac{\kappa^{\eta+2}\Gamma(\frac{n+1}{\alpha})}{\Gamma(\frac{1}{\alpha})(1 + \kappa^2)^\eta}, \quad \eta > -1.
\]

In particular we get the following expression for the mean and the variance:

\[
\mu_Y = E(Y) = \theta + \sigma \left( \frac{1}{\kappa} - \kappa \right) \frac{\Gamma(\frac{2}{\alpha})}{\Gamma(\frac{1}{\alpha})}.
\]

\[
\sigma_Y^2 = Var(Y) = \sigma^2 \frac{\Gamma(\frac{2}{\alpha})}{\Gamma(\frac{1}{\alpha}) \kappa^2(1 + \kappa^2)} - \sigma^2 \frac{\Gamma^2(\frac{2}{\alpha})}{\Gamma^2(\frac{1}{\alpha})} \frac{(1 - \kappa^2)^2}{\kappa^2}.
\]

Proposition 2.2 is useful in deriving the skewness

\[
\gamma_1 = E \left[ \left( \frac{Y - \mu_Y}{\sigma_Y} \right)^3 \right]
\]
and the (excess) kurtosis

\begin{equation}
\gamma_2 = E \left[ \left( \frac{Y - \mu_Y}{\sigma_Y} \right)^4 \right] - 3
\end{equation}

of EP laws (see Ayebo (2002)).

**Proposition 2.3.** If \( X \sim \mathcal{EP}_\alpha(\theta, \sigma, \kappa) \), then the skewness of \( X \) is

\begin{equation}
\gamma_1 = \frac{(1 - \kappa^8)\Gamma^2(\frac{1}{\alpha})\Gamma(\frac{1}{\alpha}) - 3(1 - \kappa^2)(1 + \kappa^6)\Gamma(\frac{1}{\alpha})\Gamma(\frac{2}{\alpha})\Gamma(\frac{3}{\alpha}) + 2(1 - \kappa^2)^3(1 + \kappa^2)\Gamma^3(\frac{2}{\alpha})}{(1 + \kappa^2) \left( \sqrt{\Gamma(\frac{1}{\alpha})\Gamma(\frac{3}{\alpha})(1 + \kappa^6)} - \Gamma^2(\frac{2}{\alpha})(1 - \kappa^2)^2 \right)^3}
\end{equation}

and the kurtosis is

\begin{equation}
\gamma_2 = \frac{1 + \kappa^2}{\left[ \Gamma(\frac{1}{\alpha})\Gamma(\frac{2}{\alpha})(1 + \kappa^6) - \Gamma^2(\frac{2}{\alpha})(1 - \kappa^2)^2(1 + \kappa^2) \right]^2} \times 
\left\{ \Gamma^3 \left( \frac{1}{\alpha} \right) \Gamma \left( \frac{5}{\alpha} \right) (1 + \kappa^{10}) - 4\Gamma^2 \left( \frac{1}{\alpha} \right) \Gamma \left( \frac{2}{\alpha} \right) \Gamma \left( \frac{4}{\alpha} \right) (1 - \kappa^2) (1 - \kappa^8) 
+ 6\Gamma \left( \frac{1}{\alpha} \right) \Gamma^2 \left( \frac{2}{\alpha} \right) \Gamma \left( \frac{3}{\alpha} \right) (1 - \kappa^2)^2 (1 + \kappa^6) - 3\Gamma^4 \left( \frac{2}{\alpha} \right) (1 - \kappa^2)^4 (1 + \kappa^2) \right\} - 3.
\end{equation}

Note that in the special case \( \alpha = 1 \) (Laplace distribution) we have

\[ \gamma_1 = 2 \frac{1 - \kappa^3}{(1 + \kappa^2 + \kappa^6)^2} \quad \text{and} \quad \gamma_2 = 6 - \frac{12}{(1/\kappa^2 + \kappa^2)^2}, \]

which coincide with the parameters of the asymmetric Laplace distribution, see Kotz et al. (2001).

### 2.3. Representations.

Recall that in the symmetric case \( X \sim \mathcal{EP}_\alpha(0, 1, 1) \) we have the representation \( X \overset{d}{=} IW^{\frac{1}{\alpha}} \), where \( W \) has gamma distribution with density

\begin{equation}
\begin{aligned}
f(x) &= \frac{1}{\Gamma(\frac{1}{\alpha})} x^{\frac{\alpha}{2} - 1} \exp(-x), \quad x > 0,
\end{aligned}
\end{equation}

and \( I \) is an independent random sign (see Johnson (1979)). A straightforward conditioning argument (see Ayebo (2002)) shows that in the skew case \( Y \sim \mathcal{EP}_\alpha(\theta, \sigma, \kappa) \) we have

\begin{equation}
Y \overset{d}{=} \theta + \sigma IW^{\frac{1}{\alpha}},
\end{equation}
where $W$ has gamma distribution with density (2.16) and

\begin{equation}
I = \begin{cases} 
-\kappa & \text{with prob. } \frac{\kappa^2}{1+\kappa^2}, \\
\frac{1}{\kappa} & \text{with prob. } \frac{1}{1+\kappa^2}.
\end{cases}
\end{equation}

This representation is useful in generating random variates from this distribution.

**An $\mathcal{EP}_\alpha(\theta, \sigma, \kappa)$ generator.**

- Generate a gamma $G(\frac{1}{\alpha}, 1)$ random variate $W$.
- Generate a Standard Uniform random variate $U$.
- If $U < \frac{1}{1+\kappa^2}$, set $I \leftarrow \frac{1}{\kappa}$
  else set $I \leftarrow -\kappa$
- Set $Y \leftarrow \theta + \sigma I W^\frac{1}{\alpha}$
- RETURN $Y$.

2.4. **Maximum entropy property.** The entropy of a one-dimensional r.v. $X$ with density (or probability function) $f$ is defined as

\begin{equation}
H(X) = E[-\log f(X)]
\end{equation}

and measures the uncertainty associated with the distribution of $X$ (see Jaynes (1957)). A straightforward calculation shows that the entropy of $X \sim \mathcal{EP}_\alpha(\theta, \sigma, \kappa)$ is

\begin{equation}
H(X) = \log \sigma + \log \Gamma \left( \frac{1}{\alpha} \right) + \log \left( \kappa + \frac{1}{\kappa} \right) + \frac{1}{\alpha} - \log \alpha,
\end{equation}

see Ayebo (2002).

The concept of entropy has been successfully applied in a variety of fields including statistical mechanics, statistics, stock market analysis, queuing theory, image analysis, reliability estimation [see, e.g., Kapur (1993)]. Finding the maximum entropy distribution can be viewed as a general inference procedure (see Jaynes (1957)). When $\alpha = 1$, the distribution that maximizes the entropy subject to the conditions

\begin{equation}
EX = c_1 \in \mathbb{R} \quad \text{and} \quad E|X| = c_2 > |c_1|
\end{equation}
is the skew Laplace distribution $\mathcal{EP}_1(0, \sigma, \kappa)$, see Kotz et al. (2002). In our generalization to the case $\alpha \neq 1$, we use the notation

\begin{equation}
X^{<\alpha>} = |x|^{\alpha} \text{sign}(x) = \begin{cases} 
x^{\alpha} & \text{if } x > 0, \\
0 & \text{if } x = 0, \\
-(\alpha)^{\alpha} & \text{if } x < 0.
\end{cases}
\end{equation}

**Proposition 2.4.** Consider the class $C$ of all continuous random variables with non-vanishing densities $(-\infty, \infty)$ such that

\begin{equation}
EX^{<\alpha>} = c_1 \in \mathbb{R} \quad \text{and} \quad E|X|^{\alpha} = c_2 > 0 \quad \text{for } X \in C,
\end{equation}

where

\begin{equation}
|c_1| < c_2.
\end{equation}

Then, the maximum entropy is attained for the $\mathcal{EP}$ r.v. $X^*$ with density (2.1), where $\theta = 0$,

\begin{equation}
\kappa = \left(\frac{c_2 - c_1}{c_2 + c_1}\right)^{\frac{1}{2(\alpha+1)}},
\end{equation}

and

\begin{equation}
\sigma = \left\{\frac{\alpha}{2} (c_2 - c_1)^{\frac{\alpha}{2(\alpha+1)}} (c_2 + c_1)^{\frac{\alpha}{2(\alpha+1)}} \left[(c_2 + c_1)^{\frac{1}{\alpha+1}} + (c_2 - c_1)^{\frac{1}{\alpha+1}}\right]\right\}^{\frac{1}{\alpha}}.
\end{equation}

Moreover, the maximum entropy is

\begin{equation}
\max_{X \in C} H(X) = H(X^*) = \frac{1}{\alpha} \log \Gamma\left(\frac{\frac{1}{\alpha}}{2}\right) + \frac{1 + \alpha}{\alpha} \log[(c_1 + c_2)^{\frac{1}{\alpha+1}} + (c_2 - c_1)^{\frac{1}{\alpha+1}}].
\end{equation}

**Remark 2.5.** If we drop the condition $EX^{<\alpha>} = c_1$ then the entropy is maximized by a symmetric $\mathcal{EP}$ distribution, see Ayebo (2002).

3. **Estimation**

In this section we discuss maximum likelihood estimation of the skew exponential power parameters. Assuming that the mode of the distribution is 0 ($\theta = 0$) and that $\alpha$ is known, we focus on estimators of $\sigma$ and $\kappa$ and their asymptotic behavior. We also briefly discuss a numerical procedure for estimating the shape parameter $\alpha$. 

Let \( X_1, \ldots, X_n \) be an i.i.d. random sample from an \( \mathcal{EP}_\alpha(0, \sigma, \kappa) \) distribution with the density \( f_{\alpha,0,\sigma,\kappa} \) given by (2.1), and let \( x_1, \ldots, x_n \) be their particular realization. Then, the log-likelihood function is

\[
\log L(\kappa, \sigma, \alpha) = n \left( \log \frac{\alpha}{\Gamma(1/\alpha)} + \frac{\kappa}{1 + \kappa^2} - \log \sigma - \frac{\kappa\alpha \bar{x}_\alpha^+}{\sigma^\alpha} - \frac{1}{\kappa^\alpha\sigma^\alpha \bar{x}_\alpha^-} \right),
\]

where

\[
\bar{x}_\alpha^+ = \frac{1}{n} \sum_{i=1}^{n} (x_i^+)^\alpha, \quad \bar{x}_\alpha^- = \frac{1}{n} \sum_{i=1}^{n} (x_i^-)^\alpha,
\]

and \( x_i^+ \) and \( x_i^- \) are given by (2.2).

### 3.1. Fisher information matrix

We start with the Fisher information matrix

\[
I(\sigma, \kappa) = \left[ -E \left( \frac{\partial^2}{\partial \gamma_i \partial \gamma_j} \log f_{\alpha,0,\sigma,\kappa}(X) \right) \right]_{i,j=1,2},
\]

where \( X \) has the \( \mathcal{EP}_\alpha(0, \sigma, \kappa) \) distribution with the vector-parameter \( \gamma = (\gamma_1, \gamma_2)' = (\sigma, \kappa)' \). Routine albeit lengthy calculations (see Ayebo (2002)) produce

\[
I(\sigma, \kappa) = \begin{bmatrix}
\frac{\alpha}{\sigma^2} & \frac{\alpha \kappa^2 - 1}{\sigma \kappa^2 + 1} \\
\frac{\alpha \kappa^2 - 1}{\sigma \kappa^2 + 1} & \frac{\alpha}{\kappa^2} + \frac{4}{(1+\kappa^2)^2}
\end{bmatrix}.
\]

### 3.2. Case 1: The value of \( \sigma \) is unknown

By (3.1) we need to maximize the function

\[
Q(\sigma) = - \log \sigma - \frac{\kappa \alpha}{\sigma^\alpha} \bar{x}_\alpha^+ - \frac{1}{\kappa^\alpha \sigma^\alpha} \bar{x}_\alpha^-.
\]

It is easy to see that

\[
\hat{\sigma}_n = \left[ \alpha \left( \kappa \alpha \bar{X}_\alpha^+ + \frac{1}{\kappa^\alpha \sigma^\alpha} \bar{X}_\alpha^- \right) \right]^{\frac{1}{\alpha}}
\]

is a unique maximum likelihood estimator (MLE) of \( \sigma \).

**Proposition 3.1.** Let \( X_1, \ldots, X_n \) be i.i.d. r.v.’s from the \( \mathcal{EP}_\alpha(0, \sigma, \kappa) \) distribution with an unknown value of \( \sigma \). The MLE of \( \sigma \) is given by (3.5) and is

(i) Consistent;

(ii) Asymptotically normal, where \( \sqrt{n}(\hat{\sigma}_n - \sigma) \xrightarrow{d} N(0, \sigma^2/\alpha) \);

(iii) Asymptotically efficient, that is the asymptotic variance \( \sigma^2/\alpha \) coincides with the reciprocal of the Fisher information \( I(\sigma) \).
3.3. Case 2: The value of $\kappa$ is unknown. In view of (3.1), we need to maximize

\begin{equation}
Q(\kappa) = \log \kappa - \log(1 + \kappa^2) - \frac{\kappa^\alpha}{\sigma^\alpha \bar{x}_\alpha^+} - \frac{1}{\kappa^\alpha \sigma^\alpha \bar{x}_\alpha^-},
\end{equation}

where $\bar{x}_\alpha^+$ and $\bar{x}_\alpha^-$ are given by (3.2) as before. Our first result shows that there exists a unique MLE of $\kappa$.

**Proposition 3.2.** If not both $\bar{x}_\alpha^+$ and $\bar{x}_\alpha^-$ are equal to zero, then there exists a unique $\hat{\kappa}_n \in (0, \infty)$ that maximizes the function $Q(\kappa)$ in (3.6). The value of $\hat{\kappa}_n$ is a unique solution of the equation

\begin{equation}
\kappa^\alpha(1 - \kappa^2) + \frac{\alpha}{\sigma^\alpha}(1 + \kappa^2)(\bar{x}_\alpha^- - \bar{x}_\alpha^+ \kappa^2) = 0.
\end{equation}

Moreover, if $0 < \bar{x}_\alpha^- < \bar{x}_\alpha^+$, then

$$1 < \hat{\kappa}_n < \left( \frac{\bar{x}_\alpha^-}{\bar{x}_\alpha^+} \right)^{\frac{1}{2\alpha}},$$

if $0 < \bar{x}_\alpha^- < \bar{x}_\alpha^+$, then

$$\left( \frac{\bar{x}_\alpha^-}{\bar{x}_\alpha^+} \right)^{\frac{1}{2\alpha}} < \hat{\kappa}_n < 1,$$

and if $0 < \bar{x}_\alpha^- = \bar{x}_\alpha^+$, then $\hat{\kappa}_n = 1$.

We see that the likelihood function is maximized by a unique MLE $\hat{\kappa}_n$ that in practice can be obtained by (numerically) solving the equation

\begin{equation}
y^\alpha(1 - y^2) - \frac{\bar{x}_\alpha^+}{\sigma^\alpha}y^{2\alpha}(1 + y^2) + \frac{\bar{x}_\alpha^-}{\sigma^\alpha} \alpha(1 + y^2) = 0.
\end{equation}

The properties of the MLE are presented below.

**Proposition 3.3.** Let $X_1, \ldots, X_n$ be i.i.d. variables from an $\mathcal{EP}_\alpha(0, \sigma, \kappa)$ distribution where the values of $\alpha$ and $\sigma$ are known. Then the MLE of $\kappa$ is the unique solution $\hat{\kappa}_n$ of the equation (3.8) and is

(i) Consistent;

(ii) Asymptotically normal with $\sqrt{n}(\hat{\kappa}_n - \kappa) \rightarrow N(0, \sigma^2_\kappa)$, where

\begin{equation}
\sigma^2_\kappa = \frac{\kappa^2(1 + \kappa^2)^2}{\alpha(1 + \kappa^2)^2 + 4\kappa^2};
\end{equation}

(iii) Asymptotically efficient, that is the asymptotic variance (3.9) coincides with the reciprocal of the Fisher information $I(\kappa)$. 
3.4. Case 3: The values of $\kappa$ and $\sigma$ are unknown. Here, we need to maximize the function

\begin{equation}
Q(\kappa, \sigma) = \log \kappa - \log(1 + \kappa^2) - \log \sigma - \frac{\kappa^\alpha}{\sigma^\alpha \bar{x}_\alpha^+} - \frac{1}{\kappa^\alpha \sigma^\alpha \bar{x}_\alpha^-},
\end{equation}

where $\bar{x}_\alpha^+$ and $\bar{x}_\alpha^-$ are given by (3.2) as before. Note that for any fixed value of $\kappa > 0$, the value of $Q(\kappa, \sigma)$ is maximized by

\begin{equation}
\sigma(\kappa) = \alpha^\frac{1}{\alpha} \left[ \kappa^\alpha \bar{x}_\alpha^+ + \frac{1}{\kappa^\alpha \bar{x}_\alpha^-} \right]^\frac{1}{\alpha},
\end{equation}

as we verified earlier. When we substitute this into (3.10), we find that

\begin{equation}
Q(\kappa, \sigma) \leq Q(\kappa, \sigma(\kappa)) = \log \kappa - \log(1 + \kappa^2) - \frac{1}{\alpha} \log \alpha - \frac{1}{\alpha} \log \left[ \kappa^\alpha \bar{x}_\alpha^+ + \frac{1}{\kappa^\alpha \bar{x}_\alpha^-} \right] - \frac{1}{\alpha}.
\end{equation}

Assume first that $\bar{x}_\alpha^+ > 0$ and $\bar{x}_\alpha^- = 0$, so that all sample values $x_1, \ldots, x_n$ are greater or equal than zero. Then, the right-hand-side of (3.12) is decreasing in $\kappa$ on $(0, \infty)$, so that the least upper bound of $Q(\kappa, \sigma)$ is

\[ \lim_{\kappa \to 0^+} Q(\kappa, \sigma(\kappa)) = -\frac{1}{\alpha} \log \alpha - \frac{1}{\alpha} \log \bar{x}_\alpha^+, \]

corresponding to the values of $\kappa = 0$ and $\sigma = \sigma(0) = 0$. Although these are not permitted values of the parameters, as $\kappa \to 0^+$ and $\sigma(\kappa) = \alpha^\frac{1}{\alpha} \kappa(\bar{x}_\alpha^+) \frac{1}{\alpha} \to 0^+$ the $\mathcal{E}_\alpha(0, \sigma(\kappa), \kappa)$ density converges to

\[ f(x) = \frac{\alpha}{\Gamma(\frac{1}{\alpha})} \frac{1}{(\alpha \bar{x}_\alpha^+) \frac{1}{\alpha}} \begin{cases} 
\exp\left(-\frac{1}{\alpha \bar{x}_\alpha^+} x^\alpha\right) & \text{if } x \geq 0, \\
0 & \text{if } x < 0,
\end{cases} \]

which is the one-sided exponential power density (1.1) with shape parameter $\alpha$ and scale parameter $(\alpha \bar{x}_\alpha^+) \frac{1}{\alpha}$. It makes an intuitive sense to conclude that the underlying distribution is concentrated on $(0, \infty)$ if all sample values are positive.

Similarly, when all sample values are less than or equal to zero ($\bar{x}_\alpha^+ = 0$ and $\bar{x}_\alpha^- > 0$), the maximum likelihood approach leads to the conclusion that the underlying density is

\[ f(x) = \frac{\alpha}{\Gamma(\frac{1}{\alpha})} \frac{1}{(\alpha \bar{x}_\alpha^-) \frac{1}{\alpha}} \begin{cases} 
0 & \text{if } x \geq 0, \\
\exp\left(-\frac{1}{\alpha \bar{x}_\alpha^-} (-x)^\alpha\right) & \text{if } x < 0,
\end{cases} \]

This distribution concentrated on $(-\infty, 0)$ corresponds to the random variable $-X$, where $X$ has a one-sided (positive) exponential power distribution (1.1) with shape parameter $\alpha$ and scale parameter $(\alpha \bar{x}_\alpha^-) \frac{1}{\alpha}$. 
Finally, if some sample values are positive and some are negative, we have the following result.

**Proposition 3.4.** Let \( X_1, \ldots, X_n \) be i.i.d. variables from the \( \mathcal{EP}_\alpha(0, \sigma, \kappa) \) distribution with unknown values of \( \sigma \) and \( \kappa \). Then, if \( \bar{X}_\alpha^+ > 0 \) and \( \bar{X}_\alpha^- > 0 \), the MLE’s of \( \sigma \) and \( \kappa \) are given by

\[
\hat{\sigma}_n = \left[ \alpha (\bar{X}_\alpha^+ \bar{X}_\alpha^-)^{\frac{\alpha}{2(\alpha+1)}} \left( [\bar{X}_\alpha^+]^{\frac{1}{\alpha+1}} + [\bar{X}_\alpha^-]^{\frac{1}{\alpha+1}} \right) \right]^\frac{1}{\alpha}
\]

and

\[
\hat{\kappa}_n = \left[ \frac{\bar{X}_\alpha^-}{\bar{X}_\alpha^+} \right]^{\frac{1}{2(\alpha+1)}},
\]

where \( \bar{X}_\alpha^+ \) and \( \bar{X}_\alpha^- \) are as in (3.2).

**Remark 3.5.** Note that for \( \alpha = 1 \), these estimators reduce to the MLE’s for the Laplace case (see, e.g., Kotz et al. (2001), Theorem 3.5.2). Also observe that, as in the Laplace case, we obtain explicit expressions for the MLE’s, unlike the case where \( \sigma \) is known, where one needs to use a numerical search to obtain the MLE of \( \kappa \).

**Remark 3.6.** We note a similarity of estimators (3.13) - (3.14) with the expressions for the parameters \( \sigma \) and \( \kappa \) that appear in the maximum entropy Property 2.4. In fact, the estimators obtained by maximization of the entropy are exactly the same as the MLE’s under the \( \mathcal{EP} \) model. More precisely, suppose that we have a random sample \( X_1, \ldots, X_n \) from some continuous distribution on \( \mathbb{R} \), and want to recover the distribution of the \( X_i \)'s by assuming the maximum entropy distribution under the conditions

\[
EX^{<\alpha>} = c_1 \quad \text{and} \quad E|X|^\alpha = c_2.
\]

If \( c_1 \) and \( c_2 \) are approximated from the data by the method of moments:

\[
c_1 = \frac{1}{n} \sum_{i=1}^{n} X_i^{<\alpha>} = \bar{X}_\alpha^+ - \bar{X}_\alpha^-, \quad c_2 = \frac{1}{n} \sum_{i=1}^{n} |X_i|^\alpha = \bar{X}_\alpha^+ + \bar{X}_\alpha^-,
\]

then by Proposition 2.4 we obtain the \( \mathcal{EP}_\alpha(0, \sigma, \kappa) \) distribution with the values of \( \kappa \) and \( \sigma \) given by (2.25) and (2.26). If we substitute (3.15) into (2.25) and (2.26), we obtain the MLE’s (3.13) - (3.14)! Thus, under the \( \mathcal{EP} \) model, the MLE’s and the estimators obtained by the maximum entropy principle, coincide.
The following result provides the consistency and asymptotic normality of the MLE’s.

**Proposition 3.7.** Let $X_1, \ldots, X_n$ be i.i.d. variables from the $\mathcal{E}P_\alpha(0, \sigma, \kappa)$ distribution with unknown values of $\sigma$ and $\kappa$. Then, the MLE’s of $\sigma$ and $\kappa$ given in Proposition 3.4 are

(i) Consistent;

(ii) Asymptotically bivariate normal, with the asymptotic covariance matrix

\[
\Sigma_{MLE} = \frac{\sigma^2 (1 + \kappa^2)^2}{4\alpha(\alpha + 1)} \begin{bmatrix}
\frac{\alpha}{\kappa^2} + \frac{4}{(1+\kappa^2)^2} & -\frac{\alpha}{\sigma \kappa} \kappa^2 - \frac{1}{\kappa^2 + 1} \\
-\frac{\alpha}{\sigma \kappa} \kappa^2 - \frac{1}{\kappa^2 + 1} & \frac{\alpha}{\sigma^2}
\end{bmatrix},
\]

(iii) Asymptotically efficient, that is, the above asymptotic covariance matrix coincides with the inverse of the Fisher information matrix.

3.5. **Case 4: The values of $\sigma$, $\kappa$ and $\alpha$ are unknown.** The estimation of $\alpha$ requires a numerical search. First, fix $\alpha$ and substitute the expressions for $\hat{\sigma}_n$ and $\hat{\kappa}_n$ into the likelihood function, to obtain

\[
L = \left( \frac{\alpha}{\Gamma\left(\frac{1}{\alpha}\right)} \frac{\hat{\kappa}_n}{1 + \hat{\kappa}_n^2} \right)^n \exp\left( -\frac{\hat{\kappa}_n^2}{\hat{\kappa}_n^2} \frac{1}{\hat{\sigma}_n^2} + \frac{1}{\hat{\kappa}_n} \frac{1}{\hat{\sigma}_n^2} \right)
\]

This leads to the problem of maximizing the function

\[
L(\alpha) = \left[ \frac{\alpha}{\Gamma\left(\frac{1}{\alpha}\right)} \frac{\hat{\kappa}(\alpha)}{1 + [\hat{\kappa}(\alpha)]^2} \frac{1}{\hat{\sigma}(\alpha)} \right]^n \exp\left( -\frac{1}{\alpha} \right)
\]

with respect to $\alpha > 0$, where $\hat{\kappa}(\alpha)$ is given by the right-hand-side of the expression for $\hat{\kappa}_n$ and $\hat{\sigma}(\alpha)$ is given by the right-hand-side of the expression for $\hat{\sigma}_n$. When optimizing function $L$ in (3.17) numerically, one should restrict the range of $\alpha$’s to a finite interval, typically including $\alpha = 1$ and $\alpha = 2$ corresponding to the Laplace and normal cases.
4. Applications

In this section we present skew exponential power models of currency exchange rates. Although there is no agreement regarding the best theoretical model, a general consensus is that currency exchange rates are *leptokurtic* - their empirical distributions are fat-tailed with sharp peaks at the origin. It is also generally accepted that the currency exchange rates are increasingly leptokurtic with decreasing time intervals: while daily changes have fat tails, quarterly changes are nearly normal. Some authors also believe that currency exchange rates are asymmetric (see, e.g., So (1987)).

Various distributions were proposed to describe exchange rates, including stable Paretian (see, e.g., Westerfield (1977), McFarland et al. (1982, 1987), So (1987), Koedijk et al. (1990), Nolan (2001)), Student-t distribution (see, e.g., Boothe and Glassman (1987), Koedijk et al. (1990)), mixture of normals (see, e.g., Boothe and Glassman (1987), Tucker and Pond (1988)), and asymmetric Laplace (see Kozubowski and Podgórski (2001)).

In a recent study, Chenyao et al. (1996) examined the goodness of fit of normal, stable Paretian, Student-t, mixture of two normals, and (double) Weibull distributions to currency exchange rates, reporting Weibull model to be superior. Their estimates of the shape parameter for the Weibull distribution were very close to one - corresponding to the Laplace model. This led Kozubowski and Podgórski (2001) to fit skew Laplace distribution to currency exchange data. Here, we consider the EP models, and compare their fit with those of Laplace and normal distributions.

Our data is the same as that previously considered in Nolan (2001) and Kozubowski and Podgórski (2001). It consist of daily currency exchange rates, quoted in U.K. pounds, for fifteen currencies covering the period from 2 January 1980 to 21 May 1996. The variable of interest is the logarithm of the price ratio for two consecutive days. The data were transformed accordingly, yielding \( n = 4274 \) values for each currency. The summary statistics for the transformed data appear in Table 1, including estimates of the coefficients of skewness and kurtosis. We use the maximum likelihood estimators developed in Section 3 to fit the EP models to the data. In the numerical search for \( \alpha \) we restricted the range to the interval \([0.5, 4]\). Estimates for all parameters are presented in Table 2. Next, we compare the fit of exponential power, asymmetric Laplace and normal models to the
currency data. In Table 3, we list the values of the Kolmogorov-Smirnov distance between the data and the three fitted distributions. We conclude that the normal distribution performs distinctively worse than the other two models in all fifteen cases, while there is only a slight advantage of the EP over the AL fit.

Remark 4.1. In our analysis we assumed that the mode of the distribution is at zero. This is a reasonable assumption when the data consists of logarithmic growth rates such as currency exchange rates, stock-price changes, interest rate changes, and others. We have checked this assumption by estimating the mode of the fifteen currencies (using a non-parametric estimator of Andrews et al. (1972)). In all cases the mode was approximately zero. In case when the mode is unknown, one can estimate the mode using one of the non-parametric methods (see, e.g., Bickel (2002) and Vieu (1996) for several estimation methods) and shift the data prior to computing the estimators presented in this paper.
<table>
<thead>
<tr>
<th>Currency</th>
<th>$\alpha$</th>
<th>$\sigma$</th>
<th>$\kappa$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Australia</td>
<td>0.98162996</td>
<td>0.00532231</td>
<td>1.00144388</td>
</tr>
<tr>
<td>Austria</td>
<td>0.95173896</td>
<td>0.00332828</td>
<td>1.01581415</td>
</tr>
<tr>
<td>Belgium</td>
<td>0.85323470</td>
<td>0.00322903</td>
<td>1.00550617</td>
</tr>
<tr>
<td>Canada</td>
<td>1.14060566</td>
<td>0.00570844</td>
<td>1.00613741</td>
</tr>
<tr>
<td>Denmark</td>
<td>1.02345566</td>
<td>0.00331719</td>
<td>1.00339333</td>
</tr>
<tr>
<td>France</td>
<td>0.88418297</td>
<td>0.00263442</td>
<td>1.00605617</td>
</tr>
<tr>
<td>Germany</td>
<td>0.83953550</td>
<td>0.00243369</td>
<td>1.00260847</td>
</tr>
<tr>
<td>Italy</td>
<td>0.85227561</td>
<td>0.00274836</td>
<td>0.98859803</td>
</tr>
<tr>
<td>Japan</td>
<td>1.01787442</td>
<td>0.00402650</td>
<td>1.03063535</td>
</tr>
<tr>
<td>Netherlands</td>
<td>0.87171652</td>
<td>0.00263442</td>
<td>1.01375829</td>
</tr>
<tr>
<td>Norway</td>
<td>0.99839779</td>
<td>0.00331719</td>
<td>1.00339333</td>
</tr>
<tr>
<td>Spain</td>
<td>0.91568901</td>
<td>0.00304287</td>
<td>0.98985581</td>
</tr>
<tr>
<td>Sweden</td>
<td>0.86119393</td>
<td>0.00269830</td>
<td>0.99911649</td>
</tr>
<tr>
<td>Switzerland</td>
<td>0.98777399</td>
<td>0.00372373</td>
<td>1.01871544</td>
</tr>
<tr>
<td>U.S.</td>
<td>1.05024643</td>
<td>0.00513699</td>
<td>1.00991064</td>
</tr>
</tbody>
</table>

Table 2. Estimated values of $\alpha$, $\sigma$ and $\kappa$ of the fitted $\mathcal{E}P_\alpha(0,\sigma,\kappa)$ distributions.

5. Technical results and proofs

Lemma 5.1. Let $\gamma > -1$, $\alpha > 0$, $\sigma > 0$, $\kappa > 0$, $\theta \in \mathbb{R}$ . Then

\[
\int_{-\infty}^{\theta} (\theta - x)^\gamma \exp \left( -\frac{1}{\sigma^\alpha \kappa^\alpha} (\theta - x)^\alpha \right) dx = \frac{\sigma^{\gamma+1} \kappa^{\gamma+1}}{\alpha} \Gamma \left( \frac{\gamma + \frac{1}{\alpha}}{\alpha} \right)
\]

and

\[
\int_{\theta}^{\infty} (x - \theta)^\gamma \exp \left( -\frac{\kappa^\alpha}{\sigma^\alpha} (x - \theta)^\alpha \right) dx = \frac{\sigma^{\gamma+1}}{\alpha \kappa^{\gamma+1}} \Gamma \left( \frac{\gamma + \frac{1}{\alpha}}{\alpha} \right).
\]

Proof. Let us start with (5.1). Substituting

\[ u = \left[ \frac{\theta - x}{\sigma \kappa} \right]^\alpha, \quad du = -\alpha \frac{1}{\sigma^\alpha \kappa^\alpha} (\theta - x)^{\alpha-1} dx, \]

\[ dx = -\frac{1}{\alpha \sigma^\alpha \kappa^\alpha} (\theta - x)^{\alpha-1} du. \]

Therefore, we have

\[
\int_{-\infty}^{\theta} (\theta - x)^\gamma \exp \left( -\frac{1}{\sigma^\alpha \kappa^\alpha} (\theta - x)^\alpha \right) dx = \frac{\sigma^{\gamma+1} \kappa^{\gamma+1}}{\alpha} \Gamma \left( \frac{\gamma + \frac{1}{\alpha}}{\alpha} \right).
\]

The proof of (5.2) is similar.
the left-hand-side becomes

$$\frac{\sigma^{\gamma+1}\kappa^{\gamma+1}}{\alpha} \int_0^\infty u^{\frac{\gamma+1}{\alpha} - 1} \exp(-u)du.$$ 

Since the integral above is the gamma function $\Gamma \left(\frac{\gamma+1}{\alpha}\right)$ (provided $\gamma > -1$), we obtain the right-hand-side of (5.1). The proof of (5.2) is similar. \qed

**Proof of Proposition 2.4.** Similarly to the Laplace case $\alpha = 1$ (see Kotz et al. (2002)), we find that the maximum entropy is attained by the density

$$p(x) = e^{a_0} e^{\alpha_1 x^{\alpha_1} + \alpha_2 |x|^{\alpha_2}} = e^{a_0} \begin{cases} e^{(a_1 + a_2)x^{\alpha_1}} & \text{if } x \geq 0 \\ e^{-(a_1 - a_2)(-x)^{\alpha_2}} & \text{if } x < 0 \end{cases},$$

provided that (5.3) integrates to 1 on $\mathbb{R}$ and satisfies conditions (2.21). Thus, it is enough to find the constants $a_0, a_1, a_2$ for which these conditions hold. First, note that the integrability of $p$ implies that $a_1 + a_2 < 0$ and $a_1 - a_2 > 0$, so that $a_2 < 0$. Write

$$a_1 = \frac{1}{2\sigma^\alpha} \left( \frac{1}{\kappa^\alpha} - \kappa^\alpha \right) \in \mathbb{R} \quad \text{and} \quad a_2 = -\frac{1}{2\sigma^\alpha} \left( \frac{1}{\kappa^\alpha} + \kappa^\alpha \right) < 0.$$
for some $\sigma > 0$ and $\kappa > 0$, so that the density (5.3) takes the form

$$p(x) = e^{a_0} \begin{cases} 
  e^{-\frac{a_0}{\sigma} |x|^\alpha}, & \text{if } x \geq 0 \\
  e^{-\frac{1}{\pi^\alpha \sigma^\alpha} |x|^\alpha}, & \text{if } x < 0.
\end{cases}$$

Comparing the above with (2.1), we conclude that the entropy is indeed maximized by an EPA distribution.

Next, we find the parameters $\sigma$ and $\kappa$. Using Proposition 2.2 and Lemma 5.1 we obtain

$$E|X|^\alpha = \frac{\sigma^\alpha}{\alpha} \frac{\kappa}{1 + \kappa^2} \left\{ \frac{1}{\kappa^{\alpha+1}} + \kappa^{\alpha+1} \right\} = c_2,$$

and

$$EX^{<\alpha>} = \frac{\sigma^\alpha}{\alpha} \frac{\kappa}{1 + \kappa^2} \left\{ \frac{1}{\kappa^{\alpha+1}} - \kappa^{\alpha+1} \right\} = c_1.$$

Solving these equations for $\kappa$ and $\sigma$ leads to (2.25) and (2.26) after some simple algebra. The actual value of the entropy (2.27) follows from 2.20.

Proof of Proposition 3.1. Write

$$\hat{\sigma}_n = h \left( \frac{1}{n} \sum_{i=1}^n W_i \right),$$

where

$$W_i = \kappa^\alpha (X_i^+)\alpha + \frac{1}{\kappa^\alpha} (X_i^-)\alpha = \begin{cases} 
  \kappa^\alpha X_i^\alpha, & \text{if } X_i \geq 0, \\
  \frac{1}{\kappa^\alpha} (-X_i)\alpha, & \text{if } X_i < 0,
\end{cases}$$

and

$$h(x) = \frac{1}{\alpha} x^{\frac{1}{\alpha}}, \quad x \geq 0,$$

and note that the variables $W_i$ are i.i.d. with mean $EW_i = \frac{\sigma^\alpha}{\alpha}$ and variance $Var W_i = \frac{\sigma^{2\alpha}}{\alpha}$ (by Lemma 5.1). Thus, by the law of large numbers and the continuity of $h$, we have

$$\hat{\sigma}_n = h \left( \frac{1}{n} \sum_{i=1}^n W_i \right) \overset{d}{\to} h \left( \frac{\sigma^\alpha}{\alpha} \right) = \alpha^{\frac{1}{\alpha}} \left( \frac{\sigma^\alpha}{\alpha} \right)^{\frac{1}{\alpha}} = \sigma.$$

This proves the consistency. Next, by the central limit theorem,

$$n^{\frac{1}{2}} \left( \frac{1}{n} \sum_{i=1}^n W_i - \frac{\sigma^\alpha}{\alpha} \right) \overset{d}{\to} N \left( 0, \frac{\sigma^{2\alpha}}{\alpha} \right),$$
where the right-hand-side denotes a normal variable with mean 0 and variance $\sigma^2/\alpha$. Thus, by the continuity of $h$ and standard large sample theory results (see, e.g., Serfling (1980)), we have

$$
(5.8) \quad n^{1/2} \left( h \left( \frac{1}{n} \sum_{i=1}^{n} W_i \right) - h \left( \frac{\sigma^\alpha}{\alpha} \right) \right) \xrightarrow{d} N(0, \eta^2),
$$

where

$$
(5.9) \quad \eta^2 = \left[ h'(x)|_{x=\frac{\sigma^\alpha}{\alpha}} \right]^2 \frac{\sigma^{2\alpha}}{\alpha} = (\sigma^{1-\alpha})^2 \left( \frac{\sigma^{2\alpha}}{\alpha} \right) = \frac{\sigma^2}{\alpha},
$$

which proves Part (ii). The asymptotic efficiency is obtained by noting that the asymptotic variance $\sigma^2/\alpha$ is the same as the reciprocal of the Fisher information $I(\sigma)$ given by the first entry in the Fisher information matrix (3.4).

*Proof of Proposition 3.2.* For simplicity, denote

$$
A = \frac{1}{\sigma^\alpha \bar{x}_\alpha^+} \quad \text{and} \quad B = \frac{1}{\sigma^\alpha \bar{x}_\alpha^-}.
$$

We shall consider several cases.

*Case 1: $A > 0, B = 0$ (all sample values are greater than or equal to zero).* Here, we need to maximize the function

$$
Q(\kappa) = \log \kappa - \log(1 + \kappa^2) - A \kappa^\alpha
$$

with respect to $\kappa \in (0, \infty)$. The derivative of $Q$ is

$$
u(\kappa) = \frac{dQ}{d\kappa} = \frac{1}{\kappa(1 + \kappa^2)} (u_1(\kappa) - u_2(\kappa)),
$$

where

$$u_1(\kappa) = 1 - \kappa^2 \quad \text{and} \quad u_2(\kappa) = A \alpha \kappa^\alpha(1 + \kappa^2).
$$

Note that on the interval $(1, \infty)$, the derivative $u(\kappa)$ is negative and thus the function $Q(\kappa)$ is decreasing. On the interval $(0, 1)$, the function $u_1$ is strictly decreasing with $u_1(0) = 1$ and $u_1(1) = 0$, while the function $u_2$ is strictly increasing with $u_2(0) = 0$ and $u_2(1) = 2\alpha A > 0$. Consequently, there is a unique $\hat{\kappa} \in (0, 1)$ such that $u_1(\hat{\kappa}) = u_2(\hat{\kappa})$.
and the derivative $u(\kappa)$ is positive on $(0, \hat{\kappa})$ and negative on $(\hat{\kappa}, \infty)$. This quantity is the unique solution of the equation

$$1 - \kappa^2 - A\alpha \kappa^\alpha (1 + \kappa^2) = 0,$$

which coincides with (3.7) if $B = 0$. This concludes Case 1.

**Case 2:** $A = 0$, $B > 0$ (*all sample values are less than or equal to zero*). Here, we need to maximize the function

$$Q(\kappa) = \log \kappa - \log(1 + \kappa^2) - \frac{B}{\kappa^\alpha},$$

whose derivative is

$$u(\kappa) = \frac{1}{\kappa} - \frac{2\kappa}{1 + \kappa^2} + \frac{\alpha B}{\kappa^{\alpha+1}}, \quad \kappa \in (0, \infty).$$

Consider $z = 1/\kappa \in (0, \infty)$ and let

$$v(z) = u \left( \frac{1}{z} \right) = \frac{2z}{z^2 + 1} + \alpha B z^{\alpha+1} = \frac{z}{z^2 + 1} (v_1(z) - v_2(z)),$$

where

$$v_1(z) = \alpha B z^\alpha (1 + z^2) \quad \text{and} \quad v_2(z) = 1 - z^2.$$

On the interval $(1, \infty)$, we have $v(z) > 0$ so that $u(\kappa) > 0$ for $\kappa \in (0, 1)$. Consider the values $z \in (0, 1)$. The function $v_1$ is increasing on $(0, 1)$ with $v_1(0) = 0$, $v_1(1) = 2\alpha B > 0$. The function $v_2$ is decreasing on $(0, 1)$ with $v_2(0) = 1$ and $v_2(1) = 0$. Clearly, there is a unique $z^* \in (0, 1)$ such that $v(z)$ is positive on $(z^*, \infty)$ and negative on $(0, z^*)$. Consequently, $u(\kappa)$ is positive on $(0, \hat{\kappa})$ and negative on $(\hat{\kappa}, \infty)$, where $\hat{\kappa} = 1/z^* \in (1, \infty)$. The value $z^*$ satisfies the equation

$$\alpha B z^\alpha (1 + z^2) - (1 - z^2) = 0,$$

so that $\hat{\kappa} = 1/z^*$ satisfies

$$\alpha B (\kappa^2 + 1) - \kappa^\alpha (\kappa^2 - 1) = 0,$$

which is equivalent to (3.7) with $A = 0$. This concludes Case 2.

**Case 3:** $0 < B < A$. Write the derivative of $Q$ as

$$\frac{dQ(\kappa)}{d\kappa} = \frac{\kappa^\alpha (1 - \kappa^2) + \alpha (1 + \kappa^2) (B - A\kappa^2 \alpha)}{(1 + \kappa^2) \kappa^{\alpha+1}}.$$
Note that on the interval \((0, (B/A)^{\frac{1}{2\alpha}})\), the derivative is positive while on the interval \((1, \infty)\) the derivative is negative. Therefore, the function \(Q\) attains its global maximum value somewhere within the interval \([(B/A)^{\frac{1}{2\alpha}}, 1]\). We claim that this maximum occurs at
\[
\hat{\kappa}_n \in \left(\left(\frac{B}{A}\right)^{\frac{1}{2\alpha}}, 1\right).
\] (5.11)

Indeed, write
\[
\frac{dQ(\kappa)}{d\kappa} = h_1(\kappa) - h_2(\kappa),
\]
where
\[
h_1(\kappa) = \frac{1}{\kappa} - \frac{2\kappa}{1 + \kappa^2} \quad \text{and} \quad h_2(\kappa) = \frac{B\alpha}{\kappa^{\alpha+1}} - A\alpha\kappa^{\alpha-1}.
\]
Observe that the function \(h_1\) is decreasing on \((0, \kappa_1)\) and increasing on \((\kappa_1, \infty)\), where
\[
\kappa_1 = \sqrt{2 + \sqrt{5}} > 1.
\]
Moreover, we have
\[
\lim_{\kappa \to 0^+} h_1(\kappa) = \infty, \quad h_1(1) = 0, \quad h_1(\kappa_1) < 0, \quad \lim_{\kappa \to \infty} h_1(\kappa) = 0.
\]
On the other hand, the function \(h_2\) is decreasing on the interval \((0, 1)\) with
\[
h_2\left(\left(\frac{B}{A}\right)^{\frac{1}{2\alpha}}\right) = 0 \quad \text{and} \quad h_2(1) < 0.
\]
In view of the above facts, we conclude that there exists a unique \(\hat{\kappa}_n\) satisfying (5.11) that maximizes the function \(Q\) on \((0, \infty)\). This concludes Case 3.

**Case 4:** \(0 < A = B\). Here, the derivative (5.10) is positive for \(\kappa \in (0, 1)\) and negative for \(\kappa \in (1, \infty)\), so that \(\hat{\kappa}_n = 1\) is a unique MLE of \(\kappa\).

**Case 5:** \(0 < A < B\). This is the most complex case. Clearly, for \(\kappa \in (0, 1)\) the derivative (5.10) is positive and for \(\kappa > (B/A)^{\frac{1}{2\alpha}}\), the derivative (5.10) is negative. Consequently, the function \(Q\) attains its maximum value somewhere in the interval \([1, (B/A)^{\frac{1}{2\alpha}}]\). This value should occur where the derivative is zero,
\[
\kappa^\alpha(1 - \kappa^2) + \alpha(1 + \kappa^2)(B - A\kappa^{2\alpha}) = 0.
\] (5.12)
We show below that this equation has a unique solution. Denoting
\[ y = \kappa^\alpha, \quad x = \kappa^2, \]
we observe that (5.12) is equivalent to the system
\[
\begin{align*}
    y(1 - x) + \alpha(1 + x)(B - Ay^2) &= 0 \\
    y &= x^{\frac{\alpha}{2}}.
\end{align*}
\]
We claim that the system (5.13) admits a unique solution in the region \( x, y > 1 \). Indeed, the first equation is quadratic in \( y \) and can be solved easily for \( y \) in terms of \( x \) leading to
\[
y = \frac{1}{2A\alpha} \left( \sqrt{\left(\frac{x-1}{x+1}\right)^2 + 4AB\alpha^2 - \frac{x-1}{x+1}} \right) = h(u(x)),
\]
where this time
\[
u(x) = \frac{x-1}{x+1}, \quad x \in [1, \infty),
\]
and
\[
h(u) = \frac{1}{2A\alpha} (\sqrt{u^2 + 4AB\alpha^2} - u) = \frac{2B\alpha}{\sqrt{u^2 + 4AB\alpha^2} + u}, \quad u \in (0, \infty).
\]
Note that the function \( u \) is increasing on \((1, \infty)\) with \( u(1) = 0 \) and \( \lim_{x \to \infty} u(x) = 1 \), while the function \( h \) is decreasing on \((0, \infty)\) with
\[
h(0) = \sqrt{\frac{B}{A}} > 1.
\]
Thus, the function \( v(x) = h(u(x)) \) is decreasing on \((1, \infty)\) with
\[
v(1) = h(u(1)) = h(0) = \sqrt{\frac{B}{A}} > 1.
\]
On the other hand, the function \( y = x^{\frac{\alpha}{2}} \) is increasing on \((1, \infty)\) with \( y = 1 \) when \( x = 1 \) and \( \lim_{x \to \infty} x^{\frac{\alpha}{2}} = \infty \). Consequently, the system of equations (5.13), which can be written as
\[
\begin{align*}
    y = v(x) \\
    y &= x^{\frac{\alpha}{2}},
\end{align*}
\]
will have exactly one solution in the region \( x, y > 1 \). This concludes Case 5.
Lemma 5.2. Let $X$ have an $E\mathcal{P}_\alpha(0, \sigma, \kappa)$ distribution, and let

$$W = [W_1, W_2]' = [(X^+)^\alpha, (X^-)^\alpha]'$$

Then the mean vector and the covariance matrix of $W$ are

\[(5.14)\quad EW = \begin{bmatrix} \frac{\sigma^\alpha}{\alpha \kappa^\alpha (1 + \kappa^2)} \\ \frac{\sigma^\alpha \kappa^{\alpha+2}}{\alpha (1 + \kappa^2)} \end{bmatrix}\]

and

\[(5.15)\quad \Sigma_W = \left(\frac{\sigma}{\kappa}\right)^2 \frac{1}{\alpha^2} \left(\frac{1}{1 + \kappa^2}\right)^2 \begin{bmatrix} (\alpha + 1)(1 + \kappa^2) - 1 & -\kappa^{2(\alpha+1)} \\ -\kappa^{2(\alpha+1)} & \kappa^{4(\alpha+1)} \left[\frac{1 + \kappa^2}{\kappa^2} (1 + \alpha) - 1\right] \end{bmatrix},\]

respectively.

**Proof.** The formulas follow easily from Proposition 2.2, see Ayebo (2002) for details. □

**Proof of Proposition 3.3.** Since the MLE $\hat{\kappa}_n$ is a unique solution of (3.8), it can be written as

$$\hat{\kappa}_n = H(\bar{x}_a^+, \bar{x}_a^-),$$

where $H(\cdot, \cdot)$ is a continuous and differentiable function satisfying the equation

$$F(\bar{x}_a^+, \bar{x}_a^-, H(\bar{x}_a^+, \bar{x}_a^-)) = 0$$

with

$$F(y_1, y_2, y_3) = \frac{\sigma^\alpha}{\alpha^2} y_3 (1 - y_3^2) - \frac{\alpha}{\sigma^\alpha} y_1 y_3^2 (1 + y_3^2) + \frac{\alpha}{\sigma^\alpha} y_2 (1 + y_3^2).$$

(i) To establish the consistency of the MLE, note that by the law of large numbers, we have

$$\left(\bar{X}_a^+, \bar{X}_a^-\right)' \overset{d}{\to} (\mu_1, \mu_2)' = \left[\frac{\sigma^\alpha}{\alpha \kappa^\alpha (1 + \kappa^2)} - \frac{\sigma^\alpha \kappa^{2\alpha}}{\alpha (1 + \kappa^2)}\right],'$$

see Lemma 5.2. Thus, by the continuity of $H$, we obtain

$$\hat{\kappa}_n = H(\bar{X}_a^+, \bar{X}_a^-)' \overset{d}{\to} H(\mu_1, \mu_2) = \kappa,$$
since the quantities
\[(5.16)\]
\[y_1 = \mu_1 = \frac{\sigma^\alpha}{\alpha \kappa^\alpha (1 + \kappa^2)}, \quad y_2 = \mu_2 = \frac{\sigma^\alpha \kappa^{2\alpha}}{\alpha (1 + \kappa^2)}, \quad y_3 = \kappa\]
satisfy the equation \(F(y_1, y_2, y_3) = 0\).

(ii) According to the classical (bivariate) central limit theorem, we have
\[\sqrt{n}[(\bar{X}_\alpha^+, \bar{X}_\alpha^-)' - (\mu_1, \mu_2)'] \to N(0, \Sigma_W),\]
where the right-hand-side denotes a bivariate normal variable with mean zero and variance-covariance matrix \(\Sigma_W\) given by (5.15). Now, by standard large sample theory results (see, e.g., Serfling (1980)) it follows that as \(n \to \infty\), we have
\[\sqrt{n}[H(\bar{X}_\alpha^+, \bar{X}_\alpha^-)' - H(\mu_1, \mu_2)] \to N(0, \Omega),\]
where \(\Omega = D\Sigma_W D'\) and \(D\) is the vector of partial derivatives
\[(5.17)\]
\[D = \left[\frac{\partial H}{\partial y_1}, \frac{\partial H}{\partial y_2}\right]_{(y_1, y_2) = (\mu_1, \mu_2)}.\]

Since the function \(H\) satisfies the equation
\[F(y_1, y_2, H(y_1, y_2)) = 0,\]
we have
\[(5.18)\]
\[\frac{\partial}{\partial y_1} H(y_1, y_2) = \frac{\partial}{\partial y_3} F(y_1, y_2, y_3)|_{y_3 = H(y_1, y_2)},\]
\[(5.19)\]
\[\frac{\partial}{\partial y_2} H(y_1, y_2) = \frac{\partial}{\partial y_3} F(y_1, y_2, y_3)|_{y_3 = H(y_1, y_2)}.\]

Routine calculations produce
\[\frac{\partial}{\partial y_1} F(y_1, y_2, y_3) = -\frac{\alpha}{\sigma^\alpha} y_3^{2\alpha} (1 + y_3^2),\]
\[\frac{\partial}{\partial y_2} F(y_1, y_2, y_3) = \frac{\alpha}{\sigma^\alpha} (1 + y_3^2),\]
\[\frac{\partial}{\partial y_3} F(y_1, y_2, y_3) = \alpha y_3^{\alpha - 1} - (\alpha + 2) y_3^{\alpha + 1} - \frac{\alpha}{\sigma^\alpha} y_1 (2\alpha y_3^{2\alpha - 1} + (2\alpha + 2) y_3^{2\alpha + 1}) + \frac{2\alpha}{\sigma^\alpha} y_2 y_3.\]

Substituting (5.16) into the above derivatives, we obtain
\[\frac{\partial}{\partial y_1} F(y_1, y_2, y_3)|(y_1, y_2, y_3) = (\mu_1, \mu_2, H(\mu_1, \mu_2)) = -\frac{\alpha}{\sigma^\alpha} \kappa^{2\alpha} (1 + \kappa^2),\]
$$\frac{\partial}{\partial y} F(y_1, y_2, y_3) |_{(y_1, y_2, y_3) = (\mu_1, \mu_2, H(\mu_1, \mu_2))} = \frac{\alpha}{\sigma^2} (1 + \kappa^2),$$

$$\frac{\partial}{\partial y_3} F(y_1, y_2, y_3) |_{(y_1, y_2, y_3) = (\mu_1, \mu_2, H(\mu_1, \mu_2))} = -\alpha \kappa^{\alpha-1} - \kappa^{\alpha+1} (4 + 2 \alpha) - \alpha \kappa^{\alpha+3}.$$ 

Consequently, the vector of partial derivatives (5.17) takes the form

$$D = \begin{bmatrix} \frac{\partial H}{\partial y_1} & \frac{\partial H}{\partial y_2} \end{bmatrix} |_{(y_1, y_2) = (\mu_1, \mu_2)} = \frac{\alpha}{\sigma^2} \frac{(1 + \kappa^2)^2}{\alpha \kappa^{\alpha-1} + (4 + 2 \alpha) \kappa^{\alpha+1} + \alpha \kappa^{\alpha+3}} [\kappa^{2\alpha}, -1].$$

After some algebra, we find that the product $\Omega = D\Sigma_W D'$ coincides with (3.9). This concludes the proof of asymptotic normality.

(iii) To establish the asymptotic efficiency, note that the asymptotic variance is the reciprocal of the Fisher information (see Fisher information matrix (3.4)).

**Proof of Proposition 3.4.** When we substitute $\sigma(\kappa)$ given by (3.11) into $Q(\kappa, \sigma)$ in (3.10), we obtain the following function to be maximized with respect to $\kappa \in (0, \infty)$:

$$u(\kappa) = 2 \log \kappa - \log(1 + \kappa^2) - \frac{1}{\alpha} \log(\kappa^{2\alpha} \bar{x}_\alpha^+ + \bar{x}_\alpha^-).$$

Setting

$$\frac{du(\kappa)}{d\kappa} = \frac{2}{\kappa} - \frac{2 \kappa}{1 + \kappa^2} - \frac{2 \alpha \kappa^{2\alpha-1} \bar{x}_\alpha^+}{\alpha \kappa^{2\alpha} \bar{x}_\alpha^+ + \bar{x}_\alpha^-} > 0$$

and solving for $\kappa$, we obtain

$$\kappa^{2\alpha+2} < \frac{\bar{x}_\alpha^-}{\bar{x}_\alpha^+}.$$ 

Thus, the function $u$ increases on $(0, \hat{\kappa}_n)$ and decreases on $(\hat{\kappa}_n, \infty)$, where $\hat{\kappa}_n$ is given by (3.14). Substituting this value into (3.11), after some algebra we obtain the MLE of $\sigma$ as given by (3.13).

**Proof of Proposition 3.7.** We start with Part (i). Let $W$ be the bivariate vector defined in Lemma 5.2. Then, by the law of large numbers,

$$\frac{1}{n} \sum_{i=1}^n W_i \xrightarrow{d} EW,$$

where $EW$ is given by (5.14). Note that $(\hat{\sigma}_n, \hat{\kappa}_n) = G(\frac{1}{n} \sum_{i=1}^n W_i)$, where

$$G(x, y) = (G_1(x, y), G_2(x, y))$$

with

$$G_1(x, y) = \left[ \alpha (xy)^{\frac{1}{2(\alpha+1)}} \left( \frac{1}{x^{\alpha+1}} + \frac{1}{y^{\alpha+1}} \right) \right]^{\frac{1}{\alpha}}$$

and

$$G_2(x, y) = \left( \frac{y}{x} \right)^{\frac{1}{2(\alpha+1)}}.$$
Then, since $G$ is continuous, we have,

$$G \left( \frac{1}{n} \sum_{i=1}^{n} W_i \right) \xrightarrow{d} G(EW) = (G_1(EW), G_2(EW)).$$

Since the right-hand-side above simplifies to $(\sigma, \kappa)$, we obtain the consistency.

We now move to Part (ii). By (bivariate) Central Limit Theorem, we have

$$\sqrt{n} \left( \frac{1}{n} \sum_{i=1}^{n} W_i - EW \right) \xrightarrow{d} N(0, \Sigma_W),$$

where $EW$ is given by (5.14) and the right-hand-side is a bivariate normal distribution with mean zero and variance $\Sigma_W$ given by (5.15). Then, by standard large sample theory results (see, e.g., Serfling (1980), p. 122), we have

$$\sqrt{n} \left( G \left( \frac{1}{n} \sum_{i=1}^{n} W_i \right) - G(EW) \right) \xrightarrow{d} N(0, \Omega),$$

where $\Omega = D\Sigma_W D'$ and

$$D = \left[ \frac{\partial G_i}{\partial x_j} \right]_{i,j=1}^{x_1,x_2=EW} \left[ \frac{\partial G_i}{\partial x_j} \right]^{2}_{i,j=1}$$

is the matrix of partial derivatives of the vector-valued function $G$ given in (5.20). Rather lengthy computations (facilitated by logarithmic differentiation) produce

$$D = \kappa^\alpha \frac{1}{\sigma'^\alpha} \left[ \begin{array}{cc} \frac{\sigma^2}{2} (1 + \kappa^2) + \sigma & \frac{\sigma^2}{2} \kappa^2 \frac{1}{\kappa^{2\alpha}} + \frac{\sigma}{\kappa^{2\alpha}} \\ -\kappa \left( 1 + \kappa^2 \right) \alpha & \kappa \frac{1 + \kappa^2}{\kappa^{2\alpha}} \frac{\alpha}{\kappa^{2\alpha}} \end{array} \right].$$

After straightforward but laborious matrix multiplications we find that $D\Sigma_W D'$ reduces to (3.16) given in the statement of the proposition.

Finally, to establish Part (iii), take the inverse of the Fisher information matrix (3.3) and verify that it coincides with the asymptotic covariance matrix $\Omega$.

**References**


