To appear in *Quantitative Finance*, Vol. 00, No. 00, Month 20XX, 1–25

Some analytical results on bivariate stable distributions with an application in operational risk

L. TAFAKORI†, M. BEE*‡ and A.R. SOLTANI §

†Department of Mathematical Sciences, School of Science, RMIT University, Australia
‡Department of Economics and Management, University of Trento, Italy
§Department of Statistics and Operations Research, Kuwait University, Kuwait and Department of Statistics, School of Science, Shiraz University, Shiraz, Iran

(*Received 00 Month 20XX; in final form 00 Month 20XX*)

The multivariate stable distributions are widely applicable as they can accommodate both skewness and heavy tails. Although one-dimensional stable distributions are well known, there are many open questions in the multivariate regime, since the tractability of the multivariate Gaussian universe does not extend to non-Gaussian multivariate stable distributions. In this work, we provide the Laplace transform of a bivariate stable distributions and its certain cut in the first quadrant. Given the lack of a closed-form likelihood function, we propose to estimate the parameters by means of Approximate Maximum Likelihood, a simulation-based method with desirable asymptotic properties. Simulation experiments and an application to truncated operational losses illustrate the applicability of the model.

Keywords: Characteristic function; One-sided Bivariate Laplace transform; Bivariate stable distributions; Approximate Maximum Likelihood.

1. Introduction

Classical tools for the modeling of financial data in risk management or portfolio optimization are dependent on the assumption that stock returns and risk factor returns have a normal distribution. However, both empirical evidences and further investigations by Mandelbrot (1997), Fama (1965), DuMouchel (1973), McCulloch (1986), Cheng and Rachev (1995), Nolan et al. (2001), Cont (2001) have well explained that many financial assets exhibit a number of features that contradict the normality assumption, such as asymmetry, skewness and heavy tails. One can use Elliptical distributions as one of the solutions. Multivariate Student-t and multivariate elliptical stable distributions are examples of elliptical distributions, which can describe the heavy-tailedness. Although elliptical distributions are often used for modeling of financial data, they cannot accommodate asymmetry and, hence, it can lead to an underestimation of risks. Frahm and Jaekel (2005) use the log-returns of the NASDAQ and S&P 500 indexes to illustrate asymmetry in financial data. Embrechts et al. (2015, Example 6.3) obtain similar results for the Dow Jones index. To overcome the aforementioned difficulties, two possible models are multivariate generalized hyperbolic distributions (Eberlein and Keller 1995, Embrechts et al. 2015) and multivariate (non-elliptical) stable distributions. We consider using the latter, which can describe skewed and heavy-tailed time series.

The well-known stable distribution was first introduced by Lévy (1925) during his investigations on the behaviour of sums of independent random variables. A univariate stable distribution is pa-

*Corresponding author. Email: marco.bee@unitn.it; ORCID: 0000-0002-9579-3650.

rameterized by four parameters which are location $[\gamma, -\infty < \gamma < \infty]$, scale $[\lambda, 0 < \lambda]$, skewness $[\beta, -1 \leq \beta \leq 1]$ and exponent $[\alpha, 0 < \alpha \leq 2]$. The lack of closed form for the stable density and its distribution function remain as major problems and drawbacks, but this difficulty can be solved by numerical approximation in the univariate case (Nolan 1997). For a general multivariate stable distribution, however, the situation is much more complex. Since the marginals do not have finite variance, one cannot define a covariance matrix in the usual way because none of the integrals would converge. Instead, the dependence structure of a stable distribution on $\mathbb{R}^d$ is determined by an arbitrary complicated spectral measure. This possible complexity of the dependence structure is one of the main reasons why multivariate stable distributions have not been employed in many applications, even though Embrechts et al. (2015) emphasize their application in market risk management.

The celebrated equation in Section 7.1 in Uchaikin and Zolotarev (1999) for the Laplace transform of a stable density function restricted to the positive real line has found useful applications in univariate truncated stable densities or cut-off stable densities, according to the terminology in Zolotarev (1986). In this work we extend the results for the truncated bivariate stable distributions. Let $X(\alpha, \beta, \gamma, \lambda)$ denotes a stable random variable where its density function is denoted by $q(x; \alpha, \beta, \gamma, \lambda)$, and $q(x; \alpha, \beta)$ stands for the case $\gamma = 0$ and $\lambda = 1$. The Laplace transform (one-sided) of $q(x; \alpha, \beta)$, $0 < \alpha \leq 2$, $-1 \leq \beta \leq 1$ is defined by

$$\tilde{q}(y; \alpha, \beta) = \int_0^\infty e^{-yx} q(x; \alpha, \beta) dx, \quad y \geq 0, \tag{1}$$

is given in Theorem 2.6.2 in Zolotarev (1986), for $\alpha \neq 1$:

$$\tilde{q}(y; \alpha, \beta) = \frac{1}{\pi} \int_0^\infty e^{-(yu)^\alpha} \frac{\sin(\pi\rho)}{u^2 + 2u\cos(\pi\rho) + 1} du, \ y \geq 0, \tag{2}$$

where $\rho = \frac{(1+\beta)}{2}$. Further, $\tilde{q}(y; \alpha, \beta, 0, \lambda) = \tilde{q}(\lambda^{1/\alpha} y; \alpha, \beta)$, where $\tilde{q}(y; \alpha, 1) = e^{y^\alpha}, \ \ 0 < \alpha < 1$, and for $\alpha > 1, \tilde{q}(y; \alpha, -1)$ is expressed in terms of the Mittage-Leffler function. The factor

$$g_\rho(u) = \frac{\sin(\pi\rho)}{\pi\rho(u^2 + 2u\cos(\pi\rho) + 1)}, \ \ u > 0, 0 < \rho < 1, \tag{3}$$

is the density function of a non-negative random variable that appears in characterizing geometric stable random variables (Kozubowski 2000, Tafakori and Soltani 2017). Further, Soltani and Tafakori (2013) introduce new heavy-tailed Cauchy-type distributions based on this factor on the positive real line to model continuous erratic data. Finally, theoretical contributions that include introducing new one-sided kernels linked to the generalized Linnik distribution and an extension for the mixture representation in Kotz and Ostrovskii (1996) are also considered.

A non-negative random variable X is said to be truncated strictly stable (TSS), or cut-off stable according to Zolotarev (1986), if it possesses the density function

$$q^\star(x; \alpha, \beta) = c(\alpha, \beta) q(x; \alpha, \beta), \ \ x \geq 0,$$

where

$$c(\alpha, \beta) = \frac{1}{1 - G(0, \alpha, \beta)},$$

and $q(x; \alpha, \beta)$, $G(x, \alpha, \beta)$ stand for the density of a stable random variable and its distribution function, respectively. By taking $y = 0$ in (2) and using (3), we observe that the normalization factor $c(\alpha, \beta)$ is equal to $1/\rho$. Therefore, the Laplace transform of a TSS random variable Y is given by

$$\mathbb{E}(e^{-yY}) = \frac{1}{\pi\rho}\int_0^\infty e^{-(yu)^\alpha}\frac{\sin(\pi\rho)}{u^2+2u\cos(\pi\rho)+1}du,\ y \geq 0. \tag{4}$$

The class of discrete stable distributions introduced by Steutel and Van Harn (1979) is closely related to the stable distributions. One of the main characterizations for a discrete stable random variable is to view it as a Poisson random variable $N(\nu)$ whose mean ν is a strictly stable random variable with characteristic exponent $0 < \alpha < 1$ and skewness parameter $\beta = 1$ (Devroye 1993). This is the only stable random variable which is supported on the positive axis, and hence can be taken as the mean of a Poisson distribution.

Soltani et al. (2009) consider a Poisson distribution with a random mean and take the mean to be a truncated stable random variable. Such a construction gives a fairly general class of discrete distributions with infinite second moments induced by stable random variables, that include discrete stable random variables. They obtain the probability generating function of the class of discrete distributions induced by stable laws (DIS) by using the equation in Zolotarev (1986) for the one-sided Laplace transform of a stable random variable. That can be used as substitute for the Poisson distributions in circumstances where arrivals or occurrences, which are rare in nature, become frequent. They shed light on the effectiveness of these distributions in a real dataset of daily attacks to certain computer ports, blackhole data, that is part of the Internet Motion Sensor (IMS) project (Bailey et al. (2005)), where the IMS consists of distributed blackhole sensors, each monitoring certain unused IP addresses. Zolotarev (1986) elegantly uses TSS to characterize certain functionals of stable random variables and Soltani and Shirvani (2010) work on characterization and simulation of this kind of random variables. Furthermore, the corresponding Laplace transform of a TSS random variable X can be written as Equation (2). The work by Zolotarev (1986) in deriving Equation (2), where complex integration is cleverly applied for real integration, is indeed extraordinary.

Surprisingly, a formulation for the Laplace transform of the joint density function of a bivariate stable random vector is not yet established and studied. Accordingly, the first motivation of this paper is to focus and formalize some interesting analytical results on this matter. We use the equation in Zolotarev (1986) for the characteristic function of a bivariate stable density, which is given at every point in $\mathbb{R}^2$ in polar coordinate form.

The second motivation is related to the possibility of using the bivariate truncated stable distribution as a model in financial applications. This requires the development of an estimation method. Given the analytical intractability of the density and the additional complications caused by truncation, we will use the Approximate Maximum Likelihood (AMLE) estimation approach (Rubio and Johansen 2013), a simulation-based method that exploits a frequentist interpretation of Approximate Bayesian Computation (ABC) techniques.

The remainder of this paper is structured as follows. In Section 2, we thoroughly describe preliminaries which serve as the motivation of this paper. The derivations for the Laplace transform of the bivariate stable distributions and its truncated are presented in Section 3. Section 4 details the AMLE method. Section 5 contains the results of simulation experiments and an empirical application to operational risk data. Section 6 concludes the paper.

2. Preliminaries

In this section we work on the characteristic function of a bivariate stable random vector given in Uchaikin and Zolotarev (1999). We use the notations and notions given in Zolotarev (1986). A random vector $\boldsymbol{X} = (X_1, \ldots, X_d)$ is said to be a stable random vector in $\mathbb{R}^d$ if for any positive numbers A and B there is a positive number C and a vector $\boldsymbol{D} \in \mathbb{R}^d$ such that

$$A\boldsymbol{X}^{(1)} + B\boldsymbol{X}^{(2)} = C\boldsymbol{X} + \boldsymbol{D},$$

where $\boldsymbol{X}^{(1)}$ and $\boldsymbol{X}^{(2)}$ are independent copies of $\boldsymbol{X}$. Let $\boldsymbol{Z} = (Z_1, Z_2)$ be a two-dimensional centered random vector on a plane. In polar coordinates form, it is represented by (R, ϕ), where R is the magnitude and ϕ is the phase, that is,

$$R = \sqrt{Z_1^2 + Z_2^2}, \quad \phi = \arctan\left(\frac{Z_2}{Z_1}\right).$$

We denote the probability density of ϕ, the angle between the vector $\boldsymbol{Z}$ and the x-axis, by $w(\varphi)$,

$$w(\varphi)d\varphi = P(\phi \in d\varphi).$$

The probability that the amplitude R exceeds a value r, given a fixed $\phi = \varphi$, is expressed by the relation

$$P(R > r|\phi = \varphi) = 1 - F(r|\phi = \varphi). \tag{5}$$

Further, a two dimensional random vector $\boldsymbol{Z}$ for which this conditional distribution is indeed a power distribution on $[\epsilon(\varphi), +\infty)$ that is written as follows

$$P(R > r|\varphi) = \begin{cases} c(\varphi)r^{-\alpha}, & r > \epsilon(\varphi) \\ 1, & r < \epsilon(\varphi), \end{cases} \tag{6}$$

where the positive α does not depend on φ, and $\epsilon(\varphi)$ is determined by the normalization, namely,

$$1 = P(R > \epsilon(\varphi)|\varphi) = c(\varphi)\epsilon(\varphi)^{-\alpha},$$

which yields

$$\epsilon(\varphi) = [c(\varphi)]^{\frac{1}{\alpha}}.$$

Differentiating both sides of Equation (5) yields

$$P(r|\varphi) = \begin{cases} \alpha c(\varphi)r^{-\alpha-2}, & r > [c(\varphi)]^{\frac{1}{\alpha}} \\ 0, & r < [c(\varphi)]^{\frac{1}{\alpha}}. \end{cases}$$

Then a bivariate stable random vector $\boldsymbol{X}$ is the limit (in distribution) of $(1/b_n)\sum_{i=1}^{n} \boldsymbol{Z}_i,\ b_n = bn^{1/\alpha}$, where $\boldsymbol{Z}_1, \boldsymbol{Z}_2, \ldots$ are independent copies of $\boldsymbol{Z}$. Moreover, Uchaikin and Zolotarev (1999) give the characteristic function of a centered bivariate random vector, as:

$$g_2(\boldsymbol{k}; \alpha, \beta(\theta), \lambda(\theta)) = e^{-\lambda(\theta)\boldsymbol{k}^{\alpha}\left[1 - i\beta(\theta)\tan\left(\frac{\alpha\pi}{2}\right)\right]}, \tag{7}$$

where the polar coordinates of the vector $\boldsymbol{k}$ is denoted by (k, θ),

$$\lambda(\theta) = \Gamma(1-\alpha)\cos\left(\frac{\alpha\pi}{2}\right)\int_0^{2\pi} W(\phi)\mid \cos(\theta - \phi)\mid^{\alpha} d\phi,$$

and

$$\beta(\theta) = \frac{\int_0^{2\pi} W(\phi) \mid \cos(\theta-\phi) \mid^\alpha \operatorname{sign}(\cos(\theta-\phi)) d\phi}{\int_0^{2\pi} W(\phi) \mid \cos(\theta-\phi) \mid^\alpha d\phi},$$

with $W(\varphi) \equiv w(\varphi)c(\varphi)$ a non-negative function integrable on $[0, 2\pi)$. The functions $\beta(.)$ and $\lambda(.)$ stand for the skewness and scale parameter. If we have radially symmetric assumption ((see, e.g., Nolan (2013)), then the spectral measure in this case is a uniform distribution on the unit sphere in $\mathbb{R}^d$ and $\beta(\theta) = 0$. Therefore, we note that for radially symmetric distributions, $w(\varphi) = \frac{1}{2\pi}$ and $c(\varphi) = c$. Consequently,

$$\lambda(\theta) \equiv \lambda = 2^{\alpha+1} B\left(\frac{\alpha+1}{2}, \frac{\alpha+1}{2}\right) \frac{c}{4\pi} \Gamma(1-\alpha) \cos\left(\frac{\alpha\pi}{2}\right),$$

and $\beta(\theta) = 0$.
Equation (7) can be evaluated for a general spectral measure, as we show in the following. Denoting $\lambda(\theta) = \lambda(\theta) \cos(\beta(\theta)\frac{\alpha\pi}{2})$, and $\beta(\theta) = \beta(\theta)\frac{\tan(\beta(\theta)\frac{\alpha\pi}{2})}{\tan(\frac{\alpha\pi}{2})}$, we obtain

$$\begin{aligned} g_2(\boldsymbol{k}; \alpha, \beta(\theta), \lambda(\theta)) &= e^{-\lambda(\theta)\boldsymbol{k}^\alpha \left[1 - i\beta(\theta)\tan\left(\frac{\alpha\pi}{2}\right)\right]} \\ &= e^{-\lambda(\theta)\cos\left(\beta(\theta)\frac{\alpha\pi}{2}\right)\boldsymbol{k}^\alpha \left[1 - i\beta(\theta)\frac{\tan\left(\beta(\theta)\frac{\alpha\pi}{2}\right)}{\tan\left(\frac{\alpha\pi}{2}\right)}\tan\left(\frac{\alpha\pi}{2}\right)\right]} \\ &= e^{-\lambda(\theta)\boldsymbol{k}^\alpha \left[\cos\left(\beta(\theta)\frac{\alpha\pi}{2}\right) - i\beta(\theta)\sin\left(\beta(\theta)\frac{\alpha\pi}{2}\right)\right]} \\ &= e^{-\lambda(\theta)\boldsymbol{k}^\alpha} e^{\left[-i\beta(\theta)\frac{\alpha\pi}{2}\right]}. \end{aligned}$$

In what follows, we are going to work with this representation of the characteristic function of a bivariate stable random vector. The equation in Uchaikin and Zolotarev (1999) for the characteristic function is not much popular and commonly used. In general, the characteristic function is in terms of the Cartesian coordinates. The classical equation given in Kanter et al. (1975), where the characteristic function is specified by a positive measure on the unit circle and a vector in $\mathbb{R}^2$ is usually employed and applied. Nevertheless, as we will see in Section 3, the equation in Zolotarev (1986) appears to be appropriate in derivation of the Laplace transform of the truncated density.

3. Theoretical properties of bivariate stable distributions

The characteristic functions uniquely determine the corresponding densities of stable laws; however, it is difficult to calculate the densities by using the inversion theorem directly, because dealing with improper integrals of oscillating functions is a challenge. We work with the Laplace transform since the methodology of integral transforms with various kernel types is acknowledged to be among the most powerful and efficient tools of the analysis ((Titchmarsh 1937)). The concept of stable laws in this work extended to the case of multidimensional (and even infinite-dimensional) spaces. We start with the consideration of strictly stable distributions as the most important for financial and physical applications.

Definition 3.1 *The bivariate random vector $\boldsymbol{X} = (X_1, X_2)$ in $\mathbb{R}^2$ has a bivariate strictly stable distribution in polar coordinates form if its density function satisfies the following identity*

$$q_2(x_1, x_2, \alpha) = \frac{1}{2\pi^2}\Re\left[\int_0^{\pi}\int_0^{\infty} e^{-i(x_1 k\cos\theta + x_2 k\sin\theta) - \lambda(\theta)k^\alpha e^{\frac{-i\alpha\beta(\theta)\pi}{2}}} k dk d\theta\right]$$
$$= \frac{1}{2\pi^2}\Re\left[\int_0^{\pi}\int_0^{\infty} e^{i(x_1 k\cos\theta + x_2 k\sin\theta) - \lambda(\theta)k^\alpha e^{\frac{i\alpha\beta(\theta)\pi}{2}}} k dk d\theta\right],$$

where $\beta(\theta)$ and $\lambda(\theta)$ are defined in Section 2 and $\Re$ is the real part.

Definition 3.2 *The one-sided Laplace transform of the truncated bivariate strictly stable distribution is defined by*

$$\widetilde{q_2}(\lambda_1, \lambda_2; \alpha) = \int_0^\infty\int_0^\infty e^{-\lambda_1 x_1 - \lambda_2 x_2} q_2(x_1, x_2, \alpha) dx_1 dx_2, \quad \Re\lambda_i \geq 0, \quad i = 1, 2. \tag{8}$$

Theorem 3.1 *Let $\boldsymbol{X} = (X_1, X_2)$ be a bivariate strictly stable random vector in $\mathbb{R}^2$ conditioned on both coordinates being positive, then the following statements hold.*
(i) If distribution of $\boldsymbol{X}$ is also radially symmetric, then its density function based on polar coordinates follows as

$$q_2(x_1, x_2, \alpha) = \frac{1}{2\pi^2}\Re\left[\int_0^{\pi}\int_0^{\infty} e^{-i(x_1 k\cos\theta + x_2 k\sin\theta) - \lambda k^\alpha} k dk d\theta\right], \quad x_1 > 0, x_2 > 0. \tag{9}$$

(ii) If distribution of $\boldsymbol{X}$ is also radially symmetric, then its Laplace transform in the first quadrant is given by

$$\widetilde{q}_2(\lambda_1, \lambda_2, \alpha) = \frac{1}{\pi^2}\int_0^{\frac{\pi}{2}}\int_0^{\infty}\left[\frac{\lambda_1\lambda_2}{(k^2\cos^2\theta + \lambda_1^2)(k^2\sin^2\theta + \lambda_2^2)}\right] e^{-\lambda k^\alpha} k dk d\theta$$
$$= \frac{1}{\pi^2}\int_0^{\infty}\int_0^{\infty}\frac{1}{(1 + k_1^2)(1 + k_2^2)} e^{-\lambda(\lambda_1^2 k_1^2 + \lambda_2^2 k_2^2)^{\frac{\alpha}{2}}} dk_1 dk_2. \tag{10}$$

(iii) For the general case, i.e., $\beta(\theta) \neq 0$ and positive λ_1 and λ_2, its Laplace transform in the first quadrant is defined by

$$\widetilde{q}_2(\lambda_1, \lambda_2, \alpha) = \frac{-1}{2\pi^2}\int_0^{\pi}\int_0^{\infty}\frac{A(k,\theta)}{B(k,\theta)} e^{-\lambda(\theta)k^\alpha} k dk d\theta, \tag{11}$$

where

$$\frac{A(k,\theta)}{B(k,\theta)} = \frac{k^2\cos\theta\sin\theta + \lambda_1 k\cos\theta\cos\left(\frac{\rho(\theta)\pi}{\alpha}\right) + \lambda_2 k\sin\theta\cos\left(\frac{\rho(\theta)\pi}{\alpha}\right) + \lambda_1\lambda_2\cos\left(\frac{2\rho(\theta)\pi}{\alpha}\right)}{\left(k^2\cos^2\theta + 2\lambda_1 k\cos\theta\cos\left(\frac{\rho(\theta)\pi}{\alpha}\right) + \lambda_1^2\right)\left(k^2\sin^2\theta + 2\lambda_2 k\sin\theta\cos\left(\frac{\rho(\theta)\pi}{\alpha}\right) + \lambda_2^2\right)},$$

and $\rho(\theta) = \frac{(1+\beta(\theta))\alpha}{2}$.

Proof: The proof is provided in Appendix A.

Definition 3.3 *A non negative random vector* $\boldsymbol{W}$ *has truncated bivariate strictly stable distribution conditioned on both coordinates being positive and radially symmetric if the Laplace transform of its density function is given in (10) with normalization factor equal to* $1/4$ *which is computed by taking* $\lambda_1 = 0$ *and* $\lambda_2 = 0$ *in (10).*

Theorem 3.2 *Let the polar coordinates of the vector* $\boldsymbol{X}$ *be* r *and* ϕ *and* $\beta(\theta) \neq 0$. *Then*
(i) The conditional one-sided Laplace transform involving the density of the truncated bivariate strictly stable distribution is given by

$$\widetilde{q}_{R|\varphi=\theta_0}(y, r, \theta_0, \alpha) = \frac{1}{2\pi^2 w(\theta_0)} \int_0^{\pi}\int_0^{\infty} \frac{-C}{D} e^{-\lambda(\theta)k^{\alpha}} k dk d\theta,$$

where

$$\frac{C}{D} = \frac{y\cos\left(\frac{2\rho(\theta)\pi}{\alpha}\right) + k\cos\left(\frac{\rho(\theta)\pi}{\alpha}\right)\cos(\theta-\theta_0)}{\left(k^2\cos^2(\theta-\theta_0) + 2yk\cos(\theta-\theta_0)\cos\left(\frac{\rho(\theta)\pi}{\alpha}\right) + y^2\right)}.$$

(ii) The cuts along $y = ax_1$ *is expressed as*

$$\widetilde{q}_{2}(\lambda_1, \alpha) = \frac{1}{2\pi^2}\int_0^{\pi}\int_0^{\infty} \frac{-\left[\lambda_1\cos\left(\frac{2\rho(\theta)\pi}{\alpha}\right) + k\cos\left(\frac{\rho(\theta)\pi}{\alpha}\right)(\cos\theta + a\sin\theta)\right] e^{-\lambda(\theta)k^{\alpha}}}{E} k dk d\theta,$$

where

$$E = \lambda_1^2 + 2\lambda_1 k\cos\left(\frac{\rho(\theta)\pi}{\alpha}\right)(\cos\theta + a\sin\theta) + k^2(\cos\theta + a\sin\theta)^2.$$

Proof: The proof is provided in Appendix A.

Figure 1 shows the effects of the parameters α and λ on the shape of the bivariate Laplace transform in the first quadrant for the radially symmetric distributions. It is well known that the multivariate stable distributions possess certain properties that make them useful for economic theory and probability theory. The purpose in this step is to provide certain characterizations of the conditional one-sided Laplace transform.

4. Approximate Maximum Likelihood Estimation

Standard estimation approaches are difficult to apply in the present setup. First, even the untruncated stable distribution has no explicit density; even though, at least in the univariate setup, this problem can be overcome by numerical maximization of an approximation of the density (see, e.g., Nolan 1997), the complications caused by truncation seem to preclude the use of methods based on the density function. Moreover, the general stable case has an arbitrary spectral measure, which is usually replaced by a semi-parametric discrete spectral measure for computational reasons (Tsionas 2013, Ogata 2013, Teimouri et al. 2018). Obviously, these two features make the implementation

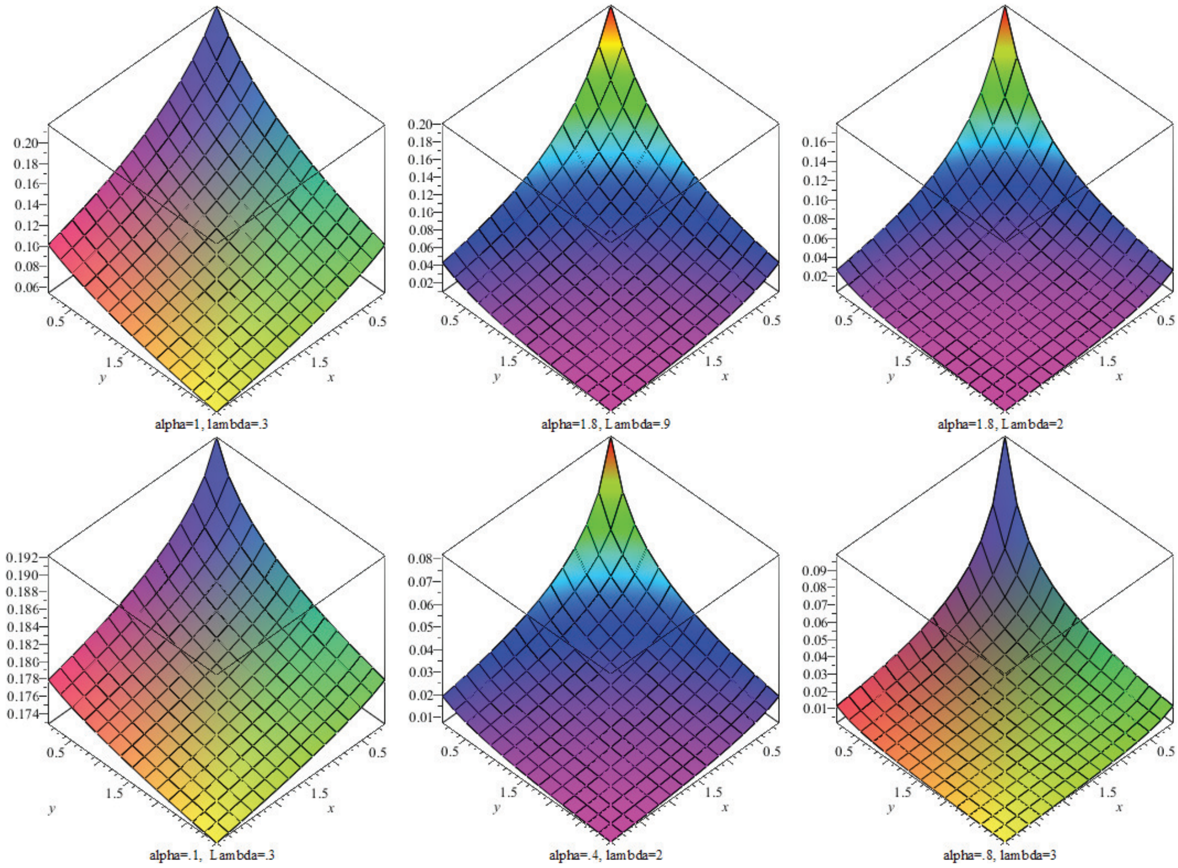

Figure 1.: Bivariate Laplace transform plots $\widetilde{q}_2(\lambda_1, \lambda_2, \alpha)$. Line 1: Left: $\alpha = 1, \lambda = 0.3$; Middle: $\alpha = 1.8, \lambda = 0.9$; Right: $\alpha = 0.1, \lambda = 0.3$. Line 2: Left: $\alpha = 1, \lambda = 3$; Middle: $\alpha = 0.4, \lambda = 2$; Right: $\alpha = 0.8, \lambda = 3$.

of maximum likelihood quite complicated.

To overcome these obstacles, we resort to Approximate Maximum Likelihood Estimation, a simulation-based method that has already proved to be effective for multivariate distributions with similar features (see Bee et al. 2015, 2017). The AMLE approach exploits the potential of Approximate Bayesian Computation techniques in a frequentist setup. In the following we briefly describe the algorithm, referring to Rubio and Johansen (2013) for details.

Given a sample $(\boldsymbol{x}_1, \dots, \boldsymbol{x}_n) \in \mathbb{R}^{d \times n}$ from a distribution with density function $f(\boldsymbol{x}; \boldsymbol{\theta})$, let $L(\boldsymbol{\theta}; \boldsymbol{x}_1, \dots, \boldsymbol{x}_n)$ be the likelihood function, where $\boldsymbol{\theta} \in \boldsymbol{\Theta} \subset \mathbb{R}^p$ is a vector of parameters. If we temporarily assume a Bayesian setup and let $\pi(\boldsymbol{\theta})$ be the prior distribution of $\boldsymbol{\theta}$, the posterior $\pi(\boldsymbol{\theta}|\boldsymbol{x})$ is given by

$$\pi(\boldsymbol{\theta}|\boldsymbol{x}) = \frac{f(\boldsymbol{x}|\boldsymbol{\theta})\pi(\boldsymbol{\theta})}{\int_{\boldsymbol{\Theta}} f(\boldsymbol{x}|\boldsymbol{t})\pi(\boldsymbol{t})d\boldsymbol{t}}. \tag{12}$$

Now consider the following approximation of the likelihood function:

$$\hat{f}_\epsilon(\boldsymbol{x}|\boldsymbol{\theta}) = \int_{\mathbb{R}^{d \times n}} K_\epsilon(\boldsymbol{x}|\boldsymbol{z}) f(\boldsymbol{z}|\boldsymbol{\theta}) d\boldsymbol{z}, \tag{13}$$

where $K_\epsilon(\boldsymbol{x}|\boldsymbol{z})$ is a normalized Markov kernel and ϵ is a scale parameter. Plugging (13) into (12) we can compute an approximation of the posterior:

$$\hat{\pi}_\epsilon(\boldsymbol{\theta}|\boldsymbol{x}) = \frac{\hat{f}_\epsilon(\boldsymbol{x}|\boldsymbol{\theta})\pi(\boldsymbol{\theta})}{\int_{\boldsymbol{\Theta}} \hat{f}_\epsilon(\boldsymbol{x}|\boldsymbol{t})\pi(\boldsymbol{t})d\boldsymbol{t}}.$$

If we assume the prior $\pi(\boldsymbol{\theta})$ to be uniform, the maximization of the likelihood is equivalent to the maximization of the posterior, provided that the latter is written in the parameterization of interest.

Let $\boldsymbol{\eta} : \mathbb{R}^{d\times n} \to \mathbb{R}^s$ be an $s \times 1$ vector of summary statistics. The kernel K_ϵ is defined on the space of these summary statistics as follows:

$$K_\epsilon(\boldsymbol{\eta}(\boldsymbol{x})|\boldsymbol{\eta}(\boldsymbol{z})) \propto \begin{cases} 1 & ||\boldsymbol{\eta}(\boldsymbol{x}) - \boldsymbol{\eta}(\boldsymbol{z})|| < \epsilon, \\ 0 & \text{otherwise}, \end{cases} \tag{14}$$

where $||\cdot||$ is a norm. The norm chosen is usually not a crucial issue, but it will be seen in the following that in the present setup the maximum norm, defined as $||\boldsymbol{x}||_\infty \stackrel{\text{def}}{=} \max_{i=1,\dots,d}|x_i|$, is more convenient for the implementation of the algorithm. Notice also that it is used by Tsionas (2013) in a similar setup.

On the other hand, the summary statistics play a key role: using $\boldsymbol{\eta}(\boldsymbol{x})$ instead of the original sample $\boldsymbol{x}$ implies no loss of information exactly if $\boldsymbol{\eta}$ is a jointly sufficient statistic for the unknown parameters of the model: in this case, $L(\boldsymbol{\theta}; \boldsymbol{x}_1, \dots, \boldsymbol{x}_n) = L(\boldsymbol{\theta}; \boldsymbol{\eta}(\boldsymbol{x}_1, \dots, \boldsymbol{x}_n))$, that is, conditioning upon the sufficient statistics is the same as conditioning upon the sample. Accordingly, the most efficient summary statistics are the sufficient statistics. Unfortunately, in many cases, included the present one, sufficient statistics are not available, so that the choice of $\boldsymbol{\eta}(\boldsymbol{x})$ has to be based on different criteria, mostly on a case-by-case basis. We will study this issue in Section 4.1.

The preceding discussion motivates the following algorithm:

Algorithm 4.1 *(AMLE)*

(i) Obtain a sample $\boldsymbol{\theta}^*_\epsilon = (\boldsymbol{\theta}^*_{\epsilon,1}, \dots, \boldsymbol{\theta}^*_{\epsilon,\ell})'$ from the approximate posterior $\hat{\pi}_\epsilon(\boldsymbol{\theta}|\boldsymbol{x})$; ℓ is commonly called ABC sample size;
(ii) Use this sample to construct a nonparametric estimator $\hat{\pi}_{\ell,\epsilon}$ of the density $\hat{\pi}_\epsilon(\boldsymbol{\theta}|\boldsymbol{x})$;
(iii) Compute the maximum of $\hat{\pi}_{\ell,\epsilon}$, $\tilde{\boldsymbol{\theta}}_{\ell,\epsilon}$. This is an approximation of the MLE $\hat{\boldsymbol{\theta}}$.

The most common implementation of Step 1 is the ABC rejection algorithm (Beaumont 2010), described by the following pseudo-code.

Algorithm 4.2 *(ABC rejection algorithm)*

(i) Simulate $\boldsymbol{\theta}^*$ from the prior distribution $\boldsymbol{\pi}(\cdot)$;
(ii) Generate $\boldsymbol{x} = (x_1, \dots, x_n)'$ from $f(\cdot|\boldsymbol{\theta}^*)$;
(iii) Use $\boldsymbol{x}$ to compute summary statistics $\boldsymbol{\eta}(\boldsymbol{x})$; accept $\boldsymbol{\theta}^*$ with probability $\propto K_\epsilon(\boldsymbol{\eta}(\boldsymbol{x})|\boldsymbol{\eta}(\boldsymbol{z}))$, otherwise return to Step 1.

Rubio and Johansen (2013) study the asymptotic properties of $\tilde{\boldsymbol{\theta}}_{\ell,\epsilon}$. They first show that, under a mild condition on K_ϵ, $\hat{\pi}_\epsilon(\boldsymbol{\theta}|\boldsymbol{x})$ converges pointwise to $\pi(\boldsymbol{\theta}|\boldsymbol{x})$ as $\epsilon \to 0$, for any $\boldsymbol{\theta} \in \boldsymbol{\Theta}$. Moreover, under the additional condition of equicontinuity of $\hat{\pi}_\epsilon(\cdot|\boldsymbol{x})$ on $\boldsymbol{\Theta}$, $\lim_{\epsilon\to 0}\hat{\pi}_\epsilon(\boldsymbol{\theta}_\epsilon|\boldsymbol{x}) = \pi(\boldsymbol{\theta}_0|\boldsymbol{x})$, where $\boldsymbol{\theta}_\epsilon$ is the maximizer of $\hat{\pi}_\epsilon(\cdot|\boldsymbol{x})$ for each $\epsilon > 0$ and $\boldsymbol{\theta}_0$ is the unique mazimizer of $\pi(\cdot|\boldsymbol{x})$.

Let now $\tilde{\boldsymbol{\theta}}_{\ell,\epsilon}$ be an estimator obtained from the sample $\boldsymbol{\theta}^*_\epsilon$ and such that $\tilde{\boldsymbol{\theta}}_{\ell,\epsilon} \to \boldsymbol{\theta}_\epsilon$ almost surely when $\ell \to \infty$. It follows that, for any $\gamma > 0$, there exists $\epsilon > 0$ such that $\lim_{\ell\to\infty}|\hat{\pi}_{\ell,\epsilon}(\tilde{\boldsymbol{\theta}}_\epsilon|\boldsymbol{x}) - \pi(\boldsymbol{\theta}_0|\boldsymbol{x})| \le \gamma$ almost surely. In other words, using a reasonably well-behaved density estimation method based on Algorithm 4.2, the maximum of the AMLE approximation can be made arbitrarily close to the maximum of the true posterior distribution, which is identical to the MLE under the conditions in Rubio and Johansen (2013). By continuity, we can conclude that $\tilde{\boldsymbol{\theta}}_{\ell,\epsilon}$ approximates $\boldsymbol{\theta}_0$ when $\ell \to \infty$ and $\epsilon \to 0$.

Having introduced the basic estimation theory, we now concentrate on its application to the truncated bivariate stable distribution.

4.1. *AMLE of the truncated bivariate stable distribution*

Let $\boldsymbol{X}$ be a bivariate stable distribution conditioned on both coordinates being positive, and let's assume a discrete spectral measure. The parameter vector is $\boldsymbol{\theta} = (\alpha, \boldsymbol{\gamma}, \boldsymbol{\mu})'$, where $\boldsymbol{\gamma} = (\gamma_1, \dots, \gamma_m)'$ and m the number of masses; let $\boldsymbol{x}_1, \dots, \boldsymbol{x}_n$ be the observed data and $\boldsymbol{\eta}^{obs}$ the corresponding summary statistics. AMLE of the bivariate truncated stable distribution works as follows:

Algorithm 4.3 *(AMLE of the bivariate truncated stable distribution)*

(i) *Simulate $\boldsymbol{\theta}^*$ from the prior distribution $\pi(\boldsymbol{\theta}) = \prod_{i=1}^p \pi(\theta_i)$, where $\pi(\theta_i)$ is $U(\theta_{iL}, \theta_{iU})$;*
(ii) *Generate $\boldsymbol{x} = (\boldsymbol{x}_1, \dots, \boldsymbol{x}_n)'$ from $f(\cdot|\boldsymbol{\theta}^*)$, where f is the bivariate truncated stable density;*
(iii) *Use $\boldsymbol{x}$ to compute summary statistics $\boldsymbol{\eta}^{sim}$; accept $\boldsymbol{\theta}^*$ with probability $\propto K_\epsilon(\boldsymbol{\eta}^{obs}|\boldsymbol{\eta}^{sim})$, otherwise return to Step 1.*
(iv) *Repeat steps 1-3 until ℓ vectors of simulated parameter values $\boldsymbol{\theta}^*_\epsilon = (\boldsymbol{\theta}^*_{\epsilon,1}, \dots, \boldsymbol{\theta}^*_{\epsilon,\ell})'$ from the approximate posterior $\tilde{\pi}_\epsilon(\boldsymbol{\theta}|\boldsymbol{x})$ are accepted; $\boldsymbol{\theta}^*_\epsilon$ is the ABC sample.*
(v) *Use $\boldsymbol{\theta}^*_\epsilon$ to find a nonparametric estimator $\hat{\pi}_{\ell,\epsilon}$ of the density $\tilde{\pi}_\epsilon(\boldsymbol{\theta}|\boldsymbol{x})$;*
(vi) *Compute the maximum of $\hat{\pi}_{\ell,\epsilon}$, $\tilde{\boldsymbol{\theta}}_{\ell,\epsilon}$, which is the approximate MLE.*

Step (iii) above requires to select only the parameter vectors such that the norm of the difference between the summary statistics computed on true and simulated data is smaller than ϵ. Typically, instead of setting ϵ, whose value is difficult to guess in advance, one chooses some large integer n_p and employs the following modified version of Algorithm 4.3.

Algorithm 4.4 *(AMLE of the bivariate truncated stable distribution)*

(i)-(ii) *Same as Algorithm 4.3;*
(iii) *Use $\boldsymbol{x}$ to compute summary statistics $\boldsymbol{\eta}^{sim}$;*
(iv) *Repat steps (i)-(iii) n_p times;*
(v) *Retain only the ℓ simulated parameter vectors corresponding to the ℓ smallest value of $||\boldsymbol{\eta}^{obs} - \boldsymbol{\eta}^{sim}||_\infty$; these ℓ vectors $\boldsymbol{\theta}^*_1, \dots, \boldsymbol{\theta}^*_\ell$ are the ABC sample $\boldsymbol{\theta}^*_\ell$;*
(vi) *Use $\boldsymbol{\theta}^*_\ell$ to find a nonparametric estimator $\hat{\pi}_{\ell,\epsilon}$ of the density $\tilde{\pi}_\epsilon(\boldsymbol{\theta}|\boldsymbol{x})$, where ϵ is a function of ℓ and n_p;*
(vii) *Compute the maximum of $\hat{\pi}_{\ell,\epsilon}$, $\tilde{\boldsymbol{\theta}}_{\ell,\epsilon}$, which is the approximate MLE.*

Two issues here are the choices of

(i) the parameters θ_{iL} and θ_{iU} $(i = 1, \dots, d)$ of the uniform distributions;
(ii) the number s and the functional forms of the summary statistics.

As for the second issue, Tsionas (2013, 2016) shows that it is convenient to use the empirical characteristic function of univariate projections. In a Bayesian setup closely related to our approach, Peters et al. (2012) also use univariate projections to construct summary statistics for the estimation of the multivariate stable distribution. The key result is that the projection of $\boldsymbol{X}$ on $\boldsymbol{\tau} \in \mathbb{R}^d$ is univariate stable (Nolan et al. 2001).

As suggested by Tsionas (2013), we use $\boldsymbol{\tau} \in S^d$, where S^d is the d-dimensional unit sphere. Hence, $\boldsymbol{\eta}^{obs}$ and $\boldsymbol{\eta}^{sim}$ at Step 4 above are the empirical characteristic functions of certain linear projections computed on true and simulated data respectively.

Given a random variable X, the empirical characteristic function (echf) estimation method is based on the minimization of

$$\tilde{Q}_n(\boldsymbol{\theta}) = \int_{\mathbb{R}} |\psi_n(r;\boldsymbol{x}) - \psi_{\boldsymbol{\theta}}(r)|^2 w(r)dr, \tag{15}$$

where $\psi_n(r;\boldsymbol{x}) = (1/n)\sum_{i=1}^n e^{(irx_i)}$ and $\psi_{\boldsymbol{\theta}}(r) = \mathrm{E}(e^{(irX)}$ are respectively the empirical and theoretical chfs and w is a weighting function. Typically, $w(r) = e^{-cr^2}$, $c > 0$. AMLE requires the evaluation of a similar integral:

$$Q_n(\boldsymbol{x}^*, \boldsymbol{x}) = \int_{\mathbb{R}} \left| \sum_{j=1}^n e^{irx_j^*} - \sum_{j=1}^n e^{irx_j} \right|^2 e^{-cr^2}dr. \tag{16}$$

The integral (15) can be solved explicitly (Xu and Knight 2010, p. 45). Analogously, the solution of (16) can be obtained in closed form (Bee and Trapin 2018):

$$\begin{aligned} Q_n(\boldsymbol{x}^*, \boldsymbol{x}) &= \int_{\mathbb{R}} \left| \sum_{j=1}^n e^{irx_j^*} - \sum_{j=1}^n e^{irx_j} \right|^2 e^{-cr^2}dr \\ &= \frac{1}{n^2}\sqrt{\frac{\pi}{c}} \left(\sum_{i=1}^n\sum_{j=1}^n e^{-\frac{1}{4c}(x_i^*-x_j^*)^2} + \sum_{i=1}^n\sum_{j=1}^n e^{-\frac{1}{4c}(x_i-x_j)^2} - 2\sum_{i=1}^n\sum_{j=1}^n e^{-\frac{1}{4c}(x_i^*-x_j)^2} \right). \end{aligned} \tag{17}$$

In the multivariate case, we can still exploit (17). Using the maximum norm, we have

$$\begin{aligned} &||\boldsymbol{\eta}^{obs} - \boldsymbol{\eta}^{sim}||_\infty = \max_{k=1,\dots,p} |\eta_k^{obs} - \eta_k^{sim}| \\ &= \max_{k=1,\dots,p} \int_{\mathbb{R}} \left| \sum_{j=1}^n e^{irx_{kj}^*} - \sum_{j=1}^n e^{irx_{kj}} \right|^2 e^{-cr^2}dr \\ &= \max_{k=1,\dots,p} \frac{1}{n^2}\sqrt{\frac{\pi}{c}} \left(\sum_{i=1}^n\sum_{j=1}^n e^{-\frac{1}{4c}(x_{ki}^*-x_{kj}^*)^2} + \sum_{i=1}^n\sum_{j=1}^n e^{-\frac{1}{4c}(x_{ki}-x_{kj})^2} - 2\sum_{i=1}^n\sum_{j=1}^n e^{-\frac{1}{4c}(x_{ki}^*-x_{kj})^2} \right), \end{aligned}$$

where x_{kj} is the j-th observation of the k-th marginal. The availability of a closed-form solution of (16) is essential for the computational feasibility of the method. Moreover, with respect to the discrete echf, the continuous echf does not require the choice of the grid of points where the echf is evaluated. Finally, simulation results in Xu and Knight (2010) and Bee and Trapin (2018) suggest that the performance of continuous echf is better than its discrete counterpart.

In classical echf estimation it is in principle possible to compute the asymptotically optimal weighting function $w(r)$ (Carrasco and Florens 2002). However, the resulting function is not necessarily exponential, so that (17) may no longer be solved in closed form. Since the explicit solution of the integral is key to obtain a computationally efficient algorithm, we prefer to trade a small efficiency loss for a viable computing time and use $w(r) = e^{-cr^2}$ with $c = 1$. Since using $c = 1$ is equivalent to assuming unit scaling, if the dispersion of the data is very large or very small, $c = 1$ may not be a good numerical value. In this case a preprocessing step, such as rescaling each coordinate by some robust measure like the interquartile range would make $c = 1$ reasonable.

5. Numerical results

5.1. *Monte Carlo experiments*

We sample $n = 500$ observations from the truncated bivariate stable distribution with $\alpha = 1.7$, $\boldsymbol{\gamma} = (0.1, 0.5, 0.5, 0.1)'$ and $\boldsymbol{\mu} = (0, 0)'$. AMLE uses the following inputs: $\boldsymbol{\theta}_L = (1.2, 0, 0, 0, 0, -0.5, -0.5)'$, $\boldsymbol{\theta}_U = (2, 1, 1, 1, 1, 0.5, 0.5)'$ and $n_p \in \{20\,000, 50\,000, 100\,000, 200\,000\}$. The choice of θ_{L1} and θ_{U1} can be based on some preliminary estimate or on a small pilot simulation. Since the latter is the most suitable approach for the remaining parameters, we use it also for θ_{L1} and θ_{U1}. To simulate the distribution, we have adapted the `mrstab` function from the `alphastable` `R` package (Teimouri et al. 2018), which is based on Modarres and Nolan (1994). Notice that the simulation procedure assumes that masses are located on the unit sphere with locations $\boldsymbol{\tau}_k = (\cos(2\pi(k-1)/m), \sin(2\pi(k-1)/m))$, $k = 1, \ldots, 4$, $m = 4$ (Nolan et al. 2001, p. 1116).

Figure 2 shows the boxplots of the simulated parameter values corresponding to the 50 smallest values of $\max_{k=1,\ldots,4} |\eta_k^{obs} - \eta_k^{sim}|$ for $n_p \in \{20\,000, 50\,000, 100\,000, 200\,000\}$, where η_k is the continuous echf of the k-th linear projection $< \boldsymbol{X}, \boldsymbol{\tau}_k >$. Remember that n_p is the total number of parameter vectors simulated from the uniform priors (step (i) of Algorithm 4.4).

The non-parametric kernel density estimator at Step (v) of Algorithm 4.3 is computed in three different ways, by means of: (i) the sample mean of the univariate ABC samples (M); (ii) the blurring mean-shift algorithm (BS); (iii) the maximum of the product of the univariate kernel densities (UKD). See Bee et al. (2017) for details on the three approaches. In all cases below, M and BS yield identical results, hence BS outcomes are not reported.

Table 1 displays the outcomes with different ABC sample sizes ℓ, for $n_p = 200\,000$, where ℓ is the number of simulated parameter vectors corresponding to the ℓ smallest values of $||\boldsymbol{\eta}^{obs} - \boldsymbol{\eta}^{sim}||_\infty$ (step (v) of Algorithm 4.4). The estimates in Table 1 are in line with the true values of the parameters. There is a small bias in some cases, but Figure 2 suggests that, according to large sample AMLE theory, the results improve, in terms of both bias and variance, as n_p increases. Estimation seems to be more problematic for the location parameters μ_1 and μ_2, analogously to the findings by Peters et al. (2012) for the bivariate stable distribution in a Bayesian setup via ABC techniques. Note, however, that the location parameters are less important in economic and financial applications, especially for risk management purposes.

Table 1.: Estimated parameter values with different ABC sample sizes ℓ, for $n_p = 200\,000$, where ℓ and n_p are defined in algorithms 4.3 and 4.4.

		α	γ_1	γ_2	γ_3	γ_4	μ_1	μ_2
$\ell = 10$	M	1.750	0.167	0.550	0.544	0.177	−0.136	−0.160
	UKD	1.696	0.110	0.472	0.632	0.288	−0.240	−0.203
$\ell = 15$	M	1.726	0.140	0.526	0.557	0.197	−0.098	−0.177
	UKD	1.696	0.110	0.472	0.632	0.288	−0.240	−0.203
$\ell = 20$	M	1.732	0.157	0.521	0.516	0.207	−0.059	−0.227
	UKD	1.784	0.145	0.536	0.652	0.238	−0.230	−0.293
$\ell = 30$	M	1.726	0.152	0.508	0.504	0.208	−0.028	−0.176
	UKD	1.784	0.145	0.536	0.652	0.238	−0.230	−0.293
$\ell = 50$	M	1.730	0.162	0.496	0.524	0.219	−0.075	−0.170
	UKD	1.784	0.145	0.536	0.652	0.238	−0.230	−0.293
$\ell = 100$	M	1.73	0.203	0.471	0.474	0.230	−0.046	−0.121
	UKD	1.768	0.161	0.452	0.523	0.219	0.116	0.002

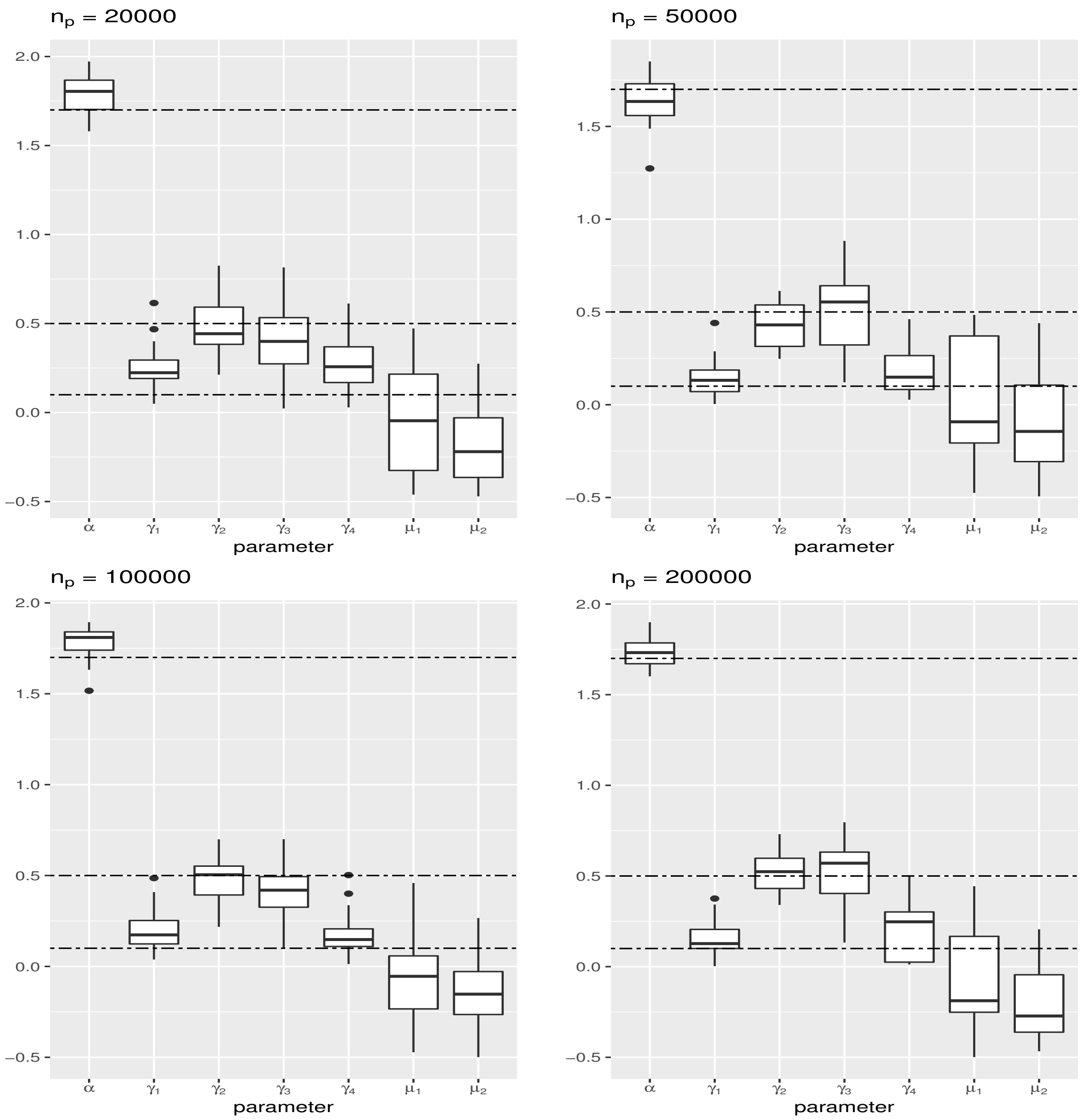

Figure 2.: Box-plots of the simulated parameter values corresponding to the 50 smallest values of $\max_{k=1,\ldots,4} |\eta_k^{obs} - \eta_k^{sim}|$.

5.2. *Empirical analysis: operational risk*

As a real-data empirical application we consider operational risk losses recorded at the Italian bank Unicredit between 2005 and 2014; for a detailed description of the data see Hambuckers et al. (2018) and Bee et al. (2019). In the present analysis we use the losses of the "Damage to Physical Assets" (*DPA*) and "Business Disruption and System Failures" (*BDSF*) business lines.

The losses are left-truncated at a known threshold *th*. For confidentiality reasons we cannot reveal the threshold; moreover, we have rescaled the data via multiplication by means of a constant *k*. Hence, without loss of generality, the data used for estimation are given by $x_1 = (DPA - th)/k$ and $x_2 = (BDSF - th)/k$, so that both variables are left-truncated at zero. Finally, since the two time series are irregularly spaced, and losses are not, in general, observed simultaneously in the two business lines, we aggregate the losses on a monthly basis. This yields a sample size $n = 114$ (monthly observations from January 2005 to June 2014). A scatterplot is displayed in Figure 3.

Some pilot simulations aimed at finding initial values for $\boldsymbol{\theta}_L$ and $\boldsymbol{\theta}_U$ lead us to use $\boldsymbol{\theta}_L = (1.01, 0, 0, 0, 0, 1.2, 0.15)'$ and $\boldsymbol{\theta}_U = (1.3, 1, 1, 1, 1, 2.2, 1.1)'$. Figure 4 shows the boxplots of the ABC samples for $\ell \in \{10, 20, 50, 100\}$, obtained with $n_p = 50\,000$. Table 2 reports the estimates for various values of ℓ. The *BS* algorithm again gives the same outcomes of the sample mean, and is

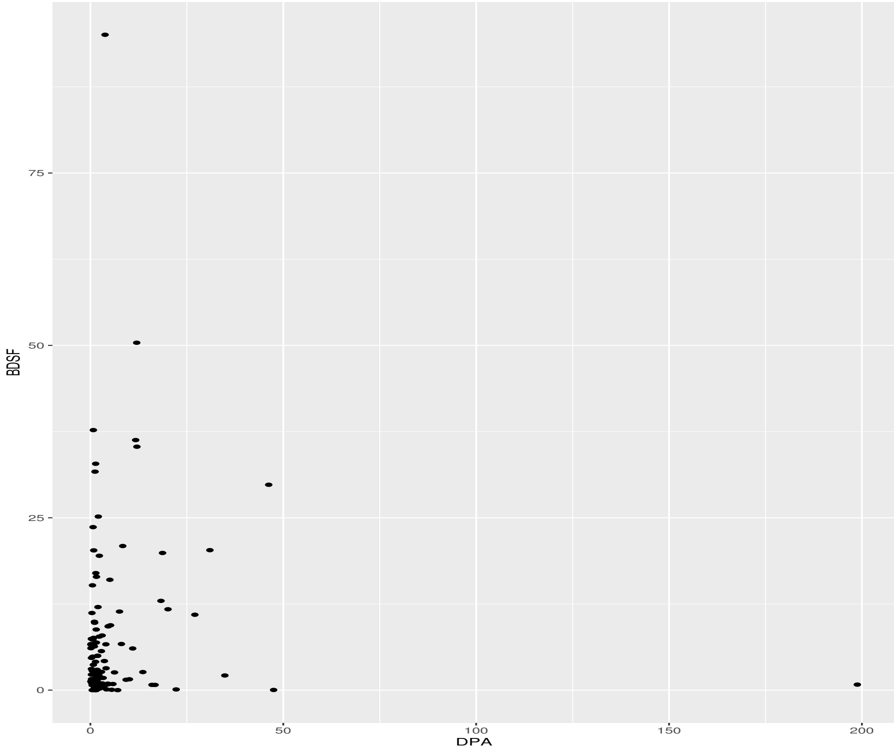

Figure 3.: Operational risk: scatterplot of the rescaled *DPA* and *BDSF* observations.

therefore omitted. For most values of n_p, the M and UKD estimates are also close to each other Hence, the results seem to be quite robust with respect to the non-parametric kernel density estimator used. Moreover, especially with the UKD method, the estimates corresponding to different n_ps are quite similar for most parameters, at least for $\ell \geq 20$, suggesting that the choice of the value of n_p is not critical either.

Table 2.: Operational risk: estimated parameter values with different ABC sample sizes ℓ, for $n_p = 50\,000$.

		α	γ_1	γ_2	γ_3	γ_4	μ_1	μ_2
$\ell = 10$	M	1.096	0.841	0.811	0.406	0.261	1.519	0.582
	UKD	1.127	0.978	0.986	0.458	0.151	1.398	0.321
$\ell = 15$	M	1.088	0.790	0.805	0.395	0.299	1.624	0.678
	UKD	1.084	0.772	0.778	0.338	0.176	1.847	0.897
$\ell = 20$	M	1.084	0.785	0.785	0.384	0.304	1.700	0.655
	UKD	1.084	0.772	0.778	0.338	0.176	1.847	0.897
$\ell = 30$	M	1.083	0.783	0.791	0.415	0.290	1.698	0.592
	UKD	1.084	0.772	0.778	0.338	0.176	1.847	0.897
$\ell = 50$	M	1.086	0.782	0.789	0.419	0.277	1.671	0.586
	UKD	1.101	0.736	0.779	0.217	0.194	1.907	0.561
$\ell = 100$	M	1.087	0.795	0.782	0.418	0.274	1.665	0.567
	UKD	1.077	0.797	0.793	0.293	0.342	1.525	0.648

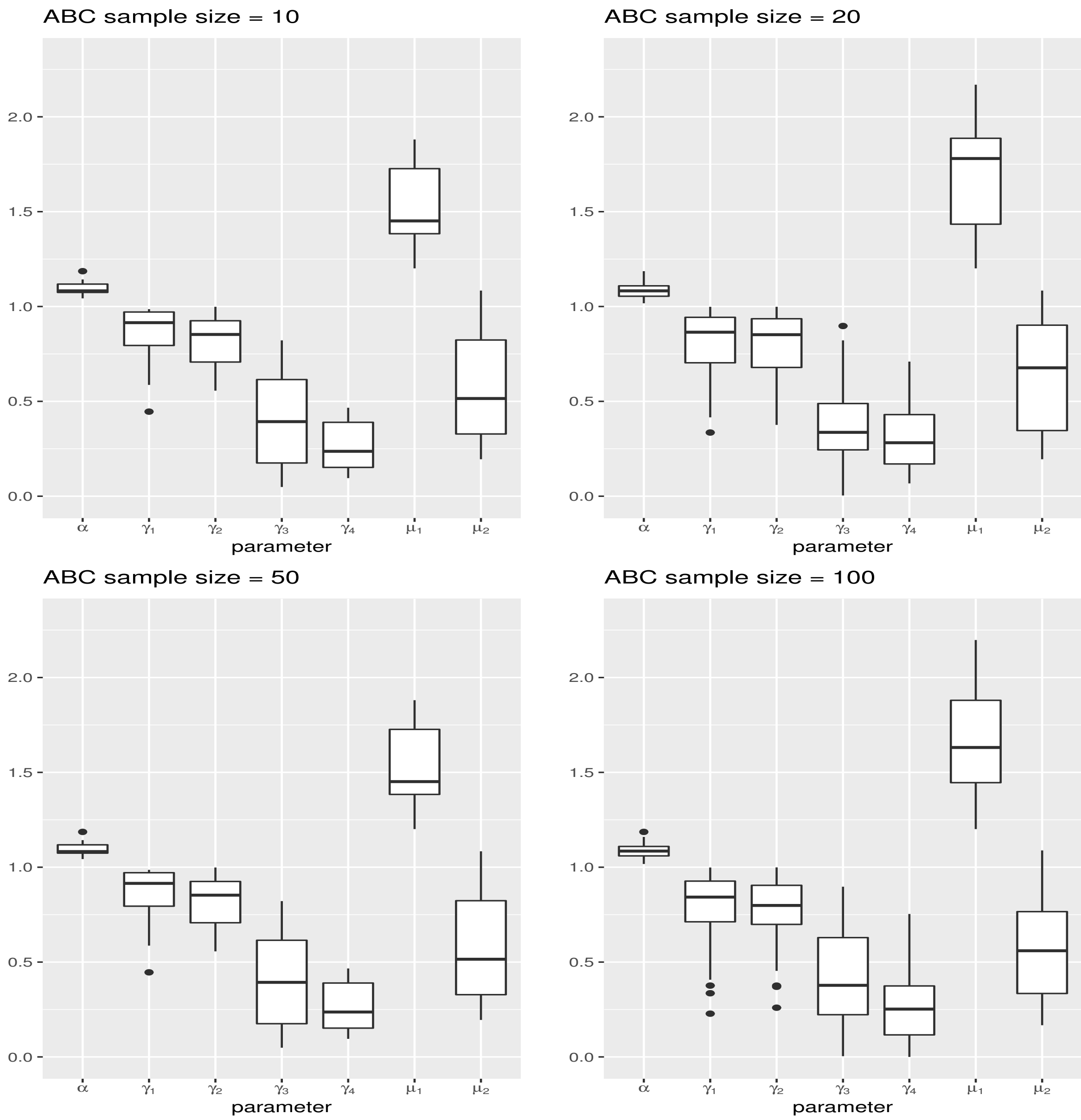

Figure 4.: Operational risk box-plots of the simulated parameter values included in the ABC sample for $\ell \in \{10, 20, 50, 100\}$.

Tail probabilities of the joint distribution can be approximated via simulation. Table 3 shows selected simulated joint probabilities (in percentage) along with empirical joint probabilities. The former are obtained as $p^S_{1,2} = \#\{x^*_{i1} > x_1, x^*_{i2} > x_2\}/B$, where $B = 10\,000$ is the number of replications and $(x^*_{i1}, x^*_{i2})'$ is the i-th bivariate observations simulated from the estimated truncated bivariate stable distribution. The estimated parameters are the M estimates obtained with $\ell = 20$ (see Table 2). The empirical joint probabilities are computed as $p^O_{1,2} = \#\{x_{i1} > x_1, x_{i2} > x_2\}/n$, where $(x_{i1}, x_{i2})'$ is the i-th observation.

Given the rather small sample size, the empirical estimates of the joint probabilities corresponding to high quantiles are not very reliable. In particular, for high quantile levels it is not surprising that $p^O_{1,2}$ are equal to zero, and it makes sense that $p^S_{1,2}$ are larger than 0. Moreover, in the bivariate truncated stable distribution the two marginals are assumed to share the same stability parameter α, an hypothesis that is unlikely to be exactly satisfied in practice.

Given these premises, the joint probabilities in Table 3 suggest a good fit. This is also confirmed by the scatterplots in Figure 5, where the top panel displays the true data and the second one a

Table 3.: Operational risk: simulated (p_{12}^S) and empirical (p_{12}^O) joint percentage tail probabilities. q_γ is the quanitle at level γ.

(x_1, x_2)	$p_{1,2}^S$	$p_{1,2}^O$
$(10, 10)$	3.5	7.89
$(15, 15)$	1.32	2.60
$(q_{0.9}, q_{0.9})$	1.51	1.75
$(q_{0.9}, q_{0.95})$	0.64	0
$(q_{0.95}, q_{0.9})$	0.54	1.75
$(q_{0.95}, q_{0.95})$	0.28	0
$(q_{0.95}, q_{0.99})$	0.20	0
$(q_{0.99}, q_{0.95})$	0.18	0

sample of the same size ($n = 114$) simulated from the fitted bivariate truncated stable. To ease the comparison, the scales of the two scatterplots are identical.

6. Conclusion

This work illustrates some analytical results for the bivariate truncated stable distribution. The results are particularly useful in financial and economic applications such as market and operational risk management, where truncated random variables are used. Because the truncated bivariate stable distribution does not have a closed-form formula for its probability density function, and since it does not necessarily have the second or even the first moment, we estimate the parameters via the Approximate Maximum Likelihood Estimation approach based on the empirical characteristic functions of certain linear combinations of a truncated bivariate stable random vector.

The properties of our procedures are illustrated via simualtion and via an empirical application on the operational risk data, which are well suited also because they are left-truncated. Given the rather small size of our sample, the joint probabilities suggest a good fit. According to our simulation experiments, only the estimation of the location parameter vector is less reliable. This issue has already been noted when using other estimation techniques (Peters et al. 2012), and would therefore deserve some additional investigation.

In future work, the theoretical results presented in this manuscript can be extended to dimensions greater than two. Further, certain results analogous to some of the main findings presented by Horrace (2005) such as the log-concavity, unimodality and the multivariate-totally-positive of order 2 properties can be developed for truncated bivariate stable distributions. Since estimation is a difficult problem that requires non-standard techniques, it would be interesting to study the performance of other approaches. In particular, a viable solution may be a technique based on the minimization of some distance between the Laplace transform developed in this article and its sample counterpart. Finally, applications to market risk data or to actuarial losses may provide further insights in the potential of the model for practical purposes.

Acknowledgements

We thank Fabio Piacenza (UniCredit SpA) for providing us with the operational risk data. We also thank an anonymous reviewer for valuable comments on an earlier version of this paper.

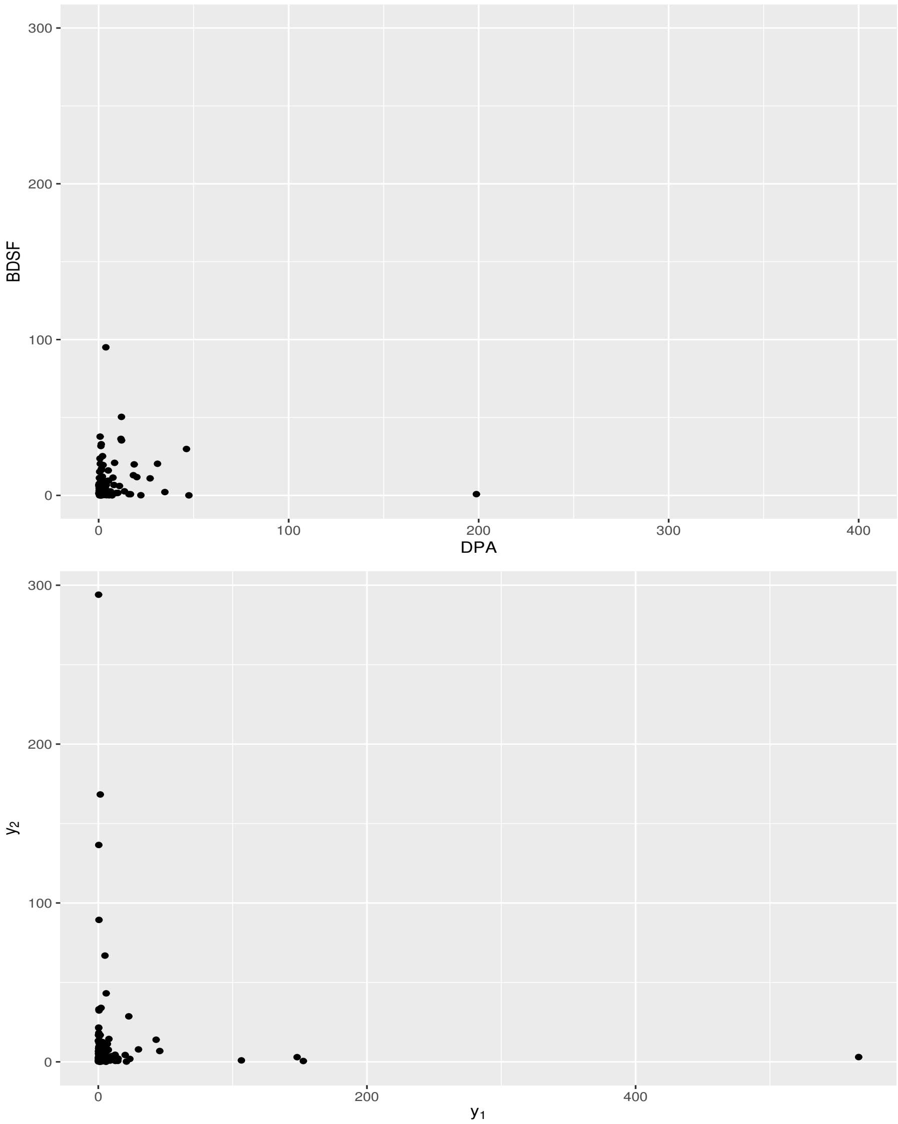

Figure 5.: Operational risk: scatterplot of the true (top panel) and simulated bivariate truncated stable (bottom panel) data.

References

Bailey, M., Cooke, E., Jahanian, F., and Nazario, J. (2005). The internet motion sensor: A distributed blackhole monitoring system. In *Conference: Proceedings of the Network and Distributed System Security Symposium, NDSS 2005, San Diego, California, USA*.

Beaumont, M. A. (2010). Approximate Bayesian Computation in evolution and ecology. *Annual Review of Ecology, Evolution, and Systematics*, 41:379–406.

Bee, M., Benedetti, R., and Espa, G. (2017). Approximate maximum likelihood estimation of the Bingham distribution. *Computational Statistics & Data Analysis*, 108:84 – 96.

Bee, M., Espa, G., and Giuliani, D. (2015). Approximate maximum likelihood estimation of the autologistic model. *Computational Statistics and Data Analysis*, 84:14–26.

Bee, M., Hambuckers, J., and Trapin, L. (2019). Estimating Value-at-Risk for the g-and-h distribution: an indirect inference approach. *Quantitative Finance*, 19(8):1255–1266.

Bee, M. and Trapin, L. (2018). A characteristic function-based approach to approximate maximum likelihood estimation. *Communications in Statistics - Theory and Methods*, 47(13):3138–3160.

Carrasco, M. and Florens, J. (2002). Efficient GMM estimation using the empirical characteristic function. Technical report, Department of Economics, University of Rochester.

Cheng, B. and Rachev, S. T. (1995). Multivariate stable futures prices. *Mathematical Finance*, 5(2):133–153.

Cont, R. (2001). Empirical properties of asset returns: stylized facts and statistical issues. *Quantitative Finance*, 1:223–236.

Devroye, L. (1993). A triptych of discrete distributions related to the stable law. *Statistics & Probability Letters*, 18(5):349–351.

DuMouchel, W. H. (1973). On the asymptotic normality of the maximum-likelihood estimate when sampling from a stable distribution. *The Annals of Statistics*, pages 948–957.

Eberlein, E. and Keller, U. (1995). Hyperbolic distributions in finance. *Bernoulli*, 1:281–299.

Embrechts, P., Frey, R., and McNeil, A. (2015). Quantitative risk management. *Princeton Series in Finance, Princeton*, 10.

Fama, E. F. (1965). The behavior of stock-market prices. *The Journal of Business*, 38(1):34–105.

Frahm, G. and Jaekel, U. (2005). Random matrix theory and robust covariance matrix estimation for financial data. *arXiv preprint physics/0503007*.

Hambuckers, J., Groll, A., and Kneib, T. (2018). Understanding the economic determinants of the severity of operational losses: A regularized Generalized Pareto regression approach. *Journal of Applied Econometrics*, 33(6):898–935.

Horrace, W. C. (2005). Some results on the multivariate truncated normal distribution. *Journal of Multivariate Analysis*, 94(1):209–221.

Kanter, M. et al. (1975). Stable densities under change of scale and total variation inequalities. *The Annals of Probability*, 3(4):697–707.

Kotz, S. and Ostrovskii, I. (1996). A mixture representation of the Linnik distribution. *Statistics & probability letters*, 26(1):61–64.

Kozubowski, T. J. (2000). Exponential mixture representation of geometric stable distributions. *Annals of the Institute of Statistical Mathematics*, 52(2):231–238.

Lévy, P. (1925). *Calcul des Probabilités*. Gauthier-Villars, Paris.

Mandelbrot, B. B. (1997). New methods in statistical economics. In *Fractals and Scaling in Finance*, pages 79–104. Springer, New York.

McCulloch, J. H. (1986). Simple consistent estimators of stable distribution parameters. *Communications in Statistics-Simulation and Computation*, 15(4):1109–1136.

Modarres, R. and Nolan, J. (1994). A method for simulating stable random vectors. *Computational Statistics*, 9(1):11–20.

Nolan, J. P. (1997). Numerical calculation of stable densities and distribution functions. *Communications in Statistics. Stochastic Models*, 13(4):759–774.

Nolan, J. P. (2013). Multivariate elliptically contoured stable distributions: theory and estimation. *Computational Statistics*, 28(5):2067–2089.

Nolan, J. P., Panorska, A. K., and McCulloch, J. H. (2001). Estimation of stable spectral measures. *Mathematical and Computer Modelling*, 34(9):1113–1122.

Ogata, H. (2013). Estimation for multivariate stable distributions with generalized empirical likelihood. *Journal of Econometrics*, 172(2):248 – 254.

Peters, G., Sisson, S., and Fan, Y. (2012). Likelihood-free Bayesian inference for α-stable models. *Computational Statistics & Data Analysis*, 56(11):3743 – 3756.

Rubio, F. J. and Johansen, A. M. (2013). A simple approach to maximum intractable likelihood estimation. *Electronic Journal of Statistics*, 7:1632–1654.

Soltani, A. and Shirvani, A. (2010). Truncated stable random variables: characterization and simulation. *Computational Statistics*, 25(1):155–161.

Soltani, A., Shirvani, A., and Alqallaf, F. (2009). A class of discrete distributions induced by stable laws. *Statistics & Probability Letters*, 79(14):1608–1614.

Soltani, A. and Tafakori, L. (2013). A class of continuous kernels and Cauchy type heavy tail distributions. *Statistics & Probability Letters*, 83(4):1018–1027.

Steutel, F. and Van Harn, K. (1979). Discrete analogues of self-decomposability and stability. *The*

Annals of Probability, pages 893–899.

Tafakori, L. and Soltani, A. (2017). A note on the Cauchy-type mixture distributions. *Journal of Statistical Computation and Simulation*, 87(9):1901–1910.

Teimouri, M., Torshizi, M., Mohammadpour, A., and Nadarajah, S. (2018). alphastable: An R Package for Modelling Multivariate Stable and Mixture of Symmetric Stable Distributions.

Titchmarsh, E. (1937). *Introduction to the Theory of Fourier Integrals*. Oxford University Press, Oxford.

Tsionas, E. G. (2013). Bayesian inference in multivariate stable distributions. SSRN Working Paper No. 2214610, Available at SSRN: https://ssrn.com/abstract=2214610.

Tsionas, E. G. (2016). Bayesian analysis of multivariate stable distributions using one-dimensional projections. *Journal of Multivariate Analysis*, 143:185 – 193.

Uchaikin, V. V. and Zolotarev, V. M. (1999). *Chance and stability: stable distributions and their applications*. de Gruyter, Berlin.

Xu, D. and Knight, J. (2010). Continuous empirical characteristic function estimation of mixtures of normal parameters. *Econometric Reviews*, 30(1):25–50.

Zolotarev, V. (1986). *One-dimensional Stable Distributions*. Translations of mathematical monographs. American Mathematical Society.

Appendix A: Proofs of Theorems 3.1 and 3.2

Proof of Theorem 3.1: (i) We establish the equation for the joint density function by using polar coordinates as follows,

$$
\begin{aligned}
q_2(x_1, x_2, \alpha) &= \frac{1}{4\pi^2}\left[\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} e^{-i(x_1z_1+x_2z_2)-\lambda(z_1^2+z_2^2)^{\alpha}} dz_1dz_2\right] \\
&= \frac{1}{4\pi^2}\left[\int_0^{2\pi}\int_0^{\infty} e^{-i(x_1k\cos\theta+x_2k\sin\theta)-\lambda k^{\alpha}} kdkd\theta\right] \\
&= \frac{1}{4\pi^2}\left[\int_0^{\pi}\int_0^{\infty} e^{-i(x_1k\cos\theta+x_2k\sin\theta)-\lambda k^{\alpha}} kdkd\theta\right] \\
&+ \frac{1}{4\pi^2}\left[\int_{\pi}^{2\pi}\int_0^{\infty} e^{-i(x_1k\cos\theta+x_2k\sin\theta)-\lambda k^{\alpha}} kdkd\theta\right].
\end{aligned}
$$

Therefore, by rewriting the sum of the last two terms in terms of I that satisfies

$$
I = \left[\int_0^{\pi}\int_0^{\infty} e^{-i(x_1k\cos\theta+x_2k\sin\theta)-\lambda k^{\alpha}} kdkd\theta\right],
$$

and the complex conjugate of I , which call it as $\bar{I}$, is defined by

$$
\begin{aligned}
\bar{I} &= \left[\int_0^{\pi}\int_0^{\infty} e^{i(x_1k\cos\theta+x_2k\sin\theta)-\lambda k^{\alpha}} kdkd\theta\right] \\
&= \left[\int_{\pi}^{2\pi}\int_0^{\infty} e^{-i(x_1k\cos\theta+x_2k\sin\theta)-\lambda k^{\alpha}} kdkd\theta\right],
\end{aligned}
$$

finally, the proof is completed by using the following identity

$$
q_2(x_1, x_2, \alpha) = 2\Re I.
$$

(ii) We substitute Equation (9) into the right-hand side of Equation (8), and change the integration order (this operation is valid since the double integral converges absolutely), thus we obtain

$$\begin{aligned}
\widetilde{q}_2(\lambda_1, \lambda_2, \alpha) &= \frac{1}{2\pi^2}\Re\int_0^{\pi}\int_0^{\infty}\int_0^{\infty}\int_0^{\infty} e^{-(\lambda_1 x_1+\lambda_2 x_2)-i(x_1 k\cos\theta+x_2 k\sin\theta)-\lambda k^{\alpha}} k dx_1 dx_2 dk d\theta \\
&= \frac{1}{2\pi^2}\Re\int_0^{\pi}\int_0^{\infty}\int_0^{\infty}\int_0^{\infty} e^{-(\lambda_1+ik\cos\theta)x_1-(\lambda_2+ik\sin\theta)x_2-\lambda k^{\alpha}} k dx_1 dx_2 dk d\theta \\
&= \frac{1}{2\pi^2}\Re\int_0^{\pi}\int_0^{\infty} \frac{1}{(\lambda_1+ik\cos\theta)(\lambda_2+ik\sin\theta)} e^{-\lambda k^{\alpha}} k dk d\theta \\
&= \frac{1}{\pi^2}\int_0^{\frac{\pi}{2}}\int_0^{\infty} \frac{\lambda_1\lambda_2}{(k^2\cos^2\theta+\lambda_1^2)(k^2\sin^2\theta+\lambda_2^2)} e^{-\lambda k^{\alpha}} k dk d\theta
\end{aligned}$$

(iii) Let

$$J = \int_0^{\infty}\int_0^{\pi} e^{-i(x_1 k\cos\theta+x_2 k\sin\theta)} e^{-\lambda(\theta)k^{\alpha} e^{\frac{-i\beta(\theta)\alpha\pi}{2}}} d\theta k dk,$$

and

$$\bar{J} = \int_0^{\infty}\int_0^{\pi} e^{i(x_1 k\cos\theta+x_2 k\sin\theta)} e^{-\lambda(\theta)k^{\alpha} e^{\frac{i\beta(\theta)\alpha\pi}{2}}} d\theta k dk,$$

then by using the subsequent properties that are $\beta(\theta) = -\beta(\theta-\pi)$, $\lambda(\theta) = \lambda(\theta-\pi)$, $J + \bar{J} = 2\Re J$, and after carrying out some algebraic calculations which yields

$$\begin{aligned}
q_2(x_1, x_2, \alpha) &= \frac{1}{4\pi^2}\left[\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} e^{-i(x_1 z_1+x_2 z_2)-\lambda(\theta)(z_1^2+z_2^2)^{\alpha/2} e^{\frac{-i\beta(\theta)\alpha\pi}{2}}} dz_1 dz_2\right] \\
&= \frac{1}{4\pi^2}\left[\int_0^{2\pi}\int_0^{\infty} e^{-i(x_1 k\cos\theta+x_2 k\sin\theta)-\lambda(\theta)k^{\alpha} e^{\frac{-i\beta(\theta)\alpha\pi}{2}}} k dk d\theta\right] \\
&= \frac{1}{4\pi^2}\left[\int_0^{\pi}\int_0^{\infty} e^{-i(x_1 k\cos\theta+x_2 k\sin\theta)-\lambda(\theta)k^{\alpha} e^{\frac{-i\beta(\theta)\alpha\pi}{2}}} k dk d\theta\right] \\
&+ \frac{1}{4\pi^2}\left[\int_{\pi}^{2\pi}\int_0^{\infty} e^{-i(x_1 k\cos\theta+x_2 k\sin\theta)-\lambda(\theta)k^{\alpha} e^{\frac{-i\beta(\theta)\alpha\pi}{2}}} k dk d\theta\right],
\end{aligned}$$

Therefore,

$$
\begin{aligned}
q_2(x_1, x_2, \alpha) &= \frac{1}{2\pi^2}\Re\left[\int_0^{\pi}\int_0^{\infty} e^{-i(x_1 k\cos\theta + x_2 k\sin\theta) - \lambda(\theta)k^{\alpha}e^{\frac{-i\beta(\theta)\alpha\pi}{2}}} k\,dk\,d\theta\right] \\
&= \frac{1}{2\pi^2}\Re\left[\int_0^{\pi}\int_0^{\infty} e^{i(x_1 k\cos\theta + x_2 k\sin\theta) - \lambda(\theta)k^{\alpha}e^{\frac{i\alpha\beta(\theta)\pi}{2}}} k\,dk\,d\theta\right] \\
&= \frac{-1}{2\pi^2}\Im\left[\int_0^{\pi}\int_0^{\infty} e^{-(x_1 k\cos\theta + x_2 k\sin\theta) - \lambda(\theta)k^{\alpha}e^{\frac{i\alpha(\beta(\theta)+1)\pi}{2}}} (ik)\,dk\,d\theta\right] \\
&= \frac{1}{2\pi^2}\Im\left[\int_0^{\pi}\int_0^{\infty} e^{-(x_1 k\cos\theta + x_2 k\sin\theta) - \lambda(\theta)k^{\alpha}e^{\frac{-i\alpha(\beta(\theta)+1)\pi}{2}}} (-ik)\,dk\,d\theta\right] \\
&= \frac{1}{2\pi^2}\Im\left[\int_0^{\pi}\int_0^{\infty} e^{-(x_1 k\cos\theta + x_2 k\sin\theta) - \lambda(\theta)k^{\alpha}e^{-i\rho(\theta)\pi}} (-ik)\,dk\,d\theta\right],
\end{aligned}
$$

where $\Im$ denote the imaginary part. Then we substitute the last equality into the right-hand side of Equation (8) and change the integration order, hence the result implies as follows,

$$
\begin{aligned}
\widetilde{q}_2(\lambda_1, \lambda_2, \alpha) &= \frac{1}{2\pi^2}\int_0^{\infty}\int_0^{\infty} e^{-(\lambda_1 x_1 + \lambda_2 x_2)} \\
&\quad \Im\left[\int_0^{\pi}\int_0^{\infty} e^{-(x_1 k e^{\frac{i\rho(\theta)\pi}{2\alpha}}\cos\theta + x_2 k e^{\frac{i\rho(\theta)\pi}{2\alpha}}\sin\theta) - \lambda(\theta)k^{\alpha}e^{-\frac{i\rho(\theta)\pi}{2}}} (-ike^{\frac{i\rho(\theta)\pi}{2\alpha}})e^{\frac{i\rho(\theta)\pi}{2\alpha}}\,dk\,d\theta\right] dx_1\,dx_2 \\
&= \frac{1}{2\pi^2}\Im\int_0^{\pi}\int_0^{\infty}\left[\int_0^{\infty} e^{-(\lambda_1 + ke^{\frac{i\rho(\theta)\pi}{2\alpha}}\cos\theta)x_1} dx_1\right]\left[\int_0^{\infty} e^{-(\lambda_2 + ke^{\frac{i\rho(\theta)\pi}{2\alpha}}\sin\theta)x_2} dx_2\right] \\
&\quad \times e^{-\lambda(\theta)k^{\alpha}e^{-\frac{i\rho(\theta)\pi}{2}}}(-ike^{\frac{i\rho(\theta)\pi}{2\alpha}})e^{\frac{i\rho(\theta)\pi}{2\alpha}}\,dk\,d\theta \\
&= \frac{1}{2\pi^2}\Im\int_0^{\pi}\int_0^{\infty}\left[\frac{e^{\frac{i\rho(\theta)\pi}{2\alpha}}}{ke^{\frac{i\rho(\theta)\pi}{2\alpha}}\cos\theta + \lambda_1}\right]\left[\frac{e^{\frac{i\rho(\theta)\pi}{2\alpha}}}{ke^{\frac{i\rho(\theta)\pi}{2\alpha}}\sin\theta + \lambda_2}\right] e^{-\lambda(\theta)k^{\alpha}e^{-\frac{i\rho(\theta)\pi}{2}}}(-ik)\,dk\,d\theta \\
&= \frac{1}{2\pi^2}\Im\int_0^{\pi}\int_0^{\infty}\left[\frac{e^{\frac{i\rho(\theta)\pi}{\alpha}}}{ke^{\frac{i\rho(\theta)\pi}{\alpha}}\cos\theta + \lambda_1}\right]\left[\frac{e^{\frac{i\rho(\theta)\pi}{\alpha}}}{ke^{\frac{i\rho(\theta)\pi}{\alpha}}\sin\theta + \lambda_2}\right] e^{-\lambda(\theta)k^{\alpha}}(-ik)\,dk\,d\theta \\
&= \frac{1}{2\pi^2}\Im\int_0^{\pi}\int_0^{\infty}\left[\frac{\lambda_1 e^{\frac{i\rho(\theta)\pi}{\alpha}} + k\cos\theta}{\left|\lambda_1 + ke^{\frac{i\rho(\theta)\pi}{\alpha}}\cos\theta\right|^2}\right]\left[\frac{\lambda_2 e^{\frac{i\rho(\theta)\pi}{\alpha}} + k\sin\theta}{\left|ke^{\frac{i\rho(\theta)\pi}{\alpha}}\sin\theta + \lambda_2\right|^2}\right] e^{-\lambda(\theta)k^{\alpha}}(-ik)\,dk\,d\theta
\end{aligned}
$$

$$\widetilde{q}_2(\lambda_1, \lambda_2, \alpha) = \frac{1}{2\pi^2} \int_0^{\pi} \int_0^{\infty} \Im \left[(-i) \frac{\lambda_1 e^{\frac{i\rho(\theta)\pi}{\alpha}} + k\cos\theta}{\left|\lambda_1 + k e^{\frac{i\rho(\theta)\pi}{\alpha}} \cos\theta\right|^2} \frac{\lambda_2 e^{\frac{i\rho(\theta)\pi}{\alpha}} + k\sin\theta}{\left|k e^{\frac{i\rho(\theta)\pi}{\alpha}} \sin\theta + \lambda_2\right|^2} \right] e^{-\lambda(\theta)k^\alpha} k dk d\theta$$
$$= \frac{-1}{2\pi^2} \int_0^{\pi} \int_0^{\infty} \frac{A(k,\theta)}{B(k,\theta)} e^{-\lambda(\theta)k^\alpha} k dk d\theta,$$

where

$$\frac{A(k,\theta)}{B(k,\theta)} = \frac{k^2\cos\theta\sin\theta + \lambda_2 k\cos\theta\cos\frac{\rho(\theta)\pi}{\alpha} + \lambda_1 k\sin\theta\cos\frac{\rho(\theta)\pi}{\alpha} + \lambda_1\lambda_2\cos\frac{2\rho(\theta)\pi}{\alpha}}{(k^2\cos^2\theta + 2\lambda_1 k\cos\theta\cos\frac{\rho(\theta)\pi}{\alpha} + \lambda_1^2)(k^2\sin^2\theta + 2\lambda_2 k\sin\theta\cos\frac{\rho(\theta)\pi}{\alpha} + \lambda_2^2)}.$$

□

Proof of Theorem 3.2: One can rewrite the expression in Definition 3.1 as,

$$q_2(r,\varphi,\alpha) = \frac{1}{2\pi^2} \Im \left[\int_0^{\pi} \int_0^{\infty} e^{-(rk\cos\theta\cos\varphi + rk\sin\theta\sin\varphi) - \lambda(\theta)k^\alpha e^{-i\rho(\theta)\pi}} (-ik) dk d\theta \right]$$
$$= \frac{1}{2\pi^2} \Im \left[\int_0^{\pi} \int_0^{\infty} e^{-rk\cos(\theta-\varphi) - \lambda(\theta)k^\alpha e^{-i\rho(\theta)\pi}} (-ik) dk d\theta \right].$$

Therefore, the corresponding Laplace transform is given by

$$\widetilde{q}_{R|\varphi=\theta_0}(y, r, \theta_0, \alpha) = \int_0^{\infty} e^{-yr} q_{R|\varphi=\theta_0}(r, \theta_0, \alpha) dr$$
$$= \frac{1}{w(\theta_0)} \int_0^{\infty} e^{-yr} q_2(r, \theta_0, \alpha) dr$$
$$= \frac{1}{2\pi^2 w(\theta_0)} \int_0^{\infty} e^{-yr} \Im \left[\int_0^{\pi} \int_0^{\infty} e^{-rk\cos(\theta-\theta_0) - \lambda(\theta)k^\alpha e^{-i\rho(\theta)\pi}} (-ik) dk d\theta \right] dr$$
$$= \frac{1}{2\pi^2 w(\theta_0)} \Im \left[\int_0^{\pi} \int_0^{\infty} \int_0^{\infty} e^{-r(y + k\cos(\theta-\theta_0)) - \lambda(\theta)k^\alpha e^{-i\rho(\theta)\pi}} (-ik) dk d\theta \right] dr$$
$$= \frac{1}{2\pi^2 w(\theta_0)} \Im \int_0^{\pi} \int_0^{\infty} \frac{1}{y + k\cos(\theta-\theta_0)} e^{-\lambda(\theta)k^\alpha e^{-i\rho(\theta)\pi}} (-ik) dk d\theta$$
$$= \frac{1}{2\pi^2 w(\theta_0)} \Im \int_0^{\pi} \int_0^{\infty} \frac{e^{-\lambda(\theta)k^\alpha}}{y + k\cos(\theta-\theta_0) e^{\frac{i\rho(\theta)\pi}{\alpha}}} (-ik e^{\frac{i\rho(\theta)\pi}{\alpha}}) e^{\frac{i\rho(\theta)\pi}{\alpha}} dk d\theta,$$

Finally, we observe that

$$\begin{aligned}
\widetilde{q}_{R|\varphi=\theta_0}(y,r,\theta_0,\alpha) &= \frac{1}{2\pi^2 w(\theta_0)}\Im\int_0^\pi\int_0^\infty \frac{-ie^{\frac{2i\rho(\theta)\pi}{\alpha}}}{y+k\cos(\theta-\theta_0)e^{\frac{i\rho(\theta)\pi}{\alpha}}}e^{-\lambda(\theta)k^\alpha}kdkd\theta \\
&= \frac{1}{2\pi^2 w(\theta_0)}\int_0^\pi\int_0^\infty \frac{-C}{D}e^{-\lambda(\theta)k^\alpha}kdkd\theta, \\
&= \frac{1}{2\pi^2 w(\theta_0)}\Im\int_0^\pi\int_0^\infty \frac{-ie^{\frac{2i\rho(\theta)\pi}{\alpha}}}{y+k\cos(\theta-\theta_0)e^{\frac{i\rho(\theta)\pi}{\alpha}}}e^{-\lambda(\theta)k^\alpha}kdkd\theta \\
&= \frac{1}{2\pi^2 w(\theta_0)}\int_0^\pi\int_0^\infty \frac{-C}{D}e^{-\lambda(\theta)k^\alpha}kdkd\theta,
\end{aligned}$$

where

$$\frac{C}{D} = \frac{y\cos(\frac{2\rho(\theta)\pi}{\alpha}) + k\cos(\frac{\rho(\theta)\pi}{\alpha})\cos(\theta-\theta_0)}{(k^2\cos^2(\theta-\theta_0)+2yk\cos(\theta-\theta_0)\cos\frac{\rho(\theta)\pi}{\alpha}+y^2)}.$$

The proof of the first part is completed. Now, in this step we continue to prove the second part of the theorem. We have

$$q_2(x_1,ax_1,\alpha) = \frac{1}{2\pi^2}\Im\int_0^\pi\int_0^\infty e^{-x_1k\cos\theta-ax_1k\sin\theta-\lambda(\theta)k^\alpha e^{-i\rho(\theta)\pi}}(-ik)\,dkd\theta,$$

where $\rho(\theta)$ is defined in Theorem 3.1, then it allows us to obtain the univariate Laplace transform as follows,

$$\begin{aligned}
\widetilde{q}_2(\lambda_1,\alpha) &= \int_0^\infty e^{-\lambda_1 x_1} q_2(x_1,ax_1,\alpha)dx_1 \\
&= \frac{1}{2\pi^2}\Im\int_0^\infty e^{-\lambda_1 x_1}\int_0^\pi\int_0^\infty e^{-x_1k\cos\theta-ax_1k\sin\theta-\lambda(\theta)k^\alpha e^{-i\rho(\theta)\pi}}(-ik)\,dkd\theta dx_1 \\
&= \frac{1}{2\pi^2}\Im\int_0^\pi\int_0^\infty\int_0^\infty e^{-(\lambda_1+k\cos\theta+ak\sin\theta)x_1}dx_1 e^{-\lambda(\theta)k^\alpha e^{-i\rho(\theta)\pi}}(-ik)\,dkd\theta \\
&= \frac{1}{2\pi^2}\Im\int_0^\pi\int_0^\infty \frac{1}{\lambda_1+k\cos\theta+ak\sin\theta}e^{-\lambda(\theta)k^\alpha e^{-i\rho(\theta)\pi}}(-ik)\,dkd\theta \\
&= \frac{1}{2\pi^2}\Im\int_0^\pi\int_0^\infty \frac{e^{\frac{i\rho(\theta)\pi}{\alpha}}}{\lambda_1+ke^{\frac{i\rho(\theta)\pi}{\alpha}}(\cos\theta+a\sin\theta)}e^{-\lambda(\theta)k^\alpha}(-ik)\,e^{\frac{i\rho(\theta)\pi}{\alpha}}dkd\theta.
\end{aligned}$$

Then by using some algebraic calculations such as,

$$\frac{1}{\lambda_1 + ke^{\frac{i\rho(\theta)\pi}{\alpha}}(\cos\theta + a\sin\theta)} = \frac{\lambda_1 + ke^{\frac{-i\rho(\theta)\pi}{\alpha}}\cos\theta + ake^{\frac{-i\rho(\theta)\pi}{\alpha}}\sin\theta}{E},$$

where

$$E = \lambda_1^2 + 2\lambda_1 k\cos\frac{\rho(\theta)\pi}{\alpha}(\cos\theta + a\sin\theta) + k^2(\cos\theta + a\sin\theta)^2,$$

and

$$\frac{ie^{2\frac{i\rho(\theta)\pi}{\alpha}}}{\lambda_1 + ke^{\frac{i\rho(\theta)\pi}{\alpha}}(\cos\theta + a\sin\theta)} = \frac{i\lambda_1 e^{\frac{2i\rho(\theta)\pi}{\alpha}} + ike^{\frac{i\rho(\theta)\pi}{\alpha}}\cos\theta + iake^{\frac{i\rho(\theta)\pi}{\alpha}}\sin\theta}{E},$$

which allows us to easily see that,

$$\Im\left(\frac{e^{2\frac{i\rho(\theta)\pi}{\alpha}}}{\lambda_1 + ke^{\frac{i\rho(\theta)\pi}{\alpha}}(\cos\theta + a\sin\theta)}\right) = \frac{\lambda_1\cos(\frac{2\rho(\theta)\pi}{\alpha}) + k\cos\theta\cos(\frac{\rho(\theta)\pi}{\alpha}) + ak\sin\theta\cos(\frac{\rho(\theta)\pi}{\alpha})}{E}.$$

Thus, the result is implied. □