Parametric estimation of a bivariate stable Lévy process

Habib Esmaeili * Claudia Klüppelberg†

January 31, 2011

Abstract

We propose a parametric model for a bivariate stable Lévy process based on a Lévy copula as a dependence model. We estimate the parameters of the full bivariate model by maximum likelihood estimation. As an observation scheme we assume that we observe all jumps larger than some $\varepsilon > 0$ and base our statistical analysis on the resulting compound Poisson process. We derive the Fisher information matrix and prove asymptotic normality of all estimates, when the truncation point ε tends to 0. A simulation study investigates the loss of efficiency because of the truncation.

AMS 2000 Subject Classifications: 62F10, 62F12, 62M05.

Keywords: Lévy copula, maximum likelihood estimation, dependence structure, Fisher information matrix, multivariate stable process, parameter estimation.

*Center for Mathematical Sciences, Technische Universität München, D-85748 Garching, Germany, Email: esmaeili@ma.tum.de, URL: http://www-m4.ma.tum.de/pers/esmaeili/

†Center for Mathematical Sciences, and Institute for Advanced Study, Technische Universität München, D-85748 Garching, Germany, Email: cklu@ma.tum.de, URL: http://www-m4.ma.tum.de

1 Introduction

The problem of parameter estimation of one-dimensional stable Lévy processes has been investigated already in the seventies of the last century by Basawa and Brockwell [3, 4]. Starting with a subordinator model, they assumed that it is possible to observe $n^{(\varepsilon)}$ jumps in a time interval $[0, t]$, all larger than a certain small $\varepsilon > 0$. Based on this observation scheme, they estimated the parameters by a maximum likelihood procedure, and investigated the distributional limits of the MLEs for $n^{(\varepsilon)} \to \infty$, which in this model if either $t \to \infty$ and/or $\varepsilon \to 0$.

The task of estimating multivariate stable processes is usually solved by estimating the parameters of the marginal processes and the spectral measure separately; cf. Nolan, Panorska and McCulloch [14] and Höpfner [10] and references therein.

The rather recent modelling of multivariate Lévy processes by their marginal processes and a Lévy copula for the dependence structure (cf. Cont and Tankov [6], Kallsen and Tankov [12], and Eder and Klüppelberg [7]) allows for the construction of new parametric models. This approach is similar to the representation of a multivariate distribution function by its marginal distributions and a copula and is valid for all multivariate Lévy processes.

Moreover, various estimation methods of the parameters of the marginal processes and the dependence structure either together or separately can be applied. Obviously, it is more efficient to estimate all parameters of a model in one go, but often the attempt fails. Problems may occur because of the complexity of the numerical optimization involved to obtain the MLEs of the parameters or, given the estimates, their asymptotic properties are not clear concerning their asymptotic covariance structure.

This is an important point in the context of Lévy processes, since these properties may depend on the observation scheme. In reality it is usually not possible to observe the continuous-time sample path, but it may be possible to observe all jumps larger than ε as in the one-dimensional problem studied by Basawa and Brockwell [4]. For a stable subordinator we obtain asymptotic normality for such an observation scheme, provided that $n^{(\varepsilon)} \to \infty$, equivalently $\varepsilon \to 0$. In a general multivariate model this is not clear at all, in particular, for the dependence parameters.

With this paper we want to start an investigation concerning statistical estimation of multivariate Lévy processes in a parametric framework. In a certain sense the present paper is a follow-up of Esmaeili and Klüppelberg [8], where we concentrated on parametric estimation of multivariate compound Poisson processes.

Since our observation scheme involves only jumps larger than ε, the observed process is a multivariate compound Poisson process. But in contrast to [8], we now assume that the Lévy process has infinite Lévy measure and we investigate asymptotic normality also for $\varepsilon \to 0$.

Our paper is organised as follows. In Section 2 we present some basic facts about Lévy copulas and recall the estimation procedure as presented in Basawa and Brockwell [3, 4] for one-dimensional α-stable subordinators in Section 3. Section 4 contains the theoretical body of our new results. In Section 4.1 we present the small jumps truncation and its consequences for the Lévy copula; Section 4.2 presents the maximum likelihood estimation for the α-stable Clayton subordinator, including an explicit calculation of the Fisher information matrix, which ensures joint asymptotic normality of all estimates. Section 5 presents a simulation study and Section 6 concludes and gives an outlook to further work.

2 Lévy processes and Lévy copulas

Let $\mathbf{S} = (\mathbf{S}(t))_{t\geq 0}$ be a Lévy process with values in $\mathbb{R}^d$ defined on a filtered probability space $(\Omega, (\mathcal{F}_t)_{t\geq 0}, \mathcal{F}, \mathcal{P})$; i.e $\mathbf{S}$ has independent and stationary increments, and we assume that it has cádlág sample paths. For each $t > 0$, the random variable $\mathbf{S}(t)$ has an infinitely divisible distribution, whose characteristic function has a Lévy-Khintchine representation:

$$E[e^{i(z,X_t)}] = \exp\left\{t\left(i(\gamma, z) - \frac{1}{2}z^\top Az + \int_{\mathbb{R}^d}\left(e^{i(z,x)} - 1 - i(z,x)1_{|x|\leq 1}\right)\Pi(dx)\right)\right\},\ z \in \mathbb{R}^d\,,$$

where $(\cdot,\cdot)$ denotes the inner product in $\mathbb{R}^d$, $\gamma \in \mathbb{R}^d$ and A is a symmetric nonnegative definite $d \times d$ matrix. The *Lévy measure* Π is a measure on $\mathbb{R}^d$ satisfying $\Pi(\{\mathbf{0}\}) = 0$ and $\int_{\mathbb{R}^d\setminus\{\mathbf{0}\}} \min\{1, |x|^2\}\Pi(dx) < \infty$. For every Lévy process its distribution is defined by (γ, A, Π), which is called the *characteristic triplet*. It is worth mentioning that the Lévy measure $\Pi(B)$ for $B \in \mathcal{B}(\mathbb{R}^d)$ is the expected number of jumps per unit time with size in B.

Brownian motion is characterised by $(0, A, 0)$ and Brownian motion with drift by $(\gamma, A, 0)$. Poisson processes and compound Poisson processes have characteristic triplet $(\gamma_1, 0, \Pi)$. The class of Lévy processes is very rich including prominent examples like stable processes, gamma processes, variance gamma processes, inverse Gaussian and normal inverse Gaussian processes. Their applications reach from finance and insurance applications to the natural sciences and engineering. A particular role is played by *subordinators*, which are Lévy processes with increasing sample paths. Other important classes are *spectrally one-sided Lévy processes*, which have only positive or only negative jumps.

We are concerned with dependence in the jump behaviour $\mathbf{S}$, which we model by an appropriate functional of the marginals of the Lévy measure Π. Since, with the exception of a compound Poisson model, all Lévy measures have a singularity in 0, we follow Cont and Tankov [6] and introduce a (survival) copula on the *tail integral*, which is called *Lévy copula* and, because of the singularity in 0, is defined for each quadrant separately; for details we refer to Kallsen and Tankov [12] and to Eder and Klüppelberg [7] for a different approach.

Throughout this paper we restrict the presentation to the positive cone $\mathbb{R}_+^d$, where only common positive jumps in all component processes happen. To extend this theory to general Lévy processes is not difficult, but notationally involved.

We present the definition of the tail integral on the positive cone $\mathbb{R}_+^d$. For a spectrally positive Lévy process this characterises the jump behaviour completely.

Definition 2.1. *Let* Π *be a Lévy measure on* $\mathbb{R}_+^d$*. The* tail integral *is a function* $\overline{\Pi} : [0,\infty]^d \to [0,\infty]$ *defined by*

$$\overline{\Pi}(x_1,\ldots,x_d) = \begin{cases} \Pi([x_1,\infty)\times\cdots\times[x_d,\infty))\,, & (x_1,\ldots,x_d)\in[0,\infty)^d\setminus\{\mathbf{0}\} \\ 0\,, & x_i=\infty \text{ for at least one } i \\ \infty\,, & (x_1,\ldots,x_d)=\mathbf{0}. \end{cases}$$

The marginal tail integrals are defined analogously for $i=1,\ldots,d$ *as* $\overline{\Pi}_i(x)=\Pi_i([x,\infty))$ *for* $x\geq 0$*.*

Also the Lévy copula is defined quadrantwise and characterises the dependence structure of a spectrally positive Lévy process completely.

Definition 2.2. *A d-dimensional* positive Lévy copula *is a measure defining function* $\mathfrak{C} : [0,\infty]^d \to [0,\infty]$ *with margins* $\mathfrak{C}_k(u)=u$ *for all* $u\in[0,\infty]$ *and* $k=1,\ldots,d$*.*

The following theorem is a version of Sklar's theorem for Lévy processes with positive jumps, proved in Tankov [17], Theorem 3.1; for the corresponding result for general Lévy processs we refer again to Kallsen and Tankov [12].

Theorem 2.3 (Sklar's Theorem for Lévy copulas)**.** *Let* $\overline{\Pi}$ *denote the tail integral of a spectrally positive d-dimensional Lévy process, whose components have Lévy measures* $\Pi_1,\ldots,\Pi_d$*. Then there exists a Lévy copula* $\mathfrak{C} : [0,\infty]^d \to [0,\infty]$ *such that for all* $x_1,x_2,\ldots,x_d\in[0,\infty]$

$$\overline{\Pi}(x_1,\ldots,x_d) = \mathfrak{C}(\overline{\Pi}_1(x_1),\ldots,\overline{\Pi}_d(x_d)). \tag{2.1}$$

If the marginal tail integrals are continuous, then this Lévy copula is unique. Otherwise, it is unique on $Ran\overline{\Pi}_1\times\cdots\times Ran\overline{\Pi}_d$*.*
Conversely, if $\mathfrak{C}$ *is a Lévy copula and* $\overline{\Pi}_1,\ldots,\overline{\Pi}_d$ *are marginal tail integrals of a spectrally positive Lévy process, then the relation* (2.1) *defines the tail integral of a d-dimensional spectrally positive Lévy process and* $\overline{\Pi}_1,\ldots,\overline{\Pi}_d$ *are tail integrals of its components.*

Remark 2.4. *In the case of multivariate stable Lévy processes the Lévy copula carries the same information as the spectral measure. By choosing a slightly different approach this was shown in Eder and Klüppelberg [7]. Note, however, that the spectral measure restricts to stable processes, whereas the Lévy copula models the dependence for all Lévy processes.*

We are concerned with the estimation of the parameters of a multivariate Lévy process and assume for simplicity that we observe all jumps larger than $\varepsilon > 0$ of a subordinator. This results in a compound Poisson process and we recall the following well-known results; see e.g. Sato [16], Theorem 21.2 and Corollary 8.8.

Proposition 2.5. *(a) A Lévy process* $\mathbf{S}$ *in* $\mathbb{R}^d$ *is compound Poisson if and only if it has a finite Lévy measure* Π *with* $\lim_{\mathbf{x}\to\mathbf{0}} \overline{\Pi}(\mathbf{x}) = \lambda$*, the intensity of the* d*-dimensional Poisson process, and jump distribution* $F(d\mathbf{x}) = \lambda^{-1}\Pi(d\mathbf{x})$ *for* $\mathbf{x} \in \mathbb{R}^d$*.*

(b) Every Lévy process is the limit of a sequence of compound Poisson processes.

3 Maximum likelihood estimation of the parameters of a one-dimensional Lévy process

3.1 Small jump truncation

With the understanding that the Lévy measure can be decomposed in positive and negative jumps we restrict ourselves to subordinators.

Let S be a one-dimensional subordinator with unbounded Lévy measure Π, without drift or Gaussian part. For all $t \geq 0$ its characteristic function has the representation $Ee^{iuS(t)} = e^{t\psi(u)}$ for $u \in \mathbb{R}$ with

$$\psi(u) \;=\; \int_{0<x<\varepsilon} \big(e^{iux} - 1\big)\Pi(dx) + \int_{x\geq\varepsilon} \big(e^{iux} - 1\big)\Pi(dx), \quad u \in \mathbb{R}\,, \tag{3.1}$$

for arbitrary $\varepsilon > 0$. The last integral in (3.1) is the characteristic exponent of a compound Poisson process with Poisson intensity $\lambda^{(\varepsilon)} \in (0,\infty)$ and jump distribution function $F^{(\varepsilon)}$

$$\lambda^{(\varepsilon)} = \int_{\varepsilon}^{\infty} \Pi(dx) \quad \text{and} \quad F^{(\varepsilon)}(dx) = \Pi(dx)/\lambda^{(\varepsilon)} \text{ on } [\varepsilon,\infty).$$

As an observation scheme we assume that we observe the whole sample path of S over a time interval $[0,t]$, but that we only observe jumps of size larger than ε. Then our observation scheme is equivalent to observing a compound Poisson process, say $S^{(\varepsilon)}$, given in its marked point process representation as $\{(T_k^{(\varepsilon)}, X_k^{(\varepsilon)}), k = 1,\ldots,n^{(\varepsilon)}\}$, where $n^{(\varepsilon)} = n^{(\varepsilon)}(t) = \text{card}\{T_k^{(\varepsilon)} \in [0,t] : k \in \mathbb{N}\}$. We also assume that $\Pi(dx) = \nu(x;\theta)dx$ where θ is a vector of parameters of the Lévy measure so that the density of $X_k^{(\varepsilon)}$ is given by $f^{(\varepsilon)}(x;\theta) = \nu(x,\theta)/\lambda^{(\varepsilon)}$ for $x \geq \varepsilon$. The likelihood function of this compound Poisson process is well-known, see e.g. Basawa and Prakasa Rao [5], and is given by

$$L^{(\varepsilon)}(\theta) \;=\; (\lambda^{(\varepsilon)})^{n^{(\varepsilon)}} e^{-\lambda^{(\varepsilon)}t} \times \prod_{i=1}^{n^{(\varepsilon)}} f^{(\varepsilon)}(x_i,\theta) \;=\; e^{-\lambda^{(\varepsilon)}t} \times \prod_{i=1}^{n^{(\varepsilon)}} \nu(x_i;\theta)\mathbf{1}_{\{x_i \geq \varepsilon\}}. \quad (3.2)$$

3.2 Asymptotic behaviour of the MLEs

MLE is a well established estimation procedure and the asymptotic properties of the estimators is well-known for iid data, but also for continuous-time stochastic processes, see e.g. Küchler and Sorensen [13] and references therein. However, this theory is usually concerned about letting the observation time, i.e. t tends to infinity. We are more interested in the case of fixed t and $\varepsilon \downarrow 0$, and here there exist to our knowledge only some specific results in the literature; see e.g. Basawa and Brockwell [3, 4] and Höpfner and Jacod [11].

We start with a general Lévy process S and base the maximum likelihood estimation on the jumps $\Delta S_v > \varepsilon$ for $v \in [0,t]$. The MLEs are, in fact, those obtained from the CPP $S^{(\varepsilon)}$ as described in Section 3.1 above. Therefore, under some regularity conditions (see e.g. Prakasa Rao [15], Section 3.11) the MLEs are consistent and asymptotically normal. In the context of a compound Poisson process the asymptotic behavior of estimators is considered for $t \to \infty$. In our set-up, however, it is also relevant to consider the performance of estimators as $\varepsilon \to 0$ with t fixed.

We investigate the asymptotic behavior of estimators for a stable Lévy process as $n^{(\varepsilon)} \to \infty$ and shall show that this covers the cases of $t \to \infty$ as well as $\varepsilon \to 0$. Asymptotic normality of the estimators has been derived in Basawa and Brockwell [3, 4]. For comparison and later reference we summarize these results in some detail.

Example 3.1. [α-stable subordinator]
Let $(S(t))_{t\geq 0}$ be a one dimensional α-stable subordinator with parameters $c > 0$ and $0 < \alpha < 1$, such that the tail integral $\overline{\Pi}(x) = cx^{-\alpha}$ for $x > 0$. Observing all jumps larger than some $\varepsilon > 0$, the resulting CPP has intensity and jump size density

$$\lambda^{(\varepsilon)} = \int_{\varepsilon}^{\infty} \Pi(dx) = c\varepsilon^{-\alpha} \qquad , \qquad f^{(\varepsilon)}(x) = \frac{\Pi(dx)/dx}{\lambda^{(\varepsilon)}} = \alpha\varepsilon^{\alpha}x^{-1-\alpha}, \quad x > \varepsilon.$$

If we observe $n^{(\varepsilon)}$ jumps larger than ε in $[0,t]$, we estimate the intensity by $\widehat{\lambda}^{(\varepsilon)} = \frac{n^{(\varepsilon)}}{t}$. Moreover, by (3.2) the loglikelihood function for $\theta = (\alpha, \log c)$ is given by

$$\ell(\alpha, c) = n^{(\varepsilon)}(\log\alpha + \log c) - e^{\log c}\varepsilon^{-\alpha}t - (1+\alpha)\sum_{i=1}^{n^{(\varepsilon)}} \log x_i\,.$$

We calculate the score functions as

$$\begin{aligned}
\frac{\partial \ell}{\partial \alpha} &= \frac{n^{(\varepsilon)}}{\alpha} + te^{\log c}\varepsilon^{-\alpha}\log\varepsilon - \sum_{i=1}^{n^{(\varepsilon)}} \log x_i\,, \\
\frac{\partial \ell}{\partial \log c} &= n^{(\varepsilon)} - e^{\log c}\varepsilon^{-\alpha} t
\end{aligned}$$

To obtain candidates for maxima we calculate

$$\begin{aligned}
e^{\log c} &= \frac{n^{(\varepsilon)}}{t\varepsilon^{-\alpha}} = \frac{\widehat{\lambda}^{(\varepsilon)}}{\varepsilon^{-\alpha}}\,, \\
\frac{1}{\alpha} &= -\frac{t}{n^{(\varepsilon)}} c\varepsilon^{-\alpha}\log\varepsilon + \frac{1}{n^{(\varepsilon)}}\sum_{i=1}^{n^{(\varepsilon)}} \log x_i \\
&= \frac{1}{n^{(\varepsilon)}}\sum_{i=1}^{n^{(\varepsilon)}} (\log x_i - \log\varepsilon) + \log\varepsilon\Big(1 - \frac{\lambda^{(\varepsilon)}}{\widehat{\lambda}^{(\varepsilon)}}\Big)\,.
\end{aligned}$$

Consequently, we have the maximum likelihood estimators

$$\begin{aligned}
\widehat{\alpha} &= \left(\frac{1}{n^{(\varepsilon)}}\sum_{i=1}^{n^{(\varepsilon)}} \Big(\log X_i^{(\varepsilon)} - \log\varepsilon\Big) + \log\varepsilon\Big(1 - \frac{\lambda^{(\varepsilon)}}{\widehat{\lambda}^{(\varepsilon)}}\Big)\right)^{-1}\,, \\
\widehat{\log c} &= \log\widehat{\lambda}^{(\varepsilon)} + \widehat{\alpha}\log\varepsilon\,.
\end{aligned}$$

Next we calculate the second derivatives as

$$\begin{aligned}
\frac{\partial^2 \ell}{\partial \alpha^2} &= -n^{(\varepsilon)}\frac{1}{\alpha^2} - ct\varepsilon^{-\alpha}(\log\varepsilon)^2 = -tc\varepsilon^{-\alpha}\left(\frac{\widehat{\lambda}^{(\varepsilon)}}{\lambda^{(\varepsilon)}}\frac{1}{\alpha^2} + (\log\varepsilon)^2\right) \\
\frac{\partial^2 \ell}{\partial \alpha\,\partial \log c} &= ct\varepsilon^{-\alpha}\log\varepsilon = \frac{\partial^2 \ell}{\partial \log c\,\partial \alpha} \\
\frac{\partial^2 \ell}{\partial (\log c)^2} &= -tc\varepsilon^{-\alpha}\,.
\end{aligned}$$

Consequently, the Fisher information matrix is given by

$$I^{(\varepsilon)}_{\alpha,\log c} = tc\varepsilon^{-\alpha}\begin{pmatrix} \frac{1}{\alpha^2} + (\log\varepsilon)^2 & -\log\varepsilon \\ -\log\varepsilon & 1 \end{pmatrix}.$$

We calculate the determinant as $\det(I^{(\varepsilon)}_{\alpha,\log c}) = c^2t^2\alpha^{-2}\varepsilon^{-2\alpha}$. Using Cramer's rule of inversion easily gives

$$(I^{(\varepsilon)}_{\alpha,\log c})^{-1} = (ct)^{-1}\varepsilon^{\alpha}\alpha^2\begin{pmatrix} 1 & \log\varepsilon \\ \log\varepsilon & \frac{1}{\alpha^2} + (\log\varepsilon)^2 \end{pmatrix}.$$

We are interested in the asymptotic behaviour of the MLEs $\widehat{\alpha}$ and $\widehat{\log c}$ based on a fixed time interval $[0,t]$ and letting $\varepsilon \to 0$. Note that we have to get the variance-covariance matrix asymptotically independent of ε. Division of $\widehat{\log c}$ by $\log\varepsilon$ changes the matrix $I^{(\varepsilon)^{-1}}_{\alpha,\log c}$ into

$$(\widetilde{I}^{(\varepsilon)}_{\alpha,\frac{\log c}{\log\varepsilon}})^{-1} = t^{-1}c^{-1}\varepsilon^{\alpha}\alpha^2 \begin{pmatrix} 1 & 1 \\ 1 & \frac{1}{\alpha^2(\log\varepsilon)^2}+1 \end{pmatrix}.$$

Since

$$\sqrt{n^{(\varepsilon)}}\left(\frac{\widehat{\lambda}^{(\varepsilon)}}{\lambda^{(\varepsilon)}}-1\right) = \sqrt{n^{(\varepsilon)}}\left(\frac{n^{(\varepsilon)}}{tc\varepsilon^{-\alpha}}-1\right) \xrightarrow{d} N(0,1)\,, \quad n^{(\varepsilon)} \to \infty\,, \tag{3.3}$$

and the regularity conditions of Section 3.11 of Prakasa Rao [15] are satisfied, classical likelihood theory ensures that

$$\sqrt{n^{(\varepsilon)}}\begin{pmatrix} \widehat{\alpha}-\alpha \\ \dfrac{\log\widehat{c}-\log c}{\log\varepsilon} \end{pmatrix} \sim \mathrm{AN}\left(\mathbf{0}, \alpha^2\begin{pmatrix} 1 & 1 \\ 1 & \frac{1}{\alpha^2}\frac{1}{(\log\varepsilon)^2}+1 \end{pmatrix}\right), \quad n^{(\varepsilon)} \to \infty\,.$$

Consistency of $\widehat{\lambda}^{(\varepsilon)}$, obtained from (3.3), and a Taylor expansion of $\log x$ around c ensures with Slutzky's theorem that for $\varepsilon \to 0$,

$$\sqrt{ct}\varepsilon^{-\alpha/2}\begin{pmatrix} \dfrac{\widehat{\alpha}}{\alpha}-1 \\ \frac{1}{\alpha\log\varepsilon}\left(\dfrac{\widehat{c}}{c}-1\right) \end{pmatrix} \xrightarrow{d} N\left(\mathbf{0}, \begin{pmatrix} 1 & 1 \\ 1 & 1 \end{pmatrix}\right) = \begin{pmatrix} N_1 \\ N_2 \end{pmatrix},$$

where N_1, N_2 are standard normal random variables with $\mathrm{Cov}(N_1,N_2)=1$, which implies that $N_1 = N_2 = N$. So the limit law is degenerate.

It has been shown in Jacod and Höpfner [11] that the natural parameterization is not (c,α), but $(\lambda^{(\varepsilon)},\alpha)$, which leads to asymptotically independent normal limits. Indeed, we have

$$\sqrt{n^{(\varepsilon)}}\begin{pmatrix} \dfrac{\widehat{\alpha}}{\alpha}-1 \\ \dfrac{\widehat{\lambda}^{(\varepsilon)}}{\lambda^{(\varepsilon)}}-1 \end{pmatrix} \xrightarrow{d} N\left(\mathbf{0}, \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}\right), \quad n^{(\varepsilon)} \to \infty\,,$$

where $n^{(\varepsilon)}$ can again be replaced by $tc\varepsilon^{-\alpha}$ and the same result holds for $t\to\infty$, equivalently, $\varepsilon \to 0$. □

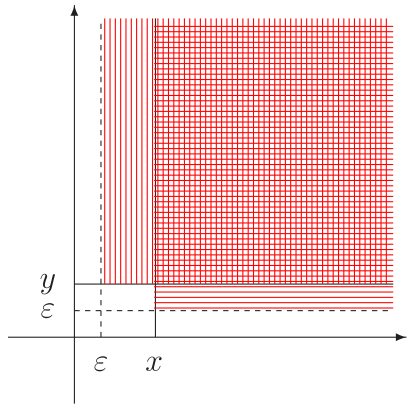

Figure 1: Illustration of the support of the bivariate tail integral $\overline{\Pi}^{(\varepsilon)}(x,y)$ and the marginal tail integrals $\overline{\Pi}_1^{(\varepsilon)}(x)$ and $\overline{\Pi}_2^{(\varepsilon)}(x)$.

4 Maximum likelihood estimation of the parameters of a bivariate Lévy processes

4.1 Small jump truncation

Let $\mathbf{S}$ be a bivariate Lévy process with unbounded Lévy measure Π in both components and marginal Lévy measures Π_1 and Π_2 corresponding to the components S_1 and S_2, respectively. It has an infinite number of jumps in the observation interval $[0,t]$. Several observation schemes are possible here concerning the truncation of the small jumps.

We consider only jumps (x,y), where both $x \geq \varepsilon$ and $y \geq \varepsilon$ at the same time. This leads to a bivariate compound Poisson model with joint jumps larger than ε.

Consider the truncated process $\mathbf{S}^{(\varepsilon)}$ with total Lévy measure

$$\Pi^{(\varepsilon)}(\mathbb{R}_+^2) = \Pi\{(x,y) \in \mathbb{R}_+^2 : x \geq \varepsilon, y \geq \varepsilon\} =: \lambda^{(\varepsilon)} < \infty \,.$$

Then there exists a representation

$$\mathbf{S}^{(\varepsilon)}(t) = \int_0^t \int_{\mathbf{x} \geq \varepsilon} \mathbf{x} M(ds \times d\mathbf{x}) = \sum_{i=1}^{N(t)} \mathbf{X}_i \,, \quad t \geq 0 \,,$$

where $\geq$ is taken componentwise and M is a Poisson random measure, which has support $[0,\infty) \times [\varepsilon,\infty)^2$ with intensity measure $ds\Pi^{(\varepsilon)}(d\mathbf{x})$ on its support; cf. Sato [16], Theorem 19.2. This means that $\mathbf{S}^{(\varepsilon)}$ is a compound Poisson process with intensity $\lambda^{(\varepsilon)}$ and leads to the observation scheme as described in Section 4 of Esmaeili and Klüppelberg [8] in detail, where now all jumps are larger than ε in both components. We now investigate the influence of the truncation on the Lévy copula.

Lemma 4.1. *Let* $\mathbf{S}$ *be a bivariate Lévy process with unbounded Lévy measure* Π *concentrated on* $\mathbb{R}_+^2$ *and Lévy copula* $\mathfrak{C}$*, which is different from the independent Lévy copula. Consider only those jumps, which are larger than* ε *in both component processes. Then the Lévy copula of the resulting CPP is given by*

$$\widetilde{\mathfrak{C}}^{(\varepsilon)}(u,v) = \mathfrak{C}(\mathfrak{C}_1^{\leftarrow}(u,\lambda_2^{(\varepsilon)}), \mathfrak{C}_2^{\leftarrow}(\lambda_1^{(\varepsilon)}, v)), \quad 0 < u, v \le \lambda^{(\varepsilon)}. \tag{4.1}$$

where $\mathfrak{C}_k^{\leftarrow}$, $k = 1, 2$ *is the inverse of* $\mathfrak{C}$ *with respect to the* k*-th argument,* $\lambda_k^{(\varepsilon)} = \overline{\Pi}_k(\varepsilon)$, $k = 1, 2$, *and* $\lambda^{(\varepsilon)} = \overline{\Pi}(\varepsilon, \varepsilon)$.

Proof. The marginal tail integrals of the CPP are given by

$$\begin{aligned} \overline{\Pi}_1^{(\varepsilon)}(x) &= \overline{\Pi}(x,\varepsilon) = \mathfrak{C}(\overline{\Pi}_1(x), \overline{\Pi}_2(\varepsilon)) = \mathfrak{C}(\overline{\Pi}_1(x), \lambda_2^{(\varepsilon)}), \quad x > \varepsilon, \\ \overline{\Pi}_2^{(\varepsilon)}(y) &= \overline{\Pi}(\varepsilon, y) = \mathfrak{C}(\overline{\Pi}_1(\varepsilon), \overline{\Pi}_2(y)) = \mathfrak{C}(\lambda_1^{(\varepsilon)}, \overline{\Pi}_2(y)), \quad y > \varepsilon, \end{aligned} \tag{4.2}$$

whereas the bivariate tail integral is

$$\overline{\Pi}^{(\varepsilon)}(x,y) = \overline{\Pi}(x,y) = \mathfrak{C}(\overline{\Pi}_1(x), \overline{\Pi}_2(y)), \quad x, y > \varepsilon. \tag{4.3}$$

Denote by $\widetilde{\mathfrak{C}}^{(\varepsilon)}$ the Lévy copula of the CPP, and from (4.3) we have

$$\widetilde{\mathfrak{C}}^{(\varepsilon)}(\overline{\Pi}_1^{(\varepsilon)}(x), \overline{\Pi}_2^{(\varepsilon)}(y)) = \mathfrak{C}(\overline{\Pi}_1(x), \overline{\Pi}_2(y)), \quad x, y > \varepsilon.$$

Together with (4.2) this implies that

$$\widetilde{\mathfrak{C}}^{(\varepsilon)}(\mathfrak{C}(\overline{\Pi}_1(x), \lambda_2^{(\varepsilon)}), \mathfrak{C}(\lambda_1^{(\varepsilon)}, \overline{\Pi}_2(y))) = \mathfrak{C}(\overline{\Pi}_1(x), \overline{\Pi}_2(y)), \quad x, y > \varepsilon.$$

Setting $u := \mathfrak{C}(\overline{\Pi}_1(x), \lambda_2^{(\varepsilon)})$ and $v := \mathfrak{C}(\lambda_1^{(\varepsilon)}, \overline{\Pi}_2(y))$, we see that for $x, y > \varepsilon$

$$\overline{\Pi}_1(x) = \mathfrak{C}_1^{\leftarrow}(u, \lambda_2^{(\varepsilon)}) \quad \text{and} \quad \overline{\Pi}_2(y) = \mathfrak{C}_2^{\leftarrow}(\lambda_1^{(\varepsilon)}, v),$$

and, hence, for $0 < u, v \le \lambda^{(\varepsilon)}$

$$\widetilde{\mathfrak{C}}^{(\varepsilon)}(u,v) = \mathfrak{C}\left(\mathfrak{C}_1^{\leftarrow}(u, \lambda_2^{(\varepsilon)}), \mathfrak{C}_2^{\leftarrow}(\lambda_1^{(\varepsilon)}, v)\right).$$

□

Proposition 4.2. *Assume that the conditions of Lemma 4.1 hold. Then*

$$\lim_{\varepsilon \to 0} \widetilde{\mathfrak{C}}^{(\varepsilon)}(u,v) = \mathfrak{C}(u,v), \quad u, v > 0.$$

Proof. Take arbitrary $u, v > 0$. Then there exists some $\varepsilon > 0$ such that $0 < u, v \leq \lambda^{(\varepsilon)}$. Invoking the Lipschitz condition for Lévy copula (Theorem 2.1, Barndorff-Nielsen and Lindner [2])and (4.1), we have

$$\begin{aligned}|\widetilde{\mathfrak{C}}^{(\varepsilon)}(u,v) - \mathfrak{C}(u,v)| &= \left|\mathfrak{C}\left(\mathfrak{C}_1^{\leftarrow}(u,\lambda_2^{(\varepsilon)}), \mathfrak{C}_2^{\leftarrow}(\lambda_1^{(\varepsilon)}, v)\right) - \mathfrak{C}(u,v)\right| \\ &\leq \left|\mathfrak{C}_1^{\leftarrow}(u,\lambda_2^{(\varepsilon)}) - u\right| + \left|\mathfrak{C}_2^{\leftarrow}(\lambda_1^{(\varepsilon)}, v) - v\right|.\end{aligned}$$

Since the Lévy copula $\mathfrak{C}$ has Lebesgue margins, i.e. $\mathfrak{C}(u,\infty) = u$ and $\mathfrak{C}(\infty, v) = v$, we have $\mathfrak{C}_1^{\leftarrow}(u,\infty) = u$ and $\mathfrak{C}_2^{\leftarrow}(\infty, v) = v$. This implies that

$$|\widetilde{\mathfrak{C}}^{(\varepsilon)}(u,v) - \mathfrak{C}(u,v)| \leq \left|\mathfrak{C}_1^{\leftarrow}(u,\lambda_2^{(\varepsilon)}) - \mathfrak{C}_1^{\leftarrow}(u,\infty)\right| + \left|\mathfrak{C}_2^{\leftarrow}(\lambda_1^{(\varepsilon)}, v) - \mathfrak{C}_2^{\leftarrow}(\infty, v)\right|.$$

The terms on the rhs tend to zero because the Lévy measure is unbounded and $\lim_{\varepsilon\to 0}\lambda_1^{(\varepsilon)} = \lim_{\varepsilon\to 0}\lambda_2^{(\varepsilon)} = \infty$. □

Now we proceed as in Esmaeili and Klüppelberg [8] and use the same notation. Denote by $(x_1, y_1), \ldots, (x_{n^{(\varepsilon)}}, y_{n^{(\varepsilon)}})$ the observed jumps larger than ε in both components, i.e. occurring at the same time during the observation interval $[0, t]$. Assume further that the dependence structure of the process $\mathbf{S} = (S_1, S_2)$ is defined by a Lévy copula $\mathfrak{C}$ with a parameter vector δ. We also assume that γ_1 and γ_2 are the parameter vectors of the marginal Lévy measures Π_1 and Π_2.

Using the notation $\nu_k(\cdot) = \lambda_k^{(\varepsilon)} f_k^{(\varepsilon)}(\cdot)$ for the marginal Lévy densities on (ε, ∞) for $k = 1, 2$ we can reformulate Theorem 4.1 of Esmaeili and Klüppelberg [8] as follows.

Theorem 4.3. *Assume an observation scheme as above for a bivariate Lévy process with only non-negative jumps. Assume that γ_1 and γ_2 are the parameters of the marginal Lévy measures Π_1 and Π_2 with Lévy densities ν_1 and ν_2, respectively, and a Lévy copula $\mathfrak{C}$ with parameter vector δ. Assume further that $\frac{\partial^2}{\partial u \partial v}\mathfrak{C}(u, v; \delta)$ exists for all $(u, v) \in (0, \infty)^2$, which is the domain of $\mathfrak{C}$. Then the full likelihood of the bivariate CPP is given by*

$$L^{(\varepsilon)}(\gamma_1, \gamma_2, \delta) = e^{-\lambda^{(\varepsilon)} t} \prod_{i=1}^{n^{(\varepsilon)}} \left[\nu_1(x_i; \gamma_1)\nu_2(y_i; \gamma_2) \frac{\partial^2}{\partial u \partial v}\mathfrak{C}(u, v; \delta)\Big|_{\substack{u=\overline{\Pi}_1(x_i;\gamma_1),\\ v=\overline{\Pi}_2(y_i;\gamma_2)}} \right] \tag{4.4}$$

where

$$\lambda^{(\varepsilon)} = \int_\varepsilon^\infty \int_\varepsilon^\infty \Pi(dx, dy) = \mathfrak{C}(\overline{\Pi}_1(\varepsilon; \gamma_1), \overline{\Pi}_2(\varepsilon; \gamma_2); \delta).$$

4.2 Asymptotic behaviour of the MLEs of a bivariate stable Clayton model

A spectrally positive Lévy process is an α-stable subordinator if and only if $0 < \alpha < 1$ and there exists a finite measure $\tilde{\rho}$ on the unit sphere $\mathcal{S}^{d-1} := \{x \in \mathbb{R}_+^d \mid \|x\| = 1\}$ in $\mathbb{R}_+^d$

(for an arbitrary norm $\|\cdot\|$) such that the Lévy measure

$$\Pi(B) = \int_{\mathcal{S}^{d-1}} \tilde{\rho}(d\xi) \int_0^\infty 1_B(r\xi) \frac{dr}{r^{1+\alpha}}, \quad B \in \mathcal{B}(\mathbb{R}_+^d),$$

(cf. Theorem 14.3(ii) and Example 21.7 in Sato [16]).

From Kallsen and Tankov [12], Theorem 4.6, it is known that a bivariate process is α-stable if and only if it has α-stable marginal processes and a homogeneous Lévy copula of order 1; i.e. $\mathfrak{C}(tu, tv) = t\, \mathfrak{C}(u, v)$. The *Clayton Lévy copula*

$$\mathfrak{C}(u, v) = \left(u^{-\delta} + v^{-\delta}\right)^{-1/\delta}, \quad u, v > 0,$$

is homogeneous of order 1. Hence it is a valid model to define a bivariate α-stable process.

Suppose S_1 and S_2 are two α-stable subordinators with same tail integrals

$$\overline{\Pi}_k(x) = cx^{-\alpha}, \quad x > 0, \text{ for } k = 1, 2.$$

Assume further that $\mathbf{S} = (S_1, S_2)$ is a bivariate α-stable process with dependence structure modeled by a Clayton Lévy copula. The joint tail integral is then given by

$$\overline{\Pi}(x, y) = \mathfrak{C}(\overline{\Pi}_1(x), \overline{\Pi}_2(y)) = c\left(x^{\alpha\delta} + y^{\alpha\delta}\right)^{-\frac{1}{\delta}}, \quad x, y > 0. \tag{4.5}$$

The bivariate Lévy density is given by

$$\nu(x, y) = c(1+\delta)\alpha^2 (xy)^{\alpha\delta - 1}\left(x^{\alpha\delta} + y^{\alpha\delta}\right)^{-\frac{1}{\delta} - 2}, \quad x, y > 0, \tag{4.6}$$

We assume the observation scheme as in Section 4.1. The Lévy measure Π will be considered on the set $[\varepsilon, \infty) \times [\varepsilon, \infty)$ with jump intensity

$$\lambda^{(\varepsilon)} = \overline{\Pi}(\varepsilon, \varepsilon) = c\left(\varepsilon^{\alpha\delta} + \varepsilon^{\alpha\delta}\right)^{-\frac{1}{\delta}} = c2^{-1/\delta}\varepsilon^{-\alpha}. \tag{4.7}$$

and marginal tail integrals

$$\overline{\Pi}_k^{(\varepsilon)}(x) = c(x^{\alpha\delta} + \varepsilon^{\alpha\delta})^{-1/\delta}, \quad k = 1, 2. \tag{4.8}$$

Moreover, for $k = 1, 2$,

$$\overline{\Pi}_k^{(\varepsilon)}(\varepsilon) = c2^{-1/\delta}\varepsilon^{-\alpha} = \lambda^{(\varepsilon)},$$

and

$$\begin{aligned} \overline{G}_k^{(\varepsilon)}(x) &= P(X > x) = P(Y > x) = \frac{\overline{\Pi}_k^{(\varepsilon)}(x)}{\lambda^{(\varepsilon)}} \\ &= \left[\frac{1}{2}\left(1 + \left(\frac{x}{\varepsilon}\right)^{\alpha\delta}\right)\right]^{-1/\delta}, \quad x > \varepsilon. \end{aligned} \tag{4.9}$$

The Lévy copula of the CPP is by Lemma 4.1 given by

$$\begin{aligned}\widetilde{\mathfrak{C}}^{(\varepsilon)}(u,v) &= \mathfrak{C}\left(\mathfrak{C}_1^{\leftarrow}(u,\lambda_2^{(\varepsilon)}),\mathfrak{C}_2^{\leftarrow}(\lambda_1^{(\varepsilon)},v)\right)\\ &= \mathfrak{C}\left(\left(u^{-\delta}-\lambda_2^{(\varepsilon)-\delta}\right)^{-1/\delta},\left(v^{-\delta}-\lambda_1^{(\varepsilon)-\delta}\right)^{-1/\delta}\right)\\ &= \left(u^{-\delta}+v^{-\delta}-2c^{-\delta}\varepsilon^{\alpha\delta}\right)^{-1/\delta}\end{aligned}$$

From the Lévy density in (4.6) and the intensity in (4.7) the joint probability density of the bivariate jumps is given by

$$g^{(\varepsilon)}(x,y)=\alpha^2(1+\delta)\varepsilon^{\alpha}2^{\frac{1}{\delta}}(xy)^{\alpha\delta-1}(x^{\alpha\delta}+y^{\alpha\delta})^{-\frac{1}{\delta}-2},\quad x,y>\varepsilon\,. \tag{4.10}$$

We note that our model is a bivariate generalized Pareto distribution (GPD); cf. Model I of Section 5.4 in Arnold et al. [1]. They present some properties of the model, and in our case X,Y are positively correlated.

We now turn to the MLE procedure. Noting that the parameterisation (c,α,δ) creates various problems taking derivatives, we propose a different choice of parameters. First we set $\alpha\delta=\theta$. Furthermore, recalling from the one-dimensional case that $\lambda^{(\varepsilon)}$ is a more natural choice than c, we decided to use the parameters $(\lambda^{(\varepsilon)},\alpha,\theta)$. Recall from (3.2) for the bivariate CPP based on observations $(x_i,y_i)>\varepsilon$ for $i=1,\dots,n^{(\varepsilon)}$,

$$\begin{aligned}L^{(\varepsilon)}(\lambda^{(\varepsilon)},\alpha,\theta) &= e^{-\lambda^{(\varepsilon)}t}\prod_{i=1}^{n^{(\varepsilon)}}\nu(x_i,y_i)\\ &= e^{-\lambda^{(\varepsilon)}t}(\lambda^{(\varepsilon)})^{n^{(\varepsilon)}}(\alpha(\alpha+\theta))^{n^{(\varepsilon)}}\varepsilon^{\alpha n^{(\varepsilon)}}2^{\frac{n^{(\varepsilon)}\alpha}{\theta}}\prod_{i=1}^{n^{(\varepsilon)}}\left[(x_iy_i)^{\theta-1}\left(x_i^{\theta}+y_i^{\theta}\right)^{-\frac{\alpha}{\theta}-2}\right].\end{aligned}$$

Then the log-likelihood is given by

$$\begin{aligned}\ell^{(\varepsilon)}(\lambda^{(\varepsilon)},\alpha,\theta) &= -\lambda^{(\varepsilon)}t+n^{(\varepsilon)}\log\lambda^{(\varepsilon)}+n^{(\varepsilon)}(\log\alpha+\log(\alpha+\theta))+\alpha n^{(\varepsilon)}\log\varepsilon+n^{(\varepsilon)}\frac{\alpha}{\theta}\log 2\\ &\quad+(\theta-1)\sum_{i=1}^{n^{(\varepsilon)}}(\log x_i+\log y_i)-(2+\frac{\alpha}{\theta})\sum_{i=1}^{n^{(\varepsilon)}}\log(x_i^{\theta}+y_i^{\theta})\,.\end{aligned}$$

Note that the last term prevents the model to belong to an exponential family, so we have to be very careful concerning exchanging differentiation and integration. For the score functions we obtain

$$\begin{aligned}
\frac{\partial \ell^{(\varepsilon)}}{\partial \lambda^{(\varepsilon)}} &= -t + \frac{n^{(\varepsilon)}}{\lambda^{(\varepsilon)}} \\
\frac{\partial \ell^{(\varepsilon)}}{\partial \alpha} &= \frac{n^{(\varepsilon)}}{\alpha} + \frac{n^{(\varepsilon)}}{\alpha+\theta} + n^{(\varepsilon)} \log \varepsilon + \frac{n^{(\varepsilon)} \log 2}{\theta} - \frac{1}{\theta} \sum_{i=1}^{n^{(\varepsilon)}} \log(x_i^\theta + y_i^\theta) \\
\frac{\partial \ell^{(\varepsilon)}}{\partial \theta} &= \frac{n^{(\varepsilon)}}{\alpha+\theta} - \frac{n^{(\varepsilon)}\alpha}{\theta^2} \log 2 + \sum_{i=1}^{n^{(\varepsilon)}} (\log x_i + \log y_i) + \frac{\alpha}{\theta^2} \sum_{i=1}^{n^{(\varepsilon)}} \log(x_i^\theta + y_i^\theta) \\
&\quad -(2 + \frac{\alpha}{\theta}) \sum_{i=1}^{n^{(\varepsilon)}} \frac{\partial}{\partial \theta} \log(x_i^\theta + y_i^\theta) \,.
\end{aligned}$$

From this we obtain the MLE $\widehat{\lambda}^{(\varepsilon)} = \frac{n^{(\varepsilon)}}{t}$, whose asymptotic properties are well-known, and note that $\widehat{\lambda}^{(\varepsilon)}$ is independent of $\widehat{\alpha}$ and $\widehat{\theta}$. So we concentrate on $\widehat{\alpha}$ and $\widehat{\theta}$.

Note first that, as a consequence of (4.9), the d.f. of $X^* = \frac{X}{\varepsilon}$ is given by

$$P(X^* > x) \;=\; P(X > \varepsilon x) \,=\, 2^{\alpha/\theta} (x^\theta + 1)^{-\alpha/\theta} \quad \text{for} \quad x > 1 \,.$$

Since also the distributions of $(X^*, Y^*) = (\frac{X}{\varepsilon}, \frac{Y}{\varepsilon})$ is independent of ε, the following quantities are independent of ε.

Lemma 4.4. *The following moments are finite.*

$$\begin{aligned}
\mathbb{E}\left[\log\left(\frac{X}{\varepsilon}\right)\right] &= 2^{\frac{\alpha}{\theta}} \int_1^\infty \frac{(1+y^\theta)^{-\frac{\alpha}{\theta}}}{y} dy \\
\mathbb{E}\left[\log\left(\frac{1}{2}\left(\left(\frac{X}{\varepsilon}\right)^\theta + \left(\frac{Y}{\varepsilon}\right)^\theta\right)\right)\right] &= \frac{\theta}{\alpha} + \frac{\theta}{\alpha+\theta} \\
\mathbb{E}\left[\frac{\partial}{\partial\theta} \log\left(\left(\frac{X}{\varepsilon}\right)^\theta + \left(\frac{Y}{\varepsilon}\right)^\theta\right)\right] &= (2 + \frac{\alpha}{\theta}) \log \varepsilon + \frac{2}{\theta} + \mathbb{E}\left[\log\left(\frac{X}{\varepsilon}\right) + \log\left(\frac{Y}{\varepsilon}\right)\right] \\
&= (\frac{2\theta}{2\theta+\alpha}) \left(\frac{1}{\theta} + 2^{\frac{\alpha}{\theta}} \int_1^\infty \frac{(y^\theta+1)^{-\frac{\alpha}{\theta}}}{y} dy\right) .
\end{aligned}$$

Proof. The first equality is a consequence of the joint density (4.10) and marginal tail distribution (4.9) with some standard analysis.

The second equality is calculated from the score function for α and (4.11).

For the last identity we calculate

$$\begin{aligned}
& (2+\frac{\alpha}{\theta})\mathbb{E}\left[\frac{\partial}{\partial\theta}\log(X^\theta+Y^\theta)\right] \\
= & \frac{1}{\alpha+\theta}-\frac{\alpha}{\theta^2}\log 2+\mathbb{E}\left[\log X+\log Y\right]+\frac{\alpha}{\theta^2}\mathbb{E}\left[\log(X^\theta+Y^\theta)\right] \\
= & \frac{1}{\alpha+\theta}-\frac{\alpha}{\theta^2}\log 2+\mathbb{E}\left[\log X+\log Y\right]+\frac{\alpha}{\theta^2}(\theta\log\varepsilon+\frac{\theta}{\alpha}+\frac{\theta}{\alpha+\theta}+\log 2) \\
= & \frac{1}{\alpha+\theta}(1+\frac{\alpha}{\theta})+\mathbb{E}\left[\log X+\log Y\right]+\frac{1}{\theta}+\frac{\alpha}{\theta}\log\varepsilon \\
= & \mathbb{E}\left[\log X+\log Y\right]+\frac{2}{\theta}+\frac{\alpha}{\theta}\log\varepsilon \\
= & (2+\frac{\alpha}{\theta})\log\varepsilon+\frac{2}{\theta}+2^{\frac{\alpha}{\theta}+1}\int_1^\infty\frac{(y^\theta+1)^{-\frac{\alpha}{\theta}}}{y}dy\,.
\end{aligned}$$

□

The following is a first step for calculating the Fisher information matrix.

Lemma 4.5. *For all $\varepsilon>0$,*

$$E\Big[\frac{\partial\ell^{(\varepsilon)}}{\partial\alpha}\Big]=E\Big[\frac{\partial\ell^{(\varepsilon)}}{\partial\theta}\Big]=0. \tag{4.11}$$

Proof. We show the result for the partial derivative with respect to α, where we use a dominated convergence argument. Since derivatives are local objects, it suffices to show that for each $\alpha_0\in(0,1)$ there exist a $\xi>0$ such that for all α in a neighbourhood of α_0, given by $N_\xi(\alpha_0):=\{\alpha\in(0,1):0<\alpha_0-\xi\le\alpha\le\alpha_0+\xi<1\}$ there exists a dominating integrable function, independent of α. We obtain

$$\left|\frac{\partial\ell^{(\varepsilon)}}{\partial\alpha}\right| \;\le\; \frac{n^{(\varepsilon)}}{\alpha_0-\xi}+\frac{n^{(\varepsilon)}}{\alpha_0-\xi+\theta}+n^{(\varepsilon)}\log\varepsilon+\frac{n^{(\varepsilon)}\log 2}{\theta}+\frac{1}{\theta}\sum_{i=1}^{n^{(\varepsilon)}}\Big|\log(x_i^\theta+y_i^\theta)\Big|.$$

The right-hand side is integrable by Lemma 4.4, which can be seen by multiplying and dividing the x_i and y_i by ε and using the second identity of Lemma 4.4.
The proof for the partial derivative with respect to θ is similar, invoking Lemma 4.4. □

Next we calculate the second derivatives

$$\begin{aligned}
\frac{\partial^2 \ell^{(\varepsilon)}}{\partial \alpha^2} &= n^{(\varepsilon)}\left(-\frac{1}{\alpha^2}-\frac{1}{(\alpha+\theta)^2}\right)\\
\frac{\partial^2 \ell^{(\varepsilon)}}{\partial \alpha \partial \theta} &= n^{(\varepsilon)}\left(-\frac{1}{(\alpha+\theta)^2}-\frac{1}{\theta^2}\log 2\right)-\frac{1}{\theta}\sum_{i=1}^{n^{(\varepsilon)}}\frac{\partial}{\partial\theta}\log(x_i^\theta+y_i^\theta)+\frac{1}{\theta^2}\sum_{i=1}^{n^{(\varepsilon)}}\log(x_i^\theta+y_i^\theta)\\
\frac{\partial^2 \ell^{(\varepsilon)}}{\partial \theta \partial \alpha} &= \frac{\partial^2 \ell^{(\varepsilon)}}{\partial \alpha \partial \theta}\\
\frac{\partial^2 \ell^{(\varepsilon)}}{\partial \theta^2} &= -n^{(\varepsilon)}\left(\frac{1}{(\alpha+\theta)^2}-\frac{2\alpha}{\theta^3}\log 2\right)-\frac{2\alpha}{\theta^3}\sum_{i=1}^{n^{(\varepsilon)}}\log(x_i^\theta+y_i^\theta)+\frac{2\alpha}{\theta^2}\sum_{i=1}^{n^{(\varepsilon)}}\frac{\partial}{\partial\theta}\log(x_i^\theta+y_i^\theta)\\
&\quad -(2+\frac{\alpha}{\theta})\sum_{i=1}^{n^{(\varepsilon)}}\frac{\partial^2}{\partial\theta^2}\log(x_i^\theta+y_i^\theta).
\end{aligned}$$

In order to calculate the Fisher information matrix we invoke Lemma 4.4. The components of the Fisher information matrix are then given by

$$\begin{aligned}
\widetilde{i}_{11} &= \mathbb{E}\left[-\frac{\partial^2}{\partial\alpha^2}\ell^{(\varepsilon)}\right]=\lambda^{(\varepsilon)}t\left[\frac{1}{\alpha^2}+\frac{1}{(\alpha+\theta)^2}\right]\\
&=: \lambda^{(\varepsilon)} i_{11}\, t\\
\widetilde{i}_{12} &= \widetilde{i}_{21}=\mathbb{E}\left[-\frac{\partial^2}{\partial\alpha\partial\theta}\ell^{(\varepsilon)}\right]\\
&= \lambda^{(\varepsilon)}t\left[\frac{1}{(\alpha+\theta)^2}-\frac{1}{\alpha\theta}-\frac{1}{\theta(\alpha+\theta)}+\frac{1}{\theta}\mathbb{E}\left[\frac{\partial}{\partial\theta}\log\left(\left(\frac{X}{\varepsilon}\right)^\theta+\left(\frac{Y}{\varepsilon}\right)^\theta\right)\right]\right]\\
&= \lambda^{(\varepsilon)}t\left[\frac{1}{(\alpha+\theta)^2}+\frac{2}{\theta(2\theta+\alpha)}-\frac{1}{\alpha\theta}-\frac{1}{\theta(\alpha+\theta)}+\frac{2^{\alpha/\theta+1}}{2\theta+\alpha}\int_1^\infty\frac{(1+u^\theta)^{-\frac{\alpha}{\theta}}}{u}du\right]\\
&=: \lambda^{(\varepsilon)} i_{12}\, t=\lambda_{21}^{(\varepsilon)} i_{12}\, t
\end{aligned}$$

$$\begin{aligned}
\widetilde{i}_{22} &= \mathbb{E}\left[-\frac{\partial^2}{\partial\theta^2}\ell^{(\varepsilon)}\right]=\lambda^{(\varepsilon)}t\left[\frac{1}{(\alpha+\theta)^2}-\frac{2\alpha\log 2}{\theta^3}+\frac{2\alpha}{\theta^3}\mathbb{E}\left(\log(X^\theta+Y^\theta)\right)\right.\\
&\quad\left.-\frac{2\alpha}{\theta^2}\mathbb{E}\left(\frac{\partial}{\partial\theta}\log(X^\theta+Y^\theta)\right)+(\frac{\alpha}{\theta}+2)\mathbb{E}\left(\frac{\partial^2}{\partial\theta^2}\log(X^\theta+Y^\theta)\right)\right]\\
&= \lambda^{(\varepsilon)}t\left[\frac{1}{(\alpha+\theta)^2}+\frac{2}{\theta^2}+\frac{2\alpha}{\theta^2(\alpha+\theta)}-\frac{4\alpha}{\theta^2(\alpha+2\theta)}-\frac{\alpha 2^{\alpha/\theta+2}}{\theta(2\theta+\alpha)}\int_1^\infty\frac{(u^\theta+1)^{-\alpha/\theta}}{u}du\right.\\
&\quad\left.+(\frac{\alpha}{\theta}+2)g(\alpha,\theta)\right]\\
&=: \lambda^{(\varepsilon)} i_{22}\, t\,,
\end{aligned}$$

where

$$g(\alpha,\theta):=\mathbb{E}\left[\frac{\partial^2}{\partial\theta^2}\log(X^\theta+Y^\theta)\right]=\mathbb{E}\left[\frac{\partial^2}{\partial\theta^2}\log\left(\left(\frac{X}{\varepsilon}\right)^\theta+\left(\frac{Y}{\varepsilon}\right)^\theta\right)\right]$$

does not depend on ε. This implies in particular that all i_{kl} are independent of ε. Consequently, the Fisher information matrix is given by

$$I^{(\varepsilon)}_{\alpha,\theta} = \lambda^{(\varepsilon)} t \begin{pmatrix} i_{11} & i_{12} \\ i_{12} & i_{22} \end{pmatrix}.$$

Recall the asymptotic normality of the estimated parameters in the one-dimensional case of Example 3.1. In our bivariate model we have additionally to those parameters the dependence parameter θ. This means that we have to check the regularity conditions (A1)-(A4) in Section 3.11 of Prakasa Rao [15] for this model. (A1) and (A2) are differentiability conditions, which are satisfied. As a prerequisite for (A3) and (A4) we need to show invertibility of the Fisher information matrix $I^{(\varepsilon)}_{\alpha,\theta}$, which we are not able to do analytically. A numerical study for a large number of values for α and θ, however, always gave a positive determinant, indicating that the inverse indeed exists. Since the Fisher information matrix depends on t only by the common factor, it is not difficult to convince ourselves that also (A3) and (A4) are satisfied. Hence, classical likelihood theory applies and ensures that

$$\sqrt{n^{(\varepsilon)}} \begin{pmatrix} \widehat{\alpha} - \alpha \\ \widehat{\theta} - \theta \end{pmatrix} \sim \text{AN}\left(\mathbf{0}, \begin{pmatrix} i_{11} & i_{12} \\ i_{12} & i_{22} \end{pmatrix}^{-1}\right), \quad n^{(\varepsilon)} \to \infty .$$

As in the one-dimensional case, we use the consistency result in (3.3) and Slutzky's theorem, which gives for $n^{(\varepsilon)} \to \infty$, equivalently, $\varepsilon \to 0$,

$$\sqrt{c 2^{-\alpha/\theta} \varepsilon^{-\alpha} t} \begin{pmatrix} \widehat{\alpha} - \alpha \\ \widehat{\theta} - \theta \end{pmatrix} \stackrel{d}{\to} N\left(\mathbf{0}, \begin{pmatrix} i_{11} & i_{12} \\ i_{12} & i_{22} \end{pmatrix}^{-1}\right).$$

In reality the parameters are estimated from the data and plugged into the rate and the i_{kl}. Moreover, the unknown expectations in the Fisher information matrix have to be either numerically calculated by the corresponding integrals or estimated by Monte Carlo simulation. In Section 5 we shall perform a simulation study and also present an example of the covariance matrix for some specific choice of parameters.

Before this we want to come back to our change of parameters and, in particular, want to discuss estimation of the parameter c of the stable margins. From (3.3) and the fact that $\widehat{\lambda}^{(\varepsilon)}$, $\widehat{\alpha}$ and $\widehat{\theta}$ are consistent, we know that for $n^{(\varepsilon)} \to \infty$,

$$\widehat{c} := \widehat{\lambda}^{(\varepsilon)} 2^{\widehat{\alpha}/\widehat{\theta}} \varepsilon^{\widehat{\alpha}}$$

is a consistent estimator of c.

We calculate as follows

$$\begin{aligned} \widehat{\log c} &= \log \widehat{\lambda}^{(\varepsilon)} + \frac{\widehat{\alpha}}{\widehat{\theta}} \log 2 + \widehat{\alpha} \log \varepsilon \\ &= \log \frac{\widehat{\lambda}^{(\varepsilon)}}{\lambda^{(\varepsilon)}} + \log c - \frac{\alpha}{\theta} \log 2 - \alpha \log \varepsilon + \frac{\widehat{\alpha}}{\widehat{\theta}} \log 2 + \widehat{\alpha} \log \varepsilon . \end{aligned}$$

Consistency of $\widehat{\lambda}^{(\varepsilon)}$ implies that

$$\begin{aligned}
\widehat{\log c} - \log c &= o_P(1) + (\widehat{\alpha} - \alpha)\log\varepsilon + \Big(\frac{\widehat{\alpha}}{\widehat{\theta}} - \frac{\alpha}{\theta}\Big)\log 2 \\
&= o_P(1) + (\widehat{\alpha} - \alpha)\log\varepsilon + \Big(\frac{1}{\widehat{\theta}} - \frac{1}{\theta}\Big)\alpha\log 2 + (\widehat{\alpha} - \alpha)\frac{\log 2}{\widehat{\theta}} \\
&= o_P(1) + (\widehat{\alpha} - \alpha)\log\varepsilon + \Big(\frac{\theta}{\widehat{\theta}} - 1\Big)\frac{\alpha}{\theta}\log 2 \\
&\quad + (\widehat{\alpha} - \alpha)\frac{\log 2}{\theta}(1 + o_P(1))\,,
\end{aligned}$$

where we have used the consistency of $\widehat{\alpha}$ and $\widehat{\theta}$. This implies for $\varepsilon \to 0$,

$$\frac{\widehat{\log c} - \log c}{\log\varepsilon} = (\widehat{\alpha} - \alpha)(1 + o_P(1))\,.$$

Consequently, analogously to the one-dimensional case, we obtain the following result.

Theorem 4.6. *Let $(\widehat{c}, \widehat{\alpha}, \widehat{\theta})$ denote the MLEs of the bivariate α-stable Clayton subordinator. Then as $\varepsilon \to 0$,*

$$\sqrt{c 2^{-\alpha/\theta}\varepsilon^{-\alpha} t}\begin{pmatrix} \dfrac{\widehat{\log c} - \log c}{\log\varepsilon} \\ \widehat{\alpha} - \alpha \\ \widehat{\theta} - \theta \end{pmatrix} \xrightarrow{d} \begin{pmatrix} N_1 \\ N_1 \\ N_2 \end{pmatrix}, \quad \varepsilon \to 0\,,$$

where $\mathrm{Cov}(N_1, N_2) = \begin{pmatrix} i_{11} & i_{12} \\ i_{12} & i_{22} \end{pmatrix}^{-1}$ *is independent of ε.*

Obviously, we can do all again a Taylor expansion to obtain the limit law for $\widehat{c}$ instead of $\widehat{\log c}$ as in the one-dimensional case.

5 Simulation study for a bivariate α-stable Clayton subordinator

We start generating data from a bivariate α-stable Clayton subordinator over a time span $[0, t]$, where we choose $t = 1$ for simplicity. Recall that our observation scheme introduced in Section 4.1. assumes that from the α-stable Clayton subordinator we only observe bivariate jumps larger than ε. Obviously, we cannot simulate a trajectory of a stable process, since we are restricted to the simulation of a finite number of jumps. For simulation purpose we choose a threshold ξ (which should be much smaller than ε) and simulate jumps larger than ξ in one component, and arbitrary in the second component. To this end we invoke Algorithm 6.15 in Cont and Tankov [6].

ε		$\delta = 2$	$\alpha = 0.5$	$c = 1$
0.001	Mean	2.0861 (0.8245)	0.5323 (0.1233)	1.0642 (0.6848)
	$\sqrt{MSE}$	0.8290 (1.3074)	0.1275 (0.0340)	0.6878(0.9855)
	MRB	0.0476	0.0658	0.0232
0.0001	Mean	2.0180 (0.4333)	0.5110 (0.0637)	1.0531 (0.5174)
	$\sqrt{MSE}$	0.4337 (0.2831)	0.0647 (0.0078)	0.5201(0.5170)
	MRB	0.0108	0.0216	0.0423
0.00001	Mean	2.0029 (0.2364)	0.5041 (0.0348)	1.0270 (0.3713)
	$\sqrt{MSE}$	0.2364 (0.0781)	0.0350 (0.0021)	0.3722 (0.2730)
	MRB	0.0015	0.0081	0.0240

Table 5.1: Estimation of the bivariate $\frac{1}{2}$-stable Clayton process with jumps truncated at different ε: the mean of MLEs of the copula and the margins parameter δ, α and c with $\sqrt{MSE}$ and standard deviations (in brackets). This is based on a simulation of the process in a unit of time, $0 \le t < 1$, for $\tau = 1000$, equivalent to truncation of small jumps at the cut-off point $\xi = \overline{\Pi}^{\leftarrow}(\tau) = 10^{-6}$.

The simulation of a bivariate stable Clayton subordinator is explained in detail in Example 6.18 of [6]. The algorithm starts by fixing a number τ determined by the required precision. This number coincides with $\lambda_1^{(\xi)}$ and fixes the average number of terms in (5.1) below.

We generate an iid sequence of standard exponential random numbers $E_1, E_2, \ldots$. Then we set $\Gamma_0^{(1)} = 0$ and $\Gamma_i^{(1)} = \Gamma_{i-1}^{(1)} + E_i$ until $\Gamma_{n^{(\xi)}}^{(1)} \le \tau$ and $\Gamma_{n^{(\xi)}+1}^{(1)} > \tau$ resulting in the jump times of a standard Poisson process $\Gamma_0^{(1)}, \Gamma_1^{(1)}, \ldots, \Gamma_{n^{(\xi)}}^{(1)}$. Besides the marginal tail integrals we also need to know for every i the conditional distribution function given for $\Gamma_i^{(1)} = u > 0$ by

$$F_{2|1}(v \mid u) = \left(1 + (u/v)^{\delta}\right)^{-1/\delta - 1}, \quad v > 0\,.$$

We simulate $\Gamma_i^{(2)}$ from the d.f. $F_{2|1}(v \mid u = \Gamma_i^{(1)})$. Finally, we simulate a sequence $U_1, U_2, \ldots$ of iid uniform random numbers on $(0, 1)$. The trajectory of the bivariate Clayton subordinator has the following representation

$$\begin{pmatrix} S_1^{(\xi)}(t) \\ S_2^{(\xi)}(t) \end{pmatrix} = \begin{pmatrix} \sum_{i=1}^{n^{(\xi)}} 1_{\{U_i \le t\}} \overline{\Pi}_1^{\leftarrow}(\Gamma_i^{(1)}) \\ \sum_{i=1}^{n^{(\xi)}} 1_{\{U_i \le t\}} \overline{\Pi}_2^{\leftarrow}(\Gamma_i^{(2)}) \end{pmatrix}, \quad 0 < t < 1\,, \tag{5.1}$$

where $(\Gamma_i^{(1)}, \Gamma_i^{(2)})$ carry the dependence structure of the Lévy copula. Note that the jump times in both components always coincide.

Table 5.1 summarizes the results of a simulation study based on 100 trajectories of the bivariate α-stable Clayton subordinator with parameters $\alpha = 0.5$, $c = 1$ and Clayton dependence parameter $\delta = 2$.

Finally, we also want to give an idea about the theoretical properties of our MLE procedure. To this end we calculate the theoretical asymptotic covariance matrix for the same set of parameters $(c, \alpha, \theta) = (1, 0.5, 1)$. Note that in this case we can calculate the integral in the Fisher information matrix explicitly. The expectation of the second derivative we obtain from a Monte Carlo simulation based on simulated (X, Y).

We conclude this section with an example of the covariance matrix $\text{Cov}(N_1, N_2)$ of the normal limit vector of the parameter estimates as given in Theorem 4.6. We do this for the model with parameters $c = 1$, $\alpha = 0.5$ and $\theta = 1$ as used for the simulation with results summarized in Table 5.1. We present the matrix resulting from two different methods. The left hand matrix has been calculated by numerical integration, whereas the right hand matrix is the result of a Monte Carlo simulation based on 1000 observations from the bivariate Pareto distribution (4.10).

Numerical integration $\qquad$ Monte Carlo simulation

$$\begin{bmatrix} 0.2492 & -0.1885 \\ -0.1885 & 1.4686 \end{bmatrix} \qquad \begin{bmatrix} 0.2487 & -0.1867 \\ -0.1867 & 1.4700 \end{bmatrix}$$

6 Conclusion and outlook

For the specific bivariate α-stable Clayton subordinator with equal marginal Lévy processs we have estimated all parameters in one go and proved asymptotic normality for $n^{(\varepsilon)} \to \infty$. Observation scheme were joint jumps larger than ε in both components and a fixed observation interval $[0, t]$. This limit result holds for $t \to \infty$ or, equivalently, for $\varepsilon \to 0$.

Since this estimation procedure requires even for a bivariate model with the same marginal processes a non-trivial numerical procedure to estimate the parameters, it seems to be advisable to investigate also two-step procedures like IFM (inference functions for margins), which we do in Esmaeili and Klüppelberg [9]. In such a procedure the parameters of the marginals may well be different, and the model of arbitrary dimension, since marginal parameters are estimated first and then estimate in a second step only the dependence structure parameters. This well-known estimation procedure in the copula framework will be investigated in a follow-up paper.

Alternatively, one can apply non-parametric estimation procedures for Lévy measures as e.g. in Ueltzhöfer and Klüppelberg [18].

Acknowledgement

CK takes pleasure to thank Alexander Lindner for some support with the analysis at the Oberwolfach Workshop on "Challenges in Statistical Theory: Complex Data Structures

and Algorithmic Optimization".

References

[1] Arnold, B.C., Castillo, E. and Sarabia, J.M. (1999) *Conditional Specification of Statistical Models.* Springer, New York.

[2] Barndorff-Nielsen, O. and Lindner, A. (2007) Lévy copulas: dynamics and transforms of Upsilon type. *Scand. J. Stat.* **34**, No. 2, 298-316(19).

[3] Basawa, I.V. and Brockwell, P.J. (1978) Inference for gamma and stable processes. *Biometrika* **65**, No. 1, 129-133.

[4] Basawa, I.V. and Brockwell, P.J. (1980) A note on estimation for gamma and stable processes. *Biometrika* **67**, No. 1, 234-236.

[5] Basawa, I.V. and Prakasa Rao, B.L.S. (1980) *Statistical Inference for Stochastic Processes.* Academic Press, New York.

[6] Cont, R. and Tankov, P. (2004) *Financial Modelling with Jump Processes.* Chapman & Hall/CRC, Boca Raton.

[7] Eder, I. and Klüppelberg, C. (2009) Pareto Lévy measures and multivariate regular variation. Submitted for publication.

[8] Esmaeili, H. and Klüppelberg, C. (2010) Parameter estimation of a bivariate compound Poisson process. *Insurance: Mathematics and Economics* **47**, 224-233.

[9] Esmaeili, H. and Klüppelberg, C. (2010) Two-step estimation of a multivariate Lévy process. Submitted for publication.

[10] Höpfner, R. (1999) On statistical models for d-dimensional stable processes, and some generalisations. *Scand. J. Stat.* **26**, 611-620.

[11] Höpfner, R. and Jacod, J. (1994) Some remarks on the joint estimation of the index and the scale parameter for stable processes. In: Mandl, P. and Huskova, M. (Eds.) *Asymptotic Statistics. Proceedings of the Fifth Prague Symposium 1993*, pp. 273-284. Physica Verlag, Heidelberg.

[12] Kallsen, J. and Tankov, P. (2006) Characterization of dependence of multidimentional Lévy processes using Lévy copulas. *J. Mult. Anal.* **97**, 1551-1572.

[13] Küchler, U. and Sørensen, M. (1997) *Exponential Families of Stochastic Processes.* Springer, New York.

[14] Nolan, J.P., Panorska, A. and McCulloch, J.H. (2001) Estimation of stable spectral measures *Mathematical and Computer Modelling* **34**, 1113-1122.

[15] Prakasa Rao, B.L.S (1987) *Asymptotic Theory of Statistical Inference.* Wiley, New York.

[16] Sato, K.-I. (1999) *Lévy Processes and Infinitely Divisible Distributions.* Cambridge University Press, Cambridge.

[17] Tankov, P. (2003) Dependence structure of spectrally positive multidimentional Lévy processes. Available at {*http://www.camp.polytechnique.fr/*}

[18] Ueltzhöfer, F. and Klüppelberg, C. (2010) An oracle inequality for penalised projection estimation of Lévy densities from high-frequency observations. Submitted for publication.