Parameter estimation of a bivariate compound Poisson process

Habib Esmaeili * Claudia Klüppelberg †

August 25, 2010

Abstract

In this article, we review the concept of a Lévy copula to describe the dependence structure of a bivariate compound Poisson process. In this first statistical approach we consider a parametric model for the Lévy copula and estimate the parameters of the full dependent model based on a maximum likelihood approach. This approach ensures that the estimated model remains in the class of multivariate compound Poisson processes. A simulation study investigates the small sample behaviour of the MLEs, where we also suggest a new simulation algorithm. Finally, we apply our method to the Danish fire insurance data.

AMS 2000 Subject Classifications: primary: 62M05, 62P05
secondary: 60G51

Keywords: dependence modelling, Lévy copula, Lévy measure, Lévy process, maximum likelihood estimation, multivariate compound Poisson process.

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

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

1 Introduction

Copulas open a convenient way to represent the dependence of a probability distribution. In fact they provide a complete characterization of possible dependence structures of a random vector with fixed margins. Moreover, using copulas, one can construct multivariate distributions with a pre-specified dependence structure from a collection of univariate laws. Modern results about copulas originate more than forty years back when Sklar [16] defined and derived the fundamental properties of a copula. Further important references are Nelson [14] and Joe [12]. Financial applications of copulas have been numerous in recent years; cf. Cherubini, Luciano and Vecchiato [6] for examples and further references.

We are considering multivariate Lévy processes, whose dependence can be modelled by a "copula" on the components of the Lévy measure. This has been suggested in Tankov [17] for subordinators, the case of general Lévy processes was treated in Kallsen and Tankov [13]; Lévy copulas can also be found in the monograph of Cont and Tankov [7].

Modelling dependence in multivariate Lévy processes by Lévy copulas offers the same flexibility for modelling the marginal Lévy processes independently of their dependence structure as we know from distributional copulas. Statistical methods, which have existed for distributional copulas for a long time, still have to be developed for Lévy copulas. The present paper is a first step.

The Lévy copula concept has been applied to insurance risk problems; more precisely, Bregman and Klüppelberg [2] have used this approach for ruin estimation in multivariate models. Eder and Klüppelberg [8] extended this work to derive the so-called quintuple law for sums of dependent Lévy processes. This describes the ruin event by stating not only the ruin probability, but also quantities like ruin time, overshoot, undershoot; i.e. they present a ladder process analysis. The notion of multivariate regular variation can also be linked to Lévy copulas, which is investigated and presented in Eder and Klüppelberg [9].

In a series of papers, Böcker and Klüppelberg [3, 4, 5] used a multivariate compound Poisson process to model operational risk in different business lines and risk types. Again dependence is modelled by a Lévy copula. Analytic approximations for the operational Value-at-Risk explain the influence of dependence on the institution's total operational risk.

In view of these economic problems, which are well recognised in academia and among practitioners, the present paper is concerned with statistical inference for bivariate compound Poisson processes. Our method is based on Sklar's theorem for Lévy copulas, which guarantees that the estimated model is again multivariate compound Poisson.

This approach, whose importance is already manifested by the above mentioned economic applications as well as in a data analysis at the end of our paper, will have far reaching implications for the estimation of multivariate Lévy processes with infinite ac-

tivity sample paths as is relevant in finance. This has been worked out in Esmaeili and Klüppelberg [11].

Our paper is organized as follows. Section 2 presents the definition of a multivariate compound Poisson process (CPP) and explains the dependence structure in three possible ways. This prepares the ground for a new simulation algorithm for multivariate compound Poisson processes and for the maximum likelihood estimation. Then we define the concept of a tail integral and a Lévy copula for such processes in Section 3. In Section 4 we derive the likelihood function for the process parameters, where we assume that we observe the continuous-time sample path. In Section 5 we suggest a new simulation algorithm for multivariate compound Poisson processes and show it at work by simulating a bivariate CPP, whose dependence structure is modelled by a Clayton Lévy copula. Finally, in Section 6 we fit a compound Poisson process to the bivariate Danish fire insurance data, and present some conclusions in Section 7.

2 The multivariate compound Poisson process

A d-dimensional *compound Poisson process* (CPP) is a Lévy process $\mathbf{S} = (\mathbf{S}(t))_{t\geq 0}$, i.e. a process with independent and stationary increments, defined on a filtered probability space $(\Omega, \mathcal{F}, (\mathcal{F}_t)_{t\geq 0}, \mathcal{P})$, with values in $\mathbb{R}^d$. It is stochastically continuous, i.e. for all $a > 0$,

$$\lim_{t\to h} P(|\mathbf{S}(t) - \mathbf{S}(h)| > a) = 0\,, \quad h \geq 0\,,$$

and as is well-known (see e.g. Sato [15], Def. 1.6), a càdlàg version exists, and we assume this property throughout. For each $t > 0$ the characteristic function has the so-called Lévy-Khintchine representation:

$$E[e^{i(\mathbf{z},\mathbf{S}(t))}] \;=\; \exp\left\{t\int_{\mathbb{R}^d}\left(e^{i(\mathbf{z},\mathbf{x})} - 1\right)\Pi(d\mathbf{x})\right\}, \quad \mathbf{z}\in\mathbb{R}^d\,, \tag{2.1}$$

where $(\cdot,\cdot)$ denotes the inner product in $\mathbb{R}^d$. The so-called Lévy measure Π is a measure on $\mathbb{R}^d$ satisfying $\Pi(\{\mathbf{0}\}) = 0$ and $\int_{\mathbb{R}^d}\Pi(d\mathbf{x}) < \infty$. Moreover, CPPs are the only Lévy processes with finite Lévy measure.

According to Sato [15], Theorem 4.3, a compound Poisson process is a stochastic process

$$\mathbf{S}(t) = \sum_{i=1}^{N(t)} \mathbf{Z}_i\,, \quad t\geq 0\,, \tag{2.2}$$

where $(N(t))_{t\geq 0}$ is a homogeneous Poisson process with intensity $\lambda > 0$ and $(\mathbf{Z}_i)_{i\in\mathbb{N}}$ is a sequence of iid random variables with values in $\mathbb{R}^d$. Moreover $(N(t))_{t\geq 0}$ and $(\mathbf{Z}_i)_{i\in\mathbb{N}}$ are independent and the $\mathbf{Z}_i$'s have no atom in $\mathbf{0}$, i.e. $P(\mathbf{Z}_1 = \mathbf{0}) = 0$.

To prepare the ground for our statistical analysis based on a Lévy copula to come, we present a bivariate CPP in more detail. In particular, we give three approaches to understand the dependence structure of such a process in more detail.

Assume that for $i \in \mathbb{N}$ the bivariate vector $\boldsymbol{Z}_i$ has df G with components Z_{1i} and Z_{2i} with dfs G_1 and G_2, respectively. It is, of course, possible that single jumps in one of the marginal processes occur, in which case the probability measure of the marks Z_{1i} and Z_{2i} have atoms in 0; i.e. they are not continuous.

In our first approach we write

$$\boldsymbol{S}(t) = \sum_{i=1}^{N(t)} (Z_{1i}, Z_{2i}) = \Big(\sum_{i=1}^{N(t)} Z_{1i}, \sum_{i=1}^{N(t)} Z_{2i} \Big), \quad t \geq 0, \tag{2.3}$$

where we set $p_1 := P(Z_{1i} = 0)$ and $p_2 := P(Z_{2i} = 0)$ and recall that possibly $p_1, p_2 > 0$. Then for almost all $\omega \in \Omega$,

$$S_1(t) = \sum_{i=1}^{N_1(t)} X_i, \quad t \geq 0 \quad \text{and} \quad S_2(t) = \sum_{i=1}^{N_2(t)} Y_i, \quad t \geq 0, \tag{2.4}$$

where X_i and Y_i take only the non-zero values of Z_1 and Z_2, respectively, and inherit the independence of $N_1(\cdot) = (1-p_1)N(\cdot)$ and $N_2(\cdot) = (1-p_2)N(\cdot)$. To make this precise, for $i = 1, 2$ and a Borel set $A \subset \mathbb{R} \setminus \{0\}$ we can write

$$P\Big(\sum_{j=1}^{N(t)} Z_{ij} \in A \Big) \;=\; P\Big(\sum_{j=1}^{N^{i1}(t)} Z_{ij}\; 1_{\{Z_{ij}>0\}} + \sum_{j=1}^{N^{i2}(t)} Z_{ij}\; 1_{\{Z_{ij}=0\}} \in A \Big), \quad t \geq 0\,, \tag{2.5}$$

where $N^{1i}(\cdot)$ and $N^{2i}(\cdot)$ count the non-zero and zero jumps, respectively. By the thinning property of the Poisson process, they are again Poisson processes. Since the last summation in (2.5) is zero, we conclude that for almost all $\omega \in \Omega$

$$\sum_{j=1}^{N^{11}(t)} Z_{1j} = \sum_{j=1}^{N_1(t)} X_j, \quad t \geq 0, \quad \text{and} \quad \sum_{j=1}^{N^{21}(t)} Z_{2j} = \sum_{j=1}^{N_2(t)} Y_j\,, \quad t \geq 0,$$

are compound Poisson processes. Here $(N_1(t))_{t\geq 0}$, $(N_2(t))_{t\geq 0}$ are Poisson processes with intensities $\lambda_1 = (1-p_1)\lambda$ and $\lambda_2 = (1-p_2)\lambda$, respectively, and $(X_i)_{i\in\mathbb{N}}$ and $(Y_i)_{i\in\mathbb{N}}$ are sequences of iid random variables with dfs given for all $x \in \mathbb{R}$ by

$$F_1(x) = P(Z_1 \leq x \mid Z_1 \neq 0) \quad \text{and} \quad F_2(y) = P(Z_2 \leq y \mid Z_2 \neq 0)$$

Unlike G_1 and G_2, the dfs F_1 and F_2 have no mass in 0.

Our second approach is based on the representation of a compound Poisson process as an integral with respect to a Poisson random measure M; cf. Sato [15], Theorems 19.2

and 19.3. For almost all $\omega \in \Omega$ we have the representation

$$\begin{aligned}
\mathbf{S}(t) &= \sum_{i=1}^{N(t)} (Z_{1i}, Z_{2i}) \\
&= \int_0^t \int_{\mathbb{R}^2 \setminus \{\mathbf{0}\}} \mathbf{z} M(ds \times d\mathbf{z}) \qquad (2.6) \\
&= \int_0^t \int_{(\mathbb{R}\setminus\{0\})\times\{0\}} \mathbf{z} M(ds \times d\mathbf{z}) + \int_0^t \int_{\{0\}\times(\mathbb{R}\setminus\{0\})} \mathbf{z} M(ds \times d\mathbf{z}) + \int_0^t \int_{(\mathbb{R}\setminus\{0\})^2} \mathbf{z} M(ds \times d\mathbf{z})
\end{aligned}$$

where M is a Poisson random measure on $[0, \infty) \times (\mathbb{R}^2 \setminus \{\mathbf{0}\})$ with intensity measure $ds\, \Pi(d\mathbf{z})$.

Corresponding to the first two integrals we can introduce two compound Poisson processes $S_1^\perp$ and $S_2^\perp$, which are called the *independent parts of* (S_1, S_2). They are independent of each other and never jump together. On the other hand, the third integral corresponds to a compound Poisson process which is supported on sets in $(\mathbb{R} \setminus \{\mathbf{0}\})^2$, and this part of (S_1, S_2) measures the simultaneous jumps of S_1 and S_2. We denote it by $(S_1^\parallel, S_2^\parallel)$, and it is the *(jump) dependent part of* (S_1, S_2). Since its components $S_1^\parallel$ and $S_2^\parallel$ always jump together, they must have the same jump intensity parameter, which we denote by $\lambda^\parallel$. Now we can decompose $(S_1(t), S_2(t))_{t\geq 0}$ for almost all $\omega \in \Omega$ into

$$\begin{aligned}
S_1(t) &= S_1(t)^\perp + S_1(t)^\parallel\,, \quad t \geq 0\,, \qquad (2.7) \\
S_2(t) &= S_2(t)^\perp + S_2(t)^\parallel\,, \quad t \geq 0\,.
\end{aligned}$$

Here we see clearly the decomposition of the bivariate compound Poisson process in single jumps in each marginal process and the process of common jumps in both components. It is clear from the properties of the Poisson random measure that the three processes $S_1^\perp$, $S_2^\perp$ and $(S_1^\parallel, S_2^\parallel)$ are compound Poisson and independent.

Our last approach is similar to the previous one, but based on a decomposition of the Lévy measure. It also prepares the ground for the following Section 3. Recall that for any Borel set $A \subseteq \mathbb{R}^2 \setminus \{0\}$ its Lévy measure $\Pi(A)$ denotes the expected number of jumps per unit time with size in A. This can be formulated as

$$\Pi(A) = E\Big[\#\{(t, (\Delta S_1(t), \Delta S_2(t))) \in (0, 1] \times A\}\Big].$$

This set, and hence Π can be decomposed into the following components:

$$\begin{aligned}
\Pi_1(A) &= E\Big[\#\{(t, (\Delta S_1(t), \Delta S_2(t))) \in (0, 1] \times A \mid \Delta S_1(t) \neq 0 \text{ and } \Delta S_2(t) = 0\}\Big], \\
\Pi_2(A) &= E\Big[\#\{(t, (\Delta S_1(t), \Delta S_2(t))) \in (0, 1] \times A \mid \Delta S_1(t) = 0 \text{ and } \Delta S_2(t) \neq 0\}\Big], \\
\Pi_3(A) &= E\Big[\#\{(t, (\Delta S_1(t), \Delta S_2(t))) \in (0, 1] \times A \mid \Delta S_1(t) \neq 0 \text{ and } \Delta S_2(t) \neq 0\}\Big].
\end{aligned}$$

Since $\Pi(A) = \Pi_1(A) + \Pi_2(A) + \Pi_3(A)$, the integral of the characteristic function in (2.1) can be decomposed into three integrals with different Lévy measures $\Pi_1(A)$, $\Pi_2(A)$ and $\Pi_3(A)$, respectively. Clearly Π_1 is supported by the set $\{(x,0) \in \mathbb{R}^2 \mid x \in \mathbb{R}\}$. We set $\Pi_1(A) = \Pi_1^{\perp}(A_1)$, where $A_1 = \{x \in \mathbb{R} \mid (x,0) \in A\}$. Then the first integral reduces to a one-dimensional integral related only to the component S_1. Similarly, for the second part $\Pi_2(A) = \Pi_2^{\perp}(A_2)$, where $A_2 = \{y \in \mathbb{R} \mid (0,y) \in A\}$; hence the second integral also reduces to a one-dimensional integral. By introducing the notation $\Pi^{\|}$ for Π_3, the characteristic function in (2.1) can be decomposed into

$$\begin{aligned} & E\big[e^{iz_1S_1(t)+iz_2S_2(t)}\big] \\ &= \exp\left\{t\int_{\mathbb{R}}(e^{iz_1x}-1)\Pi_1^{\perp}(dx) + t\int_{\mathbb{R}}(e^{iz_2y}-1)\Pi_2^{\perp}(dy) + t\int_{\mathbb{R}^2}(e^{iz_1x+iz_2y}-1)\Pi^{\|}(dx\times dy)\right\} \\ &= E\big[e^{iz_1S_1(t)^{\perp}}\big]E\big[e^{iz_2S_2(t)^{\perp}}\big]E\big[e^{iz_1S_1(t)^{\|}+iz_2S_2(t)^{\|}}\big]. \end{aligned} \tag{2.8}$$

Note that the Lévy measure $\Pi_1^{\perp}$ gives the mean number of jumps of S_1 such that S_2 does not have a jump at the same time. Similarly, the mean number of jumps for S_2, when S_1 has no jump, is measured by $\Pi_2^{\perp}$. Corresponding to $\Pi_1^{\perp}$ and $\Pi_2^{\perp}$ we find again the two processes $S_1^{\perp}$ and $S_2^{\perp}$ which we called the *independent parts of* (S_1, S_2). On the other hand, $\Pi^{\|}$ is supported by sets in $(R \setminus \{0\})^2$, and we denoted this part by $(S_1^{\|}, S_2^{\|})$, which is the *(jump) dependent part of* (S_1, S_2). This results in the same representation (2.7) as above.

This decomposition has also been presented in Cont and Tankov [7], Section 5.5 and Böcker and Klüppelberg [3], Section 3. Note that for completely dependent components we have $S_1^{\perp} = S_2^{\perp} = 0$ a.s. On the other hand, for independent components, the third part of the integral (2.6) or (2.8) is zero and this means that the components a.s. never jump together.

3 The Lévy copula

We shall present an estimation procedure for a bivariate compound Poisson process based on Lévy copulas. The reason for this is two-fold. Working with real data it may not be so easy to estimate statistically the components on the right-hand side of (2.7) so that the resulting statistical model is a bivariate compound Poisson process. Moreover, the ingredients require quite a number of parameters, which makes it desirable to find a parsimonious model. We are convinced that the notion of a Lévy copula plays here the same important role as a copula does for multivariate dfs.

Mainly for ease of notation we shall present our Lévy copula concept for spectrally non-negative CPPs only; i.e. for CPPs with non-negative jumps only. Since the Lévy

copula for a general CPP is defined for each quadrant separately, this is no restriction of the theory developed. Furthermore, the insurance claims data considered later also justify this restriction.

Lévy copulas are defined via the tail integral of a Lévy process.

Definition 3.1. *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*

$$\overline{\Pi}(x_1,\ldots,x_d) = \begin{cases} \Pi([x_1,\infty)\times\cdots\times[x_d,\infty))\,, & (x_1,\ldots,x_d)\in[0,\infty)^d \\ 0\,, & \text{if } x_i=\infty \text{ for at least one } i. \end{cases} \tag{3.1}$$

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

Next we define the Lévy copula for a spectrally positive Lévy process; for details see Nelson [14], Tankov [17] or Cont and Tankov [7].

Definition 3.2. *The* Lévy copula *of a spectrally positive Lévy process is a* d*-increasing grounded 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 notion of groundedness guarantees that $\mathfrak{C}$ defines a measure on $[0,\infty]^d$; indeed a Lévy copula is a d-dimensional measure with Lebesgue margins.

The following theorem is a version of Sklar's theorem for spectrally positive Lévy process; for a proof we refer to Tankov [17], Theorem 3.1.

Theorem 3.3 (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* $(x_1,x_2,\ldots,x_d) \in [0,\infty]^d$

$$\overline{\Pi}(x_1,\ldots,x_d) = \mathfrak{C}(\overline{\Pi}_1(x_1),\ldots,\overline{\Pi}_d(x_d)). \tag{3.2}$$

If the marginal tail integrals are continuous, then this Lévy copula is unique. Otherwise, it is unique on $Ran\overline{\Pi}_1 \times \ldots \times Ran\overline{\Pi}_d$*.*
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* (3.2) *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.*

This result opens up now a way of estimating multivariate compound Poisson processes by separating the marginal compound Poisson processes and coupling them with the dependence structure given by the Lévy copula. We shall show the procedure in details in the next section.

4 Maximum likelihood estimation of the parameters of a Lévy measure

Now the stage is set to tackle our main problem, namely the maximum likelihood estimation of the parameters of a bivariate spectrally positive CPP based on the observation of a sample path of the bivariate model in $[0,T]$ for fixed $T>0$.

Obviously, representation (2.3) suggests estimating the rate of the compound Poisson process based on the i.i.d. exponential arrival times and, independently, the bivariate distribution function of (Z_1,Z_2). Since both marginal random variables may have an atom in 0, and in the examples we are concerned about, they indeed have, we are faced with the estimation of a mixture model. This is one reason, why we base our estimation on representation (2.8). The other motivation comes from possible extensions of our estimation method to general Lévy processes; cf. Esmaeili and Klüppelberg [11].

Consequently, we assume throughout that the decomposition (2.7) holds for the observed path. We write for $t\in[0,T]$,

$$\begin{pmatrix} S_1(t) \\ S_2(t) \end{pmatrix} = \begin{pmatrix} \sum_{i=1}^{N_1(t)} X_i \\ \sum_{j=1}^{N_2(t)} Y_j \end{pmatrix} = \begin{pmatrix} S_1(t)^\perp + S_1(t)^\parallel \\ S_2(t)^\perp + S_2(t)^\parallel \end{pmatrix} = \begin{pmatrix} \sum_{i=1}^{N_1^\perp(t)} X_i^\perp + \sum_{j=1}^{N^\parallel(t)} X_j^\parallel \\ \sum_{i=1}^{N_2^\perp(t)} Y_i^\perp + \sum_{j=1}^{N^\parallel(t)} Y_j^\parallel \end{pmatrix} \tag{4.1}$$

with the familiar independence structure of the Poisson counting processes and the jump variables. Although, as described above, every bivariate CPP has three independent parts, the parts are linked by a common set of parameters in the frequency part as well as in the jump size distributions.

Our approach is an extension of the maximum likelihood method for the one-dimensional compound Poisson model; see e.g. Basawa and Prakasa Rao [1], Chapter 6.

Assume that we observe the bivariate CPP (S_1,S_2) continuously over a fixed time interval $[0,T]$. The process S_1 has frequency parameter $\lambda_1>0$ and jump size distribution F_1 and the process S_2 has frequency parameter $\lambda_2>0$ and jump size distribution F_2. Observing a CPP continuously over a time period is equivalent to observing all jump times and jump sizes in this time interval.

Let $N(T)=n$ denote the total number of jumps occurring in $[0,T]$, which decompose in the number $N_1^\perp(T)=n_1^\perp$ of jumps occurring only in the first component, the number $N_2^\perp(T)=n_2^\perp$ of jumps occurring only in the second component, and the number $N^\parallel(T)=n^\parallel$ of jumps occurring in both components. We denote by $\widetilde{x}_1,\dots,\widetilde{x}_{n_1^\perp}$ the observed jumps occurring only in the first component, by $\widetilde{y}_1,\dots,\widetilde{y}_{n_2^\perp}$ the observed jumps occurring only in the second component, and by $(x_1,y_1),\dots,(x_{n^\parallel},y_{n^\parallel})$ the observed jumps occurring in both components.

Theorem 4.1. *Assume an observation scheme as above. Assume that $\boldsymbol{\theta_1}$ is a parameter of the marginal density f_1 of the first jump component only, and $\boldsymbol{\theta_2}$ a parameter of the marginal density f_2 of the second jump component only, and that $\boldsymbol{\delta}$ is a parameter of the Lévy copula. Assume further that $\frac{\partial^2}{\partial u \partial v}\mathfrak{C}(u,v;\boldsymbol{\delta})$ exists for all $(u,v) \in (0,\lambda_1) \times (0,\lambda_2)$, which is the domain of $\mathfrak{C}$. Then the full likelihood of the bivariate CPP is given by*

$$
\begin{aligned}
& L(\lambda_1, \lambda_2, \boldsymbol{\theta_1}, \boldsymbol{\theta_2}, \boldsymbol{\delta}) \\
= \quad & (\lambda_1)^{n_1^\perp} e^{-(\lambda_1^\perp)T} \prod_{i=1}^{n_1^\perp} \left[f_1(\widetilde{x}_i; \boldsymbol{\theta_1}) \left(1 - \frac{\partial}{\partial u} \mathfrak{C}(u, \lambda_2; \boldsymbol{\delta}) \Big|_{u=\lambda_1 \overline{F}_1(\widetilde{x}_i; \boldsymbol{\theta_1})} \right) \right] \\
& \times (\lambda_2)^{n_2^\perp} e^{-(\lambda_2^\perp)T} \prod_{i=1}^{n_2^\perp} \left[f_2(\widetilde{y}_i; \boldsymbol{\theta_2}) \left(1 - \frac{\partial}{\partial v} \mathfrak{C}(\lambda_1, v; \boldsymbol{\delta}) \Big|_{v=\lambda_2 \overline{F}_2(\widetilde{y}_i; \boldsymbol{\theta_2})} \right) \right] \qquad (4.2) \\
& \times (\lambda_1 \lambda_2)^{n^\parallel} e^{-\lambda^\parallel T} \prod_{i=1}^{n^\parallel} \left[f_1(x_i; \boldsymbol{\theta_1}) f_2(y_i; \boldsymbol{\theta_2}) \frac{\partial^2}{\partial u \partial v} \mathfrak{C}(u, v; \boldsymbol{\delta}) \Big|_{u=\lambda_1 \overline{F}_1(x_i; \boldsymbol{\theta_1}), v=\lambda_2 \overline{F}_2(y_i; \boldsymbol{\theta_2})} \right]
\end{aligned}
$$

with $\lambda^\parallel = \lambda^\parallel(\boldsymbol{\delta}) = \mathfrak{C}(\lambda_1, \lambda_2, \boldsymbol{\delta})$ and $\lambda_i^\perp(\boldsymbol{\delta}) = \lambda_i - \lambda^\parallel(\boldsymbol{\delta})$ for $i = 1, 2$.

Proof. To calculate the likelihood function, we use representation (2.7) in combination with the independence as it is manifested in (2.8). This corresponds to the representation of the tail integrals for $i = 1, 2$ as

$$\overline{\Pi}_i =: \overline{\Pi}_i^\perp + \overline{\Pi}_i^\parallel ,$$

where $\overline{\Pi}_i$ denotes the marginal tail integral and $\overline{\Pi}_i^\perp$ and $\overline{\Pi}_i^\parallel$ are the tail integrals of the independent and jump dependent parts, respectively. Then, setting

$$\lambda^\parallel = \lim_{x,y \to 0^+} \overline{\Pi}(x,y) = \mathfrak{C}(\lambda_1, \lambda_2; \boldsymbol{\delta}) \quad \text{and} \quad \lambda_i^\perp = \lambda_i - \lambda^\parallel \quad \text{for} \quad i = 1, 2 ,$$

we obtain the independent parts and the jump dependent part of (S_1, S_2) as

$$
\begin{aligned}
\lambda_1^\perp \overline{F}_1^\perp(x) &= \lambda_1 \overline{F}_1(x) - \lambda^\parallel \overline{F}_1^\parallel(x) = \lambda_1 \overline{F}_1(x) - \mathfrak{C}(\lambda_1 \overline{F}_1(x), \lambda_2; \boldsymbol{\delta}) , \\
\lambda_2^\perp \overline{F}_2^\perp(y) &= \lambda_2 \overline{F}_2(y) - \lambda^\parallel \overline{F}_2^\parallel(y) = \lambda_2 \overline{F}_2(y) - \mathfrak{C}(\lambda_1, \lambda_2 \overline{F}_2(y); \boldsymbol{\delta}) , \qquad (4.3) \\
\lambda^\parallel \overline{F}^\parallel(x,y) &= \mathfrak{C}(\lambda_1 \overline{F}_1(x), \lambda_2 \overline{F}_2(y); \boldsymbol{\delta}) , \quad x, y > 0
\end{aligned}
$$

Let now $L_1(\lambda_1^\perp, \boldsymbol{\theta_2})$ be the marginal likelihood function based on the observations of the jump times and jump sizes of the first component $S_1^\perp$. To derive L_1 let $\widetilde{t}_1, \ldots, \widetilde{t}_{n_1^\perp}$ denote the jump times of $S_1^\perp$, and define the sequence of inter-arrival times $\widetilde{T}_k = \widetilde{t}_k - \widetilde{t}_{k-1}$ for $k = 1, \ldots, n_1^\perp$ with $\widetilde{t}_0 = 0$. Then the $\widetilde{T}_k$ are iid exponential random variables with mean $1/\lambda_1^\perp$ and they are independent of the observed jump sizes $\widetilde{x}_1, \ldots, \widetilde{x}_{n_1^\perp}$. The likelihood

function of the observations concerning $S_1^\perp$ is given by

$$\begin{aligned} L_1(\lambda_1^\perp, \boldsymbol{\theta_1}) &= \prod_{i=1}^{n_1^\perp} \left(\lambda_1^\perp e^{-\lambda_1^\perp \widetilde{T}_i}\right) \times \; e^{-\lambda_1^\perp (T - \widetilde{t}_{n_1^\perp})} \times \prod_{i=1}^{n_1^\perp} f_1^\perp(\widetilde{x}_i; \boldsymbol{\theta_1}) \\ &= (\lambda_1^\perp)^{n_1^\perp} e^{-\lambda_1^\perp T} \prod_{i=1}^{n_1^\perp} f_1^\perp(\widetilde{x}_i; \boldsymbol{\theta_1}) \,, \end{aligned} \tag{4.4}$$

where the density $f_1^\perp$ is found by taking the derivative in the first equation of (4.3). The second part $S_2^\perp$ is treated analogously and we obtain $L_2(\lambda_2^\perp, \boldsymbol{\theta_2})$ as (4.4) with $\lambda_1^\perp$ replaced by $\lambda_2^\perp$ and $f_1^\perp(\widetilde{x}_i, \boldsymbol{\theta_1})$ replaced by $f_2^\perp(\widetilde{y}_i, \boldsymbol{\theta_2})$. For the joint jump part of the process, that is $(S_1^\parallel, S_2^\parallel)$, we observe the number $n^\parallel = n_1 - n_1^\perp = n_2 - n_2^\perp$ of joint jumps with frequency $\lambda^\parallel$ at times $t_1, \ldots, t_{n^\parallel}$ with the observed bivariate jump sizes $(x_1, y_1), \ldots, (x_{n^\parallel}, y_{n^\parallel})$. Denote $T_k = t_k - t_{k-1}$ and $F^\parallel(x, y)$ the joint distribution of the jump sizes with joint density $f^\parallel(x, y)$. These are observations of a jump dependent CPP with frequency parameter $\lambda^\parallel$ and Lévy measure concentrated in $(0, \infty)^2$. Recall the formula for $(x, y) \in (0, \infty)^2$, which is a consequence of the formula after Theorem 5.4 on p. 148 in Cont and Tankov [7],

$$\Pi(dx, dy) = \frac{\partial^2}{\partial u \partial v} \mathfrak{C}(u, v, \boldsymbol{\delta}) \Big|_{u = \lambda_1 \overline{F}_1(x, \boldsymbol{\theta_1}), v = \lambda_2 \overline{F}_2(y, \boldsymbol{\theta_2})} \Pi_1(dx) \Pi_2(dy) \,.$$

In our case the joint density of the Lévy measure on the left hand side is given by $\lambda^\parallel f^\parallel(x, y)$. The derivative $\frac{\partial^2}{\partial u \partial v} \mathfrak{C}(u, v, \boldsymbol{\delta})$ exists by assumption. Then the likelihood of the joint jump process is given by the product in the third line of (4.2). This concludes the proof. □

Remark 4.2. Note that this estimation procedure ensures that the estimated model is again a bivariate CPP.

This applies for instance to the following parametric Lévy copula family.

Example 4.3. [Clayton Lévy copula]
The Clayton Lévy copula is defined as

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

where $\delta > 0$ is the Lévy copula parameter. We calculate

$$\begin{aligned} \frac{\partial}{\partial u} \mathfrak{C}(u, v) &= \left(1 + (\frac{u}{v})^\delta\right)^{-1/\delta - 1} , \\ \frac{\partial^2}{\partial u \partial v} \mathfrak{C}(u, v) &= (\delta + 1)(uv)^{-\delta - 1} (u^{-\delta} + v^{-\delta})^{-1/\delta - 2} \,, \\ &= (\delta + 1)(uv)^{\delta} (u^{\delta} + v^{\delta})^{-1/\delta - 2} \,, \quad u, v > 0 \,. \end{aligned}$$

We observe that the joint jump intensity is given by

$$\lambda^{\|} = (\lambda_1^{-\delta} + \lambda_2^{-\delta})^{-\frac{1}{\delta}} .$$

Two specific examples, which will be used later, are the following:
(i) For the *exponential Clayton model* the marginal jump distributions are for $i = 1, 2$ exponentially distributed with parameters $\theta_i > 0$ and densities $f_i(z; \theta_i) = \theta_i e^{-\theta_i z}$ for $z \geq 0$. The likelihood function for the continuously observed bivariate process $(S_1(t), S_2(t))_{0 \leq t \leq T}$ with the notation as in Theorem 4.1 is given by

$$\begin{aligned} L(\lambda_1, \lambda_2, \theta_1, \theta_2, \delta) &= (\theta_1 \lambda_1)^{n_1^{\perp}} e^{-\lambda_1^{\perp} T - \theta_1 \sum_{i=1}^{n_1^{\perp}} \widetilde{x}_i} \prod_{i=1}^{n_1^{\perp}} \left[1 - \left(1 + (\frac{\lambda_1}{\lambda_2})^{\delta} e^{-\delta \theta_1 \widetilde{x}_i} \right)^{-\frac{1}{\delta} - 1} \right] , \\ &\times (\theta_2 \lambda_2)^{n_2^{\perp}} e^{-\lambda_2^{\perp} T - \theta_2 \sum_{i=1}^{n_2^{\perp}} \widetilde{y}_i} \prod_{i=1}^{n_2^{\perp}} \left[1 - \left(1 + (\frac{\lambda_2}{\lambda_1})^{\delta} e^{-\delta \theta_2 \widetilde{y}_i} \right)^{-\frac{1}{\delta} - 1} \right] , \\ &\times \left((1+\delta) \theta_1 \theta_2 (\lambda_1 \lambda_2)^{\delta+1} \right)^{n^{\|}} e^{-\lambda^{\|} T - (1+\delta)(\theta_1 \sum_{i=1}^{n^{\|}} x_i + \theta_2 \sum_{i=1}^{n^{\|}} y_i)} \\ &\times \prod_{i=1}^{n^{\|}} (\lambda_1^{\delta} e^{-\theta_1 \delta x_i} + \lambda_2^{\delta} e^{-\theta_2 \delta y_i})^{-\frac{1}{\delta} - 2} . \end{aligned}$$

(ii) For the *Weibull Clayton model* the marginal jump distributions are for $i = 1, 2$ Weibull distributed with parameters $a_i, b_i > 0$ and densities $w_i(z; a_i, b_i) = \frac{b_i}{a_i^{b_i}} z^{b_i - 1} e^{-(z/a_i)^{b_i}}$ for $z \geq 0$. The likelihood function for the continuously observed bivariate process $(S_1(t), S_2(t))_{0 \leq t \leq T}$ is given by

$$\begin{aligned} & L(\lambda_1, \lambda_2, a_1, b_1, a_2, b_2, \delta) \qquad (4.5) \\ &= \left(\lambda_1 b_1 a_1^{-b_1} \right)^{n_1^{\perp}} e^{-\lambda_1^{\perp} T - \sum_{i=1}^{n_1^{\perp}} (\widetilde{x}_i / a_1)^{b_1}} \prod_{i=1}^{n_1^{\perp}} \left[\widetilde{x}_i^{b_1 - 1} \left(1 - \left(1 + \left(\frac{\lambda_1 e^{-(\widetilde{x}_i / a_1)^{b_1}}}{\lambda_2} \right)^{\delta} \right)^{-1/\delta - 1} \right) \right] \\ &\times \left(\lambda_2 b_2 a_2^{-b_2} \right)^{n_2^{\perp}} e^{-\lambda_2^{\perp} T - \sum_{i=1}^{n_2^{\perp}} (\widetilde{y}_i / a_2)^{b_2}} \prod_{i=1}^{n_2^{\perp}} \left[\widetilde{y}_i^{b_2 - 1} \left(1 - \left(1 + \left(\frac{\lambda_2 e^{-(\widetilde{y}_i / a_2)^{b_2}}}{\lambda_1} \right)^{\delta} \right)^{-1/\delta - 1} \right) \right] \\ &\times \left((1+\delta)(\lambda_1 \lambda_2)^{1+\delta} b_1 b_2 a_1^{-b_1} a_2^{-b_2} \right)^{n^{\|}} e^{-\lambda^{\|} T - (1+\delta) \sum_{i=1}^{n^{\|}} \left((x_i / a_1)^{b_1} + (y_i / a_2)^{b_2} \right)} \\ &\times \prod_{i=1}^{n^{\|}} \left[{x_i}^{b_1 - 1} {y_i}^{b_2 - 1} \left(\left(\lambda_1 e^{-(x_i / a_1)^{b_1}} \right)^{\delta} + \left(\lambda_2 e^{-(y_i / a_2)^{b_2}} \right)^{\delta} \right)^{-1/\delta - 2} \right] \end{aligned}$$

□

For a bivariate CPP, where dependence is modelled by a Lévy copula, the bivariate distribution of the joint jumps of the process exhibits a specific dependence structure, which can also be described by a distributional copula, or better by the corresponding survival copula. We explain this for the Clayton Lévy copula.

Example 4.4. [Continuation of Example 4.3]
Denote by $\overline{C}$ the survival copula of the joint jumps of $(S_1(t)^{\|}, S_2(t)^{\|})_{t\geq 0}$ given by

$$\overline{F}^{\|}(x,y) = \overline{C}\left(\overline{F}_1^{\|}(x), \overline{F}_2^{\|}(y)\right). \tag{4.6}$$

Assume further that the jump distributions F_1 and F_2 have no atom at 0. From the last equation of (4.3) we see that

$$\overline{F}_1^{\|}(x) = \lim_{y\to 0} \frac{1}{\lambda^{\|}} \mathfrak{C}(\lambda_1 \overline{F}_1(x), \lambda_2 \overline{F}_2(y))$$

and analogously for $\overline{F}_2^{\|}$. Here equation (4.6) can be rewritten as

$$\frac{1}{\lambda^{\|}} \mathfrak{C}(\lambda_1 \overline{F}_1(x), \lambda_2 \overline{F}_2(y)) = \overline{C}\left(\frac{1}{\lambda^{\|}} \mathfrak{C}(\lambda_1 \overline{F}_1(x), \lambda_2), \frac{1}{\lambda^{\|}} \mathfrak{C}(\lambda_1, \lambda_2 \overline{F}_2(y))\right).$$

For the Clayton Lévy copula $\mathfrak{C}$ the right hand side is equal to

$$\overline{C}\left(\left(\frac{(\lambda_1 \overline{F}_1(x))^{-\delta} + \lambda_2^{-\delta}}{\lambda_1^{-\delta} + \lambda_2^{-\delta}}\right)^{-\frac{1}{\delta}}, \left(\frac{\lambda_1^{-\delta} + (\lambda_2 \overline{F}_2(y))^{-\delta}}{\lambda_1^{-\delta} + \lambda_2^{-\delta}}\right)^{-\frac{1}{\delta}}\right) = \left(\frac{(\lambda_1 \overline{F}_1(x))^{-\delta} + (\lambda_2 \overline{F}_2(y))^{-\delta}}{\lambda_1^{-\delta} + \lambda_2^{-\delta}}\right)^{-\frac{1}{\delta}}$$

Abbreviating the arguments of $\overline{C}$ by u and v (note that $u, v \in (0,1)$) gives

$$(\lambda_1 \overline{F}_1(x))^{-\delta} = u^{-\delta}(\lambda_1^{-\delta} + \lambda_2^{-\delta}) - \lambda_2^{-\delta} \quad \text{and} \quad (\lambda_2 \overline{F}_2(y))^{-\delta} = v^{-\delta}(\lambda_1^{-\delta} + \lambda_2^{-\delta}) - \lambda_1^{-\delta},$$

such that

$$\begin{aligned} \overline{C}(u,v) &= \left(\frac{u^{-\delta}(\lambda_1^{-\delta} + \lambda_2^{-\delta}) - \lambda_2^{-\delta} + v^{-\delta}(\lambda_1^{-\delta} + \lambda_2^{-\delta}) - \lambda_1^{-\delta}}{\lambda_1^{-\delta} + \lambda_2^{-\delta}}\right)^{-\frac{1}{\delta}} \\ &= (u^{-\delta} + v^{-\delta} - 1)^{-\frac{1}{\delta}}, \end{aligned}$$

which is the well-known distributional Clayton copula; cf. Cont and Tankov [7], eq. (5.3) or Joe [12], Family B4 on p. 141. □

5 A simulation study

In this section we study the quality of our estimates in a small simulation study. This means that we first have to simulate sample paths of a bivariate CPP on $[0,T]$ for pre-specified $T > 0$, equivalently, we simulate the jump times and jump sizes (independently) in this time interval.

In Section 6 of Cont and Tankov [7] various simulation algorithms for Lévy processes have been suggested. We extend here their Algorithm 6.2 to a bivariate setting by invoking

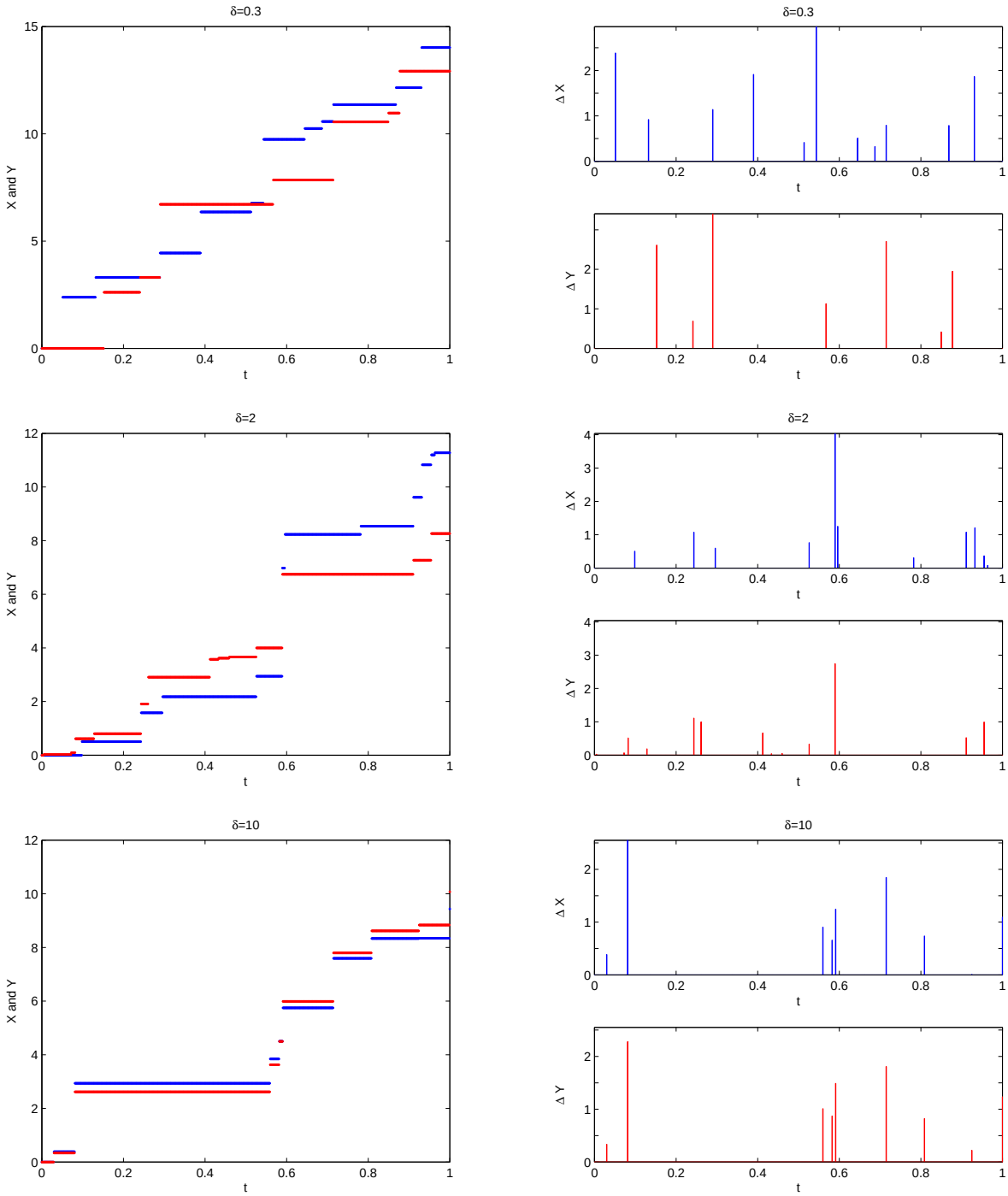

Figure 1: Simulation of three bivariate CPPs with exponentially distributed jumps and a Clayton Lévy copula with dependence parameter $\delta = 0.3$ (top), $\delta = 2$ (middle) and $\delta = 10$ (below). The left hand figures show the sample paths of the CPPs, whereas the right hand figures present the same paths as marked Poisson process.

decomposition (4.1) for given λ_1, λ_2, marginal jump distribution functions F_1, F_2 and a Lévy copula $\mathfrak{C}$.

As we work with a fully parametric bivariate model, we assume that we are given frequency parameters $\lambda_1, \lambda_2 > 0$, the parameters of the marginal jump size distributions $\boldsymbol{\theta_1} \in \mathbb{R}^{k_1}$, $\boldsymbol{\theta_2} \in \mathbb{R}^{k_2}$ for some $k_1, k_2 \in \mathbb{N}$ and, finally, the dependence parameter $\boldsymbol{\delta} \in \mathbb{R}^m$ of the Lévy copula. Moreover, we choose a time interval $[0, T]$.

Then the number of points in the first component is Poisson distributed with frequency $\lambda_1 T$, so generate a Poisson random number $N_1(T)$ with mean $\lambda_1 T$. The number of points in the second component is Poisson distributed with frequency $\lambda_2 T$, so generate a Poisson random number $N_2(T)$ with mean $\lambda_2 T$. Then $\lambda^{\|} T = \mathfrak{C}(\lambda_1, \lambda_2) T$ is the frequency parameter of the joint jumps, so simulate another Poisson random number $N^{\|}(T)$ with frequency $\lambda^{\|} T$. This implies then that $N_1^{\perp}(T) = N_1(T) - N^{\|}(T)$ and $N_2^{\perp}(T) = N_2(T) - N^{\|}(T)$.

Now conditional on these numbers, the Poisson points are uniformly distributed in the interval $[0, T]$, so simulate the correct number of $[0, T]$-uniformly distributed random variables, independently for the three components: $U_{1,i}^{\perp}$ for $i = 1, \ldots, N_1^{\perp}(T)$, $U_{2,i}^{\perp}$ for $i = 1, \ldots, N_2^{\perp}(T)$, and $U_i^{\|}$ for $i = 1, \ldots, N^{\|}(\mathrm{T})$.

Next we simulate the jump sizes. Denote by $(U_{1,i}^{\perp}, X_i^{\perp})$ for $i = 1, \ldots, N_1^{\perp}(T)$, $(U_{2,i}^{\perp}, Y_i^{\perp})$ for $i = 1, \ldots, N_2^{\perp}(T)$ and $(U_i^{\|}, X_i^{\|}, Y_i^{\|})$ for $i = 1, \ldots, N^{\|}(T)$ the marked points of the single jumps and the joint jumps, respectively, then the bivariate trajectory is given by

$$\begin{pmatrix} S_1(t) \\ S_2(t) \end{pmatrix} = \begin{pmatrix} \sum_{i=1}^{N_1^{\perp}(T)} 1_{\{U_{1,i}^{\perp} < t\}} X_i^{\perp} + \sum_{i=1}^{N^{\|}(T)} 1_{\{U_i^{\|} < t\}} X_i^{\|} \\ \sum_{i=1}^{N_2^{\perp}(T)} 1_{\{U_{2,i}^{\perp} < t\}} Y_i^{\perp} + \sum_{i=1}^{N^{\|}(T)} 1_{\{U_i^{\|} < t\}} Y_i^{\|} \end{pmatrix}, \quad 0 < t < T\,.$$

For the marks on these points given by the corresponding jump sizes we need then $N_1^{\perp}(T)$ iid jump sizes with df $F_1^{\perp}$, $N_2^{\perp}(T)$ iid jump sizes with df $F_2^{\perp}$, and $N^{\|}(T)$ bivariate jump sizes with df $F^{\|}$, all of them independent. Single jump sizes are generated by $X_i^{\perp} \stackrel{d}{=} F_1^{\perp \leftarrow}(U_i)$, $i = 1, \cdots, N_1^{\perp}(T)$ and $Y_i^{\perp} \stackrel{d}{=} F_2^{\perp \leftarrow}(U_i)$, $i = 1, \cdots, N_2^{\perp}(T)$, where for any increasing function h its generalized inverse is defined as

$$h^{\leftarrow}(u) := \inf\{s \in \mathbb{R} : h(s) \geq u\}\,,$$

(which coincides with the analytical inverse, provided h is strictly monotone).

It remains to simulate the joint jumps $(X_j^{\|}, Y_j^{\|})$ for $j = 1, \ldots, N^{\|}(T)$. We use the joint survival copula $\overline{C}$ as in (4.6). We simulate standard uniform independent random variables $U_1, \ldots, U_{N^{\|}(T)}$, $V_1, \ldots, V_{N^{\|}(T)}$ and recall that $X_j^{\|} \stackrel{d}{=} F_1^{\| \leftarrow}(U_j)$. Then the following

Value	$\widehat{\text{mean}}$	$\widehat{\text{MSE}}$	$\widehat{\text{MAE}}$	$\widehat{\text{MRB}}$
$\delta = 0.5$	0.4995	0.0036	0.0492	-0.0012
	(0.0597)	(0.0054)	(0.0339)	
$\delta = 1$	0.9896	0.0094	0.0748	0.0147
	(0.0964)	(0.0150)	(0.0617)	
$\delta = 3$	3.0583	0.0834	0.2314	-0.0321
	(0.2828)	(0.1121)	(0.1727)	
$\delta = 5$	5.0279	0.2027	0.3511	0.0147
	(0.4494)	(0.2764)	(0.2819)	

Table 1: Mean, mean squared errors (MSE), mean absolute error (MAE) and mean relative bias (MRB) are presented for 100 MLEs of the Lévy copula parameter of a bivariate exponential Clayton model (Example 5.1). Each estimate is calculated from an observed sample path of a bivariate CPP with parameters $\lambda_1 = 100, \lambda_2 = 80, \theta_1 = 1, \theta_2 = 2$ (which are assumed to be known) and unknown dependence parameter δ. The values in brackets show the standard deviation of estimates.

standard calculation for a generic pair $(X^{\|}, Y^{\|})$ is well-known:

$$\begin{aligned}
\lim_{\Delta x \to 0} P(Y^{\|} > y \mid x < X^{\|} \leq x + \Delta x) &= \lim_{\Delta x \to 0} \frac{\overline{F}^{\|}(x,y) - \overline{F}^{\|}(x+\Delta x, y)}{P(x < \Delta X^{\|} \leq x + \Delta x)} \\
= -\frac{\partial \overline{F}^{\|}(x,y)}{\partial x} \frac{1}{f_1^{\|}(x)} &= -\frac{\partial \overline{C}(\overline{F}_1^{\|}(x), \overline{F}_2^{\|}(y))}{\partial x} \frac{1}{f_1^{\|}(x)} \\
= \frac{\partial}{\partial u} \overline{C}(u, \overline{F}_2^{\|}(y)) \Big|_{u = \overline{F}_1^{\|}(x)} &=: \overline{H}_x(y) \,. \qquad (5.1)
\end{aligned}$$

Now we take the generalized inverse $H_x^{\leftarrow}$ and define $Y_j^{\|} \stackrel{d}{=} H_x^{\leftarrow}(V_j)$. Then the following calculation convinces us that this algorithm works:

$$\begin{aligned}
P(F_1^{\|\leftarrow}(U) > x, H_{X^{\|}}^{\leftarrow}(V) > y) &= P(X^{\|} > x) P(H_{X^{\|}}^{\leftarrow}(V) > y \mid X^{\|} > x) \\
&= P(X^{\|} > x) \int_x^{\infty} P(Y^{\|} > y \mid X^{\|} = t) dF_1^{\|}(t) \\
&= P(X^{\|} > x) P(Y^{\|} > y \mid X^{\|} > x) \\
&= P(X^{\|} > x, Y^{\|} > y) \,.
\end{aligned}$$

Example 5.1. [Simulation of a bivariate exponential Clayton model, continuation of Examples 4.3 and 4.4]
Let (S_1, S_2) be a bivariate CPP with exponentially distributed jump sizes, i.e. $\overline{F}_i(z) = e^{-\theta_i z}$, $z > 0$, for $i = 1, 2$, and the dependence structure of a Clayton Lévy copula $\mathfrak{C}$ with parameter $\delta > 0$. Assume further $\lambda_1, \lambda_2 > 0$ are the intensities of the marginal Poisson processes. We simulate a bivariate exponential Clayton model over the time interval $[0, 1]$.

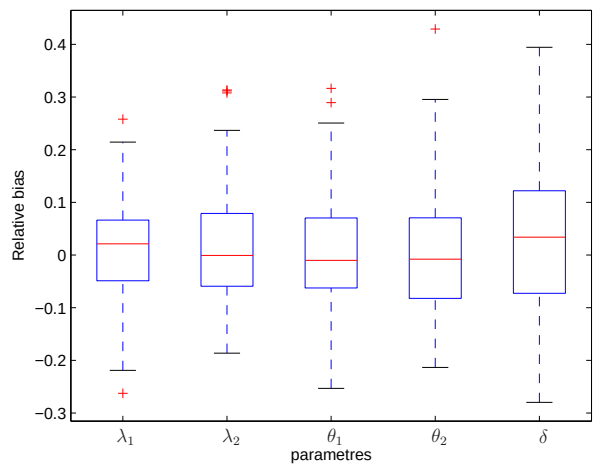

Figure 2: Box-plots of the relative bias for the estimates of the exponential Clayton model with parameter values as in Table 2.

We apply the above simulation algorithm. The distribution functions of the single jump sizes of the process are for $i = 1, 2$ given by

$$\overline{F}_i^{\perp}(z) = \frac{1}{\lambda_i^{\perp}} \left\{ \lambda_i e^{-\theta_i z} - \left(\lambda_1^{-\delta} e^{\theta_1 \delta z (2-i)} + \lambda_2^{-\delta} e^{\theta_2 \delta z (i-1)} \right)^{-\frac{1}{\delta}} \right\}, \quad z > 0\,,$$

and the bivariate distribution function for the joint jumps has the form

$$\overline{F}^{\|}(x, y) = \frac{1}{\lambda^{\|}} \left(\lambda_1^{-\delta} e^{\theta_1 \delta x} + \lambda_2^{-\delta} e^{\theta_2 \delta y} \right)^{-\frac{1}{\delta}}, \quad x, y > 0$$

with margins $\overline{F}_1^{\|}(x) = \frac{1}{\lambda^{\|}} \left(\lambda_1^{-\delta} e^{\theta_1 \delta x} + \lambda_2^{-\delta} \right)^{-\frac{1}{\delta}}$, $x > 0$, and $\overline{F}_2^{\|}(y) = \frac{1}{\lambda^{\|}} \left(\lambda_1^{-\delta} + \lambda_2^{-\delta} e^{\theta_2 \delta y} \right)^{-\frac{1}{\delta}}$, $y > 0$.

The simulation algorithm:

(a) Generate two random numbers N_1 and N_2 from Poisson distributions with parameters λ_1 and λ_2, respectively. Generate $N^{\|}$ from a Poisson distribution with parameter $\lambda^{\|} = \mathfrak{C}(\lambda_1, \lambda_2) = (\lambda_1^{-\delta} + \lambda_2^{-\delta})^{-1/\delta}$.

(b) Generate $N^{\|}$, $N_1^{\perp} = N_1 - N^{\|}$ and $N_2^{\perp} = N_2 - N^{\|}$ independent $[0, 1]$-uniformly distributed random variables. These are the Poisson points of joint and single jumps.

(c) Generate independent $U_1, \ldots, U_{N_1^{\perp}}$ and $V_1, \ldots, V_{N_2^{\perp}}$ standard uniform random variables. Then the single jump sizes of both components are found by taking the inverse of $F_1^{\perp}$ and $F_2^{\perp}$, that is, $X_i^{\perp} \stackrel{d}{=} F_1^{\perp \leftarrow}(U_i)$, $i = 1, \ldots, N_1^{\perp}$ and $Y_j^{\perp} \stackrel{d}{=} F_2^{\perp \leftarrow}(V_j)$, $j = 1, \ldots, N_2^{\perp}$.

	$\widehat{\lambda}_1$	$\widehat{\lambda}_2$	$\widehat{\theta}_1$	$\widehat{\theta}_2$	$\widehat{\delta}$
Values	100	80	1.00	2.00	1.00
$\widehat{\text{mean}}$	100.8377 (9.8302)	80.4022 (8.7985)	1.0105 (0.0979)	2.0326 (0.2158)	1.0097 (0.1197)
$\widehat{\text{MSE}}$	97.3344 (141.0848)	78.9570 (113.1887)	0.0097 (0.0168)	0.0476 (0.0714)	0.0144 (0.0202)
$\widehat{\text{MRB}}$	0.0116	0.0203	-0.0087	0.0022	0.0423

Table 2: Estimated mean, mean squared error (MSE) and mean relative bias (MRB) of 100 MLEs of an exponential Clayton model with estimated standard deviations for mean and MSE in brackets.

(d) For the bivariate jump sizes, generate new independent $[0,1]$-uniform $U_1,\dots,U_{N^\parallel}$ and $V_1,\dots,V_{N^\parallel}$ random variables. Then $X_i^\parallel \stackrel{d}{=} F_1^{\parallel\leftarrow}(U_i)$ and, given $X_i^\parallel = x$, $Y_i^\parallel \stackrel{d}{=} H_x^{\leftarrow}(V_i)$, $i=1,\dots,N^\parallel$, where for fixed $x>0$, as shown in (5.1),

$$\begin{aligned}
\overline{H}_x(y) = \frac{\partial}{\partial u}\overline{C}(u,\overline{F}_2^\parallel(y))\Big|_{u=\overline{F}_1^\parallel(x)} &= \left(1+(\frac{u}{v})^\delta - u^\delta\right)^{-1/\delta-1}\Big|_{u=\overline{F}_1^\parallel(x),v=\overline{F}_2^\parallel(y)} \\
&= \left(1+\frac{\lambda_1^{-\delta}+\lambda_2^{-\delta}e^{\theta_2\delta y}}{\lambda_1^{-\delta}e^{\theta_1\delta x}+\lambda_2^{-\delta}} - \frac{{\lambda^\parallel}^{-\delta}}{\lambda_1^{-\delta}e^{\theta_1\delta x}+\lambda_2^{-\delta}}\right)^{-1/\delta-1} \\
&= \left(\frac{\lambda_1^{-\delta}e^{\theta_1\delta x}+\lambda_2^{-\delta}e^{\theta_2\delta y}}{\lambda_1^{-\delta}e^{\theta_1\delta x}+\lambda_2^{-\delta}}\right)^{-\frac{1}{\delta}-1}, \quad y>0\,.
\end{aligned}$$

Various scenarios are depicted in Figure 1. □

Next we show the performance of the MLE estimation from Section 4 based on simulated sample paths.

Example 5.2. [Estimation of a bivariate exponential Clayton model, continuation of Examples 4.3, 4.4, and 5.1]
Let (S_1,S_2) be a bivariate CPP with exponentially distributed jump sizes, i.e. $\overline{F}_i(z) = e^{-\theta_i z}$, $z>0$, for $i=1,2$, and the dependence structure of a Clayton Lévy copula $\mathfrak{C}$ with parameter $\delta>0$. Assume further $\lambda_1,\lambda_2>0$ are the intensities of the marginal Poisson processes. We simulate 100 sample paths of a bivariate exponential Clayton model with parameters $\lambda_1=100, \lambda_2=80, \theta_1=1, \theta_2=2$ and different δ over the time interval $[0,1]$ and estimate for each sample path the parameters. The results are summarized in Tables 1 and 2 and Figure 2. Note that the critical parameter is the dependence parameter δ; cf. Figure 2. From Table 1 we note from the estimated MSE and MAE that its estimation is more precise for small δ than for large. The mean relative bias, on the other hand, remains for all δ near 0. Similar interpretations can be read off from Table 2. □

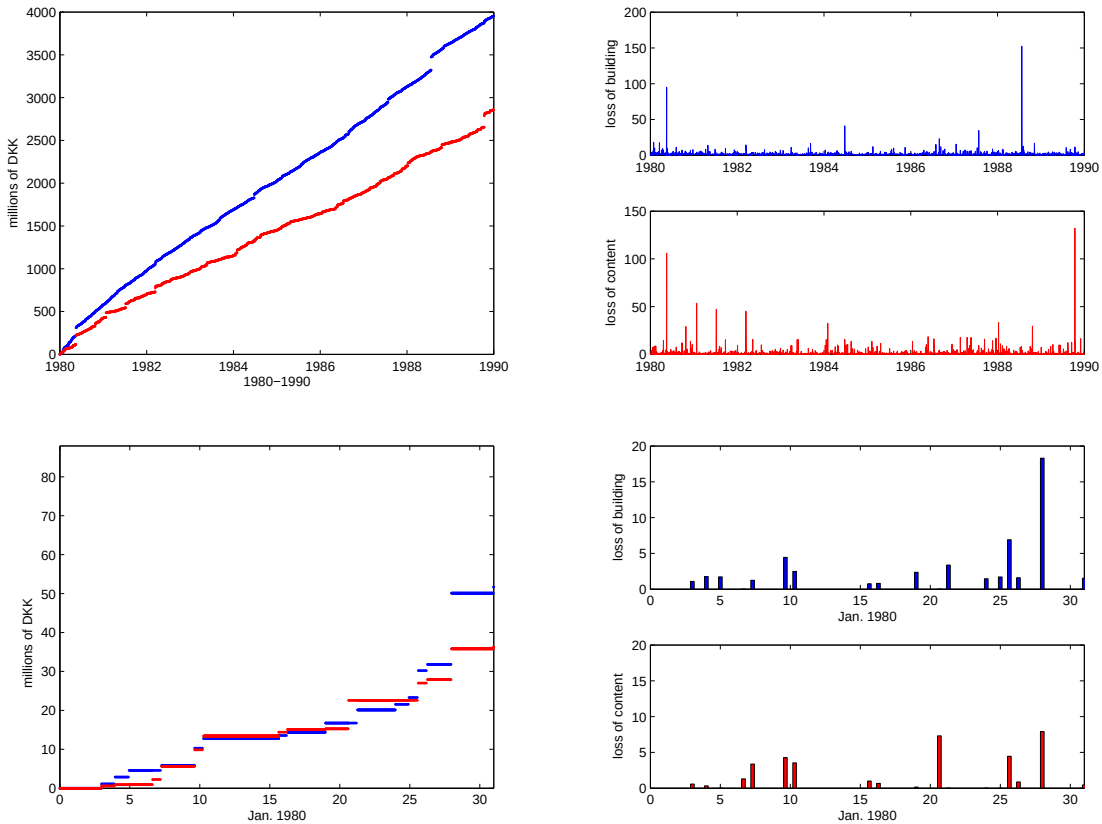

Figure 3: The Danish fire insurance data: The top figures show the total losses (left) and the individual losses (right) over the period 1980-1990. The figures below depict the data only for the one-month period of January 1980.

6 A real data analysis

In this section we fit a CPP to a bivariate data set. The data we fit is called the Danish fire insurance data and appears in aggregated form in Embrechts et al. [10], Figure 6.2.11. The data are available at *www.ma.hw.ac.uk/∼mcneil/*. As described there, the data were collected at Copenhagen Reinsurance and comprise 2167 fire losses over the period 1980 to 1990. They have been adjusted for inflation to reflect 1985 values and are expressed in millions of Danish Kroner. Every total claim has been divided into loss of building, loss of content and loss of profit. Since the last variable rarely has non-zero value, we restrict ourselves to the first two variables. Figure 3 shows the time series and the aggregated process of the data in the whole and in a one-month period of time.

We shall estimate a full bivariate parametric model based on the likelihood function of Theorem 4.1. This means that we have to specify the marginal distributions for the losses of buildings and the losses of contents, and we do this for the logarithmic data.

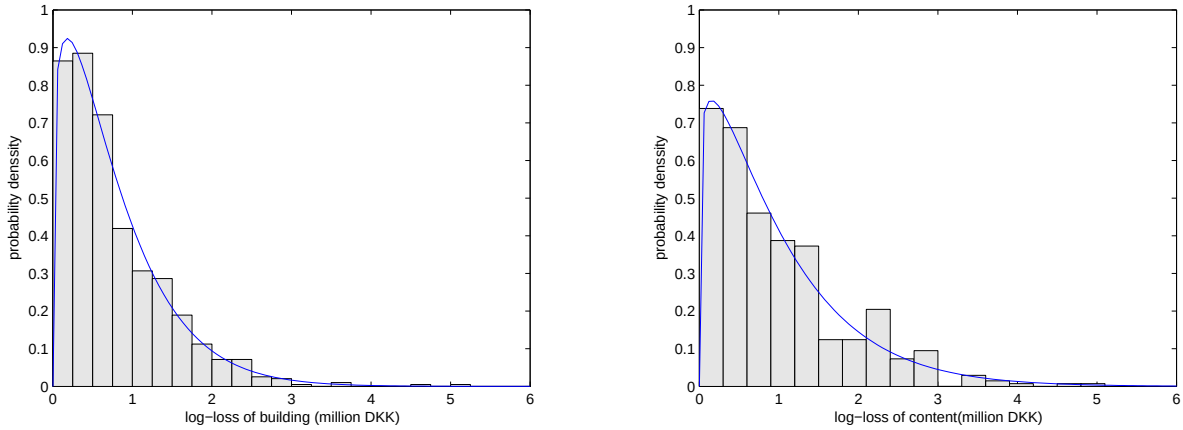

Figure 4: Histogram (and estimated Weibull density) of the logarithmic losses of buildings (left) and logarithmic losses of content based on the Danish fire insurance data larger than 1 million Danish Kroner in both variables.

As explained above the bivariate data come from originally aggregated data, where the claims (sum of losses of buildings, contents and profits) are larger than one million Danish Kroner. Due to the splitting of the data in losses of buildings and losses of contents, certain losses have become smaller than the threshold for the aggregated data, such effects also appear due to the inflation adjustment. To guarantee that the bivariate data we want to fit come from the same distribution, we have based our analysis on those data, which are larger than one million Danish Kroner after inflation adjustment in both coordinates. This amounts to 940 data points.

An explorative data analysis shows that the family of two-parameter Weibull distributions are appropriate for the log-data. We present the histograms of the log-transformed data in Figures 4 with fitted marginal Weibull densities as presented in Example 4.3(ii). The marginal parameters have been fitted by maximum likelihood estimation giving

$$f_1(x) = 1.5225(\log x)^{0.1954} \exp\left(-1.2737(\log x)^{1.1954}\right), \quad x > 1$$

$$f_2(y) = 1.0863(\log y)^{0.1289} \exp\left(-0.9622(\log y)^{1.1289}\right), \quad y > 1.$$

The corresponding QQ-plots are depicted in Figure 5.

It is worth mentioning that modelling with Lévy copulas is useful, when the dependence structure of the Poisson processes matches the dependence of the jump sizes. The reason for this is that the parameter of the Lévy copula models the dependence structure of the Lévy measure, which comprises the intensity of the jumps and the distribution of the jump sizes. By Sklar's theorem for Lévy copulas (cf. Theorem 3.3), if the data follow a bivariate compound Poisson process, this kind of dependence structure is exactly, what we expect.

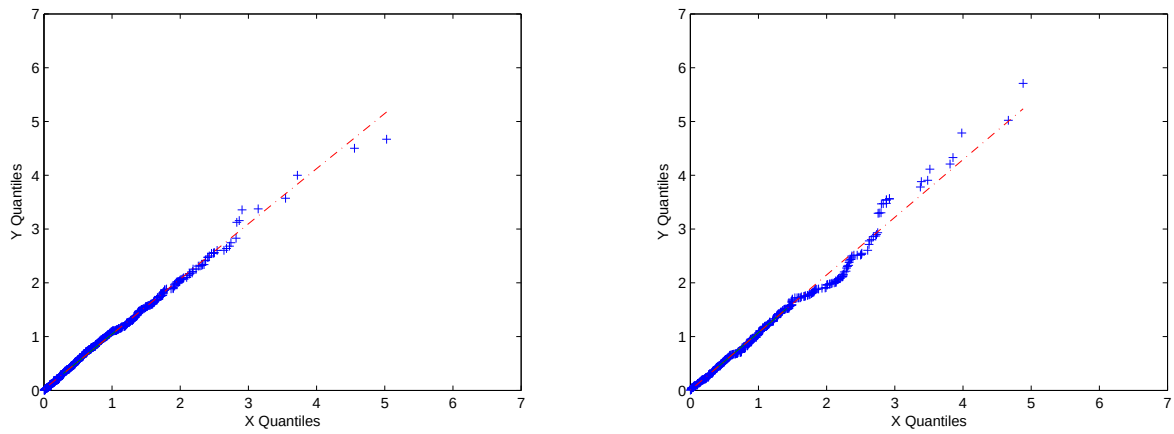

Figure 5: QQ-plot of the logarithmic Danish fire insurance data versus their estimated Weibull distributions, with parameters estimated from the data set. Left (loss of building), right (loss of content)

To check the suitability of the model for these data, we first estimate the parameter of the Clayton Lévy copula based on the point processes only. This results in solving the equation

$$(\widehat{\lambda}_1^{-\delta} + \widehat{\lambda}_2^{-\delta})^{-\frac{1}{\delta}} = \widehat{\lambda}^{\|} \,,$$

where $\widehat{\lambda}_1$, $\widehat{\lambda}_2$ and $\widehat{\lambda}^{\|}$ are the estimated intensities for each of the marginal univariate Poisson processes and the jump dependent part of the process. We obtain $\widehat{\delta} = 1.0546$.

Second, we should compare this estimator with the corresponding estimator based on the jump sizes. For this we invoke Example 4.4, which shows that the Clayton Lévy copula for a bivariate CPP implies a distributional Clayton copula for the joint jump sizes of the process. The maximum likelihood estimator of the parameter δ based on the joint jumps is obtained as $\widehat{\delta} = 0.8675$. This is close enough to convince us that a bivariate compound Poisson process is a good model for the Danish fire insurance data, and that the Clayton Lévy copula is an appropriate model.

Now we consider the full likelihood as given in equation (4.2), with two-parametric marginal Weibull distributions for the log-sizes of the claims and a Clayton Lévy copula $\mathfrak{C}$. Furthermore, we denote by λ_1 and λ_2, the intensities of losses in each component. Then the full likelihood including seven parameters is given by equation (4.5).

The resulting maximum likelihood estimates of the parameters are as follows.

Parameters	λ_1	λ_2	a_1	b_1	a_2	b_2	δ
Estimates	76.5643	44.7933	0.8302	1.1308	1.0898	1.0805	0.9531

From this table it can be seen that the estimator of the Lévy copula parameter $\widehat{\delta} = 0.9531$ for a Weibull-Clayton model is between the estimator only based on point processes

and the estimator only based on joint jumps of the process This is as expected.

7 Conclusion

We have suggested a maximum likelihood estimation procedure for a multivariate compound Poisson process, which guarantees that the estimated model is again a compound Poisson process. This is achieved on the basis of Sklar's theorem for Lévy copulas by a detailed analysis of the dependence structure. We have also suggested a new simulation algorithm for a multivariate compound Poisson process. A small simulation study has shown that the estimation procedure works well also for small sample sizes. For the Danish fire insurance data, after some explorative data analysis to find a convincing model, we have fitted a seven parameter compound Poisson process model. The use of a Lévy copula approach for the dependence modelling has proved extremely useful in this context.

References

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

[2] Bregman, Y. and Klüppelberg, C. (2005) Ruin estimation in multivariate models with Clayton dependence structure. *Scand. Act. J.* 2005 (6), 462-480.

[3] Böcker, K. and Klüppelberg, C. (2010) Multivariate models for operational risk. (PRMIA "2007 New Frontiers in Risk Management Award") *Quantitative Finance.* Doi:10.1080/14697680903358222.

[4] Böcker, K. and Klüppelberg, C. (2008) Modelling and measuring multivariate Operational Risk with Lévy copulas. *J. Operational Risk* **3**(2), 3-27.

[5] Böcker, K. and Klüppelberg, C. (2009) First order approximations to Operational Risk - dependence and consequences. To appear in: G.N. Gregoriou (ed.), *Operational Risk Toward Basel III, Best Practices and Issues in Modeling, Management and Regulation, pp. 219-245.* Wiley, New York.

[6] Cherubini, U., Luciano, E. and Vecchiato, W. (2004) *Copula Methods in Finance.* Wiley, London.

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

[8] Eder, I. and Klüppelberg, C. (2009) The first passage event for sums of dependent Lévy processes with applications to insurance risk. *Ann. Appl. Probab.* **19**(6), 2047-2079.

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

[10] Embrechts, P., Klüppelberg, C. and Mikosch, T. (1997) *Modelling Extremal Events for Insurance and Finance.* Springer, Berlin.

[11] Esmaeili, H. and Klüppelberg, K. (2010) Parametric estimation of a bivariate stable Lévy process. Submitted for publication.

[12] Joe, H. (1997) *Multivariate Models and Dependence Concepts.* Chapman& Hall, London.

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

[14] Nelsen, R. B. (1997) *An Introduction to Copulas.* Springer, New York.

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

[16] Sklar, A. (1959) Functions de répartition à n dimensions et leurs marges. *Publ. Inst. Statist. Univ. Paris* **8**, 229-231.

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