

# Capturing non-exchangeable dependence in multivariate loss processes with nested Archimedean Lévy copulas

Benjamin Avanzi<sup>a,b</sup>, Jamie Tao<sup>a</sup>, Bernard Wong<sup>\*,a</sup>, Xinda Yang<sup>a</sup>

<sup>a</sup>*School of Risk and Actuarial Studies, UNSW Australia Business School  
UNSW Sydney NSW 2052, Australia*

<sup>b</sup>*Département de Mathématiques et de Statistique, Université de Montréal  
Montréal QC H3T 1J4, Canada*

---

## Abstract

The class of spectrally positive Lévy processes is a frequent choice for modelling loss processes in areas such as insurance or operational risk. Dependence between such processes (for example, between different lines of business) can be modelled with Lévy copulas. This approach is a parsimonious, efficient, and flexible method which provides many of the advantages akin to distributional copulas for random variables.

Literature on Lévy copulas seems to have primarily focused on bivariate processes. When multivariate settings are considered, these usually exhibit an exchangeable dependence structure (whereby all subset of the processes have an identical marginal Lévy copula).

In reality, losses are not always associated in an identical way, and models allowing for non-exchangeable dependence patterns are needed. In this paper, we present an approach which enables the development of such models. Inspired by ideas and techniques from the distributional copula literature we investigate the procedure of *nesting* Archimedean Lévy copulas. We provide a detailed analysis of this construction, and derive conditions under which valid multivariate (nested) Lévy copulas are obtained. Our results are discussed and illustrated, notably with an example of model fitting to data.

*Key words:* Lévy copula, Exchangeability, Dependence, Nested copulas, Insurance claims  
JEL codes: G22, C10

---

Accepted for publication in *Annals of Actuarial Science*.

## 1. Introduction

In recent years, methods for modelling and managing the risk borne by major companies has become a prominent topic. This is especially the case for firms whose liabilities are only covered minimally by shareholders' participations, such as banks or insurance companies. Regulators, all around the world, are taking measures to encourage or even force such companies to build accurate models of the risks they face. This is not only to enable accurate assessment of their solvency and (similarly) determine appropriate levels of capitalisation, but also to encourage active management and mitigation of the firms' individual risks. Examples include Solvency II (in Europe) for insurance companies (with similar systems being developed around the world) and Basel III for banks. When considering scenarios for the financial situation of such financial institutions, it is not reasonable to assume that all the risks that are involved are independent, especially when considering those scenarios with adverse effects on the company's financial position (and which really are the ones of interest). Many circumstances will lead to strong dependence between risks or lines of business. Examples include the nature itself of the core business of the company (such as a major

---

\*Corresponding author. Tel.: +61 2 9385 2837; fax: +61 2 9385 1883.

*Email addresses:* b.avanzi@unsw.edu.au (Benjamin Avanzi), jtao@westpac.com.au (Jamie Tao), bernard.wong@unsw.edu.au (Bernard Wong), xinda.yang@unsw.edu.au (Xinda Yang)

hurricanes or earthquakes which would lead to substantial claims in different lines of insurance business), or major operational problems (such as an IT system failure or confidentiality breach in a bank).

The increased need for dependence modelling solutions led to many developments in the past few decades. Notably, copulas have been increasingly popular and are now well-known tools for modelling dependence between random variables (see, for instance, Frees and Valdez, 1998; Nelsen, 1999; McNeil et al., 2005; McNeil and Nešlehová, 2009). Copulas present many advantages, such as allowing for a very wide range of dependence structures and being formulated such that dependence structures and marginal distributions can be considered separately. The actual implementation of copula models can, however, present some challenges in practice. For instance, there is no way to determine the  $t$ -year copula that is equivalent to the  $s$ -year copula ( $s \neq t$ ) in general. This means that in the context introduced above, considering different time horizon scenarios requires copulas to be fit as many times as the number of time horizons that are considered. This can be a significant problem: for instance, 10 years of data will provide only 10 or less data points for time horizons of one year and above, which is arguably too small to ensure a reliable fit.

An alternative approach is to consider losses not only at the aggregate level for a given time horizon (the random variable approach), but dynamically as a loss process (the stochastic process approach). This makes full use of the individual data that is typically available nowadays in large companies (for instance, insurance companies might have hundreds of thousands of claims data points in a single year). With this alternative approach, 6 months of data may be enough to specify a whole stochastic process, which can subsequently be considered as a random variable at any needed time horizon (considering a random increment of the fitted process over that period of time, provided, of course, it is stationary). The challenge is then to specify and fit a dependence structure between multiple stochastic processes, a task for which copulas are inappropriate as they relate to random variables, not stochastic processes (note that we will now refer to them exclusively as ‘distributional copulas’ for that specific reason).

Fortunately, one can adopt a stochastic process approach and still retain most of the desirable features offered by the distributional copulas. If the modeller is prepared to restrict the set of candidate processes to that of Lévy processes—arguably a reasonable step as the main assumption is one of stationarity which we need anyway for further application of the model—then dependence can be introduced with the help of *Lévy copulas*, first introduced by Tankov (2003); see also Cont and Tankov (2004); Barndorff-Nielsen and Lindner (2004); Kallsen and Tankov (2006); Farkas et al. (2006); Barndorff-Nielsen and Lindner (2007). The mechanism by which Lévy copulas work is similar to that of distributional copulas: a function (the Lévy copula) ‘couples’ or ‘assembles’ the Lévy generator of the marginal processes (which fully specifies a Lévy process, akin to distribution functions which fully specify random variables) into a joint, multivariate Lévy generator (which then incorporates the desired dependence structure). Further details and benefits are discussed in Böcker and Klüppelberg (2008), or Avanzi et al. (2011).

The literature involving Lévy copulas is developing quickly and has recently been applied to insurance, operational risk and finance problems. Bäuerle and Blatter (2011) studied the optimal reinvestment and reinsurance problems for a multiline insurer, while Bregman and Klüppelberg (2005) used a Clayton Lévy copula to capture multivariate dependence and estimate the probability of ruin under this model; see also Eder and Klüppelberg (2009). In the area of operational risk modelling, applications of Lévy copulas between operational risk cells is discussed in Böcker and Klüppelberg (2008) and Böcker and Klüppelberg (2010); see also Biagini and Ulmer (2009). Esmaili and Klüppelberg (2010) and Avanzi et al. (2011) discussed and illustrated the fitting of Lévy copulas to data. Nonparametric inference is considered in Bücher and Vetter (2013).

The main concern of this paper is to facilitate the specification of *non-exchangeable* dependence structures. Such dependence arises when the associations between risks are not identical between *all* risks that are involved. The term ‘exchangeable’ comes from the fact that if all dependence relationships are identical, then the dependence structure is valid for any permutation of the risks without need of any adjustment (the ‘exchangeability’). Obviously, this is very unlikely to be observed in reality. Furthermore, while there exists some ‘elementary’ distributional (potentially non-exchangeable) copulas (such as the Gaussian copula) to tackle this problem in the realm of random variables, there is no such ‘elementary’ Lévy copula, and some simple method still needs to be developed. Indeed, most of the current Lévy copula literature either focuses on the bivariate case, or, when considering multivariate models, exchangeable dependence

structures. Inspired by recent developments with distributional copulas (specifically, nested structures and vines; see Joe, 1997; McNeil, 2008; McNeil and Nešlehová, 2009; Aas and Berg, 2009), [this paper develops ideas first presented in Tao (2011), who introduced the concept] of nested Archimedean Lévy copulas, which allows for partial non-exchangeability. We provide a detailed analysis of the nesting procedure and derive conditions under which valid multivariate Lévy copulas are obtained. Pair construction of Lévy copulas was recently discussed in Grothe and Nicklas (2013). [While this paper was under review, Grothe and Hofert (2015) was published. Their work also explored the concept of nested Archimedean Lévy copulas, and in particular developed an efficient sampling scheme under the assumption of completely monotone (combined) Archimedean Lévy generators.]

The structure of the paper is as follows. In the next section we review the relevant literature and summarise how Lévy copulas model dependence between multivariate spectrally positive Lévy processes, with particular emphasis on the class of Archimedean Lévy copulas. Section 3 contains the main theoretical contributions. Section 3 introduces the nesting procedure and presents the conditions that must be satisfied for resulting models to be valid. Trivariate and quadivariate nested Archimedean Lévy copulas are analysed in detail, and generalisation of our results to any dimension is also discussed. Simulation and parameter estimation are considered in Section 4, whereas further illustration using Danish fire data is presented in Section 5. To ensure readability, some formal definitions are given in Appendix A and detailed proofs are provided in Appendix B.

## 2. Dependence Modelling Using Lévy Copulas

This section is a brief, illustrated collection of the fundamental definitions and results about Lévy copulas. It provides an introduction to Lévy copulas and explains how they are able to capture multivariate dependence between loss processes. Note that loss processes have jumps which are positive, in contrast to general Lévy processes which can have jumps in the positive and/or negative domains. Hence, we restrict ourselves to *positive* Lévy copulas in this paper.

Note that all definitions, lemmas, theorems and examples mentioned in a given subsection are provided only at the end of that subsection to improve readability.

### 2.1. Introduction to positive Lévy copulas

Consider a multivariate Lévy process with positive jumps,  $\mathbf{S}(t)$ , representing a multivariate loss process (see for example Sato, 1999, for a formal definition). Each of the marginal processes of  $\mathbf{S}(t)$  represents the losses from one line of business, and is itself a univariate Lévy process with positive jumps. Common examples of such Lévy processes include the compound Poisson process with positive jumps, the Gamma process and the Inverse Gaussian process (see, for example, Cont and Tankov, 2004; Bertoin, 1998).

When modelling multivariate distributions, the “bottom-up” model-building approach enabled by Sklar’s theorem is intuitively appealing (see, for example, McNeil et al., 2005): assumptions are first made about the (typically well known) margins, which are then pulled together with an appropriate distributional copula. Fortunately, the same approach is possible with Lévy copulas. However, instead of joining up the probability distribution functions of (marginal) random variables, we will need to join up the *tail integrals* of the (marginal) Lévy processes. The tail integral of a multivariate Lévy process (see formal definitions in, e.g. Tankov, 2003; Cont and Tankov, 2004) with Lévy measure  $\nu(\cdots)$ ,

$$U(x_1, \dots, x_d) = \nu([x_1, \infty) \times \dots \times [x_d, \infty)), \quad x_1, \dots, x_d \in [0, \infty)^d, \quad (2.1)$$

has margins that are defined similarly to the margins of distribution functions. When all components of the tail integral are 0 except for the  $k$ -th element, one obtains the marginal tail integral

$$U_k(x_k) = U(0, \dots, 0, x_k, 0, \dots, 0). \quad (2.2)$$

The Lévy copula (see formal definitions in, e.g. Kallsen and Tankov, 2006; Bäuerle and Blatter, 2011) simply provides the link between the multivariate tail (2.1) and its margins (2.2), via Sklar’s theorem for Lévy copulas (see, e.g. Tankov, 2003; Cont and Tankov, 2004; Kallsen and Tankov, 2006).

## 2.2. Archimedean Lévy copulas

In a way similar to the construction of the Archimedean class of distribution copulas, it is possible to create a class of parametric Lévy copulas using some generator  $\psi$  (Definition 2.1); see Definition 2.2. Note that in this framework the generators  $\psi$  are defined slightly differently from that in the distribution copula framework; in particular, a convenient property is that the inverse  $\psi^{-1}$  always exists (because  $\psi$  is strictly decreasing and continuous).

It is important to note that additional conditions are required for an Archimedean Lévy generator to generate a valid Lévy copula. Lemma 2.1 below (from Bäuerle and Blatter, 2011, under different notations) provides a necessary and sufficient condition. Recall that a function with derivatives up to order  $d$  is called  $d$ -monotone if its derivatives alternate in sign, starting with a negative first derivative; a formal definition can be found in Appendix B, Definition B.1.

Table 1 summarises the Archimedean Lévy generators proposed in the literature, as well as their parameter domains. The Archimedean Lévy generators presented there are in general not the same as the distributional Archimedean generators with the same name. For instance, a Clayton Archimedean generator is a different function to the Clayton Archimedean Lévy generator. Furthermore, note that the domain and range of the Archimedean Lévy generators are both  $[0, \infty]$ , whereas the domain and range of Archimedean generators are  $[0, 1]$  and  $[0, \infty]$  respectively.

Name	$\psi(t)$	Parameter domain	Source
Clayton	$t^{-\delta_C}$	$\delta_C \in (0, \infty)$	Cont and Tankov (2004)
Gumbel	$[\log(1+t)]^{-\delta_G}$	$\delta_G \in (0, \infty)$	Bäuerle and Blatter (2011)
Ali-Mikhail-Haq	$\log\left(\frac{1-\delta_{AMH}}{t} + 1\right)$	$\delta_{AMH} \in [-1, 1)$	Bäuerle and Blatter (2011)
Archimedean type 1	$\frac{t}{1-e^{-\delta_{AI}t}}$	$\delta_{AI} \in (0, \infty)$	Avanzi et al. (2011)

Table 1: Archimedean Lévy generators

**Definition 2.1.** (Archimedean Lévy generator, Cont and Tankov, 2004; Bäuerle and Blatter, 2011) *An Archimedean Lévy generator is a strictly decreasing and continuous function  $\psi : [0, \infty] \rightarrow [0, \infty]$  which satisfies the conditions  $\psi(0) = \infty$  and  $\psi(\infty) = 0$ .*

**Definition 2.2.** (Archimedean Lévy copula, Cont and Tankov, 2004) *A  $d$ -dimensional Lévy copula  $\mathfrak{C}$  is called Archimedean if and only if it can be represented in the form*

$$\mathfrak{C}(u_1, \dots, u_d) = \psi^{-1}(\psi(u_1) + \dots + \psi(u_d)), \quad u_i \in [0, \infty]^d \quad i \in \{1, 2, \dots, d\} \quad (2.3)$$

for some Archimedean Lévy generator  $\psi$ .

**Lemma 2.1.** *The  $d$ -dimensional Archimedean construction in (2.3) yields a valid Lévy copula if and only if  $\psi^{-1}$  is  $d$ -monotone on  $(0, \infty)$ .*

## 2.3. Marginal Lévy copulas

When considering a multivariate stochastic process with Lévy copula dependence, association between any subset of the processes is described by (a reduced version of) that same Lévy copula. For example, consider a trivariate loss process  $\mathbf{S}(t) = (S_1(t), S_2(t), S_3(t))$  with trivariate Lévy copula  $\mathfrak{C}$ . The bivariate Lévy copula specifying the dependence between  $S_1(t)$  and  $S_2(t)$  is then given by

$$\mathfrak{C}(u_1, u_2, \infty) = \nu([x_1, \infty) \times [x_2, \infty) \times [0, \infty)) = U_{12}(x_1, x_2) = \mathfrak{C}_{12}(u_1, u_2). \quad (2.4)$$

This follows from Sklar's Theorem along with the definition and properties of the tail integral. Note that this is similar for the dependence structure between  $S_1(t)$  and  $S_3(t)$  and between  $S_2(t)$  and  $S_3(t)$  and for higher dimensions as well.

#### 2.4. Exchangeable Lévy copulas

Exchangeable Lévy copulas are defined in Definition 2.3 (see McNeil et al., 2005, for an analogous definition in the distributional copula literature). Such copulas are useful for modelling dependence between a set of loss processes, any pair of which exhibits very similar dependence characteristics. Arguably, this is unlikely to happen in our fields of interest, especially for high dimensional systems.

Example 2.1 provides an illustration of a trivariate process with exchangeable dependence structure. Importantly, the Clayton Lévy copula is not the only Archimedean Lévy copula that is exchangeable. In fact, a closer look at (2.3) informs us that *any* Archimedean Lévy copula, of any dimension, will always be exchangeable (as any permutation of a straight sum yields the same result).

**Definition 2.3.** (Exchangeable Lévy copula) *A Lévy copula is called an exchangeable Lévy copula if*

$$\mathfrak{C}(u_1, \dots, u_d) = \mathfrak{C}(v_1, \dots, v_d) \quad (2.5)$$

for any permutation  $(v_1, \dots, v_d)$  of  $(u_1, \dots, u_d)$ .

**Example 2.1.** *For a trivariate compound Poisson process  $\{S_1(t), S_2(t), S_3(t)\}$ , with trivariate Lévy copula  $\mathfrak{C}$ , we want to consider the dependence in frequency and severity between  $S_1(t)$  and  $S_2(t)$ . The joint tail integral  $U_{12}(x_1, x_2)$  measures those jumps which are common to  $S_1(t)$  and  $S_2(t)$  including jumps which are common to all three processes. This is different to those jumps described by  $U_{12}^{\parallel}(x_1, x_2) = \nu([x_1, \infty) \times [x_2, \infty) \times \{0\})$ , which considers common jumps between the two processes only. Expressing  $U_{12}(x_1, x_2)$  in terms of a Lévy measure (see (2.1)) yields*

$$U_{12}(x_1, x_2) = \nu([x_1, \infty) \times [x_2, \infty) \times [0, \infty)). \quad (2.6)$$

Using Sklar's theorem for Lévy copulas and the fact that  $U_i(0) = \infty$  for  $i \in \{1, 2, 3\}$  yields

$$U_{12}(x_1, x_2) = \mathfrak{C}(U_1(x_1), U_2(x_2), \infty). \quad (2.7)$$

Hence, the marginal Lévy copula  $\mathfrak{C}_{12.}$ , specifying dependence between  $S_1(t)$  and  $S_2(t)$ , can be expressed in terms of the trivariate Lévy copula  $\mathfrak{C}$ ,

$$\mathfrak{C}_{12.}(u_1, u_2) = \mathfrak{C}(u_1, u_2, \infty). \quad (2.8)$$

Similar expressions for  $\mathfrak{C}_{1.3}$  and  $\mathfrak{C}_{.23}$ , specifying the dependence structure between  $S_1(t)$  and  $S_3(t)$  and between  $S_2(t)$  and  $S_3(t)$  respectively, can be derived analogously.

For the trivariate Clayton Lévy copula,

$$\mathfrak{C}(u_1, u_2, u_3) = (u_1^{-\delta} + u_2^{-\delta} + u_3^{-\delta})^{-\frac{1}{\delta}}, \quad (2.9)$$

the marginal bivariate Lévy copulas are given by

$$\begin{cases} \mathfrak{C}_{12.}(u_1, u_2) = (u_1^{-\delta} + u_2^{-\delta})^{-\frac{1}{\delta}}, \\ \mathfrak{C}_{1.3}(u_1, u_3) = (u_1^{-\delta} + u_3^{-\delta})^{-\frac{1}{\delta}}, \\ \mathfrak{C}_{.23}(u_2, u_3) = (u_2^{-\delta} + u_3^{-\delta})^{-\frac{1}{\delta}}. \end{cases} \quad (2.10)$$

Note that the marginal bivariate Lévy copulas above, providing information regarding the dependence between different pairwise combinations, are identical. Hence, the Lévy copula (2.9) is exchangeable.

### 3. Nested Archimedean Lévy Copulas

An important limitation of the multivariate Archimedean Lévy copulas presented in the previous section is their exchangeability. For a group of loss processes with homogeneous dependence structure, exchangeable Lévy copulas are adequate. However, when more complex dependence structures arise, such as those with varying pairwise structures, exchangeable Lévy copulas can become problematic.

In this section, we will show that a *nesting* approach (such as in the distributional literature, see for instance Kurowicka and Joe, 2011) can lead to valid multivariate, non-exchangeable Lévy copula models, under certain conditions. Namely, we develop and introduce nested Archimedean Lévy copulas. Such an approach allows for non-exchangeability while still retaining the tractability of the Archimedean construction. We highlight the key ideas and present the technical conditions required for a nested Archimedean Lévy copula to be valid. For readability, proofs are included in Appendix B.

### 3.1. Construction of nested Archimedean Lévy copulas

In this section we present three theorems which describe the conditions under which the nesting procedure produces valid Lévy copulas. These three results correspond to the trivariate, quadivariate fully nested, and quadivariate partially nested cases, respectively. As discussed in Section 3.1.2, the move from two to three and then four dimensions is easily generalised to higher dimensions.

#### 3.1.1. Trivariate nested Archimedean Lévy copula

The nesting procedure is quite intuitive. The idea is to associate models in a pairwise fashion in a given sequential order. For instance, processes 1 and 2 can be first associated, and the result subsequently associated to the third process. This is illustrated in Figure 1, where the tail integrals  $u_1$  and  $u_2$  are first combined through the bivariate Lévy copula  $\mathfrak{C}_{11}$ . The next step is to then combine the resulting model ( $\mathfrak{C}_{11}$ ) with  $u_3$  with another bivariate Lévy copula  $\mathfrak{C}_{21}$ . A direct remark is that one has three choices of pairs to start with. As we will illustrate later, this choice may be restricted if one wants the resulting model to be valid.

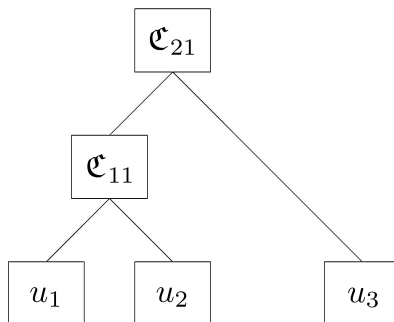


Figure 1: 3-dimensional fully nested Archimedean Lévy copula

Of course, if  $\mathfrak{C}_{11}$  and  $\mathfrak{C}_{21}$  are not the same Lévy copulas, then the dependence between the three processes is not exchangeable. In fact, in the example of Figure 1, processes 1 and 2 are associated in a certain way (according to  $\mathfrak{C}_{11}$ ), which is different from that associates process 3 with both processes 1 and 2 (according to  $\mathfrak{C}_{21}$ ). In other words, the loss processes corresponding to the tail integrals  $u_i$  for  $i = 1, 2$  will have marginal Lévy copula corresponding to the Archimedean Lévy copula generator  $\psi_{11}$ , while the loss process between the pairs  $(u_1, u_3)$  and  $(u_2, u_3)$  will both possess the same marginal Lévy copula - the one that corresponds to the Archimedean Lévy copula generator  $\psi_{21}$ .

Theorem 3.1 spells out the structure of the trivariate Lévy copula yielded by the nesting approach described above, and sets sufficient conditions for this structure to be valid. To understand the conditions, recall that a function is in class  $C^d$  if it is continuously differentiable to order  $d$ . Furthermore, a  $C^d$  function is  $d$ -monotone if its derivatives alternate in sign, up to the order  $d$  (see Definition B.1 in Appendix B for a formal definition).

**Theorem 3.1.** (Trivariate fully nested Archimedean Lévy copula) *A trivariate fully nested Archimedean Lévy copula,*

$$\mathfrak{C}(u_1, u_2, u_3) = \psi_{21}^{-1}(\psi_{21} \circ \psi_{11}^{-1}(\psi_{11}(u_1) + \psi_{11}(u_2)) + \psi_{21}(u_3)) \quad (3.1)$$

with Archimedean Lévy generators  $\psi_{11}$  and  $\psi_{21}$ , where

1.  $\psi_{21}^{-1} \in C^3$ , and is 3-monotone on  $(0, \infty)$ , and
2.  $\psi_{21} \circ \psi_{11}^{-1} \in C^2$ , and has a derivative which is 1-monotone on  $(0, \infty)$ , and
3.  $\psi_{11} \in C^1$  and  $\psi_{21} \in C^1$ ,

is a valid Lévy copula.

A full proof is provided in Appendix B.1.

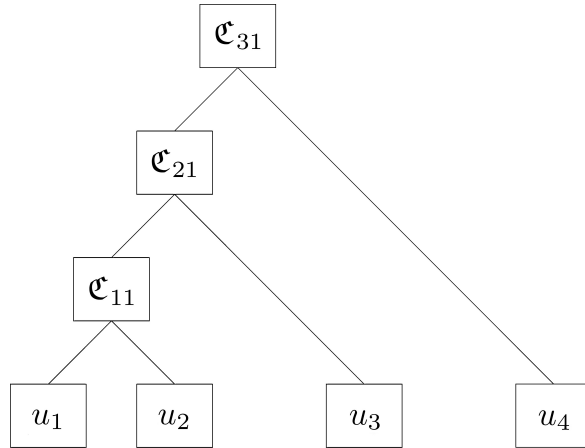


Figure 2: 4-dimensional fully nested Archimedean Lévy copula

### 3.1.2. Quadivariate and multivariate nested Archimedean Lévy copulas

With dimensions higher than 3, there is not only a choice of which processes to nest first, but also of how to nest them (different structures are available, even when the order of the pairing is determined). We distinguish two alternative approaches:

1. One can follow the nesting procedure as described in the previous section and perform the additional step by nesting the trivariate process with a fourth one—this leads to a *fully nested* structure. This is illustrated in Figure 2.
2. Alternatively, one can pair processes in a non sequential way, and nest the resulting sub-nested models—this leads to a *partially nested* structure. This is illustrated in Figure 3.

These constructions will *not* always be valid. In particular, their validity will depend on the monotonicity properties of the composite Lévy generator stemming from the nesting procedure. In the remainder of this section, we illustrate and examine these two approaches in the quadivariate case. Note that the constructions and proofs can naturally be extended to higher dimensions.

Theorem 3.2 examines the case of the fully nested quadivariate construction depicted in Figure 2. Under this construction, the loss processes corresponding to the tail integrals  $u_i$  for  $i = 1, 2$  will have marginal Lévy copula corresponding to the Archimedean Lévy copula generator  $\psi_{11}$ , while the loss process between the pairs  $(u_1, u_3)$  and  $(u_2, u_3)$  will both possess the same marginal Lévy copula - and corresponds to the Archimedean Lévy copula generator  $\psi_{21}$ .

Theorem 3.3 is concerned with the partially nested structure in Figure 3. There, the tail integrals of  $u_1$  and  $u_2$  are first coupled through the bivariate Lévy copula  $\mathfrak{C}_{11}$  and  $u_3$  and  $u_4$  through  $\mathfrak{C}_{12}$ . The construction subsequently joins together  $\mathfrak{C}_{11}$  and  $\mathfrak{C}_{12}$  with  $\mathfrak{C}_{21}$  to complete the construction.

Which of the fully or partially nested approaches should be used depends on the characteristics of the data. Specifically, pairwise association exhibited in the data should be examined with consideration of (i) which of the pairwise associations are significantly different or similar, but also of (ii) which type of nesting, once fitted, will yield a valid model. The former is important because of the partial exchangeability of the structures described above. The latter will become clearer in the next section, where the conditions for

validity (found in the theorems) are translated into interpretable conditions for some exemplary models. Both considerations are further discussed and illustrated in Section 5.

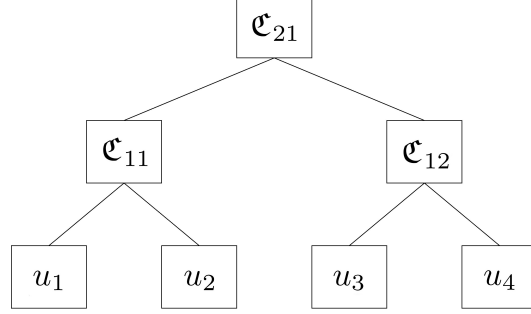


Figure 3: 4-dimensional partially nested Archimedean Lévy copula

**Theorem 3.2.** (Quadivariate fully nested Archimedean Lévy copula) *A quadivariate fully nested Archimedean Lévy copula,*

$$\mathfrak{C}(u_1, u_2, u_3, u_4) = \psi_{31}^{-1}(\psi_{31} \circ \psi_{21}^{-1}(\psi_{21} \circ \psi_{11}^{-1}(\psi_{11}(u_1) + \psi_{11}(u_2)) + \psi_{21}(u_3)) + \psi_{31}(u_4)) \quad (3.2)$$

with Archimedean Lévy generators  $\psi_{11}, \psi_{21}$  and  $\psi_{31}$ , where

1.  $\psi_{31}^{-1} \in C^4$ , and is 4-monotone on  $(0, \infty)$ , and
2.  $\psi_{31} \circ \psi_{21}^{-1} \in C^3$ , and has a derivative which is 2-monotone on  $(0, \infty)$ , and
3.  $\psi_{21} \circ \psi_{11}^{-1} \in C^2$ , and has a derivative which is 1-monotone on  $(0, \infty)$ , and
4.  $\psi_{11} \in C^1$ ,  $\psi_{21} \in C^1$  and  $\psi_{31} \in C^1$ ,

is a valid Lévy copula.

Theorem 3.2 adds a dimension to that of Theorem 3.1 and hence has a similar proof. However, a sketch of proof is still provided in Appendix B.2 in order to illustrate how to generalise these results to any dimension.

**Theorem 3.3.** (Quadivariate partially nested Archimedean Lévy copula) *A quadivariate partially nested Archimedean Lévy copula,*

$$\mathfrak{C}(u_1, u_2, u_3, u_4) = \psi_{21}^{-1}(\psi_{21} \circ \psi_{11}^{-1}(\psi_{11}(u_1) + \psi_{11}(u_2)) + \psi_{21} \circ \psi_{12}^{-1}(\psi_{12}(u_3) + \psi_{12}(u_4))) \quad (3.3)$$

with Archimedean Lévy generators  $\psi_{11}, \psi_{12}$  and  $\psi_{21}$ , where

1.  $\psi_{21}^{-1} \in C^4$ , and is 4-monotone on  $(0, \infty)$ , and
2.  $\psi_{21} \circ \psi_{11}^{-1} \in C^2$ , and has a derivative which is 1-monotone on  $(0, \infty)$ , and
3.  $\psi_{21} \circ \psi_{12}^{-1} \in C^2$ , and has a derivative which is 1-monotone on  $(0, \infty)$ , and
4.  $\psi_{11} \in C^1$  and  $\psi_{12} \in C^1$ ,

is a valid Lévy copula.

The approach to proving Theorem 3.3 is similar to that for Theorem 3.1. A full proof of Theorem 3.3 can be found in Appendix B.3.

As in the trivariate case, one can study the marginal Lévy copulas implied by a quadivariate Lévy copula. This is illustrated in Figure 4, which displays the Lévy generators corresponding to the marginal couples in both cases of full nesting (on the left hand side) and partial nesting (on the right hand side).

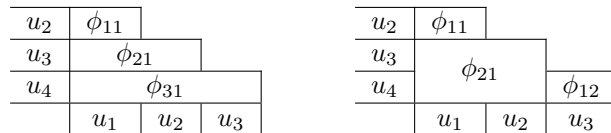


Figure 4: Bivariate Lévy copulas implied by a quadivariate fully and partially nested Lévy copula, on the left and right hand sides, respectively.

### 3.2. Examples of nested models

By examining carefully the conditions that the various Lévy generators must satisfy, the approach of nesting a construction to capture non-exchangeability can be extended to any dimension  $d$ . However, the number of ways that a nested Archimedean Lévy copula can be constructed will increase with increasing dimensions. For instance, the trivariate case in Section 3.1.1 has one distinct construction and the bivariate case in Section 3.1.2 has two distinct constructions.

Notwithstanding, validity of the Archimedean Lévy generators needs to be checked on a case by case basis. We illustrate below the effect of these conditions with some of the Lévy copulas of Table 2.

#### 3.2.1. Trivariate fully nested Clayton

Selecting a Clayton Lévy generator for both levels of nesting results in the trivariate fully nested Clayton Lévy copula leads to

$$\mathfrak{C}(u_1, u_2, u_3) = ((u_1^{-\delta_{C,11}} + u_2^{-\delta_{C,11}})^{\frac{\delta_{C,21}}{\delta_{C,11}}} + u_3^{-\delta_{C,21}})^{-\frac{1}{\delta_{C,21}}}, \quad (3.4)$$

where  $\delta_{C,11}$  and  $\delta_{C,21}$  are the parameters of the two (bivariate) Clayton Lévy copulas used in the construction.

Restrictions apply on the choice of parameters  $\delta_{C,11}$  and  $\delta_{C,21}$ . The composite Lévy generator of the trivariate nested Clayton Lévy copula,  $f(t) = \psi_{21} \circ \psi_{11}^{-1}(t)$ , takes the form

$$f(t) = t^{\frac{\delta_{C,21}}{\delta_{C,11}}}, \quad t > 0. \quad (3.5)$$

For the trivariate case, a valid Lévy copula results if this composite function has a derivative which is 1-differentiable and 1-monotone. The first derivative of this function is

$$f'(t) = \frac{\delta_{C,21}}{\delta_{C,11}} t^{\frac{\delta_{C,21}}{\delta_{C,11}} - 1}, \quad t > 0. \quad (3.6)$$

The first derivative in (3.6) is always non-negative. However, for this derivative to be 1-differentiable and 1-monotone, the power of  $t$  must be non-positive, so the second derivative is non-positive. Hence,

$$\frac{\delta_{C,21}}{\delta_{C,11}} - 1 \leq 0 \iff \delta_{C,21} \leq \delta_{C,11}, \quad (3.7)$$

which leads to the restriction

$$\delta_{C,11} \geq \delta_{C,21} > 0. \quad (3.8)$$

**Remark 3.1.** *In fact, when  $\delta_{C,11} \geq \delta_{C,21}$ , the trivariate fully nested Clayton composite generator has a completely monotone derivative. This is because the power of  $t$  will always be negative after the first derivative, causing each successive differentiation to alternate in sign.*

**Remark 3.2.** *In the case where the two parameters are equal, that is  $\delta_{C,11} = \delta_{C,21}$ , we retrieve the exchangeable trivariate Clayton Lévy copula.*

An important implication of this restriction when building a model with Clayton trivariate nested Lévy copula dependence is that the Lévy processes with a dependence captured by a larger parameter must be coupled first, and the component with dependence captured by a smaller parameter must be on the outer level of the nesting. This is similar to what was mentioned for the distributional copula case in Aas and Berg (2009) and Hofert (2008).

**Example 3.1.** *For the Clayton-Clayton fully nested Lévy copula as applied to a trivariate compound Poisson process, the implied survival copula can be derived for the jumps common to all components. If a form exists for the implied distributional copula, techniques in copula literature, which are much more developed than Lévy copulas, can be used for modelling and simulation of Lévy processes.*

Let  $\lambda_i$  denote the marginal intensity for process  $i$ , and  $\lambda_{123}^{\parallel}$  denote the intensity for jumps occurring to all three processes simultaneously. The Clayton-Clayton fully nested Lévy copula has an implied distributional survival copula  $\hat{C}_N$  given by

$$\hat{C}_N(a_1, a_2, a_3) = \left\{ \left( \left[ a_1^{-\delta_2} - \left( \frac{\lambda_3}{\lambda_{123}^{\parallel}} \right)^{-\delta_2} \right]^{\frac{\delta_1}{\delta_2}} + \left[ a_2^{-\delta_2} - \left( \frac{\lambda_3}{\lambda_{123}^{\parallel}} \right)^{-\delta_2} \right]^{\frac{\delta_1}{\delta_2}} - \left[ \left( \frac{\lambda_1}{\lambda_{123}^{\parallel}} \right)^{-\delta_1} + \left( \frac{\lambda_2}{\lambda_{123}^{\parallel}} \right)^{-\delta_1} \right]^{\frac{\delta_2}{\delta_1}} \right)^{\frac{\delta_2}{\delta_1}} + a_3^{-\delta_2} - \left[ \left( \frac{\lambda_1}{\lambda_{123}^{\parallel}} \right)^{-\delta_1} + \left( \frac{\lambda_2}{\lambda_{123}^{\parallel}} \right)^{-\delta_1} \right]^{\frac{\delta_2}{\delta_1}} \right\}^{-\frac{1}{\delta_2}}, \quad (3.9)$$

where  $a_1, a_2$  and  $a_3$  are marginal survival functions; see Appendix B.4. This is in contrast to the trivariate (plain vanilla) exchangeable case, where the survival copula  $\hat{C}_E$  is

$$\hat{C}_E(a_1, a_2, a_3) = (a_1^{-\delta} + a_2^{-\delta} + a_3^{-\delta} - 2)^{-\frac{1}{\delta}}; \quad (3.10)$$

see Cassar (2010, pages 90-91). Comparison of (3.9) with (3.10) also illustrates how the nesting procedure affects dependence not only in frequency, but also in severity.

### 3.2.2. Other same family nested models

Selecting the same Archimedean Lévy generator for both levels of nesting allows us to construct alternative trivariate fully nested Lévy copulas, such as the trivariate fully nested Gumbel Lévy copula. With a similar approach to the Clayton trivariate fully nested Lévy copula in Subsubsection 3.2.1, the associated restrictions for an Archimedean Lévy generator can be found by analysing the derivative properties of the composite Lévy generators. Table 2 summarises the parameter (sufficient) restrictions for the various trivariate fully nested Lévy copulas.

Name	Restriction
Gumbel	$0 < \delta_{G,21} \leq \delta_{G,11}$
Ali-Mikhail-Haq	$-1 \leq \delta_{AMH,21} \leq \delta_{AMH,11} < 1$
Archimedean type I	$0 < \delta_{AI,21} \leq \delta_{AI,11}$

Table 2: Restrictions for various trivariate fully nested Lévy copulas

**Remark 3.3.** Table 2 raises some questions. Is there usually a clear relationship between the level the different parameters  $\delta$  and the level of dependence induced by the associated model? Unfortunately the answer is not obvious as the concept of dependence between stochastic processes requires some care. Consider, for instance, the case of compound Poisson processes. There, both dependence in frequency and severity have to be considered. Recall that there is a direct relationship between the Lévy copula and the amount of ‘common jumps’ between processes: the larger the Lévy copula, the higher the intensity of common jumps (the Lévy copula function maps marginal intensities into a common jump intensity). In other words, the larger the Lévy copula, the stronger the dependence in frequency. For instance, if we consider bivariate Archimedean Lévy copulas (such as in Table 2), and examine

$$\frac{d}{d\delta} \mathfrak{C}(u_1, u_2) = \frac{d}{d\delta} \psi^{-1}(\psi(u_1) + \psi(u_2)) = \frac{\psi'_\delta(u_1) + \psi'_\delta(u_2)}{\psi'_\delta(\psi^{-1}(\psi(u_1) + \psi(u_2)))}, \quad (3.11)$$

where

$$\psi'_\delta(\cdot) = \frac{d}{d\delta} \psi(\cdot), \quad (3.12)$$

we see that higher  $\delta$ 's will always lead to a higher dependence level in frequency for Archimedean Lévy copulas, as long as the generator  $\psi$  is monotone in  $\delta$ . General conclusions about dependence in severity

are more difficult to draw. In some cases, such as in (3.10), the relationship is clear and well known from the distributional copula literature. Other cases may not be as clear, and may need to be checked on a case-by-case basis.

### 3.2.3. Mixed family nested models

Mixing generators for a nested Archimedean Lévy copula is possible, such as using the Clayton for the inner level of nesting and the Gumbel for the outer level of nesting. However, more complicated restrictions apply to the choice of parameters.

As an example, consider the trivariate fully nested Clayton–Gumbel Lévy copula, where  $\psi_{11}$  is a Clayton Lévy generator and  $\psi_{21}$  is a Gumbel Lévy generator. It takes the form

$$\mathfrak{C}(u_1, u_2, u_3) = \exp \left[ \left( [\log(1 + (u_1^{-\delta_{C,11}} + u_2^{-\delta_{C,11}})^{-\frac{1}{\delta_{C,11}}})]^{-\delta_{G,21}} + [\log(1 + u_3)]^{-\delta_{G,21}} \right)^{-\frac{1}{\delta_{G,21}}} \right] - 1, \quad (3.13)$$

where  $\delta_{C,11}$  is the parameter of the Clayton Lévy generator and  $\delta_{G,21}$  is the parameter of the Gumbel Lévy generator, which are subject to restrictions we will now explore.

The composite generator,  $f(t) = \psi_{21} \circ \psi_{11}^{-1}(t)$  of the Clayton–Gumbel trivariate fully nested Lévy copula is

$$f(t) = \left[ \log \left( 1 + t^{-\frac{1}{\delta_{C,11}}} \right) \right]^{-\delta_{G,21}}, \quad t \geq 0. \quad (3.14)$$

Now the first and second derivatives of this function are given as

$$f'(t) = \frac{\delta_{G,21}}{\delta_{C,11}} \left[ \log \left( 1 + t^{-\frac{1}{\delta_{C,11}}} \right) \right]^{-\delta_{G,21}-1} \frac{1}{t + t^{\frac{1}{\delta_{C,11}}+1}} \geq 0, \quad \text{and} \quad (3.15)$$

$$\begin{aligned} f''(t) &= \frac{\delta_{G,21}}{\delta_{C,11}} \left[ \log \left( 1 + t^{-\frac{1}{\delta_{C,11}}} \right) \right]^{-\delta_{G,21}-2} \frac{1}{\left( t + t^{\frac{1}{\delta_{C,11}}+1} \right)^2} \\ &\quad \times \left[ \frac{\delta_{G,21} + 1}{\delta_{C,11}} - \left[ \log \left( 1 + t^{-\frac{1}{\delta_{C,11}}} \right) \right] \left( 1 + \left( \frac{1}{\delta_{C,11}} + 1 \right) t^{\frac{1}{\delta_{C,11}}} \right) \right]. \end{aligned} \quad (3.16)$$

The first derivative (3.15) is always non-negative. The second derivative (3.16) needs to be non-positive for the composite generator condition to be satisfied. This will be the case if and only if

$$\delta_{G,21} \leq \min_t \left\{ \delta_{C,11} \left[ \log \left( 1 + t^{-\frac{1}{\delta_{C,11}}} \right) \right] \left( 1 + \left( \frac{1}{\delta_{C,11}} + 1 \right) t^{\frac{1}{\delta_{C,11}}} \right) - 1 \right\}. \quad (3.17)$$

This inequality must hold for *all* values of  $t \geq 0$ , so the minimum of the term in the curly braces with respect to  $t$  is the upper bound for  $\delta_{G,21}$ . Unlike same family generators, there is no simple relationship between  $\delta_{G,21}$  and  $\delta_{C,11}$ . Accordingly, this is not a straightforward restriction and as a result, the Clayton–Gumbel construction may yield a valid Lévy copula with limited choices of parameters.

**Remark 3.4.** *Mixing generators in Lévy copula constructions of higher dimensions will be even more restricted. In most cases, checking the validity of a general  $d$ -dimensional multivariate constructions will involve the analysis of derivatives beyond  $f''(t)$ . As can be seen with Clayton–Gumbel trivariate nested Lévy copula, the second derivative of the composite Lévy generator is already a complicated expression. Differentiating this further will become increasingly tedious.*

### 3.2.4. Quadivariate partially nested Lévy copulas

As an example, consider the Clayton quadivariate partially nested Lévy copula. This Lévy copula is constructed by selecting a Clayton Lévy generator for  $\psi_{11}, \psi_{12}$  and  $\psi_{21}$  in (3.3) and can be expressed as

$$\mathfrak{C}(u_1, u_2, u_3, u_4) = \left( (u_1^{-\delta_{C,11}} + u_2^{-\delta_{C,11}})^{\frac{\delta_{C,21}}{\delta_{C,11}}} + (u_3^{-\delta_{C,12}} + u_4^{-\delta_{C,12}})^{\frac{\delta_{C,21}}{\delta_{C,12}}} \right)^{-\frac{1}{\delta_{C,21}}}, \quad (3.18)$$

where  $\delta_{C,11}$ ,  $\delta_{C,12}$  and  $\delta_{C,21}$  are parameters.

A valid quadivariate partially nested Lévy copula requires that both composite generators,  $\psi_{21} \circ \psi_{11}^{-1}$  and  $\psi_{21} \circ \psi_{11}^{-1}$  have derivatives which are 1-differentiable and 1-monotone. Drawing from results obtained in Section 3.2.1 for the Clayton trivariate nested Lévy copula, the parameter restrictions are

$$\delta_{C,21} \leq \delta_{C,11} \quad \text{and} \quad (3.19)$$

$$\delta_{C,21} \leq \delta_{C,12}. \quad (3.20)$$

#### 4. Simulation and Parameter Estimation

Simulation schemes for univariate and bivariate Lévy processes (with Lévy copulas) were developed in Cont and Tankov (2004) and Esmaeili and Klüppelberg (2010). In this section we use the algorithm of Tao (2011, Appendix B) for trivariate nested Archimedean Lévy copulas and perform a similar analysis to that of Esmaeili and Klüppelberg (2011). This allows us to draw some conclusions on the performance of maximum likelihood estimation of model parameters, using simulated sample paths of a trivariate compound Poisson process.

Let  $(S_1, S_2, S_3)$  be a trivariate compound Poisson process with marginal intensities  $\lambda_1, \lambda_2, \lambda_3 > 0$ , exponentially distributed marginal severities given by  $\bar{F}_i(x_i) = e^{-\theta_i x}$ ,  $x > 0$  for  $i \in \{1, 2, 3\}$ , and a structure specified by a Clayton trivariate fully nested Lévy copula  $\mathfrak{C}$  with parameters  $\delta_{C,11} \geq \delta_{C,21} > 0$ .

We simulated 100 paths of a trivariate exponential fully nested Clayton model, with the parameters  $\lambda_1 = 100$ ,  $\lambda_2 = 80$ ,  $\lambda_3 = 50$ ,  $\theta_1 = 1$ ,  $\theta_2 = 2$ ,  $\theta_3 = 3$  and different  $\delta$  over the time interval  $[0, 1]$ .

The results of the maximum likelihood estimates based on the simulated paths are summarised in the tables below.

Value	$\widehat{\text{Mean}}$	$\widehat{\text{MSE}}$	$\widehat{\text{MAE}}$	$\widehat{\text{MRB}}$
$\delta_{C,11} = 2$	2.0112 (0.2082)	0.0430 (0.0552)	0.1704 (0.1189)	0.0056 (0.1041)
$\delta_{C,21} = 1$	1.0056 (0.1179)	0.0138 (0.0185)	0.0978 (0.0654)	0.0056 (0.1179)
$\delta_{C,11} = 3$	3.0178 (0.2661)	0.0704 (0.0996)	0.2127 (0.1594)	0.0059 (0.0887)
$\delta_{C,21} = 1$	0.9981 (0.1000)	0.0099 (0.0132)	0.0799 (0.0596)	-0.0019 (0.1000)
$\delta_{C,11} = 5$	4.9861 (0.5332)	0.2817 (0.4497)	0.4204 (0.3255)	-0.0028 (0.1066)
$\delta_{C,21} = 1$	1.0119 (0.1234)	0.0152 (0.0212)	0.0973 (0.0763)	0.0119 (0.1234)
$\delta_{C,11} = 3$	3.0442 (0.2629)	0.0704 (0.0921)	0.2137 (0.1579)	0.0147 (0.0876)
$\delta_{C,21} = 2$	2.0016 (0.1950)	0.0377 (0.0530)	0.1535 (0.1193)	0.0008 (0.0975)

Table 3:  $\widehat{\text{Mean}}$ ,  $\widehat{\text{MSE}}$ ,  $\widehat{\text{MAE}}$  and  $\widehat{\text{MRB}}$  (and standard deviations in brackets) for 100 MLEs of the Lévy copula parameters of the trivariate exponential fully nested Clayton model. Each estimate is calculated from an observed simulated sample path with  $\lambda_1 = 100$ ,  $\lambda_2 = 80$ ,  $\lambda_3 = 50$ ,  $\theta_1 = 1$ ,  $\theta_2 = 2$ ,  $\theta_3 = 3$  (assumed to be known) and unknown Lévy copula parameters  $\delta_{C,i1}$  for  $i \in \{1, 2\}$ .

The results in Table 3 assume that the marginal parameters are known, but that the parameters of the Lévy copula are unknown. The table shows the means of maximum likelihood estimates  $\hat{\delta}$  of  $\delta$ , as well as

$$\widehat{\text{MSE}} = \frac{1}{100} \sum_{i=1}^{100} (\hat{\delta}_i - \delta)^2; \quad \widehat{\text{MAE}} = \frac{1}{100} \sum_{i=1}^{100} |\hat{\delta}_i - \delta|; \quad \text{and} \quad \widehat{\text{MRB}} = \frac{1}{100} \sum_{i=1}^{100} \frac{\hat{\delta}_i - \delta}{\delta},$$

the mean square error, mean absolute error, and mean relative bias, respectively. All estimations achieve good results, with the maximum MRB being 1.5% approximately.

	$\hat{\lambda}_1$	$\hat{\lambda}_2$	$\hat{\lambda}_3$	$\hat{\theta}_1$	$\hat{\theta}_2$	$\hat{\theta}_3$	$\hat{\delta}_{C,11}$	$\hat{\delta}_{C,21}$
Values	100	80	50	1	2	3	2	1
$\widehat{\text{Mean}}$	96.8374	78.1351	48.9162	1.0147	2.0495	3.0703	1.9983	0.9978
$\widehat{\text{MSE}}$	9.9698	8.5595	6.7138	0.0903	0.2055	0.4301	0.2632	0.1353
$\widehat{\text{MAE}}$	108.4046	76.0100	45.7984	0.0083	0.0442	0.1881	0.0686	0.0181
$\widehat{\text{MRB}}$	123.1996	97.3085	69.0650	0.0142	0.0642	0.3542	0.0863	0.0227
	8.6831	7.1180	5.2591	0.0702	0.1665	0.3260	0.2189	0.1108
	5.7743	5.0597	4.2805	0.0584	0.1292	0.2874	0.1445	0.0769
	-0.0316	-0.0233	-0.0217	0.0147	0.0247	0.0234	-0.0009	-0.0022
	0.0997	0.1070	0.1343	0.0903	0.1027	0.1434	0.1316	0.1353

Table 4: Estimated  $\widehat{\text{Mean}}$ ,  $\widehat{\text{MSE}}$ ,  $\widehat{\text{MAE}}$  and  $\widehat{\text{MRB}}$  (and standard deviations in brackets) of 100 maximum likelihood estimates of a trivariate exponential fully nested Clayton model.

The result in Table 4 is similar to Table 3, but assumes all parameters are unknown. Figure 5 displays the relative biases in Table 4. The mean relative bias remains around 0 for all parameters in both Table 3 and Table 4 indicating a reasonably accurate estimation.

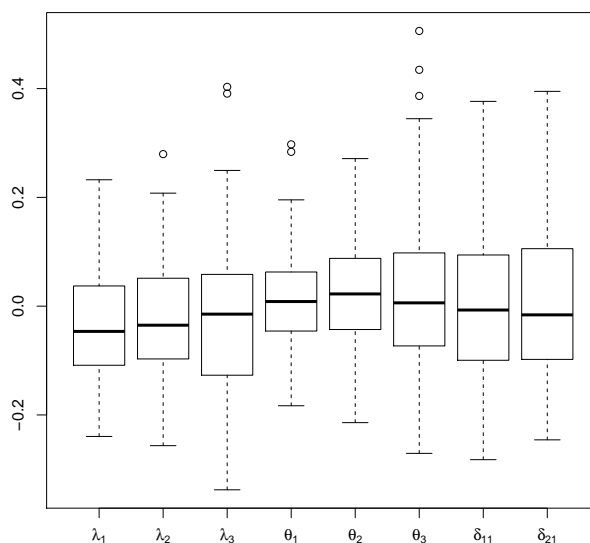


Figure 5: Boxplots of relative bias for the estimates of the trivariate exponential fully nested Clayton model with parameter values in Table 4.

**Remark 4.1.** *The simulation algorithm that has been used becomes exponentially untractable for higher dimensions. The simulation algorithm developed in Grothe and Hofert (2015) addresses this issue, and should be preferred for higher dimensions. This dimension problem carries over to the estimation procedure. To the extent of our knowledge, this has not been addressed in the literature yet, and is left for future research.*

## 5. Illustration with Danish Fire Data

In this section, we illustrate the fitting of a non-exchangeable dependence structure using the very well known Danish fire data set. Some short notes on the actual programming involved in this section are provided in Appendix C.

### 5.1. The Danish data set

The Danish fire insurance data is composed of insurance claims from the beginning of 1980 until the end of 1990, which have been adjusted to 1985 values for inflation. The same data set was considered in Esmaili and Klüppelberg (2010) in the bivariate setting by dropping the Profits line of insurance. Nadarajah and Bakar (2014) use composite models to model the payments between two lines. The Danish data is a trivariate data set of 2167 fire losses expressed in millions of Danish Kroner, in the lines of Building, Contents and Profits. Claims are only recorded where the total loss is larger than a million. That is, when the combined loss of Building, Contents and Profits exceeds one million, it is recorded.

As a result of the truncated nature of the data, an artificial negative dependence is introduced between Buildings and Contents. To ensure that this artificial dependence does not interfere with the fitting procedure, we have limited our analysis to claims which are greater than one million Danish Kroner in all three components. In this case, the size of a claim in a given line does not bear any information about claims in the other lines, since the sum is necessarily more than a million already.

Furthermore, the data does not exhibit a homogeneous arrival process, which is necessary for fitting a trivariate compound Poisson process. A transformation was applied to the arrival processes to ensure homogeneity, following the approach in Mikosch (2006). This transformation was successful and yielded processes that look homogeneous, except perhaps for the ‘Profits’ line. However, perfect homogenisation of the ‘Profits’ line cannot be expected due to a very small sample size. Overall, we consider the results reasonable for our illustrative purposes.

After these adjustments and removal of two outliers (the two largest observations), the data is reduced to 803 fire losses. The common (synchronous) claims in this set of data are directly observable as the claims data specifies the time of a claim, and claim severity across the three lines of insurance. When a claim amount is specified as 0, we assume no (Poisson) arrival in that class as there is no claim. The descriptive statistics for the reduced Danish fire data are shown Table 5.

	Building	Contents	Profits
Count	703	331	68
Mean	2.66	4.68	3.78
Standard Deviation	2.73	6.41	3.30
Skewness	4.80	4.14	2.09
Kurtosis	36.3	22.2	4.68
Minimum	1.00	1.00	1.01
First Quartile	1.32	1.46	1.67
Median	1.72	2.40	2.50
Third Quartile	2.87	4.71	4.46
Maximum	34.6	53.6	17.7

Table 5: Descriptive statistics for Danish data

The Building and Contents components have a significantly higher number of jumps compared to those in Profits. As such, there has been a primary focus on those two components in existing literature (see, for instance Esmaili and Klüppelberg, 2010). We extend those studies by including the third line available in the dataset, ‘Profits’.

### 5.2. Dependence between the three processes

Here dependence can manifest itself in either frequency (events leading to claims in more than one process) or severity (claim amounts varying in sympathy with each others).

Dependence in frequency is obvious from Table 6, which shows a decomposition of the total number of Poisson arrivals in each process. Processes include jumps that are unique to one component, jumps that are common to 2 components but not the third, and jumps that are common across all lines of business. Here,  $B$  denotes the Building component,  $C$  denotes the Contents component and  $P$  denotes Profit.  $P$  claims, although scarce (presumably because of the truncation of the data) never occur alone (unsurprisingly). Furthermore, every single  $P$  claim occurs with a  $C$  claim, which makes this pair strongly ‘frequency-dependent’. On the other hand, only about a third of the  $B$  claims are common with claims in either  $P$  or  $C$ .

Process	Number of Jumps
Unique to $B$	472
Unique to $C$	88
Unique to $P$	0
Common to $B$ and $C$ but not $P$	175
Common to $B$ and $P$ but not $C$	0
Common to $C$ and $P$ but not $B$	12
Common to $B$ , $C$ and $P$	56
<b>Total</b>	<b>803</b>

Table 6: Total number of jumps in each process

With respect to severity, Figure 6 shows the empirical copulas for the severity of *common* jumps between pairwise components. Pairwise components include the process with jumps common to two processes and not the third, as well as the process with jumps common to all components. These plots suggest some moderate positive dependence, which is confirmed by standard indicators, shown in Table 7.

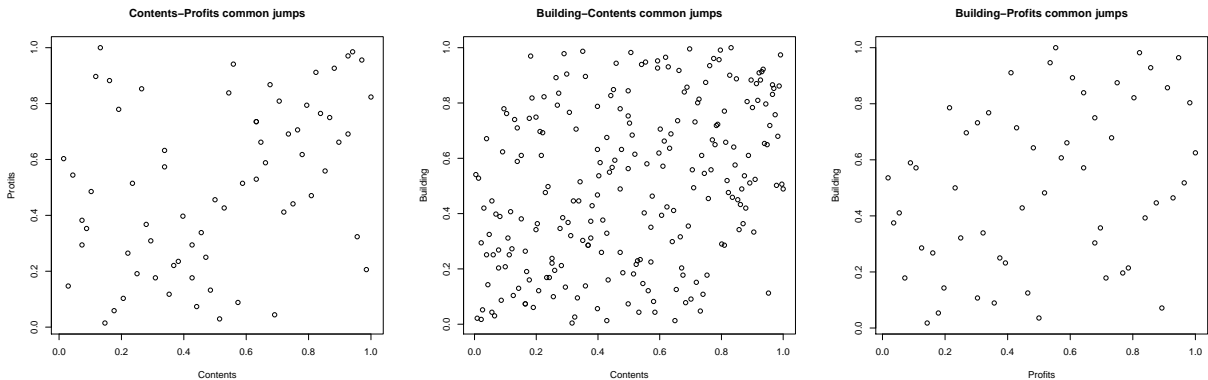


Figure 6: Empirical severity copula of common jumps between: Contents and Profits (left), Contents and Building (middle), and Profits and Building (right)

	C	P
B	0.236	0.215
C		0.260

(a) Kendall's tau

	C	P
B	0.355	0.329
C		0.366

(b) Spearman's rho

	C	P
B	0.341	0.334
C		0.323

(c) Pearson's rho

Table 7: Some standard indicators on the severities of bivariate marginal jumps

Overall, the three processes seem to be positively dependent, with the main source of dependence being common events (frequency), and with some positive dependence in severity as well. This is encouraging, as it makes the use of dependent multivariate Poisson processes an appropriate approach.

### 5.3. Fitting the marginal univariate processes

The fitting of the marginal Danish fire data processes has been studied extensively in the past literature. However, our adjustments, as well as consideration of the Profits line, require us to allocate a few lines on this step.

The estimates we obtained for the unit intensities of the three homogenised processes were 0.88, 0.41 and 0.09 (claims per day) for processes B, C and P, respectively.

We tried to fit the log of the marginal claim amounts with Lognormal, gamma, Weibull and Pareto distributions. For the C and P lines, the Weibull distribution yielded the best goodness-of-fit according to likelihood, Kolmogorov-Smirnoff and Andersen-Darling criteria. For the B line, the results were inconclusive between gamma (which had a better Andersen-Darling result) and Weibull (which had a slightly better likelihood result). In what follows, we will thus use the Weibull distribution for the C and P lines, and will defer our choice between gamma and Weibull to the trivariate likelihood fitting.

### 5.4. Fitting the trivariate process

When constructing a multivariate nested structure, one must choose which pairs to ‘couple’ first. This important choice is not trivial. Obviously, it will affect how good the fit will be, but also an inappropriate order of nesting could lead to an invalid structure (see Table 2). Often, the most strongly dependent processes must be coupled first (a similar observation was made in the distributional copula literature); see also Remark 3.3.

In order to choose a structure, we examined the goodness-of-fit of three candidate Lévy copulas (AI, Clayton and Gumbel) on the three marginal bivariate processes (using the approach suggested in Esmaeili and Klüppelberg, 2013; Avanzi et al., 2011). We ranked results using the overall maximum (full) likelihood. Initial values for the parameters were obtained using an Inference Functions of Margins (‘IFM’) method (see chapter 10 of Joe, 1997); further details about the computation of the results are available in Appendix C. Results are provided in Table 8. Note that the best severity distributions were consistent (also for sub-optimal choices of Lévy copulas): Weibull for P and C, gamma for B. Furthermore, we should note that AI wins for the couple BC only by a tiny margin; the Clayton fit was practically as good.

Couple	Lévy copula	Severity 1	Severity 2
CP	AI	Weibull (C)	Weibull (P)
BP	Clayton	gamma (B)	Weibull (P)
BC	AI	gamma (B)	Weibull (C)

Table 8: Best fits for the marginal bivariate processes

As a consequence we considered structures with AI and/or Clayton Lévy copulas. In both cases, the bivariate process with the highest dependence parameter should be coupled first. In the results of the bivariate fits, this consistently proved to be that of the pair CP, which could be expected from our analysis in Section 5.2.

Remember that the remaining two dependence relationships can be different from the first, but must necessarily be described by the same Lévy copula (see the left hand side figure in Figure 4 without the third line for a description of the trivariate fully nested structure). Thus, we investigated the choices of either AI or Clayton for the outer layer. The Clayton option did not yield reasonable results, and we finally retained fully nested trivariate AI structure with the following maximum likelihood frequency parameters

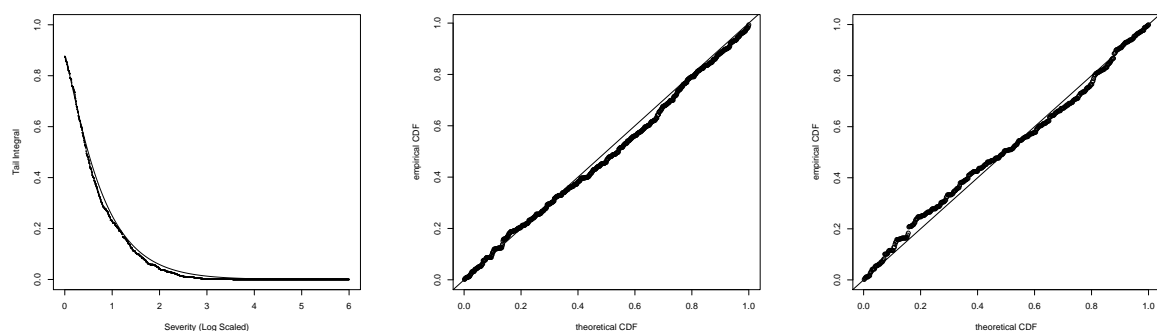
$$\hat{\lambda}_B = 0.88, \quad (\text{with s.e. } 0.03) \quad (5.1)$$

$$\hat{\lambda}_C = 0.42, \quad (\text{with s.e. } 0.02) \quad (5.2)$$

$$\hat{\lambda}_P = 0.10, \quad (\text{with s.e. } 0.01) \quad (5.3)$$

whereas the empirical frequencies were 0.88, 0.41, and 0.09, respectively. The fitted AI parameters are  $\hat{\delta}_1 = 10.28$  and  $\hat{\delta}_2 = 10.55$ , respectively. The claims severities of the processes B, C and P are gamma(1.25(s.e. 0.06),1.58(s.e. 0.08)), Weibull(1.11(s.e. 0.05),1.26(s.e. 0.06)) and Weibull(1.27(s.e. 0.12),1.50(s.e. 0.12)), respectively.

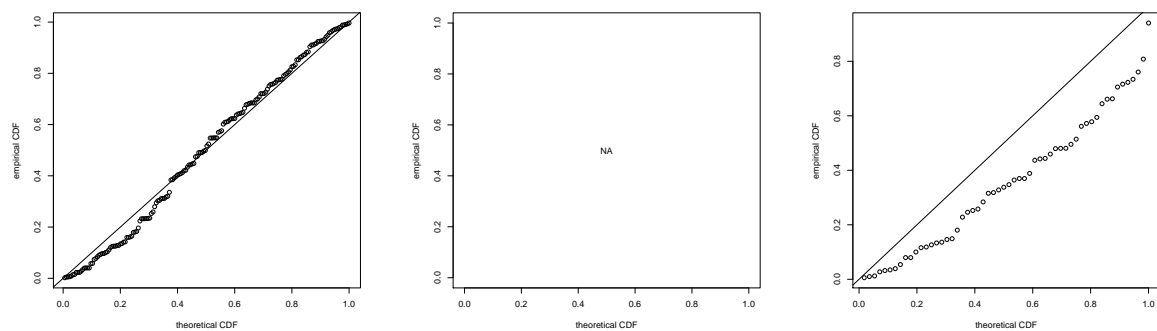
The maximum likelihood severity parameters changed from the marginal fits performed in Section 5.3 due to the inherent trade-off, during the maximisation process, between the goodness-of-fit of frequency, severity and dependence in all three processes. In this case, the fit of the severities is best illustrated with plots, which are provided in Figures 7, 8 and 9 for the processes B, C and P, respectively. The fit is quite good for B and C and slightly off for P. The latter can be explained by the relative low frequency of P claims. The other two processes carried so much weight in the maximisation process (the trade-off described above) that comparatively more concessions had to be made in the P process. Given we have only few data points for that process we believe the goodness-of-fit of this final solution is still acceptable.



(a) Empirical (dots) vs Theoretical (solid line) tail integrals

(b) Severities of all jumps

(c) Severities of unique jumps



(d) Severities when common with C but not P

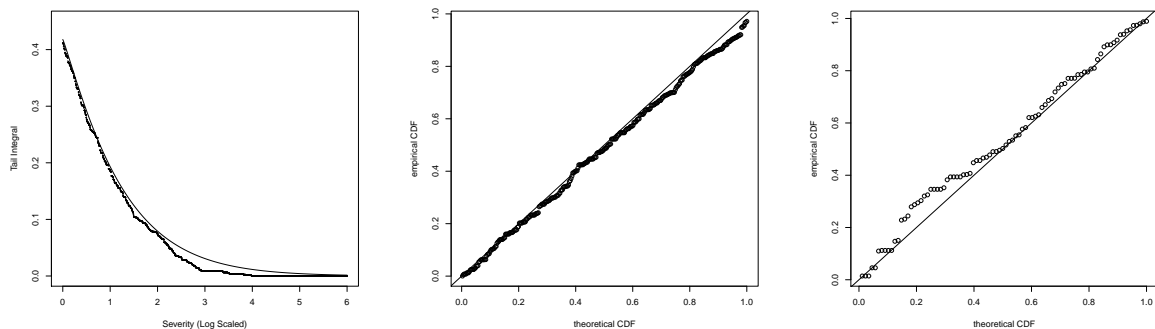
(e) Severities when common with P but not C

(f) Severities when common with C and P

Figure 7: Goodness-of-fit of the Building claims process

## Acknowledgments

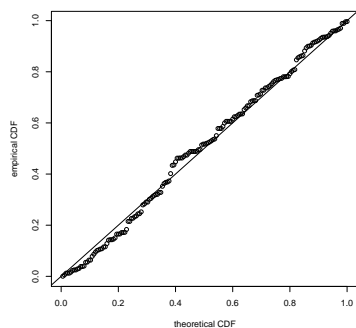
The authors are grateful to [anonymous reviewers for comments, as well as to colleagues who attended the 16<sup>th</sup> International Congress on *Insurance: Mathematics and Economics* in June 2012 in Hong Kong, and who provided useful insights about the fitting procedure.] This research was supported under an Australian Actuarial Research Grant provided by the Actuaries Institute, as well as under Australian Research Council's



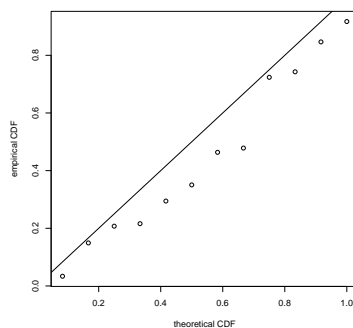
(a) Empirical (dots) vs Theoretical (solid line) tail integrals

(b) Severities of all jumps

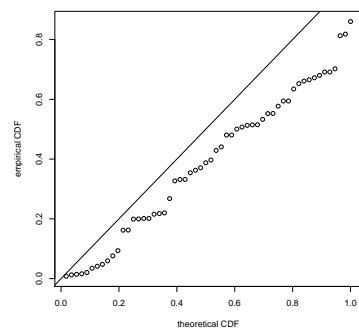
(c) Severities of unique jumps



(d) Severities when common with B but not P



(e) Severities when common with P but not B



(f) Severities when common with B and P

Figure 8: Goodness-of-fit of the Contents claims process

*Linkage Projects* funding scheme (project number LP130100723). Jamie Tao acknowledges financial support from the Brian Gray (Honours) scholarship, provided by APRA and RBA. Xinda Yang acknowledges financial support from an Australian Postgraduate Award, as well as supplementary scholarships provided by the Australian School of Business, UNSW. The views expressed herein are those of the authors and are not necessarily those of the supporting organisations.

## A. Spectrally Positive Lévy Processes

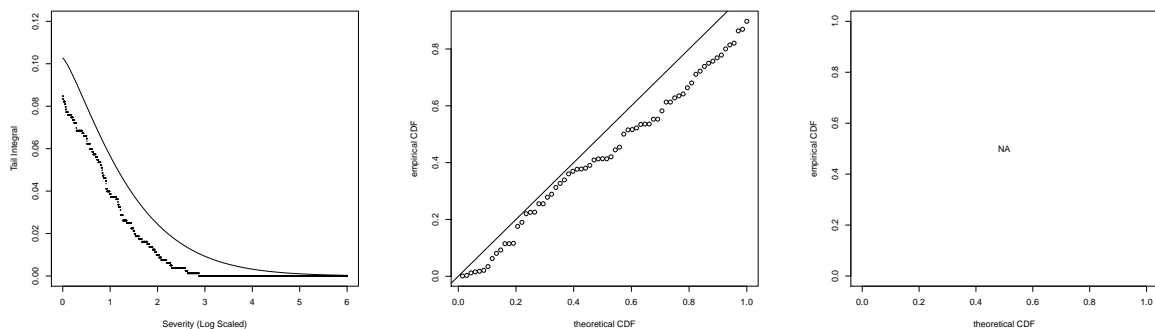
**Definition A.1.** (Lévy process, Sato, 1999) *A stochastic process  $\mathbf{S}(t)$  on  $\mathbb{R}^d$  is a Lévy process if*

1. *For every  $s, t \geq 0$ , the increment of the process  $\mathbf{S}(t+s) - \mathbf{S}(t)$  is independent of the process  $(\mathbf{S}(v), 0 \leq v \leq t)$*
2. *The distribution of  $\mathbf{S}(t+s) - \mathbf{S}(t)$  is the same as  $\mathbf{S}(s)$  (or stationary).*
3.  $\Pr(\mathbf{S}_0 = \mathbf{0}) = 1$
4. *It is stochastically continuous.*

Lévy measures are measurements of jumps which characterise a particular Lévy process, with the exception of the Wiener process. The *Lévy measure* is defined as follows.

**Definition A.2.** (Lévy measure) *For a Lévy process  $\mathbf{S}(t)$ , the Lévy measure  $\nu$  is defined by*

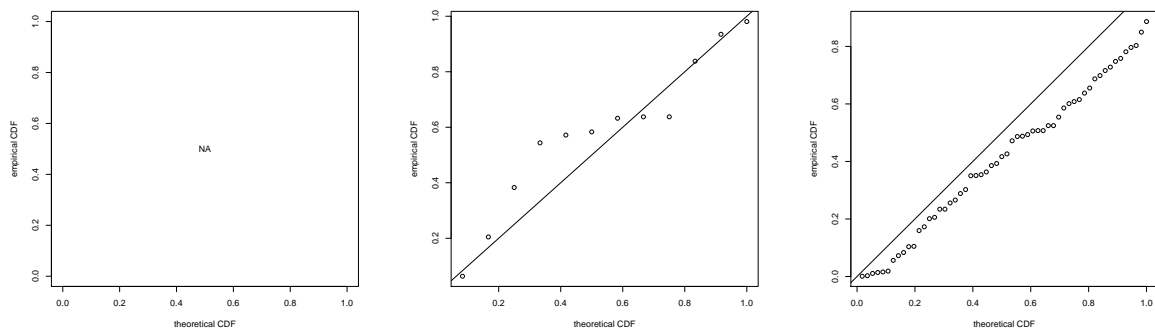
$$\nu(A) = E[\#\{t \in [0, 1] : \Delta\mathbf{S}(t) \neq \mathbf{0}, \Delta\mathbf{S}(t) \in A\}], \quad A \in \mathcal{B}(\mathbb{R}^d), \quad (\text{A.1})$$



(a) Empirical (dots) vs Theoretical (solid line) tail integrals

(b) Severities of all jumps

(c) Severities of unique jumps



(d) Severities when common with B but not C

(e) Severities when common with C but not B

(f) Severities when common with B and C

Figure 9: Goodness-of-fit of the Profits claims process

where  $\nu(A)$  is the expected number of jumps whose size belongs to  $A$  (for 1 unit of time), the expected number of jumps of size 0 are excluded (as this will result in an undefined Lévy measure), and where  $\mathcal{B}$  represents the Borel set.

## B. Detailed Proofs

The validity of any Lévy copula construction relies on whether or not it satisfies the properties of a Lévy copula. Recall from its formal definition (see, e.g. Tankov, 2003; Cont and Tankov, 2004) that these properties are 1) grounded; 2) uniform margins; and 3)  $d$ -increasing.

For the purpose of this paper, a formal proof will be completed for the trivariate and quadivariate nested Lévy copulas. Applying the same ideas as we present, the proof can readily be extended to any dimension, but the proof becomes more tedious with increasing dimensions. Firstly, we need the following definitions.

**Definition B.1.** (Monotonicity, McNeil and Nešlehová, 2009; Bäuerle and Blatter, 2011)

1. A real function  $f$  is said to be  $d$ -monotone on  $(a, b)$ , where  $a, b \in \mathbb{R}$  and  $d \geq 2$ , if it is differentiable on  $(a, b)$  up to order  $d - 2$  and the derivatives satisfy

$$(-1)^k f^{(k)}(x) \geq 0, \quad k \in \{0, 1, \dots, d - 2\}, \quad (\text{B.1})$$

for any  $x \in (a, b)$  and further if  $(-1)^{d-2} f^{(d-2)}$  is non-increasing and convex in  $(a, b)$ .

2. For  $d = 1$ ,  $f$  is called 1-monotone in  $(a, b)$  if it is non-negative and non-increasing in  $(a, b)$ .
3. If  $f$  has derivatives of all orders on  $(a, b)$  and if  $(-1)^k f^{(k)} \geq 0$  for any  $x$  in  $(a, b)$ , then  $f$  is said to be completely monotone.

**Definition B.2.** ( $d$ -increasing, Nelsen, 1999) A  $d$ -variate function  $F$  is called  $d$ -increasing if and only if for all choices  $B = [a_1, b_1] \times \dots \times [a_d, b_d]$  in the domain of  $F$ , where  $a_i \leq b_i$  for  $i \in \{1, \dots, d\}$ ,

$$V_F(B) \geq 0, \quad (\text{B.2})$$

where  $V_F(B)$ , the  $F$ -volume of  $B$  is given by  $d$ -th order difference,

$$\begin{aligned} V_F(B) &= \Delta_{a_1}^{b_1} \dots \Delta_{a_d}^{b_d} F(u_1, \dots, u_d) \\ &= \sum_{i_1=1}^2 \dots \sum_{i_d=1}^2 (-1)^{i_1 + \dots + i_d} F(u_{1i_1}, \dots, u_{di_d}), \end{aligned} \quad (\text{B.3})$$

where  $u_{j1} = a_j$  and  $u_{j2} = b_j$  for all  $j \in \{1, \dots, d\}$ .

Before we can formally prove the validity of the nested construction, we will require the following ancillary result.

**Lemma B.1.** (Non-negativity of partial derivative to difference 2) Suppose that  $\frac{\partial^d}{\partial u_1 \dots \partial u_d} \mathfrak{C}(u_1, \dots, u_d)$  exists and is continuous. Then, for any choice of a  $d$ -box  $B = [a_1, b_1] \times \dots \times [a_d, b_d]$  in the domain of  $\mathfrak{C}$  where  $a_i \leq b_i$ , we have, for all  $i \in \{1, \dots, d\}$ ,

$$\Delta_{a_1}^{b_1} \dots \Delta_{a_d}^{b_d} \mathfrak{C}(u_1, \dots, u_d) \geq 0, \quad (\text{B.4})$$

if and only if

$$\frac{\partial^d}{\partial u_1 \dots \partial u_d} \mathfrak{C}(u_1, \dots, u_d) \geq 0. \quad (\text{B.5})$$

*Proof.* For  $d = 2$  the result is well known and can be found in, for example, Ross (1983, Proposition 4.2). Extending to the case of  $d > 2$ , if  $\frac{\partial^d}{\partial u_1 \dots \partial u_d} \mathfrak{C}(u_1, \dots, u_d) \geq 0$ , then for  $a_i \leq b_i$  for all  $i \in \{1, \dots, d\}$ ,

$$\begin{aligned} &\int_{a_1}^{b_1} \dots \int_{a_{d-1}}^{b_{d-1}} \int_{a_d}^{b_d} \frac{\partial^d}{\partial u_1 \dots \partial u_d} \mathfrak{C}(u_1, \dots, u_{d-1}, u_d) du_d \geq 0 \\ &\int_{a_1}^{b_1} \dots \int_{a_{d-1}}^{b_{d-1}} \frac{\partial^{d-1}}{\partial u_1 \dots \partial u_{d-1}} (\Delta_{a_d}^{b_d} \mathfrak{C}(u_1, \dots, u_{d-1}, u_d)) du_{d-1} \geq 0 \\ &\vdots \\ &\Delta_{a_1}^{b_1} \dots \Delta_{a_{d-1}}^{b_{d-1}} \Delta_{a_d}^{b_d} \mathfrak{C}(u_1, \dots, u_{d-1}, u_d) \geq 0. \end{aligned} \quad (\text{B.6})$$

To go the other way, assume that  $\Delta_{a_1}^{b_1} \dots \Delta_{a_{d-1}}^{b_{d-1}} \Delta_{a_d}^{b_d} \mathfrak{C}(u_1, \dots, u_{d-1}, u_d) \geq 0$ . Then for  $a_i \leq b_i$  for all  $i \in \{1, \dots, d\}$ ,

$$\frac{\Delta_{a_1}^{b_1} \dots \Delta_{a_{d-1}}^{b_{d-1}} \mathfrak{C}(u_1, \dots, u_{d-1}, b_d)}{\Delta_{a_1}^{b_1} \dots \Delta_{a_{d-1}}^{b_{d-1}} (u_1 \times \dots \times u_{d-1})} \geq \frac{\Delta_{a_1}^{b_1} \dots \Delta_{a_{d-1}}^{b_{d-1}} \mathfrak{C}(u_1, \dots, u_{d-1}, a_d)}{\Delta_{a_1}^{b_1} \dots \Delta_{a_{d-1}}^{b_{d-1}} (u_1 \times \dots \times u_{d-1})}. \quad (\text{B.7})$$

Now let  $b_1$  converge to  $a_1$  to obtain

$$\frac{\partial}{\partial u_1} \left( \frac{\Delta_{a_2}^{b_2} \dots \Delta_{a_{d-1}}^{b_{d-1}} \mathfrak{C}(u_1, \dots, u_{d-1}, b_d)}{\Delta_{a_2}^{b_2} \dots \Delta_{a_{d-1}}^{b_{d-1}} (u_2 \times \dots \times u_{d-1})} \right) \geq \frac{\partial}{\partial u_1} \left( \frac{\Delta_{a_2}^{b_2} \dots \Delta_{a_{d-1}}^{b_{d-1}} \mathfrak{C}(u_1, \dots, u_{d-1}, a_d)}{\Delta_{a_2}^{b_2} \dots \Delta_{a_{d-1}}^{b_{d-1}} (u_2 \times \dots \times u_{d-1})} \right). \quad (\text{B.8})$$

Then let  $b_2$  converge to  $a_2, \dots, b_{d-1}$  converge to  $a_{d-1}$  to obtain

$$\frac{\partial}{\partial u_1 \dots \partial u_{d-1}} (\mathfrak{C}(u_1, \dots, u_{d-1}, b_d)) \geq \frac{\partial}{\partial u_1 \dots \partial u_{d-1}} (\mathfrak{C}(u_1, \dots, u_{d-1}, a_d)), \quad (\text{B.9})$$

implying that

$$\frac{\partial^d}{\partial u_1 \dots \partial u_d} \mathfrak{C}(u_1, \dots, u_d) \geq 0 \quad (\text{B.10})$$

□

### B.1. Proof of Theorem 3.1

(Grounded)

The property of being grounded means that  $\mathfrak{C}(u_1, u_2, u_3) = 0$  whenever  $u_i = 0$  for at least one  $i \in \{1, 2, 3\}$ . If  $u_1 = 0$ ,

$$\begin{aligned} \mathfrak{C}(0, u_2, u_3) &= \psi_{21}^{-1}(\psi_{21} \circ \psi_{11}^{-1}(\psi_{11}(0) + \psi_{11}(u_2)) + \psi_{21}(u_3)) \\ &= \psi_{21}^{-1}(\psi_{21} \circ \psi_{11}^{-1}(\infty) + \psi_{21}(u_3)) \\ &= \psi_{21}^{-1}(\infty) = 0. \end{aligned} \quad (\text{B.11})$$

The result for  $u_2$  has a similar proof.

If  $u_3 = 0$ ,

$$\begin{aligned} \mathfrak{C}(u_1, u_2, 0) &= \psi_{21}^{-1}(\psi_{21} \circ \psi_{11}^{-1}(\psi_{11}(u_1) + \psi_{11}(u_2)) + \psi_{21}(0)) \\ &= \psi_{21}^{-1}(\infty) = 0. \end{aligned} \quad (\text{B.12})$$

(Uniform margins)

The margin for  $u_1$  is given by

$$\begin{aligned} \mathfrak{C}(u_1, \infty, \infty) &= \psi_{21}^{-1}(\psi_{21} \circ \psi_{11}^{-1}(\psi_{11}(u_1) + \psi_{11}(\infty)) + \psi_{21}(\infty)) \\ &= \psi_{21}^{-1}(\psi_{21}(u_1) + 0) = u_1. \end{aligned} \quad (\text{B.13})$$

The margin for  $u_2$  has a similar proof.

The margin for  $u_3$  is given by

$$\begin{aligned} \mathfrak{C}(\infty, \infty, u_3) &= \psi_{21}^{-1}(\psi_{21} \circ \psi_{11}^{-1}(\psi_{11}(\infty) + \psi_{11}(\infty)) + \psi_{21}(u_3)) \\ &= \psi_{21}^{-1}(0 + \psi_{21}(u_3)) = u_3. \end{aligned} \quad (\text{B.14})$$

(3-increasing)

Let us denote  $\psi_{21}^{-1}$  as  $g$  and  $\psi_{21} \circ \psi_{11}^{-1}$  as  $h$  for notational simplicity, so (3.1) can be expressed as

$$\mathfrak{C}(u_1, u_2, u_3) = g(h(\psi_{11}(u_1) + \psi_{11}(u_2)) + \psi_{21}(u_3)). \quad (\text{B.15})$$

If  $g$  is 3-continuously differentiable and 3-monotone on  $(0, \infty)$ , then for all  $u > 0$ :  $g(u) \geq 0$ ,  $g'(u) \leq 0$ ,  $g''(u) \geq 0$ , and  $g'''(u) \leq 0$ . If  $h$  has a derivative which is 1-continuously differentiable and 1-monotone on  $(0, \infty)$ , then for all  $u > 0$ :  $h(u) \geq 0$ ,  $h'(u) \geq 0$ , and  $h''(u) \leq 0$ . Furthermore, as  $\psi_{11}$  and  $\psi_{21}$  are Archimedean Lévy generators, they are continuous and decreasing functions on  $(0, \infty)$  by definition, so for all  $u > 0$ :  $\psi_{i1}(u) \geq 0$  and  $\psi'_{i1}(u) \leq 0$  for  $i \in \{1, 2\}$ . Hence,

$$\frac{\partial^3}{\partial u_1 \partial u_2 \partial u_3} \mathfrak{C}(u_1, u_2, u_3) = \psi'_{11}(u_1) \psi'_{11}(u_2) \psi'_{21}(u_3) [g'''(\cdot)(h'(\cdot))^2 + g''(\cdot)h''(\cdot)] \geq 0, \quad (\text{B.16})$$

which is continuous, and by Lemma B.1, the nested construction has a non-negative 3rd order difference. Thus, a construction subject to the conditions specified in Theorem 3.1 is 3-increasing and the construction yields a valid Lévy copula. □

B.2. Proof of Theorem 3.2

(Grounded)

If  $u_1 = 0$ ,

$$\begin{aligned}
\mathfrak{C}(0, u_2, u_3, u_4) &= \psi_{31}^{-1}(\psi_{31} \circ \psi_{21}^{-1}(\psi_{21} \circ \psi_{11}^{-1}(\psi_{11}(0) + \psi_{11}(u_2)) + \psi_{21}(u_3)) + \psi_{31}(u_4)) \\
&= \psi_{31}^{-1}(\psi_{31} \circ \psi_{21}^{-1}(\psi_{21} \circ \psi_{11}^{-1}(\infty) + \psi_{21}(u_3)) + \psi_{31}(u_4)) \\
&= \psi_{31}^{-1}(\psi_{31} \circ \psi_{21}^{-1}(\infty) + \psi_{31}(u_4)) \\
&= \psi_{31}^{-1}(\infty) = 0.
\end{aligned} \tag{B.17}$$

The result for  $u_2$  has a similar proof.

If  $u_3 = 0$ ,

$$\begin{aligned}
\mathfrak{C}(u_1, u_2, 0, u_4) &= \psi_{31}^{-1}(\psi_{31} \circ \psi_{21}^{-1}(\psi_{21} \circ \psi_{11}^{-1}(\psi_{11}(u_1) + \psi_{11}(u_2)) + \psi_{21}(0)) + \psi_{31}(u_4)) \\
&= \psi_{31}^{-1}(\psi_{31} \circ \psi_{21}^{-1}(\infty) + \psi_{31}(u_4)) \\
&= \psi_{31}^{-1}(\infty) = 0.
\end{aligned} \tag{B.18}$$

If  $u_4 = 0$ ,

$$\begin{aligned}
\mathfrak{C}(u_1, u_2, u_3, 0) &= \psi_{31}^{-1}(\psi_{31} \circ \psi_{21}^{-1}(\psi_{21} \circ \psi_{11}^{-1}(\psi_{11}(u_1) + \psi_{11}(u_2)) + \psi_{21}(u_3)) + \psi_{31}(0)) \\
&= \psi_{31}^{-1}(\infty) = 0.
\end{aligned} \tag{B.19}$$

(Uniform margins)

The margin for  $u_1$  is given by

$$\begin{aligned}
\mathfrak{C}(u_1, \infty, \infty, \infty) &= \psi_{31}^{-1}(\psi_{31} \circ \psi_{21}^{-1}(\psi_{21} \circ \psi_{11}^{-1}(\psi_{11}(u_1) + \psi_{11}(\infty)) + \psi_{21}(\infty)) + \psi_{31}(\infty)) \\
&= \psi_{31}^{-1}(\psi_{31} \circ \psi_{21}^{-1}(\psi_{21}(u_1) + 0) + 0) \\
&= \psi_{31}^{-1}(\psi_{31}(u_1)) = u_1.
\end{aligned} \tag{B.20}$$

The margin for  $u_2$  has a similar proof.

The margin for  $u_3$  is given by

$$\begin{aligned}
\mathfrak{C}(\infty, \infty, u_3, \infty) &= \psi_{31}^{-1}(\psi_{31} \circ \psi_{21}^{-1}(\psi_{21} \circ \psi_{11}^{-1}(\psi_{11}(\infty) + \psi_{11}(\infty)) + \psi_{21}(u_3)) + \psi_{31}(\infty)) \\
&= \psi_{31}^{-1}(\psi_{31} \circ \psi_{21}^{-1}(0 + \psi_{21}(u_3)) + 0) \\
&= \psi_{31}^{-1}(\psi_{31}(u_3)) = u_3.
\end{aligned} \tag{B.21}$$

The margin for  $u_4$  is given by

$$\begin{aligned}
\mathfrak{C}(\infty, \infty, \infty, u_4) &= \psi_{31}^{-1}(\psi_{31} \circ \psi_{21}^{-1}(\psi_{21} \circ \psi_{11}^{-1}(\psi_{11}(\infty) + \psi_{11}(\infty)) + \psi_{21}(\infty)) + \psi_{31}(u_4)) \\
&= \psi_{31}^{-1}(\psi_{31}(u_4)) = u_4.
\end{aligned} \tag{B.22}$$

(4-increasing)

Let us denote  $\psi_{31}^{-1}$  as  $g$ ,  $\psi_{31} \circ \psi_{21}^{-1}$  as  $h_2$  and  $\psi_{21} \circ \psi_{11}^{-1}$  as  $h_1$  for notational simplicity, so (3.2) can be expressed as

$$\mathfrak{C}(u_1, u_2, u_3, u_4) = g(h_2(h_1(\psi_{11}(u_1) + \psi_{11}(u_2)) + \psi_{21}(u_3)) + \psi_{31}(u_4)). \tag{B.23}$$

If  $g$  is 4-continuously differentiable and 4-monotone on  $(0, \infty)$ , then for all  $u > 0$ :  $g(u) \geq 0$ ,  $g'(u) \leq 0$ ,  $g''(u) \geq 0$ ,  $g'''(u) \leq 0$  and  $g''''(u) \geq 0$ . If  $h_2$  has a derivative which is 2-continuously differentiable and 2-monotone on  $(0, \infty)$ , then for all  $u > 0$ :  $h_2(u) \geq 0$ ,  $h_2'(u) \geq 0$ ,  $h_2''(u) \leq 0$  and  $h_2''' \geq 0$ . If  $h_1$  has a derivative which is 1-continuously differentiable and 1-monotone on  $(0, \infty)$ , then for all  $u > 0$ :  $h_1(u) \geq 0$ ,  $h_1'(u) \geq 0$ ,  $h_1''(u) \leq 0$ . Furthermore, as  $\psi_{11}$ ,  $\psi_{21}$  and  $\psi_{31}$  are Archimedean Lévy generators, they are continuous and

decreasing functions on  $(0, \infty)$  by definition, so for all  $u > 0$ :  $\psi_{i1}(u) \geq 0$  and  $\psi'_{i1}(u) \leq 0$  for  $i \in \{1, 2, 3\}$ . Hence,

$$\begin{aligned} \frac{\partial^4}{\partial u_1 \partial u_2 \partial u_3 \partial u_4} \mathfrak{C}(u_1, u_2, u_3, u_4) &= \psi'_{11}(u_1) \psi'_{11}(u_2) \psi'_{21}(u_3) \psi'_{31}(u_4) [g''''(\cdot) (h'_2(\cdot))^3 (h'_1(\cdot))^2 \\ &\quad + 2g'''(\cdot) h'_2(\cdot) h''_2(\cdot) (h'_1(\cdot))^2 + g'''(\cdot) h''_2(\cdot) h'_2(\cdot) (h'_1(\cdot))^2 \\ &\quad + g''(\cdot) h_2'''(\cdot) (h'_1(\cdot))^2 + g''(\cdot) (h'_2(\cdot))^2 h''_1(\cdot) \\ &\quad + g''(\cdot) h_2''(\cdot) h''_1(\cdot)] \geq 0, \end{aligned} \quad (\text{B.24})$$

which is continuous, and by Lemma B.1, the nested construction has a non-negative 4th order difference. Thus, a construction subject to the conditions specified in Theorem 3.2 is 4-increasing and the construction yields a valid Lévy copula.

**Remark B.1.** *Note proof of Theorem 3.2 (quadivariate fully nested) is similar to that of Theorem 3.1 (trivariate fully nested), but is included here to illustrate how to generalise results to any dimensions. However, the proofs may become more tedious the larger the Lévy copula model.*

### B.3. Proof of Theorem 3.3

(Grounded)

If  $u_1 = 0$ ,

$$\begin{aligned} \mathfrak{C}(0, u_2, u_3, u_4) &= \psi_{21}^{-1}(\psi_{21} \circ \psi_{11}^{-1}(\psi_{11}(0) + \psi_{11}(u_2)) + \psi_{21} \circ \psi_{12}^{-1}(\psi_{12}(u_3) + \psi_{12}(u_4))) \\ &= \psi_{21}^{-1}(\psi_{21} \circ \psi_{11}^{-1}(\infty) + \psi_{21} \circ \psi_{12}^{-1}(\psi_{12}(u_3) + \psi_{12}(u_4))) \\ &= \psi_{21}^{-1}(\infty) = 0. \end{aligned} \quad (\text{B.25})$$

The result for  $u_2$  has a similar proof.

If  $u_3 = 0$ ,

$$\begin{aligned} \mathfrak{C}(u_1, u_2, 0, u_4) &= \psi_{21}^{-1}(\psi_{21} \circ \psi_{11}^{-1}(\psi_{11}(u_1) + \psi_{11}(u_2)) + \psi_{21} \circ \psi_{12}^{-1}(\psi_{12}(0) + \psi_{12}(u_4))) \\ &= \psi_{21}^{-1}(\psi_{21} \circ \psi_{11}^{-1}(\psi_{11}(u_1) + \psi_{11}(u_2)) + \psi_{21} \circ \psi_{12}^{-1}(\infty)) \\ &= \psi_{21}^{-1}(\infty) = 0. \end{aligned} \quad (\text{B.26})$$

The result for  $u_4$  has a similar proof.

(Uniform margins)

The margin for  $u_1$  is given by

$$\begin{aligned} \mathfrak{C}(u_1, \infty, \infty, \infty) &= \psi_{21}^{-1}(\psi_{21} \circ \psi_{11}^{-1}(\psi_{11}(u_1) + \psi_{11}(\infty)) + \psi_{21} \circ \psi_{12}^{-1}(\psi_{12}(\infty) + \psi_{12}(\infty))) \\ &= \psi_{21}^{-1}(\psi_{21} \circ \psi_{11}^{-1}(\psi_{11}(u_1)) + 0) = u_1. \end{aligned} \quad (\text{B.27})$$

The margin for  $u_2$  has a similar proof.

The margin for  $u_3$  is given by

$$\begin{aligned} \mathfrak{C}(\infty, \infty, u_3, \infty) &= \psi_{21}^{-1}(\psi_{21} \circ \psi_{11}^{-1}(\psi_{11}(\infty) + \psi_{11}(\infty)) + \psi_{21} \circ \psi_{12}^{-1}(\psi_{12}(u_3) + \psi_{12}(\infty))) \\ &= \psi_{21}^{-1}(0 + \psi_{21} \circ \psi_{12}^{-1}(\psi_{12}(u_3))) = u_3. \end{aligned} \quad (\text{B.28})$$

The margin for  $u_4$  has a similar proof.

(4-increasing)

Let us denote  $\psi_{21}^{-1}$  as  $g$ ,  $\psi_{21} \circ \psi_{12}^{-1}$  as  $h_2$  and  $\psi_{21} \circ \psi_{11}^{-1}$  as  $h_1$  for notational simplicity, so (3.3) can be expressed as

$$\mathfrak{C}(u_1, u_2, u_3, u_4) = g(h_1(\psi_{11}(u_1) + \psi_{11}(u_2)) + h_2(\psi_{12}(u_3) + \psi_{12}(u_4))). \quad (\text{B.29})$$

If  $g$  is 4-continuously differentiable and 4-monotone on  $(0, \infty)$ , then for all  $u > 0$ :  $g(u) \geq 0$ ,  $g'(u) \leq 0$ ,  $g''(u) \geq 0$ ,  $g'''(u) \leq 0$  and  $g''''(u) \geq 0$ . If  $h_i$  has a derivative which is 1-continuously differentiable and 1-monotone on  $(0, \infty)$ , then for all  $u > 0$ :  $h_i(u) \geq 0$ ,  $h'_i(u) \geq 0$ ,  $h''_i(u) \leq 0$  for  $i \in \{1, 2\}$ . Furthermore, as  $\psi_{11}$ ,  $\psi_{12}$  and  $\psi_{21}$  are Archimedean Lévy generators, they are continuous and decreasing functions on  $(0, \infty)$  by definition, so for all  $u > 0$ :  $\psi_{ij}(u) \geq 0$  and  $\psi'_{ij}(u) \leq 0$  for  $i, j \in \{1, 2\}$  and  $i + j < 4$ . Hence,

$$\begin{aligned} \frac{\partial^4}{\partial u_1 \partial u_2 \partial u_3 \partial u_4} \mathfrak{C}(u_1, u_2, u_3, u_4) &= \psi'_{11}(u_1) \psi'_{11}(u_2) \psi'_{12}(u_3) \psi'_{12}(u_4) [g''''(\cdot) (h'_1(\cdot))^2 (h'_2(\cdot))^2 \\ &\quad + g'''(\cdot) (h'_1(\cdot))^2 h''_2(\cdot) + g'''(\cdot) (h'_2(\cdot))^2 h''_1(\cdot) \\ &\quad + g''(\cdot) h'_1(\cdot) h''_2(\cdot)] \geq 0, \end{aligned} \quad (\text{B.30})$$

which is continuous, and by Lemma B.1, the partially nested construction has a non-negative 4th order difference. Thus, a construction subject to the conditions specified in Theorem 3.3 is 4-increasing and the construction yields a valid Lévy copula.  $\square$

#### B.4. Proof of (3.9)

The joint survival function of a trivariate compound Poisson process, derived by Avanzi et al. (2011) is given by

$$\bar{F}_{123}^{\parallel}(x_1, x_2, x_3) = \frac{1}{\lambda_{123}^{\parallel}} \mathfrak{C}(\lambda_1 \bar{F}_1(x_1), \lambda_2 \bar{F}_2(x_2), \lambda_3 \bar{F}_3(x_3)), \quad (\text{B.31})$$

where  $\mathfrak{C}$  is a trivariate Lévy copula,  $\lambda_{123}^{\parallel} = \mathfrak{C}(\lambda_1, \lambda_2, \lambda_3)$  is the arrival intensity of 3-common jumps,  $\lambda_i$  is the arrival intensity parameter and  $\bar{F}_i$  is the survival function of marginal component  $i$ , for  $i \in \{1, 2, 3\}$ .

By selecting the trivariate nested Lévy copula in (3.1) for  $\mathfrak{C}$ , the joint survival and marginal survival functions for the 3-common jumps can be written as

$$\bar{F}_{123}^{\parallel}(x_1, x_2, x_3) = \frac{1}{\lambda_{123}^{\parallel}} \psi_{21}^{-1} \left( \psi_{21} \circ \psi_{11}^{-1} (\psi_{11}(\lambda_1 \bar{F}_1(x_1)) + \psi_{11}(\lambda_2 \bar{F}_2(x_2))) + \psi_{21}(\lambda_3 \bar{F}_3(x_3)) \right), \quad (\text{B.32})$$

$$\bar{F}_{1;123}^{\parallel}(x_1) = \frac{1}{\lambda_{123}^{\parallel}} \psi_{21}^{-1} \left( \psi_{21} \circ \psi_{11}^{-1} (\psi_{11}(\lambda_1 \bar{F}_1(x_1)) + \psi_{11}(\lambda_2)) + \psi_{21}(\lambda_3) \right), \quad (\text{B.33})$$

$$\bar{F}_{2;123}^{\parallel}(x_2) = \frac{1}{\lambda_{123}^{\parallel}} \psi_{21}^{-1} \left( \psi_{21} \circ \psi_{11}^{-1} (\psi_{11}(\lambda_1) + \psi_{11}(\lambda_2 \bar{F}_2(x_2))) + \psi_{21}(\lambda_3) \right), \quad (\text{B.34})$$

$$\bar{F}_{3;123}^{\parallel}(x_3) = \frac{1}{\lambda_{123}^{\parallel}} \psi_{21}^{-1} \left( \psi_{21} \circ \psi_{11}^{-1} (\psi_{11}(\lambda_1) + \psi_{11}(\lambda_2)) + \psi_{21}(\lambda_3 \bar{F}_3(x_3)) \right), \quad (\text{B.35})$$

where  $\lambda_{123}^{\parallel} = \psi_{21}^{-1} \left( \psi_{21} \circ \psi_{11}^{-1} (\psi_{11}(\lambda_1) + \psi_{11}(\lambda_2)) + \psi_{21}(\lambda_3) \right)$  and  $\psi_{11}, \psi_{21}$  are Archimedean Lévy generators that meet the relevant restrictions.

Deriving the survival copula for the severity of 3-common jumps involves finding a relationship between the joint survival function in (B.32) and the marginal survival functions in (B.33)-(B.35). This can be done as follows. From  $\bar{F}_{1;123}^{\parallel}(x_1)$  and  $\bar{F}_{2;123}^{\parallel}(x_2)$ , we have

$$\begin{aligned} \left( \psi_{11}(\lambda_1 \bar{F}_1(x_1)) + \psi_{11}(\lambda_2 \bar{F}_2(x_2)) \right) &= \psi_{11} \circ \psi_{21}^{-1} \left( \psi_{21}(\lambda_{123}^{\parallel} \bar{F}_{1;123}^{\parallel}(x_1)) - \psi_{21}(\lambda_3) \right) \\ &\quad + \psi_{11} \circ \psi_{21}^{-1} \left( \psi_{21}(\lambda_{123}^{\parallel} \bar{F}_{2;123}^{\parallel}(x_2)) - \psi_{21}(\lambda_3) \right) \\ &\quad - \left( \psi_{11}(\lambda_1) + \psi_{11}(\lambda_2) \right). \end{aligned} \quad (\text{B.36})$$

From  $\bar{F}_{3;123}^{\parallel}(x_3)$ , we have

$$\psi_{21}(\lambda_{123}^{\parallel} \bar{F}_{3;123}^{\parallel}(x_3)) - \psi_{21} \circ \psi_{11}^{-1} (\psi_{11}(\lambda_1) + \psi_{11}(\lambda_2)) = \psi_{21}(\lambda_3 \bar{F}_3(x_3)). \quad (\text{B.37})$$

Hence,

$$\begin{aligned}
\bar{F}_{123}^{\parallel}(x_1, x_2, x_3) &= \frac{1}{\lambda_{123}^{\parallel}} \psi_{21}^{-1} \left( \psi_{21} \circ \psi_{11}^{-1} \left( \psi_{11} \circ \psi_{21}^{-1} \left( \psi_{21}(\lambda_{123}^{\parallel} \bar{F}_{1;123}^{\parallel}(x_1)) - \psi_{21}(\lambda_3) \right) \right. \right. \\
&\quad \left. \left. + \psi_{11} \circ \psi_{21}^{-1} \left( \psi_{21}(\lambda_{123}^{\parallel} \bar{F}_{2;123}^{\parallel}(x_2)) - \psi_{21}(\lambda_3) \right) - \left( \psi_{11}(\lambda_1) + \psi_{11}(\lambda_2) \right) \right) \right) \\
&\quad \left. + \psi_{21}(\lambda_{123}^{\parallel} \bar{F}_{3;123}^{\parallel}(x_3)) - \psi_{21} \circ \psi_{11}^{-1}(\psi_{11}(\lambda_1) + \psi_{11}(\lambda_2)) \right). \tag{B.38}
\end{aligned}$$

Rearranging (B.38) and substituting in the Clayton Archimedean Lévy generators yields the desired result.  $\square$

### C. Details on the data fitting programs

The software we used for fitting a model to the Danish fire data in Section 5 is **R** (version 2.15.3) with packages **stats4** and **moments**. Maximum likelihood estimation was used for estimating parameters. The likelihood function of a bivariate Lévy copula model has been discussed in existing literature (see, for example, Avanzi et al., 2011). The likelihood function of a trivariate Lévy copula model with three marginal compound Poisson process is

$$\begin{aligned}
\log L(\lambda_1, \lambda_2, \lambda_3, \boldsymbol{\theta}_1, \boldsymbol{\theta}_2, \boldsymbol{\theta}_3, \boldsymbol{\delta}) &= n_1 \log \lambda_1 + n_2 \log \lambda_2 + n_3 \log \lambda_3 - T\lambda \\
&+ \sum_{i=1}^{n_1} \log f_1(x_{1,i} | \boldsymbol{\theta}_1) + \sum_{i=1}^{n_2} \log f_2(x_{2,i} | \boldsymbol{\theta}_2) + \sum_{i=1}^{n_3} \log f_3(x_{3,i} | \boldsymbol{\theta}_3) \\
&+ \sum_{i=1}^{n_{123}^{\parallel}} \log \left. \frac{\partial^3 \mathfrak{C}(u_{1,i}, u_{2,i}, u_{3,i} | \boldsymbol{\delta})}{\partial u_1 \partial u_2 \partial u_3} \right|_{u_{j,i} = \lambda_j \bar{F}_j(x_{j,i}^{\parallel 123} | \boldsymbol{\theta}_j), j=1,2,3} \\
&+ \sum_{i=1}^{n_{12}^{\parallel}} \log \left( \frac{\partial^2 \mathfrak{C}(u_{1,i}, u_{2,i}, \infty | \boldsymbol{\delta})}{\partial u_1 \partial u_2} - \frac{\partial^2 \mathfrak{C}(u_{1,i}, u_{2,i}, \lambda_3 | \boldsymbol{\delta})}{\partial u_1 \partial u_2} \right) \Bigg|_{u_{j,i} = \lambda_j \bar{F}_j(x_{j,i}^{\parallel 12} | \boldsymbol{\theta}_j), j=1,2} \\
&+ \sum_{i=1}^{n_{13}^{\parallel}} \log \left( \frac{\partial^2 \mathfrak{C}(u_{1,i}, \infty, u_{3,i} | \boldsymbol{\delta})}{\partial u_1 \partial u_3} - \frac{\partial^2 \mathfrak{C}(u_{1,i}, \lambda_2, u_{3,i} | \boldsymbol{\delta})}{\partial u_1 \partial u_3} \right) \Bigg|_{u_{j,i} = \lambda_j \bar{F}_j(x_{j,i}^{\parallel 13} | \boldsymbol{\theta}_j), j=1,3} \\
&+ \sum_{i=1}^{n_{23}^{\parallel}} \log \left( \frac{\partial^2 \mathfrak{C}(\infty, u_{2,i}, u_{3,i} | \boldsymbol{\delta})}{\partial u_2 \partial u_3} - \frac{\partial^2 \mathfrak{C}(\lambda_1, u_{2,i}, u_{3,i} | \boldsymbol{\delta})}{\partial u_2 \partial u_3} \right) \Bigg|_{u_{j,i} = \lambda_j \bar{F}_j(x_{j,i}^{\parallel 23} | \boldsymbol{\theta}_j), j=2,3} \\
&+ \sum_{i=1}^{n_1^{\perp}} \log \left( 1 + \frac{d\mathfrak{C}(u_1, \lambda_2, \lambda_3 | \boldsymbol{\delta})}{du_1} - \frac{d\mathfrak{C}(u_1, \lambda_2, \infty | \boldsymbol{\delta})}{du_1} - \frac{d\mathfrak{C}(u_1, \infty, \lambda_3 | \boldsymbol{\delta})}{du_1} \right) \Bigg|_{u_{1,i} = \lambda_1 \bar{F}_1(x_{1,i}^{\perp} | \boldsymbol{\theta}_j)} \\
&+ \sum_{i=1}^{n_2^{\perp}} \log \left( 1 + \frac{d\mathfrak{C}(\lambda_1, u_2, \lambda_3 | \boldsymbol{\delta})}{du_2} - \frac{d\mathfrak{C}(\lambda_1, u_2, \infty | \boldsymbol{\delta})}{du_2} - \frac{d\mathfrak{C}(\infty, u_2, \lambda_3 | \boldsymbol{\delta})}{du_2} \right) \Bigg|_{u_{2,i} = \lambda_2 \bar{F}_2(x_{2,i}^{\perp} | \boldsymbol{\theta}_j)} \\
&+ \sum_{i=1}^{n_3^{\perp}} \log \left( 1 + \frac{d\mathfrak{C}(\lambda_1, \lambda_2, u_3 | \boldsymbol{\delta})}{du_3} - \frac{d\mathfrak{C}(\lambda_1, \infty, u_3 | \boldsymbol{\delta})}{du_3} - \frac{d\mathfrak{C}(\infty, \lambda_2, u_3 | \boldsymbol{\delta})}{du_3} \right) \Bigg|_{u_{3,i} = \lambda_3 \bar{F}_3(x_{3,i}^{\perp} | \boldsymbol{\theta}_j)}
\end{aligned} \tag{C.1}$$

where

- $T$  denotes the total length of observation;
- $n_j$  denotes the number of observations in the  $j^{\text{th}}$  marginal process,  $j = 1, 2, 3$ ;
- $\lambda = \lambda_{123}^{\parallel} + \lambda_{12}^{\parallel} + \lambda_{13}^{\parallel} + \lambda_{23}^{\parallel} + \lambda_1^{\perp} + \lambda_2^{\perp} + \lambda_3^{\perp}$ ;

- $\lambda_j$  is the marginal Poisson intensity of the  $j^{th}$  process,  $j = 1, 2, 3$ ;
- $\bar{F}_j(\cdot|\boldsymbol{\theta}_j)$  and  $f_j(\cdot|\boldsymbol{\theta}_j)$  are the survival and probability density function of the  $j^{th}$  process with a parameter vector  $\boldsymbol{\theta}_j$ ,  $j = 1, 2, 3$ ;
- $\mathfrak{C}(u_1, u_2, u_3|\boldsymbol{\delta})$  is the trivariate Lévy copula with a parameter vector  $\boldsymbol{\delta}$ ;
- $x_{j,i}^{\text{type}}$  denotes the  $i^{th}$  jump size of the  $j^{th}$  process of a certain type, where  $\perp$  denotes unique jump and  $\parallel$  denotes common jump,  $j = 1, 2, 3$ ;
- $n_j^\perp$  denotes the number of unique jumps in the  $j^{th}$  marginal process  $j = 1, 2, 3$ .

Parameters are estimated by using the `mle` function, which searches for the parameters that minimise a given negative likelihood function. The benefit of using the `mle` function is that one can easily obtain the parameters errors by using the `vcov` function. In the case of fitting a trivariate process by using a nested Lévy copula, one needs to consider the constraints of parameters of the two Archimedean Lévy copula generators to ensure valid construction of a Lévy copula. Here re-parameterisation is used with the method L-BFGS-B. In particular, the Lévy copula is re-parameterised such that the parameters of a Lévy copula are  $(\theta_2, \Delta)$  instead of  $(\theta_1, \theta_2)$ , where

$$\Delta = \delta_1 - \delta_2. \tag{C.2}$$

Then the constraint  $\Delta \geq 0$  is applied which ensures that the sufficient conditions for a valid Lévy copula are always satisfied. Please note that one needs to add an upper bound and a lower bound for all parameters that are involved.

In situations where one cannot find simple expressions for the constraints or the L-BFGS-B method cannot successfully minimise the negative likelihood function, one can use the BFGS method to estimate the parameters. The second situation happens in this study where a trivariate nested AI-AI Lévy copula model is adopted with the couple ‘contents’ and ‘profits’ being nested first. Since the BFGS method does not allow for explicit incorporation of parameters constraints, one has to manually check whether the estimated parameters satisfy the constraints when this method is selected. One can refer to Table 2 for constraints of some trivariate nested Archimedean copulas.

It is worth mentioning that one can also use other optimisation functions in R. For example, one can consider the function `constrOptim` which allows for linear constraints therefore re-parameterisation is not necessary. However, one needs to estimate the parameter errors separately.

## References

- Aas, K., Berg, D., 2009. Models for construction of multivariate dependence  $\hat{\omega}\hat{\mathbb{I}}\hat{\otimes}$  a comparison study. *The European Journal of Finance* 15 (7–8), 639–659.
- Avanzi, B., Cassar, L. C., Wong, B., 2011. Modelling dependence in insurance claims processes with Lévy copulas. *ASTIN Bulletin* 41 (2), 575–609.
- Barndorff-Nielsen, O., Lindner, A., 2007. Lévy copulas: Dynamics and transforms of Upsilon type. *Scandinavian Journal of Statistics* 34, 298–316.
- Barndorff-Nielsen, O. E., Lindner, A. M., 2004. Some aspects of Lévy copulas. Tech. Rep. 388, Sonderforschungsbereich 386, Munich.
- Bäuerle, N., Blatter, A., 2011. Optimal control and dependence modeling of insurance portfolios with Lévy dynamics. *Insurance: Mathematics and Economics* 48, 398–405.
- Bertoin, J., 1998. Lévy Processes. *Cambridge Tracts in Mathematics*. Cambridge University Press, Cambridge, UK.
- Biagini, F., Ulmer, S., 2009. Asymptotics for operational risk quantified with expected shortfall. *ASTIN Bulletin* 39 (2), 735–752.
- Böcker, K., Klüppelberg, C., 2008. Modeling and measuring multivariate operational risk with Lévy copulas. *The Journal of Operational Risk* 3(2), 3–27.
- Böcker, K., Klüppelberg, C., 2010. Multivariate models for operational risk. *Quantitative Finance* 1, 1–15.
- Bregman, Y., Klüppelberg, C., 2005. Ruin estimation in multivariate models with Clayton dependence structure. *Scandinavian Actuarial Journal* 2005 (6), 462–480.
- Bücher, A., Vetter, M., 2013. Nonparametric inference on lévy measures and copulas. *The Annals of Statistics* 41, 1485–1515.
- Cassar, L. C., 2010. Dependence modelling in multivariate compound Poisson processes with Lévy copulas., Honours Thesis, School of Risk and Actuarial Studies, UNSW Australia Business School.

- Cont, R., Tankov, P., 2004. Financial Modelling With Jump Processes. Chapman & Hall/CRC, London.
- Eder, I., Klüppelberg, C., 2009. The quintuple law for sums of dependent Lévy processes. *The Annals of Applied Probability* 19 (6), 2047–2079.
- Esmaili, H., Klüppelberg, C., 2010. Parameter estimation of a bivariate compound Poisson process. *Insurance: Mathematics and Economics* 47 (2), 224–233.
- Esmaili, H., Klüppelberg, C., 2011. Parametric estimation of a bivariate stable Lévy process. *Journal of Multivariate Analysis* 102, 918–930.
- Esmaili, H., Klüppelberg, C., 2013. Two-step estimation of a multi-variate Lévy process. *Journal of Time Series Analysis* 34, 668–690.
- Farkas, W., Reich, N., Schwab, C., May 2006. Anisotropic Stable Lévy Copula Processes - Analytical and Numerical Aspects. Tech. Rep. 2006-08, Eidgenössische Technische Hochschule, Zurich.
- Frees, E. W., Valdez, E. A., 1998. Understanding relationships using copulas. *North American Actuarial Journal* 2(1), 1–25.
- Grothe, O., Hofert, M., 2015. Construction and sampling of archimedean and nested archimedean Lévy copulas. *Journal of Multivariate Analysis* 138, 182–198.
- Grothe, O., Nicklas, S., 2013. Vine constructions of Lévy copulas. *Journal of Multivariate Analysis* 119, 1–15.
- Hofert, M., 2008. Sampling Archimedean copulas. *Computational Statistics and Data Analysis* 12, 5163–5174.
- Joe, H., 1997. *Multivariate Models and Dependence Concepts*. Chapman & Hall, London.
- Kallsen, J., Tankov, P., August 2006. Characterisation of dependence of multidimensional Lévy processes using Lévy copulas. *Journal of Multivariate Analysis* 97 (7), 1551–1572.
- Kurowicka, D., Joe, H., 2011. *Dependence Modeling Vine Copula Handbook*.
- McNeil, A. J., 2008. Sampling nested Archimedean copulas. *Journal of Statistical Computation and Simulation* 78, 567–581.
- McNeil, A. J., Frey, R., Embrechts, P., 2005. *Quantitative risk management: Concepts, techniques and tools*. Princeton University Press.
- McNeil, A. J., Nešlehová, 2009. Multivariate archimedean copulas,  $d$ -monotone functions and  $l_1$ -norm symmetric distributions. *Annals of Statistics* 37(5B), 3059–3097.
- Mikosch, T., 2006. *Non-Life Insurance Mathematics: An Introduction with Stochastic Processes*. Springer.
- Nadarajah, S., Bakar, S., 2014. New composite models for the danish fire insurance data. *Scandinavian Actuarial Journal* 2014, 180–187.
- Nelsen, R. B., 1999. *An introduction to copulas*. Springer.
- Ross, S., 1983. *Introduction to Stochastic Dynamic Programming*. Academic Press, San Diego, California.
- Sato, K.-I., 1999. *Lévy processes and infinitely divisible distributions*. Cambridge University Press, Cambridge.
- Tankov, P., 2003. *Dependence structure of spectrally positive multidimensional Lévy processes*.
- Tao, J., 2011. Capturing non-exchangeable dependence in multivariate insurance claims processes with nested Lévy copulas, Honours Thesis, School of Risk and Actuarial Studies, UNSW Australia Business School.