Technische Universität München

Department of Mathematics

Dependence modelling of operational risk with special focus on multivariate compound Poisson processes

Master Thesis

by

Jixuan Wang

Supervisor:	Prof. Dr. Claudia Klüppelberg
Advisor:	Prof. Dr. Matthias Fischer
	(cooperation with the Bayerische Landesbank)
Submission date:	25.06.2018

I hereby declare that this thesis is my own work and that no other sources have been used except those clearly indicated and referenced.

Garching, 25.06.2018

Acknowledgement

On this occasion, I would like to thank Prof. Claudia Klüppelberg and Prof. Matthias Fischer for offering me the opportunity of writing my master thesis at the Chair of Mathematical Statistics of the Technical University of Munich and at the Bayerische Landesbank (BayernLB) simultaneously. Despite his tight working schedule, Prof. Matthias Fischer always took time to advise me and to give constructive suggestions. On the other hand, Prof. Claudia Klüppelberg answered all my questions in great detail and even visited me to share her valuable expertise.

Furthermore, I am grateful for the real-world operational loss data provided by the BayernLB and the patient of Dr. Christian Dietz for thorough discussions on the data characteristics. Last but not least, I would like to thank my good friend and fellow student Christina Zou, with whom I shared the joy in mathematics during the last years and who introduced me into the fascinating topic of operational risk at first.

Abstract

Operational risk measurement has become an important research area for the financial industry in recent years. In order to accurately estimate the required capital reserves as well as to obtain a deeper understanding into this complex risk category, an appropriately specified dependence model for loss incidents attributed to different risk factors and business units is indispensable. Hence the current thesis is dedicated to exploring various proposals for dependence modelling in operational risk, and subsequently focusing on a straightforward to apply, yet flexible enough approach based on compound Poisson processes and Lévy copulas. Similar to the rationale of ordinary copulas, the Lévy measure of a multivariate Lévy process is fully characterised by its marginal components and the associated Lévy copula. Besides an in-depth theoretical treatment of bivariate models, extensive simulation and real application examples are provided.

Contents

Chapter 1

Introduction

Operational risk belongs to one of the three primary risk types encountered in the finance industry. The other two categories are market and credit risk, respectively. The commonly recognised definition of operational risk was first introduced in the second of the Basel Accords[1], which constitute the most important international regulatory framework for financial institutions and are issued by the Basel Committee on Banking Supervision. Accordingly, operational risk is the risk of loss resulting from inadequate or failed internal processes, people and systems or from external events, including legal risk[2] but excluding strategic and reputational risk. This definition is preserved in the latest finalisation of the Basel III framework from December 2017[3].

Despite the increasing attention devoted to operational risk management, the financial sector continues suffering from significant losses due to operational failures. Alone in 2017, for example, the ten largest operational losses worldwide exceeded $11.6 billion as reported by [Ris18a]. The largest loss was caused by fraudulent transactions at the Brazilian development bank totalling $2.52 billion. Secondly, employees at the Shoko Chukin Bank in Japan improperly granted $2.39 billion of loans by falsifying approval documents. In third place, the U.S. Securities and Exchange Commission brought charges against the Woodbridge Group of Companies with running a $1.22 billion Ponzi scheme.

Besides the rare events cited above, operational losses of smaller but still considerable sizes do occur at every financial institute, not least because of the progresses made in financial technology and the increasing complexity and interdependence of operations in the corresponding industry. This concern emphasises the importance of operational risk quantifications and reliable estimation methods for sufficient capital reserves against potential incidents. Hence in the present work we put our focus on dependence modelling for operational losses, which should be part of any solid operational risk model. First, a brief outline of the thesis is provided in Section 1.1, as well as an overview of various approaches for capturing dependence in Section 1.2, before the main approach based on

[1] See [Ban06].

[2] As explained in [Ban06], legal risk includes, but is not limited to, exposure to fines, penalties, or punitive damages resulting from supervisory actions, as well as private settlements.

[3] See [Ban17].

compound Poisson processes and Lévy copulas is discussed in the subsequent chapters.

1.1 Outline of the thesis

The remaining chapters of this thesis are structured as follows. To begin with, Chapter 2 reviews the definition and essential properties of Lévy processes and their associated Lévy measures. Most importantly, a version of Sklar's theorem for Lévy copulas is explained in detail, which builds the theoretical foundation for the multivariate dependence model introduced in Chapter 3. The marginal components of our dependence model are given by the familiar compound Poisson processes originating from actuarial risk theory. In order to demonstrate how Lévy copulas simultaneously shape frequency and severity interdependence among the marginal processes, the features of a bivariate model are explored in detail and illustrated through examples. Going one step further, the estimation of such a dependence model is enabled by specifying the corresponding likelihood function.

As more accurate risk exposure calculations constitute a major incentive of dependence modelling, Chapter 4 is devoted to presenting closed-form risk measure approximations for high confidence levels. The restrictions when generalising univariate results to higher dimensions and some potential extensions are discussed as well. The sensitivity of risk measure estimations towards different model components is addressed in Chapter 5, where the quality of maximum likelihood estimates and various approaches to assessing the goodness of fitted dependence structures are also investigated by means of simulation.

Drawing on real-world loss data, Chapter 6 exemplifies our modelling approach by providing the entire procedure of verifying model assumptions, estimating model parameters and evaluating the reasonableness of obtained models. The possible insufficiency of loss observations within a single financial institute and the incorporation of external loss information are briefly covered as well. Finally, concluding remarks and directions for future research follow in Chapter 7.

1.2 A literature review of dependence modelling

According to the currently effective Basel requirements for the quantitative modelling of operational risk, activities within a financial institute are divided into eight business lines, each of them exposed to seven potential loss event types. For a detailed description of the categorisation we refer to Appendix A. Hence there are 56 different combinations of business line and event type, usually called risk cells. The determination of operational risk capital is built upon the explicit estimation of frequency and severity distributions for losses assigned to each risk cell. This method is often referred to as the loss distribution approach.

Owing to the individual data availability and business organisation, financial institutes may deviate from the matrix of 56 risk cells and adopt a substructure of it for statistical

modelling. Therefore, let d denote the number of risk cells being generally considered. Inspired by the classical actuarial risk theory, the aggregate loss amount up to time t in risk cell $i \in \{1, \ldots, d\}$ is given by a compound stochastic process

$$S_i(t) \;=\; \sum_{k=1}^{N_i(t)} X_{ik}, \quad t \geq 0, \tag{1.1}$$

where $N_i(t)$ denotes the loss frequency process with starting value $N_i(0) = 0$ and X_{ik}, $k \geq 1$, are random positive loss severities from a continuous distribution. The probability of no loss occurring until time t is given by $\mathbb{P}(S_i(t) = 0) = \mathbb{P}(N_i(t) = 0)$. Under the standard assumption, the individual losses X_{ik} within the same risk cell i are i.i.d. with distribution function F_{X_i} satisfying $F_{X_i}(0) = 0$ and are independent from the number of losses $N_i(t)$.

The overall operational loss process of a financial institute is obtained by summing up over all d risk cells, that is,

$$S_+(t) \;=\; \sum_{i=1}^{d} S_i(t), \quad t \geq 0. \tag{1.2}$$

In order to estimate the necessary capital reserves against future losses, the risk measure value at risk (VaR) is typically applied. More precisely, we introduce the following definition.

Definition 1.1 (**Operational VaR**).
Let $G_{i,t}(x) = \mathbb{P}(S_i(t) \leq x)$ denote the distribution function of the aggregate loss $S_i(t)$ in risk cell $i \in \{1, \ldots, d\}$. Then the stand-alone operational VaR of risk cell i until time t at confidence level $\alpha \in (0, 1)$ is the α-quantile of $S_i(t)$ and given by the generalised inverse

$$\mathrm{VaR}_{i,t}(\alpha) \;=\; G_{i,t}^{\leftarrow}(\alpha) \;=\; \inf\left\{x \in \mathbb{R} \,|\, G_{i,t}(x) \geq \alpha\right\}. \tag{1.3}$$

Accordingly, the distribution function of $S_+(t)$ is denoted by $G_{+,t}(x)$ and the overall operational VaR of a financial institution until time t at level $\alpha \in (0, 1)$ is defined as

$$\mathrm{VaR}_{+,t}(\alpha) \;=\; G_{+,t}^{\leftarrow}(\alpha) \;=\; \inf\left\{x \in \mathbb{R} \,|\, G_{+,t}(x) \geq \alpha\right\}. \tag{1.4}$$

The standard risk measure specified by the Basel Committee is the VaR at level 99.9% for a one-year holding period[4]. In other words, the value of $\mathrm{VaR}_{+,1}(99.9\%)$ is to calculate, when assuming the time scaling $t = 1$ corresponds to one calender year. However, even the compound distribution $G_{i,1}$ of a single risk cell generally does not possess a closed-form expression, let alone the distribution $G_{+,1}$ of the overall loss process, which further involves the dependence structure among the d risk cells. For this reason, financial institutions are requested to add up the stand-alone measures $\mathrm{VaR}_{i,1}(99.9\%)$, $i \in \{1, \ldots, d\}$, for

[4]Paragraph 667 of [Ban06] states "... Whatever approach is used, a bank must demonstrate that its operational risk measure meets a soundness standard comparable to that of the internal ratings-based approach for credit risk, (i.e. comparable to a one year holding period and a 99.9th percentile confidence interval)."

calculating the overall capital reserve, unless they can provide a well-founded dependence model for the risk cells[5].

The provision of simple accumulation may be due to the fact that for any subadditive risk measure the summation of all stand-alone measures represents an upper bound for the same risk measure directly applied to the sum S_+. However, it is well-known that VaR lacks subadditivity and its potential superadditivity is particularly pronounced in case of heavy-tailed severity distributions which are commonly encountered in operational risk context. Theoretical explanations for the latter can be found in [BK08] and [CEN06], as well as for empirical evidences we refer to [CA08], [GFA09] and [MPY11].

Moreover, the assumption of the equivalence between VaR_+ and $\sum_{i=1}^{d} \mathrm{VaR}_i$ corresponds to the implicit adoption of perfectly positive dependence among the aggregate losses $S_1, \ldots, S_d$, as been proved in Proposition 7.20 in [MFE15], for instance. In contrast, empirical studies show that the dependence between aggregate losses is generally rather weak. For example, [CA08] examines international operational losses collected by the ORX[6] consortium and finds Kendall's rank correlations among the losses, aggregated either at business line or at event type level, commonly less than 0.2 and rarely exceeding 0.4. Similarly, study of the Italian DIPO[7] database by [BCP14] results in empirical Kendall's τ values ranging from -0.14 to 0.30. Further examples can be found in [Cha+04], [FVG08], [Gia+08] and [GFA09], where the authors study anonymised loss data from individual banks.

In summary, the simple addition of stand-alone VaRs could either over- or underestimate the true overall risk exposure and the comonotonic scenario among different risk cells is seen unjustified in reality. Therefore, a strong incentive to explicitly model dependence structures arises and a fruitful research on this issue emerges both in academia and practice. The latter gives the ground for the literature review in the current section.

At this point it should also be noticed that a new non-model based method, the standardised approach, is introduced in the recently published finalisation of the Basel III framework[8]. From the 1st of January 2022 on, the standardised approach shall replace all existing methodologies for measuring minimum operational risk capital requirements under Pillar I of the Basel standards. The new approach is supposed to improve the comparability and simplicity of operational risk capital calculations. On the other hand, concerns have been raised that a non-model based approach cannot sufficiently respect the complexity and firm-specific characteristics of operational losses and hence lacks risk sensitivity, for example as reported in [Coo18] and [Ris18b]. As a result, it is expected

[5]Paragraph 669(d) of [Ban06] states "Risk measures for different operational risk estimates must be added for purposes of calculating the regulatory minimum capital requirement. However, the bank may be permitted to use internally determined correlations in operational risk losses across individual operational risk estimates, provided it can demonstrate to the satisfaction of the national supervisor that its systems for determining correlations are sound, implemented with integrity, and take into account the uncertainty surrounding any such correlation estimates ...."

[6]Operational Riskdata eXchange Association.

[7]Database Italiano delle Perdite Operative.

[8]See [Ban17].

that sophisticated approaches based on mathematical modelling would retain their importance as well as be employed for assessing economic capital and Pillar II capital support. Furthermore, a reasonable dependence model should not only contribute to an accurate assessment of regulatory capital, but also improve the understanding of the overall operational risk structure within financial institutions and support risk management procedures.

The objective of the subsequent sections is to explore different approaches of relaxing the perfect dependence assumption among the d risk cells as well as the independence assumption of loss counts and loss sizes within a single risk cell. We do not strive for a full treatment of all possible dependence models, as that would fill a separate textbook. Instead we highlight the state-of-the-art techniques and summarise some practical experiences with operational loss data. In order to give the overview a clear structure, we subdivide all dependence concepts as follows:

(1) inter-cell dependence based on frequencies $N_1(t), \ldots, N_d(t)$,

(2) inter-cell dependence based on severities $X_{1k}, \ldots, X_{dk}$,

(3) inter-cell dependence based on aggregate losses $S_1(t), \ldots, S_d(t)$,

(4) inter-cell dependence based on frequencies $N_1(t), \ldots, N_d(t)$ and on severities $X_{1k}, \ldots, X_{dk}$,

(5) intra-cell dependence introduced between frequency $N_i(t)$ and severities X_{ik} for risk cell $i \in \{1, \ldots, d\}$,

(6) both inter- and intra-cell dependence.

For notational convenience, whenever the observation time horizon is regarded as fixed, for example at $t = 1$, we may omit the time index t from the notations introduced in (1.1)-(1.4). From a mathematical perspective, the total loss process of risk cell $i \in \{1, \ldots, d\}$ then reduces to a compound random variable S_i, represented as a sum of N_i random single losses.

1.2.1 Inter-cell frequency dependence

One of the most popular methods is to characterise the dependence structure among the frequencies of different risk cells via parametric copulas. Let $C_N : [0,1]^d \to [0,1]$ be a d-variate copula and let F_{N_i} denote the distribution function of the loss frequency N_i in risk cell $i \in \{1, \ldots d\}$. Then a joint distribution F_N of $N = (N_1, \ldots, N_d)^\top$ can be constructed by

$$F_N(n_1, \ldots, n_d) \;=\; C_N\left(F_{N_1}(n_1), \ldots, F_{N_d}(n_d)\right), \quad (n_1, \ldots, n_d)^\top \in \mathbb{N}_0^d. \tag{1.5}$$

Commonly utilised candidates for the marginal distribution F_{N_i} are the Poisson distribution and the negative binomial distribution, where the latter can be seen as the randomisation of the Poisson parameter through a gamma distribution and hence accounts for

over-dispersion. However, with regard to VaR estimations, the difference between Poisson and negative binomial distributed frequencies can be negligible as theoretically shown in [BK05] as well as empirically observed by [AK07] and [Val09].

The theoretical foundation of (1.5) is provided by the well-known Sklar's theorem. Since the copula C_N can be chosen arbitrarily, the current approach allows for both positive and negative dependence among $N_1, \ldots, N_d$. Nevertheless, we would like to mention the dependence structure of $N_1, \ldots, N_d$ is not solely determined by the copula, which follows from the non-uniqueness of copula for discrete random variables. Consequently, drawing inference for the parameters of the copula C_N could be tricky.

In order to circumvent the above difficulty, [SV14] employs the idea of "jittering" for modelling multivariate insurance claim numbers, which can be readily applied in operational risk context as well. The discrete frequencies $N_1, \ldots, N_d$ are jittered by subtracting an independent standard uniform random variable from each of them, such that the usual maximum likelihood inference for continuous distributions can be carried out. Rank-based dependence measures, such as Kendall's τ, are preserved within the jittering procedure.

Another variation of frequency dependence modelling via copulas is proposed by [WSZ16], in which mutual information from the entropy framework is utilised as the correlation parameters for a Gaussian copula. The mutual information $I(N_i, N_j)$ between two random variables N_i and N_j, $i \neq j$, measures the information of N_i contained in N_j and vice versa. It is symmetric among its two arguments and can be calculated through

$$\begin{aligned} I(N_i, N_j) &= H(N_j) - H(N_j|N_i) = H(N_i) - H(N_i|N_j) \\ &= H(N_i) + H(N_j) - H(N_i, N_j), \end{aligned}$$

where $H(N_i)$ and $H(N_j)$ denote the entropy of N_i and N_j, respectively, $H(N_j|N_i)$ and $H(N_i|N_j)$ are the conditional entropy, and $H(N_i, N_j)$ is the joint entropy. The value of $I(N_i, N_j)$ is always non-negative and equals to zero if and only if N_i and N_j are independent. Hence the mutual information between N_i and N_j can be considered as a measure of dependence between these variables. As the correlation parameters in a Gaussian copula have to lie in the interval $[0, 1]$, the global correlation coefficient is introduced as a standardised version of mutual information and it is given by

$$\rho_{ij}^I = \sqrt{1 - \exp[-2I(N_i, N_j)]}. \tag{1.6}$$

The authors of [WSZ16] apply the above method to calculate the operational risk capital charge for the overall Chinese banking industry.

An alternative to utilising copulas is to directly specify the joint distribution of N as a d-variate mixed Poisson distribution. More precisely, the authors of [Bad+14] and [Tan16] adopt a multivariate Erlang mixture with a common scale parameter as the mixing distribution, such that the random vector N follows a d-variate Pascal mixture distribution, that is, a negative binomial distribution with a positive integer shape parameter. All parameters are estimated by an expectation maximisation (EM) algorithm, whose M-step converges to a unique global maximum and is supposed to outperform copula-based estimations in high dimension.

In addition, the issue of left-truncated severities is addressed, as often only losses exceeding certain recording thresholds $c_1, \ldots, c_d$ are collected in practice. If the loss frequency of risk cell i is redefined as $N_i^{\text{rec}} = \sum_{k=1}^{Ni} \mathbb{1}_{\{X_{ik} > c_i\}}$, then the joint distribution of $N_1^{\text{rec}}, \ldots, N_d^{\text{rec}}$ still belongs to the class of multivariate Pascal mixtures with modified scale parameters. Moreover, in case that loss severities are discretised and satisfy the standard independent assumption as detailed after (1.1), the joint distribution of the aggregate losses $S_1, \ldots, S_d$ constitutes a compound negative binomial distribution and possesses a closed-form expression. Hence VaR calculations can be carried out through Panjer's recursion instead of Monte Carlo simulation. As numerical illustration, the above procedure is applied to the operational loss data of a North American financial institution comprising eight risk cells.

Another adaptation of Poisson mixtures to characterising multivariate loss frequencies is proposed in the lecture notes [Sch17] about an extension of the CreditRisk$^+$ framework. Interestingly, the industry model CreditRisk$^+$ from the world of credit risk management actually stems from actuarial mathematics and is now in turn utilised to analyse operational risk, whose basic model assumptions are also based upon actuarial risk theory as already indicated. More specifically, obligors and non-idiosyncratic risk factors from the extended CreditRisk$^+$ model are interpreted in the operational risk setting as business lines and event types, respectively. Furthermore, the evaluation of compound loss distributions and VaRs can be achieved via a variation of Panjer's recursion.

Besides specifying the joint distribution of $N_1, \ldots, N_d$ either directly or via copulas, intercell frequency dependence can also be replicated through a common shock structure. More precisely, losses of different risk cells are considered to be related to a series of underlying independent common shocks, such as electric failures, internal miscommunications or cybersecurity breaches. In particular, consider m independent Poisson random variables $\tilde{N}_1, \ldots, \tilde{N}_m$ with positive rate parameters $\tilde{\lambda}_1, \ldots, \tilde{\lambda}_m$, respectively. Each of these random variables represents an underlying process which can be assigned to one or more risk cells and the assignment is recorded in the indicator variables

$$\delta_{ij} = \begin{cases} 1, & \text{shock } j \text{ has an impact on cell } i, \\ 0, & \text{otherwise,} \end{cases} \qquad i \in \{1, \ldots, d\},\, j \in \{1, \ldots, m\}.$$

Then the observable frequency N_i of risk cell i has the expression

$$N_i = \sum_{j=1}^{m} \delta_{ij} \tilde{N}_j,$$

and is also Poisson distributed with mean $\lambda_i = \sum_{j=1}^{m} \delta_{ij} \tilde{\lambda}_j$. Note that only positive correlations between frequencies can be captured through this approach and an empirical support is provided by [FRS04], in which the authors observed a high number of external fraud events in case of increasing occurrence of internal fraud events.

In terms of parameter estimation, [PRT02] suggests a two-step procedure. First, the parameter λ_i of the observable frequency N_i, $i \in \{1, \ldots, d\}$, is estimated by its empirical mean, which is equivalent to the maximum likelihood estimate (MLE) in the current Poisson case. In the second step, the underlying intensities $\tilde{\lambda}_j$, $j \in \{1, \ldots, m\}$, are computed

as the solution of a constrained quadratic optimisation problem. The objective function is defined as the difference between the empirical and the theoretical covariance matrices in Frobenius norm, under the constraints of non-negative Poisson parameters and matching with the estimators from the first step. The property of equal mean and variance of a Poisson distribution is essential for the formulation of the optimisation problem.

A more flexible dependence structure is obtained through replacing the indicator variables by Bernoulli random variables $B_{ij} \sim \mathrm{Ber}(p_{ij})$, $i \in \{1, \ldots, d\}$, $j \in \{1, \ldots, m\}$. Then common shock j causes with probability $p_{ij} \in [0, 1]$ a loss in risk cell j. Furthermore, the authors of [LM03] advocate taking into account dependent severities caused by the same common shock in a similar manner.

We conclude this section about inter-cell frequency dependence by a brief discussion of its influence on the implied dependence strength among the compound losses $S_1, \ldots, S_d$. Different investigations of real loss data, for example of a French bank by [FRS04], a German bank by [AK07] and an Italian bank by [Bee05], show that even with strong frequency correlations the implied correlations between $S_1, \ldots, S_d$ are rather weak, as long as loss severities are assumed being independent. This phenomenon is argued to be particularly true for heavy-tailed severity distributions, which are commonly encountered in operational risk and dominate any frequency dependence structures. Of course, this observation also has an important implication for the overall risk measure VaR_+, and its value is expected to resemble the case of independent compound losses $S_1, \ldots, S_d$ despite potentially varying frequency correlations.

1.2.2 Inter-cell severity dependence

Obviously, a dependence structure among the single loss sizes from different risk cells can also be calibrated by means of parametric copulas. A real-life application is provided by [GH12], in which the authors in particular use pair-copula constructions to estimate the capital requirement for the French semi-cooperative banking group Caisse d'Epargne based upon its historical loss data. Both nested Archimedean and D-vine architectures are fitted to the ten risk cells being considered, whereas the bivariate building blocks are chosen from the Gumbel, Clayton, Frank, Galambos, Husler-Reiss and Tawn copulas. All compound loss distributions are built as a convolution with Poisson frequency, although the authors have also tested alternative frequency distributions such as the binomial and the negative binomial ones, and conclude the VaR estimates are insensitive to the choice of frequency distributions.

In the above example, the method of semi-parametric pseudo maximum likelihood estimation (PMLE) is employed to fit the copula parameters. In order to clarify terms, we briefly recall the three most common methods for copula estimation, as this also plays a relevant role in the subsequent sections.

The first method is of course the classical MLE, whereby the joint density is maximised simultaneously with respect to both the copula and the marginal distribution parameters.

Hence this method is often referred to as the full parametric MLE and presents the computationally most expensive one. In order to reduce computational complexity, especially in higher dimensions, the next two approaches both rely on the idea of separating the margins from the copula estimation. Depending on how the marginal distributions are treated, one differentiates between the inference function for margins (IFM) technique and the aforementioned PMLE.

The IFM is often called stepwise parametric as the marginal distributions and the copula are estimated parametrically in two successive steps. Firstly, the parameters of the margins are estimated via MLE. Then the marginal parameters are considered as fixed and plugged into the joint likelihood of the copula and the margins, which is maximised solely with respect to the copula parameter in the second step. Equivalently, the second step can be interpreted as maximising the copula density based on the so-called pseudo copula data, which are obtained through applying the estimated marginal distribution functions to the original observed loss data.

In contrast, the PMLE is called semi-parametric, as empirical marginal distribution functions are computed in the first stage and utilised to transform original data into pseudo copula data in the second stage. One rationale for this procedure is to avoid potential parametric restrictions on the margins when estimating the dependence structure, which is of course only sensible if sufficient loss data are available to ensure a good approximation through empirical distribution functions. Under mild regularity conditions on the copula family, the copula parameter estimate is shown by [GGR95] to be asymptotically normal.

1.2.3 Inter-cell aggregate loss dependence

As discussed at the end of Section 1.2.1, pure frequency dependence modelling may only result in a very limited range of aggregate loss dependence, hence another popular approach is to directly consider a dependence structure at the level of the compound losses $S_1, \ldots, S_d$.

A straightforward way for this purpose is again by means of parametric copulas. As before, let G_i, $i \in \{1, \ldots, d\}$, denote the distribution function of the compound loss in risk cell i, and let $C_S : [0,1]^d \to [0,1]$ be a d-variate copula. Then the expression

$$G(x_1, \ldots, x_d) \;=\; C_S\left(G_1(x_1), \ldots, G_d(x_d)\right), \quad (x_1, \ldots, x_d)^\top \in \mathbb{R}_+^d,$$

specifies a joint distribution G of $S = (S_1, \ldots, S_d)^\top$. Note that the marginal distributions G_i, $i \in \{1, \ldots, d\}$, are calculated by compounding the severity distribution via the frequency of the corresponding cell, whereas the standard independent assumptions following (1.1) hold. Common choices for fitting loss frequency include the Poisson, the negative binomial and the geometric distributions. In order to take rare loss occurrence in certain risk cells into account, a zero-inflated version of the aforementioned distributions can be considered. With respect to loss severity, the gamma distribution, the Weibull distribution, the lognormal distribution, the Pareto distribution as well as the generalised Pareto

distribution (GPD) are widely employed.

As already stated, the compound distributions usually do not have a closed-form expression and have to be accessed via recursion or simulation. Furthermore, in order to obtain a sufficiently large sample for copula parameter estimation, loss data are often aggregated on a quarterly or monthly basis, although the VaR estimate with respect to the annual loss amount is of primary interest for capital reserves. The implicit assumption made here is the dependence structure over a one-year time horizon corresponds to that of shorter periods.

The current copula approach is followed by many literature sources and we summarise below some variations worthy of mentioning. Instead of fitting a plain distribution to loss severities, [CR04] and [GFA09] utilise a spliced distribution with lognormal body and GPD tail. The theoretical foundation to this originates from extreme value theory (EVT), in which the well-known Pickands-Balkema-de Haan theorem ensures that the excess distribution over a high threshold can be well approximated through a GPD for all commonly encountered distribution functions. Similarly, the authors of [Gia+08] use a variation of heavy-tailed α-stable distributions to model the body of loss severities and both symmetric and skewed Student's t copulas to model dependence among the compound losses.

The application of EVT is further elaborated by [ABF12] such that the upper tail of a t copula is substituted by the upper tail of a multivariate GPD copula in a continuous way. The result constitutes a well-defined copula which is supposed to capture the heavy-tailed nature of operational losses more adequately. The authors exemplify their approach by an analysis of the SAS OpRisk Global Data, which is an external database containing worldwide publicly reported operational losses. For ease of model calibration, the thresholds chosen for the marginal spliced distributions are also used for estimating the spliced copula.

An alternative to maximum likelihood based methods is suggested in [Ang+09], where an EM algorithm is employed for frequency and severity parameter estimation in case of left truncated loss data. An empirical illustration is provided by evaluating the external dataset from the company OpVantage, in which losses exceeding $1 million are collected from public sources. Apart from this, the authors of [Val09] adopt a Bayesian model for analysing the losses of an anonymous bank, as they argue Bayesian statistics are in particular suitable when dealing with scarce operational risk data and incorporating prior information brought by experts. Parameters of both the marginal distributions as well as the Gaussian and Student's t copulas are computed via Markov chain Monte Carlo (MCMC) methods. A further utilisation of Gaussian and t copulas is incorporated by [PG09] into a graphical model. More precisely, each node in the graph represents the random total loss for a combination of business line and event type. The joint distribution of nodes within a connected subgraph is then formed via a copula whereas the interdependence between connected subgraphs is subject to hyper Markov properties.

On the other hand, the authors of [BCP14] extend the current copula approach by explicitly modelling potential zero observations in certain risk cells, as the non-occurrence of

losses should also convey information about dependence characteristics. For this purpose, a Bernoulli random variable B_i is introduced for each risk cell $i \in \{1, \dots, d\}$ and has the interpretation

$$B_i = \begin{cases} 1, & \text{if no loss occurs in cell } i, \\ 0, & \text{otherwise.} \end{cases}$$

If S_i^+ denotes the strictly positive and continuous part of the total loss S_i in risk cell i, then the total loss can be expressed as $S_i = (1 - B_i)S_i^+ \geq 0$. Additionally, let p_B denote the multivariate probability mass function of $B = (B_1, \dots, B_d)^\top$ and let $b = (b_1, \dots, b_d)^\top$ be a realisation of B. Then we introduce $D(b) = \{i \in \{1, \dots, d\} \mid b_i = 0\}$ as the set of all indices, for which the corresponding component of b is equal to 0. The $|D(b)|$-dimensional density of $\{S_i^+ \mid i \in D(b)\}$, is denoted by $g_{\{S_i^+ \mid i \in D(b)\}}$. By assuming the non-occurrence of losses is independent from the distribution of strictly positive losses, the joint density of the total losses $S_1, \dots, S_d$ can be written as

$$g_{B,S}(b, x) = p_B(b) g_{S|B}(x|b) = p_B(b) g_{\{S_i^+ \mid i \in D(b)\}}(x_i, i \in D(b)), \quad b \in \{0,1\}^d,\ x \in \mathbb{R}_+^d.$$

In this way, the dependence modelling of zero losses is separated from the dependence modelling of strictly positive losses. Whereas the latter has already been extensively discussed in the current section, in principle any d-variate copula can also be used to calibrate the joint distribution of B. However, the authors recommend elliptical copulas for reason of computational efficiency and exemplify the proposed model by analysing the aforementioned DIPO data from Italian banks.

Another analysis of the DIPO database is conducted in [MPY11] and the authors examine the impact of different fitted copulas on the overall risk capital estimate $\mathrm{VaR}_{+,1}(99.9\%)$. The dependence structure is calibrated based upon losses aggregated monthly and according to event type. Considered copulas include the elliptical as well as the Archimedean families. Simulation-based $\mathrm{VaR}_{+,1}(99.9\%)$ estimates under a copula model are found to be up to 30% higher than the value obtained through simply adding up the stand-alone estimates $\mathrm{VaR}_{i,1}(99.9\%)$, $i \in \{1, \dots, d\}$. Nonetheless, the authors argue that the observed increase in VaR estimates is attributed to two reasons, that is, the potential superadditivity of the VaR measure on the one hand and the possibly insufficient number of Monte Carlo iterations on the other hand. In order to disentangle the two effects to a certain degree, theoretical asymptotic bounds for VaRs are computed under the assumption of different underlying copulas and any resulting estimates outside the bounds should be caused by the simulation setting and are discarded.

To conclude the copula approach to modelling aggregate loss dependence, we would like to mention a few more literature references including [Cha+08], [EP08] and [Li+14a], as well as the observation described in [FVG08] and [Val09], that heavy-tailed marginal severity distributions have a much larger impact on the VaR estimation outcome than the specific choice of copula. Furthermore, differences between the Poisson and the negative binomial distribution for frequency calibration are found to be insignificant.

Without utilising copulas, the authors of [Li+14b] combine the variance-covariance approach, known from market risk management, with the concept of mutual information

in order to assess the operational risk capital for the Chinese banking industry. The suggested procedure consists of two stages. First, the stand-alone measures $\mathrm{VaR}_{i,1}(99.9\%)$, $i \in \{1, \ldots, d\}$, are estimated based on simulated annual losses. Second, the linear correlation coefficients in the common variance-covariance method are replaced by the global correlation coefficients introduced in (1.6) and the overall VaR estimate is calculated through

$$\mathrm{VaR}_{+,1}(99.9\%) = \sqrt{\sum_{i=1}^{d}\sum_{j=1}^{d} \mathrm{VaR}_{i,1}(99.9\%)\rho_{ij}^{I}\mathrm{VaR}_{j,1}(99.9\%)}.$$

As the global correlation coefficients are supposed to capture both linear and non-linear dependence across risk cells, they are considered to be superior to their liner counterparts. The authors also argue that the simple adaptation of linear correlation may lead to underestimation of VaR values.

1.2.4 Inter-cell frequency and severity dependence

The proposal of [EP08] is to model inter-cell dependence in frequency and in severity both via parametric copulas, respectively. In addition, the authors discuss the issue of possibly different resulting VaR estimations caused by differently designed risk cells, for example, either aggregated across business lines or across event types. By means of simulation they conclude that the discrepancy of the $\mathrm{VaR}_{+,1}(99.9\%)$ estimates is more sensitive to the interdependence among severities than to the interdependence among frequencies, and generally decreases with increasing dependence governed by the fitted copulas. Moreover, the Gaussian copula is found to yield a reduction of all quantile estimates compared to the Gumbel copula which allows for asymptotic upper tail dependence.

Alternatively, the joint distribution of the k-th severities from different cells can be specified via a mixed distribution instead of copulas, as this was already presented for the pure frequency dependence modelling through a Poisson mixture in Section 1.2.1. Following the idea of [Res08], the marginal severities comply with exponential distributions sharing a gamma distributed random variable as parameter, such that the joint severity has a multivariate Pareto distribution. Furthermore, frequency and severities within one cell are still assumed to be independent and the joint frequency follows a multivariate negative binomial distribution as the result of a Poisson mixture also with gamma distributed parameters.

By utilising the notion of point processes, one can explicitly characterise dependence between the k-th severities, between the k-th event inter-arrival times or between the k-th event times of different risk cells. Clearly, for this purpose the time component in the compound sum expression (1.1) is assumed to progress in a continuous manner. Following the approach in [CEN06], the frequency process $N_i(t)$ of risk cell $i \in \{1, \ldots, d\}$ with rate parameter $\lambda_i > 0$ is formulated as a Poisson point process. Given a fixed time interval $[0, T]$ and a risk cell i, let N_i^T be a Poisson random variable with mean $\lambda_i T$ and independent from i.i.d. random variables T_{ik}, $k \geq 1$, distributed according to the uniform distribution

on $[0, T]$. Then the frequency process of risk cell i can be written as the random sum $N_i(t) = \sum_{k=1}^{N_i^T} \mathbb{1}_{\{T_{ik} \leq t\}}$ for $t \in [0, T]$.

Hence the random variables T_{ik}, $k \geq 1$, precisely correspond to the loss arrival times of cell i and two kinds of elementary dependence structures under the current model setting are the following. On the one hand, the joint distribution of the arrival times $T_{1k}, \ldots, T_{dk}$ can be specified via a d-variate copula. This construction is interpreted as the presence of a common underlying effect causing losses in different risk cells at different times. On the other hand, a copula dependence structure can be imposed among the total counts of losses $N_1^T, \ldots, N_d^T$ in the interval $[0, T]$. These two constructions are exemplified in [CEN06] with a Frank copula which allows for both positive and negative dependence. In addition, the above two construction methods can be combined with each other via superposition and thinning of different Poisson processes. In order to also take dependence between loss severities into account, the random variables T_{ik} are extended to 2-dimensional random vectors $(T_{ik}, X_{ik})^\top$ for $k \geq 1$ and $i \in \{1, \ldots, d\}$.

1.2.5 Intra-cell dependence

In the current section the risk cell index $i \in \{1, \ldots, d\}$ is suppressed, as we solely consider dependence characterisations within one risk cell whose model components constitute the compound sum expression given by (1.1).

In the appendix of [FRS04], a simple concept is adopted for modelling the dependence between loss frequency N and loss severities X_k, $k \geq 1$. The Poisson distribution is chosen for frequency and the lognormal distribution for severity. Furthermore, let $(\mu, \sigma^2)^\top$ denote the lognormal parameters and let λ be the Poisson parameter estimated under the standard independence assumption between N and X_k, $k \geq 1$. Next, the independence assumption is relaxed by introducing a weight parameter $c \in [0, 1]$ which represents the proportion of the mean and the variance of the logarithm of X_k explained by N. More precisely, the conditional distribution of X_k given N is specified through a lognormal distribution with logarithmic mean $\mu(N) = (1 - c)\mu + c\frac{\mu}{\lambda}N$ and standard deviation $\sigma^2(N) = (1-c)\sigma^2 + c\frac{\sigma^2}{\lambda}N$. Hence conditional on N, the severities X_k, $k \geq 1$, are independent, whereas the parameter c controls the dependence strength between frequency and severity.

On the other hand, the authors of [GCX17] directly model the parameters μ, σ^2 and λ as random variables. Then a dependence structure is imposed by fitting either a 3-dimensional Gaussian or Student's t copula to the random distribution parameters. The authors illustrate their model calibration and VaR estimation procedure by using simulated and publicly available financial market losses, which contain remarkably more data points than common operational risk datasets. Hence the performance of the proposed methodology in operational risk context is subject to further research.

As mentioned previously, a compound random sum of form (1.1) is one of the prime components appearing in actuarial models. Therefore, below we also include several de-

pendence concepts which are originally proposed for modelling non-life insurance claims and have potential application possibilities for modelling operational risk losses. A further reason for this treatment is that the dependence between loss frequency and loss severity in operational risk management has been by far not as extensively studied as in non-life insurance context.

The first approach we consider is a joint regression model whose margins are given by univariate generalised linear models (GLM) and then linked together via a copula. As suggested in [Cza+12], the loss number N is characterised by a Poisson GLM

$$N \sim \mathrm{Poi}(\lambda) \quad \text{with} \quad \ln(\lambda) = \ln(e) + z^\top \alpha,$$

and the average loss size $\overline{X} = \frac{1}{N}\sum_{k=1}^{N} X_k$ by a gamma GLM

$$\overline{X} \sim \mathrm{Gamma}(\mu, \nu^2) \quad \text{with} \quad \ln(\mu) = z^\top \beta,$$

where z denotes the covariate vector, the offset e gives the known time length in which loss events occur, and the parameter ν is assumed to be known as well. A bivariate Gaussian copula with a single correlation parameter ρ is used to reflect the dependence between $\overline{X}$ and N. The unknown parameters α, β and ρ are estimated through a maximisation by parts (MBP) algorithm which is originally developed by [SFK05]. More precisely, the log-likelihood function is decomposed as $l(\alpha, \beta, \rho) = l_m(\alpha, \beta) + l_c(\alpha, \beta, \rho)$. The first summand l_m is independent of ρ and its maximisation provides an initial estimate for the marginal parameters $(\alpha, \beta)^\top$. Then the second summand l_c is used to estimate the copula parameter ρ as well as to update the estimate for $(\alpha, \beta)^\top$. In other words, the overall log-likelihood l is maximised by iteratively updating the estimators for $(\alpha, \beta)^\top$ and ρ.

The more recent publication [Krä+13] extends the above approach by considering Archimedean copulas to connect the marginal GLMs and by utilising the likelihood ratio test of Vuong for the selection of copula families.

A further approach involving GLMs, but without employing copulas, is proposed by [GGS16]. This approach shares similarity with the one presented at the beginning of the current section and treats the loss frequency N as a covariate in the GLM for the average loss size $\overline{X}$. If h_N and $h_{\overline{X}}$ denote the link function for N and $\overline{X}$, respectively, whereas the remaining notations stay the same, then the marginal GLMs can be written as

$$\lambda = \mathbb{E}[N|z] = h_N^{-1}(z^\top \alpha) \quad \text{and} \quad \mu_\theta = \mathbb{E}[\overline{X}|N, z] = h_{\overline{X}}^{-1}(z^\top \beta + \theta N). \tag{1.7}$$

Hence the parameter $\theta \in \mathbb{R}$ controls the degree of dependence between N and $\overline{X}$. Besides the covariate vector z describing fixed effects, the authors of [Jeo+17] extend (1.7) by adding a multivariate normally distributed covariate R to capture random effects.

Clearly, all previous stated dependence concepts based on GLMs need to be transferred with care when applying to operational risk data. First, the characterisation of the severity distribution highly relies on its expectation, which is certainly not sufficient in case of heavy-tailed operational losses. Moreover, the specification of covariates in operational risk

context is not as straightforward as in non-life insurance, in which rating factors serve as a natural choice. Nevertheless, potential impacts of economic and political environments as well as firm-specific factors on operational loss events have been reported in empirical studies and some initial work has been done in incorporating covariates into operational risk modelling. For more details we refer to Section 1.2.6 where more dependence concepts directly related to operational risk are presented.

The last proposal accounting for intra-cell dependence is found in [AT06] and [CMM10], for which the time component in the compound sum (1.1) is not considered as fixed but continuously evolving. In other words, the frequency component $N(t)$ is represented by a homogeneous Poisson process and the aggregate loss $S(t)$ accordingly by a compound Poisson process. In order to overcome the potential inconvenience when fitting a copula with discrete margins, the authors suggest to impose a copula between the single loss amount and its corresponding inter-arrival time instead of the loss number itself. Note that the inter-arrival times are i.i.d. exponential random variables under the current Poisson assumption. Besides the commonly used elliptical and Archimedean families, the Farlie-Gumbel-Morgenstern copula

$$C(u_1, u_2) \;=\; u_1 u_2 + \theta u_1 u_2 (1 - u_1)(1 - u_2), \quad (u_1, u_2)^\top \in [0, 1]^2,$$

with a single parameter $\theta \in [-1, 1]$ is highlighted due to its simplicity and tractability. This copula includes the independence copula for $\theta = 0$, as well as allows for both positive and negative dependence. Numerical examples in [CMM10] show that the dependence parameter θ has a significant impact on the ruin measures in actuarial context, so to expect a similar effect when estimating VaR in operational risk.

1.2.6 Inter- and intra-cell dependence

Of course, a model respecting both inter- and intra-cell dependence can be constructed by appropriately combining the concepts from Sections 1.2.1-1.2.4 with those from Section 1.2.5. Hence, below we shall focus on dependence models inherently developed for both inter- and intra-cell dependence.

Structural models with common factors, already widely applied in credit risk management, can also be adopted for modelling dependence in operational risk. For this purpose, the time component is treated in a discrete fashion, that is, $t \geq 1$ represents time periods of equal length, usually in annul units. Then the dependence among loss frequencies $N_i(t)$, $i \in \{1, \ldots, d\}$, and loss severities $X_{ik}(t)$, $k \geq 1$, is in particular driven by their common dependence on a set of risk factors which are modelled stochastically and may change in the course of time.

As illustration we consider a one-factor model based on Gaussian copulas as suggested in Chapter 12.5 of [CPS15]. Let $\Omega(t)$ denote the common factor, $W_i(t)$ the idiosyncratic component of $N_i(t)$, and $V_{ik}(t)$ the idiosyncratic component of $X_{ik}(t)$, respectively, and assume they are all independent random variables from the standard normal distribution. As noted before, the frequency distribution of risk cell i is given by F_{N_i} and the severity

distribution by F_{X_i}. If Φ denotes the distribution function of a standard normal random variable, then a one-factor model can be constructed as follows,

$$N_i(t) = F_{N_i}^{-1}(\Phi(\tilde{N}_i(t))) \quad \text{with} \quad \tilde{N}_i(t) = \rho_{N_i}\Omega(t) + \sqrt{1-\rho_{N_i}^2}W_i(t),$$
$$X_{ik}(t) = F_{X_i}^{-1}(\Phi(\tilde{X}_{ik}(t))) \quad \text{with} \quad \tilde{X}_{ik}(t) = \rho_{X_i}\Omega(t) + \sqrt{1-\rho_{X_i}^2}V_{ik}(t),$$

for $k \in \{1, \dots N_i(t)\}$ and $i \in \{1, \dots d\}$. Hence given $\Omega(t)$, the frequencies and severities are independent, but unconditionally they are dependent if the corresponding correlation coefficients ρ_{N_i} and ρ_{X_i} are non-zero.

From another perspective, the proposed one-factor model can also be used to identify dependent risk cells based on whether the corresponding coefficients ρ_{N_i} and ρ_{X_i} are close to zero. The obtained knowledge may help to calibrate more sophisticated dependence models. Furthermore, the extension to incorporating more than one common factor is readily achieved by assuming the joint distribution of $\Omega_1(t), \dots, \Omega_J(t)$ to be a multivariate normal distribution with zero means, unit variances, and some correlation matrix. Another modification is found in [MY09] where the authors consider Archimedean copulas instead of a Gaussian dependence structure.

Usually, the common factors are interpreted as macroeconomic variables as well as certain internal factors of each individual bank. Some factors are typically considered to affect frequencies only, for example system automations, some to affect severities only, for example changes in the legal environment, and some to affect both frequencies and severities, for example system security. Empirical evidence of such dependence is reported, amongst others, in [AB07], [Moo11], [CJY11] and [PP16].

The authors of [AB07] examine a sample of financial intermediaries and conclude that factors such as GDP, unemployment, equity indices, interest rates, foreign exchange rates, and changes in the regulatory environment have a considerable influence on operational risk. In addition, an analysis of the relationship between the unemployment rate and the operational loss events at American firms is carried out in [Moo11]. The results show a significant positive association between the unemployment rate and the loss severity on the one hand and an insignificant relation to the loss frequency on the other hand.

Apart from that, the sensitivity of operational risk with respect to firm-internal factors is explored in the empirical study [CJY11] based on the Algo FIRST database[9], which is an external database comprising publicly known loss events mainly occurred in North America. The loss incidence is found to be positively correlated with equity volatility, credit risk, a high number of anti-takeover provisions, and CEOs with a large amount of option and bonus-based compensation relative to salary. Moreover, quantifiable measures of the internal situation of a firm, such as capital ratios and number of employees, are identified to play a role in loss severities resulting from the event types of internal fraud or improper business and market practice by the more recent study [PP16], where the authors perform a multiple regression analysis on the ÖffSchOR[10] database containing

[9]Provided by the company Algorithmics at the time of the empirical study. Today Algorithmics is acquired by IBM.

[10]Öffentliche Schadenfälle OpRisk, provided by the Association of German Public Banks.

loss events occurred in German-speaking countries and culled from public sources.

Instead of imposing a dependence structure among the frequency and severity variables themselves, dependence can also be characterised by treating their underlying distribution parameters as random and modelling the corresponding joint distribution via a copula. The logic behind this is similar to that for the previous model based on common factors, that is, the distribution parameters are not constant but rather changing stochastically and jointly driven by the regulatory and economic environment or certain firm characteristics. With regard to actually fitting a model, [PSW09] advocates the use of Bayesian inference methodology which further allows for combining internal data, external data and expert opinions in the estimation procedure. The authors apply slice sampling, a special MCMC simulation algorithm, to obtain samples from the resulting posterior distribution and to estimate the model parameters.

As already indicated, dependence between severities and frequencies of different risk cells, as well as within one cell, can be achieved by explicitly incorporating covariates into a model. To this end, recall from EVT the peaks over threshold (POT) technique for characterising extreme loss events. The excesses of i.i.d. or stationary loss severities over a high threshold are independent GPD random variables and the excess times constitute a Poisson process. In order to properly depict the trends and uncertainties due to various external and internal factors which may be present in operational risk data, the classical POT method is modified by [CEN06] such that the GPD shape and scale parameters, as well as the Poisson intensity parameter, are dependent on time and covariates. More precisely, a smooth function is employed to describe the time dependence, an indicator function for the mapping to a certain risk cell, and a discontinuity indicator to capture a significant increase in loss frequency observed in the studied anonymised real loss data. The discontinuity in loss numbers is attributed to the often encountered reporting bias in operational risk, as loss events have been more frequently as well as more accurately recorded after the strengthening of regulations in the latest years.

The more recent publication [CEH16] by the same main author formalises the incorporation of covariates as a semi-parametric generalised additive model for location, scale, and shape (GAMLSS), that is,

$$h(\theta) \;=\; z_\theta^\top \beta_\theta + \sum_{j=1}^{J} h_{\theta,j}(z_{\theta,j}),$$

where θ is either the Poisson parameter, the GPD scale parameter or the GPD shape parameter; h is a monotonic link function; z_θ is a vector of linear predictors and β_θ their corresponding coefficients; $z_{\theta,j}$ is the j-th non-linear predictor for θ and $h_{\theta,j}$ a smooth function, for example composed of cubic splines. The proposed methodology is applied both to simulated data and to an external database provided by Willis Professional Risks containing losses collected from public media.

Going even one step further, the above model is extended by [Ham+16] to a so-called Markov-switching GAMLSS. Motivation for this extension is that the effect of a covariate may vary according to the state of unobservable environmental and economic factors.

Hence, the model components β_θ and $h_{\theta,j}$, $j \in \{1, \ldots, J\}$, are allowed to evolve over time and governed by an unobservable Markov chain. The applicability of the proposed approach is illustrated by fitting a model to the real losses resulting from external frauds and occurred at the Italian bank UniCredit. For computational simplicity, the authors utilise a stationary two-state model and assume the Markov chain that drives the varying dependence on covariates to be identical for frequency and severity. The estimation procedure is based on a forward recursion and the maximisation of a penalised likelihood function. Incorporated covariates include the percentage of revenue coming from fees, the Italian unemployment rate and the VIX index as a measure for market volatility.

Chapter 2

Preliminaries

The objective of this chapter is to prepare the theoretical ground for the upcoming dependence model for operational risk. In Section 2.1, the multivariate compound Poisson process is formally introduced as a special case of the more general class of Lévy processes. We present some of its essential properties which should be useful in future chapters. Furthermore, the notion of Lévy copulas is defined in Section 2.2 and shall be utilised to characterise the dependence structure among the marginal processes. As the theoretical framework of Lévy processes and Lévy copulas is quite comprehensive in its full length, and the current thesis rather focuses on their application in operational risk, we provide shortened proofs with adequate references in this chapter. Nonetheless, we detail the incentive of each selected theorem and proposition, as well as illustrate the new concept of dependence modelling via Lévy copulas by means of various examples and highlight their similarities and distinctions in comparison to ordinary copulas.

2.1 From Lévy processes to compound Poisson processes.

Definition 2.1 (**Lévy Process**).
Let $(\Omega, \mathcal{F}, \mathbb{P})$ be a probability space. A càdlàg[1] stochastic process $S(t)$, $t \geq 0$, on $(\Omega, \mathcal{F}, \mathbb{P})$ with values in $\mathbb{R}^d$ and $S(0) = 0$ a.s. is called a *Lévy process* if it satisfies the following properties:

(1) for all $n \in \mathbb{N}$ and every increasing sequence of times $0 \leq t_0 < \cdots < t_n < \infty$, the increments $S(t_0), S(t_1) - S(t_0), \ldots, S(t_n) - S(t_{n-1})$ are independent,

(2) for all $h \geq 0$ the distribution of $S(t+h) - S(t)$ is independent of $t \geq 0$,

[1]The acronym càdlàg comes from the French "continue à droite, limite à gauche", which translates to "right-continuous with left limits". Although the càdlàg assumption does not have to be imposed on the definition of Lévy processes, every Lévy process has a unique modification satisfying the càdlàg property. Hence we assume the càdlàg property without loss of generality.

(3) the process $S(t)$, $t \geq 0$, is continuous in probability, which means for all $\epsilon > 0$ there holds $\lim_{h \to 0} \mathbb{P}(|S(t+h) - S(t)| \geq \epsilon) = 0$.

An important subclass of Lévy processes are compound Poisson processes, which contribute as the foundation of the dependence model later on.

Definition 2.2 (**Compound Poisson process (CPP)**).
Let $N(t)$, $t \geq 0$, be a homogeneous Poisson process with intensity parameter $\lambda > 0$ and let X_k, $k \geq 1$, be i.i.d. random variables in $\mathbb{R}^d$ with distribution function F. Assume further the process $N(t)$ is independent from X_k for all $k \geq 1$, and the distribution F of the latter has no atom at zero. Then a d-dimensional *compound Poisson process* $S(t)$, $t \geq 0$, with intensity λ and jump size distribution F is a stochastic process given by

$$S(t) = \sum_{k=1}^{N(t)} X_k, \quad t \geq 0.$$

For notational convenience we summarise the key ingredients of a compound Poisson process into the abbreviation $S(t) \sim \mathrm{CPP}(\lambda, F)$.

Proposition 2.3. *A stochastic process $S(t)$, $t \geq 0$, is a compound Poisson process if and only if it is a Lévy process and its sample paths are piecewise constant functions.*

Proof. See Chapter 3, Proposition 3.3 in [CT04]. □

Proposition 2.4 (**Characteristic function of compound Poisson processes**).
Let $S(t) \sim CPP(\lambda, F)$ be a compound Poisson process on $\mathbb{R}^d$. Then its characteristic function has the representation

$$\mathbb{E}\left[e^{i\langle u, S(t)\rangle}\right] = \exp\left\{\lambda t \int_{\mathbb{R}^d} \left(e^{i\langle u,x\rangle} - 1\right) F(\mathrm{d}x)\right\}, \quad u \in \mathbb{R}^d. \tag{2.1}$$

Proof. Let $\phi_X(u) = \int_{\mathbb{R}^d} e^{i\langle u,x\rangle} F(\mathrm{d}x)$ denote the common characteristic function of the i.i.d. jump sizes X_k, $k \geq 1$, of $S(t)$. Then for any $n \in \mathbb{N}$, the characteristic function of the sum $\sum_{k=1}^n X_k$ is given by $\phi_X^n(u)$. By the tower property of conditional expectation and the independence between $N(t)$ and X_k for all $k \geq 1$, we calculate

$$\begin{aligned}
\mathbb{E}\left[e^{i\langle u, S(t)\rangle}\right] &= \mathbb{E}\left[\mathbb{E}\left[i\left\langle u, \sum_{k=1}^{N(t)} X_k\right\rangle \middle| N(t)\right]\right] = \sum_{n=1}^{\infty} \mathbb{E}\left[i\left\langle u, \sum_{k=1}^{n} X_k\right\rangle\right] \mathbb{P}\left(N(t) = n\right) \\
&= \sum_{n=1}^{\infty} \phi_X^n(u) \frac{e^{-\lambda t}(\lambda t)^n}{n!} = \exp\left\{\lambda t(\phi_X(u) - 1)\right\} \\
&= \exp\left\{\lambda t \int_{\mathbb{R}^d} \left(e^{i\langle u,x\rangle} - 1\right) F(\mathrm{d}x)\right\}.
\end{aligned}$$

□

Comparing (2.1) to the characteristic function $\exp\{\lambda t(e^{iu}-1)\}$ of a Poisson process, we see that a compound Poisson process $S(t)$ can be interpreted as the superposition of independent Poisson processes with the same intensity λ, but different jump sizes determined by the distribution function F. In other words, the total intensity of jump sizes in the interval $[x, x+\mathrm{d}x]$ is given by $\lambda F(\mathrm{d}x)$. This gives rise to the introduction of a new measure $\Pi(\cdot) = \lambda F(\cdot)$ on $(\mathbb{R}^d, \mathcal{B}(\mathbb{R}^d))$. Note that for simplicity we use the notation F both for the distribution as well as for the distribution function of the jump sizes. Hence formula (2.1) can be rewritten as

$$\mathbb{E}\left[e^{i\langle u,S(t)\rangle}\right] \;=\; \exp\left\{t\int_{\mathbb{R}^d}\left(e^{i\langle u,x\rangle}-1\right)\,\Pi(\mathrm{d}x)\right\}, \quad u\in\mathbb{R}^d. \tag{2.2}$$

The above formula is a special case of the so-called Lévy-Khinchin representation for Lévy processes and the measure Π is the corresponding Lévy measure having the formal definition:

Definition 2.5 (**Lévy measure**).
Let $S(t)$, $t\geq 0$, be a Lévy process on $\mathbb{R}^d$ and let $\Delta S(t) = S(t) - S(t-)$ denote its jump at time point t. The measure ν on $\mathbb{R}^d\setminus\{0\}$ defined by

$$\nu(B) \;=\; \mathbb{E}\left[\,\#\left\{\Delta S(t)\in B \,|\, t\in[0,1] \,\wedge\, \Delta S(t)\neq 0\right\}\right], \quad B\in\mathcal{B}(\mathbb{R}^d\setminus\{0\}),$$

is called the *Lévy measure* of $S(t)$.

Hence the Lévy measure $\nu(B)$ of a Borel set B is the expected number of non-trivial jumps per unit time interval with size in B. It is easy to verify that the above definition is indeed in accordance with our derivation of the measure $\Pi(\cdot) = \lambda F(\cdot)$ for the special case of compound Poisson processes. Recall the Poisson process $N(t)$ has up to time $t=1$ a finite expected number of jumps equal to its intensity parameter λ and it is independent from the jump sizes X_k for all $k\geq 1$. Moreover, all non-trivial jumps of $S(t)$ are fully characterized by the distribution function F of X_k, $k\geq 1$, and the latter are a.s. non-zero.

The distribution of a general Lévy process is also uniquely determined by its characteristic function, which is given by the Lévy-Khinchin formula:

Theorem 2.6 (**Lévy-Khinchin formula for Lévy processes**).
Let $S(t)$, $t\geq 0$, be a Lévy process on $\mathbb{R}^d$. Then there exist

(1) a positive definite matrix $A\in\mathbb{R}^{d\times d}$,

(2) a Radon measure[2] *ν on $\mathbb{R}^d\setminus\{0\}$ satisfying the conditions $\int_{|x|\leq 1}|x|^2\nu(\mathrm{d}x)<\infty$ and $\int_{|x|\geq 1}\nu(\mathrm{d}x)<\infty$,*

(3) and a vector $\gamma\in\mathbb{R}^d$,

[2]Let E be a subset of $\mathbb{R}^d$. A Radon measure on $(E,\mathcal{B}(E))$ is a measure μ such that $\mu(B)<\infty$ for every compact measurable set $B\in\mathcal{B}(E)$.

such that the characteristic function of $S(t)$ has the representation

$$\begin{aligned}\mathbb{E}\left[e^{i\langle u,S(t)\rangle}\right] &= e^{t\Psi(u)}, \quad u \in \mathbb{R}^d \\ \text{with} \quad \Psi(u) &= -\frac{1}{2}\langle u, Au\rangle + i\langle \gamma, u\rangle + \int_{\mathbb{R}^d} \left(e^{i\langle u,x\rangle} - 1 - i\langle u,x\rangle \mathbb{1}_{\{|x|\leq 1\}}\right) \nu(\mathrm{d}x).\end{aligned}$$

The triplet (A, ν, γ) is called the characteristic triplet of the process $S(t)$.

Proof. See Chapter 3, Theorem 3.1 in [CT04]. □

As before, the measure ν is referred to as the Lévy measure. The Lévy-Khinchin representation induces that a Lévy process is fully determined by its characteristic triplet. In case of a compound Poisson process $S(t) \sim \mathrm{CPP}(\lambda, F)$, we have seen that its Lévy measure is given by $\nu(\cdot) = \Pi(\cdot) = \lambda F(\cdot)$. Furthermore, it is easy to verify that the characteristic function (2.1) coincides with the Lévy-Khinchin formula if we choose $A = 0$ and $\gamma = \int_{|x|\leq 1} x\, \Pi(\mathrm{d}x)$.

In addition, the class of Lévy processes is invariant under linear transformations:

Theorem 2.7. *Let $S(t)$, $t \geq 0$, be a Lévy process on $\mathbb{R}^d$ with characteristic triplet (A, ν, γ) and let M be an $m \times d$ matrix. Then $\tilde{S}(t) = MS(t)$, $t \geq 0$, is a Lévy process on $\mathbb{R}^m$ with characteristic triplet $(\tilde{A}, \tilde{\nu}, \tilde{\gamma})$ specified through*

$$\begin{aligned}\tilde{A} &= MAM^\top, \\ \tilde{\nu}(B) &= \nu(\{x \,|\, Mx \in B\}), \quad B \in \mathcal{B}(\mathbb{R}^m), \\ \tilde{\gamma} &= M\gamma + \int_{\mathbb{R}^m} x \left(\mathbb{1}_{\{|x|\leq 1\}}(x) - \mathbb{1}_D(x)\right) \tilde{\nu}(\mathrm{d}x),\end{aligned} \tag{2.3}$$

where $D = \{My \,|\, y \in \mathbb{R}^d \wedge |y| \leq 1\}$ is the image of a closed unit ball in $\mathbb{R}^d$ under M.

Proof. See Chapter 4, Theorem 4.1 in [CT04]. □

In particular, the above theorem implies the margins of a Lévy measure can be computed in an analogous manner to the margins of a probability measure. The margins of a Lévy measure ν can be seen as the one-dimensional Lévy measures associated with the marginal processes $S_1(t), \ldots, S_d(t)$ of $S(t)$. For example, the first marginal process $S_1(t)$ is obtained by setting M equal to the first standard basis vector $(1, 0 \ldots, 0)^\top$. Then its one-dimensional Lévy measure $\nu_1(B)$ is given by

$$\nu_1(B) = \nu(B \times (-\infty, \infty) \times \cdots \times (-\infty, \infty))$$

for any Borel set $B \in \mathcal{B}(\mathbb{R} \setminus \{0\})$. Lévy measures of the remaining marginal processes are computed in a similar way.

Furthermore, we deduce that the margins of a multivariate compound Poisson process are themselves univariate compound Poisson processes. As already mentioned, the i-th marginal process can be accessed by setting M to the i-th standard basis vector. Then

the corresponding marginal Lévy measure Π_i is simply given by the projection of Π onto the i-th component. Due to the elementary choice of M, the second summand in (2.3) vanishes and only the term $\gamma_i = \int_{|x|\leq 1} x_i \Pi(\mathrm{d}x) = \int_{|x|\leq 1} x \Pi_i(\mathrm{d}x)$ remains. By inserting the triplet $(0, \Pi_i, \gamma_i)$ into the Lévy-Khinchin formula, we see that the i-th marginal process indeed has the characteristic function of a one-dimensional compound Poisson process. Similarly, it can be shown that the sum of an arbitrary subset of the margins constitutes a compound Poisson process as well.

2.2 Tail integrals and Lévy copulas

When modelling the dependence structure of multivariate Lévy processes, a natural approach is by transferring the concept of copulas for random vectors to a dependence concept for Lévy processes. The dependence among the margins, or more precisely, the dependence among the marginal jumps, of a multivariate Lévy process can be characterised by the so-called Lévy copulas, which have many common properties with the ordinary copulas, but are defined on a different domain. In order to distinguish between the two copula concepts, in the sequel we shall refer to copulas for random variables as ordinary copulas and use the regular capital letter C for notation. Copulas for Lévy processes shall be referred to as Lévy copulas and denoted by the calligraphic letter $\mathfrak{C}$.

The notion of Lévy copula was first introduced in 2003 by Tankov in [Tan03] for Lévy processes only admitting positive jumps. Shortly thereafter, in [KT06] Kallsen and Tankov presented an extension for Lévy processes with possibly negative jumps. In the following we shall concentrate on the theory for Lévy copulas and Lévy measures supported on $[0, \infty]^d$, as the main application of this dependence concept in the present thesis is the modelling of positive loss amounts in operational risk.

In order to formally define a Lévy copula, at first we need to review the notion of d-increasing functions. For this purpose, we denote a half-open interval in $(-\infty, \infty]^d$ by

$$(a, b] \;=\; (a_1, b_1] \times \cdots \times (a_d, b_d],$$

where the vectors $a = (a_1, \ldots, a_d)^\top$ and $b = (b_1, \ldots, b_d)^\top$ satisfy the property $a_i \leq b_i$ for all $i \in \{1, \ldots d\}$. We are interested in the behaviour of functions on these half-open intervals.

Definition 2.8 (**F-volume and d-increasing functions**).
Let $D \subset (-\infty, \infty]^d$ be a non-empty set and $F : D \to (-\infty, \infty]$ a d-variate function. The *F-volume* of an half-open interval $(a, b]$ in D is defined as

$$V_F((a, b]) \;=\; \sum_{j_1=1}^{2} \cdots \sum_{j_d=1}^{2} (-1)^{j_1+\cdots+j_d} \, F(u_{1j_1}, \ldots, u_{dj_d})$$

with $u_{i1} = a_i$ and $u_{i2} = b_i$ for all $i \in \{1, \ldots d\}$. The function F is called *d-increasing* if the F-volume $V_F((a, b])$ is non-negative for all half-open intervals $(a, b]$ in D.

In other words, the F-volume is the sum of function values at each vertice of the d-dimensional interval $(a, b]$, whereby a vertice contributes with positive sign, if an even number of its components belongs to the vector a, and with negative sign otherwise. Note that a d-variate probability function F always satisfies the d-increasing property, as the corresponding F-volume $V_F((a, b])$ just reflects the respective probability measure of $(a, b]$. In the two-dimensional case, the F-volume reduces to the measure of a rectangle $(a_1, b_1] \times (a_2, b_2]$ given by

$$V_F((a, b]) \;=\; F(b_1, b_2) - F(a_1, b_2) - F(b_1, a_2) + F(a_1, a_2).$$

Definition 2.9 (**Positive Lévy copulas**).
A d-dimensional *Lévy copula* for Lévy processes with positive jumps, or for short, a *positive Lévy copula*, is a function $\mathfrak{C} : [0, \infty]^d \to [0, \infty]$ such that

(1) $\mathfrak{C}$ is grounded, which means $\mathfrak{C}(u_1, \ldots, u_d) = 0$ if $u_i = 0$ for at least one $i \in \{1, \ldots d\}$,

(2) $\mathfrak{C}$ is d-increasing,

(3) $\mathfrak{C}$ has uniform margins, that is, $\mathfrak{C}_i(u_i) = \mathfrak{C}(\infty, \ldots, \infty, u_i, \infty \ldots, \infty) = u_i$ for all $i \in \{1, \ldots d\}$ and $u_i \in [0, \infty]$.

Similarly to ordinary copulas, Lévy copulas are Lipschitz continuous with Lipschitz constant equal to one:

Lemma 2.10. *Let $\mathfrak{C} : [0, \infty]^d \to [0, \infty]$ be a Lévy copula. Then for every $(u_1, \ldots, u_d)^\top$ and $(v_1, \ldots, v_d)^\top \in [0, \infty)^d$ it holds that*

$$|\mathfrak{C}(u_1, \ldots, u_d) - \mathfrak{C}(v_1, \ldots, v_d)| \;\leq\; \sum_{i=1}^{d} |u_i - v_i| \,.$$

Proof. See Lemma 3.2 in [KT06]. □

In particular, this smooth property of Lévy copulas allows us to exchange the copula $\mathfrak{C}$ with limit-taking of its arguments in subsequent sections.

Before we state the main theorem for the interaction between Lévy copulas and Lévy processes, which is an analogue to Sklar's theorem on ordinary copulas, we shall think a little more about the appropriate input arguments for a Lévy copula. Recall that ordinary copulas are introduced on the basis of distribution functions associated with probability measures, and Lévy measures play the same role for the jumps of Lévy processes as probability measures for random variables. Therefore, an appropriate input argument for Lévy copulas must be built upon the Lévy measure and the following notion of tail integrals evolves.

Definition 2.11 (**Tail integral of Lévy processes**).
Let ν be a Lévy measure on $[0, \infty)^d \setminus \{0\}$. Its *tail integral* is a function $\overline{\nu} : [0, \infty]^d \to [0, \infty]$

defined as

$$\overline{\nu}(x_1,\dots,x_d) = \begin{cases} 0, & \text{if } x_i = \infty \text{ for at least one } i \in \{1,\dots d\}, \\ \nu([x_1,\infty)\times\cdots\times[x_d,\infty)) & \text{if } (x_1,\dots,x_d)^\top \in [0,\infty)^d \setminus \{0\}, \\ \infty & \text{if } (x_1,\dots,x_d)^\top = 0. \end{cases}$$

As the Lévy measure ν is a Radon measure, its tail integral $\overline{\nu}$ is finite except at zero. In addition, the i-th margin of the tail integral $\overline{\nu}$ is given by

$$\overline{\nu}_i(x_i) \;=\; \overline{\nu}(0,\dots,0,x_i,0,\dots,0), \quad x_i > 0,$$

which corresponds to the tail integral of the i-th marginal Lévy measure ν_i, that is, $\overline{\nu}_i(x_i) = \nu_i([x_i,\infty))$.

The following theorem parallelises the dependence characterisation through copulas on the level of Lévy processes and the role of distribution functions in the ordinary Sklar's theorem is now played by tail integrals.

Theorem 2.12 (**Sklar's theorem for Lévy processes**).

(a) Let $\overline{\nu}$ denote the tail integral of a Lévy process on $[0,\infty)^d$ and let $\overline{\nu}_1,\dots,\overline{\nu}_d$ be the tail integrals of its components. Then a Lévy copula $\mathfrak{C} : [0,\infty]^d \to [0,\infty]$ exists such that for all $(x_1,\dots,x_d)^\top \in [0,\infty]^d$ the relation

$$\overline{\nu}(x_1,\dots,x_d) \;=\; \mathfrak{C}(\overline{\nu}_1(x_1),\dots,\overline{\nu}_d(x_d)) \tag{2.4}$$

holds. If the marginal tail integrals $\overline{\nu}_1,\dots,\overline{\nu}_d$ are continuous, then the Lévy copula $\mathfrak{C}$ is unique. Otherwise it is unique on the product of the range of the marginal tail integrals.

(b) Conversely, if $\mathfrak{C}$ is a d-dimensional Lévy copula and $\overline{\nu}_1,\dots,\overline{\nu}_d$ are tail integrals of one-dimensional Lévy measures, then the function $\overline{\nu}$ defined by (2.4) is the tail integral of a Lévy process on $[0,\infty)^d$ having marginal tail integrals $\overline{\nu}_1,\dots,\overline{\nu}_d$.

Proof. See Theorem 3.1 an well as Theorem 3.4 in [Tan03]. □

The above theorem shows that Lévy copulas connect multidimensional tail integrals to their margins in the same way as ordinary copulas connect multivariate probability distribution functions to their margins. Furthermore, the role of Lévy copulas is twofold. First, they provide a complete characterisation of the possible dependence structures for the jump part of multidimensional Lévy processes. Second, Lévy copulas enable the construction of a multidimensional Lévy process from a given collection of one-dimensional Lévy processes complying with a particular dependence structure.

Hereinafter we give several examples of Lévy copulas and see how some of them can actually be constructed from ordinary copulas. Of course, one fundamental dependence structure is the independence case and to this we need at first the following criterion characterising independent Lévy processes.

Lemma 2.13 (**Independence of Lévy processes**).
Let $S(t) = (S_1(t), \dots, S_d(t))^\top$ be a Lévy process on $[0,\infty)^d$ such that its characteristic triplet has the form $(0, \nu, \gamma)$. Then its margins $S_1(t), \dots, S_d(t)$ are independent if and only if ν is completely supported by the coordinate axes, which means for any Borel set $B \in \mathcal{B}([0,\infty)^d \setminus \{0\})$ it holds that

$$\nu(B) \;=\; \nu_1(B_1) + \cdots + \nu_d(B_d),$$

where for every $i \in \{1, \dots, d\}$ the measure ν_i is the Lévy measure belonging to the marginal process $S_i(t)$ and $B_i = \{x_i \,|\, (0, \dots, 0, x_i, 0, \dots, 0) \in B\}$ is the projection of the set B onto the i-th coordinate axis.

Proof. The independence of a Lévy process translates to that its marginal processes never jump at the same time. Both directions of the statement in Lemma 2.13 can be shown by elementary manipulations of the characteristic function, that is, the Lévy-Khinchin representation of $S(t)$. See also Lemma 4.1 in [Tan03]. □

Given a Lévy process $S(t)$ with discontinuous marginal tail integrals, there might be more than one Lévy copula satisfying equation (2.4) in Sklar's theorem for Lévy copulas. In general, we say $\mathfrak{C}$ is the or a Lévy copula of a process $S(t)$, if together with the tail integrals equation (2.4) is fulfilled. In this sense we can derive one possible copula corresponding to Lévy processes with independent margins:

Proposition 2.14 (**Independence Lévy copula**).
Let $S(t) = (S_1(t), \dots, S_d(t))^\top$ be a Lévy process on $[0,\infty)^d$ such that its characteristic triplet has the form $(0, \nu, \gamma)$. Then the marginal processes $S_1(t), \dots, S_d(t)$ of $S(t)$ are independent if and only if its Lévy copula (or one of them if there are many) is given by

$$\mathfrak{C}_\perp(u_1, \dots, u_d) \;=\; u_1 \mathbb{1}_{\{u_2=\infty,\dots,u_d=\infty\}} + \cdots + u_d \mathbb{1}_{\{u_1=\infty,\dots,u_{d-1}=\infty\}}.$$

Proof. As the tail integrals $\overline{\nu}_1, \dots, \overline{\nu}_d$ are finite except at zero, Lemma 2.13 implies the marginal processes $S_1(t), \dots, S_d(t)$ are independent if and only if the tail integral $\overline{\nu}$ of $S(t)$ has the form

$$\overline{\nu}(x_1, \dots, x_d) \;=\; \overline{\nu}_1(x_1) \mathbb{1}_{\{x_2=0,\dots,x_d=0\}} + \cdots + \overline{\nu}_d(x_d) \mathbb{1}_{\{x_1=0,\dots,x_{d-1}=0\}}.$$

Then the assertion directly follows from Sklar's theorem for Lévy processes. □

Besides independent margins, another important basis dependence structure is the complete dependence case. To this we proceed as before and specify first the notion of complete dependence for Lévy processes. In the second step the corresponding Lévy copula will be computed.

Definition 2.15 (**Complete dependence of Lévy processes**).

(a) A subset E of $[0,\infty)^d$ is called *increasing* if two arbitrary and different vectors $a = (a_1, \dots, a_d)^\top$ and $b = (b_1, \dots, b_d)^\top$ in E satisfy either $a_i < b_i$ for all $i \in \{1, \dots d\}$, or $a_i > b_i$ for all $i \in \{1, \dots d\}$.

(b) Let $S(t) = (S_1(t), \dots, S_d(t))^\top$ be a Lévy process on $[0,\infty)^d$. Its jumps are called *completely dependent* or *comonotonic*, if an increasing set $E \subset [0,\infty)^d \setminus \{0\}$ exists such that every non-trivial jump $\Delta S(t) = S(t) - S(t-) \neq 0$ of $S(t)$ belongs to E.

Proposition 2.16 (**Complete dependence Lévy copula**).
Let $S(t) = (S_1(t), \dots, S_d(t))^\top$ be a Lévy process on $[0,\infty)^d$. If its jumps are completely dependent, then a possible Lévy copula of $S(t)$ has the form

$$\mathfrak{C}_{\|}(u_1, \dots, u_d) \;=\; \min(u_1, \dots, u_d).$$

Conversely, if all marginal tail integrals of $S(t)$ are continuous and the Lévy copula of $S(t)$ is given by $\mathfrak{C}_{\|}$, then the jumps of $S(t)$ are completely dependent.

Proof. According to Definition 2.15, the jumps of $S(t)$ are completely dependent if and only if an increasing set $E \subset [0,\infty)^d \setminus \{0\}$ exists such that the Lévy measure ν of $S(t)$ is concentrated on E. Then the assertion follows from the property of increasing sets and Sklar's theorem for Lévy processes. For more details also see Proposition 4.3 in [Tan03]. □

Note that the complete dependence Lévy copula $\mathfrak{C}_{\|}$ has the same form as the ordinary comonotonicity copula, under which a random vector has perfectly positive dependent components. On the other hand, the ordinary independence copula is given by $C_\perp(u_1, \dots, u_d) = \prod_{i=1}^d u_i$ and has an entirely different structure than the independence Lévy copula. This is due to the fact that the independence concept defined for Lévy processes is different from the stochastically independence underlying random vectors, whereas both the ordinary and the Lévy version of the comotonicity copula are based on the idea that the marginal components are almost surely strictly increasing functions of each other.

Encouraged by the class of ordinary copulas known as Archimedean copulas, an analogous class of Archimedean Lévy copulas can be constructed from the so-called generator functions:

Proposition 2.17 (**Archimedean Lévy copulas**).
Let $\phi : [0,\infty] \to [0,\infty]$ be a strictly decreasing continuous function such that $\phi(0) = \infty$ and $\phi(\infty) = 0$. Furthermore, suppose the inverse function ϕ^{-1} has derivatives up to order d on $(0,\infty)$ and satisfies the condition $(-1)^d \frac{\mathrm{d}^d \phi^{-1}(x)}{\mathrm{d}x^d} \geq 0$. Then the d-dimensional function $\mathfrak{C} : [0,\infty]^d \to [0,\infty]$ defined through

$$\mathfrak{C}(u_1, \dots, u_d) \;=\; \phi^{-1}(\phi(u_1) + \cdots + \phi(u_d)) \tag{2.5}$$

is a positive Lévy copula.

Proof. The statement can be shown by verifying the three defining properties of a Lévy copula given in Definition 2.9. For more details see also Proposition 4.5 in [Tan03]. □

Recall that ordinary Archimedean copulas have the same defining structure as (2.5), whereas the generator function must satisfy equivalent conditions besides possessing a different domain. Writing ψ for a valid generator of an ordinary Archimedean copula, it maps $[0,1]$ onto $[0,\infty]$ with $\psi(1)=0$. For comparison, the generator functions of some common Archimedean copula families are summarised in Table 2.1[3].

Copula family	**Ordinary generator $\psi(u)$**	**Lévy generator $\phi(u)$**
Clayton	$u^{-\theta}-1,\ \theta>0$	$u^{-\theta},\ \theta>0$
Gumbel	$(-\ln u)^{\theta},\ \theta\geq 1$	$[\ln(u+1)]^{-\theta},\ \theta>0$
Ali-Mikhail-Haq	$\ln\left(\frac{1-\theta}{u}+\theta\right),\ \theta\in[-1,1)$	$\ln\left(\frac{1-\theta}{u}+1\right),\ \theta\in[-1,1)$

Table 2.1: Comparison of Archimedean generators for ordinary and Lévy copulas.

The Clayton Lévy copula was already introduced in the fundamental work [Tan03] of Lévy copulas by Tankov and hence has the longest history among the Archimedean families. Similar to its ordinary counterpart, the Clayton Lévy copula has the convenient properties that it is governed by a single parameter and comprises the complete dependence copula for $\theta\uparrow\infty$ as well as the independence copula for $\theta\downarrow 0$ as limiting cases. Since in practice often only a limited amount of data are available for the estimation of a dependence model, it is not surprising that the one-parametric Clayton family with a wide range of possible dependence strength counts to the most commonly used Lévy copulas. Its application in operational risk context will be demonstrated in the upcoming chapters. For examples in other practical fields, the interested readers are referred to [MM13] for modelling spark spreads on energy markets, to [LDD16] for maintenance optimisation, and to [Okh16] for VaR calculation of portfolios constructed on classic stocks as well as on cryptocurrencies.

We conclude this theoretical chapter with a method of constructing Lévy copulas from ordinary ones.

Proposition 2.18. *Let $C:[0,1]^d\to[0,1]$ be an ordinary copula and $f:[0,1]\to[0,\infty]$ a strictly increasing function with $f(0)=0$, $f(1)=\infty$, and non-negative derivatives up to order d on $(0,1)$. Then the function $\mathfrak{C}:[0,\infty]^d\to[0,\infty]$ defined as*

$$\mathfrak{C}(u_1,\ldots,u_d) \;=\; f\left(C\left(f^{-1}(u_1),\ldots,f^{-1}(u_d)\right)\right) \tag{2.6}$$

is a positive Lévy copula.

Proof. The properties of groundedness and uniform margins can be established through simple calculations. For the d-increasingness of $\mathfrak{C}$ we refer to Theorem 5.1 in [Tan04]. □

As many parametric families of ordinary copulas have been well studied, Proposition 2.18 substantially enlarges the available families of Lévy copulas. Hereinafter we call any func-

[3]The representation of the Gumbel and Ali-Mikhail-Haq Lévy generators is found in [BB11], primarily used to optimise insurance portfolios.

tion f satisfying the required conditions a construction function. Because we mainly concentrate on two-dimensional compound Poisson processes later on, some examples of f in order to define a valid bivariate Lévy copula are given by

$$\begin{aligned} f(x) &= \frac{x}{1-x}, \\ f(x) &= -\ln(1-x), \\ \text{and} \quad f(x) &= (x^{-\theta}-1)^{-\frac{1}{\theta}}, \quad \theta > 0. \end{aligned}$$

For illustration, consider the ordinary independence copula $C(u_1, u_2) = u_1 u_2$ and the construction function $f(x) = \frac{x}{1-x}$. By applying formula (2.6) we obtain the Lévy copula $\mathfrak{C}(u_1, u_2) = \frac{u_1 u_2}{u_1+u_2+1}$ which assembles the Ali-Mikhail-Haq Lévy copula with parameter $\theta = 0$ as given in Table 2.1.

Chapter 3

Dependence modelling via compound Poisson processes and Lévy copulas

Being well prepared by the previous chapter, we now are able to establish a multivariate model based on compound Poisson processes and Lévy copulas. After introducing the model definition in Section 3.1, Section 3.2 works out the properties of a bivariate process in detail, and the impact of Lévy copulas on both frequency and severity interdependence becomes clear. Thereafter, the associated likelihood function, either under a continuous or a discrete observation scheme, is provided and proved in Section 3.3. As a result, MLE of our dependence model is made possible and the behaviour of its resulting estimators will be studied in Chapter 5 under various distribution assumptions.

3.1 A multivariate compound Poisson model for operational risk

In line with the loss distribution approach introduced in Section 1.2, let us consider a financial institution maintaining d operational risk cells. Moreover, suppose a sufficient degree of homogeneity within each risk class is ensured such that the losses allocated to the same cell may be reasonably assumed as i.i.d and the corresponding loss frequency may be described by a homogeneous Poisson process. Then justified by Theorem 2.12, the dependence characterisation among the d risk cells can be detached from the modelling of the marginal processes. More specifically, we follow the approach of [BK10] and formalise the subsequent multivariate model.

Definition 3.1 (**Multivariate compound Poisson model**).

(a) **Marginal cell processes:** The aggregate loss amount in each risk cell $i \in \{1, \dots, d\}$ is given by a one-dimensional compound Poisson process

$$S_i(t) = \sum_{k=1}^{N_i(t)} X_{ik}, \quad t \geq 0,$$

where $N_i(t)$ is a homogeneous Poisson process with intensity parameter $\lambda_i > 0$ and X_{ik}, $k \geq 1$, are i.i.d. single losses following an absolutely continuous distribution function F_{X_i} with $F_{X_i}(0) = 0$. The loss frequency process $N_i(t)$ is independent of the loss severities X_{ik} for all $k \geq 1$.

(b) **Dependence structure:** For $i \in \{1, \ldots, d\}$, let $\overline{\Pi}_i : [0, \infty] \to [0, \infty]$ denote the tail integral associated with the marginal process $S_i(t)$, and let $\mathfrak{C} : [0, \infty]^d \to [0, \infty]$ be a Lévy copula. Then the d-variate tail integral defined by

$$\overline{\Pi}(x_1, \ldots, x_d) \;=\; \mathfrak{C}(\overline{\Pi}_1(x_1), \ldots, \overline{\Pi}_d(x_d)), \quad (x_1, \ldots, x_d)^\top \in [0, \infty]^d, \tag{3.1}$$

constitutes the tail integral of the d-dimensional compound Poisson process

$$S(t) \;=\; (S_1(t), \ldots, S_d(t))^\top, \quad t \geq 0.$$

(c) **The overall loss process:** The overall loss process is given by the one-dimensional compound Poisson process

$$S_+(t) \;=\; \sum_{i=1}^{d} S_i(t), \quad t \geq 0.$$

There are several convenient properties of the above model worth mentioning. First, under the assumption of absolutely continuous severity distributions, the tail integral of the marginal process $S_i(t) \sim \mathrm{CPP}(\lambda_i, F_{X_i})$, $i \in \{1, \ldots, d\}$, is simply given by $\overline{\Pi}_i(x_i) = \lambda_i \overline{F}_{Xi}(x_i)$ for $x_i > 0$. Second, as a consequence of Part (b) in Theorem 2.12, the d-dimensional process $S(t) = (S_1(t), \ldots, S_d(t))^\top$ defined via the tail integral in equation (3.1) indeed forms a compound Poisson process. Last but not least, due to the invariance of compound Poisson processes under linear transformations as detailed in Theorem 2.7, the overall loss process $S_+(t)$ constitutes itself a univariate compound Poisson process with tail integral

$$\overline{\Pi}_+(x) \;=\; \Pi\left(\left\{(x_1, \ldots, x_d)^\top \in [0, \infty)^d \setminus \{0\} \,\middle|\, x_1 + \cdots + x_d \geq x\right\}\right), \quad x \geq 0. \tag{3.2}$$

On the whole, Definition 3.1 presents a flexible model for characterising operational loss events, which can be readily built by taking any collection of d one-dimensional compound Poisson processes and linking them together via an arbitrary Lévy copula.

3.2 Detailed analysis of bivariate compound Poisson models

In this section we focus on bivariate compound Poisson models and give an in-depth analysis of their properties. One reason for this is to become familiar with the concept of dependence modelling through Lévy copulas, as in a low-dimensional setting we can break down the relationships between the different model components in a detailed manner. Another reason is that models of higher dimension are often most easily built based upon bivariate processes.

3.2.1 Construction and properties

In line with Definition 3.1, a bivariate compound Poisson model is specified through

$$S(t) = (S_1(t), S_2(t))^\top = \left(\sum_{j=1}^{N_1(t)} X_{1j}, \sum_{l=1}^{N_2(t)} X_{2l} \right)^\top, \quad t \geq 0, \tag{3.3}$$

where $N_1(t)$ and $N_2(t)$ denote the marginal frequency processes, and X_{1j} as well as X_{2l} are the marginal loss severities in the first and the second risk cell, respectively. On the other hand, the two-dimensional compound Poisson process $S(t)$ has by Definition 2.2 a representation through a single Poisson frequency process $N(t)$, $t \geq 0$, and i.i.d. non-negative bivariate loss severities, that is,

$$S(t) = \sum_{h=1}^{N(t)} Y_h = \sum_{h=1}^{N(t)} (Y_{1h}, Y_{2h})^\top, \quad t \geq 0. \tag{3.4}$$

Note that the components Y_{1h} and Y_{2h} of the bivariate losses are not the same as the univariate marginal losses X_{1j} and X_{2l} from equation (3.3). Bear in mind not to confuse them with each other and their relation will be developed in the course of this section.

We assume the dependence structure of the bivariate compound Poisson model, that is, the dependence between the marginal processes $S_1(t)$ and $S_2(t)$, is given by a Lévy copula $\mathfrak{C}$. Furthermore, the dependence between the margins of the bivariate losses $Y_h = (Y_{1h}, Y_{2h})^\top$, $h \geq 1$, shall comply with an ordinary survival copula C_s. Recall that a survival copula connects the marginal survival distributions to the joint survival distribution and the following lemma holds as a direct consequence of Sklar's theorem for ordinary copulas.

Lemma 3.2 (**Sklar's theorem for survival functions**).
Let $\overline{F} : \mathbb{R}^d \to [0,1]$ be a d-variate survival function and let $\overline{F}_1, \ldots, \overline{F}_d : \mathbb{R} \to [0,1]$ be the corresponding marginal survival functions. Then an ordinary copula $C : [0,1]^d \to [0,1]$ exists such that for all $(x_1, \ldots, x_d)^\top \in \mathbb{R}^d$ there holds

$$\overline{F}(x_1, \ldots, x_d) = C(\overline{F}_1(x_1), \ldots, \overline{F}_d(x_d)). \tag{3.5}$$

Conversely, if C is an ordinary copula and $\overline{F}_1, \ldots, \overline{F}_d$ are univariate survival functions, then the function $\overline{F}$ defined by (3.5) is a d-variate survival function with margins $\overline{F}_1, \ldots, \overline{F}_d$. The link between $\overline{F}$ and C is one-to-one if all survival functions $\overline{F}_1, \ldots, \overline{F}_d$ are continuous, otherwise C is unique on the product of the range of the marginal survival functions.

On the other hand, a Lévy copula operates on the level of tail integrals, which are closely linked to the severity distribution tails, that is, the survival functions. Hence it is a natural question whether there is a tangible relationship between the Lévy copula $\mathfrak{C}$ and the survival copula C_s. Before we state the result shortly, we shall first take a closer look at the domain of the survival copula C_s.

Assume the i.i.d. loss vectors $Y_h = (Y_{1h}, Y_{2h})^\top$, $h \geq 1$, have a common distribution function F_Y, and let $Y = (Y_1, Y_2)^\top$ be a generic random vector with this distribution. In other words, $F_Y(x_1, x_2) = \mathbb{P}(Y_1 \leq x_1, Y_2 \leq x_2)$ holds for $(x_1, x_2)^\top \in [0, \infty)^2$. The marginal laws shall be denoted by F_{Y_i}, $i \in \{1, 2\}$, and the corresponding survival functions are given by

$$\overline{F}_{Y_i}(x_i) \;=\; 1 - F_{Y_i}(x_i) \;=\; \mathbb{P}(Y_i > x_i), \quad x_i \geq 0,\, i \in \{1, 2\}.$$

Although the distribution function F_Y is not allowed to have an atom at zero by Definition 2.2 of compound Poisson processes, the marginal distributions F_{Y_1} and F_{Y_2} can have positive measure at zero, for which we write

$$p_i \;=\; \mathbb{P}(Y_i = 0) \;=\; F_{Y_i}(0) \in [0, 1), \quad i \in \{1, 2\}.$$

If we, in addition, assume the distribution of Y is absolutely continuous everywhere except at zero, then the marginal survival functions $\overline{F}_{Y_1}$ and $\overline{F}_{Y_2}$ have the range $[0, 1 - p_1]$ and $[0, 1 - p_2]$, respectively. As a consequence, given the marginal survival functions $\overline{F}_{Y_1}$ and $\overline{F}_{Y_2}$, every survival copula C_s for Y is unique on the rectangle $[0, 1 - p_1] \times [0, 1 - p_2]$.

Following the above notations, the subsequent relationship between the Lévy copula $\mathfrak{C}$ and the survival copula C_s can be derived.

Proposition 3.3. *Assume the Poisson frequency process in representation* (3.4) *of* $S(t)$ *has an intensity parameter* $\lambda > 0$ *and the bivariate single loss distribution* F_Y *is absolutely continuous everywhere except at zero. Then for all* $(u_1, u_2)^\top \in [0, 1 - p_1] \times [0, 1 - p_2]$*, the Lévy copula* $\mathfrak{C}$ *and the survival copula* C_s *satisfy the equation*

$$\lambda\, C_s(u_1, u_2) \;=\; \mathfrak{C}(\lambda u_1, \lambda u_2). \tag{3.6}$$

Proof. Recall from Chapter 2 that for $(x_1, x_2)^\top \in [0, \infty)^2 \setminus \{0\}$ the tail integral of the bivariate compound Poisson process $S(t) = (S_1(t), S_2(t))^\top$ is given by

$$\overline{\Pi}(x_1, x_2) \;=\; \lambda \mathbb{P}(Y_1 \geq x_1, Y_2 \geq x_2).$$

Thus the marginal tail integral $\overline{\Pi}_1$ of $\overline{\Pi}$, which also constitutes the tail integral of the marginal process $S_1(t)$, can be computed from $\overline{\Pi}$ for $x_1 > 0$ as

$$\overline{\Pi}_1(x_1) \;=\; \overline{\Pi}(x_1, 0) \;=\; \lambda \mathbb{P}(Y_1 \geq x_1, Y_2 \geq 0) \;=\; \lambda \mathbb{P}(Y_1 > x_1) \;=\; \lambda \overline{F}_{Y_1}(x_1), \tag{3.7}$$

where we have used that Y_1 is absolutely continuous on $(0, \infty)$ and Y_2 is completely supported by $[0, \infty)$. The tail integral of the marginal process $S_2(t)$ is obtained in an analogous manner and given by $\overline{\Pi}_2(x_2) = \lambda \overline{F}_{Y_2}(x_2)$ for $x_2 > 0$.

Now we use Sklar's theorem for survival functions as well as Sklar's theorem for Lévy processes to verify the claimed relationship between the survival copula C_s and the Lévy copula $\mathfrak{C}$. First, it follows from Sklar's theorem for Lévy processes that the tail integral $\overline{\Pi}$ has the representation

$$\overline{\Pi}(x_1, x_2) \;=\; \mathfrak{C}(\overline{\Pi}_1(x_1), \overline{\Pi}_2(x_2)) \;=\; \mathfrak{C}(\lambda \overline{F}_{Y_1}(x_1), \lambda \overline{F}_{Y_2}(x_2)), \quad (x_1, x_2)^\top \in (0, \infty)^2. \tag{3.8}$$

On the other hand, the tail integral $\overline{\Pi}(x_1, x_2)$ can be expressed by means of the joint survival function of $Y = (Y_1, Y_2)^\top$ through

$$\overline{\Pi}(x_1, x_2) \;=\; \lambda \mathbb{P}(Y_2 \geq x_2, Y_2 \geq x_2) \;=\; \lambda \overline{F}_Y(x_1, x_2), \quad (x_1, x_2)^\top \in (0, \infty)^2, \tag{3.9}$$

as by assumption the distribution of Y is absolutely continuous except at zero. Next we apply Sklar's theorem for survival functions with the survival copula C_s for Y and obtain

$$\overline{F}_Y(x_1, x_2) \;=\; C_s(\overline{F}_{Y_1}(x_1), \overline{F}_{Y_2}(x_2)), \quad (x_1, x_2)^\top \in (0, \infty)^2. \tag{3.10}$$

Putting everything together, from equations (3.8)-(3.10) it follows that

$$\lambda C_s(\overline{F}_{Y_1}(x_1), \overline{F}_{Y_2}(x_2)) \;=\; \mathfrak{C}(\lambda \overline{F}_{Y_1}(x_1), \lambda \overline{F}_{Y_2}(x_2)), \quad (x_1, x_2)^\top \in (0, \infty)^2.$$

Then by setting $u_1 = \overline{F}_{Y_1}(x_1)$ and $u_2 = \overline{F}_{Y_2}(x_2)$, we have verified equation (3.6) for $(u_1, u_2)^\top \in (0, 1 - p_1) \times (0, 1 - p_2)$. Furthermore, the validity of (3.6) can be extended onto the domain of $[0, 1 - p_1] \times [0, 1 - p_2]$ by the continuity of the survival copula C_s as well as the Lévy copula $\mathfrak{C}$. □

In view of the above proposition, it would be false to conclude that the relevant domain of the Lévy copula $\mathfrak{C}$ associated with a bivariate compound Poisson process is solely given by the set $[0, \lambda(1-p_1)] \times [0, \lambda(1-p_2)]$. According to Theorem 2.12, the Lévy copula is unique on the range product of the marginal tail integrals. Note the relationship $\overline{\Pi}_i(x_i) = \lambda \overline{F}_{Y_i}(x_i)$, $i \in \{1, 2\}$, only holds for $x_i > 0$, as the tail integral at zero is always fixed as infinity by Definition 2.11. Therefore, the Lévv copula $\mathfrak{C}$ is actually unique on the domain of $([0, \lambda(1 - p_1)] \cup \{\infty\}) \times ([0, \lambda(1 - p_2)] \cup \{\infty\})$. Nevertheless, with regard to the estimation of $\mathfrak{C}$, for example, its behaviour on the sets $[0, \lambda(1 - p_1)] \times \{\infty\}$, $\{\infty\} \times [0, \lambda(1 - p_2)]$ and $\{\infty\} \times \{\infty\}$ is of minor interest. This is due to the fact that a valid Lévy copula must have uniform margins, meaning that $\mathfrak{C}(u_1, \infty) = u_1$ for every $u_1 \in [0, \infty]$, and similarly, $\mathfrak{C}(\infty, u_2) = u_2$ for every $u_2 \in [0, \infty]$. Hence the behaviour of $\mathfrak{C}$ outside the range $[0, \lambda(1 - p_1)] \times [0, \lambda(1 - p_2)]$ does not contribute to the understanding of the dependence structure between the marginal processes. As a result, in the case of a bivariate compound Poisson model we assume the relevant domain of the corresponding Lévy copula is given by the rectangle $[0, \lambda(1 - p_1)] \times [0, \lambda(1 - p_2)]$ in the sequel.

As promised at the beginning of the present section, now we are going to derive the relationship between the bivariate losses $Y_h = (Y_{1h}, Y_{2h})^\top$, $h \geq 1$, and the univariate losses X_{1j}, $j \geq 1$, and X_{2l}, $l \geq 1$, of the marginal processes $S_1(t)$ and $S_2(t)$, respectively. More precisely, we assume representation (3.4) of the bivariate compound Poisson process $S(t)$ to be known, that is, the intensity parameter $\lambda > 0$ of the Poisson frequency process and the bivariate single loss distribution F_Y shall be given. Based upon these information, we want to find the parameters behind representation (3.3) of $S(t)$.

For simplicity of notation, we again consult a generic random vector $Y = (Y_1, Y_2)^\top$ with distribution function F_Y. As already mentioned, the marginal distributions F_{Y_i}, $i \in \{1, 2\}$, can have positive measure at zero. On the other hand, by Definition 2.2 the single losses of the one-dimensional compound Poisson processes $S_i(t)$, $i \in \{1, 2\}$, must not have an

atom at zero. As a result, it is a natural approach to introduce two new random variables X_1 and X_2 only taking the non-zero values of Y_1 and Y_2, respectively:

$$X_1 \overset{d}{:=} Y_1 \,|\, Y_1 > 0 \quad \text{and} \quad X_2 \overset{d}{:=} Y_2 \,|\, Y_2 > 0.$$

Recall the definition of the constants $p_i = \mathbb{P}(Y_i = 0)$, $i \in \{1, 2\}$. Then the distribution function F_{X_i} of X_i, $i \in \{1, 2\}$, is computed for $x_i \geq 0$ as

$$\begin{aligned} F_{X_i}(x_i) &= \mathbb{P}(X_i \leq x_i) = \mathbb{P}(Y_i \leq x_i \,|\, Y_i > 0) \\ &= \frac{\mathbb{P}(0 < Y_i \leq x_i)}{\mathbb{P}(Y_i > 0)} = \frac{\mathbb{P}(Y_i \leq x_i) - \mathbb{P}(Y_i = 0)}{1 - \mathbb{P}(Y_i = 0)} = \frac{F_{Y_i}(x_i) - p_i}{1 - p_i}. \end{aligned} \tag{3.11}$$

Now we claim that the single loss distributions of the marginal compound Poisson processes $S_i(t)$ are exactly given by F_{X_i}, $i \in \{1, 2\}$, and the intensities of the underlying marginal Poisson processes can be calculated by

$$\lambda_i = \lambda(1 - p_i) > 0, \quad i \in \{1, 2\}. \tag{3.12}$$

To see this, we make use of the uniqueness of the characteristic function for $S_i(t)$, $i \in \{1, 2\}$. From Section 2.1 we know the characteristic function of the one-dimensional compound Poisson process $S_i(t)$ is fully determined by its associated Lévy measure Π_i, which has a one-to-one relationship to the tail integral $\overline{\Pi}_i$. In (3.7) we have already established that the tail integrals are given by $\overline{\Pi}_i(x_i) = \lambda \overline{F}_{Y_i}(x_i)$, $i \in \{1, 2\}$. In order to obtain a representation through $\overline{F}_{X_i}$ instead of $\overline{F}_{Y_i}$, we first utilise the previous calculation (3.11) and solve for $F_{Y_i}(x_i)$, that is,

$$F_{Y_i}(x_i) = (1 - p_i) F_{X_i}(x_i) + p_i, \quad i \in \{1, 2\}.$$

Then the survival functions of Y_i, $i \in \{1, 2\}$, can be expressed in terms of the survival functions of X_i, $i \in \{1, 2\}$, respectively:

$$\overline{F}_{Y_i}(x_i) = (1 - p_i) \overline{F}_{X_i}(x_i), \quad i \in \{1, 2\}. \tag{3.13}$$

Consequently, together with the definition of $\lambda_i = \lambda(1 - p_i)$, $i \in \{1, 2\}$, it follows that

$$\overline{\Pi}_i(x_i) = \lambda \overline{F}_{Y_i}(x_i) = \lambda(1 - p_i) \overline{F}_{X_i}(x_i) = \lambda_i \overline{F}_{X_i}(x_i), \quad i \in \{1, 2\}.$$

Note the distributions of X_i, $i \in \{1, 2\}$, have by definition no atom at zero. Therefore, due to the specific form of the characteristic function of compound Poisson processes and the uniqueness of characteristic functions, the marginal process $S_i(t)$ must have F_{X_i} as severity distribution and λ_i as frequency intensity parameter for $i \in \{1, 2\}$.

3.2.2 A useful decomposition

The following decomposition lemma proves itself to be especially useful for the estimation and simulation of a bivariate compound Poisson model. Its theoretical background can be found in Section 5.5 of [CT04] and its proof is partially inspired by [Böc08].

Lemma 3.4 (Decomposition of a bivariate compound Poisson model). *Let $S(t) = (S_1(t), S_2(t))^\top$, $t \geq 0$, be a bivariate compound Poisson model in accordance with Definition 3.1 and let the dependence structure between its margins be given by a Lévy copula $\mathfrak{C}$. The marginal frequency processes $N_i(t)$, $i \in \{1,2\}$, shall have intensity parameters $\lambda_i > 0$, $i \in \{1,2\}$, and the marginal severity distribution functions are given by F_{X_i}, $i \in \{1,2\}$. Then the following holds:*

(a) Each marginal loss process $S_i(t)$, $i \in \{1,2\}$, can be decomposed into the sum of two compound Poisson processes $S_i^\perp(t)$ and $S_i^\parallel(t)$ with the representation

$$S_1(t) \;=\; S_1^\perp(t) + S_1^\parallel(t) \;=\; \sum_{j=1}^{N_1^\perp(t)} X_{1j}^\perp + \sum_{k=1}^{N^\parallel(t)} X_{1k}^\parallel, \quad t \geq 0, \tag{3.14}$$

$$S_2(t) \;=\; S_2^\perp(t) + S_2^\parallel(t) \;=\; \sum_{l=1}^{N_2^\perp(t)} X_{2l}^\perp + \sum_{k=1}^{N^\parallel(t)} X_{2k}^\parallel, \quad t \geq 0, \tag{3.15}$$

where the one-dimensional processes $S_1^\perp(t)$ and $S_2^\perp(t)$ describe the losses exclusively occurring in risk cell one and risk cell two, respectively, and the bivariate process $S^\parallel(t) = (S_1^\parallel(t), S_2^\parallel(t))^\top$ describes the simultaneous losses of both cells. The corresponding frequency processes are denoted by $N_1^\perp(t)$, $N_2^\perp(t)$ and $N^\parallel(t)$, the corresponding single losses by $X_{1j}^\perp$, $X_{2l}^\perp$ and $X_k^\parallel = (X_{1k}^\parallel, X_{2k}^\parallel)^\top$. Moreover, the processes $S_1^\perp(t)$, $S_2^\perp(t)$ and $S^\parallel(t)$ are independent from each other.

(b) The intensity parameter of the common loss frequency process $N^\parallel(t)$ is explicitly given by

$$\lambda^\parallel \;=\; \mathfrak{C}(\lambda_1, \lambda_2).$$

Subsequently, the intensities of the individual frequency processes $N_1^\perp(t)$ and $N_2^\perp(t)$ are calculated as

$$\lambda_1^\perp \;=\; \lambda_1 - \lambda^\parallel \quad \textit{and} \quad \lambda_2^\perp \;=\; \lambda_2 - \lambda^\parallel.$$

Moreover, the joint survival function of the common severities $(X_{1k}^\parallel, X_{2k}^\parallel)^\top$, $k \geq 1$, can be expressed in terms of the Lévy copula $\mathfrak{C}$ as

$$\overline{F}^\parallel(x_1, x_2) \;=\; \frac{1}{\lambda^\parallel}\mathfrak{C}(\lambda_1 \overline{F}_{X_1}(x_1), \lambda_2 \overline{F}_{X_2}(x_2)), \quad (x_1, x_2)^\top \in [0, \infty)^2. \tag{3.16}$$

Accordingly, the survival functions of the individual severities $X_{1j}^\perp$, $j \geq 1$, and $X_{2l}^\perp$, $l \geq 1$, are given by

$$\overline{F}_1^\perp(x_1) \;=\; \frac{1}{\lambda_1^\perp}\left[\lambda_1 \overline{F}_{X_1}(x_1) - \mathfrak{C}(\lambda_1 \overline{F}_{X_1}(x_1), \lambda_2)\right], \quad x_1 \geq 0,$$

$$\textit{and} \quad \overline{F}_2^\perp(x_2) \;=\; \frac{1}{\lambda_2^\perp}\left[\lambda_2 \overline{F}_{X_2}(x_2) - \mathfrak{C}(\lambda_1, \lambda_2 \overline{F}_{X_2}(x_2))\right], \quad x_2 \geq 0,$$

respectively.

Proof. A straightforward proof is based on the decomposition of the Lévy measure Π associated with the process $S(t) = (S_1(t), S_2(t))^\top$.

(a) By Definition 2.5, the Lévy measure $\Pi(B)$ of a Borel set $B \in \mathcal{B}([0,\infty)^2 \setminus \{0\})$ is the expected number of non-trivial losses per time unit with size in B, which can be formalised for the bivariate case as

$$\Pi(B) \;=\; \mathbb{E}\left[\,\#\left\{(\Delta S_1(t), \Delta S_2(t)) \in B \,|\, t \in [0,1] \,\wedge\, (\Delta S_1(t) \neq 0 \vee \Delta S_2(t) \neq 0)\right\}\right].$$

Among the non-trivial losses with size in B we differentiate between the losses with a non-zero entry in the first dimension, the losses with a non-zero entry in the second dimension, and finally the losses with non-zero entries in both dimensions. Accordingly, the measure $\Pi(B)$ can be split into

$$\Pi(B) \;=\; \tilde{\Pi}_1^\perp(B) + \tilde{\Pi}_2^\perp(B) + \Pi^\parallel(B),$$

where the three summands are given by

$$\begin{aligned}
\tilde{\Pi}_1^\perp(B) \;&=\; \mathbb{E}\left[\,\#\left\{(\Delta S_1(t),\, 0) \in B \mid t \in [0,1] \,\wedge\, \Delta S_1(t) \neq 0\right\}\right],\\
\tilde{\Pi}_2^\perp(B) \;&=\; \mathbb{E}\left[\,\#\left\{(0,\, \Delta S_2(t)) \in B \mid t \in [0,1] \,\wedge\, \Delta S_2(t) \neq 0\right\}\right]\\
\text{and}\quad \Pi^\parallel(B) \;&=\; \mathbb{E}\left[\,\#\left\{(\Delta S_1(t), \Delta S_2(t)) \in B \mid t \in [0,1] \,\wedge\, (\Delta S_1(t) \neq 0 \wedge \Delta S_2(t) \neq 0)\right\}\right].
\end{aligned}$$

By introducing the notations $B_1 = \{x_1 \,|\, (x_1, 0) \in B\}$ and $B_2 = \{x_2 \,|\, (0, x_2) \in B\}$, the measures $\tilde{\Pi}_1^\perp$ and $\tilde{\Pi}_2^\perp$ can be replaced by their one-dimensional projections:

$$\begin{aligned}
\Pi_1^\perp(B_1) \;&=\; \tilde{\Pi}_1^\perp(B) \;=\; \Pi(B_1 \times \{0\}),\\
\Pi_2^\perp(B_2) \;&=\; \tilde{\Pi}_2^\perp(B) \;=\; \Pi(\{0\} \times B_2).
\end{aligned}$$

Note the three measures $\Pi_1^\perp$, $\Pi_2^\perp$ and $\Pi^\parallel$ are well-defined Lévy measures according to Definition 2.5. Recall the characteristic function of $S(t)$ has a representation through an integral with respect to Π as given in (2.2). Together with the decomposition of Π stated above, we obtain for arbitrary $u = (u_1, u_2)^\top \in \mathbb{R}^2$ that

$$\begin{aligned}
\mathbb{E}\left[e^{i\langle u, S(t)\rangle}\right] \;&=\; \mathbb{E}\left[e^{iu_1 S_1(t) + iu_2 S_2(t)}\right]\\
&=\; \exp\left\{t \int_0^\infty \int_0^\infty \left(e^{iu_1x_1 + iu_2x_2} - 1\right) \Pi(\mathrm{d}x_1 \times \mathrm{d}x_2)\right\}\\
&=\; \exp\left\{t \int_0^\infty \int_0^\infty \left(e^{iu_1x_1 + iu_2x_2} - 1\right) (\tilde{\Pi}_1^\perp + \tilde{\Pi}_2^\perp + \Pi^\parallel)(\mathrm{d}x_1 \times \mathrm{d}x_2)\right\}\\
&=\; \exp\left\{t \int_0^\infty \left(e^{iu_1x_1} - 1\right) \Pi_1^\perp(\mathrm{d}x_1)\right\} \exp\left\{t \int_0^\infty \left(e^{iu_2x_2} - 1\right) \Pi_2^\perp(\mathrm{d}x_2)\right\}\\
&\quad\; \exp\left\{t \int_0^\infty \int_0^\infty \left(e^{iu_1x_1 + iu_2x_2} - 1\right) \Pi^\parallel(\mathrm{d}x_1 \times \mathrm{d}x_2)\right\}, \qquad (3.17)
\end{aligned}$$

where for the last equality we have used the fact that the measures $\tilde{\Pi}_1^\perp$ and $\tilde{\Pi}_2^\perp$ are solely supported by the sets $\{(x_1, 0) \,|\, x_1 \in [0,\infty)\}$ and $\{(0, x_2) \,|\, x_2 \in [0,\infty)\}$, respectively. Hence the corresponding integrals reduce to one-dimensional integrals with respect to the measures $\Pi_1^\perp$ and $\Pi_2^\perp$.

Now observe that each of the three integrals in (3.17) has the form of the characteristic function of a compound Poisson process as given in (2.2). The corresponding Lévy measures are given by the one-dimensional measures $\Pi_1^\perp$, $\Pi_2^\perp$, and the two-dimensional measure $\Pi^\parallel$, respectively. On these grounds, we first introduce a univariate compound Poisson process $S_1^\perp(t)$ determined by the measure $\Pi_1^\perp$, which only records losses in the first risk cell, when no loss occurs in the second cell at the same time. Conversely, let $S_2^\perp(t)$ denote the compound Poisson process determined by $\Pi_2^\perp$, and it only records losses in cell two, when no loss occurs in cell one. Hence in the sequel we call $S_1^\perp(t)$ and $S_2^\perp(t)$ the independent parts of $S(t)$. On the other hand, the third integral in (3.17) corresponds to a two-dimensional compound Poisson process having the Lévy measure $\Pi^\parallel$. We denote this process by $S^\parallel(t) = (S_1^\parallel(t), S_2^\parallel(t))^\top$ and call it the dependent part of $S(t)$, as $S^\parallel(t)$ describes the simultaneous losses in both risk cells.

Up to now we have established the decomposition of $(S_1(t), S_2(t))^\top$ as given in (3.14) and (3.15). Only the independence between the processes $S_1^\perp(t)$, $S_2^\perp(t)$, and $S^\parallel(t)$ remains to be shown. But this is easy to see when we once again look at the product of the three integrals in (3.17) and recall that the characteristic function of the sum of independent random variables is given by the product of the characteristic functions of the single random variables.

(b) As the marginal severity distributions F_{X_1} and F_{X_2} are absolutely continuous by assumption, the marginal intensity parameters can be recovered from the corresponding marginal tail integrals through

$$\lambda_1 = \lim_{x_1 \downarrow 0} \lambda_1 \overline{F}_{X_1}(x_1) = \lim_{x_1 \downarrow 0} \overline{\Pi}_1(x_1) \tag{3.18}$$

$$\text{and} \quad \lambda_2 = \lim_{x_2 \downarrow 0} \lambda_2 \overline{F}_{X_2}(x_2) = \lim_{x_2 \downarrow 0} \overline{\Pi}_2(x_2). \tag{3.19}$$

On the other hand, we can utilise the decomposition of the Lévy measure Π, as this was explained in Part (a) of the proof, in order to obtain another representation of the tail integrals $\overline{\Pi}_1$ and $\overline{\Pi}_2$. More precisely, it holds for $x_1 > 0$ that

$$\begin{aligned}
\overline{\Pi}_1(x_1) &= \overline{\Pi}(x_1, 0) = \Pi([x_1, \infty) \times [0, \infty)) \\
&= \Pi([x_1, \infty) \times \{0\}) + \Pi([x_1, \infty) \times (0, \infty)) \\
&= \Pi_1^\perp([x_1, \infty)) + \lim_{x_2 \downarrow 0} \Pi^\parallel([x_1, \infty) \times [x_2, \infty)) \\
&= \overline{\Pi}_1^\perp(x_1) + \lim_{x_2 \downarrow 0} \overline{\Pi}^\parallel(x_1, x_2) \\
&= \overline{\Pi}_1^\perp(x_1) + \lim_{x_2 \downarrow 0} \mathfrak{C}(\overline{\Pi}_1(x_1), \overline{\Pi}_2(x_2)),
\end{aligned} \tag{3.20}$$

whereby we have used the continuity of the Lévy measure $\Pi^\parallel$ on its support $(0, \infty)^2$ and its coincidence with the measure Π in the same domain, so that Sklar's theorem for Lévy processes applies and the second summand in the last line is replaced by a representation through the Lévy copula $\mathfrak{C}$ and the marginal tail integrals. Next, by taking the limit $x_1 \downarrow 0$ on both sides of the equation and the continuity of the Lévy

copula $\mathfrak{C}$, we obtain

$$\begin{aligned}\lim_{x_1\downarrow 0}\overline{\Pi}_1(x_1) &= \lim_{x_1\downarrow 0}\overline{\Pi}_1^{\perp}(x_1) + \lim_{x_1,x_2\downarrow 0}\mathfrak{C}(\overline{\Pi}_1(x_1),\overline{\Pi}_2(x_2))\\ &= \lim_{x_1\downarrow 0}\overline{\Pi}_1^{\perp}(x_1) + \mathfrak{C}(\lim_{x_1\downarrow 0}\overline{\Pi}_1(x_1),\lim_{x_2\downarrow 0}\overline{\Pi}_2(x_2)).\end{aligned}$$

Together with (3.18) and (3.19), it follows that

$$\lambda_1 = \lim_{x_1\downarrow 0}\overline{\Pi}_1^{\perp}(x_1) + \mathfrak{C}(\lambda_1,\lambda_2).$$

Similarly, the intensity parameter λ_2 associated with the marginal process $S_2(t)$ can be decomposed into

$$\lambda_2 = \lim_{x_2\downarrow 0}\overline{\Pi}_2^{\perp}(x_2) + \mathfrak{C}(\lambda_1,\lambda_2).$$

Now we set $\lambda^{\parallel} = \mathfrak{C}(\lambda_1,\lambda_2)$ and convince ourselves that $\lambda^{\parallel}$ is indeed the parameter of the Poisson process underlying the bivariate dependence part $S^{\parallel}(t)$ of $S(t)$. This follows from the fact that, by construction, the Lévy measure $\Pi^{\parallel}$ of the bivariate compound Poisson process $S^{\parallel}(t)$ is only supported by $(0,\infty)^2$. Therefore, the intensity parameter $\lambda^{\parallel}$ can be actually computed by taking the limit of the corresponding tail integral $\lim_{x_1,x_2\downarrow 0}\overline{\Pi}^{\parallel}(x_1,x_2)$, which results in the above definition of $\lambda^{\parallel}$. Furthermore, the frequency parameters of the univariate independent loss processes $S_1^{\perp}(t)$ and $S_2^{\perp}(t)$ are given by

$$\begin{aligned}\lambda_1^{\perp} &= \lim_{x_1\downarrow 0}\overline{\Pi}_1^{\perp}(x_1) = \lambda_1 - \mathfrak{C}(\lambda_1,\lambda_2)\\ \text{and}\quad \lambda_2^{\perp} &= \lim_{x_2\downarrow 0}\overline{\Pi}_2^{\perp}(x_2) = \lambda_2 - \mathfrak{C}(\lambda_1,\lambda_2).\end{aligned}$$

Finally, only the expressions for the survival functions of the independent as well as the dependent single losses remain to be shown. Again, note the Lévy measure $\Pi^{\parallel}$ is completely supported by $(0,\infty)^2$, thus the survival function $\overline{F}^{\parallel}$ can be retrieved from the corresponding tail integral $\overline{\Pi}^{\parallel}$ for $(x_1,x_2)^{\top}\in[0,\infty)^2$, that is,

$$\begin{aligned}\overline{F}^{\parallel}(x_1,x_2) &= \frac{1}{\lambda^{\parallel}}\overline{\Pi}^{\parallel}(x_1,x_2) = \frac{1}{\lambda^{\parallel}}\mathfrak{C}(\overline{\Pi}_1(x_1),\overline{\Pi}_2(x_2))\\ &= \frac{1}{\lambda^{\parallel}}\mathfrak{C}(\lambda_1\overline{F}_{X_1}(x_1),\lambda_2\overline{F}_{X_2}(x_2)).\end{aligned}$$

Likewise, in order to compute the severity survival function $\overline{F}_1^{\perp}$ underlying the independent marginal process $S_1^{\perp}(t)$, we replace the tail integrals in equation (3.20) by the corresponding products of intensity parameter and survival function, yielding

$$\begin{aligned}\lambda_1\overline{F}_{X_1}(x_1) &= \lambda_1^{\perp}\overline{F}_1^{\perp}(x_1) + \lim_{x_2\downarrow 0}\mathfrak{C}(\lambda_1\overline{F}_{X_1}(x_1),\lambda_2\overline{F}_{X_2}(x_2))\\ &= \lambda_1^{\perp}\overline{F}_1^{\perp}(x_1) + \mathfrak{C}(\lambda_1\overline{F}_{X_1}(x_1),\lambda_2),\quad x_1\geq 0.\end{aligned}$$

Then by simply rearranging the terms, the claimed representation of the survival function $\overline{F}_1^{\perp}$ follows. The severity survival function $\overline{F}_2^{\perp}$ of the independent loss process $S_2^{\perp}(t)$ is computed in an analogous manner. □

Let $X^{\parallel} = (X_1^{\parallel}, X_2^{\parallel})^\top$ be a generic random vector having the same distribution as the dependent severities $X^{\parallel} = (X_{1k}^{\parallel}, X_{2k}^{\parallel})^\top$, $k \geq 1$. Given formula (3.16) of the joint survival function, the marginal survival functions can be calculated as

$$\overline{F}_1^{\parallel}(x_1) = \lim_{x_2 \downarrow 0} \overline{F}^{\parallel}(x_1, x_2) = \frac{1}{\lambda^{\parallel}} \mathfrak{C}(\lambda_1 \overline{F}_{X_1}(x_1), \lambda_2), \quad x_1 \geq 0, \tag{3.21}$$

$$\text{and} \quad \overline{F}_2^{\parallel}(x_2) = \lim_{x_1 \downarrow 0} \overline{F}^{\parallel}(x_1, x_2) = \frac{1}{\lambda^{\parallel}} \mathfrak{C}(\lambda_1, \lambda_2 \overline{F}_{X_2}(x_2)), \quad x_2 \geq 0. \tag{3.22}$$

By construction, $X^{\parallel}$ has absolutely continuous margins. As a direct consequence of Sklar's theorem for survival functions, there exists a unique survival copula $C_s^{\parallel} : [0,1]^2 \to [0,1]$ satisfying

$$\overline{F}^{\parallel}(x_1, x_2) = C_s^{\parallel}(\overline{F}_1^{\parallel}(x_1), \overline{F}_2^{\parallel}(x_2)), \quad (x_1, x_2)^\top \in [0, \infty)^2. \tag{3.23}$$

At this point it is understandable to ask for the relationship between the survival copula $C_s^{\parallel}$ of the dependent severities $X^{\parallel}$ and the survival copula C_s of the severity vector Y introduced in Section 3.2.1. For this purpose, we utilise the relationship between C_s and the Lévy copula $\mathfrak{C}$ established in Proposition 3.3 and calculate

$$\lambda^{\parallel} = \mathfrak{C}(\lambda_1, \lambda_2) = \lambda\, C_s\left(\lambda_1 \lambda^{-1}, \lambda_2 \lambda^{-1}\right).$$

According to equation (3.12), the product $\lambda_i \lambda^{-1}$ is given by $1 - p_i$ for $i \in \{1, 2\}$, whereby the constant p_i is the probability of the marginal variable Y_i attaining the value zero. Hence we further rewrite $\lambda^{\parallel}$ as

$$\lambda^{\parallel} = \lambda C_s(1 - p_1, 1 - p_2) = \lambda C_s(\overline{F}_{Y_1}(0), \overline{F}_{Y_2}(0)) = \lambda \overline{F}_Y(0, 0). \tag{3.24}$$

Similarly, it holds that

$$\begin{aligned}
\mathfrak{C}(\lambda_1 \overline{F}_{X_1}(x_1), \lambda_2 \overline{F}_{X_2}(x_2)) &= \lambda\, C_s(\lambda_1 \lambda^{-1} \overline{F}_{X_1}(x_1), \lambda_2 \lambda^{-1} \overline{F}_{X_2}(x_2)) \\
&= \lambda\, C_s((1 - p_1) \overline{F}_{X_1}(x_1), (1 - p_2) \overline{F}_{X_2}(x_2)) \\
&= \lambda\, C_s(\overline{F}_{Y_1}(x_1), \overline{F}_{Y_2}(x_2)) \\
&= \lambda\, \overline{F}_Y(x_1, x_2),
\end{aligned}$$

and

$$\begin{aligned}
\mathfrak{C}(\lambda_1 \overline{F}_{X_1}(x_1), \lambda_2) &= \lambda\, C_s(\overline{F}_{Y_1}(x_1), \overline{F}_{Y_2}(0)) = \lambda\, \overline{F}_Y(x_1, 0), \\
\mathfrak{C}(\lambda_1, \lambda_2 \overline{F}_{X_2}(x_2)) &= \lambda\, C_s(\overline{F}_{Y_1}(0), \overline{F}_{Y_2}(x_2)) = \lambda\, \overline{F}_Y(0, x_2),
\end{aligned}$$

where we have used relation (3.13) between the survival functions of Y_i and X_i with $i \in \{1, 2\}$, respectively. Putting things together, the joint survival function of $X^{\parallel}$ can be rewritten as

$$\overline{F}^{\parallel}(x_1, x_2) = \frac{1}{\lambda^{\parallel}} \mathfrak{C}(\lambda_1 \overline{F}_{X_1}(x_1), \lambda_2 \overline{F}_{X_2}(x_2)) = \frac{\overline{F}_Y(x_1, x_2)}{\overline{F}_Y(0, 0)},$$

and the corresponding marginal survival functions are given by

$$\begin{aligned} \overline{F}_1^{\parallel}(x_1) &= \frac{1}{\lambda^{\parallel}}\mathfrak{C}(\lambda_1\overline{F}_{X_1}(x_1),\lambda_2) = \frac{\overline{F}_Y(x_1,0)}{\overline{F}_Y(0,0)} \\ \text{and} \quad \overline{F}_2^{\parallel}(x_2) &= \frac{1}{\lambda^{\parallel}}\mathfrak{C}(\lambda_1,\lambda_2\overline{F}_{X_2}(x_2)) = \frac{\overline{F}_Y(0,x_2)}{\overline{F}_Y(0,0)}. \end{aligned}$$

As a result, the marginal severities $X_1^{\parallel}$ and $X_2^{\parallel}$ have the same distribution as the conditional random variables

$$X_1^{\parallel} \stackrel{d}{:=} Y_1 \,|\, Y_1, Y_2 > 0 \quad \text{and} \quad X_2^{\parallel} \stackrel{d}{:=} Y_2 \,|\, Y_1, Y_2 > 0,$$

respectively. Hence the survival copula $C_s^{\parallel}$ of $X^{\parallel} = (X_1^{\parallel}, X_2^{\parallel})^\top$ must comply with

$$\frac{\overline{F}_Y(x_1,x_2)}{\overline{F}_Y(0,0)} = C_s^{\parallel}\left(\frac{\overline{F}_Y(x_1,0)}{\overline{F}_Y(0,0)}, \frac{\overline{F}_Y(0,x_2)}{\overline{F}_Y(0,0)}\right).$$

As already explained, the survival copula C_s of $Y = (Y_1, Y_2)^\top$ is only unique on the rectangle $[0, 1-p_1]\times[0, 1-p_2]$, as the marginal distributions F_{Y_1} and F_{Y_2} may have an atom at zero. Therefore, by viewing the random variables $X_1^{\parallel}$ and $X_2^{\parallel}$ as the conditional version of Y_1 and Y_2 on the requirement $Y_1, Y_2 > 0$, the survival copula $C_s^{\parallel}$ of $X^{\parallel} = (X_1^{\parallel}, X_2^{\parallel})^\top$ is precisely the normalised version of C_s onto the unit square $[0, 1]^2$.

Unfortunately, there is no general closed-form expression characterising the copula $C_s^{\parallel}$ in relation to the ordinary survival copula C_s or to the Lévy copula $\mathfrak{C}$. Only in the special case of Archimedean Lévy copulas of the form $\mathfrak{C}(u_1, u_2) = \phi^{-1}[\phi(u_1) + \phi(u_2)]$, elementary manipulations of equation (3.23) show that $C_s^{\parallel}$ has a representation in terms of the generator function ϕ as

$$C_s^{\parallel}(u_1, u_2) = \frac{1}{\lambda^{\parallel}}\,\phi^{-1}\left[\phi(\lambda^{\parallel}u_1) + \phi(\lambda^{\parallel}u_2) - \phi(\lambda^{\parallel})\right] \tag{3.25}$$

for $(u_1, u_2)^\top \in [0, 1]^2$. In particular, expression (3.25) is independent of the time t as well as the marginal severity distributions F_{X_1} and F_{X_2}.

3.2.3 Attainable range of frequency correlation

It is worth taking a moment to think about the attainable values of the frequency parameter $\lambda^{\parallel}$ underlying the dependent loss process $S^{\parallel}(t)$ and the implication thereof for the dependence structure within a bivariate model, when assuming the marginal frequency parameters λ_1 and λ_2 are already given. As the independent loss intensities are calculated as the difference $\lambda_i^{\perp} = \lambda_i - \lambda^{\parallel}$ for $i \in \{1, 2\}$, which must be non-negative, it is natural to restrain $\lambda^{\parallel}$ to the range of

$$0 \le \lambda^{\parallel} \le \min\{\lambda_1, \lambda_2\}. \tag{3.26}$$

Intuitively speaking, the case of $\lambda^{\parallel} = 0$ implies that losses in the two risk cells never occur at the same time and this reflects the understanding of independence in the framework of

Lévy processes. In the other extreme case of $\lambda^{\|} = \min\{\lambda_1, \lambda_2\}$, the expected number of simultaneous losses in a time interval is equal to the expected number of losses in the risk cell with the lower frequency parameter during the same period. The latter presents the strongest possible frequency dependence in a bivariate compound Poisson model. From a more theoretical point of view, from Lemma 3.4 we already know the value of $\lambda^{\|}$ directly depends on the underlying Lévy copula and is given by the formula $\lambda^{\|} = \mathfrak{C}(\lambda_1, \lambda_2)$. Now recall from Chapter 2 the definition of the independence and the complete dependence Lévy copula. The application of these two special copulas results in

$$\mathfrak{C}_{\perp}(\lambda_1, \lambda_2) \;=\; 0 \quad \text{and} \quad \mathfrak{C}_{\|}(\lambda_1, \lambda_2) \;=\; \min\{\lambda_1, \lambda_2\},$$

respectively. Therefore, the theoretical result matches our natural understanding of dependence and independence in compound Poisson models. Going one step further, it is easy to verify that all well-defined Lévy copulas $\mathfrak{C}$ indeed provide a value of $\lambda^{\|}$ within the bounds given in (3.26). The lower bound is trivial as all positive Lévy copulas are functions onto $[0, \infty]$. The upper bound is more interesting and can be shown by utilising Proposition 3.3 as follows:

$$\lambda^{\|} \;=\; \mathfrak{C}(\lambda_1, \lambda_2) \;=\; \lambda C_s\left(\lambda_1\lambda^{-1}, \lambda_2\lambda^{-1}\right) \;\leq\; \lambda \min\left\{\lambda_1\lambda^{-1}, \lambda_2\lambda^{-1}\right\},$$

where for the last equation we have used the upper Fréchet-Hoeffding bound for ordinary copulas. Therefore, the upper bound in (3.26) is precisely attained if the survival copula C_s of the bivariate losses $(Y_{1h}, Y_{2h})^{\top}$, $h \geq 1$, is given by the ordinary comonotonic copula and if $\lambda = \max\{\lambda_1, \lambda_2\}$ holds. The latter translates to the situation where the expected number λ of the bivariate losses, which may have one component equal to zero but not both, is equal to the expected total number of losses in the risk cell with the higher frequency parameter. This is just an equivalent interpretation of the greatest possible positive dependence between the two risk cells and once again we see how the two concepts of dependence characterisation through $\mathfrak{C}$ and C_s fit together.

As already indicated in Section 1.2, one of the popular approaches for dependence modelling in operational risk is to incorporate a dependence structure among the frequency distributions or processes of different risk cells. Hence for comparison purpose, we state below the implied Pearson's correlation coefficient between the marginal Poisson frequency processes $N_1(t)$ and $N_2(t)$ in terms of the parameters $\lambda^{\|}$, λ_1 and λ_2. Since at any time point t the frequency count $N_i(t)$, $i \in \{1, 2\}$, is Poisson distributed with parameter $\lambda_i t$, its variance is simply given by $\mathrm{Var}[N_i(t)] = \lambda_i t$. Moreover, by utilising the decomposition $N_i(t) = N_i^{\perp}(t) + N^{\|}(t)$ from Lemma 3.4, the covariance can be calculated through

$$\begin{aligned}
&\mathrm{Cov}[N_1(t), N_2(t)] \\
&\quad = \mathrm{Cov}[N_1^{\perp}(t) + N^{\|}(t), N_2^{\perp}(t) + N^{\|}(t)] \\
&\quad = \mathrm{Cov}[N_1^{\perp}(t), N_2^{\perp}(t)] + \mathrm{Cov}[N_1^{\perp}(t), N^{\|}(t)] + \mathrm{Cov}[N_2^{\perp}(t), N^{\|}(t)] + \mathrm{Cov}[N^{\|}(t), N^{\|}(t)].
\end{aligned}$$

Because the processes $N_1^{\perp}(t)$, $N_2^{\perp}(t)$ and $N^{\|}(t)$ are mutually independent, only the last summand remains and it is equal to $\mathrm{Var}[N^{\|}(t)] = \lambda^{\|} t$. Altogether we obtain the correlation coefficient

$$\mathrm{Corr}[N_1(t), N_2(t)] \;=\; \frac{\mathrm{Cov}[N_1(t), N_2(t)]}{\sqrt{\mathrm{Var} N_1(t)}\sqrt{\mathrm{Var} N_2(t)}} \;=\; \frac{\lambda^{\|}}{\sqrt{\lambda_1\lambda_2}}.$$

In particular, the correlation is independent of t and can only attain non-negative values.

3.2.4 Examples of bivariate Lévy copulas

At the end of Chapter 2 we have already listed some examples of Archimedean Lévy copulas. In the current section, we want to enlarge our toolbox and introduce several more parametric families of Lévy copulas. As the focus currently lies on the case of bivariate compound Poisson models, we state below the copulas in their bivariate form.

In Lemma 3.4 we have seen how dependence in both frequency and severity is influenced by the underlying Lévy copula. Of course, this modelling scheme shall also cover the special case in which the components of the common losses $X_k^{\|} = (X_{1k}^{\|}, X_{2k}^{\|})^\top$, $k \geq 1$, are independent, whereas dependence between the marginal frequencies is allowed. This gives rise to the pure common shock Lévy copula introduced by [ACW11].

Example 3.5 (**Pure common shock Lévy copula**).
A pure common shock Lévy copula is specified through the marginal Poisson intensity parameters $\lambda_1, \lambda_2 \geq 0$ and a third parameter $0 \leq \theta \leq \min\{\lambda_1^{-1}, \lambda_2^{-1}\}$ adjusting the strength in frequency dependence. For any $(u_1, u_2)^\top \in [0, \infty]^2$, the Lévy copula has the form

$$\begin{aligned}\mathfrak{C}(u_1, u_2) \;=\; & \theta \min\{u_1, \lambda_1\} \min\{u_2, \lambda_2\} \\ & + \left(u_1 - \theta\lambda_2 \min\{u_1, \lambda_1\}\right) \mathbb{1}_{\{u_2=\infty\}} \;+\; \left(u_2 - \theta\lambda_1 \min\{u_2, \lambda_2\}\right) \mathbb{1}_{\{u_1=\infty\}}.\end{aligned}$$

The pure common shock Lévy copula is indeed a well-defined positive Lévy copula, as it satisfies all three necessary properties therefore from Definition 2.9. By Lemma 3.4, the intensity parameter $\lambda^{\|}$ underlying the dependent loss process $S^{\|}(t)$ can be calculated as $\lambda^{\|} = \mathfrak{C}(\lambda_1, \lambda_2) = \theta\lambda_1\lambda_2$. Hence the full possible range of frequency dependence given in (3.26) is attainable through varying the parameter θ between 0 and $\min\{\lambda_1^{-1}, \lambda_2^{-1}\}$. Furthermore, the survival copula $C_s^{\|}$ associated with the dependent losses $X_k^{\|} = (X_{1k}^{\|}, X_{2k}^{\|})^\top$, $k \geq 1$, can be explicitly derived from its defining equation (3.23) and is simply given by the ordinary independence copula $C_s^{\|}(u_1, u_2) = u_1 u_2$. This shows the Lévy copula model is a very general concept and encompasses certain inter-cell frequency dependence models introduced in Section 1.2.

Besides the named Archimedean Lévy generators presented in Table 2.1, there are three more Lévy copulas of Archimedean type encountered in literature, for example see [ACW11], [BL07] and [Ket12].

Example 3.6 (**More Archimedean Lévy copulas**).

(1) Generated by $\phi(u) = \left(e^{\theta u} - 1\right)^{-1}$ with $\theta > 0$:

$$\mathfrak{C}(u_1, u_2) \;=\; \frac{1}{\theta} \ln\left[\left(\frac{1}{e^{\theta u_1} - 1} + \frac{1}{e^{\theta u_2} - 1}\right)^{-1} + 1\right].$$

(2) Generated by $\phi(u) = (e^u - 1)^{-\theta}$ with $\theta > 0$:

$$\mathfrak{C}(u_1, u_2) \;=\; \ln\left[\left(\frac{1}{(e^{u_1}-1)^{\theta}} + \frac{1}{(e^{u_2}-1)^{\theta}}\right)^{-\frac{1}{\theta}} + 1\right].$$

(3) Generated by $\phi(u) = \exp(u^{-\theta}) - 1$ with $\theta > 0$:

$$\mathfrak{C}(u_1, u_2) \;=\; \left[\ln\left(\exp(u_1^{-\theta}) + \exp(u_2^{-\theta}) - 1\right)\right]^{-\frac{1}{\theta}}.$$

The third example is sometimes called the complementary Gumbel Lévy copula, as its generator is the inverse function of the generator for the Gumbel Lévy copula, when reparameterising θ to θ^{-1} in either formulation.

At this point is should be noticed that the full range of frequency dependence given in (3.26) cannot be attained by all introduced Lévy copulas. More specifically, the Archimedean family in Example 3.6 (1) does not approach the independence Lévy copula $\mathfrak{C}_{\perp}$ as its parameter θ tends to zero, but instead the limit $\lim_{\theta\downarrow 0} \mathfrak{C}(u_1, u_2) = \frac{u_1 u_2}{u_1+u_2}$ holds. Nonetheless, this Lévy copula approaches the complete dependence Lévy copula $\mathfrak{C}_{\|}$ for $\theta \uparrow \infty$. Furthermore, the Ali-Mikhail-Haq Lévy copula from Table 2.1 tends to neither of the extreme dependence structures for limiting values of its parameter θ. In contrast, the limit relations $\lim_{\theta\downarrow -1} \mathfrak{C}(u_1, u_2) = \frac{u_1 u_2}{2+u_1+u_2}$ and $\lim_{\theta\uparrow 1} \mathfrak{C}(u_1, u_2) = \frac{u_1 u_2}{u_1+u_2}$ hold. On the other hand, the Clayton, the Gumbel, the complementary Gumbel and the second Lévy copula in Example 3.6 approach the independence copula as their corresponding parameter falls down to zero, as well as tend to the complete dependence copula as the parameter rises up to infinity.

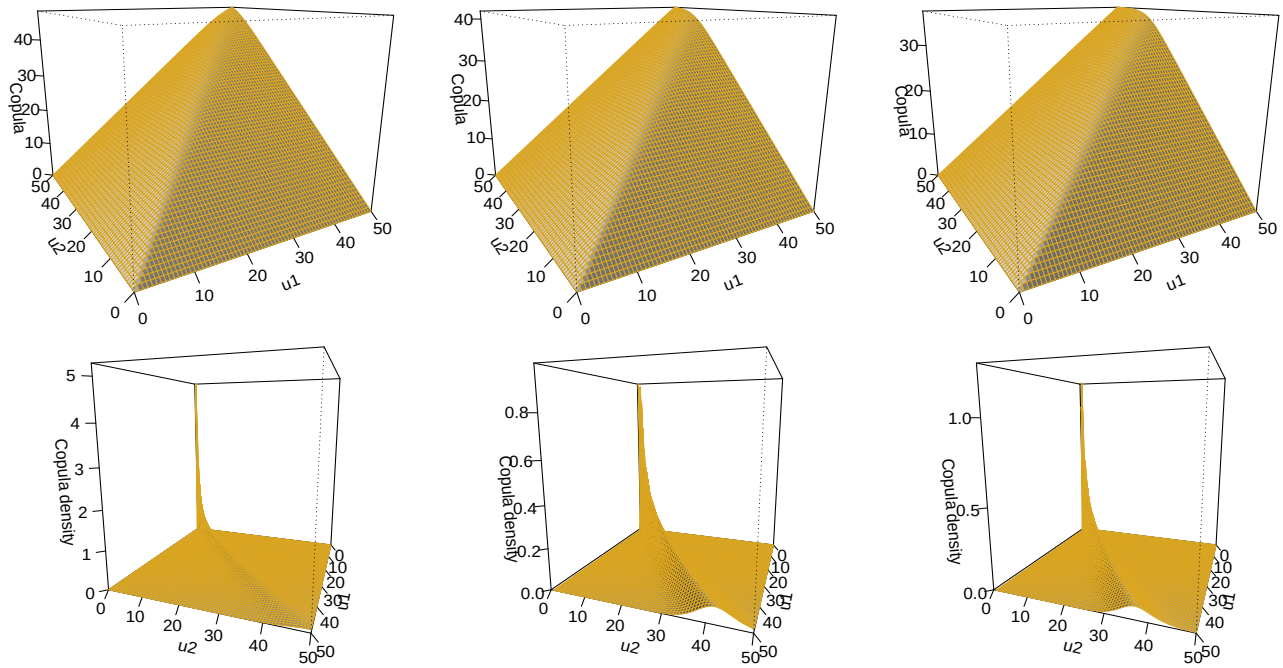
Obviously, all so far presented copulas satisfy the symmetry property $\mathfrak{C}(u_1, u_2) = \mathfrak{C}(u_2, u_1)$ for arbitrary $(u_1, u_2)^{\top} \in [0, \infty]^2$. However, it is desirable to capture potential asymmetric dependence structures between the marginal processes as well. Hence we introduce the following skewed Lévy copula from [MM13] with three parameters.

Example 3.7 (**Skewed Clayton Lévy copula**).
Given parameters $\theta > 0$, $\omega > 0$, and $0 < \kappa \leq \theta + 1$, the function $\mathfrak{C} : [0, \infty]^2 \to [0, \infty]$ defined as

$$\mathfrak{C}(u_1, u_2) \;=\; \left[\left(\omega u_2^{-\kappa} + 1\right) u_1^{-\theta} + u_2^{-\theta}\right]^{-\frac{1}{\theta}}.$$

is a valid Lévy copula.

Figure 3.1 illustrates two Clayton Lévy copulas with different asymmetry degree compared to a symmetric Clayton Lévy copula with the same parameter θ. The density plots in the bottom panel are especially informative with respect to demonstrating the diverse induced dependence structures. Note that small values of u_1 and u_2 correspond to large loss sizes in the compound Poisson setting, since severity survival functions enter as arguments into the Lévy copula. Hence all three copulas depicted in Figure 3.1 share the feature that the



(a) Symmetric Clayton Lévy copula with $\theta = 15$.

(b) Skewed Clayton Lévy copula with $\theta = 15$, $\omega = 15$ and $\kappa = 0.05$.

(c) Skewed Clayton Lévy copula with $\theta = 15$, $\omega = 100$ and $\kappa = 0.05$.

Figure 3.1: Comparison of symmetric and skewed Clayton Lévy copulas with increasing asymmetry from left to right. Top panel: copula function. Bottom panel: copula density.

induced loss severities tend to be simultaneously large in both risk cells, as the density is increasingly concentrated on small u_1 and u_2. On the other hand, for the same value of u_1 the density under a skewed copula is more concentrated on smaller values of u_2 in comparison to the symmetric copula. As a result, the same loss size in the first component is more likely to be accompanied by a larger loss in the second component.

We complete our example list of bivariate Lévy copulas by constructing two elliptical Lévy copulas from their ordinary counterparts in the spirit of Proposition 2.18.

Example 3.8 (**Elliptical Lévy copulas**).
Let f be a construction function satisfying the properties specified in Proposition 2.18.

(1) The Gaussian Lévy copula is given by

$$\mathfrak{C}(u_1, u_2) \;=\; f \circ \Phi_2 \left(\Phi^{-1} \circ f^{-1}(u_1),\, \Phi^{-1} \circ f^{-1}(u_2)\,;\, \rho\right),$$

where $\Phi(u)$ denotes the distribution function of the standard univariate normal distribution and $\Phi_2(u_1, u_2; \rho)$ is the bivariate normal distribution function with zero means, unit variances and correlation ρ.

(2) Similarly, the Student's t Lévy copula is constructed as

$$\mathfrak{C}(u_1, u_2) \;=\; f \circ \mathcal{T}_2 \left(\mathcal{T}_\nu^{-1} \circ f^{-1}(u_1),\, \mathcal{T}_\nu^{-1} \circ f^{-1}(u_2)\,;\, \nu, \rho\right),$$

where $\mathcal{T}_\nu(u)$ is the distribution function of the standard univariate t distribution with ν degrees of freedom and $\mathcal{T}_2(u_1, u_2; \nu, \rho)$ represents the bivariate t distribution function with ν degrees of freedom, zero means, unit variances and correlation ρ.

3.3 Maximum likelihood estimation of bivariate compound Poisson models

Under the assumption of parametric Lévy copulas and absolutely continuous marginal severity distributions, the likelihood function of a bivariate compound Poisson model $S(t) = (S_1(t), S_2(t))^\top$, $t \geq 0$, can be derived with the help of Lemma 3.4. In Section 3.3.1 we first treat the case of a continuous observation scheme. That is, given a fixed time interval $[0, T]$, the size and the occurrence time of all losses are known accurately. Consequently, all recorded single losses can be assigned either to the dependent part $S^{\|}(t)$ of the process $S(t)$, or to one of the independent parts $S_1^{\perp}(t)$ and $S_2^{\perp}(t)$, respectively. On the other hand, if the loss times are only known up to a short period, for example within one week, then we speak of a discrete observation scheme and the corresponding estimation method is presented in Section 3.3.2.

To begin with, we have to clarify the parameters to be estimated under the assumption of a bivariate compound Poisson model. First, we want to obtain the intensity parameters $\lambda_i > 0$, $i \in \{1, 2\}$, of the marginal Poisson frequency processes. Next, let $f_{X_i}(x_i; \theta_i)$ denote the marginal severity density of risk cell i, $i \in \{1, 2\}$, where θ_i represents the corresponding parameter vector to be estimated. Finally, the parameter $\theta_{\mathfrak{C}}$ of the Lévy copula $\mathfrak{C}(u_1, u_2; \theta_{\mathfrak{C}})$ is of interest as well.

3.3.1 MLE under a continuous observation scheme

For now suppose we can gather the following information from the observed data:

- the observation interval $[0, T]$,
- the number $n_1^{\perp}$ of losses occurred in risk cell one without losses occurring in risk cell two at the same time, and similarly, the number $n_2^{\perp}$ of the independent losses in risk cell two,
- the number $n^{\|}$ of losses simultaneously occurred in both risk cells and thus belonging to the bivariate dependent process $S^{\|}(t)$,
- the univariate severities $x_{11}^{\perp}, \ldots, x_{1n_1^{\perp}}^{\perp}$ attributed to the independent process $S_1^{\perp}(t)$ and the univariate severities $x_{21}^{\perp}, \ldots, x_{2n_2^{\perp}}^{\perp}$ attributed to the independent process $S_2^{\perp}(t)$,
- the bivariate loss severities $(x_{11}^{\|}, x_{21}^{\|})^\top, \ldots, (x_{1n^{\|}}^{\|}, x_{2n^{\|}}^{\|})^\top$ attributed to the dependent process $S^{\|}(t)$.

In order to apply MLE, we first specify the likelihood function:

Theorem 3.9 (**Likelihood function under a continuous observation scheme**). *Consider a parametric bivariate compound Poisson model as described above and assume*

for all $(u_1, u_2)^\top \in (0, \lambda_1) \times (0, \lambda_2)$ *the density* $\frac{\partial^2}{\partial u_1 \partial u_2}\mathfrak{C}(u_1, u_2; \theta_{\mathfrak{C}})$ *of the Lévy copula exists. Furthermore, let* $\lambda^{\parallel}_{\theta_{\mathfrak{C}}} = \mathfrak{C}(\lambda_1, \lambda_2; \theta_{\mathfrak{C}})$ *denote the frequency parameter of the common losses which depends on the copula parameter* $\theta_{\mathfrak{C}}$*. Accordingly, the frequency parameters of the individual losses are given by* $\lambda^{\perp}_{i\theta_{\mathfrak{C}}} = \lambda_i - \lambda^{\parallel}_{\theta_{\mathfrak{C}}}$*,* $i \in \{1, 2\}$*. Then the likelihood function of the bivariate compound Poisson model is computed as*

$$\begin{aligned}
&\mathcal{L}(\lambda_1, \lambda_2, \theta_1, \theta_2, \theta_{\mathfrak{C}}) \qquad (3.27)\\
&= (\lambda_1)^{n_1^\perp} e^{-\lambda^\perp_{1\theta_{\mathfrak{C}}} T} \prod_{j=1}^{n_1^\perp} \left[f_{X_1}(x^\perp_{1j}; \theta_1) \left(1 - \frac{\partial}{\partial u_1} \mathfrak{C}(u_1, \lambda_2; \theta_{\mathfrak{C}}) \Big|_{u_1 = \lambda_1 \overline{F}_{X_1}(x^\perp_{1j}; \theta_1)} \right) \right]\\
&\times (\lambda_2)^{n_2^\perp} e^{-\lambda^\perp_{2\theta_{\mathfrak{C}}} T} \prod_{l=1}^{n_2^\perp} \left[f_{X_2}(x^\perp_{2l}; \theta_2) \left(1 - \frac{\partial}{\partial u_2} \mathfrak{C}(\lambda_1, u_2; \theta_{\mathfrak{C}}) \Big|_{u_2 = \lambda_2 \overline{F}_{X_2}(x^\perp_{2l}; \theta_2)} \right) \right]\\
&\times (\lambda_1 \lambda_2)^{n^\parallel} e^{-\lambda^\parallel_{\theta_{\mathfrak{C}}} T} \prod_{k=1}^{n^\parallel} \left[f_{X_1}(x^\parallel_{1k}; \theta_1) f_{X_2}(x^\parallel_{2k}; \theta_2) \frac{\partial^2}{\partial u_1 \partial u_2} \mathfrak{C}(u_1, u_2; \theta_{\mathfrak{C}}) \Big|_{u_1 = \lambda_1 \overline{F}_{X_1}(x^\parallel_{1k}; \theta_1), u_2 = \lambda_2 \overline{F}_{X_2}(x^\parallel_{2k}; \theta_2)} \right].
\end{aligned}$$

Proof. The likelihood function mainly draws upon the decomposition of $S = (S_1, S_2)^\top$ into the three independent compound Poisson processes $S_1^\perp$, $S_2^\perp$ and $S^\parallel$ made possible by Lemma 3.4. Let $\mathcal{L}_{S_i^\perp}(\lambda^\perp_{i\theta_{\mathfrak{C}}}, \theta_i, \theta_{\mathfrak{C}})$, $i \in \{1, 2\}$, denote the likelihood corresponding to the process $S_i^\perp$, and let $\mathcal{L}_{S^\parallel}(\lambda^\parallel_{\theta_{\mathfrak{C}}}, \theta_1, \theta_2, \theta_{\mathfrak{C}})$ denote the likelihood of $S^\parallel$. Then the joint likelihood is readily given by the product

$$\mathcal{L}(\lambda_1, \lambda_2, \theta_1, \theta_2, \theta_{\mathfrak{C}}) \;=\; \mathcal{L}_{S_1^\perp}(\lambda^\perp_{1\theta_{\mathfrak{C}}}, \theta_1, \theta_{\mathfrak{C}})\, \mathcal{L}_{S_2^\perp}(\lambda^\perp_{2\theta_{\mathfrak{C}}}, \theta_2, \theta_{\mathfrak{C}})\, \mathcal{L}_{S^\parallel}(\lambda^\parallel_{\theta_{\mathfrak{C}}}, \theta_1, \theta_2, \theta_{\mathfrak{C}}).$$

The likelihood function of a compound Poisson process observed over a fixed period $[0, T]$ is well studied and its general form can be found at the beginning of Chapter 6, Section 4 in [BR80], for example. Here we explain the derivation of $\mathcal{L}_{S_1^\perp}$ in detail and state the similar results for $\mathcal{L}_{S_2^\perp}$ as well as for $\mathcal{L}_{S^\parallel}$ later on.

By construction of the model, the losses $x^\perp_{11}, \ldots, x^\perp_{1n_1^\perp}$ attributed to the process $S_1^\perp$ took place according to a homogeneous Poisson process with intensity $\lambda^\perp_{1\theta_{\mathfrak{C}}}$. If $t^\perp_{1,1}, \ldots, t^\perp_{1,n_1^\perp}$ are the time points at which the losses occurred, the inter-arrival times $t^\perp_{1,j} - t^\perp_{1,j-1}$ with $j = 1, \ldots, n_1^\perp$ and $t^\perp_{1,0} = 0$ are i.i.d. according to the exponential distribution with density $f(t) = \lambda^\perp_{1\theta_{\mathfrak{C}}} e^{-\lambda^\perp_{1\theta_{\mathfrak{C}}} t}$. Therefore, the Poisson frequency part of $S_1^\perp$ contributes to the likelihood through

$$e^{-\lambda^\perp_{1\theta_{\mathfrak{C}}}\left(T - t^\perp_{1,n_1^\perp}\right)} \prod_{j=1}^{n_1^\perp} \left[\lambda^\perp_{1\theta_{\mathfrak{C}}} e^{-\lambda^\perp_{1\theta_{\mathfrak{C}}}\left(t^\perp_{1,j} - t^\perp_{1,j-1}\right)} \right] \;=\; \left(\lambda^\perp_{1\theta_{\mathfrak{C}}}\right)^{n_1^\perp} e^{-\lambda^\perp_{1\theta_{\mathfrak{C}}} T},$$

whereby the first factor on the left side of the equation is the probability that no more losses occurred after $x^\perp_{1n_1^\perp}$ until the end of the observation period $[0, T]$. With regard to the severity part of $S_1^\perp$, we make use of the survival function $\overline{F}_1^\perp(x_1)$ specified in Part (b) of Lemma 3.4 and obtain the lose size density as

$$f_1^\perp(x_1) \;=\; -\frac{\partial}{\partial x_1} \overline{F}_1^\perp(x_1) \qquad (3.28)$$

$$
\begin{aligned}
&= -\frac{\partial}{\partial x_1}\left(\left(\lambda^{\perp}_{1\theta_{\mathfrak{C}}}\right)^{-1}\left[\lambda_1\overline{F}_{X_1}(x_1;\theta_1) - \mathfrak{C}(\lambda_1\overline{F}_{X_1}(x_1;\theta_1),\lambda_2;\theta_{\mathfrak{C}})\right]\right)\\
&= \left(\lambda^{\perp}_{1\theta_{\mathfrak{C}}}\right)^{-1}\lambda_1 f_{X_1}(x_1;\theta_1)\left(1-\frac{\partial}{\partial u_1}\mathfrak{C}(u_1,\lambda_2;\theta_{\mathfrak{C}})\Big|_{u_1=\lambda_1\overline{F}_{X_1}(x_1;\theta_1)}\right).
\end{aligned}
$$

Putting the severity and frequency parts of $S_1^{\perp}$ together, we arrive at its likelihood function

$$
\mathcal{L}_{S_1^{\perp}}(\lambda^{\perp}_{1\theta_{\mathfrak{C}}},\theta_1,\theta_{\mathfrak{C}}) = \left(\lambda^{\perp}_{1\theta_{\mathfrak{C}}}\right)^{n_1^{\perp}} e^{-\lambda^{\perp}_{1\theta_{\mathfrak{C}}}T}\prod_{j=1}^{n_1^{\perp}} f_1^{\perp}(x^{\perp}_{1j}),
$$

which reflects the second line of (3.27) after rearranging the terms. By symmetry, the likelihood $\mathcal{L}_{S_2^{\perp}}$ associated with the process $S_2^{\perp}$ is exactly given by the third line of (3.27).

Only the likelihood $\mathcal{L}_{S^{\parallel}}$ of the bivariate compound Poisson process $S^{\parallel} = (S_1^{\parallel}, S_2^{\parallel})^{\top}$ remains to be determined. As its frequency component follows a homogeneous Poisson process with intensity $\lambda^{\parallel}_{\theta_{\mathfrak{C}}}$, the corresponding likelihood can be derived in the same manner as above. To obtain the density of the bivariate loss sizes, we utilise their joint survival function $\overline{F}^{\parallel}(x_1,x_2)$ from Part (b) of Lemma 3.4 and calculate the twofold partial derivative

$$
\begin{aligned}
f^{\parallel}(x_1,x_2) &= \frac{\partial^2}{\partial x_1\partial x_2}\overline{F}^{\parallel}(x_1,x_2) \qquad (3.29)\\
&= \frac{\partial^2}{\partial x_1\partial x_2}\left[(\lambda^{\parallel}_{\theta_{\mathfrak{C}}})^{-1}\mathfrak{C}\left(\lambda_1\overline{F}_{X_1}(x_1;\theta_1),\lambda_2\overline{F}_{X_2}(x_2;\theta_2);\theta_{\mathfrak{C}}\right)\right]\\
&= (\lambda^{\parallel}_{\theta_{\mathfrak{C}}})^{-1}\lambda_1\lambda_2 f_{X_1}(x_1;\theta_1)f_{X_2}(x_2;\theta_2)\frac{\partial^2}{\partial u_1\partial u_2}\mathfrak{C}\left(u_1,u_2;\theta_{\mathfrak{C}}\right)|_{u_1=\lambda_1\overline{F}_{X_1}(x_1;\theta_1),u_2=\lambda_2\overline{F}_{X_2}(x_2;\theta_2)}.
\end{aligned}
$$

As a result, the likelihood function of the bivariate process $S^{\parallel}$ is given by

$$
\mathcal{L}_{S^{\parallel}}(\lambda^{\parallel}_{\theta_{\mathfrak{C}}},\theta_1,\theta_2,\theta_{\mathfrak{C}}) = \left(\lambda^{\parallel}_{\theta_{\mathfrak{C}}}\right)^{n^{\parallel}} e^{-\lambda^{\parallel}_{\theta_{\mathfrak{C}}}T}\prod_{k=1}^{n^{\parallel}} f^{\parallel}(x^{\parallel}_{1k},x^{\parallel}_{2k}),
$$

After simple reformulation, the likelihood $\mathcal{L}_{S^{\parallel}}$ resembles the last line of equation (3.27) and this completes the proof. □

In view of the practical estimation of the parameters $(\lambda_1,\lambda_2,\theta_1,\theta_2,\theta_{\mathfrak{C}})^{\top}$ underlying a bivariate compound Poisson model $S(t) = (S_1(t),S_2(t))^{\top}$, there are two common approaches. The first one is the full maximum likelihood method, requiring the maximisation of $\mathcal{L}(\lambda_1,\lambda_2,\theta_1,\theta_2,\theta_{\mathfrak{C}})$ with respect to all its arguments simultaneously. Mostly this results in a highly non-trivial optimisation problem, given the typically large dimension of the parameter vector $(\lambda_1,\lambda_2,\theta_1,\theta_2,\theta_{\mathfrak{C}})^{\top}$. As operational losses are known for possessing high skewness, high kurtosis and a heavy tail, their characterisation often involves distribution families with at least two parameters. Hence the components θ_1 and θ_2 already account for four dimensions.

Nevertheless, the full MLE is appealing due to the typically expected convenient properties of its resulting estimates. Under mild regularity conditions on the random information

matrix, [Swe80] shows that maximum likelihood estimators are weakly consistent and asymptotically normal for the observation horizon T approaching infinity. In addition, the subsequent publication [Swe83] by the same author supplements that the estimators are efficient in the sense of having asymptotically maximum probability of concentration in convex symmetric sets around the true parameter values. For more details on the existence, uniqueness and asymptotic properties of maximum likelihood estimates for stochastic processes we refer to Chapter 8 in [KS97].

The second approach resembles the IFM method, which is originally developed for the estimation of multivariate models based upon ordinary copulas. Its main idea relies on the decomposition of the joint distribution into marginal laws and copula, in order to overcome the numerical problems associated with the full MLE. The IFM concept consists of the following two steps:

(1) The marginal compound Poisson processes $S_1(t)$ and $S_2(t)$ are estimated separately. For each $i \in \{1, 2\}$, an estimate for the marginal parameter vector $(\lambda_i, \theta_i)^\top$ is obtained by maximising the marginal likelihood function

$$\mathcal{L}_{S_i}(\lambda_i, \theta_i) \;=\; (\lambda_i)^{n_i^\perp + n^\parallel}\, e^{-\lambda_i T} \prod_{j=1}^{n_i^\perp} f_{X_i}(x_{ij}^\perp; \theta_i) \prod_{k=1}^{n^\parallel} f_{X_i}(x_{ik}^\parallel; \theta_i)$$

and by utilising all observed losses $x_{i1}^\perp, \ldots, x_{in_i^\perp}^\perp; x_{i1}^\parallel, \ldots, x_{in^\parallel}^\parallel$ occurred in the corresponding risk cell i. The resulting parameters shall be denoted by $(\hat{\lambda}_1, \hat{\lambda}_2, \hat{\theta}_1, \hat{\theta}_2)^\top$.

(2) The marginal estimators are considered as fixed and plugged into the likelihood function $\mathcal{L}(\hat{\lambda}_1, \hat{\lambda}_2, \hat{\theta}_1, \hat{\theta}_2, \theta_{\mathfrak{C}})$ given in Theorem 3.9. Then the likelihood is maximised as a function solely of the Lévy copula parameter $\theta_{\mathfrak{C}}$ and this ultimately yields the estimate $\hat{\theta}_{\mathfrak{C}}$.

Clearly, the resulting estimates of the full MLE and the IFM method most likely differ from each other. The striking advantage of the latter is the reduction of numerical complexity. As the joint likelihood is maximised only over the copula parameter in the second step, a Lévy copula family with more than one parameter could be imposed to capture complexer dependence characteristics such as asymmetry. Certainly, it is also possible to combine the two concepts, for example by using the estimators obtained through IFM as starting values of a full maximum likelihood procedure. Another approach of taking the advantage of both methods is to first choose a copula family from various potentially suitable Lévy copulas via IFM, then apply the full MLE to finalise the parameter estimation. Last but not least, the likelihood function as detailed in Theorem 3.9 can be utilised within a Bayesian estimation scheme as well, in which expert knowledge is incorporated in form of prior distributions.

3.3.2 MLE under a discrete observation scheme

In practice, it is not unusual that operational losses are only reported on a weekly or even a monthly basis, and we shall call this the reporting period in the sequel. Against

this background, within one reporting period the dependent losses $(X_{1k}^{\parallel}, X_{2k}^{\parallel})^\top$ cannot be distinguished from the independent losses $X_{1j}^{\perp}$ and $X_{2l}^{\perp}$, respectively, although all loss sizes are documented exactly. As a result, the likelihood function given in Theorem 3.9 cannot be applied and a new estimation method is needed.

Following the proposal by [Vel12], suppose the entire observation horizon $[0, T]$ consists of n reporting periods with equal length. In other words, the bivariate compound Poisson model $S(t) = (S_1(t), S_2(t))^\top$ shall be observed over an equidistant grid

$$0 = t_0 < t_1 < \cdots < t_n = T \quad \text{with} \quad t_s - t_{s-1} = \Delta t = \frac{T}{n},\ s \in \{1, \ldots, n\}.$$

Moreover, losses from disjoint reporting periods are assumed to be too far apart to be caused by common reasons. Similar to Theorem 3.9, if the Lévy copula $\mathfrak{C}$ is continuously differentiable of second order on $(0, \lambda_1) \times (0, \lambda_2)$, then a closed-form likelihood function for $S(t)$ can be derived based upon the maximum loss sizes within each reporting period. Hence assume the following loss data are available:

- the number n_{is} of losses occurred in risk cell i, $i \in \{1, 2\}$, and within the reporting period $(t_s, t_{s-1}]$, $s \in \{1, \ldots, n\}$,
- the maximum loss amount x_{is} in risk cell i, $i \in \{1, 2\}$, and within the reporting period $(t_s, t_{s-1}]$, $s \in \{1, \ldots, n\}$.

Let the corresponding random variables be denoted by N_{is} and M_{is} with $i \in \{1, 2\}$ and $s \in \{1, \ldots, n\}$, respectively. Because the compound Poisson process $S(t)$ has independent and stationary increments, whose distribution only depends on the length of the interval, the random vectors $(N_{1s}, N_{2s}, M_{1s}, M_{2s})^\top$ are i.i.d. for $s \in \{1, \ldots, n\}$. Therefore, the overall likelihood can be factorised into

$$\mathcal{L}^D(\lambda_1, \lambda_2, \theta_1, \theta_2, \theta_{\mathfrak{C}}) \;=\; \prod_{s=1}^{n} \mathcal{L}^D_{n_{1s}, n_{2s}, x_{1s}, x_{2s}}(\lambda_1, \lambda_2, \theta_1, \theta_2, \theta_{\mathfrak{C}}),$$

whereby $\mathcal{L}^D_{n_{1s}, n_{2s}, x_{1s}, x_{2s}}$ is the likelihood attributed to the reporting period $(t_s, t_{s-1}]$ and evaluated at the corresponding observation. Of course, the likelihood of each reporting period has the same form and we write $\mathcal{L}^D_{n_1, n_2, x_1, x_2}$ for simpler notation. In order to make $\mathcal{L}^D_{n_1, n_2, x_1, x_2}$ tangible, we differentiate between four cases of the possible value $(n_1, n_2)^\top$ which the frequency vector $(N_{1s}, N_{2s})^\top$ can take:

$$\begin{aligned}\mathcal{L}^D_{n_1, n_2, x_1, x_2} \;=\; & H_{0,0}(\infty, \infty)\mathbb{1}_{\{n_1=0, n_2=0\}} + \frac{\partial}{\partial x_1} H_{n_1,0}(x_1, \infty)\mathbb{1}_{\{n_1>0, n_2=0\}} \\ & + \frac{\partial}{\partial x_2} H_{0,n_2}(\infty, x_2)\mathbb{1}_{\{n_1=0, n_2>0\}} + \frac{\partial^2}{\partial x_1 \partial x_2} H_{n_1,n_2}(x_1, x_2)\mathbb{1}_{\{n_1>0, n_2>0\}},\end{aligned}$$

where the function $H_{n_1,n_2}(x_1, x_2)$ denotes the joint distribution of the maximum loss amounts $(M_{1s}, M_{2s})^\top$ with given frequency values:

$$H_{n_1,n_2}(x_1, x_2) \;=\; \mathbb{P}(N_{1s} = n_1, N_{2s} = n_2, M_{1s} \le x_1, M_{2s} \le x_2).$$

In accordance with Lemma 3.4, the frequency random variable N_{is} can be decomposed into

$$N_{is} \;=\; N_{is}^{\perp} + N_s^{\parallel}, \quad i \in \{1,2\},\, s \in \{1,\dots,n\},$$

with $N_{is}^{\perp} = N_i^{\perp}(t_s) - N_i^{\perp}(t_{s-1})$ and $N_s^{\parallel} = N^{\parallel}(t_s) - N^{\perp}(t_{s-1})$. Note that in contrast to N_{is}, the values of $N_{is}^{\perp}$ and $N_s^{\parallel}$ cannot be observed under the current time discretisation assumption. Nevertheless, the consideration of the latter is essential in order to compute $H_{n_1,n_2}(x_1,x_2)$. Again let $\lambda^{\parallel} := \lambda_{\theta_{\mathfrak{C}}}^{\parallel}$ denote the frequency parameter of the dependent loss process $S^{\parallel}$ which depends on the copula parameter $\theta_{\mathfrak{C}}$. Moreover, the frequency parameter of the independent loss process $S_i^{\perp}$ is computed as $\lambda_i^{\perp} = \lambda_i - \lambda^{\parallel}$ for $i \in \{1,2\}$. Then the random variables $N_{1s}^{\perp}$, $N_{2s}^{\perp}$ and $N_s^{\parallel}$ are independently Poisson distributed with parameters $\lambda_1^{\perp}\Delta t$, $\lambda_2^{\perp}\Delta t$ and $\lambda^{\parallel}\Delta t$, respectively.

On the severity side, the maximum loss amounts can be expressed through

$$M_{is} \;=\; \max\{M_{is}^{\perp}, M_{is}^{\parallel}\}, \quad i \in \{1,2\},\, s \in \{1,\dots,n\},$$

where $M_{is}^{\perp}$ denotes the maximum of losses occurred in the interval $(t_s, t_{s-1}]$ and attributed to the independent loss process $S_i^{\perp}$, and $M_{is}^{\parallel}$ is the maximum of losses occurred in the same time interval but attributed to the marginal dependent loss process $S_i^{\parallel}$. Similar to the discretised frequency process, the maxima $M_{is}^{\perp}$ and $M_{is}^{\parallel}$ cannot be observed directly, but their consideration is necessary for deriving the likelihood. Furthermore, recall the general fact that the maximum M_m of i.i.d. random variables $X_1,\dots,X_m$ with distribution function F satisfies $\mathbb{P}(M_m \le x) = F^m(x)$.

In view of the independence between the three partial processes $S_1^{\perp}$, $S_2^{\perp}$ and $S^{\parallel}$, as well as the independence between frequency and severity, it follows for the simplest case of $n_1 = n_2 = 0$ that

$$H_{0,0}(\infty,\infty) \;=\; \mathbb{P}(N_{1s}=0, N_{2s}=0) \;=\; \mathbb{P}(N_{1s}^{\perp}=0, N_{2s}^{\perp}=0, N_s^{\parallel}=0) \;=\; e^{-(\lambda_1^{\perp}+\lambda_2^{\perp}+\lambda^{\parallel})\Delta t}.$$

If $n_1 > 0$ and $n_2 = 0$, then the joint distribution takes the form

$$\begin{aligned}
H_{n_1,0}(x_1,\infty) \;&=\; \mathbb{P}(N_{1s}=n_1, N_{2s}=0, M_{1s}\le x_1)\\
&=\; \mathbb{P}(N_{1s}^{\perp}=n_1, N_{2s}^{\perp}=0, N_s^{\parallel}=0, M_{1s}^{\perp}\le x_1)\\
&=\; \mathbb{P}(N_{1s}^{\perp}=n_1, N_{2s}^{\perp}=0, N_s^{\parallel}=0)\,\mathbb{P}(M_{1s}^{\perp}\le x_1 | N_{1s}^{\perp}=n_1)\\
&=\; \frac{(\lambda_1^{\perp}\Delta t)^{n_1}}{n_1!} e^{-(\lambda_1^{\perp}+\lambda_2^{\perp}+\lambda^{\parallel})\Delta t}\left[F_1^{\perp}(x_1)\right]^{n_1}.
\end{aligned}$$

Thus its partial derivative with respect to x_1 is given by

$$\frac{\partial}{\partial x_1} H_{n_1,0}(x_1,\infty) \;=\; \frac{(\lambda_1^{\perp}\Delta t)^{n_1}}{n_1!} e^{-(\lambda_1^{\perp}+\lambda_2^{\perp}+\lambda^{\parallel})\Delta t} n_1 \left[F_1^{\perp}(x_1)\right]^{n_1-1} f_1^{\perp}(x_1),$$

whereas the density $f_1^{\perp}(x_1)$ is already calculated in (3.28) during the derivation of the likelihood under a continuous observation scheme. Due to symmetry, the case of $n_1 = 0$ and $n_2 > 0$ follows by an analogous calculation, resulting in

$$\frac{\partial}{\partial x_2} H_{0,n_2}(\infty,x_2) \;=\; \frac{(\lambda_2^{\perp}\Delta t)^{n_2}}{n_2!} e^{-(\lambda_1^{\perp}+\lambda_2^{\perp}+\lambda^{\parallel})\Delta t} n_2 \left[F_2^{\perp}(x_2)\right]^{n_2-1} f_2^{\perp}(x_2).$$

For the last case of $n_1 > 0$ and $n_2 > 0$, the joint distribution is split into a sum by exploiting all the possible values $n^{\|}$ the dependent loss frequency $N_s^{\|}$ can take:

$$H_{n_1,n_2}(x_1,x_2) \;=\; \sum_{n^{\|}=0}^{\min\{n_1,n_2\}} \mathbb{P}(N_{1s}^{\perp} = n_1 - n^{\|}, N_{2s}^{\perp} = n_2 - n^{\|}, N_s^{\|} = n^{\|}, M_{1s} \le x_1, M_{2s} \le x_2),$$

where each summand is further computed through

$$\begin{aligned}
&\mathbb{P}(N_{1s}^{\perp} = n_1 - n^{\|}, N_{2s}^{\perp} = n_2 - n^{\|}, N_s^{\|} = n^{\|}, M_{1s}^{\perp} \le x_1, M_{2s}^{\perp} \le x_2, M_{1s}^{\|} \le x_1, M_{2s}^{\|} \le x_2) \\
&= \mathbb{P}(N_{1s}^{\perp} = n_1 - n^{\|}, N_{2s}^{\perp} = n_2 - n^{\|}, N_s^{\|} = n^{\|}) \\
&\quad \times \mathbb{P}(M_{1s}^{\perp} \le x_1 | N_{1s}^{\perp} = n_1 - n^{\|})\, \mathbb{P}(M_{2s}^{\perp} \le x_2 | N_{2s}^{\perp} = n_2 - n^{\|})\, \mathbb{P}(M_{1s}^{\|} \le x_1, M_{2s}^{\|} \le x_2 | N_s^{\|} = n^{\|}) \\
&= \frac{(\lambda_1^{\perp}\Delta t)^{n_1 - n^{\|}}}{(n_1 - n^{\|})!} e^{-\lambda_1^{\perp}\Delta t} \frac{(\lambda_2^{\perp}\Delta t)^{n_2 - n^{\|}}}{(n_2 - n^{\|})!} e^{-\lambda_2^{\perp}\Delta t} \frac{(\lambda^{\|}\Delta t)^{n^{\|}}}{n^{\|}!} e^{-\lambda^{\|}\Delta t} \\
&\quad \times \left[F_1^{\perp}(x_1)\right]^{n_1 - n^{\|}} \left[F_2^{\perp}(x_2)\right]^{n_2 - n^{\|}} \left[F^{\|}(x_1,x_2)\right]^{n^{\|}}.
\end{aligned}$$

Then the twofold partial derivative of $H_{n_1,n_2}(x_1,x_2)$ can be calculated by the product rule and in a similar manner to (3.28). The details are omitted here as they are rather lengthy and do not provide deeper insights into the subject.

In conclusion, all components of the likelihood function $\mathcal{L}^D$ under the discrete observation scheme are specified. According to [Vel12], the function $\mathcal{L}^D$ converges to the likelihood under the continuous observation scheme given in Theorem 3.9 as the number n of the reporting periods approaches infinity. Furthermore, just as in the already discussed continuous case, the practical estimation of the parameters $(\lambda_1, \lambda_2, \theta_1, \theta_2, \theta_{\mathfrak{C}})^\top$ can be achieved via either the full maximum likelihood method or the IFM approach. The latter is particularly appealing under the discrete observation scheme, as in the first step of the IFM approach all single loss sizes can be used to determine the parameters $(\lambda_1, \theta_1)^\top$ and $(\lambda_2, \theta_2)^\top$ underlying the marginal processes. In comparison, the joint maximisation of $\mathcal{L}^D(\lambda_1, \lambda_2, \theta_1, \theta_2, \theta_{\mathfrak{C}})$ with respect to all entries only makes use of the number of losses and the maximum loss sizes within each reporting period.

3.3.3 Implication of rescaled observation time unit

So far we have assumed the expected number of the marginal losses in one time unit is given by λ_i, $i \in \{1,2\}$, without explicitly defining the length of one time unit. As the risk measure currently specified by authority is the VaR for a one-year holding period, it is natural to consider one year as one time unit. However, for different reasons one may be interested in a dependence model based on alternative time units. Suppose the new time unit is the old one rescaled with a positive constant c^{-1}, for instance $c = 4$, in case we switch from yearly to quarterly modelling. If the Lévy copula under the old time unit is given by $\mathfrak{C}(u_1, u_2; \theta_{\mathfrak{C}})$, then [BL07] ensures that the Lévy copula $\tilde{\mathfrak{C}}$ under the new time unit can be derived from the old one through

$$\tilde{\mathfrak{C}}(u_1, u_2; \theta_{\mathfrak{C}}) \;=\; c^{-1}\mathfrak{C}(cu_1, cu_2; \theta_{\mathfrak{C}}), \quad (u_1, u_2)^\top \in [0,\infty]^2. \tag{3.30}$$

Note the above equation not only holds for bivariate Poisson models, but also for general Lévy processes in higher dimension. Hereinafter, we establish the relation between $\mathfrak{C}$ and $\tilde{\mathfrak{C}}$ by verifying that the corresponding likelihood functions $\mathcal{L}$ and $\tilde{\mathcal{L}}$ according to Theorem 3.9 indeed attain maximum at the same parameter values.

As the parameters θ_1 and θ_2 underlying the severity distributions remain unchanged under the new time unit, they are omitted below for simpler notation. The same principle applies to the copula parameter $\theta_{\mathfrak{C}}$. On the other hand, the rescaled marginal intensities are given by $\tilde{\lambda}_i = c^{-1}\lambda_i$, $i \in \{1,2\}$, owing to properties of homogeneous Poisson processes. So the new intensity of the dependent losses can be calculated through

$$\tilde{\lambda}^{\|} \;=\; \tilde{\mathfrak{C}}(\tilde{\lambda}_1, \tilde{\lambda}_2) = c^{-1}\mathfrak{C}(\lambda_1, \lambda_2) = c^{-1}\lambda^{\|}$$

and it is a rescaling of $\lambda^{\|}$ as expected. The intensity of the independent loss process $S_i^{\perp}$ follows as $\tilde{\lambda}_i^{\perp} = c^{-1}\lambda_i^{\perp}$ for $i \in \{1,2\}$. Furthermore, the loss data used for estimation, as they were described at the beginning of Section 3.3.1, stay the same, besides the corresponding observation interval is adjusted to $[0, \tilde{T}] = [0, cT]$.

As usual, it is more convenient to maximise the log-likelihood function instead of the likelihood function itself. Hence we apply the logarithm to the likelihood function under the rescaled time unit and obtain

$$\begin{aligned}
&\ln \tilde{\mathcal{L}}(\tilde{\lambda}_1, \tilde{\lambda}_2) \\
&\quad = n_1^{\perp} \ln \tilde{\lambda}_1 - \tilde{\lambda}_1^{\perp}\tilde{T} + \sum_{j=1}^{n_1^{\perp}} \ln f_{X_1}(x_{1j}^{\perp}) + \sum_{j=1}^{n_1^{\perp}} \ln \left(1 - \frac{\partial}{\partial u_1}\tilde{\mathfrak{C}}(u_1, \tilde{\lambda}_2)\Big|_{u_1 = \tilde{\lambda}_1 \overline{F}_{X_1}(x_{1j}^{\perp})}\right) \\
&\quad + n_2^{\perp} \ln \tilde{\lambda}_2 - \tilde{\lambda}_2^{\perp}\tilde{T} + \sum_{l=1}^{n_2^{\perp}} \ln f_{X_2}(x_{2l}^{\perp}) + \sum_{l=1}^{n_2^{\perp}} \ln \left(1 - \frac{\partial}{\partial u_2}\tilde{\mathfrak{C}}(\tilde{\lambda}_1, u_2)\Big|_{u_2 = \tilde{\lambda}_2 \overline{F}_{X_2}(x_{2l}^{\perp})}\right) \\
&\quad + n^{\|} \ln(\tilde{\lambda}_1\tilde{\lambda}_2) - \tilde{\lambda}^{\|}\tilde{T} + \sum_{k=1}^{n^{\|}} \left[\ln f_{X_1}(x_{1k}^{\|}) + \ln f_{X_2}(x_{2k}^{\|})\right] \\
&\quad + \sum_{k=1}^{n^{\|}} \ln \frac{\partial^2}{\partial u_1 \partial u_2}\tilde{\mathfrak{C}}(u_1, u_2)\Big|_{u_1 = \tilde{\lambda}_1 \overline{F}_{X_1}(x_{1k}^{\|}), u_2 = \tilde{\lambda}_2 \overline{F}_{X_2}(x_{2k}^{\|})} .
\end{aligned}$$

The partial derivatives of $\tilde{\mathfrak{C}}$ can be calculated as

$$\frac{\partial}{\partial u_i}\tilde{\mathfrak{C}}(u_1, u_2) \;=\; \frac{\partial}{\partial v_i}\mathfrak{C}(v_1, v_2)\Big|_{v_1 = cu_1, v_2 = cu_2}, \quad i \in \{1,2\},$$

and similarly, the density has the representation

$$\frac{\partial^2}{\partial u_1 \partial u_2}\tilde{\mathfrak{C}}(u_1, u_2) \;=\; c\,\frac{\partial^2}{\partial v_1 \partial v_2}\mathfrak{C}(v_1, v_2)\Big|_{v_1 = cu_1, v_2 = cu_2} .$$

Putting everything together, the log-likelihood under the new time unit satisfies

$$\ln \tilde{\mathcal{L}}(\tilde{\lambda}_1, \tilde{\lambda}_2)$$

$$
\begin{aligned}
&= n_1^{\perp}\ln(c^{-1}\lambda_1) - \lambda_1^{\perp}T + \sum_{j=1}^{n_1^{\perp}} \ln f_{X_1}(x_{1j}^{\perp}) + \sum_{j=1}^{n_1^{\perp}} \ln\left(1 - \frac{\partial}{\partial v_1}\mathfrak{C}(v_1,\lambda_2)\Big|_{v_1=\lambda_1\overline{F}_{X_1}(x_{1j}^{\perp})}\right) \\
&+ n_2^{\perp}\ln(c^{-1}\lambda_2) - \lambda_2^{\perp}T + \sum_{l=1}^{n_2^{\perp}} \ln f_{X_2}(x_{2l}^{\perp}) + \sum_{l=1}^{n_2^{\perp}} \ln\left(1 - \frac{\partial}{\partial v_2}\mathfrak{C}(\lambda_1,v_2)\Big|_{v_2=\lambda_2\overline{F}_{X_2}(x_{2l}^{\perp})}\right) \\
&+ n^{\parallel}\ln(c^{-2}\lambda_1\lambda_2) - \lambda^{\parallel}T + \sum_{k=1}^{n^{\parallel}} \left[\ln f_{X_1}(x_{1k}^{\parallel}) + \ln f_{X_2}(x_{2k}^{\parallel})\right] \\
&+ \sum_{k=1}^{n^{\parallel}} \ln\left(c\,\frac{\partial^2}{\partial v_1 \partial v_2}\mathfrak{C}(v_1,v_2)\Big|_{v_1=\lambda_1\overline{F}_{X_1}(x_{1k}^{\parallel}),v_2=\lambda_2\overline{F}_{X_2}(x_{2k}^{\parallel})}\right) \\
&= -(n_1^{\perp}+n_2^{\perp}+n^{\parallel})\ln c + \ln\mathcal{L}(\lambda_1,\lambda_2)
\end{aligned}
$$

In conclusion, the log-likelihood function $\ln\tilde{\mathcal{L}}$ only differs from the original one $\ln\mathcal{L}$ by a constant and their maximisation would deliver equivalent results. Hence the parameters under the rescaled time unit can be retrieved by properly rescaling the maximum likelihood estimates under the original time unit.

Furthermore, this is an appropriate occasion to introduce the class of homogeneous Lévy copulas. A Lévy copula is called homogeneous of order one, if it satisfies the property

$$\mathfrak{C}(u_1,u_2) = c^{-1}\mathfrak{C}(cu_1,cu_2)$$

for all $(u_1,u_2)^{\top} \in [0,\infty]^2$ and for any constant $c > 0$. By comparing the above equation to formula (3.30) for the Lévy copula after modifying the time unit by a constant, we immediately conclude that homogeneous Lévy copulas are invariant under time rescaling. Prominent examples of this special copula class are the complete dependence, the independence and the Clayton Lévy copulas.

Chapter 4

Estimation of operational risk measures

As already explained in Chapter 1, a key objective of modelling operational risk is to assess the required capital reserves in a financial institution against potential future losses. Under the current industry standards, the core principle of capital charge estimation is the VaR for a one-year ahead time horizon and measured based upon the distribution G_+ of the overall loss process $S_+ = \sum_{i=1}^{d} S_i$. A precise mathematical characterisation of operational VaR was introduced in Definition 1.1 under the general loss distribution approach, which of course equally applies to our dependence model built upon a d-dimensional compound Poisson process as detailed in Definition 3.1.

Besides VaR, the most popular alternative risk measure is given by expected shortfall (ES). In contrast to VaR, ES constitutes a coherent risk measure and in particular satisfies the subadditive property. The latter reflects the natural intuition of diversification benefit, that is, the risk exposure calculated based on the aggregate loss distribution across independent risk cells should not be larger than the sum of risk exposures calculated for each cell alone. Moreover, ES does not only state the threshold but also the expected size of potential severe losses, provided that the threshold is exceeded. Hence it is more conservative than the VaR at the same confidence level. However, the risk measure ES is only well-defined if the underlying distribution possesses finite expectation, which is not always the case regarding the heavy-tailed property of operational risk losses. In an analogous manner to Definition 1.1 for operational VaR, we make the concept of ES precise in the current context of operational risk.

Definition 4.1 (**Operational ES**).
Assume the aggregate loss $S_i(t)$ of risk cell $i \in \{1, \ldots, d\}$ has finite expectation for $t \geq 0$. Then the stand-lone operational ES of risk cell i until time $t \geq 0$ at confidence level $\alpha \in (0, 1)$ is defined as

$$\mathrm{ES}_{i,t}(\alpha) \;=\; \frac{1}{1-\alpha} \int_{\alpha}^{1} \mathrm{VaR}_{i,t}(\tilde{\alpha}) \mathrm{d}\tilde{\alpha}.$$

Accordingly, the total operational ES of a financial institution until time $t \geq 0$ at level

$\alpha \in (0,1)$ is specified through

$$\mathrm{ES}_{+,t}(\alpha) \;=\; \frac{1}{1-\alpha}\int_{\alpha}^{1} \mathrm{VaR}_{+,t}(\tilde{\alpha})\mathrm{d}\tilde{\alpha}.$$

Furthermore, ES belongs as a special case to the so-called spectral risk measures (SRMs), which build a more general coherent class of risk measures in quantitative finance and is first considered in [TW12] to quantify operational risk. Given a non-negative and non-decreasing weight function $\phi : [0,1] \to \mathbb{R}$ satisfying the normalisation $\int_0^1 \phi(s)\mathrm{d}s = 1$, a SRM can be specified as

$$\mathrm{SRM}^{\phi} \;=\; \int_0^1 \phi(s)\mathrm{VaR}(s)\,\mathrm{d}s.$$

The function ϕ is referred to as an admissible risk spectrum and its non-decreasing property ensures that the weight attached to a higher quantile being no less than it attached to a lower one. In addition, the more risk averse the user of a SRM is, the more steeply the weight ϕ rises. Hence on the contrary to VaR, SRMs allow for individual risk attitudes in operational risk exposure estimations. In order to formulate asymptotic results with respect to the confidence level α, we restrict our attention below to a subclass of admissible risk spectra, which assign non-decreasing weights to the largest $(1-\alpha)\%$ losses, and zero weight to the remaining smaller quantiles. This consideration is also reasonable in the sense that both banks and regulators are mainly concerned with the severest loss sizes. For any admissible risk spectrum ϕ as introduced above, the transformation

$$\phi^*(s) \;=\; \frac{1}{1-\alpha}\phi\left(1-\frac{1-s}{1-\alpha}\right)\mathbb{1}_{[\alpha,1]}(s). \tag{4.1}$$

constitutes a family of rescaled admissible risk spectra. Obviously, ES can be characterised as a SRM with $\phi(s) = \frac{1}{1-\alpha}\mathbb{1}_{[\alpha,1)}(s)$. Now we can define the operational SRMs for a financial institution comprising d risk cells.

Definition 4.2 (**Operational SRM**).
Let $\phi^* : [\alpha,1] \to \mathbb{R}$ denote an admissible risk spectrum as detailed in (4.1) and assume the aggregate loss $S_i(t)$ of risk cell $i \in \{1,\dots,d\}$ has finite expectation for $t \geq 0$. Then the stand-alone operational SRM associated with ϕ^* at confidence level $\alpha \in (0,1)$ and over period $[0,t]$ is given by

$$\mathrm{SRM}_{i,t}^{\phi^*}(\alpha) \;=\; \int_{\alpha}^{1} \phi^*(\tilde{\alpha})\mathrm{VaR}_{i,t}(\tilde{\alpha})\,\mathrm{d}\tilde{\alpha}$$

for risk cell i. Moreover, the total operational SRM of a financial institution until time $t \geq 0$ at level $\alpha \in (0,1)$ has the representation

$$\mathrm{SRM}_{+,t}^{\phi^*}(\alpha) \;=\; \int_{\alpha}^{1} \phi^*(\tilde{\alpha})\mathrm{VaR}_{+,t}(\tilde{\alpha})\,\mathrm{d}\tilde{\alpha}.$$

For most choices of severity and frequency distributions, the distribution G_+ of the overall loss process S_+ does not possess an analytically evaluable formula. As a result, banks often

resort to simulation methods in order to estimate operational risk measures. Nonetheless, as both regulatory and economic capital calculations are based on very high quantiles, typically at least of significance level $\alpha = 99.9\%$, a natural estimation approach is via asymptotic tail approximations. More precisely, we would like to represent the risk measures at high confidence levels through a closed-form expression in terms of the single loss distributions F_{X_i}, $i \in \{1, \ldots, d\}$. The result for VaR in the univariate case is already commonly applied in practice and well-known under the name single-loss approximation (SLA). Going one step further, the employment of Lévy copulas proves itself to be particularly convenient in generalising SLA to the multidimensional case, that is, the joint estimation of VaR_+, ES_+ or $\mathrm{SRM}_+^{\phi^*}$ for d risk cells.

In Section 4.1.1 and 4.1.2, the most relevant approximation results for the uni- and multivariate cases are briefly reviewed, respectively. We shall spare the proofs and refer the interested readers to [BK10] for detailed information on operational VaR, to [BU09] for ES, and to [TW12] for SRM. Then in Section 4.2, we make use of the compound Poisson property of the overall loss process S_+ and derive an analytical expression for the associated severity distribution F_+ in the two-dimensional setting. The implications out of this for the overall risk measure estimation is explored as well. Eventually, Section 4.3 discusses potential improvements and extensions of the previous results. In order to state the asymptotic relation $\lim\limits_{x\to\infty} \frac{f(x)}{g(x)} = 1$ between two functions f and g via a simpler notation, we introduce in this chapter the expression $f(x) \sim g(x)$ as $x \to \infty$.

4.1 Analytical approximation of operational risk measures

4.1.1 The one-dimensional case

First we recall risk measure approximations in the univariate case, that is, for a single risk cell $i \in \{1, \ldots, d\}$, whose structure is characterised by a compound Poisson process $S_i(t) \sim \mathrm{CPP}(\lambda_i, F_{X_i})$ as detailed in Part (a) of Definition 3.1. As we focus on one risk cell in the current section, the subscript i is omitted for ease of notation.

For the sake of completeness, we point out that in the one-dimensional setting, the subsequent asymptotic results do not only hold for compound Poisson processes, but also for more general compound distributions based on alternative frequency components. More precisely, the frequency process $N(t)$, $t \geq 0$, can be relaxed to a counting process with values in $\mathbb{N}_0$, which constitutes a càdlàg process with piecewise constant trajectories and sample paths moving by jump size of plus one.

Due to the independence between frequency and severity, the aggregate loss distribution function can be written as

$$G_t(x) = \mathbb{P}(S(t) \leq x)$$

$$
\begin{aligned}
&= \sum_{n=0}^{\infty} \mathbb{P}(N(t)=n)\,\mathbb{P}(S(t)\le x|N(t)=n) \\
&= \sum_{n=0}^{\infty} \mathbb{P}(N(t)=n)\,F_X^{n*}(x), \qquad\qquad t\ge 0,\, x\ge 0,
\end{aligned}
$$

where F_X^{n*} denotes the n-fold convolution of the severity distribution F_X with the special case $F_X^{0*}(x) = \mathbb{1}_{[0,\infty)}$. Under weak regularity conditions, the far out right tail of the compound distribution G_t, which is crucial for the determination of operational risk measures at high confidence levels, is related to the severity distribution F_X via the following theorem from [EKM97].

Theorem 4.3 (**Aggregate loss distribution in the subexponential case**).
If the severity distribution F_X is subexponential and for fixed $t > 0$ the frequency process $N(t)$ satisfies the condition

$$\sum_{n=0}^{\infty}(1+\epsilon)^n\,\mathbb{P}(N(t)=n) \;<\; \infty \tag{4.2}$$

for some $\epsilon > 0$, then the aggregate loss distribution G_t is subexponential with asymptotic tail behaviour

$$\overline{G}_t(x) \;\sim\; \mathbb{E}[N(t)]\overline{F}_X(x), \quad x\to\infty. \tag{4.3}$$

It is also shown in [EKM97] that condition (4.2) is fulfilled by the Poisson and the negative binomial frequency processes, which constitute the two most popular frequency modelling choices among financial institutions.

In contrast to the relaxation with respect to frequency, for the subsequent asymptotic statements to apply we strengthen the loss severity to the class of subexponential distributions denoted by $\mathcal{S}$. For a formal definition of $\mathcal{S}$ and the related classes of regularly varying distributions $\mathcal{R}$ as well as rapidly varying distributions $\mathcal{R}_\infty$, we refer to Appendix B. Nonetheless, owing to the heavy-tailed nature of operational losses, the assumption of subexponential severities does not present a substantial restriction in practice. The attribute subexponential refers to the fact that the tail of a distribution in $\mathcal{S}$ decays more slowly than any exponential tail. Important examples include the Weibull distribution with shape parameter less than one, the lognormal distribution and the Pareto distribution, which are already widely applied in the operational risk context as indicated in Section 1.2.

Given relation (4.3), it is straightforward to derive asymptotic estimations for operational VaR, ES and SRM valid at a high confidence level α near one. Although we have acknowledged ES being a special case of SRM, we separately state the approximations for ES in the subsequent theorems due to its prominent popularity after VaR in practical implementations.

Theorem 4.4 (**Operational risk measures for a single risk cell**).
Consider a one-dimensional loss process as detailed in Theorem 3.1, Part (a), and let

the conditions of Theorem 4.3 be satisfied with $\overline{F}_X \in \mathcal{S} \cap (\mathcal{R} \cup \mathcal{R}_\infty)$*. Then the following approximations hold:*

(a) The stand-alone VaR admits the asymptotic approximation

$$VaR_t(\alpha) \sim F_X^{\leftarrow}\left(1 - \frac{1-\alpha}{\mathbb{E}[N(t)]}\right), \quad \alpha \to 1. \tag{4.4}$$

(b) Assume further the severity distribution has a regularly varying tail $\overline{F}_X \in \mathcal{R}_{-\gamma}$ *with* $\gamma > 1$*, then an approximation for the operational ES can be specified as*

$$ES_t(\alpha) \sim \frac{\gamma}{\gamma - 1} VaR_t(\alpha), \quad \alpha \to 1.$$

(c) Let ϕ^* *be an admissible risk spectrum as introduced in* (4.1) *and let its unscaled equivalence satisfy for all* $s > 1$ *the condition* $\phi(1 - s^{-1}) \leq \eta s^{-\gamma^{-1}+1-\epsilon}$ *with some* $\eta > 0$ *and* $\epsilon > 0$*. Then under the same assumption for the loss severity* F_X *as in Part (b), the asymptotic SRM with respect to* ϕ^* *has the form*

$$SRM_t^{\phi^*}(\alpha) \sim k(\gamma, \phi)\, VaR_t(\alpha), \quad \alpha \to 1,$$

with the constant $k(\gamma, \phi) = \int_1^\infty s^{\gamma^{-1}-2}\, \phi(1 - s^{-1})\, \mathrm{d}s$*.*

Consequently, operational VaR, ES and SRM at high confidence levels are mainly determined by the tail of the severity distribution and the frequency expectation. As capital reserve quantification based on risk measures is of primary relevance for financial institutions, tail severity modelling should be paid with the highest attention.

Besides ES, a commonly encountered SRM with a risk spectrum satisfying the condition in Part (c) of the above theorem can be derived from the constant absolute risk aversion utility function with the Arrow-Pratt coefficient A. The corresponding weighting function is given by $\phi(s) = \frac{Ae^{-A(1-s)}}{1-e^{-A}}$, which is capable of reflecting the risk aversion of an individual financial institution.

4.1.2 The multidimensional case

The current section is attributed to risk measure estimations for a bank consisting of d risk cells, whose dependence structure is modelled by a d-variate compound Poisson process as detailed in Definition 3.1. Clearly, the individual VaR_i, ES_i and SRM_i of each risk cell $i \in \{1, \ldots, d\}$ can be separately approximated by the formulas from Theorem 4.4, if the marginal loss severities belong to the class $\mathcal{S} \cap (\mathcal{R} \cup \mathcal{R}_\infty)$. However, the Lévy copula characterising the interdependence among the risk cells plays an essential role in the estimation of the overall risk measures VaR_+, ES_+ and SRM_+ with respect to the loss process S_+. As already explained after the model specification in Section 3.1, the overall loss process S_+ itself constitutes a one-dimensional compound Poisson process with

tail integral $\overline{\Pi}_+$ given by equation (3.2). Accordingly, the Poisson frequency parameter underlying S_+ is computed through

$$\lambda_+ \;=\; \lim_{x\downarrow 0} \overline{\Pi}_+(x),$$

and the corresponding severity distribution follows as

$$F_+(x) \;=\; 1 - \overline{F}_+(x) \;=\; 1 - \lambda_+^{-1}\overline{\Pi}_+(x), \quad x \geq 0. \tag{4.5}$$

Obviously, dependence modelling by means of Lévy copulas requires the underlying margins to be Lévy processes. Hence in contrast to the univariate case, hereafter we exclusively consider compound Poisson processes, which are the only Lévy process with piecewise constant sample paths as stated by Proposition 2.3. In order not to disturb the reading flow, we postpone the reasoning why the popular alternative of negative binomial frequency process does not suit in the multivariate setting to the last section of this chapter.

Unfortunately, a universally applicable closed-form approximation for arbitrary Lévy copulas and marginal severities is not available. Notwithstanding, asymptotic VaR_+, ES_+ and SRM_+ representations in terms of Poisson frequency parameters and marginal severity distributions do exist for certain special cases, which shall be presented below and cover a quite substantial range of operational loss situations in practice.

First, we state the results for the important cases of independence and complete positive dependence, which may serve as benchmark values for the impact of dependence structures on risk exposures, provided the marginal parameters have been appropriately estimated and are regarded as fixed. If the dependence structure is described by the independence Lévy copula $\mathfrak{C}_\perp$ from Proposition 2.14, then the entire mass of the Lévy measure Π associated with the d-variate compound Poisson process $S = (S_1, \ldots, S_d)^\top$ is concentrated on the coordinate axes and losses from different risk cells almost surely never occur at the same time. Hence expression (3.2) for the tail integral of the overall loss process S_+ simplifies to $\overline{\Pi}_+(x) = \sum_{i=1}^d \overline{\Pi}_i(x)$ for $x \geq 0$.

Theorem 4.5 (Operational risk measures for independent cells).
If the dependence structure of a d-dimensional compound Poisson model is given by the independence Lévy copula, then the frequency parameter and the severity distribution of the overall loss process S_+ can be explicitly calculated as

$$\lambda_+ \;=\; \sum_{i=1}^d \lambda_i \quad \textit{and} \quad \overline{F}_+(x) \;=\; \frac{1}{\lambda_+}\sum_{i=1}^d \lambda_i \overline{F}_{X_i}(x), \quad x \geq 0,$$

respectively. Furthermore, suppose $\overline{F}_{X_1} \in \mathcal{S} \cap (\mathcal{R} \cup \mathcal{R}_\infty)$ holds and a constant $c_i \geq 0$ exists for each risk cell $i \in \{2, \ldots, d\}$, such that $\overline{F}_{X_i}(x) \sim c_i \overline{F}_{X_1}(x)$ as $x \to \infty$. Then by setting $c = \lambda_1 + \sum_{i=2}^d c_i \lambda_i$, the risk measures related to S_+ are approximated as follows:

(a) The total VaR is asymptotically equivalent to a high quantile of the severity distribution F_{X_1} and satisfies

$$VaR_{+,t}(\alpha) \;\sim\; F_+^{\leftarrow}\left(1 - \frac{1-\alpha}{\lambda_+ t}\right) \;\sim\; F_{X_1}^{\leftarrow}\left(1 - \frac{1-\alpha}{ct}\right), \quad \alpha \to 1.$$

(b) *Assume the distribution tail* $\overline{F}_{X_1}$ *is regularly varying with tail index* $\gamma_1 > 1$*, then the total ES behaves asymptotically as in the one-dimensional case, that is,*

$$ES_{+,t}(\alpha) \sim \frac{\gamma_1}{\gamma_1 - 1} VaR_{+,t}(\alpha), \quad \alpha \to 1.$$

(c) *If* ϕ^* *denotes an admissible risk spectrum fulfilling the conditions in Part (c) of Theorem 4.4, and the marginal severity of risk cell one satisfies the same condition* $\overline{F}_{X_1} \in \mathcal{R}_{-\gamma_1}$ *with* $\gamma_1 > 1$ *as in Part (b), then we obtain for the total SRM with respect to* ϕ^* *the approximation*

$$SRM^{\phi^*}_{+,t}(\alpha) \sim k(\gamma_1, \phi)\, VaR_{+,t}(\alpha), \quad \alpha \to 1,$$

with $k(\gamma_1, \phi) = \int_1^\infty s^{\gamma_1^{-1}-2}\, \phi(1 - s^{-1})\, \mathrm{d}s$.

On the other hand, if the marginal cell processes $S_1, \ldots, S_d$ are completely positively dependent, then losses always occur simultaneously across all d risk cells. Therefore, the expected number of losses per unit time is equal in all cells and the intensity parameter λ_+ of the overall loss process S_+ is readily provided by $\lambda_+ = \lambda_1 = \cdots = \lambda_d$. Furthermore, the utilisation of the complete dependence Lévy copula $\mathfrak{C}_\parallel$ as specified in Proposition 2.16 implies the entire mass of the Lévy measure Π is concentrated on

$$\begin{aligned} &\left\{(x_1, \ldots, x_d)^\top \in [0, \infty)^d \,\middle|\, \overline{\Pi}_1(x_1) = \cdots = \overline{\Pi}_d(x_d)\right\} \\ = &\left\{(x_1, \ldots, x_d)^\top \in [0, \infty)^d \,\middle|\, F_{X_1}(x_1) = \cdots = F_{X_d}(x_d)\right\}. \end{aligned}$$

Theorem 4.6 (**Operational risk measures for completely dependent cells**). *Assume the dependence structure of a d-dimensional compound Poisson model is given by the complete dependence Lévy copula. If all marginal severity distributions* $F_{X_1}, \ldots, F_{X_d}$ *are strictly increasing, the function* $h(x) = x + \sum_{i=2}^d F_{X_i}^{-1}(F_{X_1}(x))$ *is well-defined and invertible for* $x \geq 0$*. Then the severity distribution associated with* S_+ *has the closed-form tail*

$$\overline{F}_+(x) = \overline{F}_{X_1}(h^{-1}(x)), \quad x \geq 0.$$

Furthermore, if $\overline{F}_+ \in \mathcal{S} \cap (\mathcal{R} \cup \mathcal{R}_\infty)$ *holds, then Theorem 4.4 applies and enables the following approximations:*

(a) *The VaR of the overall loss precess* S_+ *asymptotically equals the sum of the stand-alone VaRs, that is,*

$$VaR_{+,t}(\alpha) \sim h\left[F_{X_1}^{-1}\left(1 - \frac{1-\alpha}{\lambda_+ t}\right)\right] \sim \sum_{i=1}^d VaR_{i,t}(\alpha), \quad \alpha \to 1.$$

(b) *Further, assume all marginal severity distributions possess a finite expectation, then the total ES can be similarly approximated by*

$$ES_{+,t}(\alpha) \sim \sum_{i=1}^d ES_{i,t}(\alpha), \quad \alpha \to 1.$$

(c) *Let ϕ^* be an admissible risk spectrum satisfying the conditions in Part (c) of Theorem 4.4. Then under the same assumption as in Part (b) for marginal severities, the total SRM with respect to ϕ^* is asymptotically given by the sum*

$$SRM_{+,t}^{\phi^*}(\alpha) \sim \sum_{i=1}^{d} SRM_{i,t}^{\phi^*}(\alpha), \quad \alpha \to 1.$$

The above result is in line with our claim in Section 1.2 that the simple summation of the stand-alone measures VaR_i, $i \in \{1, \dots d\}$, implicitly assumes complete dependence among all risk cells. However, [BK08] demonstrates that in case of extremely heavy-tailed losses, for example characterised by a Pareto distribution with tail index less than one and thus infinite expectation, the sum of individual VaRs is smaller than the overall VaR_+ calculated based on the independence Lévy copula. Hence the intuition of diversification benefit through independent risk cells must be treated with caution in the heavy-tailed situation. This shall be further illustrated in Section 5.4 with the help of simulated loss data.

The last special case is more pronounced by the constellation of marginal severities rather than the Lévy copula. More precisely, the losses in one risk cell shall possess a regularly varying distribution tail and dominate the losses in all other cells, whereas the dependence structure among the risk cells can be arbitrary.

Theorem 4.7 (**Operational risk measures in case of one dominating cell**). *Without loss of generality, assume $\overline{F}_{X_1} \in \mathcal{R}_{-\gamma_1}$ for some $\gamma_1 > 0$. Moreover, let $\tilde{\gamma} > \gamma_1$ and suppose the $\tilde{\gamma}$-th moment of the severity distribution in all other cells $i \in \{2, \dots, d\}$ is finite. Then regardless of the dependence structure between the risk cells, the asymptotic equivalence*

$$\overline{G}_{+,t}(x) \sim \mathbb{E}[N_1(t)]\overline{F}_{X_1}(x), \quad x \to \infty,$$

holds and we obtain the following risk measure estimations:

(a) *The VaR of the overall loss process S_+ is asymptotically dominated by the stand-alone VaR of the first cell, that is,*

$$VaR_{+,t}(\alpha) \sim VaR_{1,t}(\alpha), \quad \alpha \to 1.$$

(b) *Assume further the severity distribution F_{X_1} has a tail index $\gamma_1 > 1$, then the total ES satisfies a similar approximation*

$$ES_{+,t}(\alpha) \sim \frac{\gamma_1}{\gamma_1 - 1} F_{X_1}^{\leftarrow}\left(1 - \frac{1-\alpha}{\mathbb{E}[N_1(t)]}\right) \sim ES_{1,t}(\alpha), \quad \alpha \to 1.$$

(c) *Let ϕ^* be an admissible risk spectrum complying with the conditions in Part (c) of Theorem 4.4. Under the same requirement $\overline{F}_{X_1} \in \mathcal{R}_{-\gamma_1}$ with $\gamma_1 > 1$ as in Part (b),*

the overall SRM with respect to ϕ^* is asymptotically equivalent to the SRM of risk cell one, that is,

$$SRM^{\phi^*}_{+,t}(\alpha) \sim k(\gamma_1, \phi) F^{\leftarrow}_{X_1}\left(1 - \frac{1-\alpha}{\mathbb{E}[N_1(t)]}\right) \sim SRM^{\phi^*}_{1,t}(\alpha), \quad \alpha \to 1,$$

with $k(\gamma_1, \phi) = \int_1^\infty s^{\gamma_1^{-1}-2}\, \phi(1 - s^{-1})\, \mathrm{d}s$.

Note that the above theorem has an exceptional wide range of applicability, as it does not only hold for dependence modelling via Lévy copulas, but also for arbitrary dependence concepts between the marginal compound Poisson process. However, in the latter case the overall loss process S_+ is not necessarily a compound Poisson one itself.

4.2 A closed-form expression for the overall loss severity in bivariate compound Poisson models

In this section we exploit the fact that $S_+(t) \sim \mathrm{CPP}(\lambda_+, F_+)$ constitutes itself a one-dimensional compound Poisson process and derive analytical expressions for λ_+ as well as F_+ in a bivariate setting. First, recall equation (3.2) which links the tail integral $\overline{\Pi}_+$ associated with $S_+(t)$ to the Lévy measure Π of the bivariate process $S(t) = (S_1(t), S_2(t))^\top$, that is,

$$\overline{\Pi}_+(x) = \Pi\left(\left\{(x_1, x_2)^\top \in [0,\infty)^2 \setminus \{0\} \,\middle|\, x_1 + x_2 \geq x\right\}\right), \quad x \geq 0. \tag{4.6}$$

Moreover, the two-dimensional model can be written according to equation (3.4) as a random sum $S(t) = \sum_{h=1}^{N(t)} (Y_{1h}, Y_{2h})^\top$ of i.i.d. loss severities with bivariate distribution function F_Y, which are compounded via a homogeneous Poisson process $N(t)$ with intensity $\lambda > 0$. As already explained after Definition 2.5 of Lévy measure and via a slight abuse of notation, the measure $\Pi(\cdot)$ is readily given by $\lambda F_Y(\cdot)$, where $F_Y(\cdot)$ shall represent the underlying probability law of a generic random bivariate loss $Y = (Y_1, Y_2)^\top$. As the probability distribution of Y has by definition of compound Poisson processes no atom at zero, the entire mass of the Lévy measure Π is exhausted by taking the limit $\lambda_+ = \lim_{x \downarrow 0} \overline{\Pi}_+(x)$. Hence we immediately conclude that the frequency parameter λ_+ must be equal to the frequency parameter λ of the bivariate process $S(t) \sim \mathrm{CPP}(\lambda, F_Y)$.

In fact, as a jump of the process $S(t)$ almost surely manifests itself in a loss of at least one of its two components, the overall loss process admits a representation

$$S_+(t) = \sum_{h=1}^{N(t)} Y_{1h} + Y_{2h}, \quad t \geq 0.$$

In line with the Lévy measure interpretation (4.6), this yields the survival function of the overall loss severity distribution as

$$\overline{F}_+(x) = \mathbb{P}(Y_1 + Y_2 > x), \quad x \geq 0. \tag{4.7}$$

On the other hand, the associated Lévy measure $\overline{\Pi}_+$ as stated in (4.6) can be decomposed into the following three parts:

$$\begin{aligned}\overline{\Pi}_+(x) &= \Pi\left(\left\{(x_1,0)^\top \in [0,\infty)\times\{0\} \,\middle|\, x_1 \geq x\right\}\right)\\ &\quad + \Pi\left(\left\{(0,x_2)^\top \in \{0\}\times[0,\infty) \,\middle|\, x_2 \geq x\right\}\right)\\ &\quad + \Pi\left(\left\{(x_1,x_2)^\top \in (0,\infty)^2 \,\middle|\, x_1+x_2 \geq x\right\}\right)\\ &= \overline{\Pi}_1^\perp(x) + \overline{\Pi}_2^\perp(x) + \overline{\Pi}_+^\parallel(x), \end{aligned} \tag{4.8}$$

where for the last line we have employed the notation from the proof of Lemma 3.4. By taking limit on both sides of the above equation, we arrive at another representation of the overall loss frequency as

$$\begin{aligned}\lambda_+ = \lim_{x\downarrow 0}\overline{\Pi}_+(x) &= \lim_{x\downarrow 0}\overline{\Pi}_1^\perp(x) + \lim_{x\downarrow 0}\overline{\Pi}_2^\perp(x) + \lim_{x\downarrow 0}\overline{\Pi}_+^\parallel(x)\\ &= \lambda_1^\perp + \lambda_2^\perp + \lambda^\parallel\\ &= (\lambda_1 - \lambda^\parallel) + (\lambda_2 - \lambda^\parallel) + \lambda^\parallel\\ &= \lambda_1 + \lambda_2 - \lambda^\parallel, \end{aligned} \tag{4.9}$$

where λ_1 and λ_2 denote the frequency parameters of the marginal processes $S_1(t)$ and $S_2(t)$, respectively. Recall that the parameters λ_i, $i \in \{1,2\}$, are linked to the bivariate loss severities $(Y_1,Y_2)^\top$ through $\lambda_i = \lambda\overline{F}_{Y_i}(0)$. In addition, equation (3.24) connects the Poisson frequency $\lambda^\parallel$ to the joint survival function $\overline{F}_Y$ via

$$\lambda^\parallel = \lambda\overline{F}_Y(0,0) = \lambda\left(\overline{F}_{Y_1}(0) + \overline{F}_{Y_2}(0) + F_Y(0,0) - 1\right) = \lambda_1 + \lambda_2 - \lambda$$

as $F_Y(0,0) = 0$. By substituting the above expression for $\lambda^\parallel$ into equation (4.9), we obtain once again the identity $\lambda_+ = \lambda$ and everything fits together.

Now we turn our attention to deriving a representation of the overall loss severity F_+ in terms of the marginal compound Poisson processes $S_i(t) \sim \mathrm{CPP}(\lambda_i, F_{X_i})$, $i \in \{1,2\}$, and the Lévy copula $\mathfrak{C}$. It is already stated in (4.5) that the associated survival function $\overline{F}_+$ can be retrieved from the tail integral as

$$\overline{F}_+(x) = \frac{\overline{\Pi}_+(x)}{\lambda_+} = \frac{\overline{\Pi}_1^\perp(x) + \overline{\Pi}_2^\perp(x) + \overline{\Pi}_+^\parallel(x)}{\lambda_1 + \lambda_2 - \lambda^\parallel}, \tag{4.10}$$

where for the second equality we have used decomposition (4.8) for $\overline{\Pi}_+$ and decomposition (4.9) for λ_+, respectively. From Part (b) of Lemma 3.4 we know the independent tail integrals can be written as

$$\begin{aligned}\overline{\Pi}_1^\perp(x) &= \lambda_1^\perp\overline{F}_1^\perp(x) = \lambda_1\overline{F}_{X_1}(x) - \mathfrak{C}(\lambda_1\overline{F}_{X_1}(x), \lambda_2),\\ \overline{\Pi}_2^\perp(x) &= \lambda_2^\perp\overline{F}_2^\perp(x) = \lambda_2\overline{F}_{X_2}(x) - \mathfrak{C}(\lambda_1, \lambda_2\overline{F}_{X_2}(x)),\end{aligned}$$

and the intensity parameter corresponding to the dependent Lévy measure $\Pi^\parallel$ can be calculated from the Lévy copula as $\lambda^\parallel = \mathfrak{C}(\lambda_1,\lambda_2)$. As a result, the Lévy measure $\overline{\Pi}_+^\parallel$

remains the only unknown term in the last fraction of equation (4.10). As $\overline{\Pi}_+^{\|}$ precisely describes the common loss severities $(X_1^{\|}, X_2^{\|})^\top$ with distribution function $F^{\|}$, we compute

$$\begin{aligned}
\overline{\Pi}_+^{\|}(x) &= \Pi^{\|}\left(\left\{(x_1,x_2)^\top \in (0,\infty)^2 \,\middle|\, x_1+x_2 \geq x\right\}\right) \\
&= \lambda^{\|}\,\mathbb{P}(X_1^{\|}+X_2^{\|} \geq x) \\
&= \lambda^{\|}\int_0^\infty \mathbb{P}(X_2^{\|} \geq x-z|X_1^{\|}=z)\,F_1^{\|}(\mathrm{d}z) \\
&= \lambda^{\|}\int_0^x \mathbb{P}(X_2^{\|} \geq x-z|X_1^{\|}=z)\,F_1^{\|}(\mathrm{d}z) \;+\; \lambda^{\|}\int_x^\infty F_1^{\|}(\mathrm{d}z) \\
&= \lambda^{\|}\int_0^x \overline{F}_{2|1}^{\|}(x-z|z)f_1^{\|}(z)\mathrm{d}z \;+\; \lambda^{\|}\int_x^\infty f_1^{\|}(z)\mathrm{d}z,
\end{aligned}$$

where $\overline{F}_{2|1}^{\|}(x_2|x_1)$ denotes the conditional distribution of the second component $X_2^{\|}$ given the first one $X_1^{\|}$. We continue with the calculation of this conditional distribution by limit considerations, that is,

$$\begin{aligned}
\overline{F}_{2|1}^{\|}(x_2|x_1) &= \lim_{h\downarrow 0}\mathbb{P}(X_2^{\|} > x_2 \,|\, x_1 < X_1^{\|} \leq x_1+h) \;=\; \lim_{h\downarrow 0}\frac{\mathbb{P}(x_1 < X_1^{\|} \leq x_1+h, X_2^{\|} > x_2)}{\mathbb{P}(x_1 < X_1^{\|} \leq x_1+h)} \\
&= \lim_{h\downarrow 0}\frac{\overline{F}^{\|}(x_1,x_2)-\overline{F}^{\|}(x_1+h,x_2)}{F_1^{\|}(x_1+h)-F_1^{\|}(x_1)} \;=\; -\frac{\partial}{\partial x_1}\overline{F}^{\|}(x_1,x_2)\,\frac{1}{f_1^{\|}(x_1)}.
\end{aligned}$$

We have already deduced expressions for $\overline{F}^{\|}(x_1,x_2)$ and $\overline{F}_1^{\|}(x_1)$ in terms of the Lévy copula in (3.16) and (3.21), which allow to proceed with the corresponding derivatives

$$\begin{aligned}
\frac{\partial}{\partial x_1}\overline{F}^{\|}(x_1,x_2) &= \frac{\partial}{\partial x_1}\left[\left(\lambda^{\|}\right)^{-1}\mathfrak{C}\left(\lambda_1\overline{F}_{X_1}(x_1),\lambda_2\overline{F}_{X_2}(x_2)\right)\right] \\
&= -\left(\lambda^{\|}\right)^{-1}\lambda_1 f_{X_1}(x_1)\,\frac{\partial}{\partial u_1}\mathfrak{C}\left(u_1,\lambda_2\overline{F}_{X_2}(x_2)\right)\Big|_{u_1=\lambda_1\overline{F}_{X_1}(x_1)}
\end{aligned}$$

and

$$\begin{aligned}
f_1^{\|}(x_1) &= -\frac{\partial}{\partial x_1}\overline{F}_1^{\|}(x_1) \;=\; -\frac{\partial}{\partial x_1}\left[\left(\lambda^{\|}\right)^{-1}\mathfrak{C}\left(\lambda_1\overline{F}_{X_1}(x_1),\lambda_2\right)\right] \\
&= \left(\lambda^{\|}\right)^{-1}\lambda_1 f_{X_1}(x_1)\,\frac{\partial}{\partial u_1}\mathfrak{C}\left(u_1,\lambda_2\right)\Big|_{u_1=\lambda_1\overline{F}_{X_1}(x_1)},
\end{aligned}$$

respectively. Hence overall we obtain the identity

$$\overline{F}_{2|1}^{\|}(x_2|x_1) \;=\; \frac{\frac{\partial}{\partial u_1}\mathfrak{C}\left(u_1,\lambda_2\overline{F}_{X_2}(x_2)\right)\Big|_{u_1=\lambda_1\overline{F}_{X_1}(x_1)}}{\frac{\partial}{\partial u_1}\mathfrak{C}\left(u_1,\lambda_2\right)\Big|_{u_1=\lambda_1\overline{F}_{X_1}(x_1)}}. \tag{4.11}$$

All in all, we are able to express every component in expression (4.10) for the overall loss severity F_+ through the Lévy copula $\mathfrak{C}$ as well as the marginal Poisson parameters λ_i and severity distributions F_{X_i}, $i \in \{1,2\}$. Given observed loss data, the associated

parameters of the latter can be easily estimated within a maximum likelihood scheme as introduced in Section 3.3. Hence in the current bivariate setting we have an analytical access to the severity distribution F_+. In view of risk measure calculations based on the overall loss process $S_+(t) \sim \mathrm{CPP}(\lambda_+, F_+)$, simulation methods can be reduced to directly sampling from a Poisson frequency process with intensity λ_+ and the univariate severity distribution F_+. In other words, simulation of entire paths of the bivariate process $S(t) = (S_1(t), S_2(t))^\top$ is not required.

In particular, if the severity distribution F_+ belongs to the class $\mathcal{S} \cap (\mathcal{R} \cup \mathcal{R}_\infty)$, we can apply Theorem 4.4 to asymptotically estimate risk measures at high confidence levels, such that no simulation is necessary at all. Clearly, in practice it is not always straightforward to determine whether the distribution tail $\overline{F}_+$ is subexponential or even regularly varying. Note in (4.7) we established the identity $\overline{F}_+(x) = \mathbb{P}(Y_1+Y_2 > x)$. Although generally the marginal severity distributions F_{X_1} and F_{X_2} are estimated and thus their heavy-tailedness is known, the corresponding property can often be inferred for F_{Y_1} and F_{Y_2} as well. Recall the tail equivalence $\overline{F}_{Y_i} = (1-p_i)\overline{F}_{X_i}$ for $p_i \in [0,1)$ and $i \in \{1,2\}$, as this was derived in (3.13). As the classes $\mathcal{S}$ and $\mathcal{R}_{-\gamma}$, $\gamma \geq 0$, are each closed with respect to tail equivalence, subexponentiality or the regularly varying property of X_i results in the same asymptotic tail behaviour for Y_i.

Similarly to the setting in Theorem 4.7, the regularly varying characteristic of $\overline{F}_+$ can be deduced in case of one cell dominance. Without loss of generality, assume $\overline{F}_{X_1} \in \mathcal{R}_{-\gamma_1}$ for some $\gamma_1 > 0$ and it shall dominate the severity tail $\overline{F}_{X_2}$, that is,

$$\lim_{x\to\infty} \frac{\overline{F}_{X_1}(xt)}{\overline{F}_{X_1}(x)} = t^{-\gamma_1},\ t > 0, \quad \text{and} \quad \lim_{x\to\infty} \frac{\overline{F}_{X_2}(x)}{\overline{F}_{X_1}(x)} = 0 \tag{4.12}$$

hold. As just explained, the associated distribution tails $\overline{F}_{Y_1}$ and $\overline{F}_{Y_2}$ are simply a rescaled version of $\overline{F}_{X_1}$ and $\overline{F}_{X_2}$, respectively. Hence the same relation (4.12) applies to the distributions of the random variables Y_1 and Y_2. For arbitrary $x > 0$, the probability $\mathbb{P}(Y_1 + Y_2 > x)$ is bounded from above by the sum $\mathbb{P}(Y_1 > x(1-\epsilon)) + \mathbb{P}(Y_2 > x\epsilon)$ for any $\epsilon \in (0,1)$. Together with (4.12), this yields the asymptotic estimation

$$\limsup_{x\to\infty} \frac{\mathbb{P}(Y_1+Y_2 > x)}{\mathbb{P}(Y_1 > x)} \leq \lim_{x\to\infty} \frac{\mathbb{P}(Y_1 > x(1-\epsilon))}{\mathbb{P}(Y_1 > x)} + \lim_{x\to\infty} \frac{\mathbb{P}(Y_2 > x\epsilon)}{\mathbb{P}(Y_1 > x)} = (1-\epsilon)^{-\gamma_1}.$$

On the other hand, $\mathbb{P}(Y_1 + Y_2 > x)$ is bounded from below by $\mathbb{P}(Y_1 > x(1+\epsilon))$, which leads to

$$\liminf_{x\to\infty} \frac{\mathbb{P}(Y_1+Y_2 > x)}{\mathbb{P}(Y_1 > x)} \geq \lim_{x\to\infty} \frac{\mathbb{P}(Y_1 > x(1+\epsilon))}{\mathbb{P}(Y_1 > x)} = (1+\epsilon)^{-\gamma_1}.$$

By letting $\epsilon \to 0$, we obtain from the above inequalities the asymptotic equality

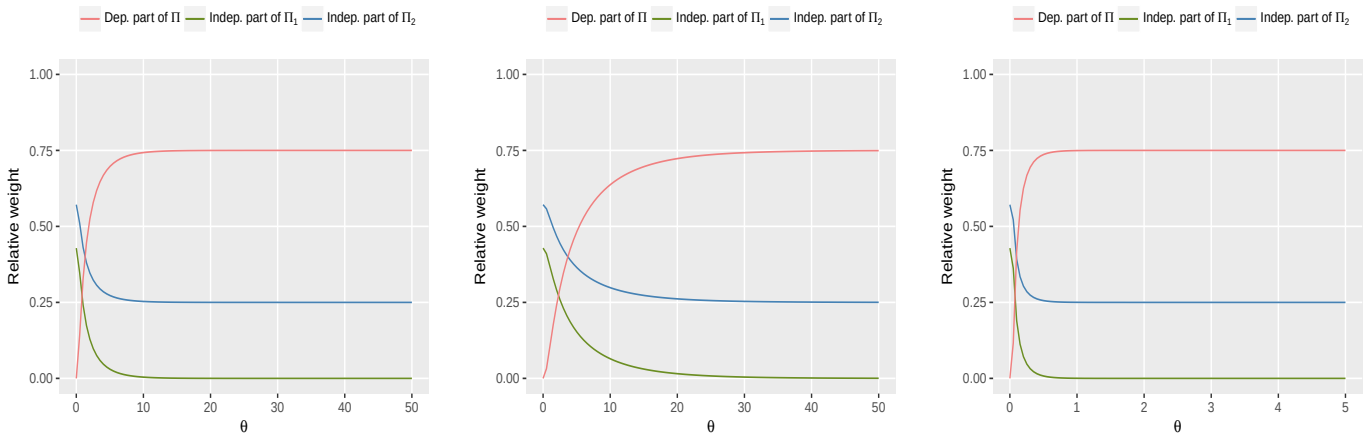
$$\lim_{x\to\infty} \frac{\overline{F}^+(x)}{\overline{F}_{Y_1}(x)} = \lim_{x\to\infty} \frac{\mathbb{P}(Y_1+Y_2 > x)}{\mathbb{P}(Y_1 > x)} = 1.$$

In other words, the distribution tail $\overline{F}_+$ is asymptotically equivalent to the dominating severity tail of the first risk cell and hence lies in the class $\mathcal{R}_{-\gamma_1}$ as well.

However, subexponentiality of marginal severities does not in general imply subexponentiality of $\overline{F}_+$. The far out right tail behaviour of $\overline{F}_+$ must be carefully examined in each particular constellation of marginal distributions and dependence structures. For more details on the heavy-tailedness of the sum of subexponential random variables, we refer the interested readers to [EG80] as well as the more recent publications [GN06], [GT09], [KT08] and [KA09].

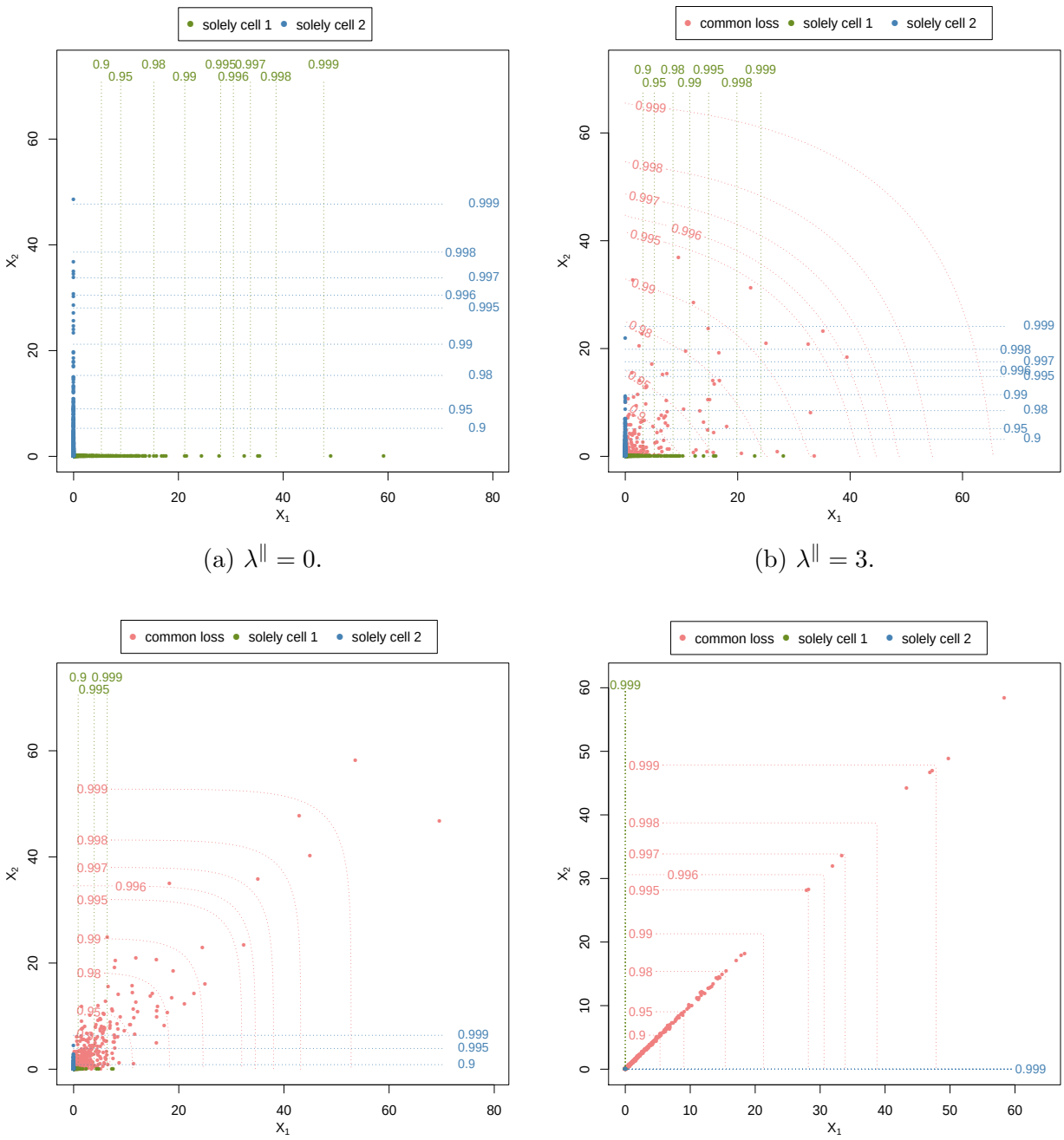
We close the current section with some visualisations of the decomposition of a bivariate compound Poisson model $S = (S_1, S_2)^\top$ into its two individual loss processes $S_1^\perp$, $S_2^\perp$, and one common loss process $S^\parallel = (S_1^\parallel, S_2^\parallel)^\top$. The corresponding Lévy measures are denoted by $\Pi_1^\perp$, $\Pi_2^\perp$, and $\Pi^\parallel$ as before. Clearly, the relative weight of each process compared to the entire measure Π directly depends on the underlying Lévy copula. As detailed in (4.9), the total mass of Π is finite and given by the Poisson intensity $\lambda_+ = \lambda$.

Figure 4.1 shows the contribution of the partial measures $\Pi_1^\perp$, $\Pi_2^\perp$ and $\Pi^\parallel$ for three different one-parametric Lévy copula families. In each subfigure, the relative weights of the partial measures are plotted as a function of the copula parameter θ, whereby the marginal Lévy measures $\Pi_1 = \Pi_1^\perp + \Pi_1^\parallel$ and $\Pi_2 = \Pi_2^\perp + \Pi_2^\parallel$ are fixed to have a total mass of 15 and 20, respectively. Note that the induced dependence strength between the marginal processes increases with the value of θ for each of the three selected copula families. Hence all subfigures have in common that the contribution of the simultaneous loss part $\Pi^\parallel$ grows from near zero in the almost independent case with small θ to the most possibly dependent case with an absolute weight of 15, resulting in a relative weight of $\frac{15}{20} = 0.75$. However, the growth rate is obviously different across the selected copulas. The Lévy copula from Example 3.6 exhibits the sharpest increase, whereas the dependence strength rises quite smoothly in case of the Gumbel Lévy copula. On the other hand, the contribution of the individual loss part $\Pi_1^\perp$ is fully exhausted by the common loss part for large values of θ, thus drops down to zero in all subfigures. In contrast, the individual loss part $\Pi_2^\perp$ retains a relative weight of $\frac{20-15}{20} = 0.25$ even in the strongest dependent situation.



(a) Clayton Lévy copula. (b) Gumbel Lévy copula. (c) Lévy copula in Example 3.6 (2).

Figure 4.1: Relative weights of the partial Lévy measures $\Pi_1^\perp$, $\Pi_2^\perp$ and $\Pi^\parallel$ with respect to the Lévy copula parameter θ.



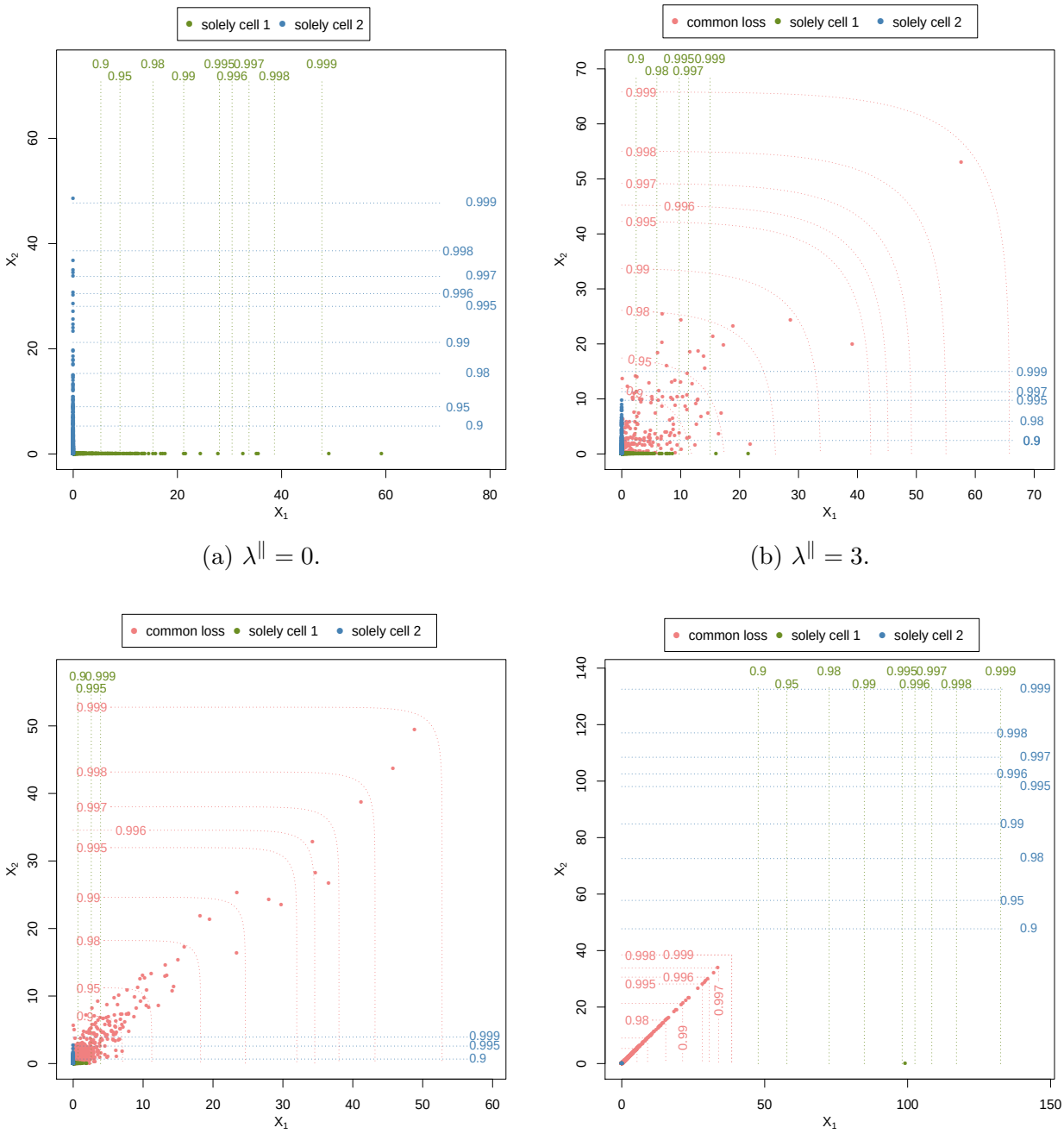
(a) $\lambda^{\|} = 0$.

(b) $\lambda^{\|} = 3$.

(c) $\lambda^{\|} = 7$.

(d) $\lambda^{\|} = 9.9$.

Figure 4.2: Simulation from the bivariate Clayton Lévy copula with different dependence strength. Theoretical quantile contour lines are superimposed on simulated single loss severities.



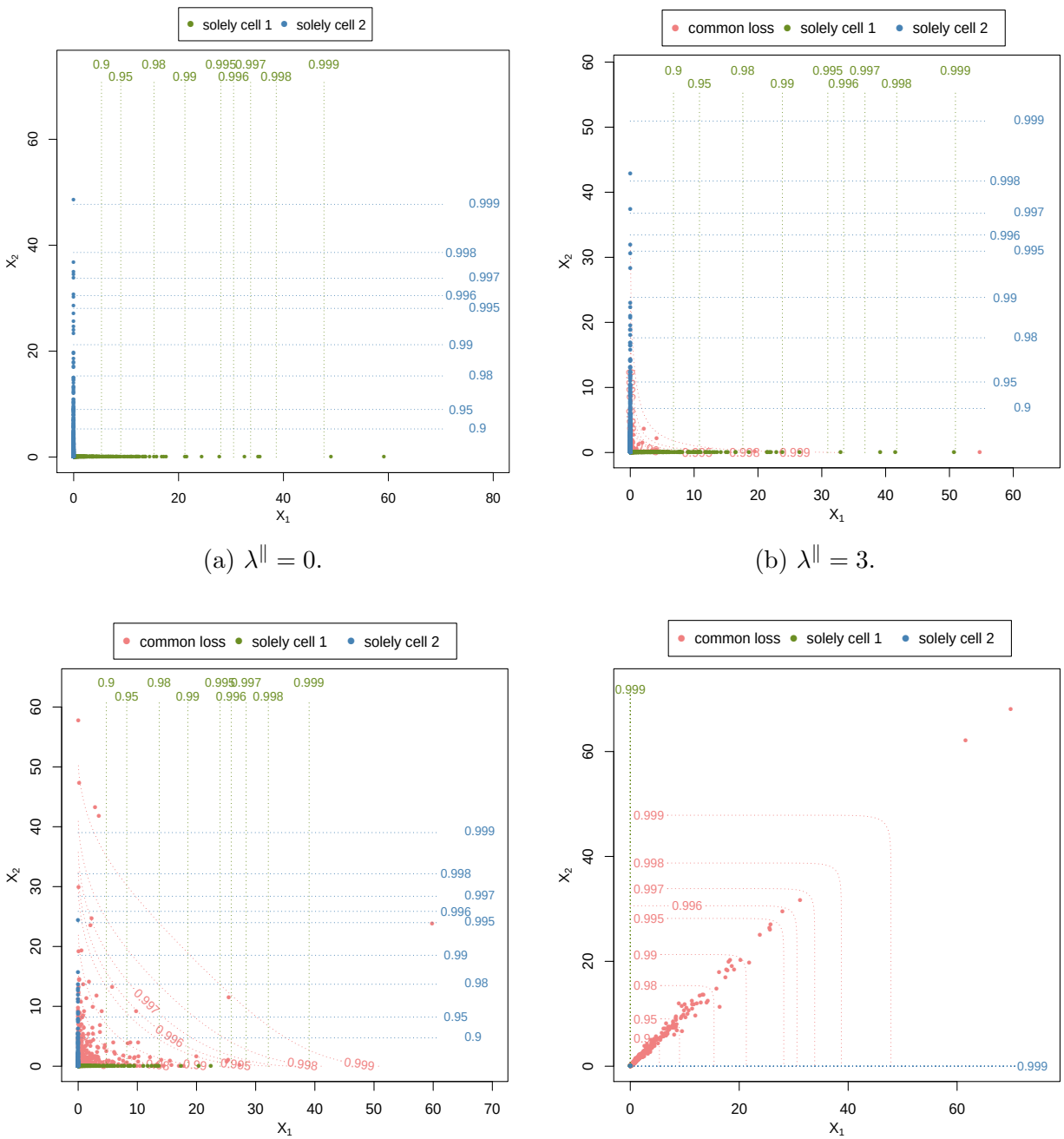
(a) $\lambda^{\|} = 0$.

(b) $\lambda^{\|} = 3$.

(c) $\lambda^{\|} = 7$.

(d) $\lambda^{\|} = 9.9$.

Figure 4.3: Simulation from the bivariate Gumbel Lévy copula with different dependence strength. Theoretical quantile contour lines are superimposed on simulated single loss severities.



(a) $\lambda^{\|} = 0$.

(b) $\lambda^{\|} = 3$.

(c) $\lambda^{\|} = 7$.

(d) $\lambda^{\|} = 9.9$.

Figure 4.4: Simulation from the bivariate Lévy copula in Example 3.6 (2) with different dependence strength. Theoretical quantile contour lines are superimposed on simulated single loss severities.

Going one step further, Figures 4.2-4.4 illustrate the interplay between the two individual loss processes $S_1^\perp$, $S_2^\perp$, and the common loss process $S^\parallel = (S_1^\parallel, S_2^\parallel)^\top$ in view of quantile estimations. The latter of course plays an important role in operational risk measure calculations. In order to highlight the impact of the dependence structure rather than the marginal parameters, in all figures the two single cell processes are assumed to have an identical Poisson intensity of 10 and a heavy-tailed Weibull severity distribution with shape parameter 0.5 and scale parameter 1. Additionally, as to make a comparison across different copula families reasonable to some extent, we select three different absolute weights of the simultaneous loss part $\Pi^\parallel$, which are readily given by the corresponding Poisson intensities $\lambda^\parallel = 3, 7, 9.9$, representing weak, medium and high dependence strength. Then we calculate the resulting copula parameters for the Clayton, the Gumbel and the Lévy copula from Example 3.6 (2), and simulate loss data over a period of $[0, 80]$ for each combination of copula families and dependence levels.

Note that the Poisson intensity parameter precisely reflects the expected number of the associated single losses in a time unit, hence the simulated individual loss severities $X_1^\perp$ and $X_2^\perp$, as well as the common loss severities $(X_1^\parallel, X_2^\parallel)^\top$, have an approximately equal sample size for the same value of $\lambda^\parallel$ across all selected copula families, respectively. Furthermore, we draw the theoretical contour lines corresponding to each of the three types of loss severities, in order to illustrate how different dependence structures may have an impact on quantile estimations. As severe loss events are of primary concern in operational risk management, the contour line at level α associated with the bivariate loss severities is calculated such that the survival probability $\mathbb{P}(X_1^\parallel > x_1, X_2^\parallel > x_2)$ is equal to $1 - \alpha$.

In all three Figures 4.2-4.4, the independence case is depicted in the upper left corner as a benchmark situation. We observe that the contour lines associated with the individual loss severities $X_1^\perp$ and $X_2^\perp$ show a decreasing trend with growing overall weight of the simultaneous loss part $\Pi^\parallel$ for all three copula families, with a single exception given by the high dependence case induced by the Gumbel Lévy copula. On the other hand, the contour lines associated with the bivariate common loss severities $(X_1^\parallel, X_2^\parallel)^\top$ exhibit an interesting convex shape in the low and medium dependence cases for the Lévy copula from Example 3.6 (2). In contrast, the contour lines derived from the Clayton and the Gumbel Lévy copulas have similar concave shapes. Nevertheless, the bivariate contour lines approach in the strongest dependent case the rectangle shape induced by the comonotonic copula across all three investigated copula families. On the whole, Figures 4.2-4.4 allow for interesting insights into the contribution of each partial processes $S_1^\perp$, $S_2^\perp$, and $S^\parallel$ to the overall quantiles under different assumptions of dependence structures. Thereby, the contribution of each set of contour lines is reflected by the relative weight of the associated Lévy measure $\Pi_1^\perp$, $\Pi_2^\perp$ or $\Pi^\parallel$, which is itself determined by the underlying Lévy copula and already illustrated in Figure 4.1.

4.3 Discussions and extensions

In comparison to the simulation approaches for risk measure estimations, the closed-form asymptotic results in Section 4.1 allow for more transparent sensitivity statements with respect to different model components, are straightforward to implement, and of course offer less time-consuming calculations. Therefore, it is not surprising that different refinements of such analytical formulas have gained great attention in both academia and practice. Since equation (4.4) can be interpreted as that one single severe loss event instead of the accumulation of small events determines the overall risk exposure, such asymptotic results are often called single-loss approximations (SLAs) in operational risk.

For instance, [BS06] derives an improved approximation for the case of large frequency expectation combined with severity random variables having finite expectation, that is, the loss severities are not extremely heavy-tailed. Despite the important role of the biggest single loss at a very high quantile level, expression (4.4) could underestimate the real risk exposure, as medium sized losses also contribute to the total VaR with a not negligible amount in this particular constellation. As a result, the authors of [BS06] add to equation (4.4) the product of severity mean and frequency mean subtracted by one, which is known as mean correction. Later on, the SLA is further refined in [Deg10], where the author not only differentiates between finite and infinite severity expectations, but also suggests distinct asymptotic expressions depending on the value of the tail index γ. The incentive of his refinements originates from analysing the relative error of the standard approximation (4.4) by the theory of second-oder subexponentiality.

To give a research example from the industry, [Opd14] and [Opd17] by the same author account for the potential divergence caused by the approximation of [Deg10] for the case of a tail index γ close to one. More precisely, the non-divergent approximation for γ exactly at one is used as an anchor to cross over the divergence zone by means of linear interpolation. Inspired by this idea, the authors of the R package OpVaR (cf. [Zou+18]) apply monotonic cubic spline interpolations to circumvent the divergence problem, whereby the author of the current thesis contributes to its implementation. Furthermore, we refer to Chapter 8 in the monograph [PS15] for a comprehensive overview of SLA refinements with theoretical deviations and detailed proofs.

Besides univariate considerations, the concept of multivariate subexponentiality also constitutes an active field of study and may be utilised to obtain alternative operational risk measure estimations. To illustrate this, we briefly discuss an analogous result to Theorem 4.3 for the two-dimensional case. As before, let X_1 and X_2 denote the severity random variables in the first and the second risk cell, respectively. Their partial sums are given by $S_1^n = \sum_j^n X_{1j}$ and $S_2^n = \sum_l^n X_{2l}$, whose joint distribution function is defined as $F_X^{n*}(x_1, x_2) = \mathbb{P}(S_1^n \leq x_1, S_2^n \leq x_2)$ for $(x_1, x_2)^\top \in [0, \infty)^\top$. We write $\overline{F_X^{n*}}(x_1, x_2)$ for the corresponding survival function. Similarly to the one-dimensional case, the joint distribution G of the random sum $(S_1 = \sum_j^N X_{1j}, S_2 = \sum_l^N X_{2l})^\top$, obtained by compounding the loss severities with a discrete random variable N, has the form $G(x_1, x_2) = \sum_{n=0}^\infty \mathbb{P}(N = n) F_X^{n*}(x_1, x_2)$. If the marginal severity distributions F_{X_1} and F_{X_2} are subexponential, then the n-fold convolution satisfies according to [DOV07] the

approximation $\overline{F_X^{n*}}(x_1, x_2) \sim n\overline{F}_X(x_1, x_2)$ as the minimum of x_1 and x_2 tends to infinity. Furthermore, in case of N fulfilling condition (4.2), the asymptotic equivalence $\overline{G}(x_1, x_2) \sim \mathbb{E}[N]\overline{F}_X(x_1, x_2)$ holds under the same limit taking. Note here the asymptotic tail probability of the random vector $(S_1, S_2)^\top$ is derived solely based on univariate subexponential assumptions about F_{X_1} and F_{X_2}. We mention the reference [CR92] and the more recent publication [OMS06] for alternative statements relying on the subexponentiality of the bivariate joint distribution F_X.

As promised at the beginning of Section 4.1.2, we dedicate the last part of this chapter to explaining the inadequacy of negative binomial processes as the frequency component in a multivariate model based on Lévy copulas. To begin with, the motivation of utilising a negative binomial process $N^{NB}(t)$, $t \geq 0$, comes from its ability of incorporating over-dispersion in loss counts. That is, the loss frequency has greater variance than its expectation. On the contrary, the loss number $N^P(t)$ up to time t described by a homogeneous Poisson process with intensity $\lambda > 0$ has equal mean and variance given by λt. Nevertheless, the negative binomial process is closely related to the Poisson one. More specifically, consider a Poisson process $\tilde{N}^P(t)$, $t \geq 0$, whose intensity parameter is not constant any more, but becomes a gamma random variable Λ, which is independent from $\tilde{N}^P(t)$ and has the density

$$f(\lambda) = \frac{b^a}{\Gamma(a)}\lambda^{a-1}e^{-b\lambda}, \quad \lambda > 0.$$

Then the probability mass function can be calculated for arbitrary $n \in \mathbb{N}_0$ as

$$\begin{aligned}\mathbb{P}(\tilde{N}^P(t) = n) &= \int_0^\infty \mathbb{P}(\tilde{N}^P(t) = n \mid \Lambda = \lambda) f(\lambda)\,\mathrm{d}\lambda = \int_0^\infty e^{-\lambda t}\frac{(\lambda t)^n}{n!} f(\lambda)\,\mathrm{d}\lambda \\ &= \binom{a+n-1}{n}\left(\frac{b}{b+t}\right)^a\left(\frac{t}{b+t}\right)^n,\end{aligned}$$

which precisely reflects the probability $\mathbb{P}(N^{NB}(t) = n)$ of a negative binomial process with parameters $a, b > 0$. For $t \geq 0$, it has expectation $\mathbb{E}[N^{NB}(t)] = ab^{-1}t$ and variance $\mathrm{Var}[N^{NB}(t)] = (1 + b^{-1}t)\mathbb{E}[N^{NB}(t)] > \mathbb{E}[N^{NB}(t)]$.

However, if we compound i.i.d. random losses X_n, $n \geq 1$, via N^{NB}, the resulting compound negative binomial process does not belong to the class of Lévy processes. The most simple way to see this is by recalling Proposition 2.3. Namely, a compound negative binomial process constitutes a stochastic process with piecewise constant sample paths, but the only Lévy process with this property is provided by the compound Poisson process as specified in Definition 2.2. To further underline this, as well as to explicitly demonstrate the infeasibility of combining negative binomial loss frequencies with Lévy copulas, we show the condition of independent and stationary increments for being a Lévy process is violated by N^{NB}.

Amongst others, the author of [Lu14] attempts to characterise N^{NB} itself as a compound Poisson process with logarithmic jump distributions. More precisely, let $M(t)$ denote a Poisson random variable with mean $-a\ln\left(\frac{b}{b+t}\right)$ and let $L(t)$ be a logarithmically dis-

tributed random variable with probability mass function

$$\mathbb{P}(L(t) = n) \;=\; -\frac{\left(\frac{t}{b+t}\right)^n}{n \ln\left(\frac{b}{b+t}\right)}, \quad n \in \mathbb{N}. \tag{4.13}$$

Then the negative binomial random variable $N^{NB}(t)$ indeed admits the representation $N^{NB}(t) = \sum_{k=1}^{M(t)} L_k(t)$ as a Poisson random sum with each summand $L_k(t)$ having the logarithmic distribution specified in (4.13). However, we immediately observe that the associated Poisson process $M(t)$, $t \geq 0$, is not homogeneous, as its intensity $-a \ln\left(\frac{b}{b+t}\right)$ does not provide a linear function of t. In addition, [Lu14] argues that the aggregate loss process of a single risk cell with negative binomial frequency $N^{NB}(t)$ could be expressed through

$$S(t) \;=\; \sum_{n=1}^{N^{NB}(t)} X_n \;=\; \sum_{k=1}^{M(t)} \left(\sum_{n=1}^{L_k(t)} X_n \right), \quad t \geq 0,$$

which shall be interpreted as a compound Poisson process with frequency $M(t)$ and severities $\sum_{n=1}^{L_k(t)} X_n$, $k \geq 1$. Still, neither the distribution of the increment $\sum_{n=1}^{L_k(t)} X_n$ is independent of the time t.

In conclusion, although a compound negative binomial process can be written in the form of a Poisson random sum, neither the compounding frequency process is a homogeneous Poisson one, nor the i.i.d. property of the corresponding summands is provided. On the other hand, the fundamental idea behind Lévy copulas is that they operate on the domain of time-independent Lévy measures, such that a dependence structure between marginal Lévy processes can be solely specified through a Lévy copula and stays invariant against the course of time. As a result, the marginal homogeneous Poisson frequency processes in our dependence model as detailed in Definition 3.1 cannot be replaced by negative binomial ones offhand, and the multivariate asymptotic risk measure approximations relying on Lévy copulas do not apply to compound negative binomial processes. Of course, one can connect univariate compound negative binomial processes by means of ordinary copulas instead. However, this would request more model parameters in comparison to a compound Poisson model based on Lévy copulas, such that its practicability is questionable in view of generally scarce operational risk data. Last but not least, note that the frequency component enters the analytical estimation formulas both in the one-dimensional case treated in Section 4.1.1 and in the multidimensional case treated in Section 4.1.2 only with its expectation instead of variance. Therefore, the potential benefit of modelling overdispersion by utilising negative binomial frequency processes is in fact insignificant with regard to risk measure calculations.

Chapter 5

Simulation study

After having presented the theory of dependence modelling and risk measure estimations based on Lévy copulas, the current chapter aims at demonstrating the practical implementation by means of simulation. First, an algorithm for sampling from an arbitrary bivariate compound Poisson model is introduced in Section 5.1. This allows generating loss data as input for MLE procedures, whose goodness is assessed in Section 5.2 for various parameterisations of marginal components and dependence structures. Next, Section 5.3 explores potential concepts for evaluating the fit of an estimated model. Since accurate and stable capital reserve estimations are of primary interest for financial institutions, Section 5.4 studies the sensitivity of risk measure values towards different model components as well as the considered confidence level α. Last but not least, the consequences of dependence structure misspecification on risk exposure outcomes are studied within a simulation example at the end of this chapter.

5.1 A flexible algorithm for sampling from bivariate compound Poisson models

Losses characterised through a bivariate compound Poisson model $S(t) = (S_1(t), S_2(t))^\top$ can be simulated by decomposing $S(t)$ into its three independent partial processes $S_1^\perp(t)$, $S_2^\perp(t)$ and $S^\parallel(t)$. Then a sample path of $S(t)$ in a prescribed time interval $[0, T]$ is obtained by recombining the losses simulated from the partial processes. The notations applied in the algorithm description below coincide with those introduced in Section 3.2.2.

Algorithm 5.1 (**Simulation of a bivariate compound Poisson model**).
Input: *a time horizon T, marginal Poisson intensity parameters λ_1 and λ_2, marginal severity distributions F_{X_1} and F_{X_2}, and a Lévy copula $\mathfrak{C}$.*
Output: *loss occurrence times and loss severities of a bivariate compound Poisson model.*

Step 1: Calculate the Poisson parameter of the dependent process $S^\parallel$ as $\lambda^\parallel = \mathfrak{C}(\lambda_1, \lambda_2)$ and the Poisson parameters of the independent processes $S_i^\perp$ as $\lambda_i^\perp = \lambda_i - \lambda^\parallel$, $i \in \{1, 2\}$.

Step 2: Simulate three independent Poisson distributed random variables $N^{\parallel} \sim Poi(\lambda^{\parallel}T)$ *and* $N_i^{\perp} \sim Poi(\lambda_i^{\perp}T)$, $i \in \{1,2\}$, *as the number of losses belonging to* $S^{\parallel}$ *and* $S_i^{\perp}$, $i \in \{1,2\}$, *respectively.*

Step 3: Simulate independent random variables $\Gamma_{1j}^{\perp}$, $j = 1,\ldots,N_1^{\perp}$, *from the Unif*$[0,T]$*-distribution as the loss arrival times of the process* $S_1^{\perp}$. *Similarly, the loss arrival times of the process* $S_2^{\perp}$ *are obtained through generation of independent random variables* $\Gamma_{2l}^{\perp}$, $l = 1,\ldots,N_2^{\perp}$, *from the Unif*$[0,T]$*-distribution as well.*

Step 4: Simulate independent random variables $\Gamma_k^{\parallel}$, $k = 1,\ldots,N^{\parallel}$, *from the Unif*$[0,T]$*-distribution as the occurrence times of the bivariate losses attributed to* $S^{\parallel}$.

Step 5: Simulate independent random variables U_{1j}, $j = 1,\ldots,N_1^{\perp}$; U_{2l}, $l = 1,\ldots,N_2^{\perp}$, *from the Unif*$[0,1]$*-distribution. Then the loss severities attributed to the process* $S_1^{\perp}$ *are given by* $F_1^{\perp\leftarrow}(U_{1j})$, $j = 1,\ldots,N_1^{\perp}$, *and the loss severities attributed to the process* $S_2^{\perp}$ *by* $F_2^{\perp\leftarrow}(U_{2l})$, $l = 1,\ldots,N_2^{\perp}$.

Step 6: Simulate independent random variables V_{1k} *and* V_{2k}, $k = 1,\ldots,N^{\parallel}$, *from the Unif*$[0,1]$*-distribution. If* $F_{2|1}^{\parallel}(x_2|x_1)$ *denotes the conditional distribution of the second component given the first one of a bivariate dependent loss, then the loss severities belonging to the process* $S^{\parallel}$ *are given by* $\left(F_1^{\parallel\leftarrow}(V_{1k}), F_{2|1}^{\parallel\leftarrow}(V_{2k}|F_1^{\parallel\leftarrow}(V_{1k}))\right)^{\top}$, $k = 1,\ldots,N^{\parallel}$.

Clearly, if the interest only lies in the accrued loss amounts up to time point T rather than the chronological loss development, Step 3 and Step 4 of the above algorithm can be omitted. However, the loss arrival times are necessary for obtaining the entire trajectory of $S(t) = (S_1(t), S_2(t))^{\top}$. Based on the simulation output, the bivariate trajectory up to time T is constructed through

$$\begin{aligned} S_1(t) &= \sum_{j=1}^{N_1^{\perp}} F_1^{\perp\leftarrow}(U_{1j})\mathbb{1}_{\{\Gamma_{1j}^{\perp}\le t\}} + \sum_{k=1}^{N^{\parallel}} F_1^{\parallel\leftarrow}(V_{1k})\mathbb{1}_{\{\Gamma_k^{\parallel}\le t\}} \\ \text{and} \quad S_2(t) &= \sum_{l=1}^{N_2^{\perp}} F_2^{\perp\leftarrow}(U_{2l})\mathbb{1}_{\{\Gamma_{2l}^{\perp}\le t\}} + \sum_{k=1}^{N^{\parallel}} F_{2|1}^{\parallel\leftarrow}(V_{2k}|F_1^{\parallel\leftarrow}(V_{1k}))\mathbb{1}_{\{\Gamma_k^{\parallel}\le t\}} \end{aligned}$$

for $t \in [0,T]$. Algorithm 5.1 mainly resembles the procedure proposed by [EK10] and its validity directly follows from Lemma 3.4. Only the following two aspects need some explanation. First, for a homogeneous Poisson process $N(t)$ it is well-known that conditionally on the number $N(T) = n$ of jumps up to time point T, the jump times in the interval $[0,T]$ are distributed like the order statistics of a sample of size n from the uniform distribution on $[0,T]$. Therefore, Step 3 and Step 4 of Algorithm 5.1 simulate the loss arrival times of $S(t)$ up to the given time horizon T correctly.

Second, the generation of loss severities in Step 5 and Step 6 follows from the commonly applied probability integral transform. Depending on the particular marginal distributions and the Lévy copula, the inverses $F_1^{\perp\leftarrow}$, $F_2^{\perp\leftarrow}$, $F_1^{\parallel\leftarrow}$ and $F_{2|1}^{\parallel\leftarrow}$ are computed either via

numerical methods or directly, provided that the corresponding distribution functions can be analytically inverted. In particular, the conditional distribution $F^{\|}_{2|1}$ can be retrieved in different ways and this constitutes the main deviation of Algorithm 5.1 from the procedure in [EK10]. The latter suggests to make use of the survival copula $C^{\|}_s$ underlying the dependent loss severities introduced in (3.23), as the survival function of $F^{\|}_{2|1}$ takes the form $\overline{F}^{\|}_{2|1}(x_2|x_1) = \frac{\partial}{\partial u_1} C^{\|}_s(u_1, \overline{F}^{\|}_2(x_2))\Big|_{u_1=\overline{F}^{\|}_1(x_1)}$. However, this approach is only feasible if the ordinary copula $C^{\|}_s$ is known in closed form, which is not always the case as already explained at the end of Section 3.2.2.

On the contrary, Step 6 of Algorithm 5.1 does not specify a particular requirement on the copula $C^{\|}_s$, indicating the calculation of $F^{\|\leftarrow}_{2|1}$ should be carried out in the most appropriate way depending on the underlying marginal distributions and the Lévy copula. More precisely, equation (4.11) provides a general expression of the corresponding survival distribution $\overline{F}^{\|}_{2|1}$, which only depends on the input parameters of Algorithm 5.1, that is, the marginal distributions and the Lévy copula $\mathfrak{C}$ itself. We illustrate this idea by giving the subsequent example based on the Clayton Lévy family.

Example 5.2. By plugging the generator function $\phi(u) = u^{-\theta}, \theta > 0$, of the Clayton Lévy copula from Table 2.1 into expression (3.25) for $C^{\|}_s$ in the case of underlying Archimedean Lévy copulas, we find out that the survival copula $C^{\|}_s$ is readily given by the ordinary Clayton copula

$$C^{\|}_s(u_1, u_2) \;=\; (u_1^{-\theta} + u_2^{-\theta} - 1)^{-\frac{1}{\theta}}.$$

Its partial derivative with respect to the first entry is given by

$$\frac{\partial}{\partial u_1} C^{\|}_s(u_1, u_2) \;=\; \left[1 + \left(\frac{u_1}{u_2}\right)^{\theta} - u_1^{\theta}\right]^{-\frac{1}{\theta}-1}. \tag{5.1}$$

Furthermore, the marginal survival functions of the bivariate dependent losses $(X^{\|}_1, X^{\|}_2)^\top$ have already been computed in equations (3.21)-(3.22) for a general Lévy copula. In the current Clayton setting they take the form

$$\overline{F}^{\|}_1(x_1) \;=\; \frac{1}{\lambda^{\|}}\mathfrak{C}(\lambda_1 \overline{F}_{X_1}(x_1), \lambda_2) \;=\; \frac{1}{\lambda^{\|}}\left[\lambda_1^{-\theta}\overline{F}_{X_1}^{-\theta}(x_1) + \lambda_2^{-\theta}\right]^{-\frac{1}{\theta}} \tag{5.2}$$

$$\text{and} \quad \overline{F}^{\|}_2(x_2) \;=\; \frac{1}{\lambda^{\|}}\mathfrak{C}(\lambda_1, \lambda_2 \overline{F}_{X_2}(x_2)) \;=\; \frac{1}{\lambda^{\|}}\left[\lambda_1^{-\theta} + \lambda_2^{-\theta}\overline{F}_{X_2}^{-\theta}(x_2)\right]^{-\frac{1}{\theta}}. \tag{5.3}$$

Then the conditional distribution $F^{\|}_{2|1}(x_2|x_1)$ can be obtained through expressing the corresponding survival function in terms of $C^{\|}_s$, that is, we insert equations (5.2)-(5.3) into (5.1) and calculate

$$\begin{aligned}\overline{F}^{\|}_{2|1}(x_2|x_1) \;&=\; \frac{\partial}{\partial u_1} C^{\|}_s(u_1, \overline{F}^{\|}_2(x_2))\Big|_{u_1=\overline{F}^{\|}_1(x_1)} \\ &=\; \left[1 + \frac{\lambda_1^{-\theta} + \lambda_2^{-\theta}\overline{F}_{X_2}^{-\theta}(x_2)}{\lambda_1^{-\theta}\overline{F}_{X_1}^{-\theta}(x_1) + \lambda_2^{-\theta}} - \frac{(\lambda^{\|})^{-\theta}}{\lambda_1^{-\theta}\overline{F}_{X_1}^{-\theta}(x_1) + \lambda_2^{-\theta}}\right]^{-\frac{1}{\theta}-1}.\end{aligned}$$

Now recall the Poisson frequency parameter of the common loss severities $(X_1^{\|}, X_2^{\|})^\top$ is related to the Lévy copula as $\lambda^{\|} = \mathfrak{C}(\lambda_1, \lambda_2) = \left(\lambda_1^{-\theta} + \lambda_2^{-\theta}\right)^{-\frac{1}{\theta}}$, which further simplifies the above equation to

$$\overline{F}_{2|1}^{\|}(x_2|x_1) = \left[\frac{\lambda_1^{-\theta}\overline{F}_{X_1}^{-\theta}(x_1) + \lambda_2^{-\theta}\overline{F}_{X_2}^{-\theta}(x_2)}{\lambda_1^{-\theta}\overline{F}_{X_1}^{-\theta}(x_1) + \lambda_2^{-\theta}}\right]^{-\frac{1}{\theta}-1}. \tag{5.4}$$

On the other hand, we can represent $\overline{F}_{2|1}^{\|}(x_2|x_1)$ without utilising $C_s^{\|}$, but directly via the Lévy copula $\mathfrak{C}$ itself. For this purpose, we first need the partial derivative of the Clayton Lévy copula

$$\frac{\partial}{\partial u_1}\mathfrak{C}(u_1, u_2) = \left[1 + \left(\frac{u_1}{u_2}\right)^{\theta}\right]^{-\frac{1}{\theta}-1},$$

then the conditional distribution follows from equation (4.11) as

$$\begin{aligned}\overline{F}_{2|1}^{\|}(x_2|x_1) &= \frac{\frac{\partial}{\partial u_1}\mathfrak{C}\left(u_1, \lambda_2\overline{F}_{X_2}(x_2)\right)\Big|_{u_1=\lambda_1\overline{F}_{X_1}(x_1)}}{\frac{\partial}{\partial u_1}\mathfrak{C}\left(u_1, \lambda_2\right)\Big|_{u_1=\lambda_1\overline{F}_{X_1}(x_1)}} = \frac{\left[1 + \left(\frac{\lambda_1\overline{F}_{X_1}(x_1)}{\lambda_2\overline{F}_{X_2}(x_2)}\right)^{\theta}\right]^{-\frac{1}{\theta}-1}}{\left[1 + \left(\frac{\lambda_1\overline{F}_{X_1}(x_1)}{\lambda_2}\right)^{\theta}\right]^{-\frac{1}{\theta}-1}} \\ &= \left[\frac{1 + \lambda_1^{\theta}\overline{F}_{X_1}^{\theta}(x_1)\lambda_2^{-\theta}\overline{F}_{X_2}^{-\theta}(x_2)}{1 + \lambda_1^{\theta}\overline{F}_{X_1}^{\theta}(x_1)\lambda_2^{-\theta}}\right]^{-\frac{1}{\theta}-1}.\end{aligned}$$

Hence by expanding both the numerator and denominator of the inner fraction in the above equation with $\lambda_1^{-\theta}\overline{F}_{X_1}^{-\theta}(x_1)$, once again we arrive at the same expression for $\overline{F}_{2|1}^{\|}(x_2|x_1)$ as detailed in (5.4).

Going one step further, assume the marginal severity in risk cell $i \in \{1, 2\}$ follows a heavy-tailed GPD(ξ_i, β_i) distribution with zero location parameter, shape $\xi_i > 0$ and scale $\beta_i > 0$. The corresponding survival function thus has the form

$$\overline{F}_{X_i}(x_i) = \left(1 + \xi_i\frac{x_i}{\beta_i}\right)^{-\frac{1}{\xi_i}}, \quad x_i > 0. \tag{5.5}$$

Then the conditional survival function $\overline{F}_{2|1}^{\|}(x_2|x_1)$ can be explicitly stated as

$$\overline{F}_{2|1}^{\|}(x_2|x_1) = \left[\frac{\lambda_1^{-\theta}\left(1 + \xi_1\frac{x_1}{\beta_1}\right)^{\frac{\theta}{\xi_1}} + \lambda_2^{-\theta}\left(1 + \xi_2\frac{x_2}{\beta_2}\right)^{\frac{\theta}{\xi_2}}}{\lambda_1^{-\theta}\left(1 + \xi_1\frac{x_1}{\beta_1}\right)^{\frac{\theta}{\xi_1}} + \lambda_2^{-\theta}}\right]^{-\frac{1}{\theta}-1}.$$

In conclusion, we demonstrated in Example 5.2 that different computations of the conditional distribution $\overline{F}_{2|1}^{\|}$ indeed result in the same representation. In addition, by choosing the Clayton Lévy copula and marginal severity distributions of a comparably simple form,

the conditional distribution $\overline{F}^{\|}_{2|1}(x_2|x_1)$ can be analytically inverted as a function of x_2 for given x_1. This convenient property of the Clayton family admittedly explains its popularity in literature. On the other hand, As equation (4.11) is free of the ordinary survival copula $C_s^{\|}$, it is of particular importance when simulating from a Lévy copula with unknown $C_s^{\|}$, or when the partial derivative of $C_s^{\|}$ is more elaborate to calculate than the partial derivative of $\mathfrak{C}$.

5.2 Assessment of maximum likelihood estimates

In this section we investigate the quality of obtained estimates for a bivariate compound Poisson model via maximising the likelihood function from Theorem 3.9. The analysis relies on simulated sample paths by employing Algorithm 5.1. To begin with, the simulation procedure shall be explained. In order to make our results also relevant to practitioners, we choose three commonly applied severity distributions in the context of operation risk to sample from and thus to produce observations for MLE. With shape parameter $a \in (0,1)$ and scale parameter $b > 0$, the Weibull distribution is heavy-tailed and denoted by $\mathrm{Weib}(a,b)$. Its distribution function has the form

$$F(x) \;=\; 1 - \exp\left[-\left(\frac{x}{b}\right)^a\right], \quad x > 0. \tag{5.6}$$

Another popular subexponential distribution is given by the lognormal distribution with two parameters $\mu \in \mathbb{R}$ and $\sigma > 0$. If Φ denotes the standard normal distribution function, then the lognormal distribution function has the representation

$$F(x) \;=\; \Phi\left(\frac{\ln x - \mu}{\sigma}\right), \quad x > 0,$$

and we utilise the abbreviation $\mathcal{LN}(\mu,\sigma)$. The last selected severity distribution is the GPD and its survival function is already specified in (5.5). The consideration of this distribution is twofold. First, a GPD with shape parameter $\xi > 0$ covers the ordinary Pareto distribution and possesses an even broader parameter space. Second and as already indicated in Section 1.2, the GPD appears as the limiting distribution for loss severities over a high threshold in the POT technique from EVT. Since operational risk measures at high confidence levels are mainly determined by severe loss events, models based on GPD provide a natural choice.

In addition, all parameter values are chosen in such a way that they belong to a realistic range for characterising operational risk losses. The details are summarised in Table 5.1 and Table 5.2.

In view of observation horizons, we consider one year as one time unit, as to be consistent with the typical capital charge calculations based on yearly aggregate loss amounts. As Paragraph 672 of the Basel document [Ban06] prescribes a minimum five-year observation period for internal loss data, we set the shortest interval length in our simulation study to be $T = 5$ as well. Loss data from even shorter periods cannot support reliable estimates

Cell 1	**Cell 2**
$\text{Weib}(a_1 = 0.16, b_1 = 4000)$	$\text{Weib}(a_2 = 0.19, b_2 = 5000)$
$\mathcal{LN}(\mu_1 = 10.3, \sigma_1 = 1.8)$	$\mathcal{LN}(\mu_2 = 9.8, \sigma_2 = 1.4)$
$\text{GPD}(\xi_1 = 1.5, \beta_1 = 6000)$	$\text{GPD}(\xi_2 = 1.3, \beta_2 = 5000)$
$\text{GPD}(\xi_1 = 0.9, \beta_1 = 6000)$	$\text{GPD}(\xi_2 = 0.9, \beta_2 = 6000)$

Table 5.1: Marginal severity distributions from the same distribution class and utilised in the sampling and estimation procedure for the assessment of MLE.

Cell 1	**Cell 2**
$\mathcal{LN}(\mu_1 = 10.3, \sigma_1 = 1.8)$	$\text{Weib}(a_2 = 0.19, b_2 = 5000)$
$\text{Weib}(a_1 = 0.19, b_1 = 5000)$	$\text{GPD}(\xi_2 = 1.3, \beta_2 = 5000)$
$\mathcal{LN}(\mu_1 = 10.3, \sigma_1 = 1.8)$	$\text{GPD}(\xi_2 = 1.3, \beta_2 = 5000)$

Table 5.2: Marginal severity distributions from different distribution classes and utilised in the sampling and estimation procedure for the assessment of MLE.

in general. Moreover, we choose two longer periods of $T = 10$ and $T = 20$ to judge the potential improvements in parameter estimates owing to an increased sample size.

Furthermore, the marginal Poisson parameters are set to be $\lambda_1 = 40$ and $\lambda_2 = 45$, which reflect reasonable frequencies of loss events observed on a yearly basis. As to make the comparison between different Lévy copula families possible to some extent, we fix the expected yearly number of simultaneous losses in the two risk cells at three levels $\lambda^{\|} = 5, 24, 38$. Then we select three different copula families, the Clayton and the Gumbel ones from Table 2.1, as well as the Archimedean copula from Example 3.6 (2). One reason for this choice of copula families is purely technical, as they are all capable to produce the desired value of $\lambda^{\|}$, which is not the general case for arbitrary Lévy copulas as explained after Example 3.6. The second reason is that we have seen in Figures 4.2-4.4 how the three chosen families reveal different severity dependence structures, even when frequency correlation is fixed at the same value. Hence it is worthwhile to explore whether the latter has an impact on the behaviour of MLE as well.

For each of the selected copula families, we calculate the implied copula parameter θ and the results are displayed in Table 5.3.

	$\lambda^{\|} = 5$	$\lambda^{\|} = 24$	$\lambda^{\|} = 38$
Clayton	$\theta_{\text{Clay,low}} = 0.3242$	$\theta_{\text{Clay,med}} = 1.2212$	$\theta_{\text{Clay,high}} = 7.0519$
Gumbel	$\theta_{\text{Gumb,low}} = 0.9317$	$\theta_{\text{Gumb,med}} = 4.3951$	$\theta_{\text{Gumb,high}} = 26.8980$
Example 3.6 (2)	$\theta_{\text{Arch,low}} = 0.0185$	$\theta_{\text{Arch,med}} = 0.0377$	$\theta_{\text{Arch,high}} = 0.1746$

Table 5.3: Lévy copula parameters utilised in the sampling and estimation procedure for the assessment of MLE.

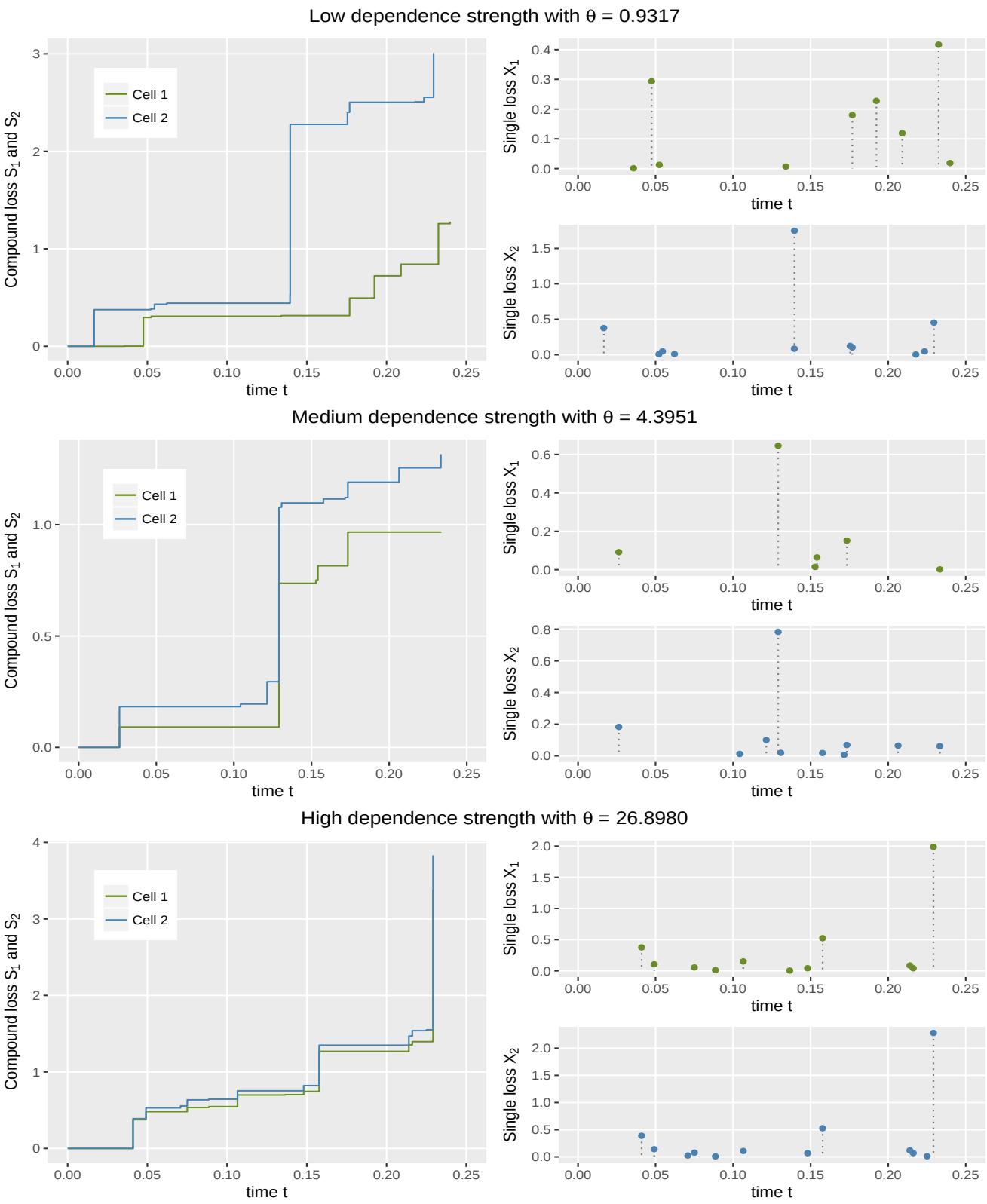


Figure 5.1: Simulation of three bivariate compound Poisson models based on the Gumbel Lévy family with different dependence strength. The left panel shows the sample paths of the compound loss processes, whereas the right panel presents the same paths as marked point processes. The marginal frequency parameters are given by $\lambda_1 = 40$, $\lambda_2 = 45$, and the severity distributions by the GPD with identical parameters $\xi_1 = \xi_2 = 0.9$, $\beta_1 = \beta_2 = 6000$ for both risk cells.

In order to evaluate the performance of MLE, we simulate 100 sample paths of a bivariate compound Poisson model for each combination of Lévy copulas from Table 5.3 and marginal severities from Table 5.1 over $T = 20$ years. The marginal severities from different distribution classes provided by Table 5.2 are only combined with the Clayton Lévy family. Then the marginal frequency and severity parameters as well as the copula parameter are estimated for each sample path by utilising the obtained loss data from the first 5 years, the first 10 years and in the end all 20 years, respectively.

To illustrate the simulation scheme, Figure 5.1 depicts simulated trajectories of three bivariate compound Poisson models based on the Gumbel Lévy family for different dependence strength as offered in Table 5.3. The underlying severity distributions are given by the GPD with identical parameters for both risk cells from Table 5.1. For a clear visualisation, all loss amounts are shown in 10,000 unit and only the time period $[0, 0.25]$ is presented. In the weak dependence case with the smallest value of the copula parameter θ, loss occurrence times of the two risk cells tend to avoid each other and the single loss sizes appear independent. With increasing value of θ, we clearly see more joint loss occurrences of the two risk cells and large losses from different cells tend to accompany each other as well.

The scatter plots in Figure 5.2 compare the single losses simultaneously occurred in the two risk cells and those solely attributed to one risk cell in more detail. The underlying model parameters are the same as for Figure 5.1, with the exception that the observation period is extended to $T = 10$. The findings from Figure 5.1 are confirmed, as more and more joint losses concentrate near the diagonal line with increasing dependence strength. In contrast, the individual losses on the axes decrease in number and reveal a narrower range from the left to the right in Figure 5.2. Moreover, the right subfigure indicates upper

Figure 5.2: Simulation of three bivariate compound Poisson models based on the Gumbel Lévy family with different dependence strength. The single loss sizes up to time horizon $T = 10$ are depicted in logarithmic scale. The marginal frequency parameters are given by $\lambda_1 = 40$, $\lambda_2 = 45$, and the severity distributions by the GPD with identical parameters $\xi_1 = \xi_2 = 0.9$, $\beta_1 = \beta_2 = 6000$ for both risk cells.

tail dependence in the bivariate common losses, as well as all severe large losses are the bivariate ones in the upper right corner rather than the univariate ones on the axes.

In order to evaluate the goodness of the obtained estimators, different performance measures are calculated. For space reasons, we relocate the detailed results in tabular format into Appendix C and summarise at this point the main findings. The MLE procedure achieves equally good results, independent of the particular combination of Lévy copula families and marginal distributions. The mean relative bias based on 20-year observations rarely exceeds 1% across all model constellations. Moreover, no systematic under- or overestimation of any parameter is identified, as the estimated biases appear evenly positively and negatively spread out near zero.

Concerning the parameter θ of the Lévy copula being the main interest in dependence modelling, the box plot in Figure 5.3 illustrates estimated relative biases of this essential parameter for all considered copula families. The true parameter values can be read off from Table 5.3 and the estimators are calculated based on losses observed up to the shortest time horizon of $T = 5$. We see that this minimum observation period prescribed in the Basel framework is already sufficient to obtain estimators of reasonable quality.

Moreover, all performance measures exhibit satisfactory improvements with extended observation horizons. For illustration, Figure 5.4 demonstrates the estimation results based on the Clayton Lévy copula with lognormal severity in risk cell one and GPD severity in risk cell two. With increasing number of observation years, the height of the boxes for all model parameters narrows significantly. As the importance of operational risk has been seriously recognised since its incorporation into the Basel regulations, reliable loss data from a period of approximately 10 years are already common practice among financial institutions. Hence we expect our simulation scenarios to reflect realistic performance of MLE when applied in practice. Although we only visualise the relative bias, other utilised performance measures provide similar interpretations and all details are reported in Appendix C.

5.3 Approaches for dependence model examination

In contrast to the univariate setting or dependence modelling via ordinary copulas, no standard procedures for goodness of fit examinations exist for dependence models based on Lévy copulas. For this reason, in the current section we investigate the ability of potential diagnostic tools, mainly by means of graphical inspections, in order to assess the fit of a bivariate compound Poisson model. Once again we consider all possible combinations of marginal severity distributions and Lévy copula families as detailed in Table 5.1 and Table 5.3, respectively. Note all parameter values are selected in such a way that they belong to reasonable scenarios for operational risk in practice. After having convinced ourselves of the reliable behaviour of MLE, now we sample only one path from each feasible bivariate model and explore below several visualisation techniques applied to these observations.

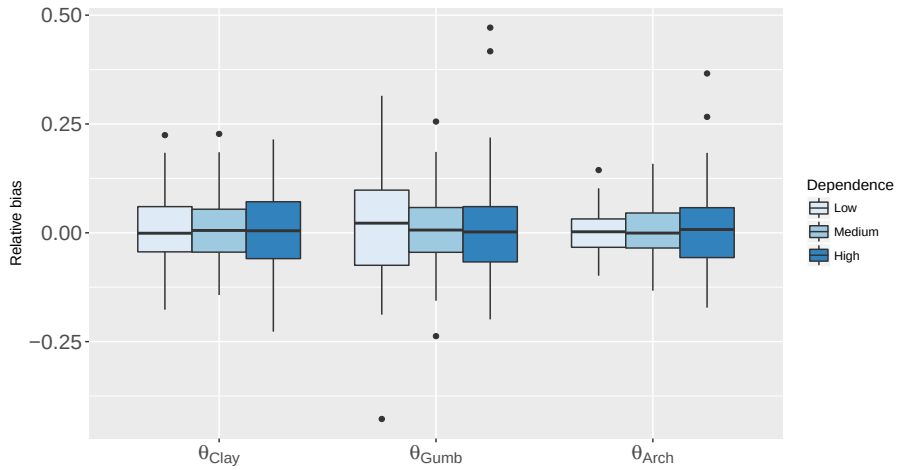


Figure 5.3: Comparison regarding relative bias of estimated copula parameters for three Lévy copula families. The MLE is performed based on marginal frequency intensities of $\lambda_1 = 40$, $\lambda_2 = 45$, and marginal GPD severities with identical parameters $\xi_1 = \xi_2 = 0.9$, $\beta_1 = \beta_2 = 6000$ for both risk cells. The observation time horizon is given by $T = 5$ and the true copula parameters can be read off from Table 5.3.

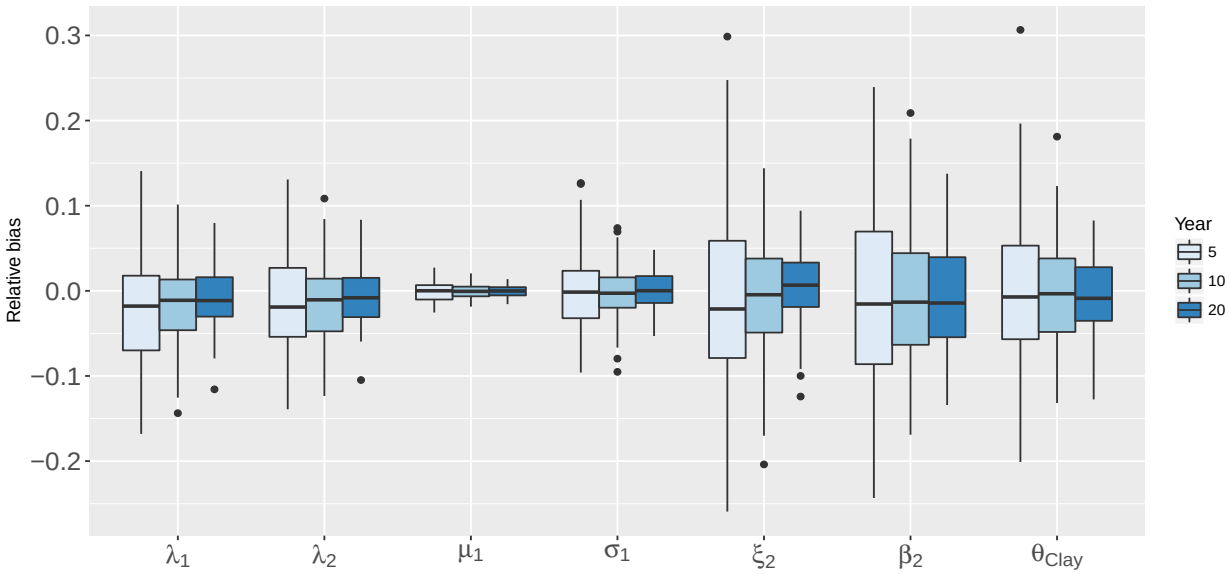


Figure 5.4: Relative bias of parameter estimates for a bivariate Clayton dependence model based on simulated observations of up to three different time horizons. The true marginal frequency intensities are given by $\lambda_1 = 40$, $\lambda_2 = 45$, the severity distributions by $\mathcal{LN}(\mu_1 = 10.3, \sigma_1 = 1.8)$, $\text{GPD}(\xi_2 = 1.3, \beta_2 = 5000)$, and the Lévy copula parameter by $\theta_{\text{Clay,high}} = 7.0519$.

Inspired by the important role of empirical distribution functions in standard goodness of fit techniques, we introduce hereafter the counterpart for tail integrals. Recall from Definition 2.5 and Definition 2.11 that the tail integral associated with a Lévy measure has the intuitive interpretation as the expected number of losses per unit time which exceed a certain loss amount in each risk cell. Therefore, the tail integral $\overline{\Pi}_1^{\perp}$ associated with the loss process $S_1^{\perp}(t)$ solely attributed to the first risk cell has the characterisation

$$\overline{\Pi}_1^{\perp}(x_1) \;=\; \mathbb{E}\left[\,\#\left\{(\Delta S_1(t),\,0) \mid t \in [0,1] \,\wedge\, \Delta S_1(t) \geq x_1\right\}\right], \quad x_1 \geq 0,$$

and it can be naturally approximated by

$$\hat{\overline{\Pi}}_{T,1}^{\perp}(x_1) \;=\; \frac{1}{T}\sum_{j=1}^{N_1^{\perp}(T)} \mathbb{1}_{\{X_{1j}^{\perp} \geq x_1\}} \;=\; \frac{1}{T}\sum_{h=1}^{N(T)} \mathbb{1}_{\{Y_{1h} \geq x_1,\, Y_{2h}=0\}},$$

where we follow the notations from Chapter 3. The stationary and independent properties of the increments of $S_1^{\perp}(t)$ as a Lévy process ensure that the same loss behaviour is to be expected over each unit time interval up to the observation horizon T, hence the standardisation by dividing through T is justified. The tail integral $\overline{\Pi}_2^{\perp}$ describing the individual losses of risk cell two is defined and estimated in an analogous manner. Finally, the bivariate joint tail integral $\overline{\Pi}^{\parallel}$ is estimated for arbitrary $(x_1, x_2)^{\top} \in (0,\infty)^2$ as

$$\hat{\overline{\Pi}}_{T}^{\parallel}(x_1, x_2) \;=\; \frac{1}{T}\sum_{k=1}^{N^{\parallel}(T)} \mathbb{1}_{\{X_{1k}^{\parallel} \geq x_1,\, X_{2k}^{\parallel} \geq x_2\}} \;=\; \frac{1}{T}\sum_{h=1}^{N(T)} \mathbb{1}_{\{Y_{1h} \geq x_1,\, Y_{2h} \geq x_2\}},$$

and its first marginal component $\overline{\Pi}_1^{\parallel}$ can be approximated through

$$\hat{\overline{\Pi}}_{T,1}^{\parallel}(x_1) \;=\; \frac{1}{T}\sum_{k=1}^{N^{\parallel}(T)} \mathbb{1}_{\{X_{1k}^{\parallel} \geq x_1\}} \;=\; \frac{1}{T}\sum_{h=1}^{N(T)} \mathbb{1}_{\{Y_{1h} \geq x_1,\, Y_{2h} > 0\}}.$$

The second marginal component $\overline{\Pi}_2^{\parallel}$ is treated similarly.

Note the tail integrals $\overline{\Pi}_i^{\perp}$ and $\overline{\Pi}_i^{\parallel}$, $i \in \{1,2\}$, comprise information about both the marginal processes and the dependence structure, as this was manifested in Section 3.2.2 where they were expressed in terms of the marginal Poisson frequencies, the marginal severity distributions and the Lévy copula. Therefore, we believe the comparison between the theoretical and empirical versions of tail integrals could contribute to the quality assessment of a fitted dependence model.

Figures 5.5-5.7 illustrate the comparison of empirical and theoretical tail integrals, whereby the underlying bivariate compound Poisson model is based on the Clayton Lévy family with different dependence strength as detailed in Table 5.3, and the marginal Weibull severities with parameters from Table 5.1. Although the marginal Poisson intensities are fixed at $\lambda_1 = 40$ and $\lambda_2 = 45$ across all three depicted models, the Poisson intensity of the common loss process $S^{\parallel} = (S_1^{\parallel}, S_2^{\parallel})^{\top}$ is given by $\lambda^{\parallel} = 5$ in Figure 5.5, by $\lambda^{\parallel} = 24$ in Figure 5.6 and by $\lambda^{\parallel} = 38$ in Figure 5.7. Recall from Section 3.2.2 that $\lambda^{\parallel}$

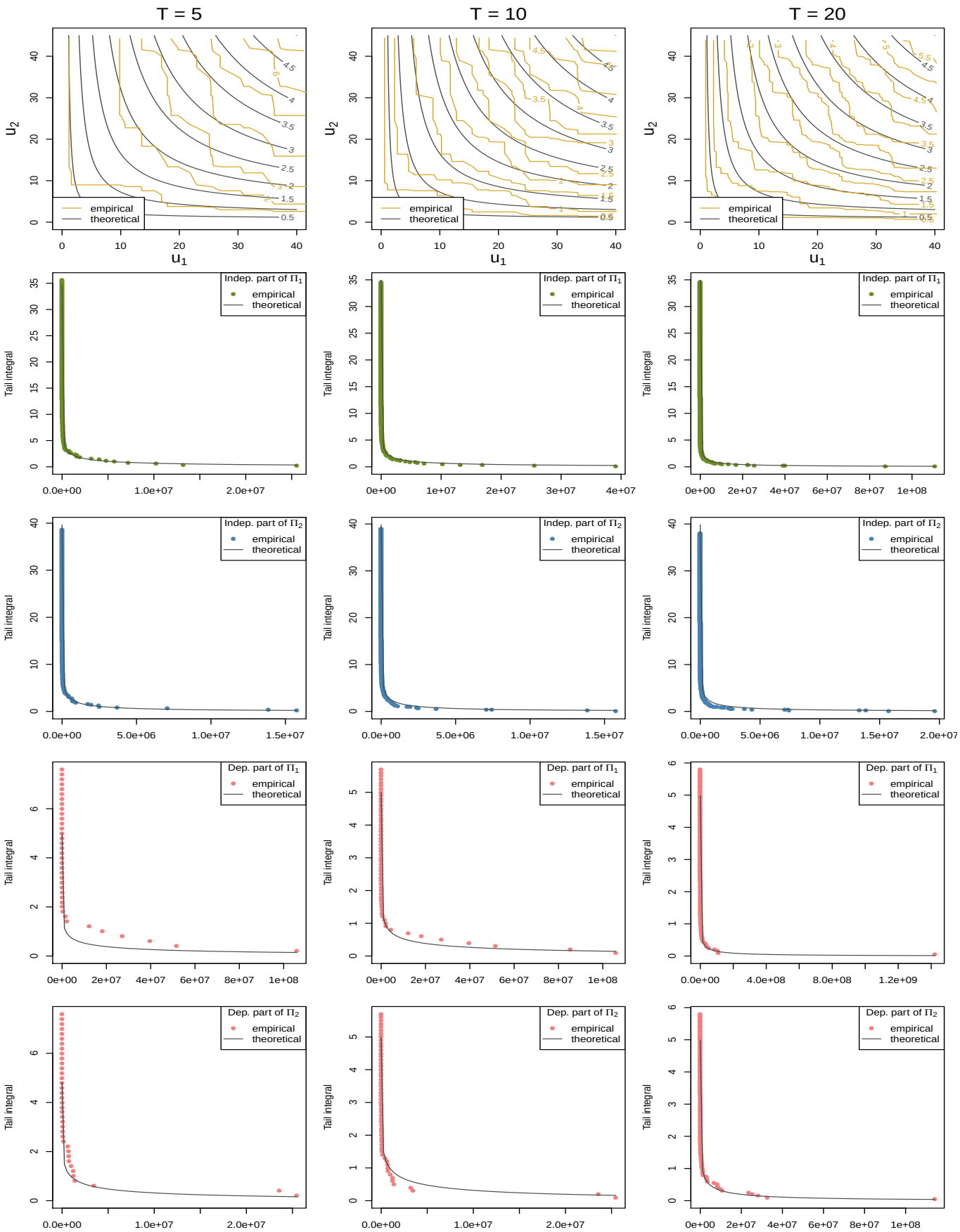


Figure 5.5: First row: contour lines of the sample Lévy copula compared with the theoretical Clayton one with low dependence strength characterised by $\theta_{\text{Clay,low}} = 0.3242$. Second to last rows: empirical tail integrals $\hat{\overline{\Pi}}_1^{\perp}$, $\hat{\overline{\Pi}}_2^{\perp}$, $\hat{\overline{\Pi}}_1^{\|}$ and $\hat{\overline{\Pi}}_2^{\|}$ compared with their theoretical counterparts.

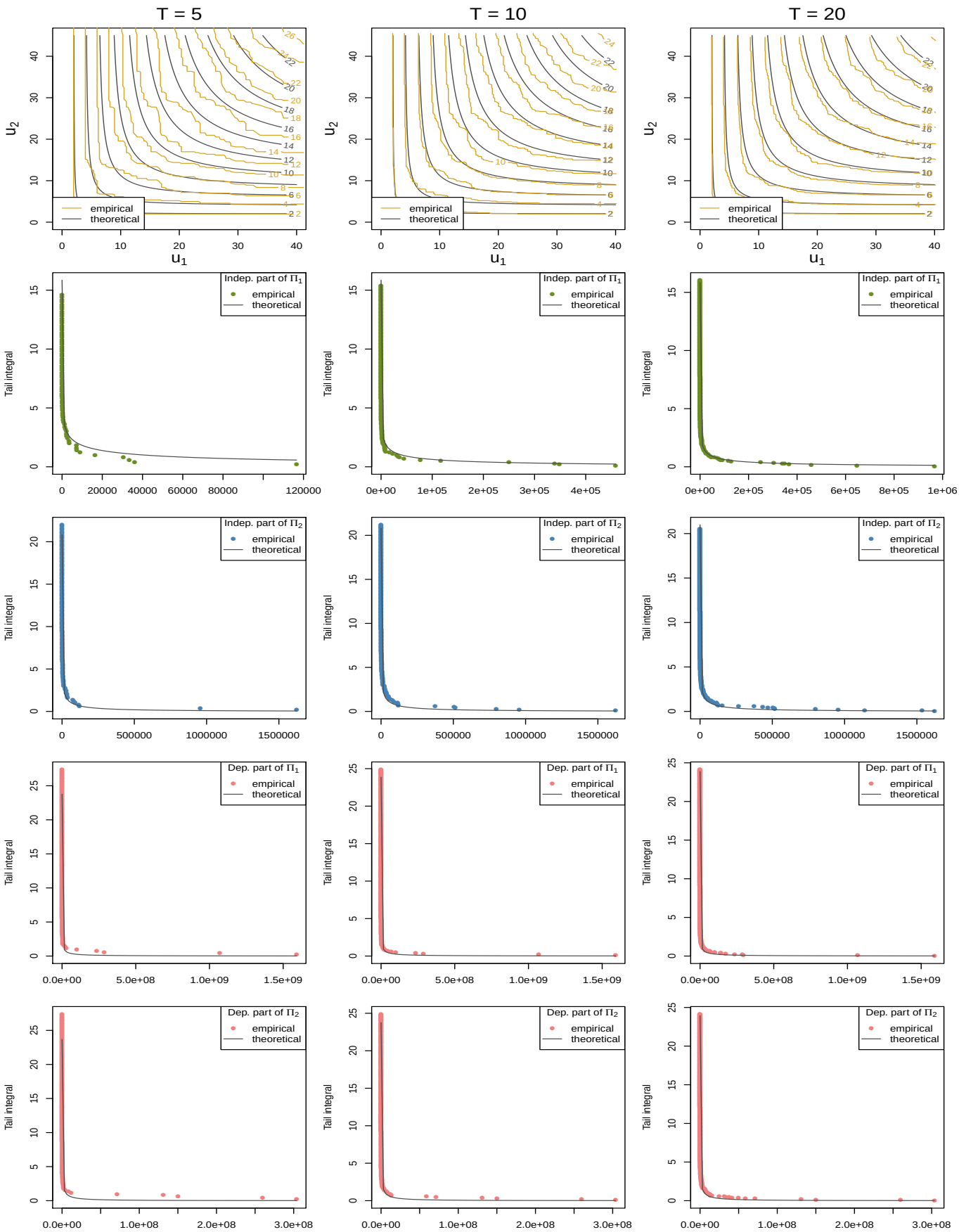


Figure 5.6: First row: contour lines of the sample Lévy copula compared with the theoretical Clayton one with medium dependence strength characterised by $\theta_{\text{Clay,med}} = 1.2212$. Second to last rows: empirical tail integrals $\hat{\overline{\Pi}}_1^{\perp}$, $\hat{\overline{\Pi}}_2^{\perp}$, $\hat{\overline{\Pi}}_1^{\parallel}$ and $\hat{\overline{\Pi}}_2^{\parallel}$ compared with their theoretical counterparts.

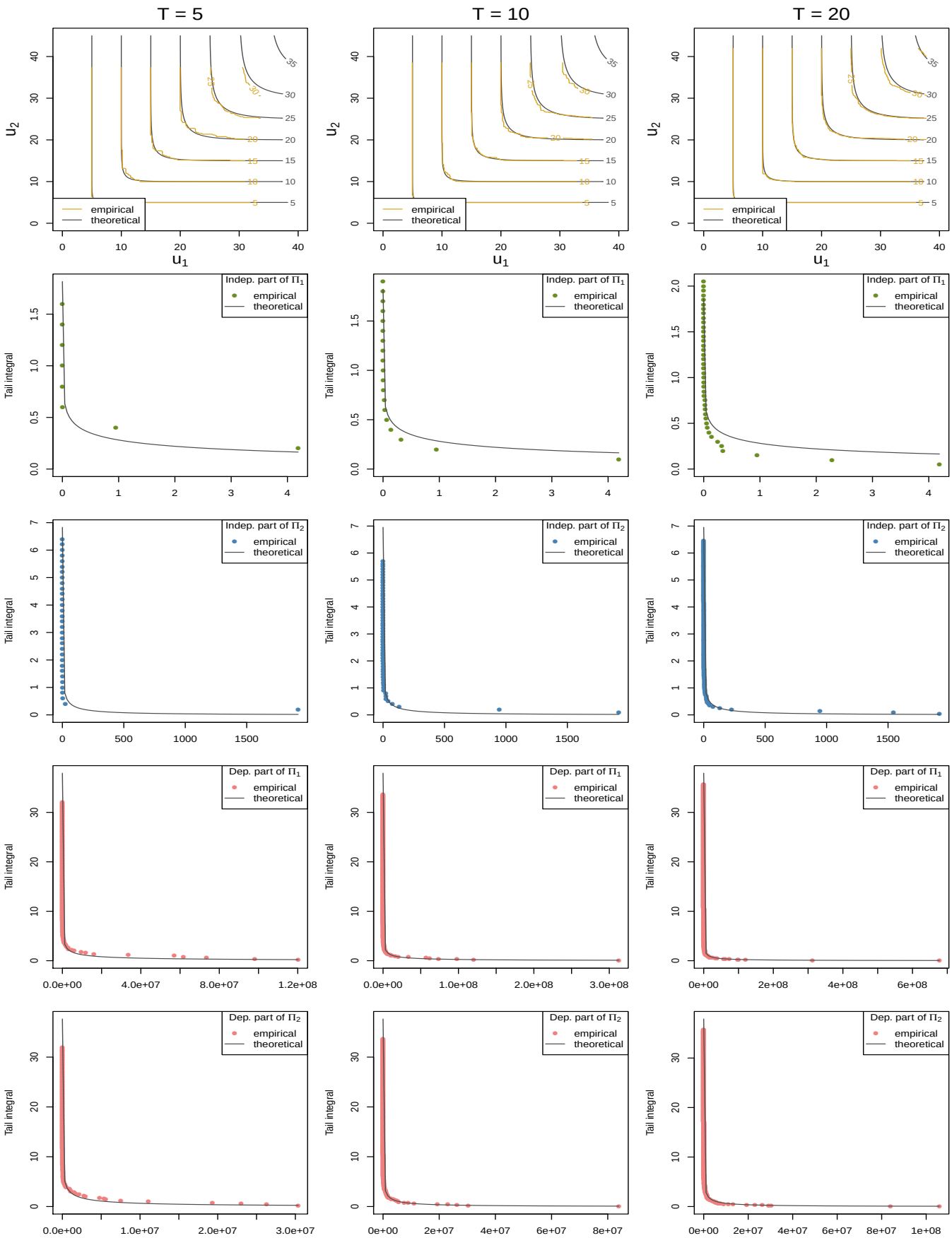


Figure 5.7: First row: contour lines of the sample Lévy copula compared with the theoretical Clayton one with high dependence strength characterised by $\theta_{\text{Clay,high}} = 7.0519$. Second to last rows: empirical tail integrals $\hat{\overline{\Pi}}_1^{\perp}$, $\hat{\overline{\Pi}}_2^{\perp}$, $\hat{\overline{\Pi}}_1^{\|}$ and $\hat{\overline{\Pi}}_2^{\|}$ compared with their theoretical counterparts.

is equal to the entire mass of the Lévy measure $\Pi^{\|}$ and precisely expresses the average number of the bivariate loss severities $X^{\|} = (X_1^{\|}, X_2^{\|})^\top$ in a unit time interval. Moreover, Figure 4.1 has already indicated that $\lambda^{\|}$ is a monotone function of the copula parameter θ under the current Clayton family. Hence the percentage expected number of common losses related to the different θ used are roughly 6%, 39%, and 81%.

As expected, all three Figures 5.5-5.7 have in common that the accordance between the empirical and theoretical integrals improves significantly with increasing observation horizon T, as more data points become available. Note that the observations up to time $T = 20$ include those from the shorter periods, as to enable a consistent comparison. However, the empirical version of the tail integrals $\overline{\Pi}_i^{\|}$, $i \in \{1, 2\}$, attributed to the common loss process does not provide satisfactory agreement with their theoretical counterparts for all considered time horizons in the last two lines of Figure 5.5. This is due to the fact that the average number of simultaneous losses in the weak dependence setting is insufficient to draw conclusions about the model fit based on the corresponding empirical tail integrals. Similar observations and interpretations can be made for the poor fit of the individual tail integral $\overline{\Pi}_1^{\perp}$ in the high dependence case as displayed in the second line of Figure 5.7.

Beyond that, in the top panel of Figures 5.5-5.7 we compare contour plots of empirical and theoretical Lévy copulas. The empirical version of a Lévy copula up to observation horizon T is naturally defined as the bivariate function satisfying

$$\hat{\overline{\Pi}}_T(x_1, x_2) \;=\; \hat{\mathfrak{C}}_T(\hat{\overline{\Pi}}_{T,1}(x_1), \hat{\overline{\Pi}}_{T,2}(x_2)),$$

where the marginal tail integrals are estimated similarly to above via

$$\hat{\overline{\Pi}}_{T,i}(x_i) \;=\; \frac{1}{T}\sum_{j=1}^{N_i^{\perp}(T)} \mathbb{1}_{\{X_{ij}^{\perp} \geq x_i\}} + \frac{1}{T}\sum_{k=1}^{N^{\|}(T)} \mathbb{1}_{\{X_{ik}^{\|} \geq x_i\}} \;=\; \frac{1}{T}\sum_{h=1}^{N(T)} \mathbb{1}_{\{Y_{ih} \geq x_i\}}, \quad i \in \{1, 2\}.$$

Since the highest average number of bivariate loss severities is offered under the strong dependence induced by the largest Clayton Lévy copula parameter θ, the best accordance between the theoretical and empirical contour lines are offered by Figure 5.7. On the

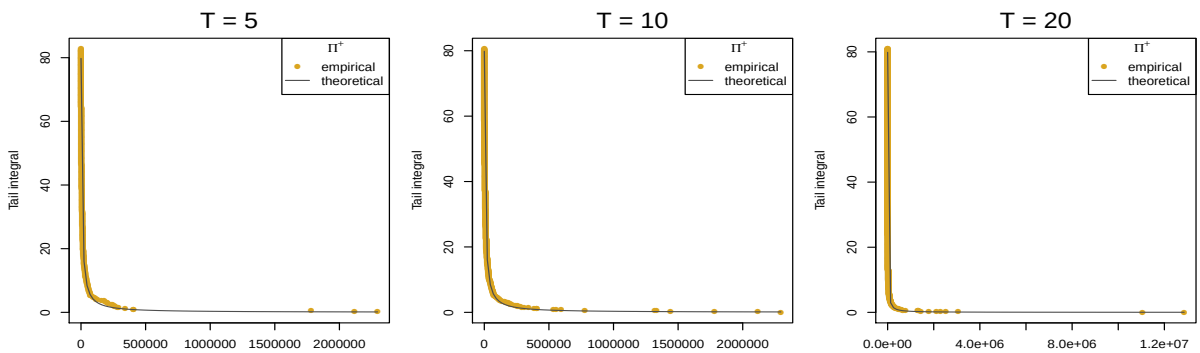


Figure 5.8: Comparison of the empirical tail integral $\hat{\overline{\Pi}}_+$ associated with the overall loss process S_+ with its theoretical counterpart up to three different observation horizons. The underlying Lévy copula is from Example 3.6 (2) with $\theta_{\text{Arch,low}} = 0.0185$.

contrary, under the weak dependence as depicted in Figure 5.5, the empirical contour lines exhibit significant deviations from the theoretical ones even up to the remotest time horizon. The kinks in the empirical copulas indicate that too few observations are present to properly assess the quality of the fitted dependence model.

At this point it should be mentioned that similar behaviours of the empirical copula contours and tail integrals as shown in Figures 5.5-5.7 have been observed during our simulation study across different model constellations, although we only picture the Clayton Lévy copula with Weibull marginal severities for illustration purpose. To conclude, the visualisation techniques based on empirical copulas and tail integrals cannot be fully trusted when assessing the fit of a dependence model, as their accordance with the theoretical counterparts appears unsatisfactorily even though the observations are sampled from the model with the true parameter values.

Since the consideration of the individual tail integrals $\overline{\Pi}_i^{\perp}$ and $\overline{\Pi}_i^{\parallel}$, $i \in \{1,2\}$, provides inconclusive results, we introduce one more visualisation option related to the tail integral $\overline{\Pi}_+$ describing the overall loss process $S_+ = S_1 + S_2$. Its theoretical representation is stated in equation (4.6) and the empirical equivalence can be naturally written as

$$\begin{aligned}\hat{\overline{\Pi}}_{T,+}(x) &= \frac{1}{T}\sum_{j=1}^{N_1^{\perp}(T)} \mathbb{1}_{\{X_{1j}^{\perp}\geq x\}} + \frac{1}{T}\sum_{l=1}^{N_2^{\perp}(T)} \mathbb{1}_{\{X_{2l}^{\perp}\geq x\}} + \frac{1}{T}\sum_{k=1}^{N^{\parallel}(T)} \mathbb{1}_{\{X_{1k}^{\parallel}+X_{2k}^{\parallel}\geq x\}} \\ &= \frac{1}{T}\sum_{h=1}^{N(T)} \mathbb{1}_{\{Y_{1h}+Y_{2h}\geq x\}}, \quad x \geq 0.\end{aligned}$$

Figure 5.8 illustrates the fit of the theoretical tail integral $\overline{\Pi}_+$ to the empirical estimates. The underlying Lévy copula is chosen from Example 3.6 (2) with the parameter indicating weak dependence from Table 5.3. In contrast to Figures 5.5-5.7, the tail integrals $\overline{\Pi}_+$ and $\hat{\overline{\Pi}}_{T,+}$ show satisfying agreement with each other for all considered time horizons. This is mainly due to the fact that $\overline{\Pi}_+$ comprises information in both the individual and simultaneous loss part, hence enables robuster estimates compared to the individual tail integrals $\overline{\Pi}_i^{\perp}$ and $\overline{\Pi}_i^{\parallel}$ with $i \in \{1,2\}$.

As the Lévy copula contributes as the most interesting part in a dependence model rather than the marginal distributions, we consider one more potential approach to evaluating the model fit featuring the empirical Levy copula $\hat{\mathfrak{C}}_T$ based on observations up to the time horizon T. It is established in [BV13] that with T approaching infinity, the difference between $\hat{\mathfrak{C}}_T$ and the true Lévy copula $\mathfrak{C}$ is asymptotically normally distributed. More precisely, the scaled difference $\sqrt{T}(\hat{\mathfrak{C}}_T(x_1,x_2) - \mathfrak{C}(x_1,x_2))$ converges for fixed $(x_1,x_2)^\top$ to a centred normal random variable whose variance is given by the true copula value $\mathfrak{C}(x_1,x_2)$. However, an observation period of near infinite length can never be encountered in practice. Hence it is worthwhile to examine whether the above asymptotic result already applies to our selected finite and realistic time horizons $T = 5, 10, 20$.

For this purpose, we choose four different vectors $(x_1,x_2)^\top$ and evaluate the Clayton Lévy family implying low, medium and high dependence strength at these points. The obtained values shall reflect the asymptotic variance of the estimates and are summarised

$(x_1, x_2)^\top \rightarrow$		$(4,4)^\top$		$(7,7)^\top$		$(3,5)^\top$		$(4,7)^\top$	
	T	$\widehat{\text{Mean}}$	$\widehat{\text{Var}}$	$\widehat{\text{Mean}}$	$\widehat{\text{Var}}$	$\widehat{\text{Mean}}$	$\widehat{\text{Var}}$	$\widehat{\text{Mean}}$	$\widehat{\text{Var}}$
$\theta_{\text{Clay,low}} = 0.3242$	5	0.0123	0.5140	-0.0418	0.8980	0.0073	0.4841	-0.0308	0.6525
	10	-0.0307	0.4513	-0.0787	0.8182	-0.0282	0.4406	-0.0524	0.6076
	20	-0.0126	0.4967	-0.0657	0.8800	-0.0154	0.4818	-0.0275	0.6433
$\theta_{\text{Clay,med}} = 1.2212$	5	0.0503	2.4272	0.0354	4.4901	0.0181	2.1789	0.0281	3.0617
	10	0.0900	2.1549	0.0039	4.4181	0.0901	2.0340	0.0562	2.8621
	20	-0.0032	2.1805	-0.0710	4.1218	0.0142	2.0170	-0.0318	2.8573
$\theta_{\text{Clay,high}} = 7.0519$	5	0.0127	3.7603	0.0960	6.1495	0.0076	3.3063	0.0466	4.2187
	10	0.0072	3.9051	0.1249	6.6045	-0.0246	3.2095	0.0577	4.3848
	20	0.0039	3.3818	0.0882	6.2391	-0.0187	2.8765	0.0315	3.7646

Table 5.4: Sample mean and variance of the standardised estimates $\sqrt{T}(\hat{\mathfrak{C}}_T(x_1, x_2) - \mathfrak{C}(x_1, x_2))$ for the Clayton Lévy copula with different dependence strengths and evaluated up to different time horizons T.

$(x_1, x_2)^\top \rightarrow$	$(4,4)^\top$	$(7,7)^\top$	$(3,5)^\top$	$(4,7)^\top$
$\theta_{\text{Clay,low}} = 0.3242$	0.4717	0.8255	0.4519	0.6162
$\theta_{\text{Clay,med}} = 1.2212$	2.2675	3.9682	2.1111	2.8622
$\theta_{\text{Clay,high}} = 7.0519$	3.6255	6.3447	2.9886	3.9892

Table 5.5: Clayton Lévy copula of three dependence strengths evaluated at four different points.

in Table 5.5. Then for each considered Lévy copula, we sample 500 times from the corresponding compound Poisson model and calculate the empirical copula values up to different time horizons. Table 5.4 reports the mean and variance of the standardised estimates $\sqrt{T}(\hat{\mathfrak{C}}_T(x_1, x_2) - \mathfrak{C}(x_1, x_2))$. Generally, we observe that the theoretical variances are well reproduced for all combinations of various dependence levels, selected points and time horizons. The sample means offer small absolute values around zero, indifferent to the length of the observation period. On the other hand, the sample variances based on 10-year data approach the theoretical asymptotic variances visibly better than only relying on 5-year data. However, no significant improvement is recognised when further extending the time horizon to 20 years.

In order to judge the performance of the normal approximation apart from the bias and variance, Figure 5.9 depicts quantile-quantile plots (Q-Q plots) for the case of strong dependence. The fine agreement between the empirical and theoretical quantiles confirms that the finite sample properties are indeed satisfying. From a statistical point of view, our small simulation study reveals the potential of constructing goodness of fit tests for Lévy copulas based on the asymptotic normal results. Note that also the covariance structure on the range of the copula must be taken into consideration when constructing such hypothesis tests.

Last but not least, if the ordinary copula $C^{\|}$ associated with the bivariate loss severities $(X_1^{\|}, X_2^{\|})^\top$ is known in closed form, as this is satisfied by the Archimedean Lévy family, and a sufficiently large sample size of bivariate loss observations is available, then goodness of fit tests and model selection criteria for ordinary copulas can also be used to assess the

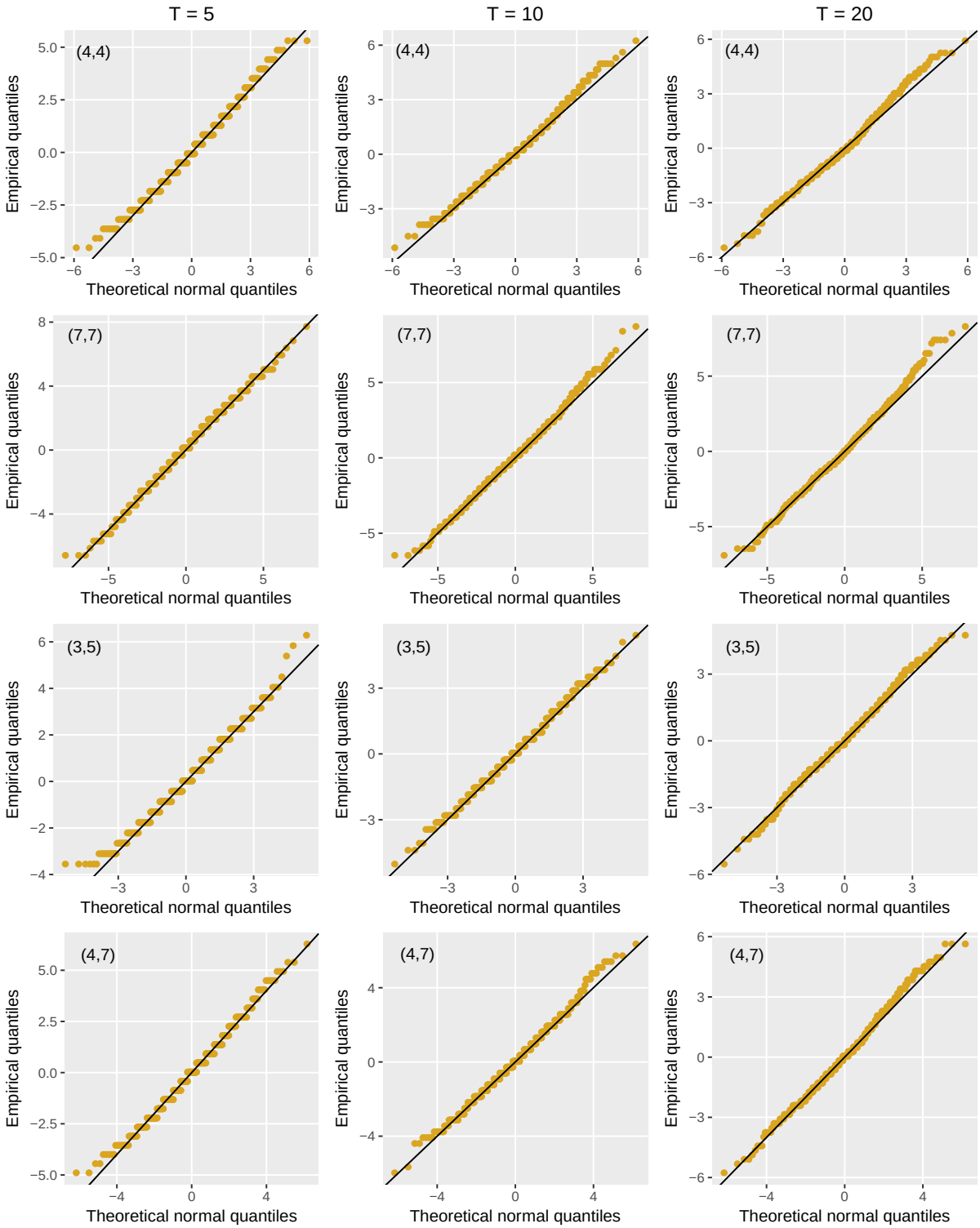


Figure 5.9: Q-Q plots of the standardised estimates $\sqrt{T}(\hat{\mathfrak{C}}_T(x_1, x_2) - \mathfrak{C}(x_1, x_2))$ with respect to asymptotic theoretical normal distributions. The underlying Lévy copula is given by the Clayton family with parameter $\theta_{\text{Clay,high}} = 7.0519$ and the evaluation point $(x_1, x_2)^\top$ is shown as label in the top left corner in each subfigure.

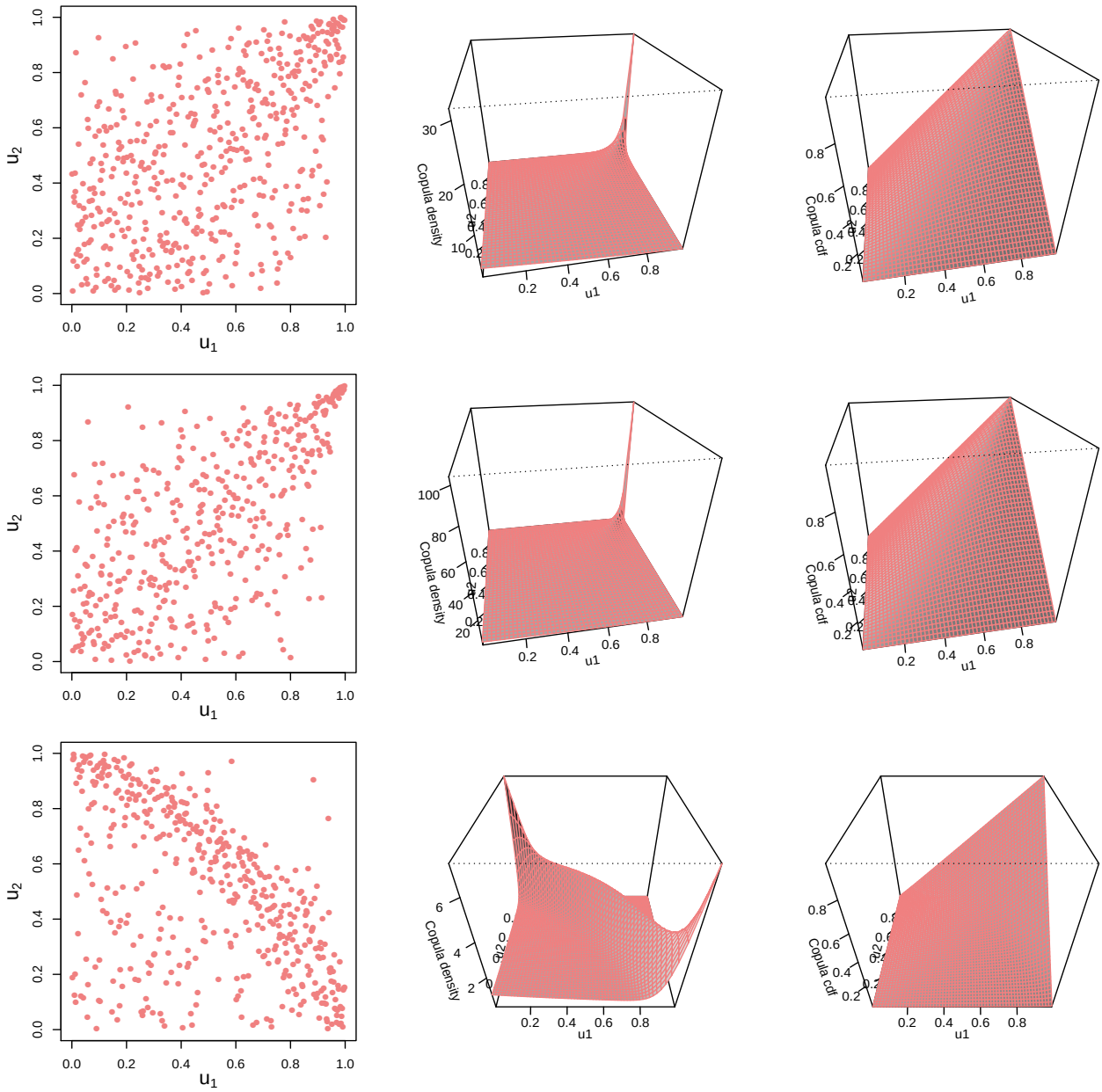


Figure 5.10: Illustration of the ordinary copula $C^{\|}$ associated with the bivariate common loss severities $(X_1^{\|}, X_2^{\|})^\top$ induced by the Clayton (first row), the Gumbel (middle row) and the Archimedean Lévy family from Example 3.6 (2) (last row). The underlying Lévy copula parameters are given by $\theta_{\text{Clay,med}} = 1.2212$, $\theta_{\text{Gumb,med}} = 4.3951$ and $\theta_{\text{Arch,med}} = 0.0377$, respectively. The left column shows bivariate losses $(X_1^{\|}, X_2^{\|})^\top$ in copula scale, the middle column presents the density of $C^{\|}$ and the right column the copula function itself.

fit of dependence structures. By way of illustration, a scatter plot of bivariate losses, a copula density plot and a copula distribution plot are shown in Figure 5.10 for all three selected Lévy copula families in the medium dependence case. The copula density and distribution functions are derived from expression (3.25) for the corresponding survival copula $C_s^{\|}$.

As the bivariate Archimedean Lévy family is symmetric in its two arguments, all induced ordinary copulas $C^{\|}$ in Figure 5.10 are symmetric with respect to the positive diagonal as well. Furthermore, recall Example 5.2 states that the survival copula $C_s^{\|}$ implied by the Clayton Lévy copula is precisely its ordinary Clayton counterpart. This explains the obvious upper tail dependence indicated by both the scatter and the density plot in the top panel. In addition, the ordinary copula $C^{\|}$ induced by the Gumbel Lévy copula shows a similar positive dependence trend as well as upper tail dependence to the Clayton one. On the contrary, the copula $C^{\|}$ underlying the Lévy family from Example 3.6 (2) shows countermonotonic behaviour among the simultaneous losses of both risk cells.

In conclusion, our observations on the implied ordinary copulas $C^{\|}$ underline the fact that Lévy copulas simultaneously model the interdependence among frequencies and severities. Despite the identical frequency correlation behind all illustrated Lévy copulas in Figure 5.10, the bivariate loss severities reveal fairly different dependence patterns.

5.4 Sensitivity of operational risk measures to model components

Whereas the asymptotic sensitivity of operational VaR, ES and SRM to the different components of a compound Poisson model can be directly deduced via the analytic approximations from Section 4.1.2 in the special cases of independence, complete dependence and one-cell dominance, the more general situation has to be studied by means of simulation. Hence in the current section we introduce several bivariate model constellations for which no asymptotic results exist and explore the impact of varying dependence strength as well as marginal distributions on the value of operational risk measures.

For reasons of comparability across difference models and of computational effort, we fix the marginal frequency parameters at $\lambda_1 = \lambda_2 = 10$. This choice of identical marginal Poisson intensities also allows for incorporating the completely positive dependence as detailed before Theorem 4.6. Furthermore, all estimated risk measure values and the corresponding confidence intervals are based on two million Monte Carlo iterations, where we employ Algorithm 5.1 for obtaining a sample path from a bivariate compound Poisson model. Nonetheless, some closed-form asymptotic approximations provided by Section 4.1.2 are also taken into consideration when appropriate and serve as benchmark outcomes.

The first model family we consider is built upon heavy-tailed Weibull distributions whose distribution function is specified in equation (5.6). The corresponding parameter values are given by $X_1 \sim \text{Weib}(a_1 = 0.5, b_1 = 2)$ and $X_2 \sim \text{Weib}(a_2 = 0.5, b_2 = 1)$. Although the

limit $\lim_{x\to\infty} \frac{\overline{F}_{X_2}(x)}{\overline{F}_{X_1}(x)} = 0$ holds, that is, the far out right distribution tail of the second risk cell is dominated by the first cell, we cannot apply the analytical results in Theorem 4.7. This is due to the fact that a Weibull distribution tail is not regularly varying.

In order to examine the sensitivity of risk measures with respect to the dependence structure, we choose the Clayton Lévy family as specified in Table 2.1 and plot estimated VaR as well ES values as a function of the dependence parameter θ_{Clay} in Figure 5.11. Recall that the Clayton family approaches the independence Lévy copula as $\theta_{\text{Clay}} \downarrow 0$ and the complete dependence Lévy copula as $\theta_{\text{Clay}} \uparrow \infty$, whereas the right limit $\theta_{\text{Clay}} = 250$ in Figure 5.11 can be considered as fairly close to the complete dependence situation. First, we observe that the overall risk measures VaR_+ and ES_+ exhibit a similar development

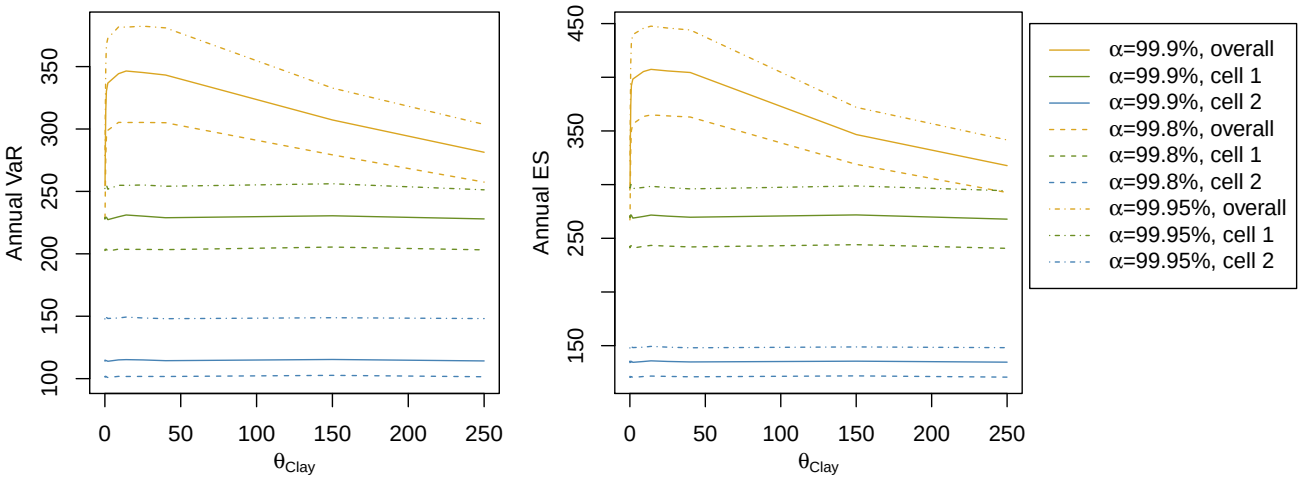


Figure 5.11: Sensitivity of risk measure estimates to the Clayton Lévy copula parameter at three different confidence levels α. The underlying marginal severity distributions are given by $\text{Weib}(a_1 = 0.5, b_1 = 2)$ and $\text{Weib}(a_2 = 0.5, b_2 = 1)$.

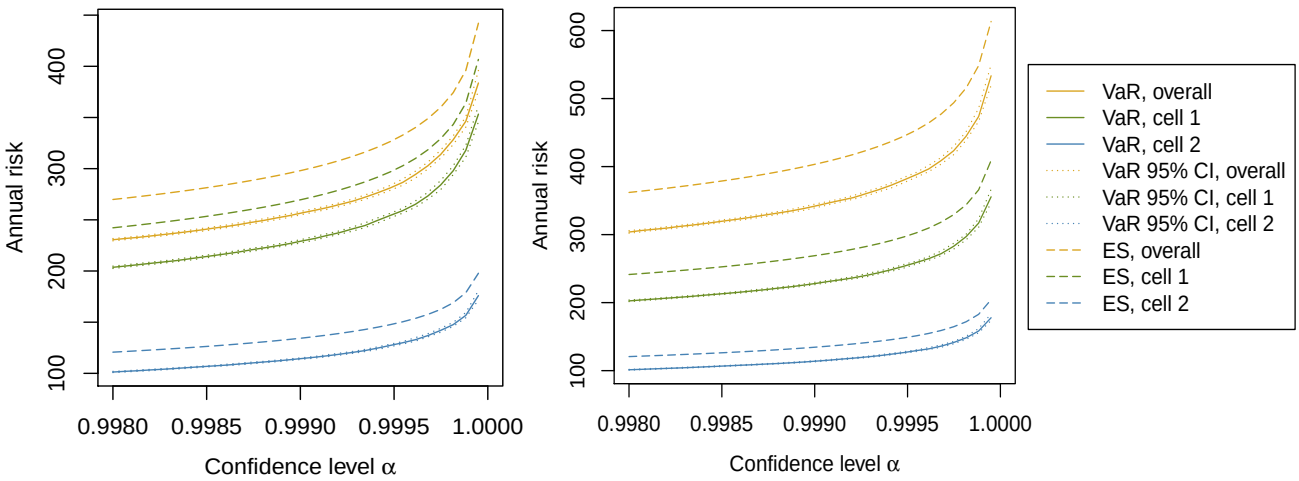


Figure 5.12: Risk measure estimates induced by the Gumbel Lévy family for two different dependence strengths: $\theta_{\text{Gumb}} = 0.13$ (left) and $\theta_{\text{Gumb}} = 17.09$ (right). The underlying marginal severity distributions are given by $\text{Weib}(a_1 = 0.5, b_1 = 2)$ and $\text{Weib}(a_2 = 0.5, b_2 = 1)$.

with increasing θ_{Clay} across all three depicted confidence levels α. Moreover, both VaR_+ and ES_+ have smaller values in the independence case than in the approximately complete dependence case. However, the latter does not constitute an upper bound for the overall risk exposure, as all yellow curves rise sharply up to a peak at roughly $\theta_{\text{Clay}} = 15$, then fall smoothly down towards the complete dependence case.

On the other hand, the stand-alone risk measure estimates are independent of the dependence structure as expected. All green and blue lines in Figure 5.11 stay constant with respect to the copula parameter θ_{Clay} and indicate the correctness of our simulation procedure. Furthermore, the annual $\text{VaR}_{1,1}(\alpha)$ of the first cell is twice as high as the annual $\text{VaR}_{2,1}(\alpha)$ of the second cell for each of the three confidence levels α. This coincides with the asymptotic result according to Theorem 4.4, as the risk measures can be approximated by $\text{VaR}_{1,1}(\alpha) \sim 2[\ln(0.1 - 0.1\alpha)]^2$ and $\text{VaR}_{2,1}(\alpha) \sim [\ln(0.1 - 0.1\alpha)]^2$, respectively.

Since financial institutions often consider not only $\alpha = 99.9\%$ for regulatory requirements, but also a broader range of quantile levels for internal business steering, Figure 5.12 visualises the sensitivity of VaR and ES with respect to α. Whereas the risk exposure exhibits a slow increase up to around $\alpha = 99.9\%$ and then exponentially increases in the left panel for weak dependence, the steep increase occurs at a higher level of roughly $\alpha = 99.95\%$ in the right panel for stronger dependence. Furthermore, the 95% confidence intervals (CIs) for VaR are quite narrow in both figures, suggesting our sample size of two million is sufficiently large.

Having examined a model example with rapidly varying marginal severities, now we move on to the regularly varying case. More precisely, we assume for both risk cells a standardised GPD with zero location and unit scale parameters $\beta_1 = \beta_2 = 1$. However, the shape parameters are set differently to $\xi_1 = 1$ and $\xi_2 = 0.8$. The core idea behind this severity construction is that the key component of a GPD is its shape parameter, which determines the heavy-tailedness of the distribution alone and the corresponding tail index is readily given by $\gamma = \xi^{-1}$. In contrast, random variables in practical applications can be easily rescaled to have standardised location and scale parameters, which should not distort our simulation results below. The practice relevance of GPDs has already been discussed in Section 1.2 as well as in Section 5.2.

The severity distribution of the first risk cell has infinite expectation and dominates the severity distribution of the second cell. In other words, the current setting satisfies the precondition of Theorem 4.7. Hence not surprisingly, the overall VaR_+ estimate is quite robust against various choices of Lévy copulas as demonstrated in Figure 5.13. The left subfigure compares two Clayton models with difference dependence strength, whereas the right subfigure features the Clayton, the Gumbel and the Archimedean family from Example 3.6 (2). The dependence parameter of the latter is abbreviated as θ_{Arch} in all subsequent illustrations. In particular, the parameters of the Lévy copulas underlying the right subfigure are chosen in such a way that the frequency correlation is identical across all three dependence models and given by 0.3. Furthermore, as stated in Theorem 4.7, the overall VaR_+ can be approximated by the stand-alone VaR_1 of the dominating risk cell. Both graphs in Figure 5.13 reveal that this analytical result indeed works reasonable

for the regulatory confidence level $\alpha = 99.9\%$, and even better as α continues rising and approaches one.

We would like to point out that the one-cell dominance case is ideal for a stable estimation of operational risk measures, as in favour of both financial institutions and regulatory authorities. Once the severity distribution of the severest risk cell is validated to be statistically reliable, the overall VaR_+ can be well approximated by a closed-form expression without explicit assumptions of dependence structures. Nonetheless, in a situation of near complete dependence, there are two approximation options offered by Theorem 4.6 and Theorem 4.7, respectively. Whereas Theorem 4.6 proposes to estimate the overall VaR_+

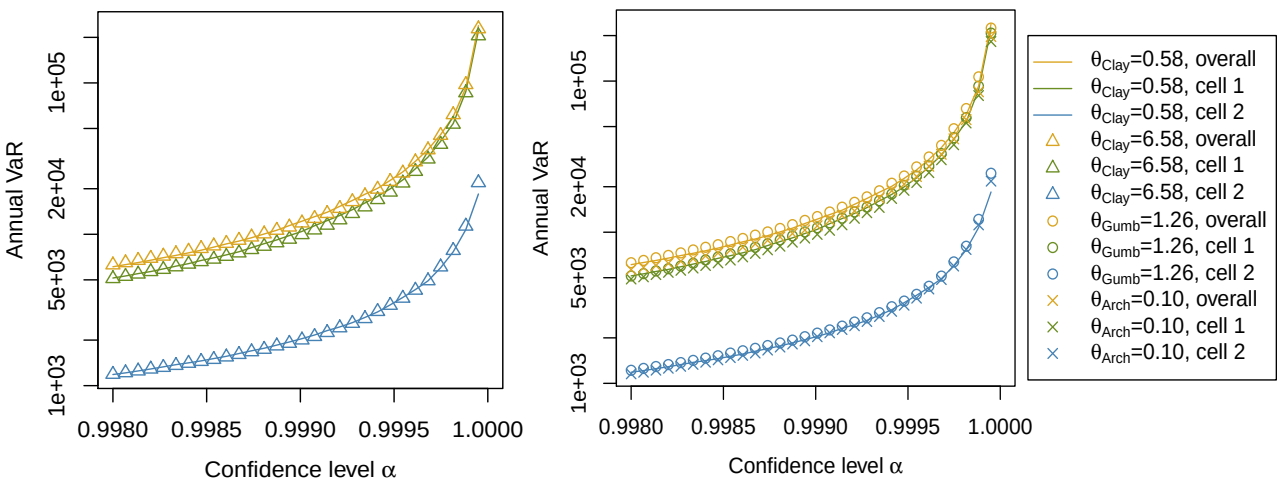


Figure 5.13: VaR estimates induced by different Lévy copula families and different dependence strengths. The underlying marginal severity distributions are given by $\mathrm{GPD}(\xi_1 = 1, \beta_1 = 1)$ and $\mathrm{GPD}(\xi_2 = 0.8, \beta_2 = 1)$. The y-axis is shown in logarithmic scale for clearer illustration.

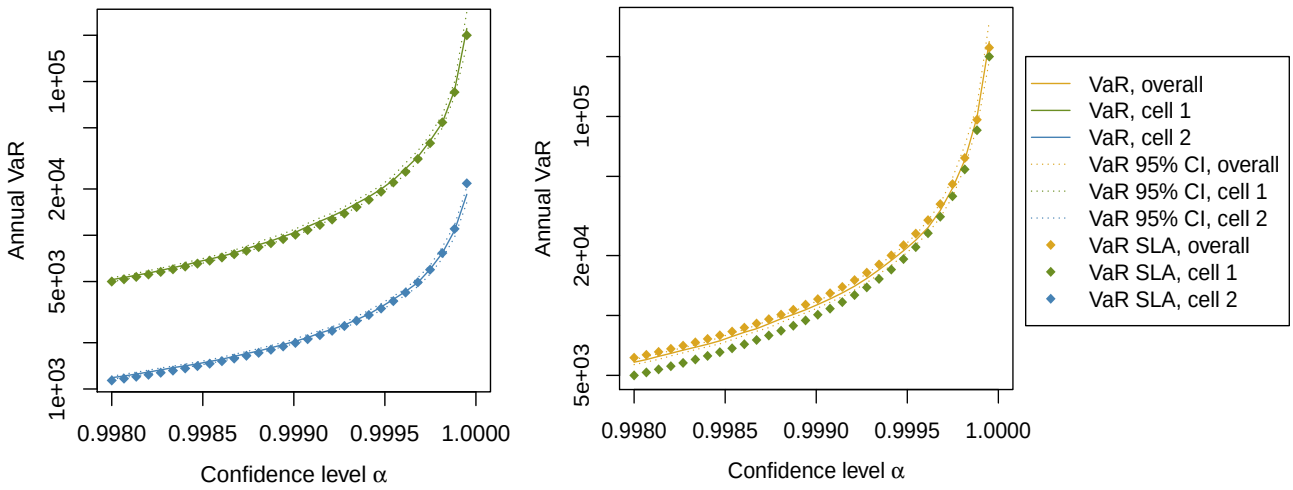


Figure 5.14: Comparison of VaR estimates obtained through simulation with analytical approximations (SLA) according to Section 4.1. The underlying marginal severity distributions are given by $\mathrm{GPD}(\xi_1 = 1, \beta_1 = 1)$ and $\mathrm{GPD}(\xi_2 = 0.8, \beta_2 = 1)$. The y-axis is shown in logarithmic scale for clearer illustration.

as the sum of the stand-alone VaR values, Theorem 4.7 suggests the approximation via the dominating marginal severity as just explained. In order to find out the more accurate method, the right panel of Figure 5.14 compares the two discussed estimation choices. The underlying data are drawn from a bivariate model based on the Clayton Lévy copula with $\theta_{\text{Clay}} = 320$. Up to a significance level around $\alpha = 99.95\%$, only the summation estimate lies within the 95% CI of the simulated result, and should therefore be preferred to the approximation through VaR_1. However, the difference becomes increasingly negligible for higher values of α. In addition, the accuracy of the stand-alone estimates VaR_1 and VaR_2 are confirmed by the left panel in Figure 5.14, as the analytical approximations lie within the 95% CI regardless the target confidence level α.

After having discussed the one-cell dominance situation in detail, now we consider a balanced severity design. That is, the marginal severities are assumed to follow a standardised GPD with identical shape parameters. Note that no closed-form risk measure approximations exist for arbitrary dependence structures in this setting, except for the independence and the complete dependence Lévy copulas. We first investigate a finite expectation case and Figure 5.15 presents estimated VaR values for $\xi_1 = \xi_2 = 0.8$. The parameters of the underlying Lévy copulas in each of the two subgraphs are precisely the same as in Figure 5.13. In contrast to Figure 5.13, Figure 5.15 clearly shows that the overall VaR_+ based on the total loss aggregated across both risk cells has varying values for different copula families and dependence strength. In other words, risk measures in a balanced severity design are sensitive to the dependence structure. Hence the Lévy copula family as well as the dependence parameter must be modelled carefully in order to obtain accurate risk measure estimates. Nonetheless, the stand-alone VaR values in Figure 5.15 are unaffected by the concrete Lévy copula, as all estimates stay within the 95% confidence interval with respect to the Clayton family with $\theta_{\text{Clay}} = 0.58$.

In the right panel of Figure 5.15, the Gumbel Lévy copula results in higher VaR_+ estimates

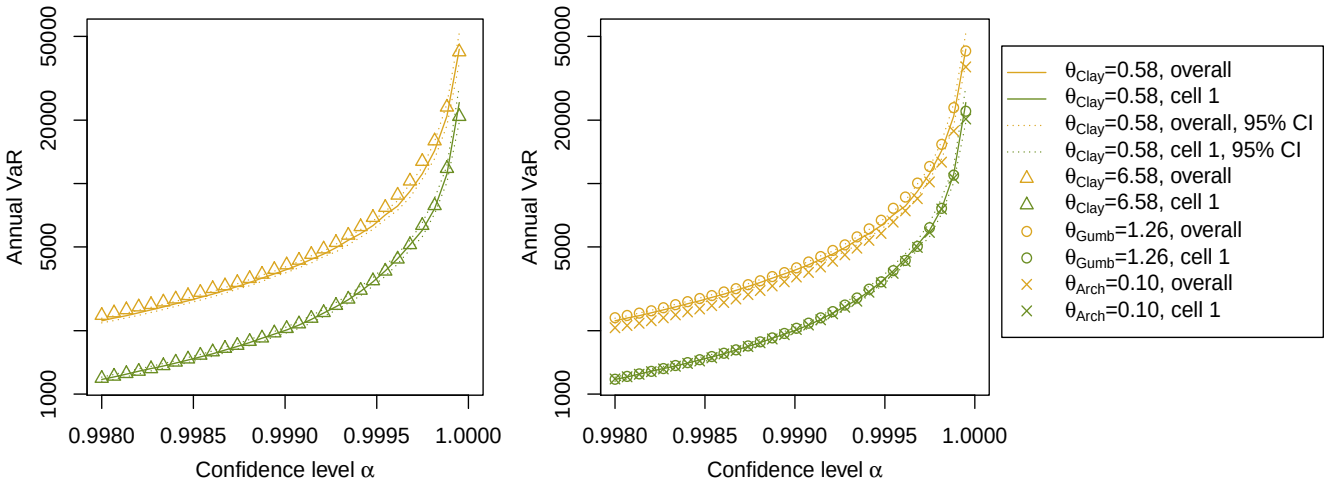


Figure 5.15: VaR estimates induced by different Lévy copula families and different dependence strengths. The marginal components have a common GPD with $\xi_1 = \xi_2 = 0.8$ and $\beta_1 = \beta_2 = 1$ as severity distribution, hence only the results for the first risk cell are depicted. The y-axis is shown in logarithmic scale for clearer illustration.

than the Clayton family, and the Archimedean family from Example 3.6 (2) offers the smallest estimates for all depicted confidence levels α. However, this is not in general the case. The common frequency correlation of the three copula families is changed from 0.3 as in Figure 5.15 to 0.001 in the left panel of Figure 5.16, and further modified to 0.99 in the right panel of Figure 5.16. For comparison purpose, the analytical approximations under the independence and the complete dependence assumptions are drawn as well. Whereas all three copula families provide nearly identical VaR_+ estimates with the independence copula in the left subfigure, the Clayton family generates the highest VaR_+ estimates in the right subfigure, followed by the Archimedean Lévy copula from 3.6 (2), which generates almost the same estimates as the complete dependence Lévy copula. Furthermore, the overall ES_+ estimates are visualised for the same selected Lévy copulas in Figure 5.17. Once again we notice the sensitivity of ES_+ towards different dependence structures and its steep increase with growing confidence level α.

In Figure 5.16, the overall VaR_+ estimate in the independence case is smaller than that under the complete dependence Lévy copula. However, it is already argued in [BK08] that subadditivity of VaR can be violated under heavy-tailed severity distributions. More precisely, the authors compare overall VaR_+ approximations for GPD severities with different shape parameters by employing Theorem 4.5 and Theorem 4.6, respectively. They conclude that in the finite expectation situation, that is, for $\xi_1 = \xi_2 < 1$, the complete dependence Lévy copula indeed produces a higher VaR_+ value than the independence one. On the other hand, subadditivity is violated for infinite expectation models with $\xi_1 = \xi_2 > 1$. Moreover, VaR_+ estimates in these two special dependence situations coincide for $\xi_1 = \xi_2 = 1$.

Within our simulation study, we extend the results in [BK08] by incorporating a medium

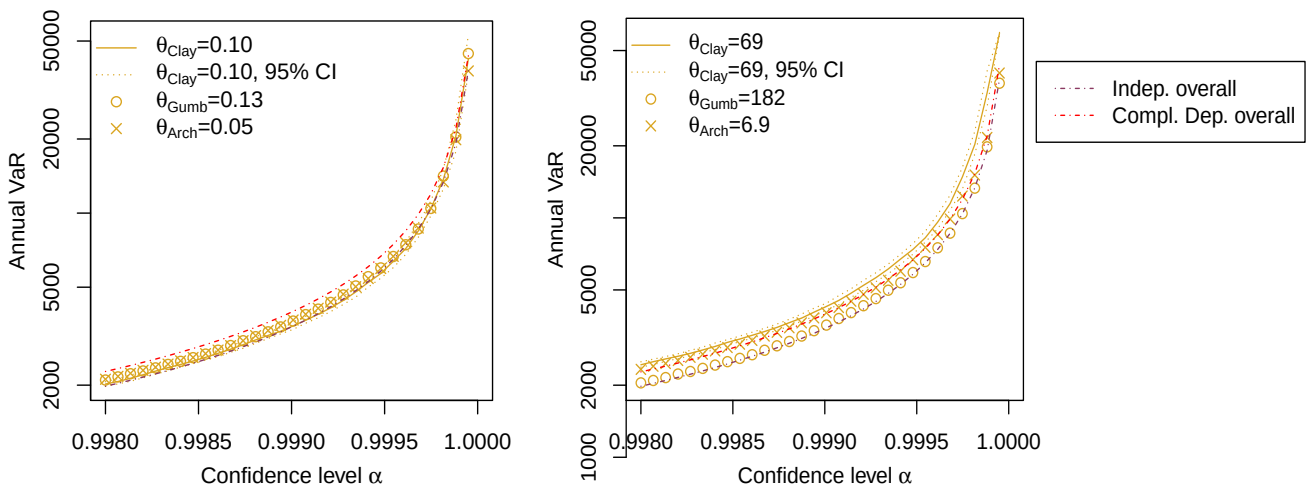


Figure 5.16: Overall VaR_+ estimates induced by different Lévy copula families and different dependence strengths. The analytical approximations for the independence and the complete dependence cases according to Section 4.1.2 are also depicted for comparison purpose. The marginal components have a common GPD with $\xi_1 = \xi_2 = 0.8$ and $\beta_1 = \beta_2 = 1$ as severity distribution and the y-axis is shown in logarithmic scale for clearer illustration.

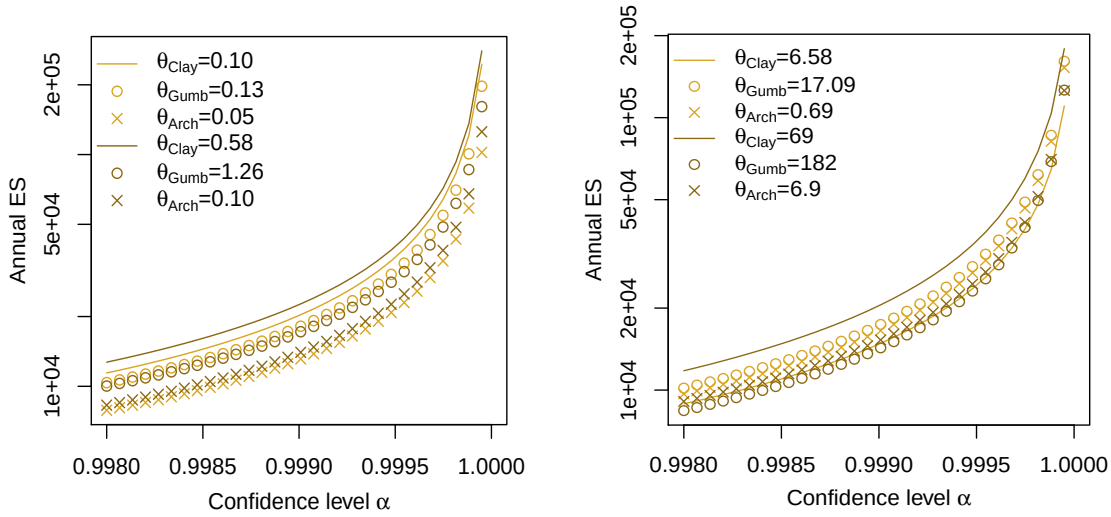


Figure 5.17: Overall ES_+ estimates induced by different Lévy copula families and different dependence strengths. The marginal components have a common GPD with $\xi_1 = \xi_2 = 0.8$ and $\beta_1 = \beta_2 = 1$ as severity distribution and the y-axis is shown in logarithmic scale for clearer illustration.

dependence structure and evaluate a broader range of potential interesting confidence levels $\alpha \in [99.8\%, 100\%)$. In order to make the results across different copula families comparable to a certain degree, the copula parameters in all four subgraphs of Figure 5.18 are determined via introducing an equal frequency correlation of 0.5.

For finite marginal severity expectations with $\xi_1 = \xi_2 = 0.8$ shown in the top left corner, the complete dependence Lévy copula indeed provides a higher value than the independence one for all depicted quantile levels. However, the overall VaR_+ estimate associated with the Gumbel family exhibits an even larger value up to roughly $\alpha = 99.95\%$. Hence the complete dependence VaR_+ cannot be regarded as an upper bond for risk measure estimations even if the severities are not extremely heavy-tailed and have finite expectations. In the threshold case of $\xi_1 = \xi_2 = 1$ shown in the top right corner, the VaR_+ values not only coincide in both ends of the dependence spectrum, but also across the three considered copula families with a medium dependence strength. Note that the VaR_+ estimates under different dependence assumptions all lie within the 95% CI associated with the Clayton family for each quantile level α. Finally, the independence VaR_+ estimate exceeds the complete dependence one in both pictures of the bottom panel. Interestingly, now the Archimedean family from Example 3.6 (2) provides the highest overall VaR_+ estimates rather than the Gumbel family as in the top left subgraph. In addition, the VaR_+ values produced by the three selected copula families indicate a more significant deviation from each other as $\xi_1 = \xi_2$ grows from 1.2 in bottom left to 1.5 in bottom right and thus is more far off the threshold one.

Beyond that, we emphasize that the heavy-tailedness of severity distributions plays a much more crucial role in view of risk measure calculations in comparison to frequency correlations. Although the latter is set equally at 0.5 for all subfigures, the obtained overall VaR_+ values exhibit obvious deviations. To further illustrate this, Figure 5.19

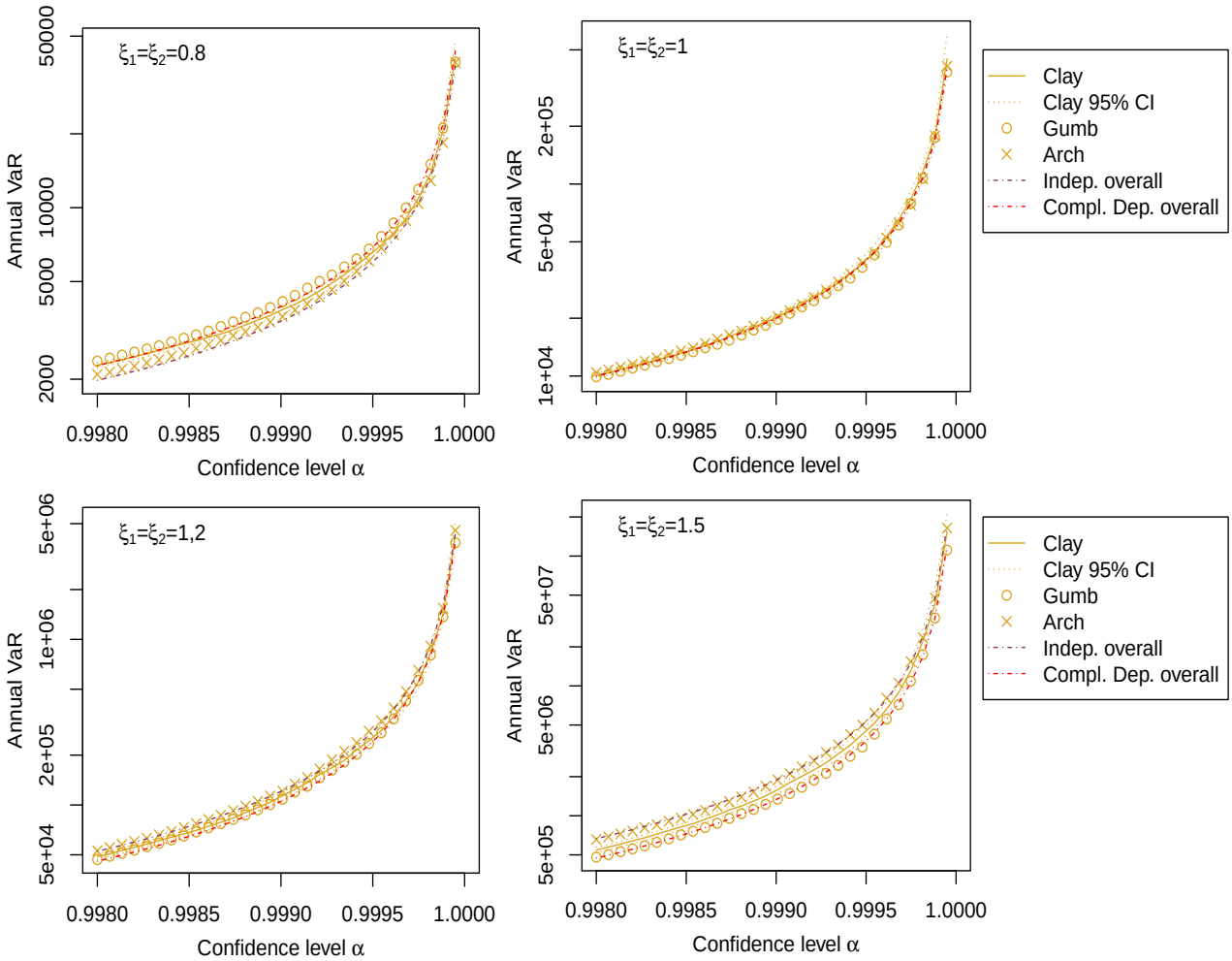


Figure 5.18: Overall VaR_+ estimates under varying shape parameters $\xi_1 = \xi_2$ of the marginal generalised Pareto severity distributions. The scale parameters are set identically at $\beta_1 = \beta_2 = 1$ in all subfigures. The different Lévy copula families are specified to have a common frequency correlation of 0.5 and the corresponding parameters are given by $\theta_{\mathrm{Clay}} = 1.00$, $\theta_{\mathrm{Gumb}} = 2.38$ and $\theta_{\mathrm{Arch}} = 0.14$. The analytical approximations for the independence and the complete dependence cases according to Section 4.1.2 are also depicted for comparison purpose. The y-axis is shown in logarithmic scale for clearer illustration.

plots the estimated VaR_+ as a function against the shape parameter $\xi_1 = \xi_2$ of the marginal GPD severities. Both the four different significant levels α corresponding to the four subgraphs and the GPD shape parameters in the x-axis belong to a realistic range for practical applications. The underlying Lévy copulas are characterised by the same dependence parameters as in Figure 5.18. As the graphs reveal an upwards linear trend with logarithmically scaled y-axis, we immediately conclude that the overall VaR_+ grows at an exponential rate with increasing $\xi_1 = \xi_2$. This pattern is consistent across all depicted confidence levels and dependence structures. Contrarily, the sensitivity of VaR_+ towards the different copula families is much less pronounced. Nonetheless, the Gumbel family provides the highest risk measure estimate for the smallest considered shape parameter $\xi_1 = \xi_2 = 0.6$, and is overtaken by the Archimedean family from Example 3.6 (2), as $\xi_1 = \xi_2$ rises to 1.5.

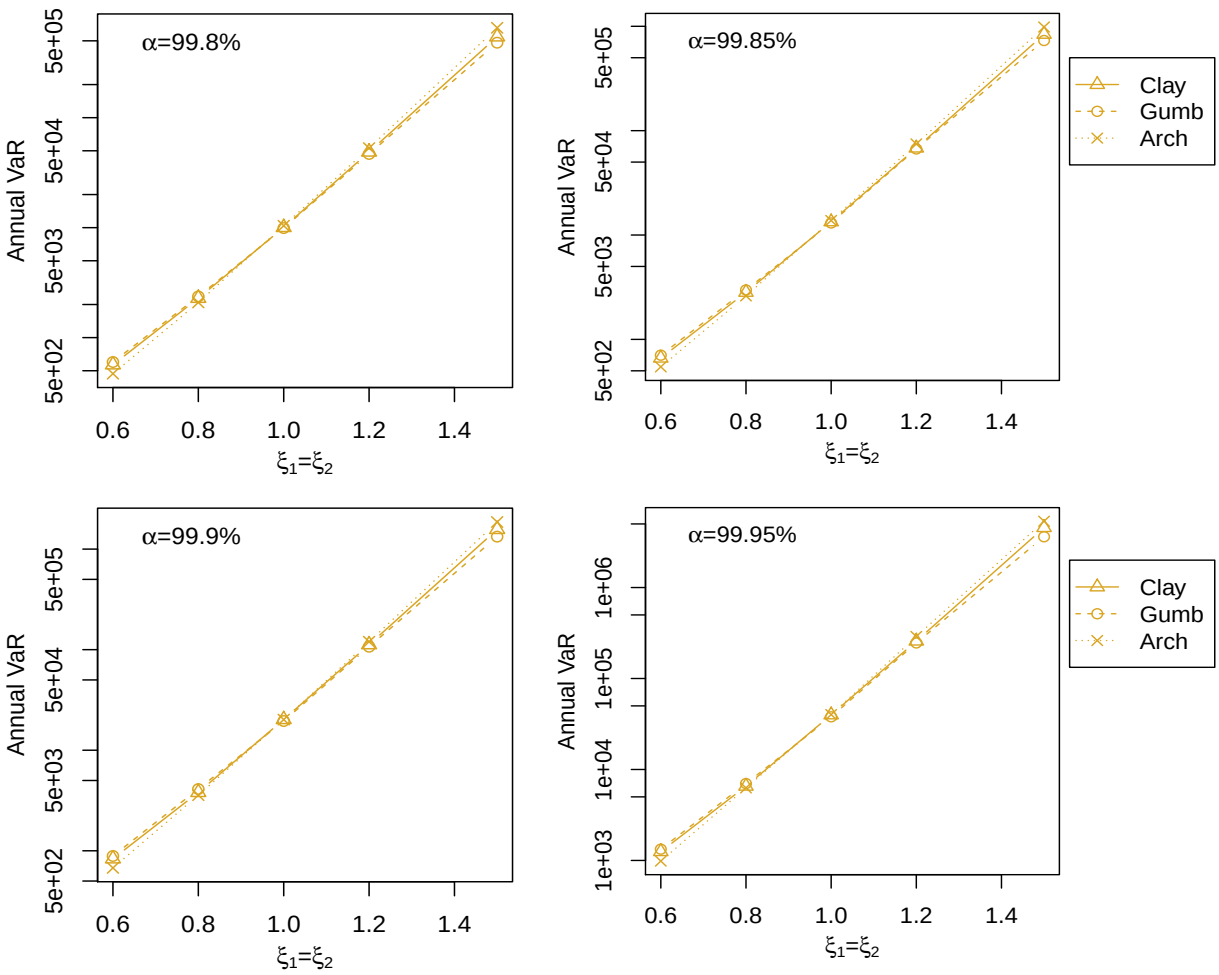


Figure 5.19: Overall VaR_+ estimates for four different confidence levels α and as a function of the shape parameter $\xi_1 = \xi_2$ underlying the marginal generalised Pareto severity distributions. The scale parameters are set identically at $\beta_1 = \beta_2 = 1$. The different Lévy copula families are specified to have a common frequency correlation of 0.5 and the corresponding parameters are given by $\theta_{\mathrm{Clay}} = 1.00$, $\theta_{\mathrm{Gumb}} = 2.38$ and $\theta_{\mathrm{Arch}} = 0.14$. The y-axis is shown in logarithmic scale for clearer illustration.

As a last observation for our simulation analysis with GPD severities, we notice the risk measure estimates are much more sensitive to the significance level α than in the less heavy-tailed Weibull case. Although Figures 5.13-5.18 are drawn with α on the x-axis and the logarithm of VaR or ES values on the y-axis, all graphs still exhibit a steeper convex shape than in Figure 5.12, where the risk measure estimates based on Weibull severities are displayed in original scale.

We close our simulation study by investigating the impact of wrongly estimated dependence structures on risk measure outcomes. More precisely, we consider three different marginal severity constellations with the true dependence structure given by the Clayton Lévy copula with $\theta_{\text{Clay}} = 1$. These three models are subsequently referred to as the true models (TMs) and the corresponding parameter values are summarised in Table 5.6. Then we simulate a sample path from each specified model up to the time horizon $T = 50$, which serves as the input for a MLE scheme in order to recover the underlying model parameters. However, when maximising the likelihood function as given in Theorem 3.9, the copula family is wrongly assumed to be the Gumbel one. The obtained false parameter estimates are also reported in Table 5.6 and accordingly denoted as the three false models (FMs).

As the parameter values of the two copula families are not directly comparable with each other, we do not calculate their relative deviations. On the other hand, the relative differences between the true and the false frequency as well as severity parameters provide

TM I	λ_1	λ_2	a_1	b_1	a_2	b_2	θ_{Clay}
	10	10	0.5	2	0.5	1	1
FM I	$\hat{\lambda}_1$	$\hat{\lambda}_2$	$\hat{a}_1$	$\hat{b}_1$	$\hat{a}_2$	$\hat{b}_2$	$\hat{\theta}_{\text{Gumb}}$
	8.9552	9.0559	0.4519	2.1822	0.4654	1.2181	2.0997
Rel. dev.	-0.1045	-0.0944	-0.0962	0.0911	-0.0693	0.2181	-
TM II	λ_1	λ_2	ξ_1	β_1	ξ_2	β_2	θ_{Clay}
	10	10	1	1	0.8	1	1
FM II	$\hat{\lambda}_1$	$\hat{\lambda}_2$	$\hat{\xi}_1$	$\hat{\beta}_1$	$\hat{\xi}_2$	$\hat{\beta}_2$	$\hat{\theta}_{\text{Gumb}}$
	10.0351	9.5624	1.2012	0.9653	1.0521	0.9601	2.1547
Rel. dev.	0.0035	-0.0438	0.2012	-0.0347	0.3151	-0.0399	-
TM III	λ_1	λ_2	ξ_1	β_1	ξ_2	β_2	θ_{Clay}
	10	10	0.8	1	0.8	1	1
FM III	$\hat{\lambda}_1$	$\hat{\lambda}_2$	$\hat{\xi}_1$	$\hat{\beta}_1$	$\hat{\xi}_2$	$\hat{\beta}_2$	$\hat{\theta}_{\text{Gumb}}$
	9.9537	10.0657	0.9409	1.0090	0.9480	1.0132	2.2542
Rel. dev.	-0.0046	0.0066	0.1762	0.0090	0.1850	0.0132	-

Table 5.6: Comparison of true model parameters with estimated parameters when falsely calibrating the dependence structure given by a Clayton Lévy copula as a Gumbel one. Model group I has Weibull severity distributions in both marginal components, whereas model groups II and III have GPD marginal severities. The relative deviations between the true and false parameters are reported in the last line for each model group.

some interesting observations. Concerning the Poisson intensities λ_1 and λ_2, their relative deviations in model groups II and III with regularly varying GPD severities are much smaller than in model group I with rapidly varying Weibull severities. Moreover, the imprecision of the estimated Weibull parameters is of approximately the same magnitude as the frequency estimates in the first model group. Unfortunately, the shape parameter estimates in model groups II and III, which constitute the most important component of a GPD, exhibit the largest relative deviations from their true values. Hence we expect the impact on risk measure computations is also more significant than with Weibull severities, which shall be examined shortly below. In contrast, the less essential scale parameter β of a GPD is estimated with a higher accuracy in false models II and III, as their relative deviations can be as small as one tenth of the relative deviations associated with the shape parameter ξ.

Nonetheless, the one-cell dominance design in true model II and the balanced severity design in true model III seem to be preserved despite the misspecification of the dependence structure. The estimate $\hat{\xi}_1$ in false model II is significantly higher than $\hat{\xi}_2$, such that losses in the first risk cell are more heavy-tailed than in the second one. On the other hand,

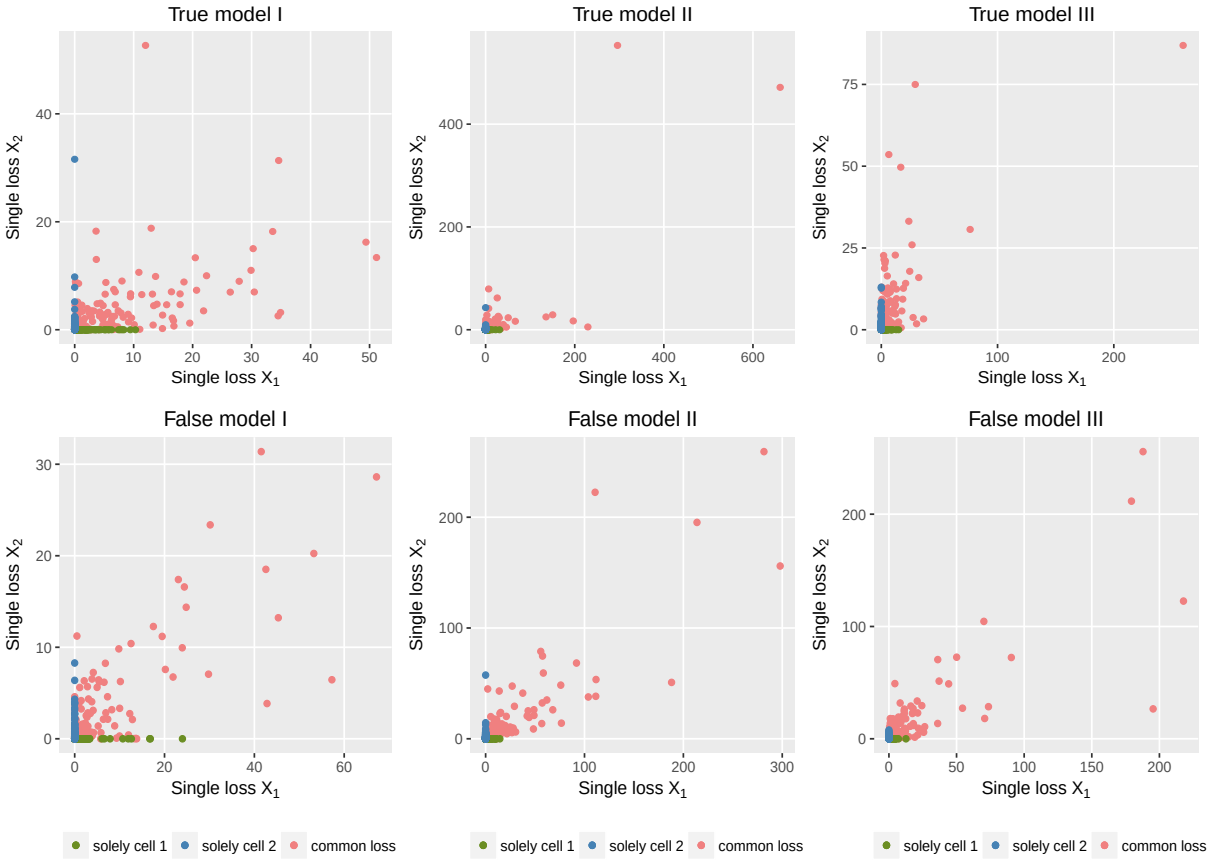


Figure 5.20: Comparison of loss severities sampled from the true models with the false models when wrongly calibrating the dependence structure given by a Clayton Lévy copula as a Gumbel one. Both the underlying marginal distribution parameters and the copula parameters are reported in Table 5.6.

the false estimates $\hat{\xi}_1$ and $\hat{\xi}_2$ for model III are nearly identical and indicate comparable marginal loss severities. Also note that model groups II and III have in common that the wrongly estimated shape parameters are bigger than their true values, which is expected to result in an overestimation of risk measures.

For illustration purpose, a scatter plot of simulated loss data is displayed in Figure 5.20 for each true and false model. The message we aim for is that the misspecification of Lévy copula families may not be easily identified by visualising the induced single loss severities, as the point clouds of the false models do not show obviously dissimilar patterns in comparison to their real counterparts.

Now we turn to evaluating the impact of inappropriately estimated dependence models on risk measure values, which presents the most relevant issue in view of capital reserve calculations. To this end, both absolute and relative deviations in VaR and ES estimates, the latter if well-defined, are depicted for each of the three model groups in Figures 5.21-5.23, respectively. Moreover, the obtained risk measure estimates themselves are reported in Table 5.7 for VaR and in Table 5.8 for ES at certain selected confidence levels.

In addition, Table 5.9 exemplarily compares the 95% CIs of VaR estimates between the true and the false models. The associated significance level is given by $\alpha = 99.9\%$, which constitutes the most commonly applied one for regulatory capital requirements. In none of the three model groups, the two CIs are overlapping, indicating massive deviations of risk measure estimates due to model misspecification. The distance between the lower bound

	99.8%	99.825%	99.85%	99.875%	99.9%	99.925%	99.95%	99.975%
TM I	293.24	299.83	307.69	317.91	329.88	346.09	371.03	413.75
FM I	458.40	470.91	485.15	504.14	527.95	557.42	604.15	689.45
TM II	6114.28	7024.06	8158.79	9753.91	12155.24	15578.86	22987.83	43047.23
FM II	30357.06	35517.04	42420.06	51586.96	66902.62	94416.58	147197.03	331672.21
TM III	2280.72	2519.75	2824.67	3210.19	3808.20	4778.93	6546.95	10946.34
FM III	6756.80	7648.88	8804.43	10459.16	12961.90	17236.27	25947.85	50918.40

Table 5.7: Comparison of total VaR_+ estimates for one time unit based on the true models with the false models when wrongly calibrating the dependence structure given by a Clayton Lévy copula as a Gumbel one. Both the underlying marginal distribution parameters and the copula parameters can be read off from Table 5.6.

	99.8%	99.825%	99.85%	99.875%	99.9%	99.925%	99.95%	99.975%
TM I	350.33	358.03	367.12	378.03	391.56	409.57	435.57	481.69
FM I	567.16	581.78	599.10	620.02	646.18	680.97	732.23	823.03
TM III	9222.02	10197.15	11451.31	13142.19	15555.15	19326.75	26191.16	44099.64
FM III	41714.22	46647.17	53054.11	61744.94	74286.60	94088.19	130679.58	225453.85

Table 5.8: Comparison of total ES_+ estimates for one time unit based on the true models with the false models when wrongly calibrating the dependence structure given by a Clayton Lévy copula as a Gumbel one. Both the underlying marginal distribution parameters and the copula parameters can be read off from Table 5.6.

TM	95% CI	FM	95% CI
I	$[327.82, 332.58]$	I	$[523.39, 532.14]$
II	$[11675.19, 12597.54]$	II	$[63631.70, 70579.16]$
III	$[3680.11, 3948.98]$	III	$[12384.29, 13554.24]$

Table 5.9: Comparison of 95% CIs associated with the overall $\mathrm{VaR}_+(99.9\%)$ estimates based on the true models with the false models when wrongly calibrating the dependence structure given by a Clayton Lévy copula as a Gumbel one. Both the underlying marginal distribution parameters and the copula parameters can be read off from Table 5.6.

of the overestimating CI and the upper bound of the true CI is the largest in model group II with the most heavy-tailed severity distributions, whereas the lower bound of the CI based on the false model is five times as high as the upper bound of the true CI.

Concerning the first model group with marginal Weibull severities depicted in Figure 5.21, the absolute deviations exhibit an increasing curve with respect to rising confidence level α for both VaR and ES, as well as for both the stand-alone and the overall values based on the total loss amount of the two risk cells. As shown by the right panel, the relative deviations in ES reveal a similarly smooth, but steeper upwards trend than the absolute deviations. However, the relative deviations in VaR have much more kinks and reveal an even sharper increase as α approaches one.

In model group II, the true marginal severity of risk cell one and the falsely estimated severities of both cells indicate an infinite expectation, hence only the resulting deviations in VaR are displayed in Figure 5.22. Interestingly, both the absolute and the relative differences in the overall VaR_+ estimates are similar to that of the dominating first cell, which has a more heavy-tailed GPD severity than the second cell in both the real and the wrong model. Whereas the second risk cell results in the lowest absolute deviations across

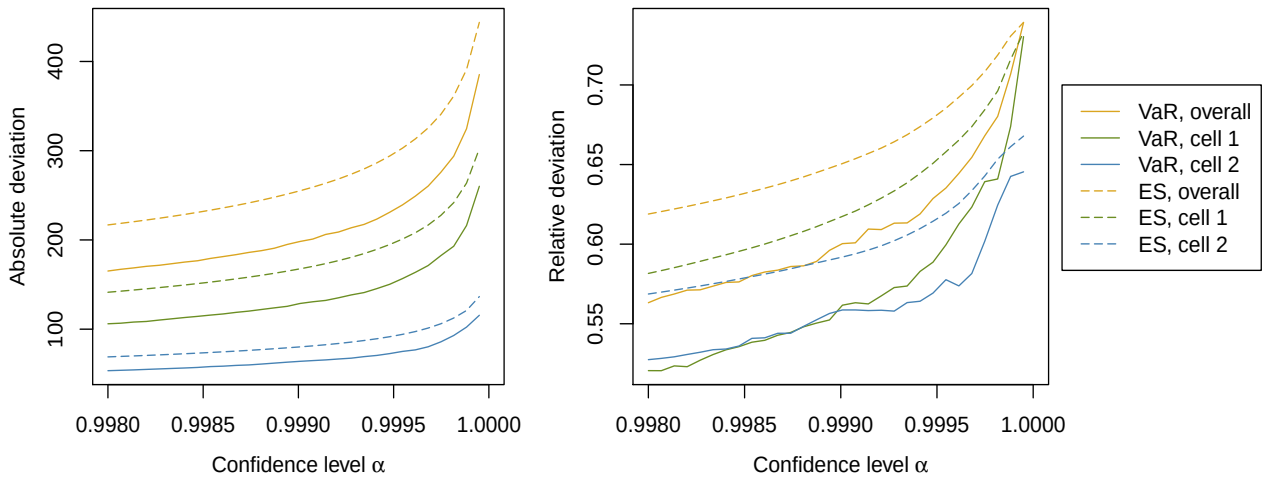


Figure 5.21: Absolute and relative deviations in risk measure estimates between the true and the false models of model group I. Both the underlying marginal distribution parameters and the copula parameters can be read off from Table 5.6.

all depicted significance levels α, it surpasses the overall VaR_+ and the individual VaR_1 of the first cell in terms of relative deviations. Although the y-axis in the left subfigure is shown in logarithmic scale, the graph still possesses an upwards convex trend, indicating that risk measure deviations grow at least exponentially with respect to α. Furthermore, the growth rate is much more dramatic than in the less heavy-tailed case with Weibull severities in Figure 5.21, where the y-axis is shown in original scale. This feature is similar to our previous observation by comparing Figures 5.13-5.18 with Figure 5.12.

Figure 5.23 illustrates the deviation in VaR and ES estimates for model group III. Recall that the underlying true model has a balanced severity design, whereby the falsely estimated shape parameters of the two risk cells have a similar value as well. Hence the

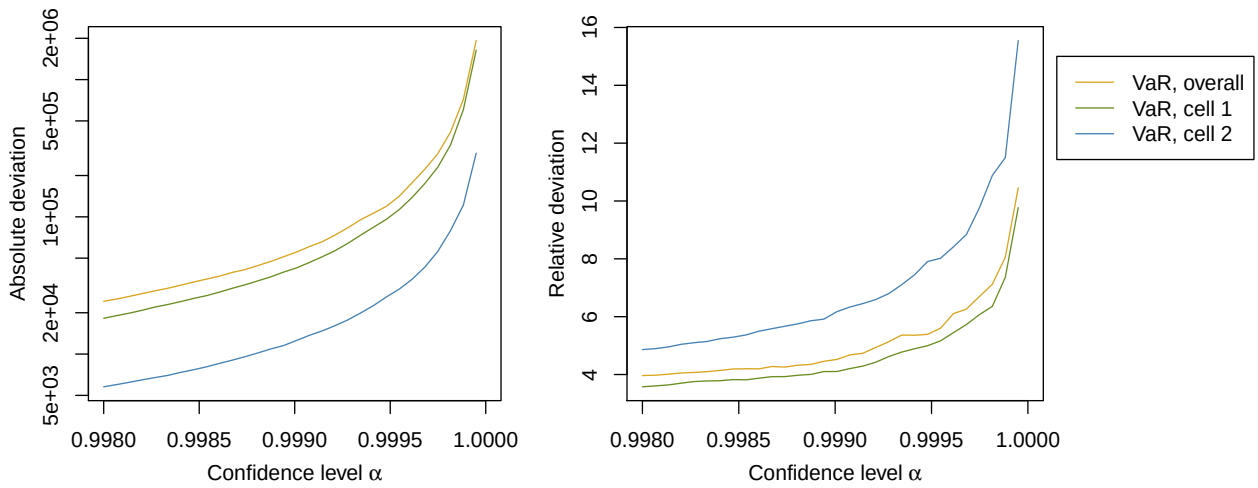


Figure 5.22: Absolute and relative deviations in risk measure estimates between the true and the false models of model group II. Both the underlying marginal distribution parameters and the copula parameters can be read off from Table 5.6. The y-axis of the left subfigure is shown in logarithmic scale for clearer illustration.

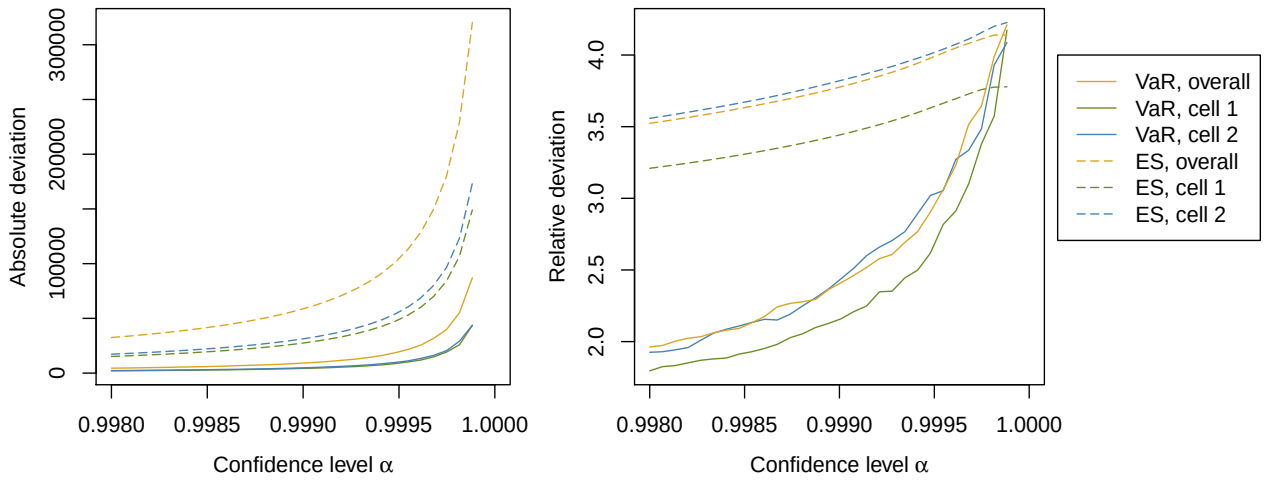


Figure 5.23: Absolute and relative deviations in risk measure estimates between the true and the false models of model group III. Both the underlying marginal distribution parameters and the copula parameters can be read off from Table 5.6.

absolute deviations in the stand-alone risk measures have similar outcomes across the two cells, and are about half the size compared to the deviation in the overall estimates VaR_+ and ES_+. However, the relative deviations in the overall risk measure outcomes are rather close to those for the second risk cell. Beyond that, ES estimates offer a much smoother increase with respect to the significance level α than VaR estimates in terms of relative deviations.

Last but not least, all three considered model groups have in common that the deviations in risk measure outcomes increase with growing confidence level α, whereas the growth rate depends on the heavy-tailedness of the underlying severity distributions. In addition, the ES deviations have a higher value than the corresponding VaR differences from both the absolute and the relative viewpoint. As already indicated within the examination of the discrepancies between the true and the false model parameters, the risk measures calculated based on all three false models overestimate their real values. This is not surprising since the scale parameters of the Weibull distributions and the shape parameters of the GPDs in the false models are estimated larger than their real counterparts, respectively. Thus the resulting loss severities have a heavier right distribution tail. To conclude, the misspecification of dependence structure by a wrongly selected Lévy copula family can cause inaccurate marginal parameter estimates as well as have an undesired impact on risk measure calculations.

Chapter 6

Real data application

In this chapter we apply our dependence modelling approach based on compound Poisson processes and Lévy copulas to three different real-life datasets. To clarify, the application example presented in Section 6.1 deals with the well-known Danish insurance claim data rather than operational loss events. The reason for this is that publicly accessible and high-quality sources of operational risk data are not yet available. The reluctance of the financial industry to share their sensitive loss information is naturally understandable. Nonetheless, we would like to exemplify our estimation methodology based on a freely available data sample, such that every modelling step and illustration can be provided in detail, as well as be tried out by interested readers. As a result, the Danish fire loss claims, which possess the same heavy-tailed distribution properties as commonly observed among operational risk incidents, seem to be the ideal choice. The multivariate version of this dataset can be found either in the package "CASdatasets" (containing various actuarial datasets) or "fitdistrplus" (developed for parametric distribution fitting) within the statistical software environment R.

In Section 6.2, both an internal and an external dataset kindly provided by the Bayerische Landesbank (BayernLB) are introduced. On the basis of these real operational loss events, we demonstrate the entire procedure from exploring potential dependence patterns, verifying model assumptions, applying MLE as detailed in Section 3.3, assessing model fit by means tested within the simulation study from Section 5.3, and finally to estimating risk measures via both simulation and approximation methods from Chapter 4.

6.1 Danish reinsurance claim dataset

The Danish insurance data were collected at the Copenhagen Reinsurance and consist of fire losses exceeding one million Danish Krone (mDKK) over the period from 1980 to 1990. Each claim is divided into a loss amount of the building coverage, of the contents coverage and of the profit coverage. As the last loss category rarely exhibits a non-zero entry, below we focus on the loss history of building and contents, which is potentially

suitable for a bivariate compound Poisson model as detailed in Definition 3.1.

In order to give our modelling approach a clear structure, we briefly state the three major steps to be followed. First we verify whether the marginal components, that is, the building and the contents claim processes, can be reasonably described by a homogeneous compound Poisson process, respectively. If this condition is satisfied, then we apply MLE by employing the likelihood function from Theorem 3.9 to fit dependence models based upon alternative Lévy copulas. Finally, we compare the fitted models by means of different diagnostic plots.

To begin with, we consider one calendar year as one time unit such that the overall observation horizon is given by $T = 11$. If the marginal claim series follow a homogeneous compound Poisson process, then the corresponding claim inter-arrival times should be i.i.d. exponentially distributed random variables. Figure 6.1 examines the exponential nature of the inter-arrival times of building and contents losses, respectively. The estimated autocorrelations shown in the top panel fairly lie within the 95% confidence interval, except for one spike in the building series. In the bottom panel, the empirical quantiles are plotted against the theoretical ones of an exponential distribution, whose rate parameter is estimated by the sample mean of the gaps between two loss arrivals. Apart from the discrete nature of empirical observations, the theoretical and empirical quantiles agree

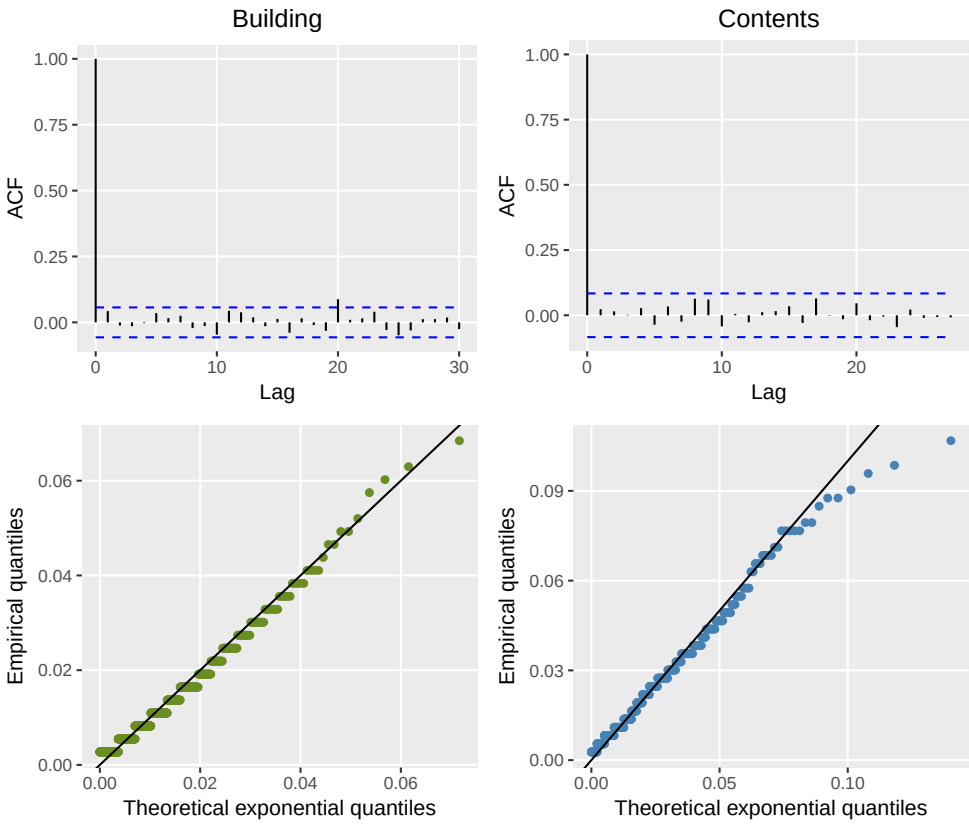


Figure 6.1: Top panel: autocorrelation functions (acfs) of marginal claim inter-arrival times. Bottom panel: Q-Q plots of the corresponding empirical quantiles against the theoretical exponential quantiles with an estimated rate parameter 108.53 for building and 49.92 for contents.

well on the positive diagonal line. Overall, there is no evidence against i.i.d. exponential claim arrival times for both the building and the contents loss processes.

Now we turn to characterising the marginal loss severities. Since the Parato-type behaviour of the fire losses in the current dataset is already well-known and the reporting limit of one mDKK provides a natural threshold, two GPDs with identical location parameter $\mu_i = 1$ for the two margins $i \in \{1, 2\}$ seem to be an appropriate choice. The associated distribution tails have the form $(1 + \xi_i \frac{x-\mu_i}{\beta_i})^{-\xi_i^{-1}}$ with shape ξ_i and scale β_i. The shape and scale parameters are estimated by means of MLE. In order to assess the GPD fit, we calculate for each loss observation x of marginal component $i \in \{1, 2\}$ the corresponding residual value $\hat{\xi}_i^{-1} \ln(1+\hat{\xi}_i \frac{x-1}{\hat{\beta}_i})$ based on estimated parameters. If the original observations follow a GPD(ξ_i, β_i, μ_i), then the residuals should be i.i.d. and follow a standard exponential distribution. In addition, the residuals have a finite second moment in contrast to the original GPD random variables with a shape parameter larger than or equal to 0.5. Therefore, we inspect the independence among loss severities via calculating the sample autocorrelation function based on the residuals, which is presented in the top panel of Figure 6.2. Furthermore, the bottom panel illustrates the match between the empirical and the theoretical exponential quantiles. Similar to the interpretation of Figure 6.1 for

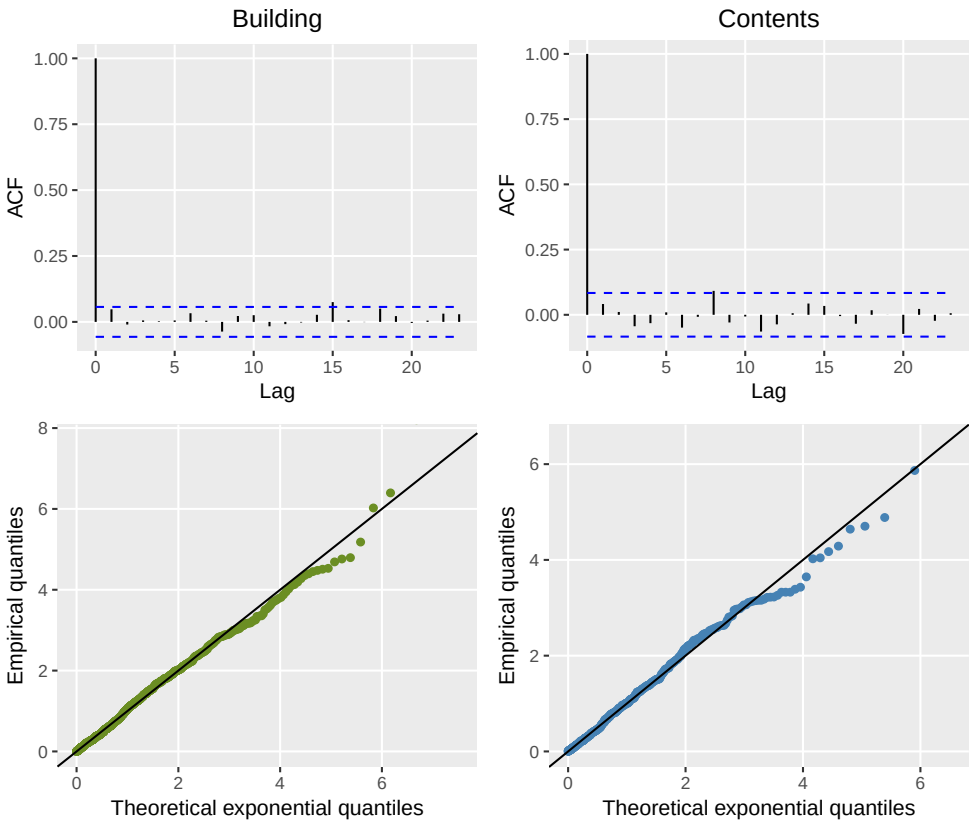


Figure 6.2: Top panel: acfs of residual values calculated from the marginal GPD fits to the loss amounts under the building and the contents coverage, respectively. The estimated parameters are given by $\hat{\xi}_1 = 0.4525$, $\hat{\beta}_1 = 1.0655$ for building and $\hat{\xi}_2 = 0.6877$, $\hat{\beta}_2 = 1.3015$ for contents. Bottom panel: Q-Q plots of the corresponding empirical quantiles against the theoretical standard exponential quantiles.

the inter-arrival times, we conclude that both building and contents claim amounts can be reasonably modelled by the GPD. Note we have also examined alternative heavy-tailed distributions not depicted here, and none of them can compete with the GPD fit.

After approving the marginal compound Poisson assumptions, now we aim at capturing the interdependence between the building and the contents loss processes by means of Lévy copulas. The potential candidates include the Clayton and the Gumbel families from Table 2.1, the Archimedean copula from Example 3.6 (2), as well as the complementary Gumbel Lévy copula from the same example. As the occurrence dates of all claim events are known exactly, we apply MLE under a continuous observation scheme as explained in Section 3.3.1. In practical implementations the corresponding log-likelihood function $\ln \mathcal{L}$ is maximised for numerical stability, and its maximum value together with the obtained estimates is summarised in Table 6.1 for different Lévy copulas.

	$\hat{\boldsymbol{\lambda}}_1$	$\hat{\boldsymbol{\lambda}}_2$	$\hat{\boldsymbol{\xi}}_1$	$\hat{\boldsymbol{\beta}}_1$	$\hat{\boldsymbol{\xi}}_2$	$\hat{\boldsymbol{\beta}}_2$	$\hat{\boldsymbol{\theta}}$	**max.** $\ln \boldsymbol{\mathcal{L}}$
Clayton	107.0989	49.2377	0.5373	1.0255	0.8110	1.2863	0.8503	1265.91
Gumbel	107.1765	50.7052	0.7466	0.9548	1.1371	1.2344	3.3625	1228.66
Example 3.6 (2)	112.5945	66.9301	0.9390	1.6044	1.6451	2.1026	0.0373	734.36
Comp. Gumbel	107.6742	50.0968	0.6717	0.9589	0.9055	1.2755	0.8448	1254.22

Table 6.1: Maximum likelihood estimates obtained by maximising the likelihood function in Theorem 3.9 for selected Lévy copula families. Note the copula parameter estimate $\hat{\theta}$ is not directly comparable across different copula families.

From Table 6.1, we immediately see that the dependence model characterised by the Archimedean Lévy copula of Example 3.6 (2) offers the lowest value of maximised log-likelihood function, which indicates its infeasibility with respect to the fire loss data. In order to draw a more comprehensive comparison between the fitted models, we employ the concept of empirical tail integrals introduced in Section 5.3. Figure 6.3 visualises their conformity with the theoretical counterparts calculated based on the estimated model parameters. The independent partial Lévy measure $\Pi_1^{\perp}$ is associated with the claims solely attributed to the building category. Similarly, the Lévy measure $\Pi_2^{\perp}$ describes contents claims without resulting in building losses. Claims attributed to both coverage categories are described by the dependent partial Lévy measure $\Pi^{\parallel} = (\Pi_1^{\parallel}, \Pi_2^{\parallel})^{\top}$.

In line with the maximised log-likelihood function values in Table 6.1, the fitted tail integrals based on the Archimedean Lévy copula from Example 3.6 (2) exhibit a poor fit to the sample version in all four subfigures. On the other hand, the Clayton, the Gumbel and its complementary family reveal equally good agreements for the individual tail integrals $\overline{\Pi}_1^{\perp}$ and $\overline{\Pi}_2^{\perp}$. However, the behaviour of the common tail integral $\overline{\Pi}^{\parallel} = (\overline{\Pi}_1^{\parallel}, \overline{\Pi}_2^{\parallel})^{\top}$ cannot be properly reproduced by any of the fitted models. Only the Clayton one, which also achieves the highest maximised log-likelihood value in Table 6.1, seems to perform slightly better than the other copulas. Our observation of Figure 6.3 constitutes no surprise, as the estimated copula parameters in Table 6.1 fall into the range from low to medium dependence strength as studied in Section 5.3. Recall Figures 5.5-5.6 where the empirical tail integrals deviate from their theoretical equivalences due to the lack of strictly positive

bivariate observations, even though the underlying data are sampled from the model with the true parameter values.

As the model suitability of the Clayton, the Gumbel and the complementary Lévy copulas cannot be fully distinguished by exploring the tail integrals in Figure 6.3, we further compare in Figure 6.4 Q-Q plots for the individual loss severities solely attributed to the building or to the contents category, respectively. The theoretical quantiles are calculated based on the corresponding survival functions $\overline{F}_1^{\perp}$ and $\overline{F}_2^{\perp}$ as specified in Lemma 3.4, Part (b). We observe deviations from the straight diagonal line in the upper quantile area of different magnitudes.

Moreover, the observed bivariate loss severities related to both the building and the contents coverage are transformed to copula scale and displayed in the upper left corner of Figure 6.5. The point cloud suggests near independence between the marginal components with a slight scarcity of points in the top left and bottom right corner, respectively. As the Archimedean Lévy copula from Example 3.6 (2) rather implies a countermonotonic relationship as revealed in the lower panel of Figure 5.10, it does not provide an appropriate fit to the fire losses. Besides the scatter plot, Figure 6.5 also presents a contour plot featuring the empirical Lévy copula as well as the three more promising fitted copula families. Once again, the sample version does not exhibit a satisfactory overlap with any

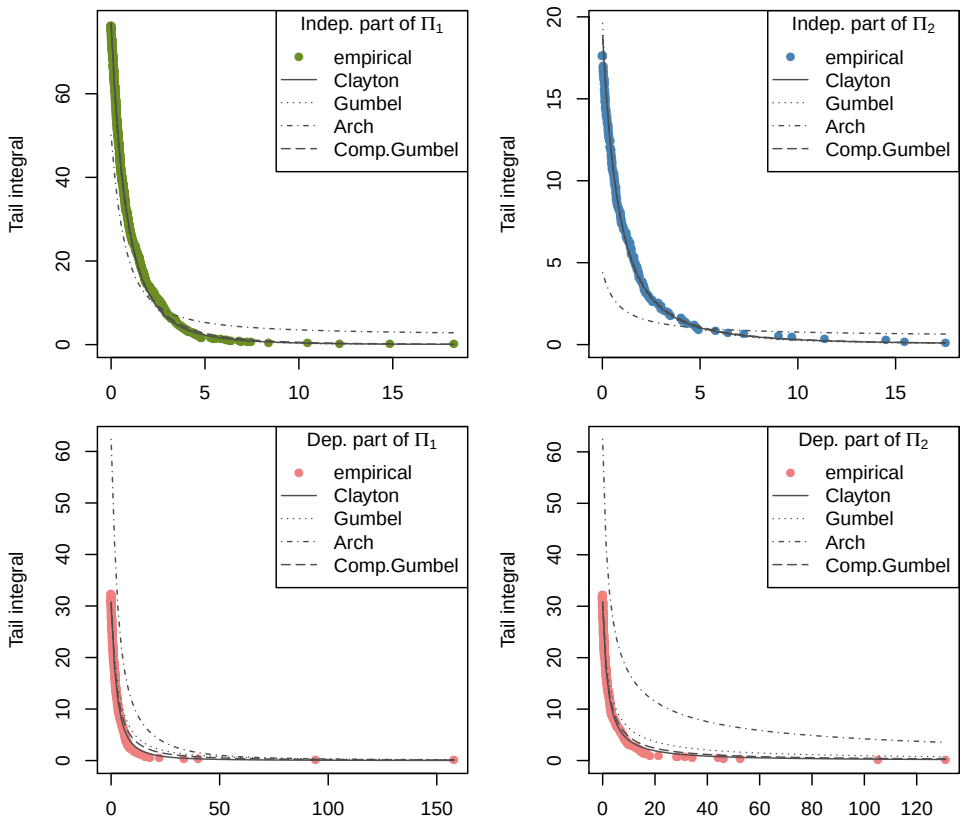


Figure 6.3: Comparison of sample tail integrals with their theoretical counterparts calculated based on different fitted Lévy copula families. The abbreviation "Arch" refers to the Archimedean Lévy copula from Example 3.6 (2).

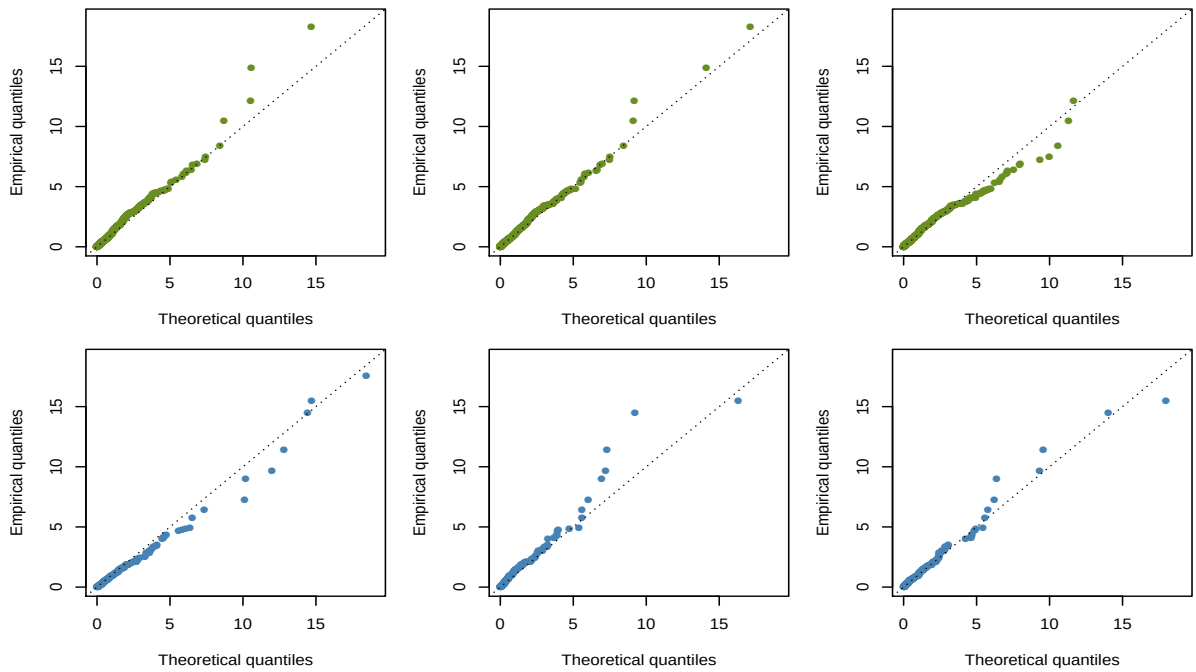
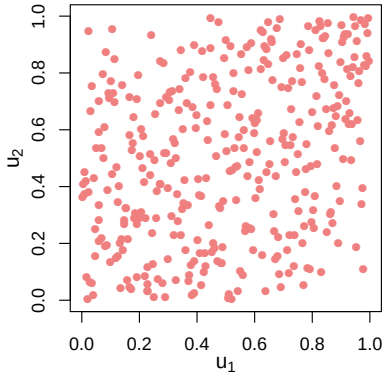


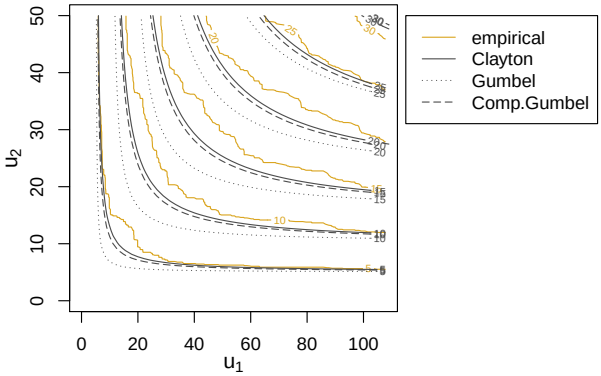
Figure 6.4: Q-Q plots of loss severities solely attributed to the building coverage (top section) and solely attributed to the contents coverage (bottom section). The left column is obtained from the estimated Clayton Lévy copula, the middle column from the Gumbel Lévy copula, and the right column from the complementary Gumbel Lévy copula.

of the theoretical counterparts owing to the current low dependence setting, similar to the reason explained before with respect to the tail integrals attributed to the common loss process. Nonetheless, the Gumbel Lévy copula with the lowest maximised log-likelihood value after the Archimedean Lévy copula from Example 3.6 (2) shows the largest deviation from the empirical contour lines. Finally, a perspective plot in the lower panel of Figure 6.5 illustrates from two different angles the shape of the theoretical Lévy copula, displayed as a 3D surface, in comparison to the sample version displayed as a 3D histogram. Although we only depict the Clayton case, similar visualisations can be obtained from other copula families.

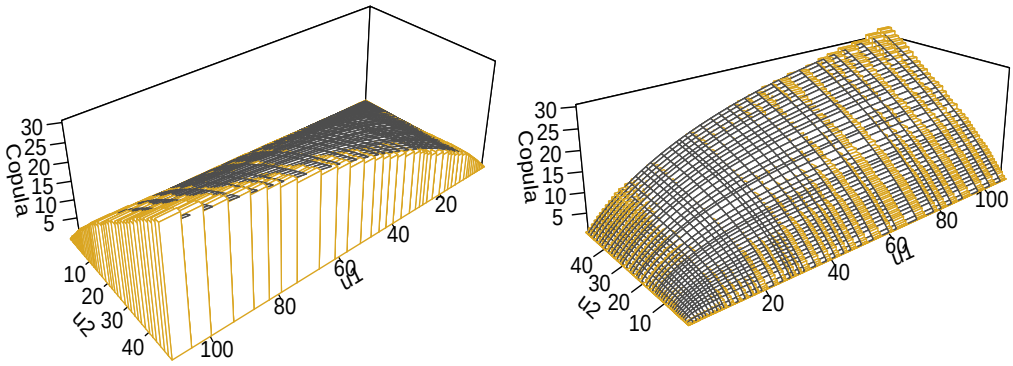
At the end we notice that the Danish insurance loss data are also analysed by means of bivariate compound Poisson models under a continuous observation scheme in [EK10] as well as under a discrete observation scheme in [Vel12]. However, in both references only the Clayton Lévy family is taken into consideration and the single losses are fitted in logarithmic scale by Weibull distributions. Hence, the current thesis contributes to the formalisation of the entire procedure of estimating a multivariate compound Poisson model, that is, from verifying the marginal compound Poisson assumptions to evaluating the fitted models based on different Lévy copulas via various diagnostic tools. Moreover, we model the loss severities in their original instead of logarithmic scale such that the obtained model can be directly utilised for operational risk measure calculations and thus is comparable with the real operational risk data introduced below.



(a) Scatter plot of losses attributed to both the building and the contents coverage in copula scale.



(b) Comparison of contour lines of the sample Lévy copula with different fitted copula families.



(c) Comparison of sample Lévy copula (yellow 3D histogram) with fitted Clayton Lévy copula (grey 3D surface).

Figure 6.5: Examination of fitted dependence models for the Danish fire losses.

6.2 Operational risk datasets

6.2.1 Internal data

It is not allowed to publish the data, but the full analysis has been made available to the supervisor and advisor.

6.2.2 External data

It is not allowed to publish the data, but the full analysis has been made available to the supervisors and advisor.

Chapter 7

Conclusion and outlook

Justified by Sklar's theorem for Lévy processes, a multivariate dependence model for operational risk loss events can be readily constructed by combining marginal compound Poisson processes through a suitable Lévy copula. The current thesis contributes to this not yet commonly known dependence concept by a detailed analysis of its model components in Chapter 3, highlighting its potential for asymptotic risk measure approximations in Chapter 4, investigating the quality of maximum likelihood estimators, the goodness of fitted dependence structures and the sensitivity of risk measure estimates in Chapter 5. Furthermore, Chapter 6 puts all previously introduced methodologies at work, whereby dependence models built upon various Lévy copulas are estimated and subsequently examined based on real loss data. In this way we hope to gain confidence in the applicability of our modelling approach also from operational risk practitioners.

In comparison to the most alternative dependence modelling techniques summarised in Section 1.2, the multivariate model based on Lévy copulas exhibits several advantages. First, the dependence structure specification can be separated from the marginal process characterisation in an analogous manner to utilising ordinary copulas. The severity model of each risk cell may be chosen from any appropriate distributions, or parameterised as a spliced one with different body and tail distribution families, such that the heterogeneity of losses attributed to different causes or business units is respected. On the other hand, various parametric Lévy copula families as well as the extreme cases of independence and complete dependence Lévy copulas are available to be selected.

Second, both frequency and severity dependence are controlled by the same Lévy copula in a calibrated model, such that relatively few parameters are required. In addition, as Lévy copulas operate on the domain of time-independent Lévy measures, the dependence characterisation stays valid as the loss occurrence progresses in the course of time. In general, the latter does not hold when imposing interdependence among compound Poisson processes via ordinary copulas, which may result in the necessity of complexer model structures with more variables. The parameter parsimony of Lévy copula models despite their flexibility is especially valuable with regard to the scarcity of reliable operational risk data compared to market or credit risk information. Moreover, the employment of Lévy

copulas enables to generalise the already well-known univariate SLA for subexponential loss severities to higher dimensions. The appeal of closed-form risk exposure estimations by contrast with simulation methods does not have to be reiterated.

Notwithstanding, the convenient property of integrating frequency and severity dependence into one single concept may also pose a limitation of Lévy copulas. Furthermore, dependence in the sense of Lévy measures is only defined for simultaneously occurring loss events, which might oversimplify real loss situations. Hence possible extensions of the current model or its combination with alternative dependence concepts constitute an interesting subject for future research. Beyond that, the issue of incorporating external data into dependence modelling and risk exposure assessment is briefly addressed at the end of Chapter 6. Since operational loss incidents of a financial institution are not least shaped by its individual procedure and control system, the challenge of rescaling external loss events with respect to representativity is evident. However, the sample size of internal observations is usually insufficient to support robust risk measure calculations at the generally high target quantile levels. As a result, it is indispensable to include external information into internal model calibration and validation schemes.

The growing participation in data consortia certainly confirms the awareness of financial institutions for sound operational risk quantifications. At the same time, a downturn in loss frequency or loss magnitude cannot be expected owing to the increasing global network of business activities and the advances made in information technology, such that operational risk exposures may become more diverse and complicated. Therefore, the incentive of implementing adequate dependence models goes far beyond the fulfilment of regulatory capital requirements, and should constitute an essential element of any risk management strategy at financial institutions. Finally, we emphasise that the assessment of the plausibility and the interpretation of a model are at least as important as fitting the model itself. Of course, all model assumptions and specifications must be periodically verified by analysts in practice and adjusted to actual loss experience where appropriate.

Appendix A

Categorisation of operational losses

Table A.1: Definition of business lines according to Annex 8 of [Ban06].

Business line	**Explanation**
Corporate Finance	Mergers and acquisitions, underwriting, privatisations, securitisation, research, debt (government, high yield), equity, syndications, IPO, secondary private placements
Trading & Sales	Fixed income, equity, foreign exchanges, commodities, credit, funding, own position securities, lending and repos, brokerage, debt, prime brokerage
Retail Banking	Retail lending and deposits, banking services, trust and estates; Private lending and deposits, banking services, trust and Retail Banking estates, investment advice; Merchant/commercial/corporate cards, private labels and retail
Commercial Banking	Project finance, real estate, export finance, trade finance, factoring, leasing, lending, guarantees, bills of exchange
Payment and Settlement	Payments and collections, funds transfer, clearing and settlement
Agency Services	Escrow, depository receipts, securities lending (customers) corporate actions; Issuer and paying agents
Asset Management	Pooled, segregated, retail, institutional, closed, open, private equity
Retail Brokerage	Execution and full service

Table A.2: Definition of event types according to Annex 9 of [Ban06].

Event type	**Explanation**
Internal fraud	Losses due to acts of a type intended to defraud, misappropriate property or circumvent regulations, the law or company policy, excluding diversity/discrimination events, which involves at least one internal party
External fraud	Losses due to acts of a type intended to defraud, misappropriate property or circumvent the law, by a third party
Employment Practices and Workplace Safety	Losses arising from acts inconsistent with employment, health or safety laws or agreements, from payment of personal injury claims, or from diversity / discrimination events
Clients, Products & Business Practices	Losses arising from an unintentional or negligent failure to meet a professional obligation to specific clients (including fiduciary and suitability requirements), or from the nature or design of a product.
Damage to Physical Assets	Losses arising from loss or damage to physical assets from natural disaster or other events.
Business disruption and system failures	Losses arising from disruption of business or system failures
Execution, Delivery & Process Management	Losses from failed transaction processing or process management, from relations with trade counterparties and vendors

Appendix B

Characterisation of distribution tails

In order to make the heavy-tailed property of operational risk losses precise, we summarise the main notions utilised in this thesis for characterising heavy-tailed probability distributions. For further discussions of the subsequent definitions we refer to Appendix 3 of [EKM97], which is a classic textbook for extreme events modelling.

Definition B.1 (**Subexponential distribution function**).
Let F be a distribution function with support $(0, \infty)$ and let F^{n*} denote its n-fold convolution. If the asymptotic relation

$$\lim_{x\to\infty} \frac{\overline{F^{n*}}(x)}{\overline{F}(x)} = n, \quad n \geq 2, \tag{B.1}$$

holds, then F is called *subexponential* and written as $\overline{F} \in \mathcal{S}$.

From relation (B.1) the following more intuitive characterisation of subexponentiality can be derived:

$$\lim_{x\to\infty} \frac{\mathbb{P}(X_1 + \cdots + X_n > x)}{\mathbb{P}(\max\{X_1, \ldots, X_n\} > x)} = 1, \quad n \geq 2.$$

Hence the distribution tail of the partial sum of n i.i.d. subexponential random variables has the same order of magnitude as the tail of the maximum among them. In other words, a severe overall loss amount is more probably owing to a single extreme loss rather than to accumulated small losses.

An important subclass of subexponential distributions comprises distributions whose far out right tail behaves like a power function:

Definition B.2 (**Regularly varying distribution tail**).
Let F be a distribution function with support $(0, \infty)$. If for some $\gamma \geq 0$ the distribution tail $\overline{F}$ satisfies the condition

$$\lim_{x\to\infty} \frac{\overline{F}(xt)}{\overline{F}(x)} = t^{-\gamma}, \quad t > 0,$$

then $\overline{F}$ is said to be *regularly varying* with index $-\gamma$ and denoted by $\overline{F} \in \mathcal{R}_{-\gamma}$. The quantity γ is often referred to as the *tail index* of F and we write $\mathcal{R} = \cup_{\gamma \geq 0} \mathcal{R}_{-\gamma}$.

Note that the k-th power moment of a random variable with $\overline{F} \in \mathcal{R}_{-\gamma}$ only exists for $k < \gamma$. A prominent example of a regularly varying distribution is given by the Pareto distribution. On the other hand, the heavy-tailed Weibull distribution and the lognormal distribution are subexponential but not regularly varying. Their tails decay faster than the tails in $\mathcal{R}$, but more slowly than any exponential tail. Hence a third distribution tail characterisation is needed:

Definition B.3 (**Rapidly varying distribution tail**).
Let F be a distribution function with support $(0, \infty)$. If the asymptotic relation

$$\lim_{x \to \infty} \frac{\overline{F}(xt)}{\overline{F}(x)} = \begin{cases} 0, & t > 1, \\ \infty, & 0 < t < 1, \end{cases}$$

holds, then $\overline{F}$ is called *rapidly varying* and denoted by $\overline{F} \in \mathcal{R}_{\infty}$.

In contrast to random variables with distribution tail in $\mathcal{R}$, all power moments of a distribution with rapidly varying tail exist and are finite.

Appendix C

Results of simulation study

In the current section we present the simulation results for Section 5.2, where the quality of obtained maximum likelihood estimates (MLEs) for a bivariate compound Poisson model based on Lévy copulas is investigated. First, the utilised performance measures shall be explained.

Let δ be either a parameter of the marginal Poisson frequency distributions, a parameter of the marginal severity distributions, or a Lévy copula parameter. A simple criterion for the goodness of its estimator $\hat{\delta}$ is given by the bias, which is defined as the expected difference between $\hat{\delta}$ and δ. In our study, the bias is estimated by the difference between the empirical mean and the true parameter value. In order to make the obtained results for parameter values of different scales comparable, the mean relative bias (MRB) is reported in the subsequent tables. The latter expresses the bias as a percentage of the true value and is estimated by

$$\widehat{\mathrm{MRB}} = \frac{1}{\delta M} \sum_{m=1}^{M} \hat{\delta}_m - \delta,$$

where $M = 100$ denotes the number of replications for each considered compound Poisson model in our simulation scheme.

A common accuracy measure is called the mean squared error (MSE), which is defined as $\mathrm{MSE}[\hat{\delta}] = \mathbb{E}[(\hat{\delta} - \delta)^2]$ and reflects the dispersion of the estimates around the true value. In particular, the MSE incorporates the concept of both bias and precision as it satisfies the decomposition $\mathrm{MSE}[\hat{\delta}] = \mathrm{Var}[\hat{\delta}] + \mathrm{Bias}[\hat{\delta}]^2$. Small variance and little bias thus lead to a highly accurate estimator. Since the MSE squares all differences, it does not have the same scale as the parameter δ. Hence we take the square root of the MSE, yielding the root mean squared error (RMSE) which is on the same scale as δ. In order to compare estimates across different model components, we again divide this measure by the true parameter value and obtain the relative RMSE (RRMSE). The corresponding empirical version is calculated as

$$\widehat{\mathrm{RRMSE}} = \frac{1}{\delta} \sqrt{\frac{1}{M} \sum_{m=1}^{M} (\hat{\delta}_m - \delta)^2}.$$

Since both the MSE and RMSE are calculated using squared differences, they could be dominated by outlying estimates far away from the true value δ. To avoid this potential problem of outliers, we also consider the absolute value of the difference between $\hat{\delta}$ and δ as a more robust measure for accuracy. The latter is commonly known as the mean absolute error (MAE) and its relative version is estimated as

$$\widehat{\mathrm{RMAE}} = \frac{1}{\delta M} \sum_{m=1}^{M} \mid \hat{\delta}_m - \delta \mid .$$

After having presented the relevant performance measures, the subsequent tables summarise the results from our simulation procedure as detailed in Section 5.2.

Table C.1: Performance measures based on 100 MLEs for the parameters of a bivariate compound Poisson model built upon the Clayton Lévy copula with different dependence strength (Top section: low dependence with true value $\theta_{\text{Clay,low}} = 0.3242$; Middle section: medium dependence with true value $\theta_{\text{Clay,med}} = 1.2212$; Bottom section: high dependence with true value $\theta_{\text{Clay,high}} = 7.0519$), and each for the time horizon of $T = 5, 10, 20$ years. The true marginal frequency parameters are given by $\lambda_1 = 40$, $\lambda_2 = 45$, and the severity distributions by $\mathcal{LN}(\mu_1 = 10.3, \sigma_1 = 1.8)$, Weib$(a_2 = 0.19, b_2 = 5000)$.

	T	$\hat{\lambda}_1$	$\hat{\lambda}_2$	$\hat{\mu}_1$	$\hat{\sigma}_1$	$\hat{a}_2$	$\hat{b}_2$	$\hat{\theta}_{\text{Clay,low}}$
	5	40.2124	45.6848	10.3202	1.7744	0.1914	5407.0411	0.3284
$\widehat{\text{Mean}}$	10	40.0461	45.3387	10.3149	1.7813	0.1904	5127.2845	0.3262
	20	39.9180	45.1435	10.3191	1.7921	0.1905	5048.6349	0.3254
	5	0.0053	0.0152	0.0020	-0.0142	0.0075	0.0814	0.0128
$\widehat{\text{MRB}}$	10	0.0012	0.0075	0.0014	-0.0104	0.0022	0.0255	0.0059
	20	-0.0021	0.0032	0.0019	-0.0044	0.0026	0.0097	0.0034
	5	0.0534	0.0539	0.0116	0.0399	0.0438	0.3323	0.0696
$\widehat{\text{RMAE}}$	10	0.0426	0.0362	0.0069	0.0306	0.0274	0.2198	0.0432
	20	0.0297	0.0258	0.0047	0.0213	0.0187	0.1529	0.0290
	5	0.0638	0.0649	0.0139	0.0494	0.0538	0.4257	0.0853
$\widehat{\text{RRMSE}}$	10	0.0504	0.0451	0.0085	0.0370	0.0344	0.2767	0.0548
	20	0.0366	0.0318	0.0059	0.0261	0.0239	0.1944	0.0390
	T	$\hat{\lambda}_1$	$\hat{\lambda}_2$	$\hat{\mu}_1$	$\hat{\sigma}_1$	$\hat{a}_2$	$\hat{b}_2$	$\hat{\theta}_{\text{Clay,med}}$
	5	39.9567	44.1770	10.2798	1.7835	0.1919	5180.8075	1.2131
$\widehat{\text{Mean}}$	10	40.1726	44.5516	10.2856	1.7943	0.1905	5132.0707	1.2242
	20	40.0598	44.8498	10.2975	1.7956	0.1898	5013.3640	1.2242
	5	-0.0011	-0.0183	-0.0020	-0.0092	0.0103	0.0362	-0.0066
$\widehat{\text{MRB}}$	10	0.0043	-0.0100	-0.0014	-0.0032	0.0028	0.0264	0.0024
	20	0.0015	-0.0033	-0.0002	-0.0025	-0.0009	0.0027	0.0025
	5	0.0593	0.0537	0.0102	0.0329	0.0347	0.2811	0.0654
$\widehat{\text{RMAE}}$	10	0.0397	0.0391	0.0067	0.0223	0.0237	0.2066	0.0501
	20	0.0317	0.0280	0.0044	0.0167	0.0170	0.1536	0.0340
	5	0.0765	0.0700	0.0126	0.0425	0.0446	0.3732	0.0835
$\widehat{\text{RRMSE}}$	10	0.0508	0.0500	0.0079	0.0289	0.0302	0.2630	0.0627
	20	0.0384	0.0351	0.0056	0.0210	0.0222	0.1888	0.0426
	T	$\hat{\lambda}_1$	$\hat{\lambda}_2$	$\hat{\mu}_1$	$\hat{\sigma}_1$	$\hat{a}_2$	$\hat{b}_2$	$\hat{\theta}_{\text{Clay,high}}$
	5	39.9308	44.8135	10.2898	1.7864	0.1918	5187.7130	7.0323
$\widehat{\text{Mean}}$	10	39.8997	44.7082	10.2979	1.7909	0.1915	5215.4336	7.0942
	20	39.9578	44.8283	10.2977	1.7937	0.1910	5153.2642	7.0786
	5	-0.0017	-0.0041	-0.0010	-0.0076	0.0093	0.0375	-0.0028
$\widehat{\text{MRB}}$	10	-0.0025	-0.0065	-0.0002	-0.0051	0.0077	0.0431	0.0060
	20	-0.0011	-0.0038	-0.0002	-0.0035	0.0052	0.0307	0.0038
	5	0.0551	0.0500	0.0090	0.0307	0.0270	0.2979	0.0628
$\widehat{\text{RMAE}}$	10	0.0374	0.0373	0.0068	0.0230	0.0183	0.2071	0.0495
	20	0.0265	0.0261	0.0054	0.0158	0.0149	0.1653	0.0361
	5	0.0696	0.0644	0.0113	0.0378	0.0337	0.3969	0.0799
$\widehat{\text{RRMSE}}$	10	0.0475	0.0468	0.0084	0.0285	0.0233	0.2778	0.0627
	20	0.0329	0.0328	0.0068	0.0197	0.0187	0.2097	0.0443

Table C.2: Performance measures based on 100 MLEs for the parameters of a bivariate compound Poisson model built upon the Clayton Lévy copula with different dependence strength (Top section: low dependence with true value $\theta_{\text{Clay,low}} = 0.3242$; Middle section: medium dependence with true value $\theta_{\text{Clay,med}} = 1.2212$; Bottom section: high dependence with true value $\theta_{\text{Clay,high}} = 7.0519$), and each for the time horizon of $T = 5, 10, 20$ years. The true marginal frequency parameters are given by $\lambda_1 = 40$, $\lambda_2 = 45$, and the severity distributions by Weib($a_1 = 0.19, b_1 = 5000$), GPD($\xi_2 = 1.3, \beta_2 = 5000$).

	T	$\hat{\lambda}_1$	$\hat{\lambda}_2$	$\hat{a}_1$	$\hat{b}_1$	$\hat{\xi}_2$	$\hat{\beta}_2$	$\hat{\theta}_{\text{Clay,low}}$
$\widehat{\text{Mean}}$	5	39.7284	44.4946	0.1919	5632.4421	1.2921	5093.9403	0.3210
	10	39.9099	44.7353	0.1907	5401.0427	1.2920	5051.7121	0.3248
	20	40.2513	44.7989	0.1909	5188.9437	1.2863	5058.2948	0.3251
$\widehat{\text{MRB}}$	5	-0.0068	-0.0112	0.0099	0.1265	-0.0061	0.0188	-0.0100
	10	-0.0023	-0.0059	0.0035	0.0802	-0.0062	0.0103	0.0018
	20	0.0063	-0.0045	0.0046	0.0378	-0.0105	0.0117	0.0028
$\widehat{\text{RMAE}}$	5	0.0614	0.0526	0.0459	0.3410	0.0805	0.1323	0.0665
	10	0.0440	0.0347	0.0303	0.2100	0.0600	0.0887	0.0425
	20	0.0298	0.0206	0.0236	0.1352	0.0446	0.0606	0.0310
$\widehat{\text{RRMSE}}$	5	0.0802	0.0632	0.0570	0.4727	0.1046	0.1671	0.0829
	10	0.0572	0.0417	0.0364	0.2749	0.0773	0.1104	0.0547
	20	0.0366	0.0248	0.0285	0.1797	0.0558	0.0795	0.0391
	T	$\hat{\lambda}_1$	$\hat{\lambda}_2$	$\hat{a}_1$	$\hat{b}_1$	$\hat{\xi}_2$	$\hat{\beta}_2$	$\hat{\theta}_{\text{Clay,med}}$
$\widehat{\text{Mean}}$	5	39.9592	44.6513	0.1902	5408.7682	1.3018	4940.3094	1.2130
	10	39.7005	44.5764	0.1897	5014.7597	1.3022	4899.4134	1.2104
	20	39.9328	44.7790	0.1900	5088.0564	1.3072	4945.4481	1.2213
$\widehat{\text{MRB}}$	5	-0.0010	-0.0077	0.0009	0.0818	0.0014	-0.0119	-0.0067
	10	-0.0075	-0.0094	-0.0013	0.0030	0.0017	-0.0201	-0.0088
	20	-0.0017	-0.0049	-0.0000	0.0176	0.0055	-0.0109	0.0001
$\widehat{\text{RMAE}}$	5	0.0568	0.0477	0.0386	0.3704	0.0786	0.1043	0.0640
	10	0.0403	0.0347	0.0243	0.2125	0.0631	0.0720	0.0487
	20	0.0246	0.0261	0.0163	0.1487	0.0418	0.0478	0.0323
$\widehat{\text{RRMSE}}$	5	0.0701	0.0605	0.0480	0.5060	0.0998	0.1272	0.0780
	10	0.0525	0.0445	0.0304	0.2901	0.0777	0.0853	0.0589
	20	0.0327	0.0323	0.0200	0.1907	0.0534	0.0623	0.0401
	T	$\hat{\lambda}_1$	$\hat{\lambda}_2$	$\hat{a}_1$	$\hat{b}_1$	$\hat{\xi}_2$	$\hat{\beta}_2$	$\hat{\theta}_{\text{Clay,high}}$
$\widehat{\text{Mean}}$	5	39.0748	44.0593	0.1920	5663.4055	1.3002	5145.6403	7.1948
	10	39.3702	44.4336	0.1907	5315.3628	1.3030	5057.1749	7.1242
	20	39.6745	44.6520	0.1903	5328.1002	1.3111	5053.1263	7.1247
$\widehat{\text{MRB}}$	5	-0.0231	-0.0209	0.0105	0.1327	0.0002	0.0291	0.0203
	10	-0.0157	-0.0126	0.0036	0.0631	0.0023	0.0114	0.0103
	20	-0.0081	-0.0077	0.0014	0.0656	0.0086	0.0106	0.0103
$\widehat{\text{RMAE}}$	5	0.0576	0.0526	0.0311	0.3176	0.0671	0.0938	0.0780
	10	0.0389	0.0345	0.0209	0.1944	0.0437	0.0707	0.0530
	20	0.0285	0.0256	0.0147	0.1544	0.0380	0.0471	0.0345
$\widehat{\text{RRMSE}}$	5	0.0720	0.0651	0.0391	0.3971	0.0848	0.1173	0.0999
	10	0.0485	0.0433	0.0258	0.2476	0.0567	0.0898	0.0669
	20	0.0353	0.0326	0.0188	0.1922	0.0479	0.0599	0.0446

Table C.3: Performance measures based on 100 MLEs for the parameters of a bivariate compound Poisson model built upon the Clayton Lévy copula with different dependence strength (Top section: low dependence with true value $\theta_{\text{Clay,low}} = 0.3242$; Middle section: medium dependence with true value $\theta_{\text{Clay,med}} = 1.2212$; Bottom section: high dependence with true value $\theta_{\text{Clay,high}} = 7.0519$), and each for the time horizon of $T = 5, 10, 20$ years. The true marginal frequency parameters are given by $\lambda_1 = 40$, $\lambda_2 = 45$, and the severity distributions by $\mathcal{LN}(\mu_1 = 10.3, \sigma_1 = 1.8)$, GPD$(\xi_2 = 1.3, \beta_2 = 5000)$.

	T	$\hat{\lambda}_1$	$\hat{\lambda}_2$	$\hat{\mu}_1$	$\hat{\sigma}_1$	$\hat{\xi}_2$	$\hat{\beta}_2$	$\hat{\theta}_{\text{Clay,low}}$
$\widehat{\text{Mean}}$	5	39.8686	45.1355	10.3061	1.7881	1.2873	5133.8954	0.3248
	10	40.0596	44.8414	10.2987	1.7896	1.2770	5099.1410	0.3218
	20	39.7667	44.9224	10.2985	1.7963	1.2773	5092.6631	0.3234
$\widehat{\text{MRB}}$	5	-0.0033	0.0030	0.0006	-0.0066	-0.0098	0.0268	0.0016
	10	0.0015	-0.0035	-0.0001	-0.0058	-0.0177	0.0198	-0.0076
	20	-0.0058	-0.0017	-0.0001	-0.0021	-0.0175	0.0185	-0.0025
$\widehat{\text{RMAE}}$	5	0.0524	0.0529	0.0108	0.0342	0.0912	0.1280	0.0791
	10	0.0393	0.0363	0.0073	0.0272	0.0679	0.0824	0.0499
	20	0.0279	0.0248	0.0041	0.0177	0.0459	0.0615	0.0357
$\widehat{\text{RRMSE}}$	5	0.0657	0.0682	0.0132	0.0433	0.1121	0.1563	0.0999
	10	0.0485	0.0480	0.0086	0.0340	0.0835	0.1065	0.0635
	20	0.0349	0.0312	0.0050	0.0213	0.0584	0.0785	0.0454
	T	$\hat{\lambda}_1$	$\hat{\lambda}_2$	$\hat{\mu}_1$	$\hat{\sigma}_1$	$\hat{\xi}_2$	$\hat{\beta}_2$	$\hat{\theta}_{\text{Clay,med}}$
$\widehat{\text{Mean}}$	5	39.8461	44.4483	10.3084	1.7873	1.3070	5063.4838	1.2288
	10	40.0318	44.7946	10.3034	1.7951	1.3126	5013.9403	1.2287
	20	39.9473	44.8117	10.3018	1.8014	1.3058	5004.9352	1.2234
$\widehat{\text{MRB}}$	5	-0.0038	-0.0123	0.0008	-0.0071	0.0054	0.0127	0.0062
	10	0.0008	-0.0046	0.0003	-0.0027	0.0097	0.0028	0.0062
	20	-0.0013	-0.0042	0.0002	0.0008	0.0045	0.0010	0.0018
$\widehat{\text{RMAE}}$	5	0.0547	0.0535	0.0092	0.0370	0.0788	0.1149	0.0601
	10	0.0359	0.0350	0.0067	0.0269	0.0517	0.0734	0.0445
	20	0.0275	0.0289	0.0044	0.0179	0.0421	0.0532	0.0342
$\widehat{\text{RRMSE}}$	5	0.0700	0.0663	0.0119	0.0464	0.1003	0.1565	0.0777
	10	0.0463	0.0457	0.0087	0.0337	0.0681	0.0951	0.0553
	20	0.0335	0.0364	0.0058	0.0219	0.0501	0.0682	0.0429
	T	$\hat{\lambda}_1$	$\hat{\lambda}_2$	$\hat{\mu}_1$	$\hat{\sigma}_1$	$\hat{\xi}_2$	$\hat{\beta}_2$	$\hat{\theta}_{\text{Clay,high}}$
$\widehat{\text{Mean}}$	5	39.1453	44.2740	10.2870	1.7974	1.2900	4947.2790	7.0393
	10	39.3099	44.3939	10.2915	1.7974	1.2937	4949.5253	7.0344
	20	39.6911	44.7503	10.2958	1.8015	1.3013	4961.7985	7.0235
$\widehat{\text{MRB}}$	5	-0.0214	-0.0161	-0.0013	-0.0015	-0.0077	-0.0105	-0.0018
	10	-0.0173	-0.0135	-0.0008	-0.0015	-0.0049	-0.0101	-0.0025
	20	-0.0077	-0.0055	-0.0004	0.0008	0.0010	-0.0076	-0.0040
$\widehat{\text{RMAE}}$	5	0.0527	0.0522	0.0096	0.0350	0.0764	0.0970	0.0696
	10	0.0387	0.0375	0.0067	0.0246	0.0508	0.0632	0.0497
	20	0.0282	0.0252	0.0053	0.0179	0.0337	0.0533	0.0353
$\widehat{\text{RRMSE}}$	5	0.0672	0.0637	0.0117	0.0455	0.0949	0.1171	0.0892
	10	0.0501	0.0475	0.0082	0.0320	0.0644	0.0774	0.0619
	20	0.0346	0.0309	0.0064	0.0226	0.0433	0.0629	0.0427

Table C.4: Performance measures based on 100 MLEs for the parameters of a bivariate compound Poisson model built upon the Clayton Lévy copula with different dependence strength (Top section: low dependence with true value $\theta_{\text{Clay,low}} = 0.3242$; Middle section: medium dependence with true value $\theta_{\text{Clay,med}} = 1.2212$; Bottom section: high dependence with true value $\theta_{\text{Clay,high}} = 7.0519$), and each for the time horizon of $T = 5, 10, 20$ years. The true marginal frequency parameters are given by $\lambda_1 = 40$, $\lambda_2 = 45$, and the severity distributions by Weib($a_1 = 0.16, b_1 = 4000$), Weib($a_2 = 0.19, b_2 = 5000$).

	T	$\hat{\lambda}_1$	$\hat{\lambda}_2$	$\hat{a}_1$	$\hat{b}_1$	$\hat{a}_2$	$\hat{b}_2$	$\hat{\theta}_{\text{Clay,low}}$
$\widehat{\text{Mean}}$	5	39.4263	45.1948	0.1615	4809.1522	0.1911	5687.2657	0.3289
	10	39.8423	45.2273	0.1615	4361.6227	0.1902	5339.0626	0.3272
	20	39.8868	45.2191	0.1605	4171.9963	0.1906	5180.3028	0.3264
$\widehat{\text{MRB}}$	5	-0.0143	0.0043	0.0091	0.2023	0.0058	0.1375	0.0145
	10	-0.0039	0.0051	0.0095	0.0904	0.0009	0.0678	0.0089
	20	-0.0028	0.0049	0.0031	0.0430	0.0030	0.0361	0.0065
$\widehat{\text{RMAE}}$	5	0.0516	0.0487	0.0451	0.4410	0.0403	0.3671	0.0675
	10	0.0382	0.0358	0.0318	0.2875	0.0309	0.2185	0.0488
	20	0.0281	0.0237	0.0222	0.2020	0.0182	0.1332	0.0334
$\widehat{\text{RRMSE}}$	5	0.0635	0.0606	0.0558	0.6125	0.0496	0.4794	0.0840
	10	0.0483	0.0441	0.0413	0.3767	0.0366	0.2786	0.0600
	20	0.0350	0.0305	0.0293	0.2486	0.0227	0.1716	0.0423
	T	$\hat{\lambda}_1$	$\hat{\lambda}_2$	$\hat{a}_1$	$\hat{b}_1$	$\hat{a}_2$	$\hat{b}_2$	$\hat{\theta}_{\text{Clay,med}}$
$\widehat{\text{Mean}}$	5	39.8399	45.2654	0.1613	4197.9251	0.1908	5156.4578	1.2241
	10	40.0144	45.1698	0.1606	4101.4673	0.1909	5178.6554	1.2179
	20	40.2196	45.3155	0.1599	4050.1897	0.1909	5122.6918	1.2188
$\widehat{\text{MRB}}$	5	-0.0040	0.0059	0.0084	0.0495	0.0043	0.0313	0.0024
	10	0.0004	0.0038	0.0037	0.0254	0.0046	0.0357	-0.0027
	20	0.0055	0.0070	-0.0004	0.0125	0.0050	0.0245	-0.0020
$\widehat{\text{RMAE}}$	5	0.0540	0.0489	0.0383	0.3289	0.0389	0.2904	0.0626
	10	0.0381	0.0401	0.0245	0.2413	0.0294	0.1959	0.0429
	20	0.0274	0.0298	0.0168	0.1807	0.0209	0.1318	0.0315
$\widehat{\text{RRMSE}}$	5	0.0666	0.0603	0.0476	0.4301	0.0491	0.3561	0.0762
	10	0.0483	0.0516	0.0319	0.3181	0.0346	0.2482	0.0530
	20	0.0355	0.0367	0.0218	0.2287	0.0257	0.1720	0.0377
	T	$\hat{\lambda}_1$	$\hat{\lambda}_2$	$\hat{a}_1$	$\hat{b}_1$	$\hat{a}_2$	$\hat{b}_2$	$\hat{\theta}_{\text{Clay,high}}$
$\widehat{\text{Mean}}$	5	40.3741	45.3980	0.1611	4281.1287	0.1912	5174.3061	7.0648
	10	40.1756	45.1960	0.1605	4150.7013	0.1902	5050.7024	7.1111
	20	40.2173	45.2234	0.1601	4050.7317	0.1899	4996.3067	7.0878
$\widehat{\text{MRB}}$	5	0.0094	0.0088	0.0070	0.0703	0.0061	0.0349	0.0018
	10	0.0044	0.0044	0.0032	0.0377	0.0009	0.0101	0.0084
	20	0.0054	0.0050	0.0008	0.0127	-0.0007	-0.0007	0.0051
$\widehat{\text{RMAE}}$	5	0.0553	0.0517	0.0256	0.3774	0.0315	0.2912	0.0674
	10	0.0351	0.0337	0.0193	0.2508	0.0198	0.1883	0.0576
	20	0.0274	0.0251	0.0144	0.1788	0.0156	0.1380	0.0380
$\widehat{\text{RRMSE}}$	5	0.0685	0.0630	0.0329	0.5088	0.0394	0.3787	0.0888
	10	0.0438	0.0411	0.0232	0.3416	0.0254	0.2553	0.0714
	20	0.0338	0.0316	0.0176	0.2242	0.0197	0.1784	0.0479

Table C.5: Performance measures based on 100 MLEs for the parameters of a bivariate compound Poisson model built upon the Clayton Lévy copula with different dependence strength (Top section: low dependence with true value $\theta_{\text{Clay,low}} = 0.3242$; Middle section: medium dependence with true value $\theta_{\text{Clay,med}} = 1.2212$; Bottom section: high dependence with true value $\theta_{\text{Clay,high}} = 7.0519$), and each for the time horizon of $T = 5, 10, 20$ years. The true marginal frequency parameters are given by $\lambda_1 = 40$, $\lambda_2 = 45$, and the severity distributions by $\mathcal{LN}(\mu_1 = 10.3, \sigma_1 = 1.8)$, $\mathcal{LN}(\mu_2 = 9.8, \sigma_2 = 1.4)$.

	T	$\hat{\lambda}_1$	$\hat{\lambda}_2$	$\hat{\mu}_1$	$\hat{\sigma}_1$	$\hat{\mu}_2$	$\hat{\sigma}_2$	$\hat{\theta}_{\text{Clay,low}}$
$\widehat{\text{Mean}}$	5	39.9164	44.9538	10.3100	1.7975	9.7996	1.4027	0.3243
	10	39.9262	44.9793	10.3103	1.7943	9.7980	1.4035	0.3251
	20	39.9969	45.1528	10.3010	1.7920	9.8027	1.4040	0.3238
$\widehat{\text{MRB}}$	5	-0.0021	-0.0010	0.0010	-0.0014	-0.0000	0.0019	0.0002
	10	-0.0018	-0.0005	0.0010	-0.0032	-0.0002	0.0025	0.0027
	20	-0.0001	0.0034	0.0001	-0.0044	0.0003	0.0029	-0.0013
$\widehat{\text{RMAE}}$	5	0.0519	0.0505	0.0096	0.0390	0.0079	0.0426	0.0634
	10	0.0358	0.0356	0.0063	0.0253	0.0057	0.0259	0.0461
	20	0.0260	0.0271	0.0049	0.0182	0.0039	0.0171	0.0323
$\widehat{\text{RRMSE}}$	5	0.0671	0.0669	0.0121	0.0496	0.0099	0.0519	0.0797
	10	0.0452	0.0457	0.0081	0.0309	0.0071	0.0345	0.0581
	20	0.0322	0.0329	0.0063	0.0235	0.0046	0.0216	0.0414
	T	$\hat{\lambda}_1$	$\hat{\lambda}_2$	$\hat{\mu}_1$	$\hat{\sigma}_1$	$\hat{\mu}_2$	$\hat{\sigma}_2$	$\hat{\theta}_{\text{Clay,med}}$
$\widehat{\text{Mean}}$	5	39.7469	44.6879	10.3087	1.7844	9.8170	1.3908	1.2182
	10	39.9191	44.8704	10.3056	1.7867	9.8082	1.3959	1.2170
	20	39.9989	44.7628	10.3034	1.7937	9.8003	1.3988	1.2188
$\widehat{\text{MRB}}$	5	-0.0063	-0.0069	0.0008	-0.0087	0.0017	-0.0066	-0.0024
	10	-0.0020	-0.0029	0.0005	-0.0074	0.0008	-0.0029	-0.0034
	20	-0.0000	-0.0053	0.0003	-0.0035	0.0000	-0.0009	-0.0020
$\widehat{\text{RMAE}}$	5	0.0500	0.0527	0.0091	0.0381	0.0076	0.0326	0.0598
	10	0.0345	0.0372	0.0069	0.0256	0.0056	0.0228	0.0401
	20	0.0263	0.0254	0.0053	0.0180	0.0041	0.0175	0.0311
$\widehat{\text{RRMSE}}$	5	0.0645	0.0662	0.0120	0.0455	0.0095	0.0397	0.0790
	10	0.0450	0.0465	0.0086	0.0317	0.0071	0.0281	0.0497
	20	0.0320	0.0310	0.0067	0.0223	0.0052	0.0207	0.0366
	T	$\hat{\lambda}_1$	$\hat{\lambda}_2$	$\hat{\mu}_1$	$\hat{\sigma}_1$	$\hat{\mu}_2$	$\hat{\sigma}_2$	$\hat{\theta}_{\text{Clay,high}}$
$\widehat{\text{Mean}}$	5	40.4191	45.6044	10.2911	1.7912	9.7902	1.3961	7.0370
	10	40.4015	45.4730	10.2984	1.7971	9.7976	1.3999	7.0508
	20	40.1179	45.2114	10.3061	1.7966	9.8011	1.3991	7.0650
$\widehat{\text{MRB}}$	5	0.0105	0.0134	-0.0009	-0.0049	-0.0010	-0.0028	-0.0021
	10	0.0100	0.0105	-0.0002	-0.0016	-0.0002	-0.0001	-0.0002
	20	0.0029	0.0047	0.0006	-0.0019	0.0001	-0.0006	0.0019
$\widehat{\text{RMAE}}$	5	0.0555	0.0488	0.0095	0.0295	0.0078	0.0307	0.0729
	10	0.0386	0.0352	0.0071	0.0194	0.0053	0.0215	0.0516
	20	0.0292	0.0272	0.0050	0.0148	0.0038	0.0154	0.0361
$\widehat{\text{RRMSE}}$	5	0.0714	0.0630	0.0116	0.0362	0.0099	0.0366	0.0933
	10	0.0484	0.0443	0.0086	0.0241	0.0068	0.0259	0.0646
	20	0.0363	0.0345	0.0062	0.0181	0.0047	0.0201	0.0462

Table C.6: Performance measures based on 100 MLEs for the parameters of a bivariate compound Poisson model built upon the Clayton Lévy copula with different dependence strength (Top section: low dependence with true value $\theta_{\text{Clay,low}} = 0.3242$; Middle section: medium dependence with true value $\theta_{\text{Clay,med}} = 1.2212$; Bottom section: high dependence with true value $\theta_{\text{Clay,high}} = 7.0519$), and each for the time horizon of $T = 5, 10, 20$ years. The true marginal frequency parameters are given by $\lambda_1 = 40$, $\lambda_2 = 45$, and the severity distributions by GPD($\xi_1 = 1.5, \beta_1 = 6000$), GPD($\xi_2 = 1.3, \beta_2 = 5000$).

	T	$\hat{\lambda}_1$	$\hat{\lambda}_2$	$\hat{\xi}_1$	$\hat{\beta}_1$	$\hat{\xi}_2$	$\hat{\beta}_2$	$\hat{\theta}_{\text{Clay,low}}$
$\widehat{\text{Mean}}$	5	39.4806	44.9344	1.4527	6075.8581	1.2809	5016.4564	0.3139
	10	39.9064	44.9220	1.4882	6053.7551	1.2952	4977.6855	0.3190
	20	40.1449	45.0540	1.4895	6019.4580	1.2896	5030.6567	0.3212
$\widehat{\text{MRB}}$	5	-0.0130	-0.0015	-0.0316	0.0126	-0.0147	0.0033	-0.0318
	10	-0.0023	-0.0017	-0.0078	0.0090	-0.0037	-0.0045	-0.0162
	20	0.0036	0.0012	-0.0070	0.0032	-0.0080	0.0061	-0.0095
$\widehat{\text{RMAE}}$	5	0.0659	0.0531	0.0844	0.1065	0.0910	0.1302	0.0733
	10	0.0363	0.0387	0.0686	0.0852	0.0646	0.0871	0.0510
	20	0.0301	0.0256	0.0513	0.0601	0.0453	0.0522	0.0407
$\widehat{\text{RRMSE}}$	5	0.0790	0.0676	0.1040	0.1307	0.1124	0.1660	0.0955
	10	0.0442	0.0485	0.0822	0.1035	0.0791	0.1107	0.0617
	20	0.0362	0.0325	0.0620	0.0762	0.0543	0.0681	0.0503
	T	$\hat{\lambda}_1$	$\hat{\lambda}_2$	$\hat{\xi}_1$	$\hat{\beta}_1$	$\hat{\xi}_2$	$\hat{\beta}_2$	$\hat{\theta}_{\text{Clay,med}}$
$\widehat{\text{Mean}}$	5	40.2708	44.8805	1.4971	6074.1477	1.2964	5030.7057	1.2304
	10	39.9647	44.6188	1.4883	6101.4901	1.3029	4979.2258	1.2275
	20	40.0979	44.7993	1.4925	6008.4340	1.2929	5024.6935	1.2259
$\widehat{\text{MRB}}$	5	0.0068	-0.0027	-0.0020	0.0124	-0.0027	0.0061	0.0075
	10	-0.0009	-0.0085	-0.0078	0.0169	0.0022	-0.0042	0.0051
	20	0.0024	-0.0045	-0.0050	0.0014	-0.0055	0.0049	0.0039
$\widehat{\text{RMAE}}$	5	0.0540	0.0501	0.0789	0.1122	0.0821	0.1142	0.0716
	10	0.0426	0.0348	0.0545	0.0892	0.0564	0.0806	0.0424
	20	0.0302	0.0249	0.0389	0.0647	0.0401	0.0517	0.0323
$\widehat{\text{RRMSE}}$	5	0.0725	0.0636	0.1018	0.1353	0.1042	0.1444	0.0880
	10	0.0543	0.0437	0.0682	0.1147	0.0702	0.1000	0.0562
	20	0.0388	0.0323	0.0484	0.0805	0.0500	0.0656	0.0403
	T	$\hat{\lambda}_1$	$\hat{\lambda}_2$	$\hat{\xi}_1$	$\hat{\beta}_1$	$\hat{\xi}_2$	$\hat{\beta}_2$	$\hat{\theta}_{\text{Clay,high}}$
$\widehat{\text{Mean}}$	5	40.1587	45.2055	1.4715	6102.6530	1.2719	5081.5584	7.1509
	10	40.1376	45.0229	1.4793	6040.4492	1.2789	5078.1190	7.0773
	20	39.9179	44.8596	1.4822	6025.7559	1.2831	5050.1588	7.0820
$\widehat{\text{MRB}}$	5	0.0040	0.0046	-0.0190	0.0171	-0.0216	0.0163	0.0140
	10	0.0034	0.0005	-0.0138	0.0067	-0.0162	0.0156	0.0036
	20	-0.0021	-0.0031	-0.0119	0.0043	-0.0130	0.0100	0.0043
$\widehat{\text{RMAE}}$	5	0.0492	0.0431	0.0645	0.1000	0.0666	0.0897	0.0647
	10	0.0354	0.0326	0.0414	0.0650	0.0452	0.0611	0.0394
	20	0.0277	0.0277	0.0307	0.0449	0.0334	0.0429	0.0316
$\widehat{\text{RRMSE}}$	5	0.0620	0.0556	0.0814	0.1247	0.0833	0.1139	0.0796
	10	0.0451	0.0415	0.0515	0.0796	0.0552	0.0774	0.0536
	20	0.0346	0.0346	0.0383	0.0582	0.0406	0.0519	0.0384

Table C.7: Performance measures based on 100 MLEs for the parameters of a bivariate compound Poisson model built upon the Clayton Lévy copula with different dependence strength (Top section: low dependence with true value $\theta_{\text{Clay,low}} = 0.3242$; Middle section: medium dependence with true value $\theta_{\text{Clay,med}} = 1.2212$; Bottom section: high dependence with true value $\theta_{\text{Clay,high}} = 7.0519$), and each for the time horizon of $T = 5, 10, 20$ years. The true marginal frequency parameters are given by $\lambda_1 = 40$, $\lambda_2 = 45$, and the severity distributions by GPD($\xi_1 = 0.9, \beta_1 = 6000$), GPD($\xi_2 = 0.9, \beta_2 = 6000$).

	T	$\hat{\lambda}_1$	$\hat{\lambda}_2$	$\hat{\xi}_1$	$\hat{\beta}_1$	$\hat{\xi}_2$	$\hat{\beta}_2$	$\hat{\theta}_{\text{Clay,low}}$
$\widehat{\text{Mean}}$	5	39.6000	44.9858	0.9056	6169.0486	0.8814	6198.6746	0.3270
	10	39.7663	45.1606	0.9013	6110.6873	0.8979	6112.7971	0.3259
	20	39.8128	45.0460	0.8956	6079.5861	0.8983	6058.4359	0.3258
$\widehat{\text{MRB}}$	5	-0.0100	-0.0003	0.0062	0.0282	-0.0206	0.0331	0.0085
	10	-0.0058	0.0036	0.0015	0.0184	-0.0024	0.0188	0.0049
	20	-0.0047	0.0010	-0.0049	0.0133	-0.0019	0.0097	0.0047
$\widehat{\text{RMAE}}$	5	0.0553	0.0463	0.1092	0.1234	0.1200	0.1053	0.0616
	10	0.0362	0.0324	0.0730	0.0907	0.0791	0.0707	0.0434
	20	0.0293	0.0214	0.0535	0.0602	0.0574	0.0509	0.0318
$\widehat{\text{RRMSE}}$	5	0.0681	0.0579	0.1436	0.1530	0.1554	0.1302	0.0791
	10	0.0465	0.0403	0.0905	0.1101	0.1019	0.0885	0.0547
	20	0.0367	0.0273	0.0660	0.0732	0.0737	0.0619	0.0412
	T	$\hat{\lambda}_1$	$\hat{\lambda}_2$	$\hat{\xi}_1$	$\hat{\beta}_1$	$\hat{\xi}_2$	$\hat{\beta}_2$	$\hat{\theta}_{\text{Clay,med}}$
$\widehat{\text{Mean}}$	5	40.4259	45.1548	0.9022	6081.2606	0.8983	6110.7307	1.2363
	10	40.2095	44.9877	0.9072	5979.5402	0.9091	6017.7748	1.2307
	20	40.1766	45.0691	0.9065	5940.8229	0.9116	5974.4792	1.2262
$\widehat{\text{MRB}}$	5	0.0106	0.0034	0.0025	0.0135	-0.0018	0.0185	0.0124
	10	0.0052	-0.0003	0.0080	-0.0034	0.0101	0.0030	0.0078
	20	0.0044	0.0015	0.0072	-0.0099	0.0129	-0.0043	0.0041
$\widehat{\text{RMAE}}$	5	0.0656	0.0553	0.0893	0.0903	0.0900	0.0915	0.0594
	10	0.0430	0.0352	0.0677	0.0701	0.0697	0.0665	0.0435
	20	0.0309	0.0303	0.0520	0.0505	0.0490	0.0453	0.0302
$\widehat{\text{RRMSE}}$	5	0.0789	0.0722	0.1146	0.1104	0.1062	0.1151	0.0752
	10	0.0521	0.0476	0.0854	0.0871	0.0876	0.0819	0.0554
	20	0.0384	0.0370	0.0652	0.0611	0.0638	0.0560	0.0375
	T	$\hat{\lambda}_1$	$\hat{\lambda}_2$	$\hat{\xi}_1$	$\hat{\beta}_1$	$\hat{\xi}_2$	$\hat{\beta}_2$	$\hat{\theta}_{\text{Clay,high}}$
$\widehat{\text{Mean}}$	5	40.7235	45.6070	0.8854	6110.4527	0.8860	6166.9081	7.1168
	10	40.1786	45.2457	0.8919	6097.7207	0.8943	6080.3753	7.1054
	20	40.1925	45.3117	0.8936	6034.8480	0.8948	6006.7217	7.0586
$\widehat{\text{MRB}}$	5	0.0181	0.0135	-0.0162	0.0184	-0.0156	0.0278	0.0092
	10	0.0045	0.0055	-0.0090	0.0163	-0.0063	0.0134	0.0076
	20	0.0048	0.0069	-0.0071	0.0058	-0.0057	0.0011	0.0010
$\widehat{\text{RMAE}}$	5	0.0531	0.0493	0.0937	0.0975	0.0903	0.1001	0.0750
	10	0.0363	0.0334	0.0651	0.0593	0.0633	0.0600	0.0487
	20	0.0279	0.0282	0.0435	0.0409	0.0417	0.0399	0.0341
$\widehat{\text{RRMSE}}$	5	0.0682	0.0630	0.1119	0.1207	0.1061	0.1233	0.0918
	10	0.0469	0.0454	0.0804	0.0741	0.0773	0.0747	0.0597
	20	0.0350	0.0344	0.0543	0.0510	0.0519	0.0512	0.0417

Table C.8: Performance measures based on 100 MLEs for the parameters of a bivariate compound Poisson model built upon the Gumbel Lévy copula with different dependence strength (Top section: low dependence with true value $\theta_{\text{Gumb,low}} = 0.9317$; Middle section: medium dependence with true value $\theta_{\text{Gumb,med}} = 4.3951$; Bottom section: high dependence with true value $\theta_{\text{Gumb,high}} = 26.8980$), and each for the time horizon of $T = 5, 10, 20$ years. The true marginal frequency parameters are given by $\lambda_1 = 40$, $\lambda_2 = 45$, and the severity distributions by Weib($a_1 = 0.16, b_1 = 4000$), Weib($a_2 = 0.19, b_2 = 5000$).

	T	$\hat{\lambda}_1$	$\hat{\lambda}_2$	$\hat{a}_1$	$\hat{b}_1$	$\hat{a}_2$	$\hat{b}_2$	$\hat{\theta}_{\text{Gumb,low}}$
$\widehat{\text{Mean}}$	5	40.2164	44.6700	0.1610	4320.0109	0.1933	5772.8305	0.9195
	10	40.2581	44.9302	0.1617	4169.4294	0.1921	5436.4003	0.9265
	20	40.0259	44.7941	0.1608	4152.3709	0.1905	5144.2345	0.9321
$\widehat{\text{MRB}}$	5	0.0054	-0.0073	0.0060	0.0800	0.0174	0.1546	-0.0131
	10	0.0065	-0.0016	0.0104	0.0424	0.0111	0.0873	-0.0056
	20	0.0006	-0.0046	0.0053	0.0381	0.0027	0.0288	0.0004
$\widehat{\text{RMAE}}$	5	0.0527	0.0612	0.0439	0.3591	0.0451	0.3214	0.0921
	10	0.0343	0.0410	0.0328	0.2775	0.0299	0.2059	0.0623
	20	0.0287	0.0305	0.0197	0.1904	0.0194	0.1403	0.0486
$\widehat{\text{RRMSE}}$	5	0.0656	0.0760	0.0538	0.5006	0.0554	0.4388	0.1182
	10	0.0430	0.0502	0.0407	0.3569	0.0373	0.2673	0.0780
	20	0.0355	0.0373	0.0251	0.2463	0.0236	0.1790	0.0617
	T	$\hat{\lambda}_1$	$\hat{\lambda}_2$	$\hat{a}_1$	$\hat{b}_1$	$\hat{a}_2$	$\hat{b}_2$	$\hat{\theta}_{\text{Gumb,med}}$
$\widehat{\text{Mean}}$	5	40.2316	45.0689	0.1609	4292.3140	0.1917	5374.2003	4.4070
	10	39.9582	44.9934	0.1601	4117.0492	0.1911	5149.1394	4.4032
	20	39.8848	44.8213	0.1602	4054.5263	0.1909	5160.2550	4.3889
$\widehat{\text{MRB}}$	5	0.0058	0.0015	0.0054	0.0731	0.0090	0.0748	0.0027
	10	-0.0010	-0.0001	0.0004	0.0293	0.0056	0.0298	0.0019
	20	-0.0029	-0.0040	0.0009	0.0136	0.0047	0.0321	-0.0014
$\widehat{\text{RMAE}}$	5	0.0553	0.0537	0.0373	0.3516	0.0311	0.3023	0.0757
	10	0.0395	0.0355	0.0257	0.2448	0.0223	0.1837	0.0505
	20	0.0266	0.0261	0.0173	0.1738	0.0150	0.1439	0.0373
$\widehat{\text{RRMSE}}$	5	0.0698	0.0643	0.0455	0.4757	0.0403	0.3742	0.0937
	10	0.0480	0.0479	0.0320	0.3201	0.0265	0.2322	0.0666
	20	0.0351	0.0342	0.0211	0.2145	0.0189	0.1749	0.0455
	T	$\hat{\lambda}_1$	$\hat{\lambda}_2$	$\hat{a}_1$	$\hat{b}_1$	$\hat{a}_2$	$\hat{b}_2$	$\hat{\theta}_{\text{Gumb,high}}$
$\widehat{\text{Mean}}$	5	39.5735	44.4862	0.1611	4361.8710	0.1913	5295.9069	27.3544
	10	39.6912	44.7412	0.1605	4181.0857	0.1904	5111.8820	27.0513
	20	39.9180	44.9589	0.1604	4163.8795	0.1904	5136.0205	27.0818
$\widehat{\text{MRB}}$	5	-0.0107	-0.0114	0.0071	0.0905	0.0067	0.0592	0.0170
	10	-0.0077	-0.0058	0.0033	0.0453	0.0023	0.0224	0.0057
	20	-0.0021	-0.0009	0.0024	0.0410	0.0019	0.0272	0.0068
$\widehat{\text{RMAE}}$	5	0.0573	0.0585	0.0285	0.3683	0.0295	0.2688	0.0814
	10	0.0452	0.0452	0.0204	0.2591	0.0194	0.1981	0.0529
	20	0.0323	0.0310	0.0146	0.1750	0.0143	0.1456	0.0392
$\widehat{\text{RRMSE}}$	5	0.0698	0.0711	0.0354	0.4827	0.0367	0.3717	0.1009
	10	0.0571	0.0555	0.0245	0.3297	0.0253	0.2567	0.0647
	20	0.0407	0.0389	0.0179	0.2212	0.0181	0.1740	0.0481

Table C.9: Performance measures based on 100 MLEs for the parameters of a bivariate compound Poisson model built upon the Gumbel Lévy copula with different dependence strength (Top section: low dependence with true value $\theta_{\text{Gumb,low}} = 0.9317$; Middle section: medium dependence with true value $\theta_{\text{Gumb,med}} = 4.3951$; Bottom section: high dependence with true value $\theta_{\text{Gumb,high}} = 26.8980$), and each for the time horizon of $T = 5, 10, 20$ years. The true marginal frequency parameters are given by $\lambda_1 = 40$, $\lambda_2 = 45$, and the severity distributions by $\mathcal{LN}(\mu_1 = 10.3, \sigma_1 = 1.8)$, $\mathcal{LN}(\mu_2 = 9.8, \sigma_2 = 1.4)$.

	T	$\hat{\lambda}_1$	$\hat{\lambda}_2$	$\hat{\mu}_1$	$\hat{\sigma}_1$	$\hat{\mu}_2$	$\hat{\sigma}_2$	$\hat{\theta}_{\text{Gumb,low}}$
$\widehat{\text{Mean}}$	5	39.9114	45.0306	10.2825	1.7873	9.7971	1.3967	0.9244
	10	39.9733	44.9614	10.2893	1.7907	9.8030	1.4011	0.9267
	20	39.9813	45.0108	10.2933	1.7965	9.8022	1.4024	0.9270
$\widehat{\text{MRB}}$	5	-0.0022	0.0007	-0.0017	-0.0071	-0.0003	-0.0024	-0.0079
	10	-0.0007	-0.0009	-0.0010	-0.0052	0.0003	0.0008	-0.0054
	20	-0.0005	0.0002	-0.0007	-0.0019	0.0002	0.0017	-0.0051
$\widehat{\text{RMAE}}$	5	0.0553	0.0509	0.0089	0.0363	0.0079	0.0313	0.0871
	10	0.0369	0.0376	0.0069	0.0283	0.0060	0.0256	0.0612
	20	0.0263	0.0265	0.0048	0.0206	0.0042	0.0184	0.0467
$\widehat{\text{RRMSE}}$	5	0.0691	0.0632	0.0109	0.0466	0.0101	0.0399	0.1073
	10	0.0457	0.0470	0.0085	0.0358	0.0072	0.0322	0.0779
	20	0.0329	0.0333	0.0059	0.0257	0.0051	0.0223	0.0581
	T	$\hat{\lambda}_1$	$\hat{\lambda}_2$	$\hat{\mu}_1$	$\hat{\sigma}_1$	$\hat{\mu}_2$	$\hat{\sigma}_2$	$\hat{\theta}_{\text{Gumb,med}}$
$\widehat{\text{Mean}}$	5	39.6337	44.5268	10.2866	1.7912	9.7896	1.3914	4.3207
	10	39.8199	44.9170	10.3117	1.7960	9.8003	1.4016	4.4197
	20	40.0314	45.0300	10.3086	1.8034	9.8095	1.4002	4.4266
$\widehat{\text{MRB}}$	5	-0.0092	-0.0105	-0.0013	-0.0049	-0.0011	-0.0062	-0.0169
	10	-0.0045	-0.0018	0.0011	-0.0022	0.0000	0.0011	0.0056
	20	0.0008	0.0007	0.0008	0.0019	0.0010	0.0001	0.0072
$\widehat{\text{RMAE}}$	5	0.0656	0.0548	0.0109	0.0246	0.0074	0.0277	0.0896
	10	0.0455	0.0394	0.0083	0.0195	0.0054	0.0198	0.0621
	20	0.0273	0.0250	0.0052	0.0130	0.0040	0.0130	0.0422
$\widehat{\text{RRMSE}}$	5	0.0829	0.0686	0.0142	0.0307	0.0096	0.0367	0.1140
	10	0.0571	0.0485	0.0101	0.0250	0.0069	0.0250	0.0785
	20	0.0342	0.0327	0.0065	0.0160	0.0052	0.0170	0.0541
	T	$\hat{\lambda}_1$	$\hat{\lambda}_2$	$\hat{\mu}_1$	$\hat{\sigma}_1$	$\hat{\mu}_2$	$\hat{\sigma}_2$	$\hat{\theta}_{\text{Gumb,high}}$
$\widehat{\text{Mean}}$	5	40.2213	45.1986	10.3040	1.8021	9.8055	1.4013	27.4689
	10	40.3078	45.2557	10.3072	1.8000	9.8079	1.3997	27.3196
	20	40.1482	45.0565	10.3095	1.8027	9.8087	1.4021	27.2211
$\widehat{\text{MRB}}$	5	0.0055	0.0044	0.0004	0.0012	0.0006	0.0010	0.0212
	10	0.0077	0.0057	0.0007	0.0000	0.0008	-0.0002	0.0157
	20	0.0037	0.0013	0.0009	0.0015	0.0009	0.0015	0.0120
$\widehat{\text{RMAE}}$	5	0.0584	0.0534	0.0094	0.0255	0.0082	0.0253	0.0837
	10	0.0391	0.0364	0.0056	0.0197	0.0048	0.0200	0.0516
	20	0.0257	0.0247	0.0042	0.0132	0.0036	0.0135	0.0363
$\widehat{\text{RRMSE}}$	5	0.0710	0.0666	0.0115	0.0310	0.0100	0.0308	0.1037
	10	0.0493	0.0473	0.0072	0.0254	0.0064	0.0262	0.0683
	20	0.0327	0.0309	0.0053	0.0168	0.0044	0.0174	0.0472

Table C.10: Performance measures based on 100 MLEs for the parameters of a bivariate compound Poisson model built upon the Gumbel Lévy copula with different dependence strength (Top section: low dependence with true value $\theta_{\text{Gumb,low}} = 0.9317$; Middle section: medium dependence with true value $\theta_{\text{Gumb,med}} = 4.3951$; Bottom section: high dependence with true value $\theta_{\text{Gumb,high}} = 26.8980$), and each for the time horizon of $T = 5, 10, 20$ years. The true marginal frequency parameters are given by $\lambda_1 = 40$, $\lambda_2 = 45$, and the severity distributions by GPD($\xi_1 = 1.5, \beta_1 = 6000$), GPD($\xi_2 = 1.3, \beta_2 = 5000$).

	T	$\hat{\lambda}_1$	$\hat{\lambda}_2$	$\hat{\xi}_1$	$\hat{\beta}_1$	$\hat{\xi}_2$	$\hat{\beta}_2$	$\hat{\theta}_{\text{Gumb,low}}$
$\widehat{\text{Mean}}$	5	39.5429	44.5132	1.4697	6283.6742	1.2766	5199.1889	0.9345
	10	39.6504	44.8438	1.4888	6154.7059	1.2906	5025.6624	0.9325
	20	39.9579	44.9658	1.4890	6084.3774	1.2990	5041.8845	0.9352
$\widehat{\text{MRB}}$	5	-0.0114	-0.0108	-0.0202	0.0473	-0.0180	0.0398	0.0029
	10	-0.0087	-0.0035	-0.0074	0.0258	-0.0072	0.0051	0.0008
	20	-0.0011	-0.0008	-0.0073	0.0141	-0.0007	0.0084	0.0037
$\widehat{\text{RMAE}}$	5	0.0594	0.0495	0.0785	0.1307	0.0878	0.1436	0.0934
	10	0.0397	0.0360	0.0545	0.0804	0.0590	0.0860	0.0628
	20	0.0290	0.0254	0.0441	0.0614	0.0356	0.0518	0.0367
$\widehat{\text{RRMSE}}$	5	0.0742	0.0634	0.1038	0.1691	0.1177	0.1849	0.1134
	10	0.0490	0.0477	0.0713	0.1106	0.0763	0.1044	0.0757
	20	0.0368	0.0324	0.0535	0.0795	0.0469	0.0652	0.0452
	T	$\hat{\lambda}_1$	$\hat{\lambda}_2$	$\hat{\xi}_1$	$\hat{\beta}_1$	$\hat{\xi}_2$	$\hat{\beta}_2$	$\hat{\theta}_{\text{Gumb,med}}$
$\widehat{\text{Mean}}$	5	40.0266	45.2881	1.4884	6080.6771	1.2812	4999.4044	4.3864
	10	39.8682	44.9094	1.4887	6081.3711	1.2816	5033.7814	4.3765
	20	39.9360	44.9620	1.4951	6068.2213	1.2934	5015.7577	4.3901
$\widehat{\text{MRB}}$	5	0.0007	0.0064	-0.0077	0.0134	-0.0144	-0.0001	-0.0020
	10	-0.0033	-0.0020	-0.0075	0.0136	-0.0141	0.0068	-0.0042
	20	-0.0016	-0.0008	-0.0032	0.0114	-0.0050	0.0032	-0.0011
$\widehat{\text{RMAE}}$	5	0.0518	0.0529	0.0644	0.1009	0.0650	0.0935	0.0649
	10	0.0369	0.0364	0.0443	0.0839	0.0448	0.0754	0.0482
	20	0.0268	0.0248	0.0307	0.0569	0.0305	0.0480	0.0354
$\widehat{\text{RRMSE}}$	5	0.0628	0.0652	0.0788	0.1262	0.0789	0.1132	0.0811
	10	0.0443	0.0444	0.0575	0.1040	0.0558	0.0945	0.0625
	20	0.0324	0.0311	0.0375	0.0699	0.0389	0.0612	0.0440
	T	$\hat{\lambda}_1$	$\hat{\lambda}_2$	$\hat{\xi}_1$	$\hat{\beta}_1$	$\hat{\xi}_2$	$\hat{\beta}_2$	$\hat{\theta}_{\text{Gumb,high}}$
$\widehat{\text{Mean}}$	5	40.0846	45.1240	1.4894	6022.6879	1.2887	5027.8557	27.3753
	10	40.0602	44.9990	1.4913	5989.0825	1.2897	5023.4766	26.9650
	20	39.9138	44.9331	1.4947	6002.2036	1.2947	4996.1381	26.9584
$\widehat{\text{MRB}}$	5	0.0021	0.0028	-0.0071	0.0038	-0.0087	0.0056	0.0177
	10	0.0015	-0.0000	-0.0058	-0.0018	-0.0079	0.0047	0.0025
	20	-0.0022	-0.0015	-0.0035	0.0004	-0.0041	-0.0008	0.0022
$\widehat{\text{RMAE}}$	5	0.0526	0.0520	0.0692	0.1001	0.0682	0.0766	0.0857
	10	0.0359	0.0359	0.0454	0.0747	0.0452	0.0647	0.0559
	20	0.0272	0.0276	0.0291	0.0456	0.0290	0.0416	0.0374
$\widehat{\text{RRMSE}}$	5	0.0637	0.0623	0.0856	0.1159	0.0839	0.0946	0.1034
	10	0.0462	0.0457	0.0583	0.0926	0.0577	0.0806	0.0701
	20	0.0344	0.0348	0.0375	0.0573	0.0377	0.0535	0.0461

Table C.11: Performance measures based on 100 MLEs for the parameters of a bivariate compound Poisson model built upon the Gumbel Lévy copula with different dependence strength (Top section: low dependence with true value $\theta_{\text{Gumb,low}} = 0.9317$; Middle section: medium dependence with true value $\theta_{\text{Gumb,med}} = 4.3951$; Bottom section: high dependence with true value $\theta_{\text{Gumb,high}} = 26.8980$), and each for the time horizon of $T = 5, 10, 20$ years. The true marginal frequency parameters are given by $\lambda_1 = 40$, $\lambda_2 = 45$, and the severity distributions by GPD($\xi_1 = 0.9, \beta_1 = 6000$), GPD($\xi_2 = 0.9, \beta_2 = 6000$).

	T	$\hat{\lambda}_1$	$\hat{\lambda}_2$	$\hat{\xi}_1$	$\hat{\beta}_1$	$\hat{\xi}_2$	$\hat{\beta}_2$	$\hat{\theta}_{\text{Gumb,low}}$
$\widehat{\text{Mean}}$	5	39.8841	45.0753	0.9180	5981.8488	0.9020	6082.4192	0.9447
	10	39.8080	45.0409	0.8995	6026.0918	0.8876	6097.8865	0.9333
	20	39.8961	45.1428	0.9043	6033.0511	0.8984	6005.1707	0.9358
$\widehat{\text{MRB}}$	5	-0.0029	0.0017	0.0200	-0.0030	0.0022	0.0137	0.0139
	10	-0.0048	0.0009	-0.0006	0.0043	-0.0137	0.0163	0.0017
	20	-0.0026	0.0032	0.0048	0.0055	-0.0018	0.0009	0.0044
$\widehat{\text{RMAE}}$	5	0.0545	0.0540	0.1129	0.1002	0.0892	0.1014	0.0990
	10	0.0363	0.0402	0.0734	0.0732	0.0722	0.0752	0.0669
	20	0.0267	0.0277	0.0505	0.0563	0.0466	0.0479	0.0455
$\widehat{\text{RRMSE}}$	5	0.0681	0.0676	0.1405	0.1270	0.1167	0.1245	0.1223
	10	0.0456	0.0477	0.0931	0.0983	0.0875	0.0922	0.0816
	20	0.0324	0.0338	0.0632	0.0695	0.0560	0.0586	0.0563
	T	$\hat{\lambda}_1$	$\hat{\lambda}_2$	$\hat{\xi}_1$	$\hat{\beta}_1$	$\hat{\xi}_2$	$\hat{\beta}_2$	$\hat{\theta}_{\text{Gumb,med}}$
$\widehat{\text{Mean}}$	5	40.1506	45.4965	0.8863	6193.1889	0.8920	6018.4122	4.4285
	10	40.0179	45.3504	0.8950	6097.5172	0.9039	5915.6334	4.4074
	20	40.1153	45.0880	0.8974	6018.5176	0.9025	5960.7398	4.4089
$\widehat{\text{MRB}}$	5	0.0038	0.0110	-0.0152	0.0322	-0.0089	0.0031	0.0076
	10	0.0004	0.0078	-0.0056	0.0163	0.0043	-0.0141	0.0028
	20	0.0029	0.0020	-0.0029	0.0031	0.0027	-0.0065	0.0031
$\widehat{\text{RMAE}}$	5	0.0539	0.0502	0.0774	0.0971	0.0788	0.0972	0.0627
	10	0.0430	0.0353	0.0574	0.0697	0.0555	0.0696	0.0510
	20	0.0288	0.0250	0.0421	0.0466	0.0386	0.0494	0.0369
$\widehat{\text{RRMSE}}$	5	0.0692	0.0658	0.0957	0.1174	0.0995	0.1179	0.0803
	10	0.0538	0.0467	0.0734	0.0871	0.0689	0.0834	0.0619
	20	0.0365	0.0333	0.0528	0.0614	0.0496	0.0611	0.0465
	T	$\hat{\lambda}_1$	$\hat{\lambda}_2$	$\hat{\xi}_1$	$\hat{\beta}_1$	$\hat{\xi}_2$	$\hat{\beta}_2$	$\hat{\theta}_{\text{Gumb,high}}$
$\widehat{\text{Mean}}$	5	39.7476	44.5191	0.8940	5991.1450	0.8921	6062.3819	27.0920
	10	39.9985	44.8812	0.8948	6002.7881	0.8936	6050.9067	26.9045
	20	39.8976	44.8008	0.8947	5996.2052	0.8940	6022.5115	26.7872
$\widehat{\text{MRB}}$	5	-0.0063	-0.0107	-0.0066	-0.0015	-0.0088	0.0104	0.0072
	10	-0.0000	-0.0026	-0.0058	0.0005	-0.0071	0.0085	0.0002
	20	-0.0026	-0.0044	-0.0058	-0.0006	-0.0067	0.0038	-0.0041
$\widehat{\text{RMAE}}$	5	0.0534	0.0514	0.0700	0.0745	0.0690	0.0845	0.0765
	10	0.0326	0.0326	0.0529	0.0514	0.0544	0.0584	0.0560
	20	0.0228	0.0234	0.0398	0.0372	0.0400	0.0414	0.0403
$\widehat{\text{RRMSE}}$	5	0.0664	0.0626	0.0853	0.0969	0.0839	0.1051	0.1053
	10	0.0402	0.0408	0.0669	0.0647	0.0682	0.0740	0.0715
	20	0.0272	0.0288	0.0489	0.0481	0.0488	0.0537	0.0501

Table C.12: Performance measures based on 100 MLEs for the parameters of a bivariate compound Poisson model built upon the Lévy copula in Example 3.6 (2) with different dependence strength (Top section: low dependence with true value $\theta_{\text{Arch,low}} = 0.0185$; Middle section: medium dependence with true value $\theta_{\text{Arch,med}} = 0.0377$; Bottom section: high dependence with true value $\theta_{\text{Arch,high}} = 0.1746$), and each for the time horizon of $T = 5, 10, 20$ years. The true marginal frequency parameters are given by $\lambda_1 = 40$, $\lambda_2 = 45$, and the severity distributions by Weib($a_1 = 0.16, b_1 = 4000$), Weib($a_2 = 0.19, b_2 = 5000$).

	T	$\hat{\lambda}_1$	$\hat{\lambda}_2$	$\hat{a}_1$	$\hat{b}_1$	$\hat{a}_2$	$\hat{b}_2$	$\hat{\theta}_{\text{Arch,low}}$
$\widehat{\text{Mean}}$	5	39.6449	45.0026	0.1601	4157.9074	0.1908	5080.3731	0.0186
	10	39.7272	45.2021	0.1603	4002.9639	0.1904	5110.5558	0.0186
	20	40.1143	45.1632	0.1597	3991.4773	0.1898	4985.5159	0.0185
$\widehat{\text{MRB}}$	5	-0.0089	0.0001	0.0004	0.0395	0.0040	0.0161	0.0073
	10	-0.0068	0.0045	0.0021	0.0007	0.0020	0.0221	0.0030
	20	0.0029	0.0036	-0.0021	-0.0021	-0.0013	-0.0029	-0.0017
$\widehat{\text{RMAE}}$	5	0.0524	0.0532	0.0460	0.3468	0.0391	0.2580	0.0387
	10	0.0415	0.0373	0.0320	0.2533	0.0269	0.1962	0.0283
	20	0.0278	0.0270	0.0220	0.1707	0.0178	0.1374	0.0200
$\widehat{\text{RRMSE}}$	5	0.0663	0.0661	0.0593	0.4414	0.0496	0.3225	0.0493
	10	0.0504	0.0476	0.0389	0.3182	0.0353	0.2780	0.0348
	20	0.0353	0.0341	0.0289	0.2153	0.0222	0.1825	0.0250
	T	$\hat{\lambda}_1$	$\hat{\lambda}_2$	$\hat{a}_1$	$\hat{b}_1$	$\hat{a}_2$	$\hat{b}_2$	$\hat{\theta}_{\text{Arch,med}}$
$\widehat{\text{Mean}}$	5	39.9833	44.5070	0.1607	4231.5093	0.1922	5278.5295	0.0382
	10	39.9394	44.7173	0.1603	4100.0270	0.1912	5213.6351	0.0380
	20	40.0007	44.7869	0.1600	4010.7836	0.1906	4992.8627	0.0379
$\widehat{\text{MRB}}$	5	-0.0004	-0.0110	0.0043	0.0579	0.0114	0.0557	0.0126
	10	-0.0015	-0.0063	0.0019	0.0250	0.0062	0.0427	0.0065
	20	0.0000	-0.0047	0.0002	0.0027	0.0029	-0.0014	0.0059
$\widehat{\text{RMAE}}$	5	0.0460	0.0478	0.0398	0.3000	0.0367	0.2645	0.0480
	10	0.0310	0.0319	0.0299	0.2229	0.0244	0.1793	0.0293
	20	0.0240	0.0244	0.0208	0.1391	0.0181	0.1221	0.0224
$\widehat{\text{RRMSE}}$	5	0.0578	0.0589	0.0518	0.3885	0.0470	0.3513	0.0579
	10	0.0390	0.0392	0.0375	0.2752	0.0298	0.2289	0.0382
	20	0.0305	0.0309	0.0258	0.1737	0.0227	0.1597	0.0286
	T	$\hat{\lambda}_1$	$\hat{\lambda}_2$	$\hat{a}_1$	$\hat{b}_1$	$\hat{a}_2$	$\hat{b}_2$	$\hat{\theta}_{\text{Arch,high}}$
$\widehat{\text{Mean}}$	5	40.0766	45.0151	0.1605	4225.7601	0.1911	5132.7539	0.1749
	10	40.2806	45.3051	0.1604	4108.0894	0.1904	4959.7071	0.1741
	20	40.1242	45.1105	0.1601	4002.7443	0.1902	4953.2431	0.1739
$\widehat{\text{MRB}}$	5	0.0019	0.0003	0.0033	0.0564	0.0056	0.0266	0.0019
	10	0.0070	0.0068	0.0023	0.0270	0.0023	-0.0081	-0.0026
	20	0.0031	0.0025	0.0004	0.0007	0.0010	-0.0094	-0.0038
$\widehat{\text{RMAE}}$	5	0.0596	0.0573	0.0390	0.2936	0.0353	0.2384	0.0640
	10	0.0366	0.0336	0.0306	0.2193	0.0257	0.1627	0.0445
	20	0.0277	0.0255	0.0214	0.1487	0.0202	0.1069	0.0329
$\widehat{\text{RRMSE}}$	5	0.0720	0.0712	0.0485	0.3755	0.0457	0.3148	0.0793
	10	0.0466	0.0448	0.0378	0.2731	0.0323	0.2035	0.0553
	20	0.0336	0.0327	0.0265	0.1862	0.0243	0.1357	0.0408

Table C.13: Performance measures based on 100 MLEs for the parameters of a bivariate compound Poisson model built upon the Lévy copula in Example 3.6 (2) with different dependence strength (Top section: low dependence with true value $\theta_{\text{Arch,low}} = 0.0185$; Middle section: medium dependence with true value $\theta_{\text{Arch,med}} = 0.0377$; Bottom section: high dependence with true value $\theta_{\text{Arch,high}} = 0.1746$), and each for the time horizon of $T = 5, 10, 20$ years. The true marginal frequency parameters are given by $\lambda_1 = 40$, $\lambda_2 = 45$, and the severity distributions by $\mathcal{LN}(\mu_1 = 10.3, \sigma_1 = 1.8)$, $\mathcal{LN}(\mu_2 = 9.8, \sigma_2 = 1.4)$.

	T	$\hat{\lambda}_1$	$\hat{\lambda}_2$	$\hat{\mu}_1$	$\hat{\sigma}_1$	$\hat{\mu}_2$	$\hat{\sigma}_2$	$\hat{\theta}_{\text{Arch,low}}$
$\widehat{\text{Mean}}$	5	40.4752	44.5744	10.2871	1.8019	9.8001	1.3982	0.0185
	10	40.1440	44.6219	10.2922	1.7962	9.8064	1.3903	0.0186
	20	40.1887	44.9048	10.2933	1.7965	9.8058	1.3971	0.0185
$\widehat{\text{MRB}}$	5	0.0119	-0.0095	-0.0013	0.0011	0.0000	-0.0013	0.0021
	10	0.0036	-0.0084	-0.0008	-0.0021	0.0007	-0.0069	0.0023
	20	0.0047	-0.0021	-0.0007	-0.0019	0.0006	-0.0021	-0.0010
$\widehat{\text{RMAE}}$	5	0.0712	0.0469	0.0095	0.0362	0.0073	0.0397	0.0412
	10	0.0436	0.0353	0.0058	0.0278	0.0058	0.0255	0.0282
	20	0.0312	0.0254	0.0044	0.0202	0.0038	0.0180	0.0207
$\widehat{\text{RRMSE}}$	5	0.0880	0.0595	0.0122	0.0436	0.0090	0.0496	0.0498
	10	0.0540	0.0445	0.0075	0.0345	0.0071	0.0314	0.0348
	20	0.0377	0.0322	0.0055	0.0256	0.0046	0.0219	0.0260
	T	$\hat{\lambda}_1$	$\hat{\lambda}_2$	$\hat{\mu}_1$	$\hat{\sigma}_1$	$\hat{\mu}_2$	$\hat{\sigma}_2$	$\hat{\theta}_{\text{Arch,med}}$
$\widehat{\text{Mean}}$	5	40.2352	45.5970	10.2898	1.7950	9.8029	1.3991	0.0375
	10	40.2857	45.2729	10.2990	1.7975	9.8029	1.4002	0.0375
	20	40.0939	45.0183	10.2991	1.7997	9.8024	1.4008	0.0377
$\widehat{\text{MRB}}$	5	0.0059	0.0133	-0.0010	-0.0028	0.0003	-0.0007	-0.0050
	10	0.0071	0.0061	-0.0001	-0.0014	0.0003	0.0001	-0.0057
	20	0.0023	0.0004	-0.0001	-0.0002	0.0002	0.0005	-0.0012
$\widehat{\text{RMAE}}$	5	0.0531	0.0436	0.0076	0.0383	0.0062	0.0292	0.0461
	10	0.0376	0.0329	0.0051	0.0258	0.0042	0.0230	0.0334
	20	0.0266	0.0217	0.0037	0.0182	0.0032	0.0182	0.0234
$\widehat{\text{RRMSE}}$	5	0.0675	0.0557	0.0094	0.0453	0.0081	0.0371	0.0585
	10	0.0483	0.0426	0.0064	0.0327	0.0055	0.0283	0.0419
	20	0.0331	0.0278	0.0048	0.0225	0.0042	0.0222	0.0298
	T	$\hat{\lambda}_1$	$\hat{\lambda}_2$	$\hat{\mu}_1$	$\hat{\sigma}_1$	$\hat{\mu}_2$	$\hat{\sigma}_2$	$\hat{\theta}_{\text{Arch,high}}$
$\widehat{\text{Mean}}$	5	39.8654	44.9796	10.3113	1.7785	9.7985	1.3988	0.1757
	10	39.8484	44.9057	10.3054	1.7954	9.7995	1.3993	0.1757
	20	39.8770	44.8999	10.3038	1.7967	9.7966	1.3984	0.1757
$\widehat{\text{MRB}}$	5	-0.0034	-0.0005	0.0011	-0.0119	-0.0002	-0.0009	0.0062
	10	-0.0038	-0.0021	0.0005	-0.0025	-0.0001	-0.0005	0.0066
	20	-0.0031	-0.0022	0.0004	-0.0018	-0.0003	-0.0011	0.0063
$\widehat{\text{RMAE}}$	5	0.0578	0.0569	0.0074	0.0331	0.0059	0.0340	0.0670
	10	0.0423	0.0390	0.0054	0.0224	0.0040	0.0233	0.0484
	20	0.0270	0.0256	0.0038	0.0155	0.0029	0.0162	0.0328
$\widehat{\text{RRMSE}}$	5	0.0709	0.0686	0.0094	0.0423	0.0072	0.0431	0.0817
	10	0.0526	0.0490	0.0069	0.0277	0.0050	0.0295	0.0609
	20	0.0354	0.0327	0.0047	0.0192	0.0036	0.0204	0.0420

Table C.14: Performance measures based on 100 MLEs for the parameters of a bivariate compound Poisson model built upon the Lévy copula in Example 3.6 (2) with different dependence strength (Top section: low dependence with true value $\theta_{\text{Arch,low}} = 0.0185$; Middle section: medium dependence with true value $\theta_{\text{Arch,med}} = 0.0377$; Bottom section: high dependence with true value $\theta_{\text{Arch,high}} = 0.1746$), and each for the time horizon of $T = 5, 10, 20$ years. The true marginal frequency parameters are given by $\lambda_1 = 40$, $\lambda_2 = 45$, and the severity distributions by GPD($\xi_1 = 1.5, \beta_1 = 6000$), GPD($\xi_2 = 1.3, \beta_2 = 5000$).

	T	$\hat{\lambda}_1$	$\hat{\lambda}_2$	$\hat{\xi}_1$	$\hat{\beta}_1$	$\hat{\xi}_2$	$\hat{\beta}_2$	$\hat{\theta}_{\text{Arch,low}}$
	5	40.0761	45.3860	1.4976	6090.5952	1.2755	5144.7178	0.0185
$\widehat{\text{Mean}}$	10	39.9029	45.0736	1.5060	6036.8552	1.2914	5036.5174	0.0185
	20	39.9969	45.0559	1.5040	6025.7539	1.2990	5002.2732	0.0185
	5	0.0019	0.0086	-0.0016	0.0151	-0.0189	0.0289	-0.0026
$\widehat{\text{MRB}}$	10	-0.0024	0.0016	0.0040	0.0061	-0.0066	0.0073	0.0011
	20	-0.0001	0.0012	0.0026	0.0043	-0.0008	0.0005	0.0005
	5	0.0540	0.0561	0.0983	0.1289	0.1030	0.1211	0.0480
$\widehat{\text{RMAE}}$	10	0.0395	0.0405	0.0673	0.0886	0.0688	0.0822	0.0306
	20	0.0297	0.0264	0.0446	0.0587	0.0502	0.0580	0.0194
	5	0.0726	0.0713	0.1211	0.1531	0.1307	0.1481	0.0595
$\widehat{\text{RRMSE}}$	10	0.0525	0.0508	0.0810	0.1121	0.0903	0.1018	0.0372
	20	0.0375	0.0336	0.0573	0.0731	0.0627	0.0732	0.0259
	T	$\hat{\lambda}_1$	$\hat{\lambda}_2$	$\hat{\xi}_1$	$\hat{\beta}_1$	$\hat{\xi}_2$	$\hat{\beta}_2$	$\hat{\theta}_{\text{Arch,med}}$
	5	40.0785	45.0461	1.4893	6114.3703	1.2846	5012.7072	0.0379
$\widehat{\text{Mean}}$	10	40.1351	45.0233	1.4911	6033.6682	1.2966	4991.4238	0.0378
	20	40.0593	44.9183	1.5029	6001.6752	1.2962	5025.8983	0.0378
	5	0.0020	0.0010	-0.0071	0.0191	-0.0119	0.0025	0.0049
$\widehat{\text{MRB}}$	10	0.0034	0.0005	-0.0059	0.0056	-0.0026	-0.0017	0.0020
	20	0.0015	-0.0018	0.0019	0.0003	-0.0029	0.0052	0.0012
	5	0.0536	0.0462	0.0820	0.1154	0.0829	0.1106	0.0488
$\widehat{\text{RMAE}}$	10	0.0346	0.0331	0.0553	0.0846	0.0649	0.0709	0.0346
	20	0.0253	0.0252	0.0365	0.0496	0.0448	0.0573	0.0258
	5	0.0652	0.0567	0.1032	0.1428	0.1056	0.1437	0.0628
$\widehat{\text{RRMSE}}$	10	0.0444	0.0413	0.0688	0.1038	0.0809	0.0915	0.0422
	20	0.0313	0.0304	0.0453	0.0630	0.0565	0.0696	0.0311
	T	$\hat{\lambda}_1$	$\hat{\lambda}_2$	$\hat{\xi}_1$	$\hat{\beta}_1$	$\hat{\xi}_2$	$\hat{\beta}_2$	$\hat{\theta}_{\text{Arch,high}}$
	5	40.1190	45.2840	1.5026	6245.7079	1.2914	5085.2836	0.1750
$\widehat{\text{Mean}}$	10	40.0096	45.0911	1.5071	6117.4158	1.2905	5048.0774	0.1750
	20	39.9587	45.0179	1.5091	6047.9958	1.2980	5021.2755	0.1748
	5	0.0030	0.0063	0.0018	0.0410	-0.0066	0.0171	0.0027
$\widehat{\text{MRB}}$	10	0.0002	0.0020	0.0047	0.0196	-0.0073	0.0096	0.0026
	20	-0.0010	0.0004	0.0061	0.0080	-0.0016	0.0043	0.0016
	5	0.0487	0.0470	0.0855	0.1131	0.0788	0.0909	0.0694
$\widehat{\text{RMAE}}$	10	0.0350	0.0330	0.0599	0.0745	0.0535	0.0567	0.0505
	20	0.0271	0.0259	0.0407	0.0471	0.0387	0.0450	0.0380
	5	0.0614	0.0581	0.1027	0.1409	0.0970	0.1174	0.0848
$\widehat{\text{RRMSE}}$	10	0.0416	0.0405	0.0748	0.0935	0.0674	0.0721	0.0608
	20	0.0341	0.0318	0.0521	0.0614	0.0489	0.0574	0.0471

Table C.15: Performance measures based on 100 MLEs for the parameters of a bivariate compound Poisson model built upon the Lévy copula in Example 3.6 (2) with different dependence strength (Top section: low dependence with true value $\theta_{\text{Arch,low}} = 0.0185$; Middle section: medium dependence with true value $\theta_{\text{Arch,med}} = 0.0377$; Bottom section: high dependence with true value $\theta_{\text{Arch,high}} = 0.1746$), and each for the time horizon of $T = 5, 10, 20$ years. The true marginal frequency parameters are given by $\lambda_1 = 40$, $\lambda_2 = 45$, and the severity distributions by GPD($\xi_1 = 0.9, \beta_1 = 6000$), GPD($\xi_2 = 0.9, \beta_2 = 6000$).

	T	$\hat{\lambda}_1$	$\hat{\lambda}_2$	$\hat{\xi}_1$	$\hat{\beta}_1$	$\hat{\xi}_2$	$\hat{\beta}_2$	$\hat{\theta}_{\text{Arch,low}}$
$\widehat{\text{Mean}}$	5	39.9680	45.2377	0.8780	6121.3737	0.8920	6058.0105	0.0185
	10	39.9527	45.5691	0.8904	6074.9159	0.8877	6066.2244	0.0184
	20	39.9034	45.3661	0.8993	5992.0966	0.8915	6030.5225	0.0185
$\widehat{\text{MRB}}$	5	-0.0008	0.0053	-0.0244	0.0202	-0.0089	0.0097	-0.0013
	10	-0.0012	0.0126	-0.0106	0.0125	-0.0137	0.0110	-0.0052
	20	-0.0024	0.0081	-0.0008	-0.0013	-0.0095	0.0051	-0.0023
$\widehat{\text{RMAE}}$	5	0.0487	0.0506	0.1222	0.1035	0.1055	0.0907	0.0375
	10	0.0408	0.0395	0.0824	0.0745	0.0797	0.0697	0.0296
	20	0.0283	0.0265	0.0655	0.0469	0.0529	0.0395	0.0219
$\widehat{\text{RRMSE}}$	5	0.0609	0.0624	0.1549	0.1266	0.1295	0.1115	0.0464
	10	0.0501	0.0485	0.1010	0.0940	0.1002	0.0824	0.0356
	20	0.0350	0.0333	0.0788	0.0588	0.0657	0.0496	0.0266
	T	$\hat{\lambda}_1$	$\hat{\lambda}_2$	$\hat{\xi}_1$	$\hat{\beta}_1$	$\hat{\xi}_2$	$\hat{\beta}_2$	$\hat{\theta}_{\text{Arch,med}}$
$\widehat{\text{Mean}}$	5	39.9966	45.0280	0.8903	5969.0081	0.8713	6203.1752	0.0379
	10	39.8900	44.9662	0.8937	5990.2169	0.8815	6103.9991	0.0379
	20	39.8081	44.8546	0.8977	5994.6307	0.8906	6044.2956	0.0379
$\widehat{\text{MRB}}$	5	-0.0001	0.0006	-0.0108	-0.0052	-0.0319	0.0339	0.0058
	10	-0.0027	-0.0008	-0.0069	-0.0016	-0.0206	0.0173	0.0048
	20	-0.0048	-0.0032	-0.0026	-0.0009	-0.0105	0.0074	0.0054
$\widehat{\text{RMAE}}$	5	0.0479	0.0449	0.1058	0.0911	0.0990	0.1013	0.0467
	10	0.0354	0.0360	0.0797	0.0608	0.0653	0.0703	0.0332
	20	0.0227	0.0250	0.0521	0.0434	0.0549	0.0523	0.0221
$\widehat{\text{RRMSE}}$	5	0.0593	0.0578	0.1339	0.1135	0.1249	0.1281	0.0588
	10	0.0443	0.0455	0.0947	0.0737	0.0837	0.0883	0.0400
	20	0.0286	0.0310	0.0655	0.0545	0.0658	0.0635	0.0288
	T	$\hat{\lambda}_1$	$\hat{\lambda}_2$	$\hat{\xi}_1$	$\hat{\beta}_1$	$\hat{\xi}_2$	$\hat{\beta}_2$	$\hat{\theta}_{\text{Arch,high}}$
$\widehat{\text{Mean}}$	5	40.3363	45.5008	0.9001	6001.3233	0.8904	5994.5807	0.1759
	10	40.3000	45.2965	0.9063	5948.6131	0.9026	5968.4655	0.1752
	20	40.1984	45.1913	0.9096	5955.9315	0.9029	5965.2039	0.1745
$\widehat{\text{MRB}}$	5	0.0084	0.0111	0.0001	0.0002	-0.0106	-0.0009	0.0074
	10	0.0075	0.0066	0.0070	-0.0086	0.0028	-0.0053	0.0036
	20	0.0050	0.0043	0.0107	-0.0073	0.0032	-0.0058	-0.0007
$\widehat{\text{RMAE}}$	5	0.0576	0.0571	0.1059	0.0858	0.1206	0.0961	0.0708
	10	0.0381	0.0369	0.0769	0.0633	0.0777	0.0642	0.0478
	20	0.0263	0.0255	0.0553	0.0472	0.0598	0.0413	0.0323
$\widehat{\text{RRMSE}}$	5	0.0720	0.0719	0.1333	0.1098	0.1492	0.1193	0.0913
	10	0.0472	0.0445	0.0978	0.0783	0.0996	0.0806	0.0592
	20	0.0337	0.0331	0.0676	0.0581	0.0759	0.0514	0.0396

Bibliography

[AT06] Hansjörg Albrecher and Jef L. Teugels. "Exponential behavior in the presence of dependence in risk theory". In: *Journal of Applied Probability* 43.1 (2006), pp. 257–273.

[AB07] Linda Allen and Turan G. Bali. "Cyclicality in catastrophic and operational risk measurements". In: *Journal of Banking & Finance* 31.4 (2007), pp. 1191–1235.

[Ang+09] Carla Angela, Rossella Bisignani, Giovanni Masala, and Marco Micocci. "Advanced operational risk modelling in banks and insurance companies". In: *Investment Management and Financial Innovations* 6.3 (2009), pp. 73–83.

[AK07] Falko Aue and Michael Kalkbrener. *LDA at work*. Working paper. Risk Analytics & Instruments, Risk and Capital Management, Deutsche Bank AG, 2007. URL: `http://kalkbrener.at/Selected_publications_files/AueKalkbrener06.pdf`.

[ABF12] Stefan Aulbach, Verena Bayer, and Michael Falk. "A multivariate piecing-together approach with an application to operational loss data". In: *Bernoulli* 18.2 (2012), pp. 455–475.

[ACW11] Benjamin Avanzi, Luke C. Cassar, and Bernard Wong. "Modelling dependence in insurance claims processes with Lévy copulas". In: *ASTIN Bulletin* 41.2 (2011), pp. 575–609.

[Bad+14] Andrei L. Badescu, Lan Gong, X. Sheldon Lin, and Dameng Tang. "Modeling correlated frequencies with application in operational risk management". In: *The Journal of Operational Risk* 10.1 (2014), pp. 1–43.

[Ban06] Basel Committee on Banking Supervision. *International convergence of capital measurement and capital standards, a revised framework*. Bank for International Settlements, June 2006. URL: `https://www.bis.org/publ/bcbs128.htm`.

[Ban17] Basel Committee on Banking Supervision. *Basel III: finalising post-crisis reforms*. Bank for International Settlements, Dec. 2017. URL: `https://www.bis.org/bcbs/publ/d424.htm`.

[BL07] Ole E. Barndorff-Nielsen and Alexander M. Lindner. "Lévy copulas: dynamics and transforms of Upsilon-type". In: *Scandinavian Journal of Statistics* 34 (2007), pp. 298–316.

[BR80] Ishwar V. Basawa and B.L.S. Prakasa Rao. *Statistical inference for stochastic processes*. Academic Press Inc., 1980.

[BB11] Nicole Bäuerle and Anja Blatter. "Optimal control and dependence modeling of insurance portfolios with Lévy dynamics". In: *Insurance: Mathematics and Economics* 48 (2011), pp. 398–405.

[Bee05] Marco Bee. *Copula-based multivariate models with applications to risk management and insurance*. Working paper. Department of Economics, University of Trento, 2005. URL: `https://papers.ssrn.com/sol3/papers.cfm?abstract_id=795964`.

[BU09] Francesca Biagini and Sascha Ulmer. "Asymptotics for operational risk quantified with expected shortfall". In: *ASTIN Bulletin* 39.2 (2009), pp. 735–752.

[Böc08] Klaus Böcker. "Quantifying risk: modelling and estimation". PhD thesis. Department of Mathematics, Technical University of Munich, 2008. URL: `https://mediatum.ub.tum.de/node?id=676423`.

[BK05] Klaus Böcker and Claudia Klüppelberg. "Operational VaR: a closed-form approximation". In: *RISK Magazine* (2005), pp. 90–93.

[BK08] Klaus Böcker and Claudia Klüppelberg. *Modelling and measuring multivariate operational risk with Lévy copulas*. Working paper. Risk Analytics and Methods, UniCredit Group; Center for Mathematical Sciences, Munich University of Technology, 2008. URL: `https://mediatum.ub.tum.de/node?id=1072574`.

[BK10] Klaus Böcker and Claudia Klüppelberg. "Multivariate models for operational risk". In: *Quantitative Finance* 10.8 (2010), pp. 855–869.

[BS06] Klaus Böcker and Jacob Sprittulla. "Operational VAR: meaningful means". In: *RISK Magazine* (2006), pp. 96–98.

[BCP14] Eike C. Brechmann, Claudia Czado, and Sandra Paterlini. "Flexible dependence modeling of operational risk losses and its impact on total capital requirements". In: *Journal of Banking & Finance* 40 (2014), pp. 271–285.

[BV13] Axel Bücher and Mathias Vetter. "Nonparametric inference on Lévy measures and copulas". In: *The Annals of Statistics* 41.3 (2013), pp. 1485–1515.

[Cha+04] Ariane Chapelle, Yves Crama, Georges Hübner, and Jean-Philippe Peters. *Basel II and operational risk: implications for risk measurement and management in the financial sector*. Conference. National Bank of Belgium conference on "Efficiency and Stability in an Evolving Financial System", May 17–18, 2004. URL: `https://papers.ssrn.com/sol3/papers.cfm?abstract_id=1691592`.

[Cha+08] Ariane Chapelle, Yves Crama, Georges Hübner, and Jean-Philippe Peters. "Practical methods for measuring and managing operational risk in the financial sector: a clinical study". In: *Journal of Banking & Finance* 32 (2008), pp. 1049–1061.

[CEH16] Valérie Chavez-Demoulin, Paul Embrechts, and Marius Hofert. "An extreme value approach for modeling operational risk losses depending on covariates". In: *The Journal of Risk and Insurance* 83.3 (2016), pp. 735–776.

[CEN06] Valérie Chavez-Demoulin, Paul Embrechts, and Johanna Nešlehová. "Quantitative models for operational risk: extremes, dependence and aggregation". In: *Journal of Banking & Finance* 30.10 (2006), pp. 2635–2658.

[CJY11] Anna Chernobai, Philippe Jorion, and Fan Yu. "The determinants of operational risk in U.S. financial institutions". In: *Journal of Financial and Quantitative Analysis* 46.6 (2011), pp. 1683–1725.

[CR04] Annalisa Di Clemente and Claudio Romano. "A copula-extreme value theory approach for modelling operational risk". In: *Operational risk modelling and analysis: theory and practice*. Ed. by Marcelo Cruz. Risk Books, 2004, pp. 189–208.

[CR92] Daren B.H. Cline and Sidney I. Resnick. "Multivariate subexponential distributions". In: *Stochastic Processes and their Applications* 42 (1992), pp. 49–72.

[CT04] Rama Cont and Peter Tankov. *Financial modelling with jump processes*. Chapman & Hall/CRC, 2004.

[Coo18] KPMG International Cooperative. *Basel 4: the way ahead, operational risk, the new standardised approach*. Feb. 21, 2018. URL: `https://assets.kpmg.com/content/dam/kpmg/xx/pdf/2018/02/basel-4-the-way-ahead.pdf` (visited on 04/08/2018).

[CA08] Eric Cope and Gianluca Antonini. "Observed correlations and dependencies among operational losses in the ORX consortium database". In: *The Journal of Operational Risk* 3.4 (2008), pp. 47–74.

[CMM10] Héléne Cossette, Etienne Marceau, and Fouad Marri. "Analysis of ruin measures for the classical compound Poisson risk model with dependence". In: *Scandinavian Actuarial Journal* 3 (2010), pp. 221–245.

[CPS15] Marcelo G. Cruz, Gareth W. Peters, and Pavel V. Shevchenko. *Fundamental aspects of operational risk and insurance analytics: a handbook of operational risk*. John Wiley & Sons, Inc., 2015.

[Cza+12] Claudia Czado, Rainer Kastenmeier, Eike C. Brechmann, and Aleksey Min. "A mixed copula model for insurance claims and claim sizes". In: *Scandinavian Actuarial Journal* 4 (2012), pp. 278–305.

[DOV07] D. J. Daley, Edward Omey, and Rein Vesilo. "The tail behaviour of a random sum of subexponential random variables and vectors". In: *Extremes* 10 (2007), pp. 21–39.

[Deg10] Matthias Degen. "The calculation of minimum regulatory capital using single-loss approximations". In: *The Journal of Operational Risk* 5.4 (2010), pp. 3–17.

[EG80] Paul Embrechts and Charles M. Goldie. "On closure and factorization properties of subexponential and related distributions". In: *Journal of the Australian Mathematical Society* 29.2 (1980), pp. 243–256.

[EKM97] Paul Embrechts, Claudia Klüppelberg, and Thomas Mikosch. *Modelling extremal events for insurance and finance*. Springer-Verlag, 1997.

[EP08] Paul Embrechts and Giovanni Puccetti. "Aggregating risk across matrix structured loss data: the case of operational risk". In: *The Journal of Operational Risk* 3.2 (2008), pp. 29–44.

[EK10] Habib Esmaeili and Claudia Klüppelberg. "Parameter estimation of a bivariate compound Poisson process". In: *Insurance: Mathematics and Economics* 47 (2010), pp. 224–233.

[FVG08] Dean Fantazzini, Luciana Dalla Valle, and Paolo Giudici. "Copulae and operational risks". In: *International Journal of Risk Assessment and Management* 9.3 (2008), pp. 238–257.

[FRS04] Antoine Frachot, Thierry Roncalli, and Eric Salomon. *The correlation problem in operational risk*. Working paper. Groupe de Recherche Opérationnelle, Group Risk Management, Crédit Agricole SA, 2004. URL: `https://papers.ssrn.com/sol3/papers.cfm?abstract_id=1032594`.

[GGS16] Jose Garrido, Christian Genest, and Juliana Schulz. "Generalized linear models for dependent frequency and severity of insurance claims". In: *Insurance: Mathematics and Economics* 70 (2016), pp. 205–215.

[GN06] Jaap Geluk and Kai W. Ng. "Tail behaviour of negatively associated heavy-tailed sums". In: *Journal of Applied Probability* 43.2 (2006), pp. 587–593.

[GT09] Jaap Geluk and Qihe Tang. "Asymptotic tail probabilities of sums of dependent subexponential random variables". In: *Journal of Theoretical Probability* 22 (2009), pp. 871–882.

[GGR95] Christian Genest, Kilani Ghoudi, and Louis-Paul Rivest. "A semiparametric estimation procedure of dependence parameters in multivariate families of distributions". In: *Biometrika* 82.3 (1995), pp. 543–552.

[Gia+08] Rosella Giacometti, Svetlozar Rachev, Anna Chernobai, and Marida Bertocchi. "Aggregation issues in operational risk". In: *The Journal of Operational Risk* 3.3 (2008), pp. 3–23.

[GFA09] Elise Gourier, Walter Farkas, and Donato Abbate. "Operational risk quantification using extreme value theory and copulas: from theory to practice". In: *Journal of Operational Risk* 4.3 (2009), pp. 1–24.

[GH12] Dominique Guégan and Bertrand K. Hassani. *Multivariate VaRs for operational risk capital computation: a vine structure approach*. Working paper. Documents de travail du Centre d'Economie de la Sorbonne, July 17, 2012. URL: `https://www.inderscienceonline.com/doi/full/10.1504/IJRAM.2013.057104`.

[GCX17] Sabyasachi Guharay, K. C. Chang, and Jie Xu. *Flexible estimation of risk metric using copula model for the joint severity-frequency loss framework.* Conference. 20th International Conference on Information Fusion, July 10–13, 2017. URL: `https://ieeexplore.ieee.org/document/8009872/`.

[Ham+16] Julien Hambuckers, Thomas Kneib, Roland Langrock, and A. Sohn. *A Markov-switching generalized additive model for compound Poisson processes, with applications to operational losses models.* Working paper. Georg-August-Universität Göttingen, Bielefeld University, Sept. 21, 2016. URL: `https://www.uni-goettingen.de/de/12_hambuckers_09_2016/549704.html`.

[Jeo+17] Himchan Jeong, Emiliano A. Valdez, Jae Youn Ahn, and Sojung Park. *Generalized linear mixed models for dependent compound risk models.* Working paper. University of Connecticut, Ewha Womans University, Seoul National University, Sept. 15, 2017. URL: `https://ssrn.com/abstract=3045360`.

[Kal+11] Malaak Kallache, Mathieu Vrac, Philippe Naveau, and Paul-Antoine Michelangeli. "Nonstationary probabilistic downscaling of extreme precipitation". In: *Journal of Geophysical Research* 116 (2011).

[KT06] Jan Kallsen and Peter Tankov. "Characterization of dependence of multidimensional Lévy processes using Lévy copulas". In: *Journal of Multivariate Analysis* 97 (2006), pp. 1551–1572.

[Ket12] Paul C. Kettler. *Lévy-copula-driven financial processes.* Working paper. Catríona Research Foundation, 2012. URL: `http://www.paulcarlislekettler.net/docs/LCDFP.pdf`.

[KT08] Bangwon Ko and Qihe Tang. "Sums of dependent nonnegative random variables with subexponential tails". In: *Journal of Applied Probability* 45.1 (2008), pp. 85–94.

[KA09] Dominik Kortschak and Hansjörg Albrecher. "Asymptotic results for the sum of dependent non-identically distributed random variables". In: *Methodology and Computing in Applied Probability* 11.3 (2009), pp. 279–306.

[Krä+13] Nicole Krämer, Eike C. Brechmann, Daniel Silvestrini, and Claudia Czado. "Total loss estimation using copula-based regression models". In: *Insurance: Mathematics and Economics* 53 (2013), pp. 829–839.

[KS97] Uwe Küchler and Michael Sørensen. *Exponential families of stochastic processes.* Springer-Verlag, 1997.

[LDD16] Heping Li, Estelle Deloux, and Laurence Dieulle. "A condition-based maintenance policy for multi-component systems with Lévy copulas dependence". In: *Reliability Engineering and System Safety* 149 (2016), pp. 44–55.

[Li+14a] Jianping Li, Xiaoqian Zhu, Jianming Chen, Lijun Gao, Jichuang Feng, Dengsheng Wu, and Xiaolei Sun. "Operational risk aggregation across business lines based on frequency dependence and loss dependence". In: *Mathematical Problems in Engineering* (2014), Article ID 404208.

[Li+14b] Jianping Li, Xiaoqian Zhu, Yongjia Xie, Jianming Chen, Lijun Gao, Jichuang Feng, and Wujiang Shi. "The mutual-information-based variance-covariance approach: an application to operational risk aggregation in Chinese banking". In: *Journal of Operational Risk* 9.3 (2014), pp. 3–19.

[LM03] Filip Lindskog and Alexander J. Mcneil. "Common Poisson shock models: applications to insurance and credit risk modelling". In: *ASTIN Bulletin* 33.2 (2003), pp. 209–238.

[Lu14] Zhaoyang Lu. "Asymptotic results for over-dispersed operational risk by using the asymptotic expansion method". In: *Journal of Systems Science and Complexity* 3 (2014), pp. 524–536.

[MFE15] Alexander J. McNeil, Rüdiger Frey, and Paul Embrechts. *Quantitative risk management: concepts, techniques and tools*. Revised edition. Princeton University Press, 2015.

[MM13] Thilo Meyer-Brandis and Michael Morgan. "A dynamic Lévy-copula model for the spark spread". In: *Quantitative Energy Finance*. Ed. by Fred Espen Benth, Valery A. Kholodnyi, and Peter Laurence. Springer, 2013, pp. 237–257.

[MVL09] Paul-Antoine Michelangeli, Mathieu Vrac, and Harilaos Loukos. "Probabilistic downscaling approaches: application to wind cumulative distribution functions". In: *Geophysical Research Letters* 36 (2009).

[MPY11] Stefan Mittnik, Sandra Paterlini, and Tina Yener. *Operational-risk dependencies and the determination of risk capital*. Working paper 3. Center for Quantitative Risk Analysis (CEQURA), Department of Statistics, University of Munich, 2011. URL: `https://papers.ssrn.com/sol3/papers.cfm?abstract_id=1905600`.

[MY09] Stefan Mittnik and Tina Yener. *Estimating operational risk capital for correlated, rare events*. Working paper 1. Center for Quantitative Risk Analysis (CEQURA), Department of Statistics, University of Munich, 2009. URL: `http://www.cequra.uni-muenchen.de/download/cequra_wp_01.pdf`.

[Moo11] Imad Moosa. "Operational risk as a function of the state of the economy". In: *Economic Modelling* 28 (2011), pp. 2137–2142.

[Okh16] Ostap Okhrin. "Lévy copulae for financial returns". In: *Dependence Modeling* 4 (2016), pp. 288–305.

[OMS06] Edward Omey, Fermin Mallor, and Javier Santos. "Multivariate subexponential distributions and random sums of random vectors". In: *Advances in Applied Probability* 38.4 (2006), pp. 1028–1046.

[Opd14] John Douglas Opdyke. *Estimating operational risk capital with greater accuracy, precision, and robustness*. Working paper. Operational Risk Modeling, Quantitative Methods Group, GE Capital, 2014. URL: `https://arxiv.org/abs/1406.0389`.

[Opd17] John Douglas Opdyke. *Fast, accurate, straightforward extreme quantiles of compound loss distributions*. Working paper. Operational Risk Modeling, Quantitative Methods Group, GE Capital, 2017. URL: https://arxiv.org/abs/1610.03718.

[PS15] Gareth W. Peters and Pavel V. Shevchenko. *Advances in heavy tailed risk modeling. A handbook of operational risk*. John Wiley & Sons, Inc., 2015.

[PSW09] Gareth W. Peters, Pavel V. Shevchenko, and Mario V. Wüthrich. "Dynamic operational risk: modeling dependence and combining different sources of information". In: *The Journal of Operational Risk* 4.2 (2009), pp. 69–104.

[PG09] Danae Politou and Paolo Giudici. "Modelling operational risk losses with graphical models and copula functions". In: *Methodology and Computing in Applied Probability* 11.1 (2009), pp. 65–93.

[PRT02] Miro R. Powojowski, Diane Reynolds, and Hans J.H. Tuenter. "Dependent events and operational risk". In: *ALGO Research Quarterly* 5.2 (2002), pp. 65–73.

[PP16] Jörg Prokop and Suren Pakhchanyan. *Determinants of internal fraud and improper business or market practices in financial institutions*. Working paper. Department of Business Administration, Economics, Law, Finance, and Banking, University of Oldenburg, 2016.

[Res08] Ganna Reshetar. *Dependence of operational losses and the capital at risk*. Working paper. Swiss Banking Institute, University of Zurich, Jan. 7, 2008. URL: https://papers.ssrn.com/sol3/papers.cfm?abstract_id=1081256.

[Ris18a] Risk.net. *Top 10 op risk losses of 2017: crisis-era fines abate*. Jan. 16, 2018. URL: https://www.risk.net/comment/5384771/top-10-op-risk-losses-of-2017-crisis-era-fines-abate (visited on 03/16/2018).

[Ris18b] Steve Marlin Risk.net. *Op risk modelling to survive move to SMA*. Feb. 8, 2018. URL: https://www.risk.net/risk-management/5403881/op-risk-modelling-to-survive-move-to-sma (visited on 04/08/2018).

[Sch17] Uwe Schmock. *Modelling dependent credit risks with extensions of CreditRisk$^+$ and application to operational risk*. Lecture notes. Institute for Stochastics and Mathematical Methods in Economics, Vienna University of Technology, Mar. 28, 2017. URL: https://fam.tuwien.ac.at/~schmock/notes/ExtensionsCreditRiskPlus.pdf.

[SV14] Peng Shi and Emiliano A. Valdez. "Multivariate negative binomial models for insurance claim counts". In: *Insurance: Mathematics and Economics* 55 (2014), pp. 18–29.

[SFK05] Peter X.-K. Song, Yanquin Fan, and John D. Kalbfleish. "Maximization by parts in likelihood inference". In: *Journal of the American Statistical Association* 100.472 (2005), pp. 1145–1167.

[Swe80] Trevor J. Sweeting. "Uniform asymptotic normality of the maximum likelihood estimator". In: *The Annals of Statistics* 8.6 (1980), pp. 1375–1381.

[Swe83] Trevor J. Sweeting. "On estimator efficiency in stochastic processes". In: *Stochastic Processes and their Applications* 15.1 (1983), pp. 93–98.

[Tan16] Dameng Tang. "Applications of Pascal mixture models to insurance and risk management". PhD thesis. Graduate Department of Statistical Sciences, University of Toronto, 2016. URL: `https://tspace.library.utoronto.ca/bitstream/1807/73218/1/Tang_Dameng_201606_PhD_thesis.pdf`.

[Tan03] Peter Tankov. *Dependence structure of spectrally positive multidimensional Lévy processes*. Working paper. Centre de Mathématiques Appliquées, 2003. URL: `http://www.proba.jussieu.fr/pageperso/tankov/`.

[Tan04] Peter Tankov. "Lévy processes in finance: inverse problems and dependence modelling". PhD thesis. École polytechnique, Mathématiques Appliquées, 2004. URL: `https://pastel.archives-ouvertes.fr/tel-00007944/document`.

[TW12] Bin Tong and Chongfeng Wu. "Asymptotics for operational risk quantified with a spectral risk measure". In: *Journal of Operational Risk* 7.3 (2012), pp. 91–116.

[Val09] Luciana Dalla Valle. "Bayesian copulae distributions, with application to operational risk management". In: *Methodology and Computing in Applied Probability* 11.1 (2009), pp. 95–115.

[Vel12] J. L. van Velsen. *Parameter estimation of a Lévy copula of a discretely observed bivariate compound Poisson process with an application to operational risk modelling*. Working paper. EC Modelling, Duisenberg school of finance, 2012. URL: `https://arxiv.org/pdf/1212.0092.pdf`.

[WSZ16] Wenzhou Wang, Limeng Shi, and Xiaoqian Zhu. "Operational risk aggregation based on business line dependence: a mutual information approach". In: *Discrete Dynamics in Nature and Society* (2016), Article ID 6546318.

[Zou+18] Christina Zou, Marius Pfeuffer, Matthias Fischer, Kristina Dehler, Nicole Derfuss, Benedikt Graswald, Linda Moestel, Jixuan Wang, and Leonie Wicht. *OpVaR: Statistical methods for modeling operational risk*. R package version 1.0.5. 2018. URL: `https://CRAN.R-project.org/package=OpVaR`.