A copula transformation in multivariate mixed discrete-continuous models

Jae Youn Ahn[†1], Sebastian Fuchs[‡2,**], and Rosy Oh[§3,**]

[1]*Ewha Womans University, Republic of Korea*
[2]*Universität Salzburg, Austria*
[3]*Ewha Womans University, Republic of Korea*
[**]*Corresponding authors*

Abstract

Copulas allow a flexible and simultaneous modeling of complicated dependence structures together with various marginal distributions. Especially if the density function can be represented as the product of the marginal density functions and the copula density function, this leads to both an intuitive interpretation of the conditional distribution and convenient estimation procedures. However, this is no longer the case for copula models with mixed discrete and continuous marginal distributions, because the corresponding density function cannot be decomposed so nicely. In this paper, we introduce a copula transformation method that allows to represent the density function of a distribution with mixed discrete and continuous marginals as the product of the marginal probability mass/density functions and the copula density function. With the proposed method, conditional distributions can be described analytically and the computational complexity in the estimation procedure can be reduced depending on the type of copula used.

1 Introduction

Along with random effect methods, copula methods are a widely used tool to model multivariate distributions. In case of both, the univariate marginal distribution functions and the copula associated with a $(d+1)$-dimensional distribution function H, are absolutely continuous, the density h of H fulfills

$$h(x, y_1, \ldots, y_d) = c\big(F(x), G_1(y_1), \ldots, G_d(y_d)\big) f(x) \prod_{i=1}^{d} g_i(y_i), \tag{1}$$

where c denotes the copula density, F and G_i denote the univariate marginal distribution functions and f and g_i denote its corresponding density functions, $i \in \{1, \ldots, d\}$. The copula representation (1) enables the dependence structure to be separated from the marginal distributions. Such a complete separation not only provides a meaningful interpretation of the

[†]jaeyahn@ewha.ac.kr

[‡]sebastian.fuchs@sbg.ac.at

[§]rosy.oh5@gmail.com

model but also allows a convenient estimation [Nelsen, 2006; Joe, 2014; Durante and Sempi, 2015]. However, this convenience is lost when discrete distributions appear in the model. Then, in contrast to (1), the copula representation of the multivariate distribution no longer provides a complete separation of the dependence structure from the marginal distributions in the density function. Thus, the interpretation of the dependence structure becomes difficult [Genest and Nešlehová, 2007], and the traditional statistical estimation procedures cannot be directly applied [Song and Song, 2007]. For instance, when modeling d discrete random variables, the evaluation of the likelihood function requires the calculation of 2^d terms[1] which provoke computational difficulties in the estimation procedure and complicates an interpretation of the dependence structure [Smith and Khaled, 2012; Zilko and Kurowicka, 2016]. Copula methods for mixed - discrete and continuous - marginals, mixed copula models for short, suffer from similar difficulties.

In this paper we investigate mixed copula models with a single discrete and several absolutely continuous variables and mainly focus on problems related to interpretation difficulties of the (conditional) dependence structure and computational difficulties in estimation.

Often the dependence structure in the copula model is explained in terms of conditional distributions. For instance, a wide range of copula families including Archimedean copulas and some elliptical copula families are closed under the operation of conditioning [Mesfioui and Quessy, 2008; Ding, 2016]. For such copula families, the conditional distribution has both an analytical and an intuitive interpretation. However, this convenience is lost in copula models with discrete marginals. We note that the interpretability of conditional distributions when conditioning with respect to a discrete random variable is important in several research areas including case-control studies in medicine [He et al., 2012; de Leon and Wu, 2011] and frequency-severity models in insurance [Czado et al., 2012; Krämer et al., 2013].

A second problem that occurs in mixed copula models is the computational complexity in the calculation of density functions [Kadhem and Nikoloulopoulos, 2019] since statistical estimation procedures require the evaluation of the corresponding joint density function for multiple terms. In the case of implicit copula models, for instance, the calculation of the density function requires numerical integration [Nikoloulopoulos and Karlis, 2009; Kadhem and Nikoloulopoulos, 2019]. Such numerical difficulties may complicate the estimation procedure mainly due to the computational difficulties in calculating the likelihood functions and subsequent derivatives [Song and Song, 2007].

In the present paper we aim at providing a handy representation of the mixed copula model so that both interpretation and calculation of the density function remain intuitive and simple. To this end, we start with a rather naive question: Can we reformulate a mixed copula model with some discrete distribution function F and some absolutely continuous distribution functions $G_1, \ldots, G_n$, and find some closely related distribution whose density satisfies (1)? An answer to that question requires a copula transformation that is presented in Section 2. With the rather appealing form of the density in (1), we expect that the proposed distribution may provide a meaningful interpretation and excellence in computation in the estimation

---

[1] When pair copula construction is adapted, the computational burden of evaluating n-dimensional discrete random variables only requires $2n(n-1)$ terms [Panagiotelis et al., 2012].

procedure. A numerical analysis comparing the mixed copula model with the modified version is presented in Section 3. As an application, we apply the proposed method to the collective risk model, CRM for short, (Section 4) which plays a crucial role in insurance. The CRM models the aggregate claim amount of a portfolio where the number of claims is random. In particular, for the prediction of the fair premium, modeling the dependence structure in the CRM is important. There are several ways to do this, whereby in this paper we focus on two CRMs that were presented in the recent actuarial literature: two part frequency-severity model (see, e.g., Frees et al. [2014]; Shi et al. [2015]; Garrido et al. [2016]; Park et al. [2018]) and copula-based CRM (see, e.g., Cossette et al. [2019]; Oh et al. [2020]). These two methods were developed independently in different mathematical settings which makes the comparison of the two models difficult. However, with the proposed copula transformation applied to the copula-based CRM, we provide an example demonstrating the linkage between the two models.

Throughout the paper we write $\mathbb{I} := [0,1]$, $\mathbb{N}_0 := \mathbb{N} \cup \{0\}$ and let $d \geq 2$ be an integer which will be kept fixed. Bold symbols are used for vectors, e.g., $\mathbf{y} = (y_1, \ldots y_d) \in \mathbb{R}^d$, or vectors of functions, e.g., $\mathbf{f} = (f_1, \ldots, f_d)$. We denote by ζ^d the d-dimensional Lebesgue measure; in case of $d = 1$ we simply write ζ.

2 Transformation of the mixed model

In this section, we consider a multivariate mixed model with a single discrete and several absolutely continuous variables and present a modification of this model that allows both a meaningful interpretation and a convenient estimation.

First, fix some probability space $(\Omega, \mathcal{A}, \mathbb{P})$ and consider a random variable N (on this probability space) with distribution function F such that N follows a discrete distribution on $\mathbb{N}_0$. Additionally, we consider a d-dimensional random vector $\mathbf{Y}$ (on the same probability space) whose margins Y_i, $i \in \{1, \ldots, d\}$, follow an absolutely continuous distribution with ζ–densities g_i and distribution functions G_i, $i \in \{1, \ldots, d\}$. Then, by Sklar's Theorem (see, e.g., **??**), there exists some $(d+1)$-dimensional copula C such that the distribution function H of $(N, \mathbf{Y})$ satisfies

$$H(n, \mathbf{y}) = C\big(F(n), \mathbf{G}(\mathbf{y})\big) \tag{2}$$

for every $(n, \mathbf{y}) \in \mathbb{N}_0 \times \mathbb{R}^d$. Note that the copula C fails to be unique, in general. In the following, we assume that C is absolutely continuous, i.e. there exists some ζ^{d+1}- density c of C. Then, H has $(\mu \otimes \zeta^d)$- density (where μ denotes the counting measure on the power set of $\mathbb{N}_0$) h satisfying

$$
\begin{aligned}
h(n, \mathbf{y}) &= \left(\frac{\partial^d}{\partial \mathbf{y}} C\big(F(n), \mathbf{G}(\mathbf{y})\big) - \frac{\partial^d}{\partial \mathbf{y}} C\big(F(n-1), \mathbf{G}(\mathbf{y})\big) \right) \prod_{i=1}^d g_i(y_i) \\
&= \left(\int_{(F(n-1), F(n)]} c\big(u, \mathbf{G}(\mathbf{y})\big) \, \mathrm{d}\zeta(u) \right) \prod_{i=1}^d g_i(y_i)
\end{aligned}
\tag{3}
$$

for $(\mu \otimes \zeta^d)$–almost all $(n, \mathbf{y}) \in \mathbb{N}_0 \times \mathbb{R}^d$. While a model with a density function fulfilling (1) allows a meaningful interpretation of the underlying dependence structure and the univariate

3

marginal distributions, such a direct interpretation turns out to be difficult in a situation like (3). In addition to that, the estimation in (3) is usually cumbersome since the calculation of the likelihood function can be quite involved and may require numerical integration. Therefore, we are interested in the following naive question (Q1): Can we find a $(\mu \otimes \zeta^d)$- density h^* of the following form

$$(\text{Q1}) \qquad h^*(n, \mathbf{y}) = c\big(F(n), \mathbf{G}(\mathbf{y})\big) \, \mathbb{P}[\{N = n\}] \prod_{i=1}^{d} g_i(y_i)$$

for $(\mu \otimes \zeta^d)$–almost all $(n, \mathbf{y}) \in \mathbb{N}_0 \times \mathbb{R}^d$, and, if yes, what are the distributional properties of h^*?

We additionally focus on the conditional version of the distribution function H of $(N, \mathbf{Y})$ in (2). Then, for $\mathbb{P}_N$–almost every $n \in \mathbb{N}_0$, the conditional joint density $h(.|n)$ of $\mathbf{Y}$ given $N = n$ equals

$$
\begin{aligned}
h(\mathbf{y}|n) &= \frac{1}{\mathbb{P}[\{N = n\}]} \left(\frac{\partial^d}{\partial \mathbf{y}} C\big(F(n), \mathbf{G}(\mathbf{y})\big) - \frac{\partial^d}{\partial \mathbf{y}} C\big(F(n-1), \mathbf{G}(\mathbf{y})\big) \right) \prod_{i=1}^{d} g_i(y_i) \quad (4) \\
&= \frac{1}{\mathbb{P}[\{N = n\}]} \left(\int_{(F(n-1), F(n)]} c\big(u, \mathbf{G}(\mathbf{y})\big) \, \mathrm{d}\zeta(u) \right) \prod_{i=1}^{d} g_i(y_i)
\end{aligned}
$$

for ζ^d–almost all $\mathbf{y} \in \mathbb{R}^d$. In contrast to the case $(N, \mathbf{Y})$ would be absolutely continuous, equation (4) does not lead to a practicable interpretation. Therefore, we are also interested in the following naive question (Q2) related to (Q1): For $\mathbb{P}_N$–almost every $n \in \mathbb{N}_0$, can we find a ζ^d- density $h^*(.|n)$ conditional on $N = n$ of the following form

$$(\text{Q2}) \qquad h^*(\mathbf{y}|n) = c\big(F(n), \mathbf{G}(\mathbf{y})\big) \prod_{i=1}^{d} g_i(y_i)$$

for ζ^d–almost all $\mathbf{y} \in \mathbb{R}^d$, and, if yes, what are the distributional properties of $h^*(.|n)$?

Remark. 2.1. If we assume $c(u, \mathbf{G}(\mathbf{y})) = c(F(n), \mathbf{G}(\mathbf{y}))$ on $(F(n-1), F(n)] \times \mathbb{R}^d$ for all $n \in \mathbb{N}_0$, then the identities $h^*(n, \mathbf{y}) = h(n, \mathbf{y})$ and $h^*(\mathbf{y}|n) = h(\mathbf{y}|n)$ hold for $(\mu \otimes \zeta^d)$–almost all $(n, \mathbf{y}) \in \mathbb{N}_0 \times \mathbb{R}^d$. $\qquad \square$

In the following, we use a generalization of the idea described in Remark 2.1 and construct a 'copula'-density that is, with respect to the first coordinate, partially constant and hence a step function.

2.1 Transformation of the copula

For $\alpha \in (0, 1]$, we define the map $\lceil . \rceil_{\alpha, F} : (0, 1] \to \mathbb{R}$ by letting

$$\lceil u \rceil_{\alpha, F} \; := \; \sum_{n \in \mathbb{N}_0} F_\alpha(n) \, \mathbb{1}_{(F(n-1), F(n)]}(u) \tag{5}$$

4

where $F_\alpha : \mathbb{N}_0 \to \mathbb{I}$ is given by $F_\alpha(n) := (1 - \alpha)\, F(n - 1) + \alpha\, F(n)$, and canonically extend $\lceil . \rceil_{\alpha,F}$ to $\mathbb{I}$ by putting $\lceil 0 \rceil_{\alpha,F} := 0$. Then, $0 \le \lceil u \rceil_{\alpha,F} \le 1$ for all $u \in \mathbb{I}$.

Remark. 2.2. We note that the identity $\lceil u \rceil_{\alpha,F} = (F_\alpha \circ F^\leftarrow)(u)$ holds for all $u \in (0, 1)$ where $F^\leftarrow : (0, 1) \to \mathbb{R}$ denotes the *pseudo inverse* of F given by $F^\leftarrow(u) := \inf\{x \in \mathbb{R} : F(x) \ge u\}$. □

For the copula C with density function c, we further define the map $\mathfrak{c}_{\alpha,F,C} : \mathbb{I} \times \mathbb{I}^d \to \mathbb{R}$ by letting

$$\mathfrak{c}_{\alpha,F,C}(u, \mathbf{v}) := c\big(\lceil u \rceil_{\alpha,F}, \mathbf{v}\big). \tag{6}$$

Then $\mathfrak{c}_{\alpha,F,C}$ is positive, measurable and a ζ^{d+1}-density: Indeed, since $\partial_1 C(u, \mathbf{1}) = 1$ holds for ζ-almost all $u \in \mathbb{I}$ (where $\partial_1 C$ denotes the partial derivative of C with respect to the first coordinate), we obtain

$$\int_{\mathbb{I} \times \mathbb{I}^d} \mathfrak{c}_{\alpha,F,C}(u, \mathbf{v})\, \mathrm{d}\zeta^{d+1}(u, \mathbf{v}) = \int_{\mathbb{I} \times \mathbb{I}^d} c\big(\lceil u \rceil_{\alpha,F}, \mathbf{v}\big)\, \mathrm{d}\zeta^{d+1}(u, \mathbf{v}) = \int_{\mathbb{I}} \partial_1 C(\lceil u \rceil_{\alpha,F}, \mathbf{1})\, \mathrm{d}\zeta(u) = 1.$$

The following result is immediate from Equations (5) and (6):

Corollary. 2.3. *Consider* $\mathbf{v} \in \mathbb{I}^d$.

- *The identity* $\mathfrak{c}_{\alpha,F,C}(u, \mathbf{v}) = c(F_\alpha(n), \mathbf{v})$ *holds for all* $n \in \mathbb{N}_0$ *and all* $u \in (F(n-1), F(n)]$.

- *The map* $\mathfrak{c}_{\alpha,F,C}(., \mathbf{v})$ *is a positive step function.*

The next example, in which we consider a pertubation of the independence copula Π (see, e.g., **??**), illustrates the construction principle and shows that $\mathfrak{c}_{\alpha,F,C}$ fails to be a copula density, in general:

Example. 2.4. For $\theta \in [-1, 1]$, consider the copula $C : \mathbb{I} \times \mathbb{I}^d \to \mathbb{I}$ given by

$$C(u, \mathbf{v}) := \Pi(u, \mathbf{v}) + \theta\, u(1 - u)\, v_1(1 - v_1) \prod_{i=2}^{d} v_i$$

and the distribution function $F : \mathbb{R} \to \mathbb{I}$ given by

$$F(x) := \frac{1}{3}\, \mathbb{1}_{[1,2)}(x) + \frac{2}{3}\, \mathbb{1}_{[2,3)}(x) + \mathbb{1}_{[3,\infty)}(x).$$

Then we have $c(u, \mathbf{v}) = 1 + \theta\, (1 - 2u)(1 - 2v_1)$ for all $(u, \mathbf{v}) \in \mathbb{I} \times \mathbb{I}^d$ and

$$F_\alpha(n) = \begin{cases} \alpha/3 & n = 1; \\ 1/3 + \alpha/3 & n = 2; \\ 2/3 + \alpha/3 & n = 3; \\ 1 & n \ge 4; \end{cases} \quad \text{and} \quad \lceil u \rceil_{\alpha,F} = \begin{cases} 0 & u = 0; \\ \alpha/3 & u \in \left(0, \frac{1}{3}\right]; \\ 1/3 + \alpha/3 & u \in \left(\frac{1}{3}, \frac{2}{3}\right]; \\ 2/3 + \alpha/3 & u \in \left(\frac{2}{3}, 1\right]. \end{cases}$$

5

Thus, the density $\mathfrak{c}_{\alpha,F,C}$ is a step function satisfying

$$\mathfrak{c}_{\alpha,F,C}(u,\mathbf{v}) = 1 + \theta\left(1 - 2v_1\right) \begin{cases} 1 & u = 0; \\ \frac{3-2\alpha}{3} & u \in \left(0, \frac{1}{3}\right]; \\ \frac{1-2\alpha}{3} & u \in \left(\frac{1}{3}, \frac{2}{3}\right]; \\ \frac{-1-2\alpha}{3} & u \in \left(\frac{2}{3}, 1\right], \end{cases}$$

for all $\mathbf{v} \in \mathbb{I}^d$. Since, for every $\alpha \neq 1/2$ and $\theta \neq 0$ and every $v_1 \in (0,1)$,

$$\int_{\mathbb{I}\times[0,v_1]\times\mathbb{I}^{d-1}} \mathfrak{c}_{\alpha,F,C}(s,\mathbf{t})\, \mathrm{d}\zeta^{d+1}(s,\mathbf{t}) = v_1 + \theta\, v_1\,(1-v_1)\left(\frac{1-2\alpha}{3}\right) \neq v_1,$$

$\mathfrak{c}_{\alpha,F,C}$ fails to be a copula density. $\qquad\square$

Obviously, if $\mathfrak{c}_{\alpha,F,C}$ fails to be a copula density, then $\mathfrak{c}_{\alpha,F,C} \neq c$. Note that the converse implication is not true, in general. For instance, take $\alpha = 1/2$ and $\theta \neq 0$ in Example 2.4. Then, it is straightforward to show that $\mathfrak{c}_{\alpha,F,C}$ is a copula density but fails to coincide with c. Nevertheless, there exist copulas satisfying $\mathfrak{c}_{\alpha,F,C} = c$.

Example. 2.5. Consider a d-dimensional copula A with ζ^d-density a and define the map $C : \mathbb{I} \times \mathbb{I}^d \to \mathbb{I}$ by letting $C(u,\mathbf{v}) := u\, A(\mathbf{v})$. Then, by [?, Theorem 6.6.3], C is a $(d+1)$-dimensional copula with ζ^{d+1}-density c satisfying $\mathfrak{c}_{\alpha,F,C}(u,\mathbf{v}) = c(\lceil u \rceil_{\alpha,F}, \mathbf{v}) = a(\mathbf{v}) = c(u,\mathbf{v})$ for all $(u,\mathbf{v}) \in \mathbb{I} \times \mathbb{I}^d$. $\qquad\square$

Since $\mathfrak{c}_{\alpha,F,C}$ is a ζ^{d+1}-density, the map $\mathfrak{C}_{\alpha,F,C} : \mathbb{I} \times \mathbb{I}^d \to \mathbb{R}$ given by

$$\mathfrak{C}_{\alpha,F,C}(u,\mathbf{v}) := \int_{[0,u]\times[\mathbf{0},\mathbf{v}]} \mathfrak{c}_{\alpha,F,C}(s,\mathbf{t})\, \mathrm{d}\zeta^{d+1}(s,\mathbf{t}) \tag{7}$$

is a distribution function on $\mathbb{I}^{d+1}$ following an absolutely continuous distribution. In the next lemma we gather some properties of $\mathfrak{C}_{\alpha,F,C}$ that turn out to be quite useful.

Lemma. 2.6. *The identity*

$$\mathfrak{C}_{\alpha,F,C}(u,\mathbf{v}) = \sum_{k=0}^{n-1} \partial_1 C(F_\alpha(k),\mathbf{v})\, \mathbb{P}[\{N=k\}] + \partial_1 C(F_\alpha(n),\mathbf{v})\left(u - F(n-1)\right)$$

holds for all $n \in \mathbb{N}_0$ and all $(u,\mathbf{v}) \in (F(n-1), F(n)] \times \mathbb{I}^d$. In particular, we have

$$\mathfrak{C}_{\alpha,F,C}(u,\mathbf{1}) = u \quad \text{and} \quad \mathfrak{C}_{\alpha,F,C}(1,\mathbf{v}) = E\left[\partial_1 C\left(F_\alpha(N),\mathbf{v}\right)\right]$$

for all $u \in \mathbb{I}$ and all $\mathbf{v} \in \mathbb{I}^d$.

Proof. For every $n \in \mathbb{N}_0$ and every $u \in (F(n-1), F(n)]$, Corollary 2.3 yields

$$\int_{[0,u]} \mathfrak{c}_{\alpha,F,C}(s,\mathbf{v})\, \mathrm{d}\zeta(s) = \int_{[0,u]} c\left(\lceil s \rceil_{\alpha,F},\mathbf{v}\right)\, \mathrm{d}\zeta(s)$$

6

$$= \sum_{k=0}^{n-1} \int_{(F(k-1),F(k)]} c(F_\alpha(k),\mathbf{v})\,\mathrm{d}\zeta(s) + \int_{(F(n-1),u]} c(F_\alpha(n),\mathbf{v})\,\mathrm{d}\zeta(s)$$

$$= \sum_{k=0}^{n-1} c(F_\alpha(k),\mathbf{v})\big(F(k) - F(k-1)\big) + c(F_\alpha(n),\mathbf{v})\big(u - F(n-1)\big)$$

$$= \sum_{k=0}^{n-1} c(F_\alpha(k),\mathbf{v})\,\mathbb{P}[\{N=k\}] + c(F_\alpha(n),\mathbf{v})\big(u - F(n-1)\big)$$

for ζ^d-almost all $\mathbf{v} \in \mathbb{I}^d$. This proves the assertion. $\qquad\square$

From a probabilistic viewpoint, if a random vector $(U,\mathbf{V})$ is distributed according to $\mathfrak{C}_{\alpha,F,C}$, then it follows from Lemma 2.6 that U is uniformly distributed, i.e.

$$\mathbb{P}[\{U \le u\}] = u$$

for all $u \in \mathbb{I}$, and that the identity

$$\mathbb{P}[\{\mathbf{V} \le \mathbf{v}\}] = E\big[\partial_1 C\big(F_\alpha(N),\mathbf{v}\big)\big]$$

holds for all $\mathbf{v} \in \mathbb{I}^d$.

We illustrate the construction principle by completing Example 2.4.

Example. 2.7. For $\theta \in [-1,1]$, consider the copula C and the distribution function F discussed in Example 2.4. Then

$$\mathfrak{C}_{\alpha,F,C}(u,\mathbf{v}) \;=\; \Pi(u,\mathbf{v}) + \theta\,v_1(1-v_1)\prod_{i=2}^{d} v_i \begin{cases} 0 & u=0 \\ \left(\frac{3-2\alpha}{3}\right)u & u \in \left(0,\frac{1}{3}\right] \\ \frac{2}{9} + \left(\frac{1-2\alpha}{3}\right)u & u \in \left(\frac{1}{3},\frac{2}{3}\right] \\ \frac{6}{9} + \left(\frac{-1-2\alpha}{3}\right)u & u \in \left(\frac{2}{3},1\right] \end{cases}$$

for all $\mathbf{v} \in \mathbb{I}^d$ and hence $\mathfrak{C}_{\alpha,F,C}(u,\mathbf{1}) = u$ for every $u \in \mathbb{I}$ as well as

$$\mathfrak{C}_{\alpha,F,C}(1,\mathbf{v}) = E\big[\partial_1 C\big(F_\alpha(N),\mathbf{v}\big)\big] = \prod_{i=1}^{d} v_i + \theta\,v_1(1-v_1)\prod_{i=2}^{d} v_i \left(\frac{1-2\alpha}{3}\right) \in \mathbb{I}$$

for every $\mathbf{v} \in \mathbb{I}^d$. Since, for every $\alpha \ne 1/2$ and $\theta \ne 0$ and every $v_1 \in (0,1)$

$$\mathfrak{C}_{\alpha,F,C}\big(1,(v_1,\mathbf{1})\big) = v_1 + \theta\,v_1(1-v_1)\left(\frac{1-2\alpha}{3}\right) \ne v_1,$$

$\mathfrak{C}_{\alpha,F,C}$ fails to be a copula. $\qquad\square$

Choosing $\theta = 0$ in the previous example yields $C = \Pi$ and hence $\mathfrak{C}_{\alpha,F,C} = \Pi = C$. A more general result is given by the following example which extends Example 2.5.

Example. 2.8. Consider a d-dimensional copula A with ζ^d-density and the copula $C : \mathbb{I} \times \mathbb{I}^d \to \mathbb{I}$ given by $C(u,\mathbf{v}) := u\,A(\mathbf{v})$. Then, $\mathfrak{C}_{\alpha,F,C} = C$. In particular, $\mathfrak{C}_{\alpha,F,\Pi} = \Pi$. $\qquad\square$

2.2 Transformation of the random vector

Although the copula transformation $\mathfrak{C}_{\alpha,F,C}$ of C fails to be a copula, in general, it is a distribution function on $\mathbb{I}^{d+1}$ whose first coordinate is distributed uniformly. In this subsection we will use this copula transformation in combination with Sklar's theorem to construct a distribution function that helps answering questions (Q1) and (Q2).

For $\alpha \in (0,1]$, the copula C with density function c, the discrete distribution function F and absolutely continuous distribution functions $G_1,\ldots,G_d$, we define the function $\mathfrak{H}_{\alpha,F,\mathbf{G},C}$: $\mathbb{R} \times \mathbb{R}^d \to \mathbb{I}$ by letting

$$\mathfrak{H}_{\alpha,F,\mathbf{G},C}(x,\mathbf{y}) \quad := \quad \mathfrak{C}_{\alpha,F,C}\big(F(x),\mathbf{G}(\mathbf{y})\big).$$

Then it is straightforward to verify that $\mathfrak{H}_{\alpha,F,\mathbf{G},C}$ is a distribution function satisfying

$$\lim_{\mathbf{t}\to\infty} \mathfrak{H}_{\alpha,F,\mathbf{G},C}(x,\mathbf{t}) = F(x) \quad \text{and} \quad \lim_{s\to\infty} \mathfrak{H}_{\alpha,F,\mathbf{G},C}(s,\mathbf{y}) = E\big[\partial_1 C\big(F_\alpha(N),\mathbf{G}(\mathbf{y})\big)\big].$$

The next result is immediate from Corollary 2.3 and solves question (Q1).

Theorem. 2.9. *The $(\mu \otimes \zeta^d)$- density $\mathfrak{h}_{\alpha,F,\mathbf{G},C}$ of $\mathfrak{H}_{\alpha,F,\mathbf{G},C}$ satisfies*

$$\mathfrak{h}_{\alpha,F,\mathbf{G},C}(n,\mathbf{y}) = c\big(F_\alpha(n),\mathbf{G}(\mathbf{y})\big)\,\mathbb{P}[\{N=n\}]\,\prod_{i=1}^{d} g_i(y_i)$$

for $(\mu \otimes \zeta^d)$- almost all $(n,\mathbf{y}) \in \mathbb{N}_0 \times \mathbb{R}^d$.

From a probabilistic viewpoint, if a random vector $(M,\mathbf{T})$ is distributed according to $\mathfrak{H}_{\alpha,F,\mathbf{G},C}$, then Lemma 2.6 shows that M is a random variable whose distribution function F_M equals

$$F_M = F$$

and that the distribution function $F_{\mathbf{T}}$ of $\mathbf{T}$ satisfies

$$F_{\mathbf{T}} = E\big[\partial_1 C\big(F_\alpha(N),\mathbf{G}(.)\big)\big].$$

Note that $F_{T_i} \neq G_i$, $i \in \{1,\ldots,d\}$, in general; see Example 2.7. Now, we answer question (Q2).

Theorem. 2.10. *Consider a random vector $(M,\mathbf{T})$ distributed according to $\mathfrak{H}_{\alpha,F,\mathbf{G},C}$. Then, for $\mathbb{P}_M$- almost every $n \in \mathbb{N}_0$, the conditional joint density $\mathfrak{h}_{\alpha,F,\mathbf{G},C}(.|n)$ of $\mathbf{T}$ given $M = n$ satisfies*

$$\mathfrak{h}_{\alpha,F,\mathbf{G},C}(\mathbf{t}|n) = c\big(F_\alpha(n),\mathbf{G}(\mathbf{t})\big)\,\prod_{i=1}^{d} g_i(t_i) \tag{8}$$

for ζ^d- almost all $\mathbf{t} \in \mathbb{R}^d$.

Remark. 2.11. In the special case $\alpha = 1$, the results in Theorems 2.9 and 2.10 reduce to

$$\mathfrak{h}_{\alpha,F,\mathbf{G},C}(n,\mathbf{y}) = c\big(F(n),\mathbf{G}(\mathbf{y})\big)\,\mathbb{P}[\{N=n\}]\,\prod_{i=1}^{d} g_i(y_i)$$

and

$$\mathfrak{h}_{\alpha,F,\mathbf{G},C}(\mathbf{y}|n) = c\big(F(n), \mathbf{G}(\mathbf{y})\big) \prod_{i=1}^{d} g_i(y_i)$$

for $(\mu \otimes \zeta^d)$–almost all $(n, \mathbf{y}) \in \mathbb{N}_0 \times \mathbb{R}^d$. $\qquad\qquad\qquad\qquad\qquad\qquad\square$

3 Numerical Analysis

In this section, we perform a numerical study illustrating the impact of the copula transformation method suggested in Section 2.1 by measuring the distance between the copula C and its transformed version $\mathfrak{C}_{\alpha,F,C}$.

As distance measure we use the Kullback-Leibler divergence (KL divergence for short); see, e.g., Kullback and Leibler [1951]; Kullback [1997]. For two k-dimensional joint distribution functions P and Q having p and q as probability density functions, the KL divergence from P to Q is defined as

$$D(P,Q) := \int_{\mathbb{R}^k} p(\mathbf{x}) \log \left(\frac{p(\mathbf{x})}{q(\mathbf{x})} \right) \, \mathrm{d}\zeta^k(\mathbf{x}).$$

We note that $D(P,Q) \geq 0$ where equality holds if and only if $P = Q$. In addition, since $D(P,Q) \neq D(Q,P)$, KL divergence fails to be symmetric, in general.

In the first part of our numerical analysis, we consider bivariate copulas C_θ from various parametric copula families (Gaussian, Student t, Clayton and Gumbel) and put

$$P = C_\theta \qquad \text{and} \qquad Q = \mathfrak{C}_{\alpha,F,C_\theta}$$

where F denotes a Poisson distribution with mean ζ. The parameter θ is choosen in such way that it corresponds to a certain value of bivariate Kendall's tau and hence indicates the degree of dependence represented by C_θ. We measure the KL divergence from P to Q under various scenarios which are combinations of $\alpha = 0.25, 0.5, 0.75, 1.0$, $\zeta = 0.1, 0.5, 1.0, 5.0, 10.0$ and θ given in Table 1.

Kendall's tau		-0.8	-0.3	-0.1	0	0.1	0.3	0.8
θ	Gaussian / Student t	-0.951	-0.454	-0.156	0.000	0.156	0.454	0.951
	Clayton	-	-	-	0.000	0.222	0.857	8.000
	Gumbel	-	-	-	1.000	1.111	1.429	5.000

Table 1: Copula parameter θ corresponding to bivariate Kendall's tau

For each scenario, we further calculate the values of Spearman's rho

$$\rho(P) \qquad \text{and} \qquad \rho(Q)$$

to illustrate how the dependence structure of the copula changes with the proposed transformation; recall that Spearman's rho of a bivariate distribution function H with marginals F and G is defined as

$$\rho(H) := 12 \int_{\mathbb{R}^2} F(x)G(y) \, \mathrm{d}H(x,y) - 3$$

9

In the second part of the numerical analysis, we leave the bivariate setting and consider 3-dimensional copulas. Here, we restrict ourselves to positive dependence (symmetric Gaussian and Clayton copulas) whereas the parameter θ is chosen in such a way that it corresponds to a certain value of the bivariate Kendall's tau as in Table 1.

The results of the numerical analysis are summarized in Table 2 to 11:
- Table 2 to 5: KL divergence for Gaussian, Student t, Clayton and Gumbel copula in dimension 2.
- Table 6 to 9: Spearmans's rho for Gaussian, Student t, Clayton and Gumbel copula in dimension 2.
- Table 10 and 11: KL divergence for the 3-dimensional Gaussian and Clayton copula.

From Table 2 and 3, one may observe that, for each α and ζ, the values of KL divergence are symmetric about 0 with respect to θ which is due to the fact that the density functions of two Gaussian copulas (t-copulas) having parameters of opposite signs, $\pm\rho$, are reflection of each other over the horizon line $x = 0.5$. Similarly, from Table 6 and 7, one may also observe that, for each α and ζ, the values $\rho(P)$ and $\rho(Q)$ are symmetric about 0.

As expected from the definition of the transformation $\mathfrak{C}_{\alpha,F,C_\theta}$, we observe that copulas with weaker dependence tend to have smaller KL divergence and smaller discrepancy between $\rho(P)$ and $\rho(Q)$. We also observe that for each θ, as ζ increases, KL divergence decreases and the discrepancy is diminished, which is also an expected result from the definition of the transformation $\mathfrak{C}_{\alpha,F,C_\theta}$. However, one interesting phenomenon discovered in this numerical analysis is that, around $\alpha = 0.5$, we observe the smallest KL divergence for each combination of ζ and θ. Finally, as can be shown in Table 10 and 11, we note that 3-dimensional copulas show similar patterns as in 2-dimensional copulas.

4 Application to Collective Risk Model

In this section, we apply the proposed copula transformation method to the collective risk model (CRM, for short). In the CRM, the *aggregate severity* in insurance portfolio is modelled as the random sum of individual severities. Specifically, for the nonnegative integer valued random variable N and the positive random variables Y_j, the aggregate severity S is defined by

$$S := \begin{cases} \sum_{j=1}^{N} Y_j, & \text{if } N = n \in \mathbb{N}; \\ 0, & \text{if } N = 0. \end{cases}$$

Note that the aggregate severity can be expressed as $S = MN$ where M is the *average severity* given by

$$M := \begin{cases} \dfrac{1}{N} \sum_{j=1}^{N} Y_j, & \text{if } N = n \in \mathbb{N}; \\ 0, & \text{if } N = 0. \end{cases}$$

We first review two CRMs in the insurance literature. The first model that we will revisit is so called the *two part* CRM where dependence between frequency and severity is induced by using the frequency as an explanatory variable of the severities; see, e.g, Frees et al. [2014];

Shi et al. [2015]; Garrido et al. [2016]; Park et al. [2018]. As a result, in this model the distribution of the aggregate severity can be easily determined. However, it is known that the dependence structure in the two part model is quite limited; see, e.g., Liu and Wang [2017]; Shi et al. [2020]. The second model that we will revisit is the *copula-based* CRM, which can cover the full spectrum of dependencies by describing the dependence for the frequency and the individual severities based on copula function [Cossette et al., 2019; Oh et al., 2020]. Note that the description of the aggregate severity, which is the main concern of insurance industry, under the copula-based CRM can be inconvenient as will be explained below. Finally, we apply the proposed copula transformation to the copula based CRM which enables convenient handling of the distribution of the aggregate severity, and we provide an example where the linkage between the two CRMs is demonstrated.

4.1 Two part CRM for frequency and aggregate severity

Model. 4.1. The *two part CRM for frequency and aggregate severity*

$$(N, M)$$

is defined within the framework of the exponential dispersion family (EDF) as follows (see Frees et al. [2014]; Garrido et al. [2016]):
(i) We specify the frequency component N as

$$N \sim F \tag{9}$$

where F can be any discrete distribution function on $\mathbb{N}_0$.
(ii) We specify the conditional distribution of the average severity conditional on the number of claims $N = n \in \mathbb{N}$ as

$$M \big| N = n \overset{\text{i.i.d.}}{\sim} \text{ED}\left(\mu_n, \sigma_n^2\right) \tag{10}$$

where $\text{ED}\left(\mu_n, \sigma_n^2\right)$ is the reproductive exponential dispersion model with mean μ_n and dispersion parameter σ_n^2; see, e.g., Jørgensen [1987]; Jorgensen [1997]. Here, the mean parameter μ_n is implicitly given by $\eta_1\left(\mu_n\right) = \beta_1 + \psi_1(n)$ for properly chosen function ψ_1 and link function η_1, and the dispersion parameter is given by

$$\sigma_n^2 := \frac{\sigma_0^2}{n}. \tag{11}$$

The choice of the dispersion parameter in (11) can be justified by the following distributional assumption on the individual severities

$$Y_1, \ldots, Y_n | N = n \overset{\text{i.i.d.}}{\sim} \text{ED}\left(\mu_n, \sigma_0^2\right) \qquad n \in \mathbb{N} \tag{12}$$

which implies (10) by the convolutionary property of EDF. As a result, one may replace the description of the average severity in Model 4.1 with the description of the individual severities in (12). We call such model as the *two part CRM* for frequency and the individual severities. However, while convenient in many ways, the conditional independence assumption in (12) is a rather restrictive dependence assumption of frequency and individual severities as pointed out in Liu and Wang [2017] and Shi et al. [2020]. Therefore, the two part CRM with

the functional form of the dispersion parameter in (11) can accommodate only restrictive dependence structures of frequency and individual severities.

Alternatively, one may choose a more complicated functional form of σ_n as mentioned in Lee et al. [2019] to consider more general dependence structures in the CRM. Depending on the purpose of the data analysis, one may use, for instance, advanced regression modeling strategies such as non-parametric regression or additive modeling; see, e.g., Hastie and Tibshirani [1990]; Faraway [2005]. However, in such a case the important linkage between the average severity in (10) and the individual severities in (12) is violated, in general.

4.2 The copula-based CRM for the frequency and the individual severities

We now consider the copula-based CRM for the frequency and the individual severities discussed in Cossette et al. [2019] and Oh et al. [2020] where a wider variety of dependence structures is possible depending on the particular choice of the used copula family.

Model. 4.2. The *copula-based CRM* for frequency and individual severities

$$(N, Y_1, Y_2, \ldots, Y_N)$$

is defined as follows:
(i)   We specify the frequency component N as

$$N \sim F \tag{13}$$

where F can be any discrete distribution function on $\mathbb{N}_0$; compare (9).
(ii) We specify the conditional distribution of the vector of individual severities $(Y_1, \cdots, Y_n)$ conditional on the number of claims $N = n \in \mathbb{N}$ as

$$(Y_1, \ldots, Y_n) \big| N = n \sim W_{(n)} \tag{14}$$

for some distribution function $W_{(n)}$ given by

$$W_{(n)}(y_1, \ldots, y_n) := \frac{C_{(n)}\big(F(n), G(y_1), \ldots, G(y_n)\big) - C_{(n)}\big(F(n-1), G(y_1), \ldots, G(y_n)\big)}{\mathbb{P}[\{N = n\}]}$$

with $C_{(n)}$ being an absolutely continuous $(n+1)$-dimensional copula.

In this model, the density function of $(Y_1, Y_2, \cdots, Y_n)$ at point $(y_1, y_2, \cdots, y_n)$ conditional on $N = n \in \mathbb{N}$ satisfies

$$\frac{\partial^n}{\partial y_1 \ldots \partial y_n} W_{(n)}(y_1, \ldots, y_n) = \left(\int_{(F(n-1), F(n)]} c_{(n)}\big(u, G(y_1), \ldots, G(y_n)\big) \, \mathrm{d}\zeta(u) \right) \frac{\prod_{i=1}^n g(y_i)}{\mathbb{P}[\{N = n\}]}. \tag{15}$$

Hence, loosely speaking, the copula-based CRM can be understood as

$$(N, Y_1, \ldots, Y_N) \sim C_{(N)}(F, G, \ldots, G). \tag{16}$$

We refer to Cossette et al. [2019] and Oh et al. [2020] for the natural linkage between (15) and the density function in (16).

The copula-based CRM allows modeling both, the dependence among the individual severities and the dependence between frequency and severities. Hence, in terms of the dependence structure, the copula-based CRM provides a wider range of dependence than the two part CRM. But while the main interest of insurance industry lies in the aggregate severity S, the copula-based CRM does not allow an analytical interpretation of S, in general. For an analysis of S one may consider to use the conditional distribution in (14). However, mainly due to the non-continuous nature of the frequency N, an analytical interpretation of the conditional distribution in (14) or equivalently in (15) is difficult, in general.

4.3 The transformed copula-based CRM for the frequency and the individual severities

In the following, we provide an example where the proposed copula transformation method allows an analytical interpretation of the conditional dependence among severities in (14) as well as the dependence between frequency and individual severities. The following model is a modification of the previous copula-based CRM where the copula C is replaced by its transformed version $\mathfrak{C}_{\alpha,F,C_{(n)}}$.

Model. 4.3. For $\alpha \in (0,1)$, the *transformed copula-based CRM* for

$$(N, Y_1, Y_2, \ldots, Y_N)$$

is defined as follows:

(i) We specify the frequency component N as

$$N \sim F \tag{17}$$

 where F can be any discrete distribution function on $\mathbb{N}_0$; compare (9) and (13).

(ii) We specify the conditional distribution of the vector of individual severities $(Y_1, \ldots, Y_n)$ conditional on the number of claims $N = n \in \mathbb{N}$ as

$$(Y_1, \ldots, Y_n)\big| N = n \sim W_{(n)}^* \tag{18}$$

 for some distribution function $W_{(n)}^*$ given by

$$W_{(n)}^*(y_1, \ldots, y_n) := \frac{\mathfrak{C}_{\alpha,F,C_{(n)}}\big(F(n), G(y_1), \ldots, G(y_n)\big) - \mathfrak{C}_{\alpha,F,C_{(n)}}\big(F(n-1), G(y_1), \ldots, G(y_n)\big)}{\mathbb{P}[\{N = n\}]} \tag{19}$$

 with $C_{(n)}$ being an absolutely continuous $(n+1)$-dimensional copula.

Note that, due to Theorem 2.10, the conditional distribution function in (18) can be represented as

$$W_{(n)}^*(y_1, \ldots, y_n) = \partial_1 C_{(n)}\big(F_\alpha(n), G(y_1), \ldots, G(y_n)\big)$$

and the corresponding density function at $(y_1, y_2, \cdots, y_n)$ conditional on $N = n \in \mathbb{N}$ satisfies

$$\frac{\partial^n}{\partial y_1 \ldots \partial y_n} W_{(n)}^*(y_1, \ldots, y_n) = c_{(n)}\big(F_\alpha(n), G(y_1), \cdots, G(y_n)\big) \prod_{i=1}^n g(y_i) \tag{20}$$

13

Similar to the case of copula-based CRM, the transformed copula-based CRM can be understood as

$$(N, Y_1, \ldots, Y_N) \sim \mathfrak{C}_{\alpha, F, C_{(N)}}(F, G, \ldots, G) \tag{21}$$

Now we are ready to provide an example which shows a link between two part CRM and copula-based CRM. First, we define the following matrices for $\rho_1, \rho_2 \in [-1, 1]$:

(i) For $k \in \mathbb{N}$ and $l \in \{1, 2\}$, define the $k \times k$ matrix $\mathbf{\Sigma}_{\rho_2}^{[k,l]}$ by letting

$$\left[\mathbf{\Sigma}_{\rho_2}^{[k,1]}\right]_{i,j} := \begin{cases} 1, & \text{if } i = j; \\ \rho_2, & \text{if } i \neq j; \end{cases}$$

and

$$\left[\mathbf{\Sigma}_{\rho_2}^{[k,2]}\right]_{i,j} := \begin{cases} 1, & \text{if } i = j; \\ \rho_2^{|i-j|}, & \text{if } i \neq j. \end{cases}$$

(ii) For $k \in \mathbb{N}$ and $l \in \{1, 2\}$, define the $(k+1) \times (k+1)$ matrix $\mathbf{\Sigma}_{\rho_1, \rho_2}^{[k,l]}$ by letting

$$\mathbf{\Sigma}_{\rho_1, \rho_2}^{[k,l]} := \begin{cases} \begin{pmatrix} 1 & \rho_1 (\mathbf{1}_k)^{\mathrm{T}} \\ \rho_1 \mathbf{1}_k & \mathbf{\Sigma}_{\rho_2}^{[k,l]} \end{pmatrix}, & k \in \mathbb{N}; \\ 1, & k = 0; \end{cases}$$

where $\mathbf{1}_k$ is a column vector with entries 1 of length k.

As shown in Oh et al. [2020], the condition

$$(\rho_1, \rho_2) \in \left\{ (\rho_1, \rho_2) \in (-1, 1)^2 \,\middle|\, \rho_1^2 < \rho_2 < 1 \right\} \tag{22}$$

implies positive definiteness of the matrix $\mathbf{\Sigma}_{\rho_1, \rho_2}^{[k,1]}$ for every $k \in \mathbb{N}_0$. Similarly, using the well known result on the Schur complement of a block matrix [Haynsworth, 1968], the matrix $\mathbf{\Sigma}_{\rho_1, \rho_2}^{[k,2]}$ is positive definite for any $\rho_1, \rho_2 \in (-1, 1)$ satisfying

$$1 - \rho_1^2 (k (1 - \rho_2) + 2\rho_2)(1 - \rho_2) > 0. \tag{23}$$

We denote by $C\big(\cdot; \mathbf{\Sigma}_{\rho_1, \rho_2}^{[n,l]}\big)$ the $(n+1)$-dimensional Gaussian copula with correlation matrix $\mathbf{\Sigma}_{\rho_1, \rho_2}^{[n,l]}$ and by $c\big(\cdot; \mathbf{\Sigma}_{\rho_1, \rho_2}^{[n,l]}\big)$ its density.

In the sequel, we consider Model 4.3 for $(N, Y_1, Y_2, \cdots, Y_N)$ assuming a symmetric dependence structure for the individual severities:

Assumption. 4.4.
- ρ_1 and ρ_2 satisfy condition (22).
- $C_{(n)}$ is an $(n+1)$-dimensional Gaussian copula with correlation matrix $\mathbf{\Sigma}_{\rho_1, \rho_2}^{[n,1]}$.
- G is a normal distribution with mean ξ and variance σ^2.

In this case, the conditional density function of $(Y_1, Y_2, \cdots, Y_n)$ conditional on $N = n \in \mathbb{N}$ satisfies

$$\frac{\partial^n}{\partial y_1 \ldots \partial y_n} W_{(n)}^*(y_1, \cdots, y_n) = c_{(n)} \left(F_\alpha(n), \Phi_{\xi, \sigma^2}(y_1), \ldots, \Phi_{\xi, \sigma^2}(y_n); \mathbf{\Sigma}_{\rho_1, \rho_2}^{[n,1]} \right) \prod_{i=1}^n \phi_{\xi, \sigma^2}(y_i) \tag{24}$$

and part i of Lemma A.1 shows that, for $n \in \mathbb{N}$, $(Y_1, \cdots, Y_n)$ conditional on $N = n$ follows a multivariate normal distribution with mean $\left(\xi + \sigma \rho_1 \Phi_{0,1}^{-1}(F_\alpha(n))\right) \mathbf{1}_n$ and covariance matrix

$$\sigma^2 \left(\mathbf{\Sigma}_{\rho_2}^{[n,1]} - \rho_1^2 \boldsymbol{J}_{n \times n} \right)$$

where $\boldsymbol{J}_{n \times n}$ denotes an $n \times n$ matrix with entries 1. By the convolutionary property of the multivariate normal distribution, we then obtain

$$\frac{Y_1 + \cdots + Y_n}{n} \bigg| N = n \sim \mathrm{N}\left(\mu_n, \sigma_n^2\right)$$

where

$$\mu_n = \xi + \sigma \rho_1 \Phi_{0,1}^{-1}(F_\alpha(n)) \quad \text{and} \quad \sigma_n^2 = \frac{1}{n} \sigma^2 \big((n-1)\rho_2 - n\rho_1^2 + 1\big)$$

As a result, the distribution of S can be expressed in closed form which in turn allows a closed form expression for subsequent statistics of S. Specifically, for $s \geq 0$, we have

$$\mathbb{P}\left[\{S \leq s\}\right] = \sum_{n=0}^{\infty} \mathbb{P}\left[\{S \leq s\} | \{N = n\}\right] \mathbb{P}\left[\{N = n\}\right]$$

$$= F(0) + \sum_{n=1}^{\infty} \Phi\left(\frac{s/n - \mu_n}{\sigma_n}\right) \mathbb{P}\left[\{N = n\}\right]$$

Additionally, we obtain

$$E\left[S\right] = E\left[N \, E\left[M | N\right]\right] = \sum_{n=1}^{\infty} n\mu_n \, \mathbb{P}\left[\{N = n\}\right]$$

and

$$\mathrm{Var}\left[S\right] = E\left[N^2 \, \mathrm{Var}\left[M | N\right]\right] + \mathrm{Var}\left[N \, E\left[M | N\right]\right]$$

$$= \sum_{n=1}^{\infty} \left(n^2 \sigma_n^2 + n^2 \mu_n^2\right) \mathbb{P}\left[\{N = n\}\right] - \left(\sum_{n=1}^{\infty} n\mu_n \, \mathbb{P}\left[\{N = n\}\right]\right)^2.$$

Note that, while the dependence structure of the conditional severities in (12) under two part CRM is restricted to conditional independence, the dependence structure of the conditional severities in (19) under transformed copula-based CRM allows more general dependence structures. In Lemma A.1, we also provide the condition where two part CRM and transformed copula-based CRM are equivalent.

Finally, we consider Model 4.3 for $(N, Y_1, Y_2, \cdots, Y_N)$ assuming an autoregressive dependence structure for the individual severities.

Assumption. 4.5.
- F is a discrete distribution function with finite support on $\mathbb{N}_0$ having $\kappa_0 \in \mathbb{N}$ as the essential supremum of F.
- $k = \kappa_0$ satisfies (23).
- $C_{(n)}$ is an $(n+1)$-dimensional Gaussian copula with correlation matrix $\mathbf{\Sigma}_{\rho_1,\rho_2}^{[n,2]}$ for $n \leq \kappa_0$.

- G is a normal distribution with mean ξ and variance σ^2.

Following the same procedure as above for $n \in \mathbb{N}$ yields

$$\frac{Y_1 + \cdots + Y_N}{n} \,\Big|\, N = n \sim \mathrm{N}\left(\mu_n, \sigma_n^2\right) \tag{25}$$

where

$$\mu_n = \xi + \sigma\rho_1\Phi^{-1}(F_\alpha(n)) \quad \text{and} \quad \sigma_n^2 = \frac{\left(1 - n\rho_1^2\right)}{n} + \frac{2}{n^2}\frac{\rho_2^2}{1 - \rho_2^2}\left(\rho_2^{n-1} - 1\right).$$

Again, the distribution of S can be expressed in closed form which in turn allows a closed form expression for subsequent statistics of S.

Acknowledgements

References

Cossette, H., Marceau, E., and Mtalai, I. (2019). Collective risk models with dependence. *Insurance: Mathematics and Economics*, 87:153–168.

Czado, C., Kastenmeier, R., Brechmann, E. C., and Min, A. (2012). A mixed copula model for insurance claims and claim sizes. *Scand. Actuar. J.*, (4):278–305.

de Leon, A. R. and Wu, B. (2011). Copula-based regression models for a bivariate mixed discrete and continuous outcome. *Statistics in Medicine*, 30(2):175–185.

Ding, P. (2016). On the conditional distribution of the multivariate t distribution. *The American Statistician*, 70(3):293–295.

Durante, F. and Sempi, C. (2015). *Principles of copula theory*. CRC press.

Faraway, J. J. (2005). *Extending the linear model with R: generalized linear, mixed effects and nonparametric regression models*. CRC press.

Frees, E. W., Derrig, R. A., and Meyers, G. (2014). *Predictive Modeling Applications in Actuarial Science*, volume 1. Cambridge University Press.

Garrido, J., Genest, C., and Schulz, J. (2016). Generalized linear models for dependent frequency and severity of insurance claims. *Insurance: Mathematics and Economics*, 70:205 – 215.

Genest, C. and Nešlehová, J. (2007). A primer on copulas for count data. *ASTIN Bulletin: The Journal of the IAA*, 37(2):475–515.

Hastie, T. J. and Tibshirani, R. J. (1990). *Generalized additive models*, volume 43. CRC Press.

Haynsworth, E. V. (1968). On the schur complement. Technical report, BASEL UNIV (SWITZERLAND) MATHEMATICS INST.

He, J., Li, H., Edmondson, A. C., Rader, D. J., and Li, M. (2012). A gaussian copula approach for the analysis of secondary phenotypes in case–control genetic association studies. *Biostatistics*, 13(3):497–508.

Joe, H. (2014). *Dependence Modeling with Copulas*. CRC Press.

Jørgensen, B. (1987). Exponential dispersion models. *Journal of the Royal Statistical Society: Series B (Methodological)*, 49(2):127–145.

Jorgensen, B. (1997). *The theory of dispersion models*. CRC Press.

Kadhem, S. H. and Nikoloulopoulos, A. K. (2019). Factor copula models for mixed data. *arXiv preprint arXiv:1907.07395*.

Krämer, N., Brechmann, E. C., Silvestrini, D., and Czado, C. (2013). Total loss estimation using copula-based regression models. *Insurance: Mathematics and Economics*, 53(3):829–839.

Kullback, S. (1997). *Information theory and statistics*. Courier Corporation.

Kullback, S. and Leibler, R. A. (1951). On information and sufficiency. *The annals of mathematical statistics*, 22(1):79–86.

Lee, W., Park, S. C., and Ahn, J. Y. (2019). Investigating dependence between frequency and severity via simple generalized linear models. *Journal of the Korean Statistical Society*, 48(1):13–28.

Liu, H. and Wang, R. (2017). Collective risk models with dependence uncertainty. *ASTIN Bulletin, Forthcoming*.

Mesfioui, M. and Quessy, J.-F. (2008). Dependence structure of conditional archimedean copulas. *Journal of Multivariate Analysis*, 99(3):372–385.

Nelsen, R. B. (2006). *An introduction to copulas*. Springer Series in Statistics. Springer, New York, second edition.

Nikoloulopoulos, A. K. and Karlis, D. (2009). Modeling multivariate count data using copulas. *Communications in Statistics-Simulation and Computation*, 39(1):172–187.

Oh, R., Ahn, J. Y., and Lee, W. (2020). On copula-based collective risk models: from elliptical copulas to vine copulas. *Scandinavian Actuarial Journal*, 0(0):1–33.

Panagiotelis, A., Czado, C., and Joe, H. (2012). Pair copula constructions for multivariate discrete data. *Journal of the American Statistical Association*, 107(499):1063–1072.

Park, S. C., Kim, J. H., and Ahn, J. Y. (2018). Does hunger for bonuses drive the dependence between claim frequency and severity? *Insurance: Mathematics and economics*, 83:32–46.

Shi, P., Feng, X., and Ivantsova, A. (2015). Dependent frequency–severity modeling of insurance claims. *Insurance: Mathematics and Economics*, 64:417–428.

Shi, P., Zhao, Z., et al. (2020). Regression for copula-linked compound distributions with applications in modeling aggregate insurance claims. *Annals of Applied Statistics*, 14(1):357–380.

Smith, M. S. and Khaled, M. A. (2012). Estimation of copula models with discrete margins via bayesian data augmentation. *Journal of the American Statistical Association*, 107(497):290–303.

Song, X.-K. and Song, P. X.-K. (2007). *Correlated data analysis: modeling, analytics, and applications*. Springer Science & Business Media.

Zilko, A. A. and Kurowicka, D. (2016). Copula in a multivariate mixed discrete–continuous model. *Computational Statistics & Data Analysis*, 103:28–55.

A Appendix

Lemma. A.1. *Consider $\alpha \in (0,1)$, a random vector $(N, Y_1, \ldots, Y_N)$ from a transformed copula-based CRM (Model 4.3) satisfying Assumption 4.4. Then, we have following properties:*

i. *For $n \in \mathbb{N}$, consider a random vector $(Z_0, Z_1, \cdots, Z_n)$ following a multivariate normal distribution with mean $\left(0, \xi, \cdots, \xi\right)^{\mathrm{T}}$ and covariance matrix*

$$\mathrm{diag}\left(1, \sigma, \ldots, \sigma\right) \boldsymbol{\Sigma}_{\rho_1, \rho_2}^{[n,1]} \mathrm{diag}\left(1, \sigma, \ldots, \sigma\right)$$

where $\boldsymbol{J}_{n \times n}$ denotes an $n \times n$ matrix with entries 1. Then, the conditional distribution of $(Y_1, \cdots, Y_n)$ conditional on $N = n$ equals the conditional distribution of $(Z_1, \cdots, Z_n)$ conditional on $Z_0 = \Phi_{0,1}^{-1}\left(F_\alpha(n)\right)$ which satisfies

$$\begin{aligned} (Z_1, \cdots, Z_n) \mid Z_0 &= \Phi_{0,1}^{-1}\left(F_\alpha(n)\right) \\ &\sim \mathrm{MVN}\left(\left(\xi + \sigma\rho_1\Phi_{0,1}^{-1}\left(F_\alpha(n)\right)\right)\boldsymbol{1}_n, \sigma^2\left(\boldsymbol{\Sigma}_{\rho_2}^{[n,1]} - \rho_1^2\boldsymbol{J}_{n \times n}\right)\right) \end{aligned} \tag{26}$$

ii. *Consider a random vector $(X, Y_1, \cdots, Y_X)$ from a two part CRM (Model 4.1) and assume that $X \stackrel{\mathrm{d}}{=} N$ shares the same distribution function F with $N \sim F$. If we further assume $\rho_1 = \rho_2^2$ and that μ_n and σ_0^2 in (12) satisfying*

$$\mu_n = \sigma\rho_1\Phi^{-1}\left(F_\alpha(n)\right) \qquad \text{and} \qquad \sigma_0^2 = \sigma^2\left(1 - \rho_1^2\right)$$

for $X = n \in \mathbb{N}$, then the two random vectors $(N, Y_1, \cdots, Y_N)$ and $(X, Y_1, \cdots, Y_X)$ have the same distribution.

Proof.   We first prove part i. By Oh et al. [2020], the matrix $\boldsymbol{\Sigma}_{\rho_1,\rho_2}^{[n,1]}$ is positive definite. Furthermore, since the corresponding marginals and copula function of a multivariate normal distribution are normal distributions and Gaussian copula, respectively, the conditional density function of $(Z_1,\ldots,Z_n)$ at point $(z_1,\ldots,z_n)$ conditional on $Z_0 = z_0$ equals

$$c_{(n)}\left(\Phi_{0,1}(z_0),\Phi_{\xi,\sigma^2}(z_1),\ldots,\Phi_{\xi,\sigma^2}(z_n);\boldsymbol{\Sigma}_{\rho_1,\rho_2}^{[n,1]}\right)\prod_{i=1}^{n}\phi_{\xi,\sigma^2}(z_i). \qquad (27)$$

On the other hand, the conditional distribution of $(Z_1,\cdots,Z_n)$ conditional on $Z_0 = z_0$ satisfies

$$(Z_1,\cdots,Z_n)\mid Z_0 = z_0 \sim \mathrm{MVN}\left((\xi+\rho_1\sigma z_0)\,\mathbf{1}_n,\sigma^2\left(\boldsymbol{\Sigma}_{\rho_2}^{[n,1]}-\rho_1^2\boldsymbol{J}_{n\times n}\right)\right) \qquad (28)$$

Since (27) and (28) describe the same distribution, the conditional distribution of $(Y_1,\ldots,Y_n)$ conditional on $N = n$ and the conditional distribution of $(Z_1,\ldots,Z_n)$ conditional on $Z_0 = \Phi_{0,1}^{-1}(F_\alpha(n))$ coincide which follows from (24). This proves part i. The proof of part ii is immediate from part i. $\square$

Table 2: KL divergence $D(P, Q)$ from P to Q under various parameter settings for Gaussian copula

(a) $\alpha = 0.25$

		ζ				
		0.1	0.5	1.0	5.0	10.0
	-0.951	4.532	1.932	0.998	0.149	0.072
	-0.454	0.124	0.053	0.028	0.004	0.002
	-0.156	0.012	0.005	0.003	0	0
θ	0	0	0	0	0	0
	0.156	0.012	0.005	0.003	0	0
	0.454	0.124	0.054	0.028	0.004	0.002
	0.951	4.526	1.932	1.004	0.149	0.071

(b) $\alpha = 0.5$

		ζ				
		0.1	0.5	1.0	5.0	10.0
	-0.951	3.142	1.422	0.731	0.089	0.042
	-0.454	0.087	0.039	0.020	0.002	0.001
	-0.156	0.008	0.004	0.002	0	0
θ	0	0	0	0	0	0
	0.156	0.008	0.004	0.002	0	0
	0.454	0.086	0.039	0.020	0.002	0.001
	0.951	3.156	1.417	0.726	0.088	0.041

(c) $\alpha = 0.75$

		ζ				
		0.1	0.5	1.0	5.0	10.0
	-0.951	4.962	2.246	1.153	0.153	0.072
	-0.454	0.135	0.062	0.031	0.004	0.002
	-0.156	0.013	0.006	0.003	0	0
θ	0	0	0	0	0	0
	0.156	0.013	0.006	0.003	0	0
	0.454	0.136	0.062	0.032	0.004	0.002
	0.951	4.955	2.244	1.150	0.153	0.072

(d) $\alpha = 1.0$

		ζ				
		0.1	0.5	1.0	5.0	10.0
	-0.951	13.068	4.369	2.174	0.334	0.161
	-0.454	0.359	0.119	0.060	0.009	0.004
	-0.156	0.034	0.012	0.006	0.001	0
θ	0	0	0	0	0	0
	0.156	0.035	0.012	0.006	0.001	0
	0.454	0.358	0.120	0.060	0.009	0.004
	0.951	13.074	4.375	2.166	0.335	0.162

Table 3: KL divergence $D(P,Q)$ from P to Q under various parameter settings for Student t copula

(a) $\alpha = 0.25$

		ζ				
		0.1	0.5	1.0	5.0	10.0
	-0.951	2.559	1.450	0.878	0.165	0.081
	-0.454	0.139	0.069	0.040	0.006	0.003
	-0.156	0.026	0.017	0.012	0.002	0.001
θ	0	0	0	0	0	0
	0.156	0.025	0.017	0.012	0.002	0.001
	0.454	0.138	0.069	0.040	0.006	0.003
	0.951	2.562	1.447	0.879	0.166	0.081

(b) $\alpha = 0.5$

		ζ				
		0.1	0.5	1.0	5.0	10.0
	-0.951	2.025	1.079	0.638	0.103	0.047
	-0.454	0.112	0.060	0.035	0.004	0.002
	-0.156	0.026	0.018	0.011	0.001	0
	0	0	0	0	0	0
	0.156	0.026	0.017	0.012	0.001	0
	0.454	0.113	0.060	0.035	0.004	0.002
	0.951	2.027	1.078	0.642	0.103	0.048

(c) $\alpha = 0.75$

		ζ				
		0.1	0.5	1.0	5.0	10.0
	-0.951	2.472	1.419	0.893	0.174	0.082
	-0.454	0.151	0.086	0.051	0.007	0.003
	-0.156	0.028	0.023	0.016	0.002	0.001
θ	0	0	0	0	0	0
	0.156	0.028	0.022	0.016	0.002	0.001
	0.454	0.151	0.085	0.050	0.006	0.003
	0.951	2.477	1.423	0.892	0.171	0.082

(d) $\alpha = 1.0$

		ζ				
		0.1	0.5	1.0	5.0	10.0
	-0.951	4.576	2.454	1.568	0.361	0.181
	-0.454	0.328	0.146	0.085	0.013	0.006
	-0.156	0.057	0.033	0.024	0.004	0.002
θ	0	0	0	0	0	0
	0.156	0.057	0.032	0.023	0.004	0.002
	0.454	0.327	0.146	0.085	0.013	0.006
	0.951	4.572	2.447	1.565	0.360	0.181

Table 4: KL divergence $D(P, Q)$ from P to Q under various parameter settings for Clayton copula

(a) $\alpha = 0.25$

		ζ				
		0.1	0.5	1.0	5.0	10.0
	0	0	0	0	0	0
θ	0.222	0.019	0.013	0.008	0.001	0
	0.857	0.192	0.131	0.083	0.010	0.004
	8	5.041	3.459	2.306	0.462	0.224

(b) $\alpha = 0.5$

		ζ				
		0.1	0.5	1.0	5.0	10.0
	0	0	0	0	0	0
θ	0.222	0.020	0.013	0.008	0.001	0
	0.857	0.196	0.131	0.082	0.007	0.003
	8	3.953	2.697	1.770	0.313	0.143

(c) $\alpha = 0.75$

		ζ				
		0.1	0.5	1.0	5.0	10.0
	0	0	0	0	0	0
θ	0.222	0.027	0.019	0.012	0.001	0
	0.857	0.268	0.183	0.114	0.012	0.005
	8	4.892	3.354	2.257	0.481	0.235

(d) $\alpha = 1.0$

		ζ				
		0.1	0.5	1.0	5.0	10.0
	0	0	0	0	0	0
θ	0.222	0.036	0.025	0.016	0.002	0.001
	0.857	0.357	0.242	0.156	0.020	0.009
	8	6.447	4.482	3.114	0.821	0.441

Table 5: KL divergence $D(P, Q)$ from P to Q under various parameter settings for Gumbel copula

(a) $\alpha = 0.25$

		ζ				
		0.1	0.5	1.0	5.0	10.0
θ	1	0.247	0.168	0.107	0.013	0.006
	1.111	0.293	0.197	0.126	0.016	0.007
	1.429	0.434	0.295	0.188	0.024	0.010
	5	2.749	1.872	1.226	0.211	0.097

(b) $\alpha = 0.5$

		ζ				
		0.1	0.5	1.0	5.0	10.0
θ	1	0.247	0.166	0.103	0.009	0.004
	1.111	0.289	0.196	0.123	0.011	0.004
	1.429	0.422	0.282	0.179	0.017	0.006
	5	2.239	1.513	0.977	0.144	0.062

(c) $\alpha = 0.75$

		ζ				
		0.1	0.5	1.0	5.0	10.0
θ	1	0.341	0.230	0.146	0.016	0.006
	1.111	0.399	0.271	0.171	0.018	0.007
	1.429	0.571	0.387	0.246	0.029	0.011
	5	2.864	1.955	1.284	0.231	0.105

(d) $\alpha = 1.0$

		ζ				
		0.1	0.5	1.0	5.0	10.0
θ	1	0.452	0.308	0.199	0.027	0.012
	1.111	0.530	0.357	0.231	0.032	0.014
	1.429	0.757	0.517	0.337	0.049	0.022
	5	3.774	2.597	1.762	0.399	0.198

Table 6: Spearman's rho values $\rho(P)/\rho(Q)$ under various parameter settings for Gaussian copula

(a) $\alpha = 0.25$

		ζ				
		0.1	0.5	1.0	5.0	10.0
	-0.951	-0.946 / -0.259	-0.946 / -0.746	-0.946 / -0.884	-0.946 / -0.939	-0.946 / -0.943
	-0.454	-0.438 / -0.148	-0.437 / -0.358	-0.437 / -0.411	-0.439 / -0.438	-0.439 / -0.437
	-0.156	-0.150 / -0.053	-0.149 / -0.123	-0.150/ -0.147	-0.149 / -0.149	-0.151 / -0.149
θ	0	0 / -0.001	-0.001 / 0	-0.001 / 0	0 / 0.001	0.001 / 0.006
	0.156	0.150 / 0.053	0.148 / 0.118	0.147 / 0.143	0.149 / 0.145	0.149 / 0.148
	0.454	0.436 / 0.156	0.438 / 0.366	0.437 / 0.413	0.438 / 0.437	0.437 / 0.437
	0.951	0.946 / 0.258	0.947 / 0.747	0.947 / 0.881	0.946 / 0.934	0.946 / 0.945

(b) $\alpha = 0.5$

		ζ				
		0.1	0.5	1.0	5.0	10.0
	-0.951	-0.946 / -0.256	-0.946 / -0.745	-0.946 / -0.876	-0.946 / -0.934	-0.947 / -0.942
	-0.454	-0.437 / -0.128	-0.438 / -0.304	-0.439 / -0.369	-0.437 / -0.428	-0.438 / -0.433
	-0.156	-0.150 / -0.041	-0.150 / -0.103	-0.148 / -0.121	-0.149 / -0.149	-0.149 / -0.149
θ	0	0 / 0	0.001 / 0	0.001 / -0.001	0 / -0.003	0 / -0.004
	0.156	0.149 / 0.038	0.149 / 0.104	0.149 / 0.121	0.148 / 0.145	0.149 / 0.147
	0.454	0.439 / 0.125	0.437 / 0.308	0.438 / 0.364	0.436 / 0.426	0.437 / 0.435
	0.951	0.946 / 0.256	0.946 / 0.749	0.946 / 0.881	0.946 / 0.938	0.946 / 0.944

(c) $\alpha = 0.75$

		ζ				
		0.1	0.5	1.0	5.0	10.0
	-0.951	-0.946 / -0.26	-0.946 / -0.739	-0.946 / -0.872	-0.947 / -0.936	-0.946 / -0.939
	-0.454	-0.437 / -0.104	-0.437 / -0.280	-0.436 / -0.353	-0.436 / -0.425	-0.438 / -0.430
	-0.156	-0.150 / -0.035	-0.148 / -0.097	-0.150 / -0.116	-0.150/ -0.144	-0.149 / -0.149
θ	0	0 / 0.004	-0.002 / 0.003	-0.001 / 0.001	0.002 / -0.001	0.001 / 0.003
	0.156	0.148 / 0.034	0.150 / 0.087	0.149 / 0.116	0.148 / 0.144	0.150 / 0.137
	0.454	0.437 / 0.105	0.438 / 0.278	0.436 / 0.356	0.437 / 0.420	0.437 / 0.434
	0.951	0.947 / 0.263	0.946 / 0.737	0.946 / 0.871	0.946 / 0.933	0.947 / 0.943

(d) $\alpha = 1.0$

		ζ				
		0.1	0.5	1.0	5.0	10.0
	-0.951	-0.946 / -0.259	-0.946 / -0.737	-0.947 / -0.873	-0.946 / -0.934	-0.946 / -0.942
	-0.454	-0.439 / -0.099	-0.438 / -0.271	-0.438 / -0.355	-0.438 / -0.425	-0.438 / -0.430
	-0.156	-0.152 / -0.027	-0.149 / -0.089	-0.151/ -0.114	-0.150 / -0.141	-0.148 / -0.144
θ	0	0.001 / 0.002	-0.001 / 0.003	-0.001 / -0.006	-0.001 / 0	0 / -0.002
	0.156	0.149 / 0.033	0.150 / 0.092	0.149 / 0.122	0.150 / 0.144	0.150 / 0.148
	0.454	0.437 / 0.099	0.437 / 0.276	0.437 / 0.348	0.437 / 0.423	0.438 / 0.431
	0.951	0.947 / 0.249	0.946 / 0.737	0.946 / 0.877	0.946 / 0.937	0.947 / 0.941

Table 7: Spearman's rho values $\rho(P) / \rho(Q)$ under various parameter settings for Student t copula

(a) $\alpha = 0.25$

		ζ				
		0.1	0.5	1.0	5.0	10.0
	-0.951	-0.942 / -0.258	-0.942 / -0.743	-0.942 / -0.879	-0.942 / -0.935	-0.942 / -0.938
	-0.454	-0.427 / -0.144	-0.427 / -0.353	-0.427 / -0.407	-0.427 / -0.425	-0.426 / -0.428
	-0.156	-0.144 / -0.050	-0.145 / -0.115	-0.144 / -0.136	-0.144 / -0.143	-0.148 / -0.143
θ	0	-0.001 / 0.001	-0.002 / -0.002	-0.002 / -0.005	0 / 0.004	0 / -0.004
	0.156	0.147 / 0.053	0.146 / 0.121	0.146 / 0.141	0.147 / 0.144	0.145 / 0.146
	0.454	0.428 / 0.145	0.428 / 0.352	0.426 / 0.407	0.429 / 0.429	0.427 / 0.430
	0.951	0.942 / 0.256	0.942 / 0.742	0.942 / 0.871	0.942 / 0.936	0.943 / 0.936

(b) $\alpha = 0.5$

		ζ				
		0.1	0.5	1.0	5.0	10.0
	-0.951	-0.942 / -0.258	-0.942 / -0.739	-0.942 / -0.869	-0.942 / -0.930	-0.942 / -0.937
	-0.454	-0.426 / -0.122	-0.427 / -0.302	-0.426 / -0.364	-0.427 / -0.419	-0.425 / -0.425
	-0.156	-0.146 / -0.043	-0.146 / -0.106	-0.146 / -0.120	-0.145 / -0.144	-0.145 / -0.141
θ	0	0 / -0.002	0.001 / -0.003	-0.001 / 0	-0.001 / -0.004	0.001 / 0.002
	0.156	0.145 / 0.048	0.145 / 0.101	0.147 / 0.123	0.146 / 0.141	0.145 / 0.141
	0.454	0.426 / 0.125	0.426 / 0.304	0.427 / 0.366	0.427 / 0.418	0.427 / 0.420
	0.951	0.942 / 0.249	0.942 / 0.737	0.942 / 0.871	0.942 / 0.935	0.942 / 0.933

(c) $\alpha = 0.75$

		ζ				
		0.1	0.5	1.0	5.0	10.0
	-0.951	-0.942 / -0.256	-0.942 / -0.731	-0.942 / -0.865	-0.942 / -0.932	-0.942 / -0.935
	-0.454	-0.428 / -0.117	-0.428 / -0.286	-0.426 / -0.351	-0.428 / -0.415	-0.428 / -0.421
	-0.156	-0.145 / -0.039	-0.147 / -0.096	-0.145 / -0.121	-0.145 / -0.144	-0.145 / -0.145
θ	0	0 / 0.001	0 / 0	0.001 / 0	0 / 0.004	0.001 / 0.003
	0.156	0.146 / 0.035	0.146 / 0.089	0.146 / 0.121	0.146 / 0.140	0.144 / 0.146
	0.454	0.429 / 0.111	0.429 / 0.277	0.428 / 0.352	0.429 / 0.411	0.427 / 0.424
	0.951	0.942 / 0.257	0.942 / 0.725	0.942 / 0.866	0.942 / 0.932	0.942 / 0.934

(d) $\alpha = 1.0$

		ζ				
		0.1	0.5	1.0	5.0	10.0
	-0.951	-0.942 / -0.255	-0.942 / -0.730	-0.942 / -0.867	-0.942 / -0.930	-0.942 / -0.935
	-0.454	-0.428 / -0.108	-0.428 / -0.283	-0.428 / -0.355	-0.427 / -0.415	-0.428 / -0.417
	-0.156	-0.147 / -0.039	-0.146 / -0.095	-0.146 / -0.118	-0.146 / -0.141	-0.145 / -0.136
θ	0	0 / 0.001	0.001 / 0.001	-0.001 / 0.001	0 / 0.003	-0.002 / -0.001
	0.156	0.146 / 0.033	0.147 / 0.096	0.147 / 0.119	0.147 / 0.143	0.146 / 0.143
	0.454	0.428 / 0.115	0.428 / 0.282	0.427 / 0.354	0.426 / 0.416	0.427 / 0.421
	0.951	0.942 / 0.252	0.942 / 0.729	0.942 / 0.866	0.942 / 0.933	0.942 / 0.939

Table 8: Spearman's rho values $\rho(P) / \rho(Q)$ under various parameter settings for Clayton copula

(a) $\alpha = 0.25$

		ζ				
		0.1	0.5	1.0	5.0	10.0
θ	0	-0.001 / 0.004	0 / 0	-0.001 / -0.001	0 / 0.001	-0.001 / -0.004
	0.222	0.148 / 0.035	0.149 / 0.117	0.149 / 0.153	0.149 / 0.158	0.149 / 0.156
	0.857	0.434 / 0.117	0.435 / 0.375	0.435 / 0.454	0.434 / 0.461	0.434 / 0.454
	8	0.941 / 0.262	0.941 / 0.738	0.941 / 0.874	0.941 / 0.945	0.941 / 0.939

(b) $\alpha = 0.5$

		ζ				
		0.1	0.5	1.0	5.0	10.0
θ	0	-0.001 / -0.006	0 / 0.001	0.001 / 0.001	0 / -0.003	0 / -0.001
	0.222	0.149 / 0.021	0.148 / 0.071	0.149 / 0.104	0.149 / 0.147	0.150 / 0.151
	0.857	0.434 / 0.060	0.435 / 0.241	0.433 / 0.337	0.433 / 0.427	0.435 / 0.429
	8	0.941 / 0.248	0.941 / 0.729	0.941 / 0.864	0.941 / 0.928	0.941 / 0.936

(c) $\alpha = 0.75$

		ζ				
		0.1	0.5	1.0	5.0	10.0
θ	0	-0.002 / 0	0 / -0.001	-0.001 / -0.005	0.002 / 0.003	0 / -0.003
	0.222	0.148 / 0.013	0.147 / 0.049	0.150 / 0.080	0.148 / 0.136	0.150 / 0.136
	0.857	0.434 / 0.030	0.435 / 0.162	0.435 / 0.271	0.434 / 0.393	0.433 / 0.411
	8	0.941 / 0.202	0.941 / 0.698	0.941 / 0.839	0.941 / 0.920	0.941 / 0.924

(d) $\alpha = 1.0$

		ζ				
		0.1	0.5	1.0	5.0	10.0
θ	0	0.001 / 0.003	0 / 0.004	0.001 / -0.004	-0.001 / 0.009	-0.001 / -0.001
	0.222	0.149 / 0.003	0.150 / 0.031	0.148 / 0.062	0.149 / 0.123	0.149 / 0.134
	0.857	0.435 / 0.012	0.434 / 0.111	0.435 / 0.214	0.435 / 0.368	0.433 / 0.390
	8	0.941 / 0.067	0.941 / 0.611	0.941 / 0.798	0.941 / 0.904	0.941 / 0.924

Table 9: Spearman's rho values $\rho(P) / \rho(Q)$ under various parameter settings for Gumbel copula

(a) $\alpha = 0.25$

		ζ				
		0.1	0.5	1.0	5.0	10.0
θ	1	0.479 / 0.137	0.480 / 0.414	0.478 / 0.508	0.479 / 0.506	0.479 / 0.499
	1.111	0.510 / 0.146	0.510 / 0.439	0.509 / 0.534	0.510 / 0.535	0.508 / 0.533
	1.429	0.584 / 0.173	0.583 / 0.516	0.586 / 0.613	0.585 / 0.602	0.585 / 0.602
	5	0.884 / 0.252	0.885 / 0.734	0.885 / 0.847	0.885 / 0.891	0.884 / 0.896

(b) $\alpha = 0.5$

		ζ				
		0.1	0.5	1.0	5.0	10.0
θ	1	0.479 / 0.072	0.477 / 0.270	0.478 / 0.374	0.477 / 0.469	0.479 / 0.472
	1.111	0.510 / 0.083	0.511 / 0.296	0.508 / 0.414	0.509 / 0.497	0.511 / 0.504
	1.429	0.585 / 0.096	0.583 / 0.360	0.586 / 0.491	0.585 / 0.572	0.585 / 0.585
	5	0.885 / 0.231	0.884 / 0.695	0.884 / 0.820	0.885 / 0.874	0.885 / 0.877

(c) $\alpha = 0.75$

		ζ				
		0.1	0.5	1.0	5.0	10.0
θ	1	0.478 / 0.041	0.479 / 0.185	0.478 / 0.301	0.480 / 0.441	0.478 / 0.455
	1.111	0.509 / 0.043	0.510 / 0.197	0.509 / 0.333	0.509 / 0.464	0.510 / 0.477
	1.429	0.585 / 0.048	0.584 / 0.259	0.584 / 0.401	0.585 / 0.542	0.584 / 0.558
	5	0.885 / 0.151	0.885 / 0.603	0.884 / 0.772	0.885 / 0.857	0.885 / 0.870

(d) $\alpha = 1.0$

		ζ				
		0.1	0.5	1.0	5.0	10.0
θ	1	0.479 / 0.013	0.479 / 0.122	0.478 / 0.247	0.476 / 0.411	0.479 / 0.430
	1.111	0.509 / 0.014	0.510 / 0.130	0.508 / 0.265	0.509 / 0.444	0.510 / 0.465
	1.429	0.583 / 0.016	0.585 / 0.168	0.584 / 0.328	0.584 / 0.512	0.585 / 0.538
	5	0.884 / 0.044	0.884 / 0.466	0.884 / 0.713	0.885 / 0.832	0.884 / 0.848

Table 10: KL divergence $D(P, Q)$ from P to Q under various parameter settings for Gaussian copula of dimension 3

(a) $\alpha = 0.25$

		ζ				
		0.1	0.5	1.0	5.0	10.0
	0	0	0	0	0	0
	0.156	0.022	0.009	0.005	0.001	0
θ	0.454	0.190	0.081	0.042	0.006	0.003
	0.951	6.091	2.597	1.347	0.200	0.096

(b) $\alpha = 0.5$

		ζ				
		0.1	0.5	1.0	5.0	10.0
	0	0	0	0	0	0
	0.156	0.015	0.007	0.003	0	0
θ	0.454	0.132	0.059	0.030	0.004	0.002
	0.951	4.224	1.907	0.977	0.118	0.056

(c) $\alpha = 0.75$

		ζ				
		0.1	0.5	1.0	5.0	10.0
	0	0	0	0	0	0
	0.156	0.023	0.011	0.005	0.001	0
θ	0.454	0.207	0.094	0.048	0.006	0.003
	0.951	6.675	3.023	1.553	0.204	0.097

(d) $\alpha = 1.0$

		ζ				
		0.1	0.5	1.0	5.0	10.0
	0	0	0	0	0	0
	0.156	0.061	0.020	0.010	0.002	0.001
θ	0.454	0.547	0.182	0.092	0.014	0.007
	0.951	17.564	5.877	2.914	0.451	0.218

Table 11: KL divergence $D(P, Q)$ from P to Q under various parameter settings for Clayton copula of dimension 3

(a) $\alpha = 0.25$

		ζ				
		0.1	0.5	1.0	5.0	10.0
θ	0	0	0	0	0	0
	0.222	0.033	0.023	0.014	0.002	0.001
	0.857	0.298	0.203	0.130	0.016	0.007
	8	8.277	5.678	3.806	0.795	0.381

(b) $\alpha = 0.5$

		ζ				
		0.1	0.5	1.0	5.0	10.0
θ	0	0	0	0	0	0
	0.222	0.033	0.022	0.014	0.001	0
	0.857	0.270	0.182	0.114	0.011	0.004
	8	4.757	3.257	2.176	0.435	0.206

(c) $\alpha = 0.75$

		ζ				
		0.1	0.5	1.0	5.0	10.0
θ	0	0	0	0	0	0
	0.222	0.047	0.032	0.020	0.002	0.001
	0.857	0.369	0.251	0.160	0.018	0.007
	8	5.235	3.611	2.466	0.583	0.303

(d) $\alpha = 1.0$

		ζ				
		0.1	0.5	1.0	5.0	10.0
θ	0	0	0	0	0	0
	0.222	0.061	0.042	0.027	0.003	0.001
	0.857	0.489	0.335	0.214	0.032	0.014
	8	6.816	4.771	3.383	0.976	0.546