Université de Montréal

Densités de Copules Archimédiennes Hiérarchiques

par
David-Olivier Pham

Département de Mathématiques
Faculté des arts et des sciences

Mémoire présenté à la Faculté des études supérieures
en vue de l'obtention du grade de Maître ès sciences (M.Sc.)
en mathématiques orientées en actuariat

Avril, 2012

Université de Montréal
Faculté des études supérieures

Ce mémoire intitulé:

Densités de Copules Archimédiennes Hiérarchiques

présenté par:

David-Olivier Pham

a été évalué par un jury composé des personnes suivantes:

Bédard Mylène,	président-rapporteur
Morales Manuel,	directeur de recherche
Doray Louis,	membre du jury

Mémoire accepté le: ....

ABSTRACT

Nested Archimedean copulas recently gained interest since they generalize the well-known class of Archimedean copulas to allow for partial asymmetry. Sampling algorithms and strategies have been well investigated for nested Archimedean copulas. However, for likelihood based inference such as estimation or goodness-of-fit testing it is important to have the density. The present work fills this gap. After a short introduction on copula and nested Archimedean copulas, a general formula for the derivatives of the nodes and inner generators appearing in nested Archimedean copulas is developed. This leads to a tractable formula for the density of nested Archimedean copulas. Various examples including famous Archimedean families and transformations of such are given. Furthermore, a numerically efficient way to evaluate the log-density is presented.

RÉSUMÉ

Les copulas archimédiennes hiérarchiques ont récemment gagné en intérêt puisqu'elles généralisent la famille de copules archimédiennes, car elles introduisent une asymétrie partielle. Des algorithmes d'échantillonnages et des méthodes ont largement été développés pour de telles copules. Néanmoins, concernant l'estimation par maximum de vraisemblance et les tests d'adéquations, il est important d'avoir à disposition la densité de ces variables aléatoires. Ce travail remplie ce manque. Après une courte introduction aux copules et aux copules archimédiennes hiérarchiques, une équation générale sur les dérivées des noeuds et générateurs internes apparaissant dans la densité des copules archimédiennes hiérarchique. sera dérivée. Il en suit une formule tractable pour la densité des copules archimédiennes hiérarchiques. Des exemples incluant les familles archimédiennes usuelles ainsi que leur transformations sont présentés. De plus, une méthode numérique efficiente pour évaluer le logarithme des densités est présentée.

Keywords: Nested Archimedean copulas, generator derivatives, likelihood-based inference.

MSC2010: 62H99, 65C60, 62H12, 62F10.

CONTENTS

LIST OF TABLES

LIST OF FIGURES

ACKNOWLEDGMENTS

I would like to thank Johanna Nešlehová and Christian Genest for having introduced me to the wonderful world of mathematical statistics and copulas. I am also indebted to Paul Embrechts, who accepted to supervise my master's thesis, and to Manuel Morales, who provided me with a glimpse of the research environment in Montreal. This work could not have been done without the support, enthusiasm and expertise of Marius Hofert. He carefully presented me the problem, helped me throughout the writing of this thesis, and above all shared his experience and knowledge about copulas and his actual research results.

PREFACE

In recent years, there have been several critics about the uses of too sophisticated mathematical and statistical models, especially in finance. People do argue against the restrictive hypothesis of the celebrated Black-Scholes formula for option pricing, however, there also is a lot of concern about the use of copulas. Copulas are tools to describe stochastic dependencies, and during recent years, a certain class of copulas was used in the financial industries to price a large amount of financial derivatives. Indeed, Embrechts (2009) provided three reasons to explain the recent interest of copulas in research: finance, finance, finance. However, although there has been a large amount of evidences and cautions against the uses of this class of copulas, people still used it. It resulted in an inadequate pricing of financial products and finally contributed in a crash of the financial market, a worldwide recession and a large amount of loss for people who did not listen to the advices.

Since then, there have been discussions among academics and profesionals about using complex numerical models in the industries. However, it is undeniable that sophisticated statistical methods have their place there. Copulas, for instance, provide a useful and successful tool to understand and forecast multivariate random behaviors in hydrology, in medical and climate researches.

I am on the side of those who think that it can only be helpful to society to have at its disposition robust and well understood mathematical tools. This belief motivates the subject of this thesis. Hierarchical models using Archimedean copulas have recently gained interest in order to describe multidimensional phenomena, nevertheless, densities of such model are not known. This work tries to fill the gap.

Chapter 1 starts with a little introduction to copulas. Chapter 2 presents known results about (nested) Archimedean copulas and sets the framework for the next chapter. Chapter 3 is the core of this thesis. It begins by giving the general strategy and results from analysis to compute the densities of nested Archimedean copulas. Examples and numerical implementation are also covered. It is worth to underscore that numerical implementation of copulas in high dimension is a tremendously hard task. Fortunately, the R package `nacopula` helps a lot.

CHAPTER 1

COPULAS

In this chapter, we present general results about copulas and multivariate stochastic dependencies and cover results needed for this work. For more details, the interested reader may consult the more detailed literature, for instance the excellent textbooks McNeil et al. (2005), Nelsen (2007) or the mathematically self-contained PhD thesis Hofert (2010).

1.1 Preliminaries

In the study of univariate random behavior, one can naturally ask the question, given the level p between zero and one, which observation q is needed so that $\mathbb{P}(X \leq q) = p$, where X is a real-valued random variable. In the case where the distribution function is bijective, the answer is trivially given by the inverse function. However, when this is not the case, one has to pay attention to what is allowed during the computations and what happens when one uses this inverse. It leads us to the following definition, where $\mathbb{R}$ is the set of real number and I is the unit cube $[0, 1]$.

Definition 1.1 (Generalized Inverse)**.** Let $F : \bar{\mathbb{R}} \to I$ be an increasing function with the convention that $F(\pm\infty)$ is interpreted in the sense of the limit of $F(x)$ as x tends to $\pm\infty$. Then the *generalized inverse* $F^{-} : I \to \bar{\mathbb{R}}$ of F is defined by

$$F^{-}(y) = \inf \{x \in \mathbb{R} : F(x) \geq y\}$$

with the convention that $\inf \emptyset = \infty$.

Note that $F^{-}(y) = -\infty$ if and only if $F(x) \geq y$ for all $x \in \mathbb{R}$ and $F^{-}(y) = \infty$ if and only if $F(x) < y$ for all $x \in \mathbb{R}$. As stressed before, the generalized inverse must be manipulated with care. The main properties are precisely stated in Embrechts and Hofert (2011). The following proposition gives a useful result.

Proposition 1.2. Let $F : \bar{\mathbb{R}} \to I$ be the distribution function of X. Then the inequality $F(x) \geq y$ stands if and only if $x \geq F^{-}(y)$.

Proof

See Embrechts and Hofert (2011). □

The following proposition is the key to sampling univariate random variables with distribution function F. It says that if one can sample the uniform distribution, sampling from F can theoretically be performed. It is known as the *inversion method*.

Proposition 1.3. Let X be an univariate random variable following the distribution function F. If $U \sim \mathrm{U}[0,1]$, then $F^{-}(U) \sim F$. Moreover, if F is continuous, then $F(X) \sim \mathrm{U}[0,1]$.

Proof
By Proposition 1.2, $\mathbb{P}\{F^{-}(U) \leq x\} = \mathbb{P}\{U \leq F(x)\} = F(x)$. For the second part of the proposition, since F is continuous, $\mathbb{P}\{F(X) \leq u\} = 1 - \mathbb{P}\{F(X) \geq u\}$ and Proposition 1.2 implies

$$\begin{aligned}\mathbb{P}\{F(X) \leq u\} &= 1 - \mathbb{P}\{F(X) \geq u\} = 1 - \mathbb{P}\{X \geq F^{-}(u))\} \\ &= \mathbb{P}\{X \leq F^{-}(u)\} = F\{F^{-}(u)\} = u,\end{aligned}$$

for all $u \in (0,1)$. □

The second statement is a motivation to use copulas defined on the unit hypercube. With these tools and facts at hand, we are now prepared to study copulas.

1.2 Definitions and Basic Properties of Copulas

In this section, we present definitions and properties of copulas in order to understand Sklar's Theorem which is the foundation of copula theory. In the following, for two vectors $\boldsymbol{a}, \boldsymbol{b}$ of $\mathbb{R}^d$, the inequality $\boldsymbol{a} \leq \boldsymbol{b}$ means that $a_j \leq b_j$ for all $j \in \{1, \ldots, d\}$.

Definition 1.4 (Copula)**.** A d-dimensional *copula* is a d-dimensional distribution function with standard uniform univariate margins. It implies that a copula is a function C from I^d to I (where I stands for the unit interval $[0,1]$), that fulfills the following conditions.

(i) C is *grounded*, that is, $C(u) = 0$ if $u_j = 0$ for any $j \in \{1, \ldots, d\}$;

(ii) C has *uniform margins*, meaning $C(u) = u_j$ for all $u_j \in I$ if all the other component of $\boldsymbol{u}$ are equal to 1.

(iii) C is d-increasing, that is, for all $\boldsymbol{a}, \boldsymbol{b} \in I^d$ with $\boldsymbol{a} \leq \boldsymbol{b}$, C respects

$$\Delta_{(\boldsymbol{a},\boldsymbol{b}]} C = \sum_{i_1=1}^{2} \cdots \sum_{i_d=1}^{2} (-1)^{i_1 + \cdots + i_d} C(x_{1i_1}, \ldots, x_{di_d}) \geq 0,$$

where $x_{j1} = a_j$ and $x_{j2} = b_j$ for $j \in \{1, \ldots, d\}$.

An interesting observation is that for any $d \geq 3$ and d-dimensional copula C, each k-dimensional margin of C is a k-dimensional copula, for $k = 2, \ldots, d-1$.

The d-increasing property of copulas expresses the fact that these functions can only give nonnegative measures to volumes in I^d and from the combination of properties (i) and (iii), it stands $C(\boldsymbol{u}) \geq 0$ for all $u \in I^d$. As often, the best way to understand a concept is by working with an example.

Example 1.5. The function $\Pi : I^d \to I$ defined by $\Pi(u) = \prod_{i=1}^d u_i$ fulfills the three conditions. It is grounded, because if $u_j = 0$ for any $j \in \{1,\dots,d\}$ then

$$C(\boldsymbol{u}) = 0 \cdot \prod_{k \in \{1,\dots,d\}\setminus\{j\}} u_j = 0$$

and has uniforms margins, because $C(1,\dots,1,u_j,1,\dots,1) = u_j$. For the last property, one observes that for any suitable $\boldsymbol{a},\boldsymbol{b} \in I^d$ with $\boldsymbol{a} \le \boldsymbol{b}$,

$$\Delta_{(\boldsymbol{a},\boldsymbol{b}]}\Pi = \prod_{i=1}^d (b_j - a_j) \ge 0.$$

Note that Π is called the *independence copula* since one can show that the copula characterizing an independent random vector is Π. Two other examples of well-known copulas are given by the *Clayton copula*

$$C_\theta^{\mathrm{Cl}}(\boldsymbol{u}) = \left(\sum_{j=1}^d u_j^{-\theta} - 1\right)^{-1/\theta}, \quad \boldsymbol{u} \in I^d, \quad 0 < \theta < \infty,$$

and by the *Gumbel* (or *Gumbel-Hogugard copula*)

$$C_\theta^{\mathrm{Gu}}(\boldsymbol{u}) = \exp\left[-\left\{\sum_{j=1}^d (-\ln u_j)^\theta\right\}^{1/\theta}\right], \quad \boldsymbol{u} \in I^d, \quad 1 \le \theta < \infty.$$

The fact that these two functions are copulas will be shown later. One observe in passing that the three above copulas can be expressed as

$$C(\boldsymbol{u}) = \psi\{\psi^{-1}(u_1) + \cdots + \psi^{-1}(u_d)\} = \psi\left\{\sum_{s=1}^d \psi^{-1}(u_s)\right\}, \tag{1.1}$$

for some function ψ. We call copulas respecting the above equation *Archimedean copula.* The next chapter gives conditions on the function ψ such that a function defined as (1.1) yields a proper copula.

The functions $M(\boldsymbol{u}) = \min_{1 \le j \le d}\{u_j\}$ and $W(\boldsymbol{u}) = \max\{\sum_{j=1}^d u_j - d + 1, 0\}$ are respectively referred to as the *upper* and *lower Fréchet-Hoeffding bounds* and the following theorem explains why. This result is attributed to Fréchet (1935) and Hoeffding (1940).

Theorem 1.6 (Fréchet-Hoeffding bounds)**.** For $d \ge 2$ and any copula C, the following inequality stands

$$W(\boldsymbol{u}) \le C(\boldsymbol{u}) \le M(u), \quad \boldsymbol{u} \in I^d, \tag{1.2}$$

where $W(\boldsymbol{u}) = \max\{\sum_{j=1}^d u_j - d + 1, 0\}$ and $M(\boldsymbol{u}) = \min_{1 \le j \le d}\{u_j\}$. Moreover, M is a copula in any dimension $d \ge 2$, whereas W is a copula only when $d = 2$.

Proof

The proof is in fact quite elegant, because of its simplicity. If the copula C is considered as a d-dimensional distribution function with standard uniform univariate margins, then using some elementary properties of probability measure gives the result. Indeed using the inequality $\mathbb{P}(A) \leq \mathbb{P}(B)$, where A, B are measurable sets, and $A \subset B$, with $A = \cap_{i=1}^{d}\{U_i < u_i\}$ and $B = \{U_j < u_j\}$ for some $j \in \{1,\ldots,d\}$ yields the right inequality of (1.2). Indeed, for all $j \in \{1,\ldots,d\}$,

$$C(\boldsymbol{u}) = C(u_1,\ldots,u_d) = \mathbb{P}(\cap_{i=1}^{d}\{U_i \leq u_i\}) \leq \mathbb{P}(U_j \leq u_j) = u_j,$$

where the last equality stands because C has standard uniform margins. Hence $C(\boldsymbol{u})$ is less than or equal to any u_j, and in particular the smallest of the u_j. To show the left inequality of (1.2), observe that for any measurable sets A, B,

$$\mathbb{P}(A \cap B) = \mathbb{P}(A) + \mathbb{P}(B) - \mathbb{P}(A \cup B) \geq \mathbb{P}(A) + \mathbb{P}(B) - 1.$$

Proceeding by simple iteration of the previous inequality, it results

$$C(\boldsymbol{u}) = \mathbb{P}(\cap_{i=1}^{d}\{U_i \leq u_i\}) \geq \sum_{j=1}^{d} \mathbb{P}(U_j \leq u_j) - d + 1.$$

Combined with the fact that $C(\boldsymbol{u}) \geq 0$, the inequality follows. For the last part of the statement, observe that if $U \sim \mathrm{U}[0,1]$ then $(U,\ldots,U) \sim M$ and that $(U, 1-U) \sim W$. Moreover, the probability measure under W of the hypercube $(1/2,1]^d$ is given by

$$\begin{aligned}
\Delta_{[\mathbf{1/2},\mathbf{1}]}W &= \sum_{i_1=1}^{2} \cdots \sum_{i_d=1}^{2} (-1)^{i_1+\cdots+i_d} W(x_{1i_1},\ldots,x_{di_d}) \\
&= \max\{1+\cdots+1-d+1,0\} - d\max\Big\{\frac{1}{2}+1+\cdots+1-d+1,0\Big\} \\
&\quad + \binom{d}{2}\max\Big\{\frac{1}{2}+\frac{1}{2}+1+\cdots+1-d+1,0\Big\} \\
&\qquad + \cdots + (-1)^d \max\Big\{\frac{1}{2}+\cdots+\frac{1}{2}-d+1,0\Big\} \\
&= 1-d/2,
\end{aligned}$$

which is negative if $d > 2$. Hence W cannot be a distribution function for $d > 2$. $\square$

1.3 Sklar's Theorem

This section is devoted to the theorem at the core of copula theory. It elegantly describes the relationship between a multivariate random vector and its lower-dimensional margins. This theorem referred to as *Sklar's Theorem* comes from a letter from Sklar to Fréchet in which the latter asked about the previous relationship. By introducing copulas, Sklar answered to the question for one-dimensional margins. His result was then published by Fréchet as Sklar (1959). In the following theorem, the range of a function F is denoted as $\operatorname{ran} F$.

Theorem 1.7 (Sklar (1959))**.** Let H be a d-dimensional distribution function with margins F_j, $j \in \{1,\ldots,d\}$. Then, there exists a copula C such that

$$H(\boldsymbol{x}) = H(x_1,\ldots,x_d) = C\{F_1(x_1),\ldots,F_d(x_d)\}, \quad \boldsymbol{x} \in \mathbb{R}^d. \tag{1.3}$$

Additionally, C is uniquely determined on $\Pi_{j=1}^d \operatorname{ran} F_j$, $j \in \{1,\ldots,d\}$, and is given by

$$C(\boldsymbol{u}) = H\{F_1^-(u_1),\ldots,F_d^-(u_d)\}, \quad \boldsymbol{u} \in \prod_{j=1}^d \operatorname{ran} F_j. \tag{1.4}$$

Conversely, given a copula C and univariate distribution functions F_j, $j \in \{1,\ldots,d\}$, H defined by (1.3) is a distribution function with margins F_j, $j \in \{1,\ldots,d\}$.

Proof
See Sklar (1996) for the classical proof or Rüschendorf (2009) for a modern proof. □

It is worth to emphasize that this result allows us to decompose any multivariate distribution function into its margins and in a copula, that is one can study multivariate distribution functions independently of the margins. Note that the normalization of the margins to $[0,1]$ is arbitrary; historically it was even proposed to use $[-1/2,1/2]$ instead of I. Moreover, Sklar's Theorem provides an elegantly simple tool to construct multivariate distributions and is therefore used for sampling purposes, for example.

Example 1.8. There is a well-known tricky question asked to undergraduate students during their first lectures in statistics: If a bivariate distribution has standard normal margins, is the distribution a bivariate normal distribution? Sklar's Theorem provides the answer to this question. Indeed, to construct any multivariate random vector with normal margins which is not itself a multivariate normal, it suffices to take any copula different from the so-called bivariate Gaussian copula C_ρ^{Ga}, given by

$$C_\rho^{\mathrm{Ga}}(u_1,u_2) = \int_{-\infty}^{\Phi^{-1}(u_1)} \int_{-\infty}^{\Phi^{-1}(u_2)} \frac{1}{2\pi\sqrt{1-\rho^2}} \exp\left\{ \frac{-(z_1^2 - 2\rho z_1 z_2 + z_2^2)}{2(1-\rho^2)} \right\} dz_1 dz_2,$$

with $-1 \le \rho \le 1$, and $\Phi(x)$ denotes the cumulative distribution of the standardized normal distribution.

1.4 Random Vectors and Copulas

This section focuses on the link between copulas and random vectors. Sklar's Theorem asserts that each random vector $\boldsymbol{X}$, respectively multivariate distribution function

H, is associated to at least one copula and in the continuous case, the theorem also provides a way to construct it explicitly. Moreover, the following proposition gives a stochastic representation for the copula.

Proposition 1.9. Let $\boldsymbol{X} \sim H$, such that $X_j \sim F_j$, and F_j being continuous, $j \in \{1, \dots, d\}$. Then the copula associated to $\boldsymbol{X}$, respectively to H, is the distribution function of the random vector $(F_1(X_1), \dots, F_d(X_d))^\mathsf{T}$.

Proof

The key point of this proof is the continuity of the margins. Indeed, for all $j \in \{1, \dots, d\}$, Proposition 1.3 states that $F_j(X_j) \sim \mathrm{U}[0,1]$, thus $F_j(X_j)$ and X_j are continuously distributed. Hence the inequality in their distribution function can be substituted by strict ones and vice-versa. Thanks to Proposition 1.2, it stands

$$\mathbb{P}(\cap_{j=1}^d \{F_j(X_j) < u\}) = \mathbb{P}(\cap_{j=1}^d \{X_j < F_j^-(u_j)\}) = \mathbb{P}(\cap_{j=1}^d \{X_j \leq F_j^-(u_j)\}).$$

□

The previous proposition thus motivates the following definition.

Definition 1.10. Let $\boldsymbol{X} \sim H$ be a d-dimensional random vector with continuous margins F_j, $j \in \{1, \dots, d\}$. The *copula of* $\boldsymbol{X}$, respectively the *copula of* H, is defined as the distribution function of the random vector $(F_1(X_1), \dots, F_d(X_d))^\mathsf{T}$.

Note that the previous theorem and definition provide a powerful tool to build and sample implicit multivariate copulas. Indeed, if the random vector $\boldsymbol{X} \in \mathbb{R}^d$ follows a distribution function H, with continuous margins F_j, $j \in \{1, \dots, d\}$, then a sample from the copula of H can be created in the following manner. First draw a sample $\boldsymbol{x}_i = (x_{1i}, \dots, x_{di})^\mathsf{T}$, with $i \in \{1, \dots, n\}$ from H and return $\boldsymbol{u}_i = (F_1(x_{1i}), \dots, F_d(x_{di}))^\mathsf{T}$. This is actually how Gaussian copulas C_P^{Ga} can be sampled. See (McNeil et al., 2005, p. 193) for more details.

Observe that copulas are invariant under strictly increasing transformation on the range of the underlying random variables. Moreover, there exist specific copulas for certain kind of dependence. Indeed, a random vector with continuous margins has independent components if and only if the underlying copula C is Π. If $d = 2$, then X_2 is almost surely a strictly decreasing function in X_1 if and only if the underlying copula C is the lower Fréchet bound W. For any $d \in \mathbb{N}$, there exist $d-1$ almost surely strictly increasing functions linking the first margin of a random vector to its other margins if and only if the underlying copula C is the upper Fréchet bound M.

The last observations are due to Schweizer and Wolff (1981). These authors establish a link between copula theory and the investigation of dependencies between random variables for which copulas are mostly applied today. The invariance of copulas under strictly increasing transformation implies that copulas are only concerned with the dependence structures of a random vector, that is independently of the margins. According to Fisher (1997), this is one of the main reasons why copulas are studied.

Moreover, note that if a bivariate random vector $\boldsymbol{X}$ has W as copula, then it is called *countermonotone* and corresponds to perfect negative dependence. If $\boldsymbol{X}$ has copula M instead, then it is called *comonotone* and corresponds to perfect positive dependence.

1.5 Density and Conditional Distributions

When studying multivariate phenomena, one eventually deals with conditional probability distributions. In the two-dimensional case, one heuristically observes

$$\begin{aligned} C_{U_2|U_1}(u_2|u_1) = \mathbb{P}(U_2 \leq u_2 \,|\, U_1 = u_1) &= \lim_{h \to 0} \frac{1}{h}\{C(u_1+h, u_2) - C(u_1, u_2)\} \\ &= \frac{\partial}{\partial u_1} C(u_1, u_2). \end{aligned}$$

The partial derivative is justified by the fact that copulas are increasing continuous functions in each argument, hence almost everywhere differentiable, see (Hofert, 2010, p. 25) for a proof in the multivariate case. Additionally, as one would expect from a conditional distribution, the range of first order partial derivatives of the copula $C(u_1, u_2)$ is within the unit interval, that is,

$$0 \leq \frac{\partial}{\partial u_j} C(u_1, u_2) \leq 1, \quad j \in \{1, 2\}.$$

A nice statistical interpretation of conditional distributions is the following. Let the random vector $(X_1, X_2)^\mathsf{T}$ have continuous margins with unique copula C. Then the quantity $1 - C_{U_2|U_1}(q|p)$ is the probability that X_2 exceeds its qth quantile given that X_1 attains its pth quantile.

As mentioned before, densities are required to perform maximum likelihood estimation, goodness-of-fit or hypothesis tests, hence it is desirable to be able to calculate them. However, it is common that multivariate distributions, including the Fréchet bounds M and W, do not possess a density function and copulas are not an exception. Nevertheless, when the density exists, as for example with the bivariate Gaussian or Gumbel copula, it is given by

$$c(\boldsymbol{u}) = c(u_1, \ldots, u_d) = \frac{\partial^d C(u_1, \ldots, u_d)}{\partial u_1 \ldots \partial u_d} = \frac{\partial}{\partial \boldsymbol{u}} C(\boldsymbol{u}).$$

Moreover, one could even be more precise when the margins F_j, $j \in \{1, \ldots, d\}$, and the distribution H are absolutely continuous. Indeed, using the previous equation with the representation $C(\boldsymbol{u}) = H\{F_1^-(u_1), \ldots, F_d^-(u_d)\}$, one gets

$$c(\boldsymbol{u}) = \frac{h\{F_1^-(u_1), \ldots, F_d^-(u_d)\}}{f_1\{F^-(u_1)\} \cdots f_d\{F^-(u_d)\}},$$

where h is the density of H, and f_j the density of F_j, $j \in \{1, \ldots, d\}$. This may give the strange situation that there exists a closed form for the density but not for the copula. An example is given with the Gaussian copula; see, for example, (Hofert, 2010, p. 48) for the closed form of the density.

1.6 Symmetries of Copulas

In the following we briefly describe two concepts of symmetries of random variables.

Definition 1.11. A d-dimensional random vector $\boldsymbol{X}$ is called

(1) *radially symmetric* (for $d=1$ simply *symmetric*) around $\boldsymbol{a} \in \mathbb{R}^d$, if $\boldsymbol{X}-\boldsymbol{a}$ and $\boldsymbol{a}-\boldsymbol{X}$ share the same distribution, that is

$$\mathbb{P}(X_1 - a_1 \leq x_1, \dots, X_d - a_d \leq x_d) = \mathbb{P}(a_1 - X_1 \leq x_1, \dots, a_d - X_d \leq x_d),$$

for all $\boldsymbol{x} \in \mathbb{R}^d$ which are continuity points of the probability distribution function.

(2) *exchangeable* if $(X_1, \dots, X_d)^\mathsf{T}$ and $(X_{j_1}, \dots, X_{j_d})^\mathsf{T}$ are equally distributed for all permutations $(j_1, \dots, j_d)$ of $\{1, \dots, d\}$. Analogously, a copula C with the property

$$C(u_1, \dots, u_d) = C(u_{j_1}, \dots, u_{j_d})$$

for all permutations $(j_1, \dots, j_d)$ of $\{1, \dots, d\}$, is called *symmetric* or *exchangeable.*

One observes that if a random vector $\boldsymbol{X}$ is independently and identically distributed then it also has the property of exchangeability. Moreover, the exchangeability of $\boldsymbol{X}$ implies that its margins are identically distributed. However, the converses are false in general.

As a counter example, suppose that $\boldsymbol{X}$ has identically distributed margins. Then for the first statement, it suffices to choose the copula of $\boldsymbol{X}$ to be any but the independence copula Π, for example, the Gumbel copula on page 3 with $\theta > 1$.

The following proposition is a multivariate generalization of the results in (Nelsen, 2007, p. 37)

Proposition 1.12. Let H be a distribution function with continuous margins F_j, $j \in \{1, \dots, d\}$, with copula C and $\boldsymbol{X} \sim H$.

(i) If X_j are symmetric about a_j, $j \in \{1, \dots, d\}$, then $\boldsymbol{X}$ is radially symmetric about $\boldsymbol{a} \in \mathbb{R}^d$ if and only if $C = \hat{C}$, where $\hat{C}$ is the distribution function of $(1-U_1, \dots, 1-U_d)^\mathsf{T}$ for $\boldsymbol{U} \sim C$.

(ii) $\boldsymbol{X}$ is exchangeable if and only if all the margins are equal and if C is exchangeable.

One observes that exchangeability and radial symmetry are not equivalent, nor does either of these notions implies the other. For example, consider the bivariate random vector $(U_1, U_2)^\mathsf{T}$ following the Clayton copula, given in Example 1.5. Then $(U_1, U_2)^\mathsf{T}$ is exchangeable but not radially symmetric. For the converse statement, a radially symmetric copula which is not exchangeable can be constructed via $C(u_1, u_2) = u_1 u_2 + u_1(1-u_1)\sin(\pi u_2)/\pi$.

1.7 Dependence Measures

In order to fully appreciate the flexibility and the advantages of copulas, we briefly introduce some measures of dependence. These measures are among the most popular in the bivariate settings, that is why we will focus on this case.

The most common measure of association is without doubt Pearson's linear correlation coefficient given by

$$\rho = \rho_{12} = \rho_{X_1,X_2} = \frac{\mathrm{Cov}(X_1, X_2)}{\sqrt{\mathrm{Var}(X_1)}\sqrt{\mathrm{Var}(X_2)}},$$

under the condition that X_1 and X_2 have finite second moment. However, this measure has a lot of weaknesses which are listed in Embrechts et al. (2002) and (McNeil et al., 2005, p. 201). In order to avoid them, the two following measures have been introduced.

Definition 1.13 (Spearman's rho and Kendall's tau)**.** Let $X_j \sim F_j$, $j \in \{1,2\}$, be continuously distributed random variables with copula C. Then *Spearman's rho* is defined by

$$\rho_S = \rho_{S,12} = \rho_{S,X_1,X_2} = \rho_{S_C} = \rho_{F_1(X_1),F_2(X_2)} = 12\int_{I^2} uv\, dC(u,v) - 3,$$

that is ρ_S is the Pearson correlation coefficient of the random variables $F_1(X_1)$ and $F_2(X_2)$.

If $(X_1', X_2')^\mathsf{T}$ is an independent and identically distributed copy of $(X_1, X_2)^\mathsf{T}$ then *Kendall's tau* is defined by

$$\tau = \tau_{12} = \tau_{X_1,X_2} = \tau_C = \mathbb{E}[\mathrm{sign}\{(X_1 - X_1')(X_2 - X_2')\}] = 4\int_{I^2} C(u,v)\, dC(u,v) - 1,$$

where $\mathrm{sign}(x) = 1(0 < x < \infty) - 1(-\infty < x < 0)$.

For modeling purposes, these two measures have at least three advantages over Pearson's correlations: They are defined for every pair $(X_1, X_2)^\mathsf{T}$ of random variables; they depend on the copula of $(X_1, X_2)^\mathsf{T}$ only; for every value $\hat{\kappa}$ between -1 and 1 of these measures, there exists a copula C such that the theoretical value of the measure associated to C is equal to its empiric counterpart.

In order to assess extreme events, Sibuya (1959) introduced the concept of *tail dependence*. It is a measure of the extremal dependence between two random variables, meaning, the strength of dependence in the tails of their bivariate distribution.

Definition 1.14. Let $X_j \sim F_j$, $j \in \{1,2\}$, be continuously distributed random variables. The *lower tail-dependence coefficient*, respectively the *upper tail-dependence coefficient*, of X_1 and X_2 are defined as

$$\lambda_l = \lambda_{l,12} = \lambda_{l,X_1,X_2} = \lambda_{l,C} = \lim_{u \to 0} \mathbb{P}\{X_2 \leq F_2^-(u) | X_1 \leq F_1^-(u)\},$$
$$\lambda_u = \lambda_{u,12} = \lambda_{u,X_1,X_2} = \lambda_{u,C} = \lim_{u \to 1} \mathbb{P}\{X_2 > F_2^-(u) | X_1 > F_1^-(u)\},$$

provided that the limits exist. If $\lambda_l \in (0,1]$, respectively $\lambda_u \in (0,1]$, then X_1 and X_2 are *lower tail dependent*, respectively *upper tail dependent*. If $\lambda_l = 0$, respectively $\lambda_u = 0$, then X_1 and X_2 are *lower tail independent*, respectively *upper tail independent*.

One can show that the tail dependence is a property of the underlying copula and does not depend on the marginal distributions.

CHAPTER 2

ARCHIMEDEAN AND NESTED ARCHIMEDEAN COPULAS

In contrast to Gaussian copulas, Archimedean copulas are not constructed via Sklar's Theorem from known multivariate distributions. Rather, one starts with a given functional representation and asks what are the conditions to be a proper copula. As a result, Archimedean copulas are explicit, which is one of the main advantages of this class of copulas. Another advantage of Archimedean copulas is their ability to model different kind of dependencies: from tail dependencies to not radially symmetrical relationships, which is not available for Gaussian copulas. Moreover, many known copula families are Archimedean, which emphasizes the importance of this class. The main drawback of Archimedean copulas in high dimensions is the exchangeability of the components. However, this restriction can be relaxed thanks to nested Archimedean copulas as we will see.

2.1 Archimedean Copulas

We start with the definition of an Archimedean copula. Recall that $I = [0, 1]$.

Definition 2.1 (Archimedean generator and copula)**.** An *Archimedean generator*, or simply *generator*, is a function $\psi : [0, \infty] \to I$ which

(i) is continuous and decreasing;

(ii) is strictly decreasing on $[0, \inf\{t : \psi(t) = 0\}]$;

(iii) satisfies $\psi(0) = 1$ and $\psi(t) \to 0$ as $t \to \infty$.

The set of all such functions is denoted $\boldsymbol{\Psi}$. An Archimedean generator $\psi \in \boldsymbol{\Psi}$ is called *strict* if $\psi(t)$ is positive for all nonnegative t. A d-dimensional copula C is called *Archimedean* if it permits the representation

$$C(\boldsymbol{u}) = C(\boldsymbol{u}; \psi) = \psi\{\psi^{-1}(u_1) + \cdots + \psi^{-1}(u_d)\} = \psi\left\{\sum_{j=1}^{d} \psi^{-1}(u_j)\right\}. \qquad (2.1)$$

Because of the symmetric functional form (2.1), a random vector following an Archimedean copula generated by $\psi \in \boldsymbol{\Psi}$ is exchangeable. In passing, observe that if $\psi \in \boldsymbol{\Psi}$ generates a copula, then the function $\psi(ct)$, where $c > 0$, generates the same copula, hence the choice of the generator of a copula is only up to a positive constant.

2.1.1 Bivariate Archimedean Copulas

In the literature, there are different notations for Archimedean copulas. Traditionally, a bivariate Archimedean copula is presented in the form

$$C(u_1,u_2) = \varphi^{[-1]}\{\varphi(u_1)+\varphi(u_2)\}, \quad (u_1,u_2) \in I^2. \tag{2.2}$$

for a continuous, strictly decreasing function $\varphi : I \to [0,\infty]$ with $\varphi(1) = 0$, where the *pseudo-inverse* $\varphi^{[-1]} : [0,\infty] \to I$ is defined by

$$\varphi^{[-1]}(t) = \varphi^{-1}(t)\cdot 1\{0 \le t \le \varphi(0)\},$$

where φ^{-1} stands for the usual inverse. In this context, φ is referred to as Archimedean generator, see, for example, (McNeil et al., 2005, p.222) or (Nelsen, 2007, p.112). However, it turns out that ψ is more convenient to work with. This notation can be found in (Joe, 1997, p.86) or (Hofert, 2010, p.52).

In two dimensions it was established by Schweizer and Sklar (1963) that a function of type (2.2) is a copula if and only if φ is convex, or equivalently any function C of type (2.1) is a copula if and only $\psi \in \mathbf{\Psi}$ is convex.

The following heuristic argument shows a reason why functions of type (2.2) are used to model bivariate random vectors. If U and V are standard uniform random variables, then the function

$$C(u,v) = \Pi(u,v) = uv, \quad (u,v) \in I^2$$

is perhaps the simplest copula that could exist and this distribution function happens exceptionally when the underlying random variables U and V are independent. If more generally, the random variables are dependent, the goal is to transform the probability distribution function such that U and V become independent after the manipulations. More precisely, consider $\lambda : I \to I$, a continuous strictly increasing function such that $\lambda(0) = 0$, $\lambda(1) = 1$, and suppose that

$$\lambda\{\mathbb{P}(U \le u, V \le v)\} = \lambda\{\mathbb{P}(U \le u)\}\lambda\{\mathbb{P}(V \le v)\} = \lambda(u)\lambda(v), \quad (u,v) \in I^2. \tag{2.3}$$

If $\lambda(t) = t$ then U and V are independent as in the common definition. Setting $\varphi(t) = -\log\lambda(t)$ for $0 < t \le 1$, it results by applying $-\log$ to Equation (2.3)

$$\varphi\{C(u,v)\} = \varphi(u)+\varphi(v),$$

where C is the repartition function of $(U,V)^\mathsf{T}$. Applying $\varphi^{[-1]}$ to both sides of the previous equation yields Equation (2.2).

The term *Archimedean* has a long and interesting history. Inspired by the work of Ling (1965), Genest and MacKay (1986) applied the term Archimedean to copula theory. Ling (1965) studied functions called triangular norms or *t-norms*, which consist of two-place functions with uniform margins, which are grounded, increasing in each component, symmetrical with respect to their arguments and associative, that is, $T\{T(u,v),w\} = T\{u,T(v,w)\}$ for a two-place function T.

Note that a t-norm is a copula if and only if it is 2-increasing and that a copula is a t-norm if and only if it is associative.

T-norms appeared as triangle inequalities in so-called probabilistic metric spaces where the metric $d(p,q)$ is defined as a distribution function $F_{pq}(x)$ on the positive real line. $F_{pq}(x)$ is interpreted as the probability that two points p,q in the metric space are less than x. In this context the term *Archimedean* t-norm appeared and defined a t-norm T satisfying a version of the *Archimedean property*, namely for every $0 \leq u,v \leq 1$ there exists $n \in \mathbb{N}$, such that $T^n(u,\ldots,u) < v$, where $T^2(u_1,u_2) = T(u_1,u_2)$ and $T^n(u_1,\ldots,u_n) = T\{T^{n-1}(u_1,\ldots,u_{n+1}),u_n\}$.

Moreover, Ling (1965) showed that every continuous t-norm admits the representation as in (2.2) and together with the result of Schweizer and Sklar (1963) stating that every function of type (2.2) is a t-norm with φ and $\varphi^{[-1]}$, it follows that a continuous Archimedean t-norm is a copula if and only if φ is convex. Hence due to this connection, copulas of type (2.2) also share the Archimedean property.

Further, Ling (1965) provided a characterization of bivariate Archimedean copulas: A 2-dimensional copula C is a bivariate Archimedean copula if it is associative and if $C(u,u) < u$ for all $u \in (0,1)$.

Moreover, Genest and MacKay (1986) used a nice result from the Norwegian mathematician Abel to give a criteria to determine whether a copula is Archimedean or not. This criteria involves partial derivatives of the copula and the derivative of the generator. The exact formulation can be found in Genest and MacKay (1986).

Note that famous copulas named parametric Archimedean copulas were already studied in Schweizer and Sklar (1961) without their actual name, and thus it appears that the set of Archimedean copulas is the first class of copulas that have been studied.

2.1.2 Multivariate Archimedean Copulas

There has been a great amount of research in the bivariate settings in order to enhance the flexibility of the Archimedean copulas. Nevertheless, in higher dimensions, Archimedean copulas of the form (2.1) lose a little bit of their attractiveness because of their functional symmetry.

As in the bivariate settings, the first question to be answered is what are necessary and sufficient conditions such that ψ generates a proper copula. This answer has only been answered recently and is provided by Malov (2001).

Theorem 2.2 (Malov (2001))**.** Let $\psi \in \mathbf{\Psi}$ and $d \geq 2$. Then, the function given by (2.1) is a copula if and only if ψ is d-monotone, that is, ψ admits derivatives up to the order $d-2$ satisfying

$$(-1)^k \psi^{(k)}(t) \geq 0, \quad t \geq 0, \; k \in \{0,\ldots,d-2\}$$

and $(-1)^{d-2}\psi^{(d-2)}(t)$ is decreasing and convex on $(0,\infty)$.

This result is a generalization of an older well-known result from Kimberling (1974) derived in the context of t-norms. It states that the function given by (2.1) is a copula for all $d \geq 0$ if and only if ψ is completely monotone, that is, ψ has derivatives of all orders and $(-1)^k \psi^{(k)}(t)$ is nonnegative for all nonnegative t and all $k \in \mathbb{N}$. We denote $\boldsymbol{\Psi}_\infty$ the set of all completely monotone generators $\psi \in \boldsymbol{\Psi}$.

Additionally, *Bernstein's Theorem* states that a function ψ is completely monotone on $[0,\infty]$ with $\psi(0) = 1$ if and only if ψ is the *Laplace-Stieltjes* transform of a distribution function F on $[0,\infty]$ and Kimberling (1974) gives a reason to consider as candidates for generators of copulas the set of Laplace-Stieltjes transforms of a distribution functions F on $[0,\infty]$. For such a function F, the Laplace-Stieltjes is defined as

$$\mathcal{LS}[F](t) = \int_0^\infty e^{-tx} dF(x).$$

A nice feature of a generator $\psi \in \boldsymbol{\Psi}$ such that $\psi = \mathcal{LS}[F]$ is that ψ is strict and that $F(0) = 0$. The following proposition inspired from Marshall and Olkin (1988) shows how Laplace-Stieltjes transforms are used to construct random vectors whose distributions are multivariate Archimedean copulas. We state this result because it shows how these kinds of copulas can be simulated.

Proposition 2.3. Let F_0 be a distribution function on $[0,\infty]$ satisfying $F_0(0) = 0$ and let $\psi_0 = \mathcal{LS}[F_0]$. Let $V_0 \sim F_0$ and $E_j \overset{\text{iid}}{\sim} \text{Exp}(1)$ independently of V_0 and $G_0(u_j; v_0)$ denotes $\exp\{-v\psi^{-1}(u_j)\}$, $j \in \{1 \ldots, d\}$. Then

$$\boldsymbol{U} = \left(\psi_0\left\{\frac{E_1}{V_0}\right\}, \ldots, \psi_0\left\{\frac{E_d}{V_0}\right\}\right)^\mathsf{T}$$

is an Archimedean copula with generator $\psi_0 \in \boldsymbol{\Psi}_\infty$.

Proof

We have

$$\begin{aligned}
\mathbb{P}(\boldsymbol{U} \leq \boldsymbol{u}) &= \int_0^\infty \mathbb{P}(\boldsymbol{U} \leq \boldsymbol{u} | V_0 = v_0)\, dF_0(v_0) = \int_0^\infty \prod_{i=1}^d G_0(u_i; v_0)\, dF_0(v_0) \\
&= \int_0^\infty \exp[-v_0\{\psi_0^{-1}(u_1) + \cdots + \psi_0^{-1}(u_d)\}]\, dF_0(v_0) \\
&= \psi_0\{\psi_0^{-1}(u_1) + \cdots + \psi_0^{-1}(u_d)\}.
\end{aligned}$$

□

For statistical applications it is desirable to be able to evaluate the density of a multivariate model (for parameter estimation and goodness-of-fit testing, for example). For Archimedean copulas, the density (if it exists) is theoretically trivial to write down; for (2.1), one obtains

$$c(\boldsymbol{u}) = \psi^{(d)}\left\{\sum_{j=1}^d \psi^{-1}(u_j)\right\} \prod_{j=1}^d (\psi^{-1})'(u_j), \ \boldsymbol{u} \in (0,1)^d.$$

However, the generator derivatives $\psi^{(d)}$ are non-trivial to access theoretically and, even more, computationally. This issue has recently been solved for several well-known Archimedean copulas and transformations of such; see Hofert et al. (2011b) or Section 2.3.

Concerning measures of association, Genest and MacKay (1986) provided a semi-closed form to link the generator of a copula C and the Kendall's tau of C, while Joe and Hu (1996) derived a formula for the lower and upper tail coefficients λ_l and λ_u. These are given by

$$\tau = 4\int_0^1 \frac{\psi^{-1}(t)}{(\psi^{-1})'(t)}\,dt + 1, \tag{2.4}$$
$$\lambda_l = \lim_{t\to\infty} \frac{\psi(2t)}{\psi(t)} = 2\lim_{t\to\infty} \frac{\psi'(2t)}{\psi'(t)},$$
$$\lambda_u = 2 - \lim_{t\to 0_+} \frac{1-\psi(2t)}{1-\psi(t)} = 2 - 2\lim_{t\to 0_+} \frac{\psi'(2t)}{\psi'(t)}.$$

As stated before, the interest of using copulas in higher dimensions is limited because of the functional symmetry. A way to add more flexibility to Archimedean copulas in high dimension is given by *Khoudraji-transformed Archimedean copulas*, defined as

$$C(\boldsymbol{u}) = C_\psi(u_1^{\alpha_1}, \ldots, u_d^{\alpha_d})\,\Pi(u_1^{1-\alpha_1}, \ldots, u_d^{1-\alpha_d}), \quad u \in I^d,$$

where C_ψ is an Archimedean copula generated by $\psi \in \boldsymbol{\Psi}$, Π is the independence copula, and $\alpha_j \in [0,1]$ for $j \in \{1,\ldots,d\}$. Another technique to introduce asymmetry is given by nested Archimedean copulas, which is the subject of the next section.

2.2 Nested Archimedean Copulas

In simple terms, a nested Archimedean copula is an Archimedean copula whose arguments can be themselves Archimedean copulas. This construction allows for asymmetries among the pairs of components.

Definition 2.4. A d-dimensional copula C is called *nested Archimedean* if it is an Archimedean copula with arguments possibly replaced by other nested Archimedean copulas. If C is recursively given by (2.1) for $d = 2$ and

$$C(\boldsymbol{u};\psi_0,\ldots,\psi_{d-2}) = \psi_0[\psi_0^{-1}(u_1) + \psi_0^{-1}\{C(u_2,\ldots,u_d;\psi_1,\ldots,\psi_{d-2})\}]$$

for $d \geq 3$, then C is called a fully nested Archimedean copula with $d-1$ nesting levels or hierarchies. Otherwise, C is called a *partially-nested Archimedean copula*

In the present work, we are mostly interested in a partially-nested Archimedean copula of the form

$$C(\boldsymbol{u}) = C_0\{C_1(\boldsymbol{u_1}),\ldots,C_{d_0}(\boldsymbol{u}_{d_0})\}, \quad \boldsymbol{u} \in I^d, \tag{2.5}$$

where, for all $s_1 \in \{1, \ldots, d_0\}$, each C_{s_1} is a d_{s_1}-dimensional copula generated by ψ_{s_1}, $\sum_{s_1=1}^{d_0} d_{s_1} = d$ and $\boldsymbol{u}_{s_1} = (u_{s_1 1}, \ldots, u_{s_1 d_{s_1}})^\top$. Its structure can be depicted in the form a tree as in Figure 2.1.

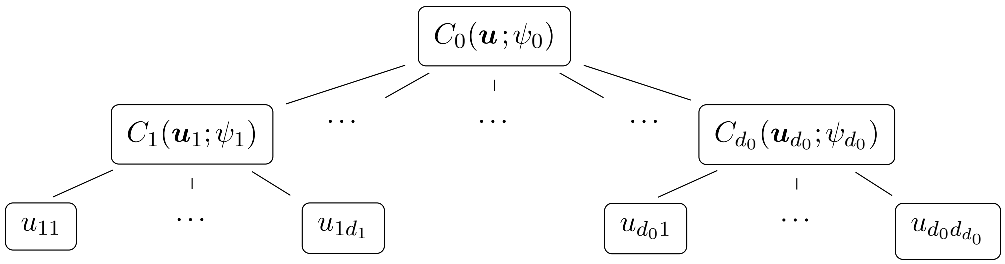

Figure 2.1: Tree structure of a d-dimensional partially nested Archimedean copula with $\sum_{s_1=1}^{d_0} d_{s_1} = d$.

Observe that the copula (2.5) has the following alternative representations. If one denotes

$$\mathring{\psi}_{0s_1} = \psi_0^{-1} \circ \psi_{s_1}$$

then Equation (2.5) can be developed into

$$\begin{aligned} C(\boldsymbol{u}) &= \psi_0\left(\sum_{s_1=1}^{d_0} \psi_0^{-1}\left[\psi_{s_1}\left\{ \sum_{s_2=1}^{d_{s_1}} \psi_{s_1}^{-1}(u_{s_1 s_2}) \right\}\right]\right) = \psi_0\left[\sum_{s_1=1}^{d_0} \mathring{\psi}_{0s_1}\{t_{s_1}(\boldsymbol{u}_{s_1})\}\right] \\ &= \int_0^\infty \exp\left(-v_0\left[\sum_{s_1=1}^{d_0} \mathring{\psi}_{0s_1}\{t_{s_1}(\boldsymbol{u}_{s_1})\}\right]\right) dF_0(v_0) \\ &= \int_0^\infty \prod_{s_1=1}^{d_0} \psi_{0s_1}\{t_{s_1}(\boldsymbol{u}_{s_1}); v_0\}\, dF_0(v_0), \end{aligned}$$

where, for all $s_1 \in \{1, \ldots, d_0\}$

$$\psi_{0s_1}(t; v_0) = \exp[-v_0 \psi_0\{\psi_{s_1}(t)\}] = \exp\{-v_0 \mathring{\psi}_{0s_1}(t)\}$$

and

$$t_{s_1}(\boldsymbol{u}_{s_1}) = \sum_{s_2=1}^{d_{s_1}} \psi_{s_1}^{-1}(u_{s_1 s_2}).$$

We refer to ψ_{0s_1} as *inner generator*. It is a proper generator in t for each $v > 0$ as a composition of the completely monotone function $f(t) = \exp(-vt)$ with $\mathring{\psi}_{0s_1}$ which has a completely monotone derivative, since $\theta_0 \leq \theta_{s_1}$.

The copula C_0 is referred to as *root* (or *outer*) *copula* and each C_s, $s \in \{1, \ldots, d_0\}$, as *sector copula*. Model (2.5) provides an intuitive hierarchical structure, since, for example,

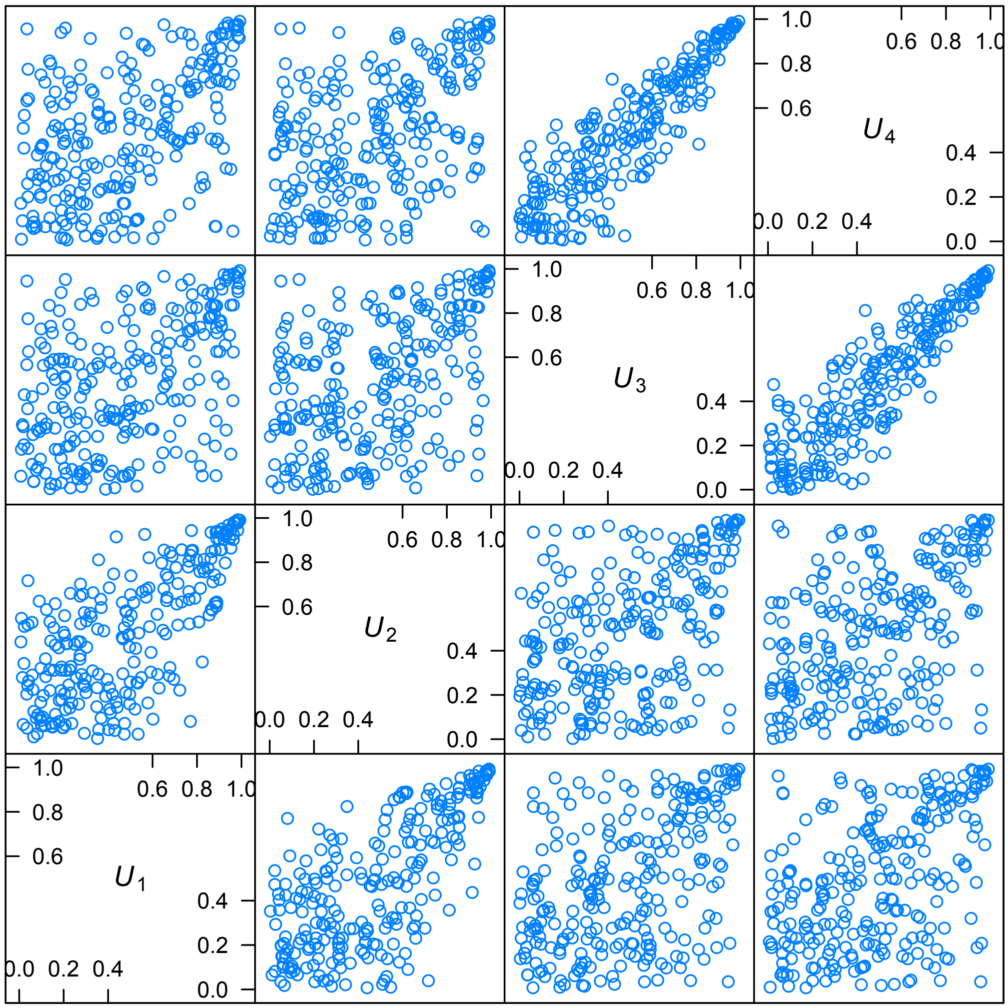

Figure 2.2: Pairwise scatterplots of 250 simulated points from a four-dimensional hierarchical nested Archimedean copula $C_0\{C_1(u_1,u_2),C_2(u_3,u_4)\}$ with Gumbel generators. $\boldsymbol{\theta} = (\theta_0,\theta_1,\theta_2)^\mathsf{T}$ corresponds to Kendall's tau $\boldsymbol{\tau} = (\tau_0,\tau_1,\tau_2)^\mathsf{T} = (0.3, 0.5, 0.7)^\mathsf{T}$

if $\boldsymbol{U} \sim C$ the pair $(U_{sj}, U_{sk})^\mathsf{T}$ $(j \neq k)$ has joint copula C_s whereas the pair $(U_{rj}, U_{sk})^\mathsf{T}$ $(r \neq s)$ follows the root copula C_0. Thanks to the R package `nacopula`, a scatterplot of 250 realization of a random variable $\boldsymbol{U} \sim C_0\{C_1(u_1, u_2), C_2(u_3, u_4)\}$ is provided in Figure **??**.

More complicated nesting structures can be constructed. For example, one can obviously expand copula (2.5), by replacing each (Archimedean) copula C_{s_1} by a nested Archimedean copula of the form (2.5). Hofert and Scherer (2011) used (2.5) to price Credit Default Swaps (CDOs), complex financial derivatives, which take effect when entities default on their loans.

The class of nested Archimedean copulas was first considered in (Joe, 1997, pp. 87) in the three- and four-dimensional case and later by McNeil (2008) in the general d-dimensional cases. McNeil (2008) and Hofert (2011a) derive an explicit stochastic representation for nested Archimedean copulas which allows for fast sampling similar to the Marshall-Olkin algorithm for Archimedean copulas. Hofert (2011b) provides efficient sampling strategies for the most important ingredients to this algorithm, the random variables responsible for introducing hierarchical dependencies. An implementation for several well-known Archimedean families (and transformations of such) is provided by the R package `nacopula`; see Hofert and Mächler (2011).

A sufficient condition under which (2.5) is indeed a proper copula is that the *node*

$$\mathring{\psi}_{0s} = \psi_0^{-1} \circ \psi_s$$

has a completely monotone derivative; see McNeil (2008). Note that this *sufficient nesting condition* is indeed only sufficient but not necessary. For example, if $\psi_0(t) = -\log(1-(1-e^{-\theta_0})\exp(-t))/\theta_0$ denotes the generator of a Frank copula and $\psi_1(t) = (1+t)^{-1/\theta_1}$ the generator of a Clayton copula, then $C(\boldsymbol{u}) = C_0(u_1, C_1(u_2, u_3))$ is a valid (nested Archimedean) copula for all θ_0, θ_1 such that $\theta_0/(1-e^{-\theta_0}) - 1 \leq \theta_1$ although $\mathring{\psi}_{01}$ is not completely monotone for all parameters θ_0, θ_1. Following this example, one could ask the following question: If one sets the outer generator ψ_0, what is the set of possible compatible generators ψ_s so that $\mathring{\psi}_{0s}$ has completely monotone derivative? Hering et al. (2010) gives the answer and the set is given by

$$\begin{aligned} M_{\psi_0} = \Big\{ \psi_s | \psi_s(x) = \psi_0\{\mu x + \int_0^\infty (1-e^{xt})\,\nu(dt)\}, \\ \text{where } \mu \geq 0 \text{ and } \nu \text{ is a measure on } (0,\infty) \text{ satysfing} \int_0^\infty \min\{1,t\}\,\nu(dt) < 0, \\ \text{and either } \mu > 0 \text{ or } \nu\{(0,1)\} = \infty, \text{ or both} \Big\}. \end{aligned}$$

As mentioned before, although nesting is possible in more complicated ways, we will mainly focus on nested Archimedean copulas of Type (2.5) (with some sectors possibly shrunk to single arguments of C_0) and assume that the Archimedean generators ψ_s, $s \in \{0, \ldots, d_0\}$, belong to Ψ_∞, the set of all Archimedean generators which are completely monotone.

Table 2.I: Generator and its dth derivaties of the Gumbel, Clayton, Ali-Mikail-Haq , Frank and Joe copulas.

Family	Parameter	Generator $\psi(t)$	Derivative $(-1)^d\psi^{(d)}(t)$
G	$\theta \in [1,\infty)$	$\exp(-t^{\frac{1}{\theta}})$	$\psi(t)\sum_{k=1}^{d} s_{dk}(1/\theta)\ t^{k/\theta-d}$
C	$\theta \in (0,\infty)$	$(1+t)^{-\frac{1}{\theta}}$	$(d-1+\frac{1}{\theta})_d\,(1+t)^{-(d+\frac{1}{\theta})}$
AMH	$\theta \in [0,1)$	$(1-\theta)(e^t-\theta)^{-1}$	$(1-\theta)\mathrm{Li}_{-d}(\theta e^{-t})/\theta$
J	$\theta \in [1,\infty)$	$1-(1-e^{-t})^{\frac{1}{\theta}}$	$\frac{1}{\theta}\sum_{k=1}^{d} a_{dk}^{J}(\theta)\,e^{-kt}(1-e^{-t})^{k+1/\theta}$
F	$\theta \in (0,\infty)$	$-\log\{1-(1-e^{-\theta})e^{-t}\}/\theta$	$\mathrm{Li}_{-(d-1)}\{(1-e^{-\theta})e^{-t}\}/\theta$

2.3 Parametric Archimedean Families

Among the most widely used parametric Archimedean families are those of Ali-Mikhail-Haq, Clayton, Frank, Gumbel, and Joe. These five families already provide a great flexibility in statistical modeling. A sample of one thousand points are provided in Figure 2.3, 2.4 and 2.5 in order to show on the one hand, how bivariate random variables can differ even with the same measure of association and on the other hand to show the differences of these families. Ali-Mikhail-Haq copula is provided but the range of its Kendall's tau is restricted to $[0,1/3)$.

These one-parameter families can easily be extended to allow for more parameters. A great amount of work and effort have been done to construct new copulas, by creating generators with more parameters. Hofert (2010) provided a transformation called the *tilted outer power transformation* which consists of using the generator $\tilde{\psi}$ defined as

$$\tilde{\psi}(t) = \psi\{(c^\alpha + t)^\alpha - c\},$$

where $\psi \in \mathbf{\Psi}_\infty$, $\alpha \geq 1$ and $c \geq 0$. This yields a proper generator because the function $(c^\alpha + t)^\alpha - c$ is nonnegative and has a completely monotone derivative. Since the composition of two completely monotone functions is still completely monotone, $\tilde{\psi}$ is indeed completely monotone. If $c = 0$, then it yields the commonly called *outer power* transformations. This paradoxical nomenclature comes from the fact that this transformation was first applied to φ.

Using a slightly different version of (2.4), Hofert (2010) derived a semi-closed formula for the Kendall's tau of the copula generated by $\tilde{\psi}$ and showed that with the Clayton copula, one could set parameters such that the theoretical values of both tails coefficients match their respective estimators. Alternatively, one could also fit a parameter in order to reproduce the value of one of the tails coefficients and then fit the second in order to match the theoretical value of the Kendall's tau to the observations.

It is interesting to observe that the densities of the Frank, Gumbel, Joe and Ali-Mikhail-Haq copulas have only been discovered recently in Hofert et al. (2011b). Numerical implementation of these also requires a lot of attention, since in high dimensions, the density function numerically explodes for a large range of inputs.

Table 2.II: Inverse of the generator and its first derivative of the Gumbel, Clayton, Ali-Mikail-Haq , Frank and Joe copula.

Family	Inverse $\psi^{-1}(t)$	Derivative of the Inverse $(\psi^{-1})'(t)$
G	$(-\log t)^{\theta}$	$\theta(-\log t)^{\theta-1}/t$
C	$t^{-\theta}-1$	$\theta t^{-(1+\theta)}$
AMH	$\log\{1-\theta(1-t)\}-\log(t)$	$(1-\theta)/[t\{1-\theta(1-t)\}]$
J	$-\log\{1-(1-t)^{\theta}\}$	$\theta(1-t)^{\theta-1}/\{1-(1-t)^{\theta}\}$
F	$-\log\{1-\exp(-\theta t)\}+\log\{1-\exp(-\theta)\}$	$\theta\exp(-\theta t)/\{1-\exp(-\theta t)\}$

Table 2.III: Inner generator and the distribution $F=\mathcal{LS}^{-1}[\psi]$ of the Gumbel, Clayton, Ali-Mikail-Haq , Frank and Joe copula.

Family	$\psi_{0s_1}(t;v_0),\ \theta_0\leq\theta_{s_1}$	$F=\mathcal{LS}^{-1}[\psi]$
G	$\exp(-v_0 t^{\theta_0/\theta_{s_1}})$	$\mathrm{S}(1/\theta,1,cos^{\theta}\{\pi/2\theta)\},I_{\{\theta=1\}};1)$
C	$\exp\{-v_0((1+t)^{\theta_0/\theta_{s_1}}-1)\}$	$\Gamma(\theta_0/\theta_s,1)$
AMH	$(1-\theta_{s_1})^{v_0}\{(1-\theta_0)\exp(t)-(\theta_{s_1}-\theta_0)\}^{-v_0}$	$\mathrm{Geo}(1-\theta)$
J	$\{1-(1-e^{-t})^{\theta_0/\theta_{s_1}}\}^{v_0}$	$\mathrm{Sibuya}(\alpha)$
F	$(1-e^{-\theta_0})^{-v_0}\,[1-\{1-(1-e^{\theta_{s_1}})e^{-t}\}]^{1/\theta}$	$\mathrm{Log}(1-e^{-\theta})$

Tables 2.I and 2.III present known one-parameter Archimedean families with generators $\psi\in\mathbf{\Psi}_{\infty}$. In these tables, one can find:

– an abbreviation of the name of the family of copula, that is "G" (Gumbel), "C" (Clayton), "AMH" (Ali-Mikhail-Haq), "J" (Joe), "F" (Frank);

– the parameter space of the family;

– the families' generator, their inverse and their derivatives;

– the explicit form of $\psi_{0s_1}(t;v_0)=\exp[-v_0\psi_0^{-1}\{\psi_{s_1}(t)\}]$ where ψ_0 and ψ_{s_1} are from the same family and the parameter respect $\theta_0\leq\theta_{s_1}$;

– the distribution $F=\mathcal{LS}^{-1}[\psi]$ whose Laplace-Stieltjes transform yields the generator ψ.

Observe that the function $\psi_{0s}(t;v_0)$ will play a central role when the density of the nested Archimedean copula (2.5) will be computed. The condition $\theta_0\leq\theta_s$ is a compulsory restriction if one desires $\mathring{\psi}_{0s}=\psi_0^{-1}\circ\psi_s$ to be completely monotone.

For Gumbel's family, the function $s_{dk}(x)$ appearing in the derivative of ψ is defined as

$$s_{dk}(x) = (-1)^{d-k} \sum_{j=k}^{d} x^j s(d,j) S(j,k)$$

where s and S are the *Stirling numbers of the first and second kind*, respectively, given by the recurrence relations

$$\begin{aligned} s(n+1,k) &= s(n,k-1) - ns(n,k), \\ S(n+1,k) &= S(n,k-1) + kS(n,k), \end{aligned}$$

for all $k \in \mathbb{N}$, $n \in \mathbb{N}_0$, with $s(0,0) = S(0,0) = 1$ and $s(n,0) = s(0,n) = S(n,0) = S(0,n) = 0$ for all $n \in \mathbb{N}$. Observe that the function $s_{nk}(\theta)$ is numerically challenging to evaluate, however this issue is resolved in the R package `nacopula`. The distribution $F = \mathcal{LS}^{-1}[\psi]$ of the Gumbel family is the *stable distribution* $\mathrm{S}(\alpha,\beta,\gamma,\delta,;1)$. An heuristic definition of a stable distribution is that if X and Y are of the same type, that is, one is a location scale transformation of the other, then their sum $X+Y$ is still of the same type; see (Nolan, 2011, p. 4) for more details. It is also usual to define the stable distribution $\mathrm{S}(\alpha,\beta,\gamma,\delta,;1)$ by its characteristic function

$$\phi(t) = \exp[i\delta t - \gamma^\alpha |t|^\alpha \{1 - i\beta\, \mathrm{sign}(t) w(t,\alpha)\}]$$

where $w(t,\alpha) = 1(\alpha \neq 1)\tan(\alpha\pi/2) - 1(\alpha = 1) 2\log(|t|)/\pi$. Stable laws have been extensively used to model financial (log-)returns thanks to their flexibility. Note that the Gumbel family is the only Archimedean *extreme-value copulas* that is $C^t(\boldsymbol{u}^{1/t}) = C(\boldsymbol{u})$ for all positive t, see (Nelsen, 2007, p. 143). Interestingly, Π, the independence copula, can be considered as member of the Gumbel family. Indeed, the Gumbel copula with parameter $\theta = 1$ has as generator $\psi(t) = \exp(-t)$, which is the generator of Π.

Concerning the Clayton copula, $(d)_k$ denotes the so-called falling factorial which is given by

$$d(d-1)\cdot(d-k+1) = \frac{d!}{(d-k)!} = k!\binom{d}{k} = \frac{\Gamma(d+1)}{\Gamma(d-k+1)},$$

where

$$\Gamma(\alpha) = \int_0^\infty x^{\alpha-1} e^{-x} dx$$

is the celebrated Euler's gamma function. $\Gamma(\alpha,\beta)$ denotes the *gamma distribution* with shape parameter $\alpha \geq 0$ and scale $\beta > 0$. This law is often represented with its density function given by

$$f(x) = \frac{x^{\alpha-1}}{\Gamma(\alpha)} \beta^\alpha e^{-\beta x}.$$

The Clayton copula appeared in Clayton (1978), however bears the name of *Cook and Johnson* copula in Genest and MacKay (1986).

The Ali-Mikhail-Haq copula appeared in Ali et al. (1978) in the context of survival

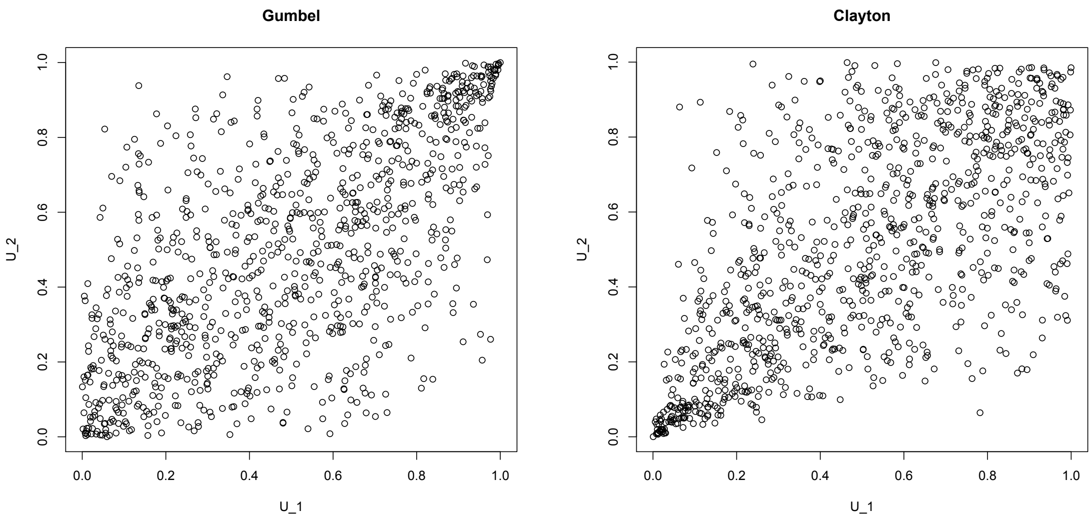

Figure 2.3: One thousand simulated points from the Gumbel and Clayton copula with parameter θ corresponding to a Kendall's tau of 0.5.

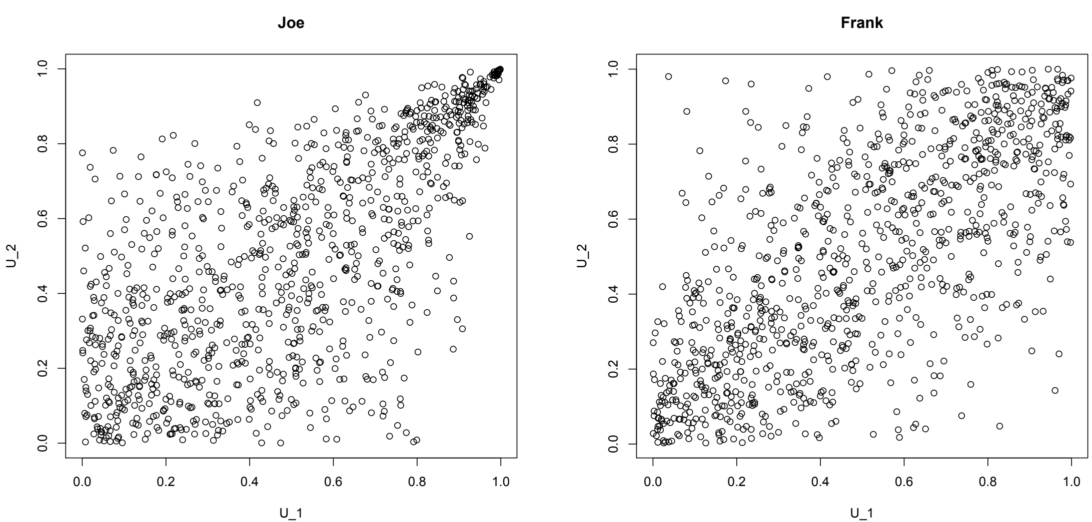

Figure 2.4: One thousand simulated points from the Joe and Frank copula with parameter θ corresponding to a Kendall's tau of 0.5.

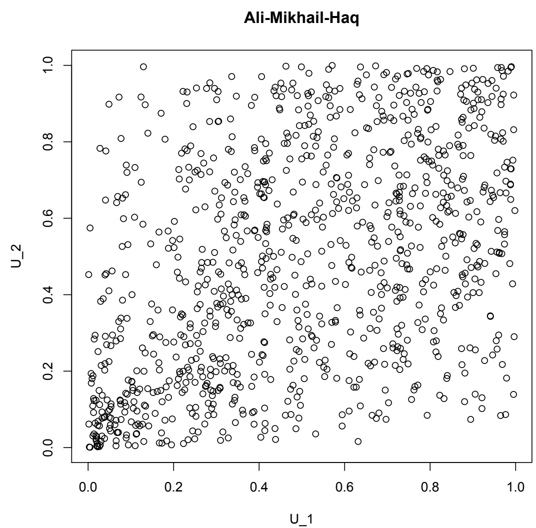

Figure 2.5: One thousand simulated points from the Ali-Mikhail-Haq Copula with parameter θ corresponding to a Kendall's tau of 0.3

analysis. The function $\mathrm{Li}_{-(d-1)}(x)$ denotes the Polylogarithm function, also known as *Jonquière's function*, and is defined for $\alpha \in \mathbb{Z}$ as

$$\mathrm{Li}_\alpha(x) = \sum_{k=1}^{\infty} x^k / k^\alpha.$$

For $d \in \mathbb{N}$, Wood (1992) derived the alternative expression

$$\mathrm{Li}_{-d}(x) = (-1)^{d+1} \sum_{k=1}^{d+1} (k-1)! S(d+1,k) \left(\frac{-1}{1-x} \right)^k.$$

The inverse Laplace-transform of ψ is given by the *geometric distribution* $\mathrm{Geo}(p)$ with probability of success $p \in (0,1]$. This discrete law is usually given by its probability mass function $p_k = (1-p)^{k-1} p$ for $k \in \mathbb{N}$.

Concerning the Joe copula, it appeared in Joe (1993). The function $a_{dk}^J(\theta)$ appearing in the generator's derivative is defined by

$$a_{dk}^J(\theta) = S(d,k)(k-1-1/\theta)_{k-1} = S(d,k) \frac{\Gamma(k-1/\theta)}{\Gamma(1-1/\theta)}, \quad k \in \{1,\dots,d\}.$$

The distributional law originating Joe copula's generator is the discrete *Sibuya distribution* $\mathrm{Sibuya}(p)$ with probability mass function $p_k = \binom{p}{k}(-1)^{k-1}$ for $k \in \mathbb{N}$.

Finally, Frank's family was studied in Frank (1979) where the author also shows that this family is the only radially symmetric family among the Archimedean copulas. The distribution $F = \mathcal{LS}^{-1}[\psi]$ is the *logarithmic law* $\mathrm{Log}(p)$, with parameter $p \in (0,1]$ and probability mass function $p_k = p^k/\{-k\log(1-p)\}$, for $k \in \mathbb{N}$.

The interested reader can consult Hofert et al. (2011b), Hofert (2010) and the references therein for explicit or semi-explicit forms for Kendall's tau or tail-dependence coefficients, as well as for more details about the content of this section.

CHAPTER 3

DENSITIES OF NESTED ARCHIMEDEAN COPULAS

In this chapter, we first recall the notation and restate some useful results. We then tackle the problem by first deriving a convenient form for nested Archimedean copulas of Type (3.1) described below. This will allow us to compute their densities. The general strategy is presented in Section 3.2 and all necessary details for several well-known Archimedean families are provided. Section 3.4 addresses numerical evaluation of the log-density. Section 3.5 studies the behavior of the density when we set a third level of nesting.

3.1 Introduction

There has recently been interest in multivariate hierarchical models, that is, models that are able to capture different dependencies between and within different groups of random variables. One such class of models is based on nested Archimedean copulas. A *partially nested Archimedean copula* C with two nesting levels and d_0 groups, is given by

$$C(\boldsymbol{u}) = C_0\{C_1(\boldsymbol{u}_1), \dots, C_{d_0}(\boldsymbol{u}_{d_0})\},\ \boldsymbol{u} = (\boldsymbol{u}_1, \dots, \boldsymbol{u}_{d_0})^{\mathsf{T}}, \tag{3.1}$$

where d_0 denotes the dimension of C_0 and each copula C_s, $s \in \{0, \dots, d_0\}$, is Archimedean with some *generator* ψ_s, that is,

$$C_s(\boldsymbol{u}_s) = \psi_s\{\psi_s^{-1}(u_{s1}) + \dots + \psi_s^{-1}(u_{sd_s})\} = \psi_s\{t_s(\boldsymbol{u}_s)\},\ s \in \{0, \dots, d_0\}; \tag{3.2}$$

here and in the following we denote

$$t_s(\boldsymbol{u}_s) = \sum_{j=1}^{d_s} \psi_s^{-1}(u_{sj}).$$

The copula C_0 is referred to as *root* (or *outer*) *copula* and each C_s, $s \in \{1, \dots, d_0\}$, as *sector copula*. For Archimedean copulas, the density (if it exists) is theoretically trivial to write down; for (3.2), one obtains

$$c_s(\boldsymbol{u}_s) = \psi_s^{(d)}\{t_s(\boldsymbol{u}_s)\} \prod_{j=1}^{d_s} (\psi_s^{-1})'(u_{sj}),\ \boldsymbol{u}_s \in (0,1)^d.$$

However, as already stressed, the appearing generator derivatives $\psi_s^{(d)}$ are non-trivial to access theoretically and, even more, computationally. This issue has recently been solved for several well-known Archimedean copulas and transformations of such; see Hofert et al. (2011b). Our goal is to extend these results to corresponding nested Archimedean copulas. Note that this is more challenging because differentiating (3.1) is more complicated due to the inner derivatives that appear when applying the Chain Rule; in contrast to

Archimedean copulas, these inner derivatives depend on variables with respect to which one has to differentiate again. Already in low dimensions the corresponding formulas for the density c become challenging to write down and, even more, to evaluate in a numerically stable way.

3.2 Inner generator derivatives and densities for two-level nested Archimedean copulas

3.2.1 The basic idea

Let C be a d-dimensional nested Archimedean copula of Type (3.1) (with some child copulas possibly shrunk to single arguments of C_0) and assume the sufficient nesting condition to hold; for the families Ali–Mikhail–Haq, Clayton, Frank, Gumbel, and Joe, this is fulfilled as long as all generators belong to the same family and $\theta_0 \le \theta_s$, $s \in \{1, \dots, d_0\}$. This condition implies that each copula C_s, $s \in \{1, \dots, d_0\}$, is more concordant than C_0.

One of the main ingredients we need in the following is the function

$$\psi_{0s}(t; v) = \exp\{-v \mathring{\psi}_{0s}(t)\} \tag{3.3}$$

which we refer to as *inner generator*. It is a proper generator in t for each $v > 0$ as a composition of the completely monotone function $f(t) = \exp(-vt)$ with $\mathring{\psi}_{0s}$ which has a completely monotone derivative. The corresponding inverse is $\psi_{0s}^{-1}(u; v) = \mathring{\psi}_{0s}^{-1}\{(-\log u)/v\} = \psi_s^{-1}[\psi_0\{(-\log u)/v\}]$, so that

$$\psi_s^{-1}(u) = \psi_{0s}^{-1}\{G_0(u; v); v\},$$

where $G_0(u; v) = \exp\{-v\psi_0^{-1}(u)\}$. Note that G_0 is a distribution function in $u \in [0, 1]$ for any $v > 0$. With $F_0 = \mathcal{LS}^{-1}[\psi_0]$, we thus obtain

$$C(\boldsymbol{u}) = C_0\{C_1(\boldsymbol{u}_1), \dots, C_{d_0}(\boldsymbol{u}_{d_0})\} = \int_0^\infty \exp\Big[-v_0 \sum_{s=1}^{d_0} \mathring{\psi}_{0s}\{t_s(\boldsymbol{u}_s)\}\Big]\, dF_0(v_0)$$

$$= \int_0^\infty \prod_{s=1}^{d_0} \psi_{0s}\{t_s(\boldsymbol{u}_s); v_0\}\, dF_0(v_0) \tag{3.4}$$

$$= \int_0^\infty \prod_{s=1}^{d_0} \psi_{0s}\Big[\sum_{j=1}^{d_s} \psi_{0s}^{-1}\{G_0(u_{sj}; v_0); v_0\}; v_0\Big]\, dF_0(v_0). \tag{3.5}$$

Representation (3.5) provides a simple F_0-mixture representation of C of the form

$$C(\boldsymbol{u}) = \int_0^\infty \prod_{s=1}^{d_0} C_{0s}\{G_0(u_{s1}; v_0), \dots, G_0(u_{sd_s}; v_0); v_0\}\, dF_0(v_0) \tag{3.6}$$

$$= \mathbb{E}\Big[\prod_{s=1}^{d_0} C_{0s}\{G_0(u_{s1}; V_0), \dots, G_0(u_{sd_s}; V_0); V_0\}\Big], \tag{3.7}$$

where C_{0s} denotes the Archimedean copula generated by ψ_{0s}, $s \in \{1, \dots, d_0\}$.

Now let us turn to the density and note that it exists by our assumption of having completely monotone generators. In general, the density is given by

$$c(\boldsymbol{u}) = \frac{\partial^d}{\partial u_{d_0 d_{d_0}} \cdots \partial u_{11}} C(\boldsymbol{u})$$

but instead of differentiating (3.1) directly, the idea is to use Representation (3.6). By differentiating under the integral sign one obtains

$$c(\boldsymbol{u}) = \int_0^\infty \prod_{s=1}^{d_0} \Big[c_{0s}\{G_0(u_{s1}; v_0), \dots, G_0(u_{sd_s}; v_0); v_0\} \prod_{j=1}^{d_s} g_0(u_{sj}; v_0) \Big] \, dF_0(v_0),$$

where c_{0s} denotes the density of C_{0s}, $s \in \{1, \dots, d_0\}$, and g_0 the one of G_0. For the cost of one integral (which will be computed explicitly below!), one can easily compute c (theoretically) as a F_0-mixture of the densities of the multivariate distributions specified by C_{0s} with margins G_0. This is especially advantageous in large dimensions as the complexity of the problem does not (again, theoretically) depend on the sizes of the child copulas too much, rather on the number of children.

From a numerical point of view, it is more efficient to work with (3.4) instead of (3.6). By doing so, c allows for the representation

$$\begin{aligned} c(\boldsymbol{u}) &= \int_0^\infty \prod_{s=1}^{d_0} \psi_{0s}^{(d_s)}\{t_s(\boldsymbol{u}_s); v_0\} \, dF_0(v_0) \cdot \prod_{s=1}^{d_0} \prod_{j=1}^{d_s} (\psi_s^{-1})'(u_{sj}) \\ &= \mathbb{E}\Big[\prod_{s=1}^{d_0} \psi_{0s}^{(d_s)}\{t_s(\boldsymbol{u}_s); V_0\} \Big] \cdot \prod_{s=1}^{d_0} \prod_{j=1}^{d_s} (\psi_s^{-1})'(u_{sj}). \end{aligned} \tag{3.8}$$

From Equation (3.8), we identify the following key challenges:

Challenge 1 Find the derivatives of the inner generators $\psi_{0s}(t; v_0)$;

Challenge 2 Compute their product;

Challenge 3 Integrate it with respect to the mixture distribution function $F_0 = \mathcal{LS}^{-1}[\psi_0]$.

All three challenges will be solved in Section 3.2.3 with the help of the tools presented in the following section.

3.2.2 The tools needed: Faà di Bruno's formula and Bell polynomials

One formula which proves to be useful here, is the expression of the nth derivative of a composition of functions, named after the mathematician Faà di Bruno. For suitable functions f and g, Faà di Bruno's formula states that

$$(f \circ g)^{(n)}(x) = \sum_{k=1}^{n} f^{(k)}\{g(x)\} \sum_{\boldsymbol{j} \in \mathcal{P}_{n,k}} \binom{n}{j_1, \dots, j_{n-k+1}} \prod_{l=1}^{n-k+1} \Big(\frac{g^{(l)}(x)}{l!} \Big)^{j_l}, \tag{3.9}$$

where $\binom{n}{j_1,\ldots,j_n} = \frac{n!}{j_1!\cdot\ldots\cdot j_n!}$ denotes a multinomial coefficient, $\boldsymbol{j} = (j_1,\ldots,j_n)^\mathsf{T} \in \mathbb{N}_0^n$, and

$$\mathcal{P}_{n,k} = \left\{ \boldsymbol{j} \in \mathbb{N}_0^{n-k+1} : \sum_{i=1}^{n-k+1} i j_i = n \text{ and } \sum_{i=1}^{n-k+1} j_i = k \right\}. \tag{3.10}$$

Alternatively, one can use *Bell polynomials* to reformulate (3.9). These are defined by

$$B_{n,k}(x_1,\ldots,x_{n-k+1}) = \sum_{\boldsymbol{j}\in\mathcal{P}_{n,k}} \binom{n}{j_1,\ldots,j_{n-k+1}} \prod_{l=1}^{n-k+1} \left(\frac{x_l}{l!}\right)^{j_l}. \tag{3.11}$$

This implies that (3.9) can be written as

$$(f\circ g)^{(n)}(x) = \sum_{k=1}^{n} f^{(k)}\{g(x)\} B_{n,k}\{g'(x), g''(x), \ldots, g^{(n-k+1)}(x)\} \tag{3.12}$$

In the sections to come, we frequently need the following results. We recall that $(x)_n = x\cdot(x-1)\cdot\ldots\cdot(x-n+1)$ denotes the *falling factorial*, and $s(n,k)$ and $S(n,k)$ denote the *Stirling numbers of the first* and *second kind.*

Lemma 3.1. Let $B_{n,k}$ be the Bell polynomial as in (3.11) and $n \in \mathbb{N}$. Then

(1) For $\boldsymbol{j} \in \mathcal{P}_{n,k}$, $\sum_{l=1}^{n-k+1}(x-l)j_l = xk - n$;

(2) $B_{n,k}(x,\ldots,x) = S(n,k)x^k$, $k \in \{0,\ldots,n\}$;

(3) $B_{n,k}(-x,\ldots,(-1)^{n-k+1}x) = (-1)^n S(n,k)x^k$, $k \in \{0,\ldots,n\}$;

(4) $\operatorname{sign}\left[B_{n,k}\{g'(x), g''(x), \ldots, g^{(n-k+1)}(x)\}\right] = (-1)^{n-k}$ for all x if g' is completely monotone.

Proof

(1) $\sum_{l=1}^{n-k+1}(x-l)j_l = x\sum_{l=1}^{n-k+1} j_l - \sum_{l=1}^{n-k+1} l j_l = xk - n.$

(2) The identity $B_{n,k}(1,\ldots,1) = S(n,k)$ can be found, for example, in (Comtet, 1974, p. 135) or (Charalambides, 2005, p. 87). It then follows that

$$\begin{aligned} B_{n,k}(x,\ldots,x) &= \sum_{\boldsymbol{j}\in\mathcal{P}_{n,k}} \binom{n}{j_1,\ldots,j_{n-k+1}} \prod_{l=1}^{n-k+1} \left(\frac{x}{l!}\right)^{j_l} \\ &= x^k B_{n,k}(1,\ldots,1) = S(n,k)x^k, \end{aligned}$$

since $\sum_{l=1}^{n-k+1} j_l = k$ by definition of $\mathcal{P}_{n,k}$.

(3) By definition of $\mathcal{P}_{n,k}$ it follows from $\sum_{l=1}^{n-k+1} lj_l = n$ and (2) that

$$B_{n,k}(-x,\ldots,(-1)^{n-k+1}x) = \sum_{\boldsymbol{j}\in\mathcal{P}_{n,k}} \binom{n}{j_1,\ldots,j_{n-k+1}} \prod_{l=1}^{n-k+1} \left\{\frac{(-1)^l x}{l!}\right\}^{j_l}$$
$$= (-1)^n B_{n,k}(x,\ldots,x) = (-1)^n S(n,k)x^k.$$

(4) By (3.11),

$$B_{n,k}\{g'(x),g''(x),\ldots,g^{(n-k+1)}(x)\} = \sum_{\boldsymbol{j}\in\mathcal{P}_{n,k}} \binom{n}{j_1,\ldots,j_{n-k+1}} \prod_{l=1}^{n-k+1} \left\{\frac{g^{(l)}(x)}{l!}\right\}^{j_l}. \tag{3.13}$$

Note that $\text{sign}(g^{(l)}(x)) = (-1)^{l-1}$ for all x and $l \in \{1,\ldots,n-k+1\}$, so that the sign of the lth factor in the product in (3.13) is $(-1)^{(l-1)j_l}$. This implies that

$$\text{sign} \prod_{l=1}^{n-k+1} \left\{\frac{g^{(l)}(x)}{l!}\right\}^{j_l} = (-1)^{\sum_{l=1}^{n-k+1}(l-1)j_l} = (-1)^{\sum_{l=1}^{n-k+1} lj_l - \sum_{l=1}^{n-k+1} j_l}.$$

Now since we sum over $\boldsymbol{j} \in \mathcal{P}_{n,k}$, we see from (3.10) that

$$\text{sign} \prod_{l=1}^{n-k+1} \left\{\frac{g^{(l)}(x)}{l!}\right\}^{j_l} = (-1)^{n-k}$$

and thus the whole sum in (3.13) has this sign.

□

Proposition 3.2. Let

$$s_{nk}(x) = \sum_{l=k}^{n} s(n,l)S(l,k)x^l = (-1)^n \sum_{l=k}^{n} |s(n,l)|S(l,k)(-x)^l.$$

Then

(1) For all $k \in \{0,\ldots,n\}$, $B_{n,k}((x)_1 y^{x-1},\ldots,(x)_{n-k+1} y^{x-(n-k+1)}) = y^{xk-n} s_{nk}(x)$;

(2) $\sum_{k=1}^{n} (-1)_k s_{nk}(x) = (-x)_n$;

(3) If $x \in (0,1]$, $\text{sign}\{s_{nk}(x)\} = (-1)^{n-k}$.

Proof

(1) Let $h(t) = \exp\{(yt)^x\}$. It follows from Faà di Bruno's formula ((3.9) and (3.12) with

$f(x) = \exp(x)$ and $g(t) = (yt)^x$) and Lemma 3.1 Part (1) that

$$\begin{aligned}
h^{(n)}(t) &= h(t) \sum_{k=1}^{n} \sum_{\boldsymbol{j} \in \mathcal{P}_{n,k}} \binom{n}{j_1, \dots, j_{n-k+1}} \prod_{l=1}^{n-k+1} \left(\frac{(x)_l y^x t^{x-l}}{l!} \right)^{j_l} \\
&= h(t) \sum_{k=1}^{n} \sum_{\boldsymbol{j} \in \mathcal{P}_{n,k}} \binom{n}{j_1, \dots, j_{n-k+1}} \prod_{l=1}^{n-k+1} \left(\frac{(x)_l y^x}{l!} \right)^{j_l} t^{xk-n} \\
&= h(t) \sum_{k=1}^{n} y^n \sum_{\boldsymbol{j} \in \mathcal{P}_{n,k}} \binom{n}{j_1, \dots, j_{n-k+1}} \prod_{l=1}^{n-k+1} \left(\frac{(x)_l y^{x-l}}{l!} \right)^{j_l} t^{xk-n} \\
&= \frac{h(t)}{t^n} \sum_{k=1}^{n} y^n B_{n,k}((x)_1 y^{x-1}, \dots, (x)_{n-k+1} y^{x-(n-k+1)}) t^{xk}.
\end{aligned}$$

On the other hand, we may differentiate the series expansion of h and use the identity $e^{-x} \sum_{k=0}^{\infty} k^l x^k / k! = \sum_{k=0}^{l} S(l,k) x^k$, see Boyadzhiev (2009), with x being $(yt)^x$. Furthermore, note that for $n \in \mathbb{N}$ (in particular $n \neq 0$), the Stirling numbers of the first kind satisfy

$$(x)_n = \sum_{j=1}^{n} s(n,j) x^j. \tag{3.14}$$

Applying these results leads to

$$\begin{aligned}
h^{(n)}(t) &= \sum_{k=0}^{\infty} (xk)_n \frac{t^{xk-n}}{k!} y^{xk} = \frac{1}{t^n} \sum_{k=0}^{\infty} \left\{ \sum_{l=1}^{n} s(n,l)(xk)^l \right\} \frac{(yt)^{xk}}{k!} \\
&= \frac{1}{t^n} \sum_{l=1}^{n} s(n,l) x^l \sum_{k=0}^{\infty} k^l \frac{(yt)^{xk}}{k!} = \frac{\exp\{(yt)^x\}}{t^n} \sum_{l=1}^{n} s(n,l) x^l \sum_{k=0}^{l} S(l,k)(yt)^{xk} \\
&= \frac{h(t)}{t^n} \sum_{k=0}^{n} \left\{ y^{xk} \sum_{l=k}^{n} x^l s(n,l) S(l,k) \right\} t^{xk} = \frac{h(t)}{t^n} \sum_{k=1}^{n} \left\{ y^{xk} \sum_{l=k}^{n} x^l s(n,l) S(l,k) \right\} t^{xk}.
\end{aligned}$$

Comparing the two representations for $h^{(n)}$ leads to the result as stated.

(2) Since $\sum_{k=1}^{j} (x)_k S(j,k) = x^j$ and by (3.14), one obtains that

$$\sum_{k=1}^{n} (-1)_k s_{nk}(x) = \sum_{j=1}^{n} s(n,j) x^j \sum_{k=1}^{j} (-1)_k S(j,k) = \sum_{j=1}^{n} s(n,j)(-x)^j = (-x)_n.$$

(3) For all $x \in (0,1]$, $s_{nk}(x) = (-1)^{n-k} p(n;k) n!/k!$, where the probability mass function $p(n;k) > 0$ (in $n \in \{k, k+1, \dots\}$) corresponds to the distribution function whose Laplace-Stieltjes transform is the inner generator appearing in a nested Joe copula; consider (Hofert, 2010, p. 99) with $V_0 = k$ and $\theta_0/\theta_1 = x$ to see this. This representation implies that $\operatorname{sign}\{s_{nk}(x)\} = (-1)^{n-k}$.

□

3.2.3 The main result

We are now able to derive a general formula for the derivatives of the inner generators and also for the density of nested Archimedean copulas of Type (3.1). It will follow from Faà di Bruno's formula that the derivatives of the inner generators $\psi_{0s}(t;v_0)$ are the inner generators themselves times a polynomial in $-v_0$. The product of these derivatives can then be computed as a Cauchy product. Interpreting the appearing quantities correctly allows us to compute the expectation with respect to F_0 via the derivatives of ψ_0. This solves all three of the above challenges.

Theorem 3.3 (Main theorem)**.** Let $\psi_s \in \Psi_\infty$, $s \in \{0,\dots,d_0\}$, such that $\mathring{\psi}_{0s}$ has a completely monotone derivative for all $s \in \{1,\dots,d_0\}$.

(1) For all $n \in \mathbb{N}$,

$$\psi_{0s}^{(n)}(t;v_0) = \psi_{0s}(t;v_0)\sum_{k=1}^{n} a_{s,nk}(t)(-v_0)^k, \tag{3.15}$$

where

$$a_{s,nk}(t) = B_{n,k}\{\mathring{\psi}_{0s}'(t),\dots,\mathring{\psi}_{0s}^{(n-k+1)}(t)\} \tag{3.16}$$

with $\operatorname{sign}\{a_{s,nk}(t)\} = (-1)^{n-k}$; if $\psi_s = \psi_0$ and $n = k = 1$ then $a_{s,nk}(t) = 1$ for all t.

(2) The density of (3.1) is given by

$$c(\boldsymbol{u}) = \Big[\sum_{k=d_0}^{d} b_{\boldsymbol{d},k}^{d_0}\{\boldsymbol{t}(\boldsymbol{u})\}\psi_0^{(k)}\{t(\boldsymbol{u})\}\Big]\cdot\prod_{s=1}^{d_0}\prod_{j=1}^{d_s}(\psi_s^{-1})'(u_{sj}), \tag{3.17}$$

where

$$\begin{aligned}\boldsymbol{t}(\boldsymbol{u}) &= (t_1(\boldsymbol{u}_1),\dots,t_{d_0}(\boldsymbol{u}_{d_0}))^{\mathsf{T}},\\ b_{\boldsymbol{d},k}^{d_0}\{\boldsymbol{t}(\boldsymbol{u})\} &= \sum_{\boldsymbol{j}\in\mathcal{Q}_{\boldsymbol{d},k}^{d_0}}\prod_{s=1}^{d_0} a_{s,d_sj_s}\{t_s(\boldsymbol{u}_s)\},\\ t(\boldsymbol{u}) &= \psi_0^{-1}\{C(\boldsymbol{u})\},\end{aligned} \tag{3.18}$$

with $\boldsymbol{d} = (d_1,\dots,d_{d_0})^{\mathsf{T}}$ and

$$\mathcal{Q}_{\boldsymbol{d},k}^{d_0} = \Big\{\boldsymbol{j}\in\mathbb{N}^{d_0} : \sum_{s=1}^{d_0} j_s = k,\ j_s \le d_s,\ s\in\{1,\dots,d_0\}\Big\};$$

that is, $b_{\boldsymbol{d},k}^{d_0}$ is a coefficient in the Cauchy product of the polynomials $\sum_{k=1}^{d_s} a_{s,d_sk}(t)(-v_0)^k$.

Proof

(1) Apply (3.12) with $f(x) = \exp(-v_0 x)$ and $g(t) = \mathring{\psi}_{0s}(t)$. For the statement about the signs, apply Lemma 3.1, Part (4). For the last statement, note that by Lemma 3.1, Part (2), $a_{s,11}(t) = B_{1,1}\{\mathring{\psi}'_{0s}(t)\} = B_{1,1}(1) = S(1,1)\cdot 1^1 = 1$ if $\psi_s = \psi_0$.

(2) Given the form (3.15) we see that the product appearing as integrand in (3.8) can be computed via

$$\begin{aligned}\prod_{s=1}^{d_0} \psi_{0s}^{(d_s)}\{t_s(\boldsymbol{u}_s); v_0\} &= \prod_{s=1}^{d_0}\sum_{k=1}^{d_s} a_{s,d_s k}\{t_s(\boldsymbol{u}_s)\}(-v_0)^k \cdot \prod_{s=1}^{d_0}\psi_{0s}\{t_s(\boldsymbol{u}_s); v_0\} \\ &= \sum_{k=d_0}^{d} b_{\boldsymbol{d},k}^{d_0}\{\boldsymbol{t}(\boldsymbol{u})\}(-v_0)^k \cdot \prod_{s=1}^{d_0}\psi_{0s}\{t_s(\boldsymbol{u}_s); v_0\}.\end{aligned}$$

Now note that $\prod_{s=1}^{d_0}\psi_{0s}\{t_s(\boldsymbol{u}_s); v_0\} = \exp(-v_0 t(\boldsymbol{u}))$. Hence, we obtain

$$\prod_{s=1}^{d_0} \psi_{0s}^{(d_s)}\{t_s(\boldsymbol{u}_s); v_0\} = \sum_{k=d_0}^{d} b_{\boldsymbol{d},k}^{d_0}\{\boldsymbol{t}(\boldsymbol{u})\}(-v_0)^k \exp\{-v_0 t(\boldsymbol{u})\}.$$

By replacing v_0 by V_0 and taking the expectation, one obtains

$$\begin{aligned}\mathbb{E}\Big[\prod_{s=1}^{d_0} \psi_{0s}^{(d_s)}\{t_s(\boldsymbol{u}_s); V_0\}\Big] &= \sum_{k=d_0}^{d} b_{\boldsymbol{d},k}^{d_0}\{\boldsymbol{t}(\boldsymbol{u})\}\mathbb{E}[(-V_0)^k \exp\{-V_0 t(\boldsymbol{u})\}] \\ &= \sum_{k=d_0}^{d} b_{\boldsymbol{d},k}^{d_0}\{\boldsymbol{t}(\boldsymbol{u})\}\psi_0^{(k)}\{t(\boldsymbol{u})\},\end{aligned}$$

so that, by (3.8), the density is of the form as stated.

□

Remark 3.4.

(1) We see from Theorem 3.3, Part (1) that all derivatives of $\psi_{0s}(t; v_0)$ are of similar form in v_0, namely $\psi_{0s}(t; v_0)$ times a polynomial in $-v_0$ where the coefficients $a_{s,d_s k}\{t_s(\boldsymbol{u}_s)\}$ are the Bell polynomials evaluated at the derivatives of the nodes $\mathring{\psi}_{0s}$. This structure is crucial for solving Challenge 3 since it allows one to compute the expectation with respect to F_0 explicitly.

(2) We see from Theorem 3.3 Part (2) how the (log-)density can be evaluated in general. It involves the sign-adjusted derivatives of ψ_0 which are known in many cases; see Hofert et al. (2011a). Furthermore, the quantities $b_{\boldsymbol{d},k}^{d_0}$, $k \in \{d_0, \dots, d\}$, have to be computed. The remaining parts are comparably trivial to obtain.

(3) If there are *degenerate* child copulas, that is, there exists a subset $\mathcal{S}$ of indices such that $d_s = 1$ for all $s \in \mathcal{S}$, then a straightforward application of Theorem 3.3, Part (1)

shows that

$$c(\boldsymbol{u}) = \Big[\sum_{k=d_0'}^{d-d_{\mathcal{S}}} b_{\boldsymbol{d}',k}^{d_0'} \{\boldsymbol{t}(\boldsymbol{u})\} \psi_0^{(k+d_{\mathcal{S}})}\{t(\boldsymbol{u})\} \Big] \cdot \prod_{s=1}^{d_0} \prod_{j=1}^{d_s} (\psi_s^{-1})'(u_{sj}),$$

where $d_{\mathcal{S}} = \sum_{s \in \mathcal{S}} d_s = |\mathcal{S}|$, $d_0' = d_0 - d_{\mathcal{S}}$ and $\boldsymbol{d}'$ is the vector containing all the dimensions d_s for $s \notin \mathcal{S}$.

3.3 Example - Families and transformations

3.3.1 Tilted outer power families, Clayton and Gumbel copulas

In order to construct and sample new nested Archimedean copulas it turns out to be useful to consider certain generator transformations; see Hofert (2010) for more details. One such transformation leads to *tilted outer power generators*

$$\psi_s(t) = \psi\{(c^{\theta_s} + t)^{1/\theta_s} - c\}, \tag{3.19}$$

for a generator $\psi \in \Psi_\infty$, $c \in [0, \infty)$, $\theta_s \in [1, \infty)$, and $s \in \{0, \dots, d_0\}$. Note that generators of this form are elements of Ψ_∞ as a composition of the completely monotone function ψ with a function with completely monotone derivative; see (Feller, 1971, p. 441). It follows from Equation (3.12) and Proposition 3.2 Part (1) (with $x = 1/\theta_0$ and $y = c^{\theta_0} + t$) that the derivatives of ψ_0 are

$$\psi_0^{(n)}(t) = \sum_{k=1}^{n} \psi^{(k)}\{(c^{\theta_0} + t)^{1/\theta_0} - c\}(c^{\theta_0} + t)^{k/\theta_0 - n} s_{nk}(1/\theta_0). \tag{3.20}$$

The inner generator and its derivatives

For nesting generators of Type (3.19), note that the nodes are given by

$$\mathring{\psi}_{0s}(t) = (c^{\theta_s} + t)^{\alpha_s} - c^{\theta_0},$$

where $\alpha_s = \theta_0/\theta_s$. This implies that tilted outer power generators of Type (3.19) fulfill the sufficient nesting condition if $\theta_0 \leq \theta_s$. Furthermore,

$$\mathring{\psi}_{0s}^{(k)}(t) = (\alpha_s)_k (c^{\theta_s} + t)^{\alpha_s - k}, \ k \in \mathbb{N}.$$

By Proposition 3.2 Part (1) (with $x = \alpha_s$ and $y = c^{\theta_s} + t$), this implies that

$$a_{s,nk}(t) = B_{n,k}\{\mathring{\psi}_{0s}'(t), \dots, \mathring{\psi}_{0s}^{(n-k+1)}(t)\} = (c^{\theta_s} + t)^{\alpha_s k - n} s_{nk}(\alpha_s).$$

By Theorem 3.3 Part (1), this implies that the inner generator

$$\psi_{0s}(t; v_0) = \exp[-v_0\{(c^{\theta_s} + t)^{\alpha_s} - c^{\theta_0}\}],$$

has derivatives

$$\psi_{0s}^{(n)}(t;v_0)=\psi_{0s}(t;v_0)\sum_{k=1}^{n}a_{s,nk}(t)(-v_0)^k=\psi_{0s}(t;v_0)\sum_{k=1}^{n}(c^{\theta_s}+t)^{\alpha_s k-n}s_{nk}(\alpha_s)(-v_0)^k, \tag{3.21}$$

which also reveals the structure of the coefficients $a_{s,nk}$ in (3.15), namely,

$$a_{s,nk}(t)=(c^{\theta_s}+t)^{\alpha_s k-n}s_{nk}(\alpha_s).$$

The density

Note that $(\psi_s^{-1})'(u)=\theta_s(\psi^{-1})'(u)\{c+\psi^{-1}(u)\}^{\theta_s-1}$. By Equation (3.20), Theorem 3.3 Part (2), and slight simplifications, we thus obtain

$$c(\boldsymbol{u})=\biggl(\sum_{k=d_0}^{d}b_{\boldsymbol{d},k}^{d_0}\{\boldsymbol{t}(\boldsymbol{u})\}\biggl[\sum_{j=1}^{k}\psi^{(j)}\bigl[\{c^{\theta_0}+t(\boldsymbol{u})\}^{1/\theta_0}-c\bigr]\{c^{\theta_0}+t(\boldsymbol{u})\}^{j/\theta_0-k}s_{kj}(1/\theta_0)\biggr]\biggr)$$
$$\cdot\prod_{s=1}^{d_0}\theta_s^{d_s}\prod_{j=1}^{d_s}(\psi^{-1})'(u_{sj})\{c+\psi^{-1}(u_{sj})\}^{\theta_s-1}. \tag{3.22}$$

Remark 3.5 (Clayton and Gumbel copulas)**.**

(1) By taking $\psi(t)=1/(1+t)$ and $c=1$ we see that the tilted outer power generator (3.19) is $\psi_s(t)=(1+t)^{-1/\theta_s}$, that is, a generator of the Clayton family. As a special case of this section, we thus obtain the inner generator derivatives and the densities of nested Clayton copulas. Concerning the former, we obtain from (3.21) that

$$\psi_{0s}^{(n)}(t;v_0)=\psi_{0s}(t;v_0)\sum_{k=1}^{n}s_{nk}(\alpha_s)(1+t)^{\alpha_s k-n}(-v_0)^k.$$

Concerning the latter, plugging in the corresponding quantities in (3.22) and simplifying the terms (in particular, the power of $1+\psi_0^{-1}\{C(\boldsymbol{u})\}$ can be taken out of the inner sum), we obtain

$$c(\boldsymbol{u})=\biggl[\sum_{k=d_0}^{d}(-1)^{d-k}b_{\boldsymbol{d},k}^{d_0}\{\boldsymbol{t}(\boldsymbol{u})\}\{1+t(\boldsymbol{u})\}^{-(k+1/\theta_0)}\sum_{j=1}^{k}(-1)^{k-j}s_{kj}(1/\theta_0)\biggr]$$
$$\cdot\prod_{s=1}^{d_0}\theta_s^{d_s}\biggl(\prod_{j=1}^{d_s}u_{sj}\biggr)^{-(1+\theta_s)}.$$

By Proposition 3.2 Part (3), we can further simplify this expression and obtain

$$c(\boldsymbol{u}) = (-1)^d \Big(\sum_{k=d_0}^{d} b_{\boldsymbol{d},k}^{d_0} \{\boldsymbol{t}(\boldsymbol{u})\} (-1/\theta_0)_k \{1+t(\boldsymbol{u})\}^{-(k+1/\theta_0)} \Big)$$
$$\cdot \prod_{s=1}^{d_0} \theta_s^{d_s} \Big(\prod_{j=1}^{d_s} u_{sj} \Big)^{-(1+\theta_s)}$$

for the density of nested Clayton copulas of Type (3.1). This formula also follows directly from Theorem 3.3 Part (2) by plugging in the generator derivatives $\psi_0^{(k)}(t) = (-1/\theta_0)_k(1+t)^{-(k+1/\theta_0)}$ and simplifying the expressions.

(2) Interestingly, the inner generator derivatives and the densities of nested Gumbel copulas of Type (3.1) also follow as a special case of nested tilted outer power families. To see this take $\psi(t) = \exp(-t)$ (the generator of the independence copula) and consider a zero tilt (so $c = 0$). It follows from (3.21) that

$$\psi_{0s}^{(n)}(t;v_0) = \psi_{0s}(t;v_0) \sum_{k=1}^{n} s_{nk}(\alpha_s) t^{\alpha_s k-n} (-v_0)^k.$$

Concerning the density, a short calculation shows that

$$c(\boldsymbol{u}) = (-1)^d \frac{C(\boldsymbol{u})}{\Pi(\boldsymbol{u})} \Big(\sum_{k=d_0}^{d} b_{\boldsymbol{d},k}^{d_0} \{\boldsymbol{t}(\boldsymbol{u})\} \Big[\sum_{j=1}^{k} \{-t(\boldsymbol{u})^{1/\theta_0}\}^j s_{kj}(1/\theta_0) \Big] \Big)$$
$$\cdot \prod_{s=1}^{d_0} \theta_s^{d_s} \Big(\prod_{j=1}^{d_s} -\log u_{sj} \Big)^{\theta_s - 1},$$

where C is (3.1) and Π is the independence copula (hence the product of its arguments). As before, this result can also be obtained directly from 3.3 Part (2) based on Gumbel's generator derivatives $\psi_0^{(k)}(t) = \{\psi_0(t)/t^k\} \sum_{j=1}^{k} s_{kj}(1/\theta_0)(-1/\theta_0)^j$ as derived in Hofert et al. (2011a).

3.3.2 Ali–Mikhail–Haq copulas

The inner generator and its derivatives

A nested Archimedean copula of Type (3.1) with all components C_s, $s \in \{0,\dots,d_0\}$, belonging to the Ali–Mikhail–Haq family is a valid copula according to the sufficient nesting condition if $\theta_0 \le \theta_s$ for all $s \in \{1,\dots,d_0\}$. The generator $\psi_{0s}(t;v_0)$ is given by

$$\psi_{0s}(t;v_0) = \Big\{ \frac{1-\theta_s}{(1-\theta_0)\exp(t) - (\theta_s - \theta_0)} \Big\}^{v_0} = \Big\{ \frac{1-\theta_{0s}}{\exp(t) - \theta_{0s}} \Big\}^{v_0},$$

where $\theta_{0s} = (\theta_s - \theta_0)/(1-\theta_0) \in [0,1)$ and $v_0 \in \mathbb{N}$. To compute the derivatives of this inner generator, let $f(x) = x^{-v_0}$ and $g(t) = \exp(t) - \theta_{0s}$. It follows from (3.12), Lemma 3.1, Part (2), and by simplifying the terms that

$$\begin{aligned}\psi_{0s}^{(n)}(t;v_0) &= (1-\theta_{0s})^{v_0}\sum_{l=1}^{n}(-v_0)_l\{\exp(t)-\theta_{0s}\}^{-v_0-l}B_{n,l}\{\exp(t),\dots,\exp(t)\}\\ &= (1-\theta_{0s})^{v_0}\sum_{l=1}^{n}(-v_0)_l\{\exp(t)-\theta_{0s}\}^{-v_0-l}S(n,l)\exp(tl)\\ &= \psi_{0s}(t;v_0)\sum_{l=1}^{n}S(n,l)\left(\frac{1}{1-\theta_{0s}\exp(-t)}\right)^{l}(-v_0)_l.\end{aligned}$$

It now follows from (3.14) and by interchanging the order of summations that

$$\begin{aligned}\psi_{0s}^{(n)}(t;v_0) &= \psi_{0s}(t;v_0)\sum_{l=1}^{n}S(n,l)\left\{\frac{1}{1-\theta_{0s}\exp(-t)}\right\}^{l}\sum_{k=1}^{l}s(l,k)(-v_0)^k\\ &= \psi_{0s}(t;v_0)\sum_{k=1}^{n}\left[\sum_{l=k}^{n}S(n,l)s(l,k)\left\{\frac{1}{1-\theta_{0s}\exp(-t)}\right\}^{l}\right](-v_0)^k\\ &= \psi_{0s}(t;v_0)\sum_{k=1}^{n}s_{nk}\left[1/\{1-\theta_{0s}\exp(-t)\}\right](-v_0)^k, \end{aligned} \tag{3.23}$$

which reveals the structure of the coefficients $a_{s,nk}$ in (3.15), namely,

$$a_{s,nk}(t) = s_{nk}\left[1/\{1-\theta_{0s}\exp(-t)\}\right].$$

Remark 3.6. Note that we could have used Theorem 3.3 Part (1) as well. Since $\mathring{\psi}'_{0s}(t) = \sum_{k=0}^{\infty}(\theta_{0s}\exp(-t))^k$ a short calculation shows that

$$\mathring{\psi}_{0s}^{(n)}(t) = (-1)^{n-1}\operatorname{Li}_{-(n-1)}\{\theta_{0s}\exp(-t)\}$$

for all $n\in\mathbb{N}\backslash\{1\}$ where $\operatorname{Li}_s(z)$ denotes the *polylogarithm of order s at z*. The case $n=1$ can be incorporated as follows:

$$\mathring{\psi}_{0s}^{(n)}(t) = \left\{\frac{1}{\theta_{0s}\exp(-t)}\right\}^{\mathbb{1}_{\{n=1\}}}(-1)^{n-1}\operatorname{Li}_{-(n-1)}\{\theta_{0s}\exp(-t)\}.$$

Now plugging these quantities in (3.15) and replacing $\theta_{0s}\exp(-t)$ by some $x\in[0,1)$ we can compare the coefficients with (3.23) and obtain a rather remarkable formula which relates polylogarithms and Bell polynomials with (generalized) binomial coefficients:

$$B_{n,k}\left\{\left(x^{-\mathbb{1}_{\{j=1\}}}(-1)^{j-1}\operatorname{Li}_{-(j-1)}(x)\right)_{j\in\{1,\dots,n-k+1\}}\right\} = s_{nk}\{1/(1-x)\},\quad x\in[0,1).$$

The density

Hofert et al. (2011a) showed that

$$\psi_0^{(k)}(t) = (-1)^k \frac{1-\theta_0}{\theta_0} \operatorname{Li}_{-k}\{\theta_0 \exp(-t)\},\ t \in (0,\infty),\ k \in \mathbb{N}_0.$$

It follows from Theorem 3.3 Part (2) that

$$\begin{aligned} c(\boldsymbol{u}) &= \Big[\sum_{k=d_0}^{d} b_{\boldsymbol{d},k}^{d_0}\{\boldsymbol{t}(\boldsymbol{u})\}\psi_0^{(k)}\{t(\boldsymbol{u})\}\Big] \cdot \prod_{s=1}^{d_0}\prod_{j=1}^{d_s}(\psi_s^{-1})'(u_{sj}) \\ &= (-1)^d \frac{1-\theta_0}{\theta_0}\Big(\sum_{k=d_0}^{d} b_{\boldsymbol{d},k}^{d_0}\{\boldsymbol{t}(\boldsymbol{u})\}(-1)^k \operatorname{Li}_{-k}\big[\theta_0 \exp\{-t(\boldsymbol{u})\}\big]\Big) \\ &\quad \cdot \prod_{s=1}^{d_0}(1-\theta_s)^{d_s}\Big[\prod_{j=1}^{d_s} u_{sj}\{1-\theta_s(1-u_{sj})\}\Big]^{-1}. \end{aligned}$$

3.3.3 Joe copula

The inner generator and its derivatives

Nested Joe copulas of Type (3.1) are valid copulas if $\theta_0 \le \theta_s$ for all $s \in \{1,\dots,d_0\}$. The generator $\psi_{0s}(t;v_0)$ is given by

$$\psi_{0s}(t;v_0) = \big[1-\{1-\exp(-t)\}^{\alpha_s}\big]^{v_0}, \tag{3.24}$$

where $\alpha_s = \theta_0/\theta_s$, $s \in \{1,\dots,d_0\}$, and $v_0 \in \mathbb{N}$. The nodes are given by

$$\mathring{\psi}_{0s}(t) = f\{g(t)\},$$

where $f(x) = -\log(x)$ and $g(x) = \psi_s(x) = (1-(1-\exp(-x))^{\alpha_s})$. Formula (3.12) gives

$$\mathring{\psi}_{0s}^{(n)}(t) = \sum_{k=1}^{n} f^{(k)}\{g(t)\}B_{n,k}\{g'(t),\dots,g^{(n-k+1)}(t)\}.$$

Now

$$\begin{aligned} f^{(k)}(x) &= (-1)^k x^{-k}, \\ g^{(k)}(x) &= (-1)^k \alpha_s \{1-\exp(-x)\}^{\alpha_s} P_{d,1/\alpha_s}^{\mathrm{J}}\big[\exp(-x)/\{1-\exp(-x)\}\big], \end{aligned}$$

where $P_{d,1/\alpha_s}^{\mathrm{J}}$ is a polynomial defined in Hofert et al. (2011a); see also below in the density section. Similar to the Ali–Mikhail–Haq family, Joe's family is thus an example where Formula (3.15) does not seem to be of help in computing $\psi_{0s}^{(n)}$.

Let us now consider a different approach. It follows from (3.24) that

$$\psi_{0s}^{(n)}(t;v_0) = \sum_{l=0}^{v_0} \binom{v_0}{l} (-1)^l \frac{d^n}{dt^n} \{1-\exp(-t)\}^{\alpha_s l}.$$

Together with (3.12) (take $f(x) = (1-x)^{\alpha_s l}$ and $g(x) = \exp(-x)$) and Lemma 3.1 3 we obtain

$$\begin{aligned}
\frac{d^n}{dt^n}\{1-\exp(-t)\}^{\alpha_s l} &= \sum_{m=1}^{n} (-1)^m (\alpha_s l)_m \{1-\exp(-t)\}^{\alpha_s l - m} \\
&\quad \cdot B_{n,m}\{-g(t), \ldots, (-1)^{n-m+1} g(t)\} \\
&= \sum_{m=1}^{n} (-1)^m (\alpha_s l)_m \{1-\exp(-t)\}^{\alpha_s l - m} (-1)^n g(t)^m S(n,m) \\
&= (-1)^n \{1-\exp(-t)\}^{\alpha_s l} \sum_{m=1}^{n} S(n,m)(\alpha_s l)_m \left\{-\frac{\exp(-t)}{1-\exp(-t)}\right\}^m
\end{aligned}$$

This yields

$$\begin{aligned}
\psi_{0s}^{(n)}(t;v_0) &= \sum_{l=0}^{v_0} \binom{v_0}{l} (-1)^{n+l} \{1-\exp(-t)\}^{\alpha_s l} \sum_{m=1}^{n} S(n,m)(\alpha_s l)_m \left\{-\frac{\exp(-t)}{1-\exp(-t)}\right\}^m \\
&= (-1)^n \sum_{m=1}^{n} S(n,m) \left\{-\frac{\exp(-t)}{1-\exp(-t)}\right\}^m \sum_{l=0}^{v_0} \binom{v_0}{l} (-\{1-\exp(-t)\}^{\alpha_s})^l (\alpha_s l)_m.
\end{aligned} \tag{3.25}$$

In order to obtain Representation (3.15) in terms of the polynomial in v_0 for this derivative, focus on the last sum. With $x = -\{1-\exp(-t)\}^{\alpha_s}$ and $(\alpha_s l)_m = \sum_{j=1}^{m} s(m,j)(\alpha_s l)^j$ we obtain

$$\begin{aligned}
\sum_{l=0}^{v_0} \binom{v_0}{l} [-\{1-\exp(-t)\}^{\alpha_s}]^l (\alpha_s l)_m &= \sum_{l=0}^{v_0} \binom{v_0}{l} x^l (\alpha_s l)_m = \sum_{l=0}^{v_0} \binom{v_0}{l} x^l \sum_{j=1}^{m} s(m,j)(\alpha_s l)^j \\
&= \sum_{j=1}^{m} s(m,j) \alpha_s^j \sum_{l=1}^{v_0} \binom{v_0}{l} l^j x^l,.
\end{aligned} \tag{3.26}$$

where the term with index $l = 0$ in the last formula can be omitted since $j \in \mathbb{N}$. An application of Grunert's Formula (see (Gould, 2010, Formula (2.44)) and note that $B_{k,k}^r = k!S(r,k)$ in that reference), $(v_0)_l = \sum_{k=1}^{l} s(l,k) v_0^k$, and interchanging the order of summation leads to

$$\sum_{l=1}^{v_0}\binom{v_0}{l} l^j x^l = (1+x)^{v_0}\sum_{l=1}^{j} S(j,l)(v_0)_l\Big(\frac{x}{1+x}\Big)^l$$
$$= (1+x)^{v_0}\sum_{k=1}^{j} v_0^k \sum_{l=k}^{j} s(l,k)S(j,l)\Big(\frac{x}{1+x}\Big)^l. \qquad (3.27)$$

Putting this result in (3.26) and first interchanging the order of summations in j and k and afterwards the one in j and l leads to

$$\sum_{l=0}^{v_0}\binom{v_0}{l} x^l(\alpha_s l)_m = (1+x)^{v_0}\sum_{j=1}^{m} s(m,j)\alpha_s^j \sum_{k=1}^{j} v_0^k \sum_{l=k}^{j} s(l,k)S(j,l)\Big(\frac{x}{1+x}\Big)^l$$
$$= (1+x)^{v_0}\sum_{k=1}^{m} v_0^k \sum_{j=k}^{m} s(m,j)\alpha_s^j \sum_{l=k}^{j} s(l,k)S(j,l)\Big(\frac{x}{1+x}\Big)^l$$
$$= (1+x)^{v_0}\sum_{k=1}^{m} v_0^k \sum_{l=k}^{m} s(l,k)\Big(\frac{x}{1+x}\Big)^l \sum_{j=l}^{m} s(m,j)S(j,l)\alpha_s^j$$
$$= \psi_{0s}(t;v_0)\sum_{k=1}^{m} v_0^k \sum_{l=k}^{m} s(l,k)s_{ml}(\alpha_s)\Big(\frac{x}{1+x}\Big)^l.$$

Plugging this result in (3.25) and interchanging the order of summations in m and k we obtain

$$\psi_{0s}^{(n)}(t;v_0) = \psi_{0s}(t;v_0)(-1)^n \sum_{m=1}^{n} S(n,m)\Big(-\frac{e^{-t}}{1-e^{-t}}\Big)^m \sum_{k=1}^{m} v_0^k \sum_{l=k}^{m} s(l,k)s_{ml}(\alpha_s)\Big(\frac{x}{1+x}\Big)^l$$
$$= \psi_{0s}(t;v_0)\sum_{k=1}^{n} a_{s,nk}(t)(-v_0)^k$$

for

$$a_{s,nk}(t) = \Big[(-1)^{n-k}\sum_{m=k}^{n} S(n,m)\Big\{-\frac{\exp(-t)}{1-\exp(-t)}\Big\}^m \sum_{l=k}^{m} s(l,k)s_{ml}(\alpha_s)\Big(\frac{x}{1+x}\Big)^l\Big]$$
$$= \Big[(-1)^{n-k}\sum_{m=k}^{n} S(n,m)\Big\{-\frac{\exp(-t)}{1-\exp(-t)}\Big\}^m \sum_{l=k}^{m} s(l,k)s_{ml}(\alpha_s)\Big(\frac{\psi^{\mathrm{J}}_{1/\alpha_s}(t)-1}{\psi^{\mathrm{J}}_{1/\alpha_s}(t)}\Big)^l\Big],$$

where $\psi^{\mathrm{J}}_{1/\alpha_s}(t) = 1-\{1-\exp(-t)\}^{\alpha_s}$ denotes Joe's generator with parameter $1/\alpha_s = \theta_s/\theta_0$. This is precisely the form as given in (3.15).

The density

For Joe's family, it follows from Hofert et al. (2011a) that

$$\psi_0^{(k)}(t) = (-1)^k \frac{\{1-\exp(-t)\}^{1/\theta_0}}{\theta_0} P^{\mathrm{J}}_{k,\theta_0}\left\{\frac{\exp(-t)}{1-\exp(-t)}\right\},\ t\in(0,\infty),\ n\in\mathbb{N},$$

where $P^{\mathrm{J}}_{k,\theta_0}(x) = \sum_{l=1}^{k} S(k,l)(l-1-1/\theta_0)_{l-1}x^l$. We obtain from Theorem 3.3 Part (2) that

$$\begin{aligned} c(\boldsymbol{u}) &= \left[\sum_{k=d_0}^{d} b^{d_0}_{\boldsymbol{d},k}\{\boldsymbol{t}(\boldsymbol{u})\}\psi_0^{(k)}\{t(\boldsymbol{u})\}\right]\cdot\prod_{s=1}^{d_0}\prod_{j=1}^{d_s}(\psi_s^{-1})'(u_{sj}) \\ &= \frac{(-1)^d}{\theta_0}\big[1-\exp\{-t(\boldsymbol{u})\}\big]^{1/\theta_0}\left(\sum_{k=d_0}^{d}(-1)^k b^{d_0}_{\boldsymbol{d},k}\{\boldsymbol{t}(\boldsymbol{u})\}P^{\mathrm{J}}_{k,\theta_0}\left[\frac{\exp\{t(\boldsymbol{u})\}}{1-\exp\{-t(\boldsymbol{u})\}}\right]\right) \\ &\quad\cdot\prod_{s=1}^{d_0}\theta_s^{d_s}\prod_{j=1}^{d_s}\frac{(1-u_{sj})^{\theta_s-1}}{1-(1-u_{sj})^{\theta_s}}. \end{aligned}$$

Note that $\exp\{-t(\boldsymbol{u})\} = \prod_{s=1}^{d_0}[1-\{1-C_s(\boldsymbol{u}_s)\}^{\theta_0}]$.

3.3.4 Frank Copula

The inner generator and its derivatives

Nested Frank copulas of Type (3.1) are valid copulas according to the sufficient nesting condition if $\theta_0 \le \theta_s$ for all $s\in\{1,\dots,d_0\}$. The generator $\psi_{0s}(t;v_0)$ is given by

$$\psi_{0s}(t;v_0) = \left[\frac{1-\{1-p_s\exp(-t)\}^{\alpha_s}}{p_0}\right]^{v_0},$$

where $\alpha_s = \theta_0/\theta_s$, $p_j = 1-e^{-\theta_j}$, $j\in\{0,s\}$, and $v_0\in\mathbb{N}$. Note that this inner generator is a shifted (and appropriately scaled) inner Joe generator, that is,

$$\psi_{0s}(t;v_0) = \frac{\psi^{\mathrm{J}}_{0s}(h+t;v_0)}{\psi^{\mathrm{J}}_{0s}(h;v_0)},$$

where $h = -\log p_s$; see (Hofert, 2010, p. 104) for more details about such generators. In particular, with the representation for the generator derivatives for the inner Joe generator, this implies that

$$\begin{aligned} \psi_{0s}^{(n)}(t;v_0) &= \frac{\psi^{\mathrm{J}\,(n)}_{0s}(h+t;v_0)}{\psi^{\mathrm{J}}_{0s}(h;v_0)} = \frac{\psi^{\mathrm{J}}_{0s}(h+t;v_0)}{\psi^{\mathrm{J}}_{0s}(h;v_0)}\sum_{k=1}^{n} a^{\mathrm{J}}_{s,nk}(t+h)(-v_0)^k \\ &= \psi_{0s}(t;v_0)\sum_{k=1}^{n} a^{\mathrm{J}}_{s,nk}(t+h)(-v_0)^k \end{aligned}$$

and thus that $a_{s,nk}(t) = a^{\mathrm{J}}_{s,nk}(t+h)$, that is, the coefficients of the polynomial in $-v_0$ for the derivatives of the inner Frank generator are the ones of the inner Joe generator, appropriately shifted.

The density

It follows from Hofert et al. (2011a) that

$$\psi_0^{(k)}(t) = (-1)^k \frac{1}{\theta_0} \operatorname{Li}_{-(k-1)}\{p_0 \exp(-t)\},\ t \in (0,\infty),\ k \in \mathbb{N}_0.$$

Theorem 3.3 Part (2) then implies that

$$\begin{aligned} c(\boldsymbol{u}) &= \Big[\sum_{k=d_0}^{d} b^{d_0}_{\boldsymbol{d},k}\{\boldsymbol{t}(\boldsymbol{u})\}\psi_0^{(k)}\{t(\boldsymbol{u})\}\Big] \cdot \prod_{s=1}^{d_0}\prod_{j=1}^{d_s}(\psi_s^{-1})'(u_{sj}) \\ &= (-1)^d \Bigg(\sum_{k=d_0}^{d} b^{d_0}_{\boldsymbol{d},k}\{\boldsymbol{t}(\boldsymbol{u})\}(-1)^k \operatorname{Li}_{-(k-1)}\left[p_0 \exp\{-t(\boldsymbol{u})\}\right]\Bigg) \\ &\quad \cdot \prod_{s=1}^{d_0} \theta_s^{d_s} \prod_{j=1}^{d_s} \frac{\exp(-\theta_s u_{sj})}{1-\exp(-\theta_s u_{sj})}. \end{aligned}$$

3.3.5 A nested Ali–Mikhail–Haq ∘ Clayton copula

The inner generator and its derivatives

If ψ_0 is the generator of an Ali–Mikhail–Haq copula and ψ_s, $s \in \{1,\dots,d_0\}$, generate Clayton copulas, then (Hofert, 2010, p. 115) showed that the sufficient nesting condition holds if $\theta_s \in [1,\infty)$, $s \in \{1,\dots,d_0\}$. Hence one can build nested Archimedean copulas of Type (3.1) with the root copula C_0 being of Ali–Mikhail–Haq and the child copulas C_s being of Clayton type under this condition (referred to as Ali–Mikhail–Haq ∘ Clayton copulas). In this case, a short calculation shows that

$$\psi_{0s}(t;v_0) = \psi\{(1+t)^{1/\theta_s} - 1\}$$

for $\psi(t) = \{1+(1-\theta_0)t\}^{-v_0}$. We can thus apply (3.20) with $c=1$ to see that

$$\psi_{0s}^{(n)}(t;v_0) = \sum_{j=1}^{n} \psi^{(j)}\{(1+t)^{1/\theta_s} - 1\}(1+t)^{j/\theta_s - n} s_{nj}(1/\theta_s), \tag{3.28}$$

where

$$\begin{aligned} \psi^{(j)}(t) &= (-v_0)_j\{1+(1-\theta_0)t\}^{-(v_0+j)}(1-\theta_0)^j \\ &= \psi(t)\{1+(1-\theta_0)t\}^{-j}(1-\theta_0)^j \sum_{k=1}^{j} s(j,k)(-v_0)^k \end{aligned}$$

and hence

$$\psi^{(j)}((1+t)^{1/\theta_s}-1) = (1-\theta_0)^j \psi_{0s}(t;v_0)\psi_{0s}(t;j)\sum_{k=1}^{j} s(j,k)(-v_0)^k.$$

Plugging this result into (3.28) and interchanging the order of the two summations, we obtain

$$\begin{aligned}\psi_{0s}^{(n)}(t;v_0) &= \psi_{0s}(t;v_0)\sum_{j=1}^{n} s_{nj}(1/\theta_s)\psi_{0s}(t;j)(1-\theta_0)^j(1+t)^{j/\theta_s-n}\sum_{k=1}^{j} s(j,k)(-v_0)^k \\ &= \psi_{0s}(t;v_0)\sum_{k=1}^{n}\Big\{\sum_{j=k}^{n} s(j,k)s_{nj}(1/\theta_s)\psi_{0s}(t;j)(1-\theta_0)^j(1+t)^{j/\theta_s-n}\Big\}(-v_0)^k,\end{aligned}$$

which provides the structure of the coefficients $a_{s,nk}$ in (3.15), namely,

$$a_{s,nk}(t) = \sum_{j=k}^{n} s(j,k)s_{nj}(1/\theta_s)\Big\{\frac{1-\theta_0}{\theta_0+(1-\theta_0)(1+t)^{1/\theta_s}}\Big\}^j (1+t)^{j/\theta_s-n}.$$

The density

It is clear from Theorem 3.3 Part (2) that the density for the nested Ali–Mikhail–Haq $\circ$ Clayton copula basically consists of the corresponding pieces of the Ali–Mikhail–Haq and the Clayton density we have already seen earlier. It is given by

$$\begin{aligned}c(\boldsymbol{u}) &= \Big[\sum_{k=d_0}^{d} b_{\boldsymbol{d},k}^{d_0}\{\boldsymbol{t}(\boldsymbol{u})\}\psi_0^{(k)}\{t(\boldsymbol{u})\}\Big]\cdot\prod_{s=1}^{d_0}\prod_{j=1}^{d_s}(\psi_s^{-1})'(u_{sj}) \\ &= (-1)^d\frac{1-\theta_0}{\theta_0}\Big(\sum_{k=d_0}^{d} b_{\boldsymbol{d},k}^{d_0}\{\boldsymbol{t}(\boldsymbol{u})\}(-1)^k \operatorname{Li}_{-k}\big[\theta_0\exp\{-t(\boldsymbol{u})\}\big]\Big)\prod_{s=1}^{d_0}\theta_s^{d_s}\Big(\prod_{j=1}^{d_s}u_{sj}\Big)^{-(1+\theta_s)}.\end{aligned}$$

Note that $\exp\{t(\boldsymbol{u})\} = \prod_{s=1}^{d_0}\dfrac{C_s(\boldsymbol{u}_s)}{1-\theta_0\{1-C_s(\boldsymbol{u}_s)\}}$.

3.4 Numerical evaluation

3.4.1 The log-density

In statistical applications one typically aims at computing the log-density. From a numerical point of view, this is typically not as trivial as computing the density and taking the logarithm afterwards. Often, the density cannot be computed without running into numerical problems, hence taking the logarithm of the density faces the same problems. However, an intelligent implementation of the log-density is possible (and often even required), see the implementation in the R package `nacopula`.

We now briefly explain how one can efficiently compute the log-density of a nested

Archimedean copula of Type (3.1). Recall from (3.17) that

$$c(\boldsymbol{u}) = \Big[\sum_{k=d_0}^{d} b_{\boldsymbol{d},k}^{d_0}\{\boldsymbol{t}(\boldsymbol{u})\}\psi_0^{(k)}\{t(\boldsymbol{u})\}\Big]\cdot\prod_{s=1}^{d_0}\prod_{j=1}^{d_s}(\psi_s^{-1})'(u_{sj}).$$

Let us first think about the signs of the terms $b_{\boldsymbol{d},k}^{d_0}\{\boldsymbol{t}(\boldsymbol{u})\}$, $k\in\{d_0,\dots,d\}$. By Theorem 3.3, Part (1) we know that $\operatorname{sign}\big(a_{s,d_s j_s}\{t_s(\boldsymbol{u}_s)\}\big)=(-1)^{d_s-j_s}$, thus

$$\operatorname{sign}\prod_{s=1}^{d_0} a_{s,d_s j_s}\{t_s(\boldsymbol{u}_s)\} = (-1)^{\sum_{s=1}^{d_0} d_s - \sum_{s=1}^{d_0} j_s}.$$

Recall from (3.18) the structure of $b_{\boldsymbol{d},k}^{d_0}\{\boldsymbol{t}(\boldsymbol{u})\}$, which is the sum in $\boldsymbol{j}\in\mathcal{Q}_{\boldsymbol{d},k}^{d_0}$ over $\prod_{s=1}^{d_0} a_{s,d_s j_s}\{t_s(\boldsymbol{u}_s)\}$. For such $\boldsymbol{j}$, it follows from the definition of $\mathcal{Q}_{\boldsymbol{d},k}^{d_0}$ that $\sum_{s=1}^{d_0} j_s = k$. Furthermore, note that $\sum_{s=1}^{d_0} d_s = d$, hence

$$\operatorname{sign} b_{\boldsymbol{d},k}^{d_0}\{\boldsymbol{t}(\boldsymbol{u})\} = (-1)^{d-k}.$$

This implies that

$$c(\boldsymbol{u}) = \Big[\sum_{k=d_0}^{d} (-1)^{d-k} b_{\boldsymbol{d},k}^{d_0}\{\boldsymbol{t}(\boldsymbol{u})\}(-1)^k\psi_0^{(k)}\{t(\boldsymbol{u})\}\Big]\cdot\prod_{s=1}^{d_0}\prod_{j=1}^{d_s}(-\psi_s^{-1})'(u_{sj}) \tag{3.29}$$

where we note that $\prod_{s=1}^{d_0}\prod_{k=1}^{d_s}(-1)=(-1)^d$.

We see from (3.29) that all appearing quantities are positive which is quite convenient for computing the log-density

$$\log c(\boldsymbol{u}) = \log\Big[\sum_{k=d_0}^{d} (-1)^{d-k} b_{\boldsymbol{d},k}^{d_0}\{\boldsymbol{t}(\boldsymbol{u})\}(-1)^k\psi_0^{(k)}\{t(\boldsymbol{u})\}\Big] + \sum_{s=1}^{d_0}\sum_{j=1}^{d_s}\log(-\psi_s^{-1})'(u_{sj}).$$

Since the latter double sum is typically trivial to compute, let us focus on the first sum. To compute the (intelligent) logarithm of this sum, let

$$x_k = \log\big[(-1)^{d-k} b_{\boldsymbol{d},k}^{d_0}\{\boldsymbol{t}(\boldsymbol{u})\}\big] + \log\big[(-1)^k\psi_0^{(k)}\{t(\boldsymbol{u})\}\big],\ k\in\{d_0,\dots,d\},$$

and note that

$$\begin{aligned}\log\sum_{k=d_0}^{d} (-1)^{d-k} b_{\boldsymbol{d},k}^{d_0}\{\boldsymbol{t}(\boldsymbol{u})\}(-1)^k\psi_0^{(k)}\{t(\boldsymbol{u})\} &= \log\sum_{k=d_0}^{d}\exp(x_k)\\ &= x_{\max} + \log\sum_{k=d_0}^{d}\exp(x_k - x_{\max}),\end{aligned}$$

where $x_{\max} = \max_{d_0\le k\le d} x_k$. Since all summands in the latter sum are in (0,1], the corresponding logarithm can easily be computed. It remains to discuss how the x_k,

$k \in \{d_0, \dots, d\}$, can be computed.

For computing the coefficients x_k, $k \in \{d_0, \dots, d\}$, efficient implementations for the functions $\log\{(-1)^k \psi_0^{(k)}(t)\}$ in the R package `nacopula` can be used. Computing the quantities $\log\big[(-1)^{d-k} b_{\boldsymbol{d},k}^{d_0}\{\boldsymbol{t}(\boldsymbol{u})\}\big]$ is more challenging. Recall from (3.18) that

$$b_{\boldsymbol{d},k}^{d_0}\{\boldsymbol{t}(\boldsymbol{u})\} = \sum_{\boldsymbol{j} \in \mathcal{Q}_{\boldsymbol{d},k}^{d_0}} \prod_{s=1}^{d_0} a_{s,d_s j_s}\{t_s(\boldsymbol{u}_s)\},$$

where $a_{s,d_s j_s}\{t_s(\boldsymbol{u}_s)\}$ is given in (3.16). For computing the function $s_{d_s j_s}$ that often appears in $a_{s,d_s j_s}\{t_s(\boldsymbol{u}_s)\}$, the function `coeffG` in `nacopula` can be used; to be more precise, `(-1)^(`d_s`-`j_s`) * nacopula:::coeffG(`d_s`,x)` computes $s_{d_s j_s}(\mathtt{x})$. For the summation over the set $\mathcal{Q}_{\boldsymbol{d},k}^{d_0}$, the R package `partitions` provides the function `blockparts`. With the instruction `blockparts(`$\boldsymbol{d}$`-rep(1L,`d_0`),`k`-`d_0`)+1L` one can then obtain a matrix with d_0 rows where each column gives one $\boldsymbol{j} \in \mathcal{Q}_{\boldsymbol{d},k}^{d_0}$.

3.4.2 The -log-likelihood of two-parameter nested Clayton and Gumbel copulas

In this section, we compute the -log-likelihood (based on a sample of size $n = 100$) of two nested Clayton and two nested Gumbel copulas with parameters θ_0 and θ_1 such that Kendall's tau equals 0.25 and 0.5, respectively. In order to be able to provide graphical insights, we focus on two-parameter nested copulas of the form

$$C(\boldsymbol{u}) = C_0\{u_1, C_1(u_2, \dots, u_d)\} \tag{3.30}$$

where $d \in \{3, 10\}$.

Note that we obtain nested Archimedean copulas of Type (3.30) from (3.1) by artificially thinking of u_1 as a child copula $\psi_0\{\psi_0^{-1}(u_1)\}$ of dimension 1, that is, as a degenerate child copula. For such s, note that $a_{s,d_s j_s}\{t_s(\boldsymbol{u}_s)\} = a_{s,11}\{t_s(\boldsymbol{u}_s)\}$ and $t_s(\boldsymbol{u}_s)$ equals ψ_0^{-1} at the corresponding (one-dimensional) argument, which is u_1 in (3.30). It follows from the last statement in Theorem 3.3, Part (1) that $a_{s,d_s j_s}\{t_s(\boldsymbol{u}_s)\} = 1$ for degenerate children. This implies that these terms drop out of the product in (3.18). The set $\mathcal{Q}_{\boldsymbol{d},k}^{d_0}$ shrinks accordingly since $1 \le j_s \le d_s = 1$ for degenerate children s.

Figures 3.1 and 3.2 display the -log-likelihoods (as wireframe and level plots) for the nested Clayton copulas as described above based on a sample of size $n = 100$ (note the restriction $\theta_0 \le \theta_1$). The parameters from which the samples were drawn and the minima based on the grid points as displayed in the wireframe plots are included. It is interesting to see the behavior of the -log-likelihood in the child parameter θ_1 when the dimension of the child copula C_1 is increased (with C_0 and its dimension fixed). As can be seen from Figures 3.1 and 3.2, the -log-likelihood is easier to minimize in θ_1-direction. This behavior was already observed by Hofert et al. (2011a) for Archimedean copulas and can be expressed by the empirical observation that the mean squared error behaves like $1/(nd)$ which is decreasing in d for fixed n. The same is visible for the -log-likelihoods of the nested Clayton copulas here which is not a surprise since the marginal copula for $u_1 = 1$ is the Archimedean copula C_1. The same behavior can be seen from Figures 3.3 and 3.4 which show the -log-likelihoods (as wireframe and level plots) for the nested Gumbel copulas.

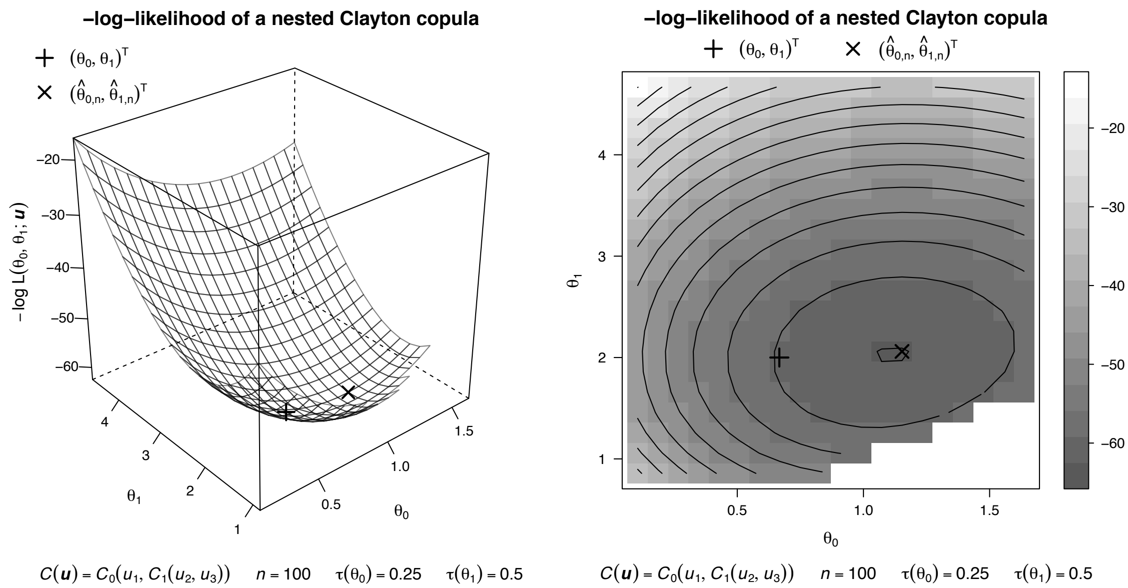

Figure 3.1: Wireframe (left) and level plot (right) of the -log-likelihood of a three-dimensional nested Clayton copula $C(\boldsymbol{u}) = C_0\{u_1, C_1(u_2, u_3)\}$ with parameters $\theta_0 = 2/3$ (Kendall's tau equals 0.25) and $\theta_1 = 2$ (Kendall's tau equals 0.5) based on a sample of size $n = 100$.

3.5 Densities for three- (and higher-) level nested Archimedean copulas

In this section, a density formula analogous to (3.17) is derived for three-level nested Archimedean copulas and extensions to higher nesting levels are briefly addressed.

When working with three or more nesting levels, it turns out to be convenient to (slightly) change the notation used in the previous sections. Consider a three-level nested Archimedean copula of the form

$$C(\boldsymbol{u}) = C_1\big[C_{11}\{C_{111}(\boldsymbol{u}_{111}), \dots, C_{11d_{11}}(\boldsymbol{u}_{11d_{11}})\}, \dots, C_{1d_1}\{C_{1d_11}(\boldsymbol{u}_{1d_11}), \dots, C_{1d_1d_{1d_1}}(\boldsymbol{u}_{1d_1d_{1d_1}})\}\big], \tag{3.31}$$

where $\boldsymbol{u}_{s_1s_2s_3} = (u_{s_1s_2s_31}, \dots, u_{s_1s_2s_3d_{s_1s_2s_3}})^\top$ denotes the argument of $C_{s_1s_2s_3}$ (the copula generated by $\psi_{s_1s_2s_3}$), $d_{s_1s_2s_3}$ denotes the dimension of $C_{s_1s_2s_3}$, and $d_{s_1s_2}$ denotes the dimension of $C_{s_1s_2}$ (the copula generated by $\psi_{s_1s_2}$). Here and in the following, s_1 always equals 1, $s_2 \in \{1, \dots, d_{s_1}\}$, and $s_3 \in \{1, \dots, d_{s_1s_2}\}$. Note that it is convenient to think of

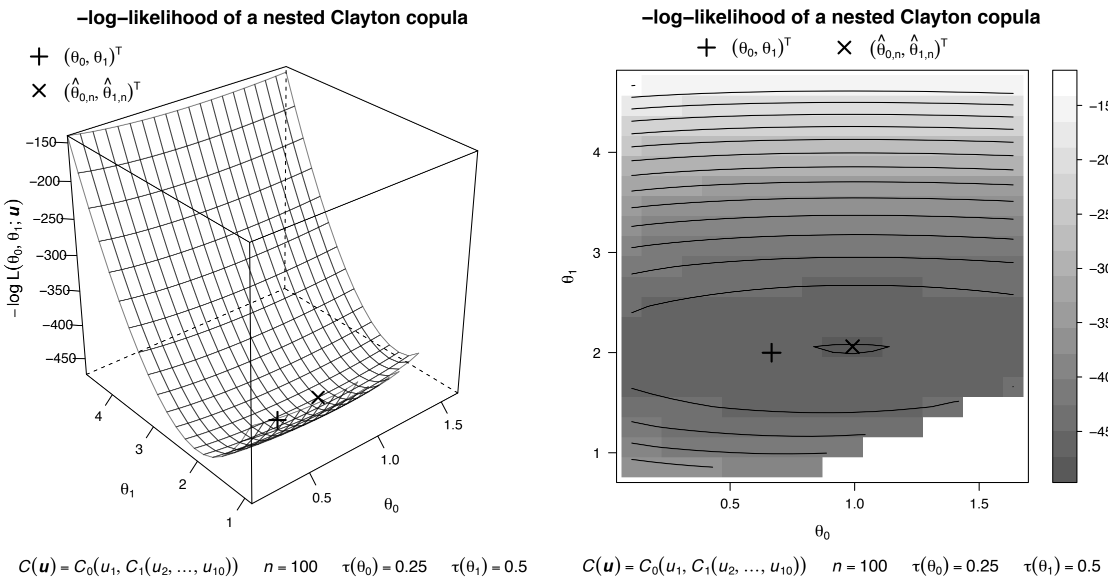

Figure 3.2: Wireframe (left) and level plot (right) of the -log-likelihood of a ten-dimensional nested Clayton copula $C(\boldsymbol{u}) = C_0\{u_1, C_1(u_2, \ldots, u_{10})\}$ with parameters $\theta_0 = 2/3$ (Kendall's tau equals 0.25) and $\theta_1 = 2$ (Kendall's tau equals 0.5) based on a sample of size $n = 100$.

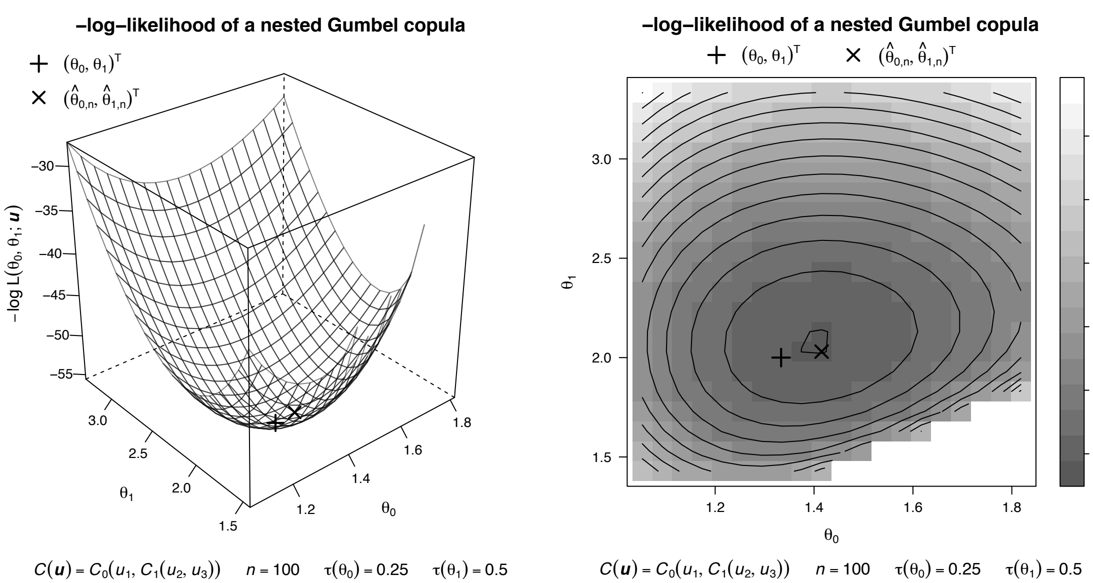

Figure 3.3: Wireframe (left) and level plot (right) of the -log-likelihood of a three-dimensional nested Gumbel copula $C(\boldsymbol{u}) = C_0\{u_1, C_1(u_2, u_3)\}$ with parameters $\theta_0 = 4/3$ (Kendall's tau equals 0.25) and $\theta_1 = 2$ (Kendall's tau equals 0.5) based on a sample of size $n = 100$.

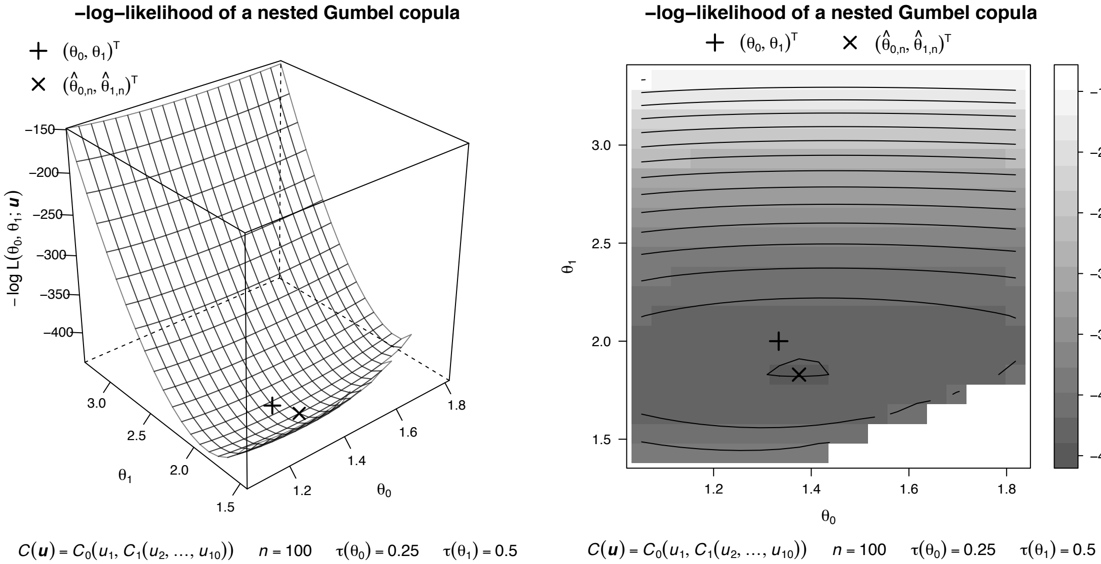

Figure 3.4: Wireframe (left) and level plot (right) of the -log-likelihood of a ten-dimensional nested Gumbel copula $C(\boldsymbol{u}) = C_0\{u_1, C_1(u_2, \ldots, u_{10})\}$ with parameters $\theta_0 = 4/3$ (Kendall's tau equals 0.25) and $\theta_1 = 2$ (Kendall's tau equals 0.5) based on a sample of size $n = 100$.

(3.31) as a tree; see Figure 3.5. Furthermore, let

$$
\begin{aligned}
t_{s_1s_2s_3}(\boldsymbol{u}_{s_1s_2s_3}) &= \sum_{s_4=1}^{d_{s_1s_2s_3}} \psi^{-1}_{s_1s_2s_3}(u_{s_1s_2s_3s_4}) = \psi^{-1}_{s_1s_2s_3}\{C_{s_1s_2s_3}(\boldsymbol{u}_{s_1s_2s_3})\},\\
\boldsymbol{u}_{s_1s_2} &= (\boldsymbol{u}^\mathsf{T}_{s_1s_21}, \ldots, \boldsymbol{u}^\mathsf{T}_{s_1s_2d_{s_1s_2}})^\mathsf{T},\\
\boldsymbol{t}_{s_1s_2}(\boldsymbol{u}_{s_1s_2}) &= (t_{s_1s_21}(\boldsymbol{u}_{s_1s_21}), \ldots, t_{s_1s_2d_{s_1s_2}}(\boldsymbol{u}_{s_1s_2d_{s_1s_2}}))^\mathsf{T},\\
C^*_{s_1s_2}(\boldsymbol{u}_{s_1s_2}) &= C_{s_1s_2}\{C_{s_1s_21}(\boldsymbol{u}_{s_1s_21}), \ldots, C_{s_1s_2d_{s_1s_2}}(\boldsymbol{u}_{s_1s_2d_{s_1s_2}})\}\\
&= \psi_{s_1s_2}\Bigg[\sum_{s_3=1}^{d_{s_1s_2}} \mathring{\psi}_{s_1s_2,s_1s_2s_3}\{t_{s_1s_2s_3}(\boldsymbol{u}_{s_1s_2s_3})\}\Bigg],\\
t^*_{s_1s_2}(\boldsymbol{u}_{s_1s_2}) &= \psi^{-1}_{s_1s_2}\{C^*_{s_1s_2}(\boldsymbol{u}_{s_1s_2})\} = \sum_{s_3=1}^{d_{s_1s_2}} \mathring{\psi}_{s_1s_2,s_1s_2s_3}\{t_{s_1s_2s_3}(\boldsymbol{u}_{s_1s_2s_3})\},
\end{aligned}
$$

where

$$\mathring{\psi}_{s_1s_2,s_1s_2s_3} = \psi^{-1}_{s_1s_2} \circ \psi_{s_1s_2s_3}$$

and $C^*_{s_1s_2}$ denotes the (marginal) nested Archimedean copula with root $C_{s_1s_2}$. Note that the dimension of the root copula $C_{s_1s_2}$ of the nested Archimedean copula $C^*_{s_1s_2}$ is $d_{s_1s_2}$ which is in general not equal to the dimension $d_{s_1s_2\cdot} = \sum_{s_3=1}^{d_{s_1s_2}} d_{s_1s_2s_3}$ of $C^*_{s_1s_2}$. Furthermore, the root copula C_1 $(= C_{s_1})$ of the nested Archimedean copula C has d_1 $(= d_{s_1})$ arguments, the s_2th of which has $d_{s_1 s_2}$-many arguments. Overall, $d_{s_1\cdot\cdot} = \sum_{s_2=1}^{d_{s_1}} \sum_{s_3=1}^{d_{s_1 s_2}} d_{s_1 s_2 s_3}$ equals d, the dimension of C.

In order to compute the density c of C, we use a similar idea as in Section 3.2.1.

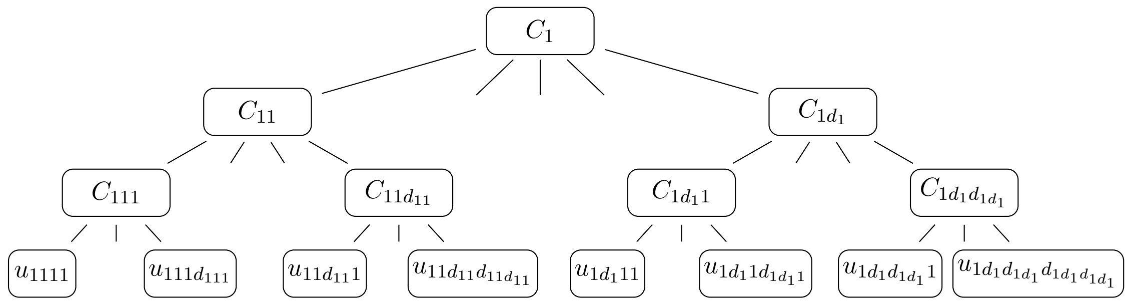

Figure 3.5: Tree structure (some arguments are omitted) for a three-level nested Archimedean copula of Type (3.31).

By replacing $\psi_1(=\psi_{s_1}) = \mathcal{LS}[F_1]$ with the corresponding integral, we obtain

$$\begin{aligned} C(\boldsymbol{u}) &= \int_0^\infty \prod_{s_2=1}^{d_1} \exp\big[-v_1 \psi_{s_1}^{-1}\{C^*_{s_1 s_2}(\boldsymbol{u}_{s_1 s_2})\}\big]\, dF_1(v_1) \\ &= \int_0^\infty \prod_{s_2=1}^{d_1} \psi_{s_1, s_1 s_2}\{t^*_{s_1 s_2}(\boldsymbol{u}_{s_1 s_2}); v_1\}\, dF_1(v_1) \end{aligned}$$

where

$$\psi_{s_1, s_1 s_2}(t; v_1) = \exp\{-v_1 \mathring{\psi}_{s_1, s_1 s_2}(t)\}$$

and thus

$$c(\boldsymbol{u}) = \int_0^\infty \prod_{s_2=1}^{d_1} \frac{\partial}{\partial \boldsymbol{u}_{s_1 s_2}} \psi_{s_1, s_1 s_2}\{t^*_{s_1 s_2}(\boldsymbol{u}_{s_1 s_2}); v_1\}\, dF_1(v_1), \tag{3.32}$$

where $\frac{\partial}{\partial \boldsymbol{u}_{s_1 s_2}}$ denotes the derivative with respect to all components of $\boldsymbol{u}_{s_1 s_2}$ (which is a vector of length $d_{s_1 s_2 \cdot}$).

Similar as in Section 3.2.1, we observe the following key challenges:

Challenge 1 Find the derivatives in the integrand;

Challenge 2 Compute their product;

Challenge 3 Integrate it with respect to the mixture distribution function $F_1 = \mathcal{LS}^{-1}[\psi_1]$.

We will first solve Challenge 1 by considering a multivariate version of Faà di Bruno's formula. For suitable functions $f : \mathbb{R} \to \mathbb{R}$ and $g : \mathbb{R}^n \to \mathbb{R}$, it follows from Hardy (2006) that

$$\frac{\partial}{\partial \boldsymbol{x}} f\{g(\boldsymbol{x})\} = \sum_{k=1}^{n} f^{(k)}\{g(\boldsymbol{x})\} \sum_{\pi:|\pi|=k} \prod_{B\in\pi} \frac{\partial^{|B|}}{\prod_{i\in B}\partial x_i} g(\boldsymbol{x}),\ \boldsymbol{x} = (x_1,\dots,x_n)^\mathsf{T}, \tag{3.33}$$

where the last sum extends over all partitions π of $\{1,\dots,n\}$ with k elements and the last product over all blocks B of π. Observe that, if $x_1 = \cdots = x_n = x$, then the univariate Faà di Bruno's formula (3.12) can be restated as

$$(f\circ g)^{(n)}(x) = \sum_{k=1}^{n} f^{(k)}\{g(x)\} \sum_{\pi:|\pi|=k} \prod_{B\in\pi} g^{(|B|)}(x),$$

where π is a partition of $\{1,\dots,n\}$. Comparing this identity with (3.12) yields

$$B_{n,k}\{g'(x),\dots,g^{(n-k+1)}(x)\} = \sum_{\pi:|\pi|=k} \prod_{B\in\pi} g^{(|B|)}(x). \tag{3.34}$$

This will be used in the following lemma, which is a special case of (3.33) with stronger assumptions on the function g. It will then lead us to a solution for Challenge 1 by choosing suitable functions f and g.

Lemma 3.7. Suppose there exists a partition $\{B_1,\dots,B_m\}$ of $\{1,\dots,n\}$ with $|B_l| = d_l$ for $l\in\{1,\dots,m\}$ (with $\sum_{l=1}^{m} d_l = n$), such that for any indices $k_1\in B_i$ and $k_2\in B_j$, for $i,j\in\{1,\dots,m\}$ with $i\neq j$, the partial derivative of $g(x_1,\dots,x_n)$ with respect to x_{k_1} and x_{k_2} equals zero, that is

$$\frac{\partial^2}{\partial x_{k_1}\partial x_{k_2}} g(\boldsymbol{x}) = 0,\ \forall k_1\in B_i,\ \forall k_2\in B_j,\ \forall i,j\in\{1,\dots,m\},\ i\neq j. \tag{3.35}$$

Moreover, suppose that for any $l\in\{1,\dots,m\}$ and any subset B of B_l, there exist functions h_{l1} and h_{l2} such that

$$\frac{\partial^{|B|}}{\prod_{i\in B}\partial x_i} g(\boldsymbol{x}) = h_{l1}^{(|B|)}\{h_{l2}(\boldsymbol{x})\} \prod_{i\in B} \frac{\partial}{\partial x_i} h_{l2}(\boldsymbol{x}). \tag{3.36}$$

Then one has, for any suitable function f,

$$\begin{aligned}\frac{\partial}{\partial \boldsymbol{x}} f\{g(\boldsymbol{x})\} = &\left\{\prod_{l=1}^{m}\prod_{i\in B_l}\frac{\partial}{\partial x_i} h_{l2}(\boldsymbol{x})\right\}\\ &\cdot \sum_{k=1}^{n} f^{(k)}\{g(\boldsymbol{x})\} \sum_{\boldsymbol{j}\in\mathcal{Q}^m_{\boldsymbol{d},k}} \prod_{l=1}^{m} B_{d_l,j_l}[h'_{l1}\{h_{l2}(\boldsymbol{x})\},\dots,h_{l1}^{(d_l-j_l+1)}\{h_{l2}(\boldsymbol{x})\}],\end{aligned}$$

where $\mathcal{Q}^m_{\boldsymbol{d},k}$ is defined as in Theorem 3.3, Part (2) and $\boldsymbol{d} = (d_1,\dots,d_m)^\mathsf{T}$.

Proof

Due to Equation (3.35), observe that the second sum in (3.33) may be rewritten as

$$\sum_{\pi:|\pi|=k}\prod_{B\in\pi}\frac{\partial^{|B|}}{\prod_{i\in B}\partial x_i}g(\boldsymbol{x})=\sum_{\boldsymbol{j}\in\mathcal{Q}^m_{\boldsymbol{d},k}}\prod_{l=1}^m\Big\{\sum_{\pi_l:|\pi_l|=j_l}\prod_{B\in\pi_l}\frac{\partial^{|B|}}{\prod_{i\in B}\partial x_i}g(\boldsymbol{x})\Big\},$$

where π_l is a partition of B_l with j_l elements, $l\in\{1,\dots,m\}$; note that $j_l\le d_l$, because $|B_l|=d_l$. Both sides are equal because the remaining terms in the sum on the left-hand side are zero as derivatives with respect to two or more variables belonging to different partitions vanish. Combining this result with (3.36), it follows from (3.33) that

$$\frac{\partial}{\partial\boldsymbol{x}}f\{g(\boldsymbol{x})\}=\sum_{k=1}^n f^{(k)}\{g(\boldsymbol{x})\}\sum_{\boldsymbol{j}\in\mathcal{Q}^m_{\boldsymbol{d},k}}\prod_{l=1}^m\Big\{\sum_{\pi_l:|\pi_l|=j_l}\Big(\prod_{B\in\pi_l}\Big[h_{l1}^{(|B|)}\{h_{l2}(\boldsymbol{x})\}\prod_{i\in B}\frac{\partial}{\partial x_i}h_{2l}(\boldsymbol{x})\Big]\Big)\Big\}.$$

Observe that for any partition π_l of B_l, $l\in\{1,\dots,m\}$, we have that

$$\prod_{B\in\pi_l}\prod_{i\in B}\frac{\partial}{\partial x_i}h_{l2}(\boldsymbol{x})=\prod_{i\in B_l}\frac{\partial}{\partial x_i}h_{l2}(\boldsymbol{x}),$$

so that

$$\begin{aligned}\frac{\partial}{\partial\boldsymbol{x}}f\{g(\boldsymbol{x})\}&=\sum_{k=1}^n f^{(k)}\{g(\boldsymbol{x})\}\sum_{\boldsymbol{j}\in\mathcal{Q}^m_{\boldsymbol{d},k}}\prod_{l=1}^m\Big[\Big\{\prod_{i\in B_l}\frac{\partial}{\partial x_i}h_{l2}(\boldsymbol{x})\Big\}\sum_{\pi_l:|\pi_l|=j_l}\prod_{B\in\pi_l}h_{l1}^{(|B|)}\{h_{l2}(\boldsymbol{x})\}\Big]\\&=\Big\{\prod_{l=1}^m\prod_{i\in B_l}\frac{\partial}{\partial x_i}h_{l2}(\boldsymbol{x})\Big\}\sum_{k=1}^n f^{(k)}\{g(\boldsymbol{x})\}\sum_{\boldsymbol{j}\in\mathcal{Q}^m_{\boldsymbol{d},k}}\prod_{l=1}^m\sum_{\pi_l:|\pi_l|=j_l}\prod_{B\in\pi_l}h_{l1}^{(|B|)}\{h_{l2}(\boldsymbol{x})\}.\end{aligned}$$

Finally, applying Identity (3.34) completes the proof. □

Remark 3.8. For the two nesting levels case, it is interesting to observe that although it is less intuitive, Lemma 3.7 could have been applied directly to suitable functions in order to obtain Equation (3.17).

We are now in the position to solve Challenge 1. By applying Lemma 3.7 with $\boldsymbol{x}=\boldsymbol{u}_{s_1s_2}$, $n=d_{s_1s_2\cdot}$, $m=d_{s_1s_2}$, $B_l=\{s_1s_2l1,\dots,s_1s_2ld_{s_1s_2l}\}$ (slightly abusing the notation), $\boldsymbol{d}_{s_1s_2}=(d_{s_1s_21},\dots,d_{s_1s_2d_{s_1s_2}})^\top$, $f(t)=\psi_{s_1,s_1s_2}(t;v_1)$, $g(\boldsymbol{x})=t^*_{s_1s_2}(\boldsymbol{u}_{s_1s_2})$, $h_{l1}(t)=\mathring{\psi}_{s_1s_2,s_1s_2l}(t)$, and $h_{l2}(\boldsymbol{u})=t_{s_1s_2l}(\boldsymbol{u}_{s_1s_2l})=\sum_{s_4=1}^{d_{s_1s_2l}}\psi^{-1}_{s_1s_2l}(u_{s_1s_2ls_4})$, the derivative in the integrand in (3.32) is given by

$$\frac{\partial}{\partial \boldsymbol{u}_{s_1s_2}}\psi_{s_1,s_1s_2}\{t^*_{s_1s_2}(\boldsymbol{u}_{s_1s_2});v_1\}$$
$$=\left\{\prod_{s_3=1}^{d_{s_1s_2}}\prod_{s_4=1}^{d_{s_1s_2s_3}}(\psi^{-1}_{s_1s_2s_3})'(u_{s_1s_2s_3s_4})\right\}\sum_{l=1}^{d_{s_1s_2\cdot}}\psi^{(l)}_{s_1,s_1s_2}\{t^*_{s_1s_2}(\boldsymbol{u}_{s_1s_2});v_1\}$$
$$\cdot\left(\sum_{\boldsymbol{j}\in\mathcal{Q}^{d_{s_1s_2}}_{\boldsymbol{d}_{s_1s_2},l}}\prod_{s_3=1}^{d_{s_1s_2}}B_{d_{s_1s_2s_3},j_{s_3}}\left[(\mathring{\psi}^{(k)}_{s_1s_2,s_1s_2s_3}\{t_{s_1s_2s_3}(\boldsymbol{u}_{s_1s_2s_3})\})_{k\in\{1,\ldots,d_{s_1s_2s_3}-j_{s_3}+1\}}\right]\right).$$

By applying Theorem 3.3, Part (1), this derivative can be written as

$$\frac{\partial}{\partial \boldsymbol{u}_{s_1s_2}}\psi_{s_1,s_1s_2}\{t^*_{s_1s_2}(\boldsymbol{u}_{s_1s_2});v_1\}$$
$$=\left\{\prod_{s_3=1}^{d_{s_1s_2}}\prod_{s_4=1}^{d_{s_1s_2s_3}}(\psi^{-1}_{s_1s_2s_3})'(u_{s_1s_2s_3s_4})\right\}\psi_{s_1,s_2s_2}\{t^*_{s_1s_2}(\boldsymbol{u}_{s_1s_2});v_1\}$$
$$\cdot\sum_{l=1}^{d_{s_1s_2\cdot}}\sum_{k=1}^{l}a_{s_1s_2,lk}\{t^*_{s_1s_2}(\boldsymbol{u}_{s_1s_2})\}(-v_1)^k$$
$$\cdot\left(\sum_{\boldsymbol{j}\in\mathcal{Q}^{d_{s_1s_2}}_{\boldsymbol{d}_{s_1s_2},l}}\prod_{s_3=1}^{d_{s_1s_2}}B_{d_{s_1s_2s_3},j_{s_3}}\left[(\mathring{\psi}^{(k)}_{s_1s_2,s_1s_2s_3}\{t_{s_1s_2s_3}(\boldsymbol{u}_{s_1s_2s_3})\})_{k\in\{1,\ldots,d_{s_1s_2s_3}-j_{s_3}+1\}}\right]\right).$$

Interchanging the order of summations yields

$$\frac{\partial}{\partial \boldsymbol{u}_{s_1s_2}}\psi_{s_1,s_1s_2}\{t^*_{s_1s_2}(\boldsymbol{u}_{s_1s_2});v_1\}$$
$$=\left\{\prod_{s_3=1}^{d_{s_1s_2}}\prod_{s_4=1}^{d_{s_1s_2s_3}}(\psi^{-1}_{s_1s_2s_3})'(u_{s_1s_2s_3s_4})\right\}\psi_{s_1,s_2s_2}\{t^*_{s_1s_2}(\boldsymbol{u}_{s_1s_2});v_1\}$$
$$\cdot\sum_{k=1}^{d_{s_1s_2\cdot}}(-v_1)^k\sum_{l=k}^{d_{s_1s_2\cdot}}\left\{a_{s_1s_2,lk}\{t^*_{s_1s_2}(\boldsymbol{u}_{s_1s_2})\}\right.$$
$$\left.\cdot\left(\sum_{\boldsymbol{j}\in\mathcal{Q}^{d_{s_1s_2}}_{\boldsymbol{d}_{s_1s_2},l}}\prod_{s_3=1}^{d_{s_1s_2}}B_{d_{s_1s_2s_3},j_{s_3}}\left[(\mathring{\psi}^{(k)}_{s_1s_2,s_1s_2s_3}\{t_{s_1s_2s_3}(\boldsymbol{u}_{s_1s_2s_3})\})_{k\in\{1,\ldots,d_{s_1s_2s_3}-j_{s_3}+1\}}\right]\right)\right\}$$

which can be written as

$$
\begin{aligned}
&\frac{\partial}{\partial \boldsymbol{u}_{s_1 s_2}} \psi_{s_1, s_1 s_2}\{t^*_{s_1 s_2}(\boldsymbol{u}_{s_1 s_2}); v_1\} \\
&= \left\{ \prod_{s_3=1}^{d_{s_1 s_2}} \prod_{s_4=1}^{d_{s_1 s_2 s_3}} (\psi^{-1}_{s_1 s_2 s_3})'(u_{s_1 s_2 s_3 s_4}) \right\} \psi_{s_1, s_2 s_2}\{t^*_{s_1 s_2}(\boldsymbol{u}_{s_1 s_2}); v_1\} \\
&\quad \cdot \sum_{k=1}^{d_{s_1 s_2 \cdot}} a_{s_1 s_2, d_{s_1 s_2 \cdot} k}\{\boldsymbol{t}_{s_1 s_2}(\boldsymbol{u}_{s_1 s_2})\}(-v_1)^k,
\end{aligned}
\tag{3.37}
$$

for

$$
\begin{aligned}
&a_{s_1 s_2, d_{s_1 s_2 \cdot} k}\{\boldsymbol{t}_{s_1 s_2}(\boldsymbol{u}_{s_1 s_2})\} \\
&= \sum_{l=k}^{d_{s_1 s_2 \cdot}} \left\{ a_{s_1 s_2, lk}\{t^*_{s_1 s_2}(\boldsymbol{u}_{s_1 s_2})\} \right. \\
&\quad \left. \cdot \left(\sum_{\boldsymbol{j} \in \mathcal{Q}^{d_{s_1 s_2}}_{\boldsymbol{d}_{s_1 s_2}, l}} \prod_{s_3=1}^{d_{s_1 s_2}} B_{d_{s_1 s_2 s_3}, j_{s_3}} \left[(\mathring{\psi}^{(k)}_{s_1 s_2, s_1 s_2 s_3}\{t_{s_1 s_2 s_3}(\boldsymbol{u}_{s_1 s_2 s_3})\})_{k \in \{1, \ldots, d_{s_1 s_2 s_3} - j_{s_3} + 1\}} \right] \right) \right\}.
\end{aligned}
\tag{3.38}
$$

With this notation, the connection with Theorem 3.3, Part (1) is clearly visible. Finally, in order to solve Challenges 2 and 3, one introduces

$$
b^{d_{s_1}}_{\boldsymbol{d}_{s_1}, k}\{\boldsymbol{t}(\boldsymbol{u})\} = \sum_{\boldsymbol{j} \in \mathcal{Q}^{d_1}_{\boldsymbol{d}_{s_1}, k}} \prod_{s_2=1}^{d_{s_1}} a_{s_1 s_2, d_{s_1 s_2 \cdot} j_{s_2}}\{\boldsymbol{t}_{s_1 s_2}(\boldsymbol{u}_{s_1 s_2})\},
\tag{3.39}
$$

with

$$
\begin{aligned}
\boldsymbol{t}(\boldsymbol{u}) &= (\boldsymbol{t}_1(\boldsymbol{u}_1)^\top, \ldots, \boldsymbol{t}_{d_1}(\boldsymbol{u}_{d_1})^\top)^\top, \\
\boldsymbol{t}_{s_1}(\boldsymbol{u}_{s_1}) &= (\boldsymbol{t}_{s_1 1}(\boldsymbol{u}_{s_1 1})^\top, \ldots, \boldsymbol{t}_{s_1 d_{s_1}}(\boldsymbol{u}_{s_1 d_{s_1}})^\top)^\top, \\
\boldsymbol{d}_{s_1} &= (\boldsymbol{d}^\top_{s_1 1}, \ldots, \boldsymbol{d}^\top_{s_1 d_{s_1}})^\top.
\end{aligned}
$$

Similarly to Theorem 3.3, Part (2), one then obtains (with $t(\boldsymbol{u}) = \psi_1^{-1}\{C(\boldsymbol{u})\}$ as before)

$$\begin{aligned} c(\boldsymbol{u}) &= \int_0^\infty \prod_{s_2=1}^{d_{s_1}} \frac{\partial}{\partial \boldsymbol{u}_{s_1 s_2}} \psi_{s_1,s_1s_2}\{t^*_{s_1s_2}(\boldsymbol{u}_{s_1s_2}); v_1\}\, dF_1(v_1) \\ &= \Big\{ \prod_{s_2=1}^{d_{s_1}} \prod_{s_3=1}^{d_{s_1s_2}} \prod_{s_4=1}^{d_{s_1s_2s_3}} (\psi^{-1}_{s_1s_2s_3})'(u_{s_1s_2s_3s_4}) \Big\} \sum_{k=d_{s_1}}^{d} b^{d_{s_1}}_{\boldsymbol{d}_{s_1},k}\{\boldsymbol{t}(\boldsymbol{u})\} \\ &\quad \cdot \int_0^\infty \Big[\prod_{s_2=1}^{d_{s_1}} \psi_{s_1,s_1s_2}\{t^*_{s_1s_2}(\boldsymbol{u}_{s_1s_2}); v_1\} \Big] (-v_1)^k dF_1(v_1) \end{aligned}$$

Finally, by replacing the integral by $\psi_1^{(k)}$, one gets

$$c(\boldsymbol{u}) = \Big\{ \prod_{s_2=1}^{d_{s_1}} \prod_{s_3=1}^{d_{s_1s_2}} \prod_{s_4=1}^{d_{s_1s_2s_3}} (\psi^{-1}_{s_1s_2s_3})'(u_{s_1s_2s_3s_4}) \Big\} \sum_{k=d_{s_1}}^{d} b^{d_{s_1}}_{\boldsymbol{d}_{s_1},k}\{\boldsymbol{t}(\boldsymbol{u})\}\, \psi_1^{(k)}\{t(\boldsymbol{u})\}. \tag{3.40}$$

Remark 3.9. The pattern to compute the density of nested Archimedean copulas with more than three levels can be deduced from the previous computations, the following heuristic argument shows how. In order to understand the reasoning, it is useful to remind ourselves that the structure of nested Archimedean copulas can be depicted by trees; see Figure 3.5 for a tree representation of (3.31). Let L denote the number of levels (with (3.31) having $L = 3$). Thanks to our notation, we can easily identify a certain branch of the tree with the corresponding sequence of indices. Each time a nesting level is added, for each branch $s_1 s_2 \cdots s_l$, $l \in \{1, \dots, L\}$, finite sequences of coefficients a's and b's (similar to $a_{s_1s_2,d_{s_1s_2}\cdot k}$ and $b^{d_{s_1}}_{\boldsymbol{d}_{s_1},k}$ above) will appear and their structure can be deduced from Equation (3.38). More precisely, as in Equations (3.18) and (3.39), the sequence of b's can always be interpreted as the coefficients of the Cauchy product of the polynomials with a's as coefficients.

The structure of the a's at each branch $s_1 \dots s_l$ is more complicated. For any branch $s_1 \dots s_L$, that is on the ultimate level of nesting, the a's are simply the Bell polynomials applied to the function $\mathring{\psi}_{s_1 \dots s_{L-1}, s_1 \cdots s_L}$ and its derivatives. For any other branch $s_1 \dots s_l$, $l \in \{1, \dots, L-1\}$, the coefficients a's are the (Euclidean) inner product of the vector of all Bell polynomials applied to the function $\mathring{\psi}_{s_1 \dots s_l, s_1 \dots s_{l+1}}$ and its derivatives with the vector of all coefficients b's, the exact structure of Equation (3.38). In Equation (3.38), the level l is equal to $1 = L - 2$ and the term $a_{s_1s_2,lk}\{t^*_{s_1s_2}(\boldsymbol{u}_{s_1s_2})\}$ stands for the Bell polynomial applied to $\mathring{\psi}_{s_1 \dots s_l, s_1 \dots s_{l+1}}$ and its derivatives, while the lth member of the sequence of b's is defined by

$$\sum_{\boldsymbol{j} \in \mathcal{Q}^{d_{s_1s_2}}_{\boldsymbol{d}_{s_1s_2},l}} \prod_{s_3=1}^{d_{s_1s_2}} B_{d_{s_1s_2s_3}, j_{s_3}} \Big[\big(\mathring{\psi}^{(k)}_{s_1s_2,s_1s_2s_3}\{t_{s_1s_2s_3}(\boldsymbol{u}_{s_1s_2s_3})\} \big)_{k \in \{1, \dots, d_{s_1s_2s_3} - j_{s_3} + 1\}} \Big],$$

the term appearing in (3.38).

CHAPTER 4

CONCLUSION

After a short introduction to copulas and to the Archimedean family in particular, general formulas for the derivatives of the nodes, the inner generators, and for the densities of two-level and three-level nested Archimedean copulas were presented. For the densities, the main idea is to first conveniently use an integral representation for nested Archimedean copulas, to differentiate this representation, to simplify the appearing expressions, and finally to reinterpret the integrals in terms of the generator derivatives of the root copula. In contrast to trying to differentiate the nested structure directly, this approach leads to formulas that are tractable. The main results (for two nesting levels) were computed for various examples including a generator transformation and well-known nested Archimedean families. Furthermore, an efficient way to compute the log-density was presented. Extensions to higher nesting levels were briefly addressed. Results on Bell polynomials and Faà di Bruno's formula were also introduced and are the essential mathematical tools needed in order to complete this work. Finally, further studies, in particular in simulation, should be conducted in order to access the behavior of Model (3.1) when it is fitted to real data.

BIBLIOGRAPHY

Ali, M. M., N. N. Mikhail, and M. S. Haq (1978). A Class of Bivariate Distributions Including the Bivariate Logistic. *Journal of Multivariate Analysis 8*, 405–412.

Boyadzhiev, K. N. (2009). Exponential Polynomials, Stirling Numbers, and Evaluation of Some Gamma Integrals. *Abstract and Applied Analysis 2009*.

Charalambides, C. A. (2005). *Combinatorial Methods in Discrete Distributions*. Wiley-Interscience.

Clayton, D. G. (1978). A model for association in bivariate life tables and its application in epidemiological studies of familial tendency in chronic disease incidence. *Biometrika 65*(1), 141–151.

Comtet, L. (1974). *Advanced Combinatorics: The Art of Finite and Infinite Expansions*. Dordrecht, Holland / Boston, U.S.A.: D. Reidel Publishing Company.

Embrechts, P. (2009). Copulas: A personal view. *Journal of Risk and Insurance 76*(3), 639–650.

Embrechts, P. and M. Hofert (2011). A note on generalized inverses. submitted.

Embrechts, P., A. J. McNeil, and D. Straumann (2002). Correlation and Dependence in Risk Management: Properties and Pitfalls. In M. Dempster (Ed.), *Risk Management: Value at Risk and Beyond*, pp. 176–223. Cambridge University Press.

Feller, W. (1971). *An Introduction to Probability Theory and Its Applications* (2 ed.), Volume 2. Wiley.

Fisher, N. I. (1997). Copulas. *Encyclopedia of Statistical Sciences 1*, 159–163.

Frank, M. J. (1979). On the simultaneous associativity of $F(x,y)$ and $x+y-F(x,y)$. *Aequationes Mathematicae 19*, 194–226.

Fréchet, M. (1935). Généralisations du théorème des probabilités totales. *Fundamenta Mathematica 25*, 379–387.

Genest, C. and R. J. MacKay (1986). Copules archimédiennes et familles de lois bidimensionnelles dont les marges sont données. *The Canadian Journal of Statistics 14*, 145–159.

Gould, H. (2010). Combinatorial numbers and associated identities: Table 1: Stirling numbers.

Hardy, M. (2006). Combinatorics of Partial Derivatives. *The Electronic Journal of Combinatorics 13(1)*.

Hering, C., M. Hofert, J.-F. Mai, and M. Scherer (2010). Constructing nested Archimedean copulas with Lévy subordinators. *Journal of Multivariate Analysis 101*(6), 1428–1433.

Hoeffding, W. (1940). Massstabinvariante Korrelationstheorie. *Schriften des Mathematischen Seminars und des Instituts für Angewandte Mathematik der Universität Berlin 5*, 181–233.

Hofert, M. (2010). *Sampling Nested Archimedean Copulas with Applications to CDO Pricing*. Südwestdeutscher Verlag für Hochschulschriften AG & Co. KG. PhD thesis.

Hofert, M. (2011a). A stochastic representation and sampling algorithm for nested Archimedean copulas. *Journal of Statistical Computation and Simulation*. in press.

Hofert, M. (2011b). Efficiently sampling nested Archimedean copulas. *Computational Statistics & Data Analysis 55*, 57–70.

Hofert, M. and M. Mächler (2011). Nested Archimedean Copulas Meet R: The nacopula Package. *Journal of Statistical Software 39*(9), 1–20.

Hofert, M., M. Mächler, and A. J. McNeil (2011a). Estimators for Archimedean copulas in high dimensions: A comparison. in progress.

Hofert, M., M. Mächler, and A. J. McNeil (2011b). Likelihood inference for Archimedean copulas. submitted.

Hofert, M. and M. Scherer (2011). CDO pricing with nested Archimedean copulas. *Quantitative Finance 11*(5), 775–787.

Joe, H. (1993). Parametric families of multivariate distributions with given margins. *Journal of Multivariate Analysis 46*, 262–282.

Joe, H. (1997). *Multivariate Models and Dependence Concepts*. Chapman & Hall/CRC.

Joe, H. and T. Hu (1996). Multivariate Distributions from Mixtures of Max-Infinitely Divisible Distributions. *Journal of Multivariate Analysis 57*, 240–265.

Kimberling, C. H. (1974). A probabilistic interpretation of complete monotonicity. *Aequationes Mathematicae 10*, 152–164.

Ling, C. H. (1965). Representation of associative functions. *Publicationes Mathematicae Debrecen 12*, 189–212.

Malov, S. V. (2001). On finite-dimensional Archimedean copulas. In N. Balakrishnan, I. Ibragimov, and V. Nevzorov (Eds.), *Asymptotic Methods in Probability and Statistics with Applications*, pp. 19–35. Birkhäuser.

Marshall, A. W. and I. Olkin (1988). Families of Multivariate Distributions. *Journal of the American Statistical Association 83*(403), 834–841.

McNeil, A. J. (2008). Sampling nested Archimedean copulas. *Journal of Statistical Computation and Simulation 78*(6), 567–581.

McNeil, A. J., R. Frey, and P. Embrechts (2005). *Quantitative Risk Management: Concepts, Techniques, Tools.* Princeton University Press.

Nelsen, R. B. (2007). *An Introduction to Copulas.* Springer.

Nolan, J. P. (2011). *Stable Distributions – Models for Heavy Tailed Data.* Birkhäuser.

Rüschendorf, L. (2009). On the distributional transform, Sklar's Theorem, and the empirical copula process. *Journal of Statistical Planning and Inference 139*(11), 3921–3927.

Schweizer, B. and A. Sklar (1961). Associative functions and statistical triangle inequalities. *Publicationes Mathematicae Debrecen 8*, 169–186.

Schweizer, B. and A. Sklar (1963). Associative functions and abstract semigroups. *Publicationes Mathematicae Debrecen 10*, 69–81.

Schweizer, B. and E. F. Wolff (1981). On nonparametric measures of dependence for random variables. *The Annals of Statistics 9*, 879–885.

Sibuya, M. (1959). Bivariate extreme statistics, I. *Annals of the Institute of Statistical Mathematics 11*(2), 195–210.

Sklar, A. (1959). Fonctions de répartition à n dimensions et leurs marges. *Publications de L'Institut de Statistique de L'Université de Paris 8*, 229–231.

Sklar, A. (1996). Random variables, distribution functions, and copulas – a personal look backward and forward. *Distributions with Fixed Marginals and Related Topics 28*, 1–14.

Wood, D. (1992). The computation of polylogarithms.