Manuel Úbeda Flores
Enrique de Amo Artero
Fabrizio Durante
Juan Fernández Sánchez *Editors*

Copulas and Dependence Models with Applications

Contributions in Honor of Roger B. Nelsen

Springer

Copulas and Dependence Models with Applications

Manuel Úbeda Flores • Enrique de Amo Artero
Fabrizio Durante • Juan Fernández Sánchez
Editors

Copulas and Dependence Models with Applications

Contributions in Honor of Roger B. Nelsen

Springer

Editors
Manuel Úbeda Flores
Department of Mathematics
University of Almería
Almería, Spain

Enrique de Amo Artero
Department of Mathematics
University of Almería
Almería, Spain

Fabrizio Durante
Dipartimento di Scienze dell'Economia
Università del Salento
Lecce, Italia

Juan Fernández Sánchez
Research Group of Mathematical Analysis
University of Almería
Almería, Spain

ISBN 978-3-319-64220-8 ISBN 978-3-319-64221-5 (eBook)
DOI 10.1007/978-3-319-64221-5

Library of Congress Control Number: 2017952950

Mathematics Subject Classification (2010): 62Exx, 62Fxx, 62Hxx, 60Exx, 28Axx, 62Pxx

Printed on acid-free paper

This Springer imprint is published by Springer Nature
The registered company is Springer International Publishing AG
The registered company address is: Gewerbestrasse 11, 6330 Cham, Switzerland

Preface

This edited volume has been written to celebrate the 75th birthday of Prof. Roger B. Nelsen, Professor Emeritus of Mathematics at Lewis & Clark College in Portland, Oregon, USA. In addition to his monograph *An Introduction to Copulas*, Prof. Nelsen has authored or coauthored 11 books published by the Mathematical Association of America (mainly devoted to ”proofs without words”) and numerous research papers; in particular, dealing with Copula Theory, dependence and association measures. He has been universally recognized as one of the most influential scientist in the copula community, as witnessed by a number of invited talks he has given around the world along his career. For a historical overview of the Prof. Nelsen’s main achievements as well as their developments, you may read the recent interview published by the journal *Dependence Modeling*.

This edited volume agrees with the main research areas investigated by Prof. Nelsen during his long and inspiring career and, in particular, to Copula Theory. A copula is a multivariate probability distribution whose univariate marginals are uniform on the unit interval and is used to describe the dependence between random variables. Motivated by an open problem posed by Maurice Fréchet, Abe Sklar introduced in 1959 the concept of copula to explain the relationship between multivariate distribution functions and their univariate marginals. Since then, copulas have been used in various fields, primarily in Probability Theory and Statistics, but also in Probabilistic Metric Spaces, Fuzzy Set Theory and Aggregation Theory, and in many applications to Economics, Finance, Biosciences, Environmental Sciences, just to cite few of them.

In the first edition of his celebrated *An Introduction to Copulas* (1999), Prof. Nelsen wrote:

> The study of copulas and the role they play in probability, statistics, and stochastic processes is a subject still in its infancy. There are many open problems and much work to be done.

After almost twenty years, we can see that his words are still up-to-date, as can be inferred from the various results presented here.

This volume includes 15 invited contributions dealing with different results related to copulas, quasi-copulas and related concepts, as well as to various aspects of dependence modelling and their main applications. Authored by well-known researchers in the field, the book contains both original contributions and surveys, with particular emphasis on classical topics, such as distributions with fixed marginals, constructions of copula models, etc. Most of the present contributions will be also presented and discussed in a scientific conference to be held in Almería, Spain, in July 2017.

The purpose of this edited volume is to give the copula community the opportunity of knowing some of the more significant contributions of Prof. Nelsen and provide some new contributions and reviews related to this field. The primary audience of this book are researchers as well as practitioners, in stochastic models of dependence from a variety of different perspectives.

The Editors apologize for the inevitable non-uniformity in the form and the style of the contributions and for typos that might be inadvertently in. However, each contribution is self-consistent and the possible different mathematical notations are auto-explained in each chapter and in the corresponding references.

A brief summary of the contents of the chapters is mentioned below.

Chapter 1 by de Amo, Díaz Carrillo, and Fernández Sánchez reviews problems related to copula sections and some constructions of copulas with some additional information, including copulas with given diagonal and/or opposite diagonal sections, semilinear copulas, biconic copulas, copulas with quadratic sections, among others.

Next chapter, by Cherubini and Mulinacci, is devoted to the Gumbel-Marshall-Olkin distribution, a generalization of the Marshall-Olkin distribution, an important model to study default risks. The extension keeps the main structure of the model and it is suitable for credit risk applications.

Chapter 3 by De Baets and De Meyer is devoted to the study of the degree of asymmetry of a (quasi-)copula with respect to a continuous strictly increasing curve in the unit square, generalizing the seminal work of Prof. Nelsen in this matter. The authors also derive the maximum and the minimum value of asymmetry of a copula with respect to a curve.

In Chapter 4, Di Lascio, Durante, and Pappadà review some clustering methods which use a dissimilarity measure based either on concordance or tail dependence or risk measures and a clustering procedure based on the likelihood of the copula. Illustrations of the above methods are provided.

Chapter 5, written by Erdely, deals with the study of a gluing copula approach to decompose the underlying copula into totally ordered copulas that, when combined, may lead to a non-monotone regression function. The idea arises because most common parametric families of copulas are totally ordered and, in many cases, they lead to monotone regression functions.

Chapter 6, by Genest and Nešlehová, presents a review of two of the most known families of copulas: Gumbel and Galambos copulas. The authors of this chapter recall some of their main properties and show their “connections” in any dimension.

The next chapter, by Jaworski, provides a study of systemic risk management by using copula methods. In particular, the compatibility of the modified Conditional Value at Risk (CoVaR) with the concordance ordering of copulas is considered. Moreover, the modified CoVaR is investigated in several known families of copulas.

In Chapter 8, tail asymmetry and dependence properties and measures for copulas are summarized by Joe. With this purpose, new bivariate parametric families of copulas are presented and some of their dependence and asymmetry properties are determined.

Chapter 9, by Klement, Kolesárová, Mesiar, and Saminger-Platz, deals with the study of some copula constructions by means of ultramodular functions. Interesting connections with Schur concavity and functional inequalities are also stressed.

In the next chapter, Mayor, Suñer, and Torrens review two families of discrete operations (or operations defined on finite chains): On the one hand, discrete triangular norms, that are generalizations of copulas applied in fuzzy logic and approximate reasoning; on the other hand, discrete copulas, with applications in probability, statistics and economy.

Chapter 11 by Quesada-Molina reviews the long-standing collaboration of Prof. Roger Nelsen with him and some of his colleagues. The author recalls, from the first, his numerous meetings with Prof. Nelsen, highlighting how some of the most influential articles they coauthored were brewing, including those that study copulas with prescribed sections, distribution functions of copulas, best-possible bounds on sets of copulas, properties of quasi-copulas.

Next chapter, authored by Rüschendorf, reviews and elaborates on several developments of improved Fréchet-Hoeffding bounds for the distribution of a random vector which assume some restriction on the dependence structure additional to the information on the marginals. Improved VaR bounds for the joint portfolio of risk vectors are obtained.

In Chapter 13, written by Sempi, a complete survey on the concept of quasi-copula is analyzed, with special attention to its different characterizations, main properties, recent results and the important contributions of Prof. Nelsen in this matter. Some open problems are also listed.

Chapter 14, by Trutschnig, deals with some problems related to complete dependence, putting the focus on dependence measures that detect functional dependence, on a problem related to joint-default maximization, on a question from uniform distribution theory, and on the relationship between the two most well-known measures of concordance, Kendall's tau and Spearman's rho. A short excursion to topology showing that complete dependence is not at all atypical complements the chapter.

Last chapter, by Úbeda-Flores and Fernández-Sánchez, presents a complete proof of Sklar's theorem following seminal ideas by Abe Sklar. The authors also provide an alternative proof by using the Zorn's lemma, and review other proofs appeared in the literature.

The Editors want to dedicate this book to Prof. Roger Nelsen, for his important contributions to the Theory of Copulas, and they also wish that this edited volume will contribute to promote new theoretical and practical results. Moreover, the Editors would like to thank each one of the authors for all their contributions, because without their enthusiastic efforts the volume would have not been possible.

The Editors are also grateful to Springer-Verlag for giving them the opportunity of publishing this editorial work. Finally, they also acknowledge the support by the Ministerio de Economía y Competitividad (Spain) under research project MTM2014-60594-P and FEDER.

March 2017

Almería, Spain *Enrique de Amo Artero*
Juan Fernández Sánchez
Manuel Úbeda Flores

Lecce, Italy *Fabrizio Durante*

Contents

The original version of this book was revised. An erratum to this book can be found at
https://doi.org/10.1007/978-3-319-64221-5_16

List of Contributors

Enrique de Amo
Departamento de Matemáticas, Universidad de Almería, Almería, Spain
e-mail: edeamo@ual.es

Umberto Cherubini
Department of Statistics, University of Bologna, Bologna, Italy
e-mail: umberto.cherubini@unibo.it

Bernard De Baets
KERMIT, Department of Mathematical Modelling, Statistics and Bioinformatics, Ghent University, Gent, Belgium
e-mail: bernard.debaets@ugent.be

Hans De Meyer
Department of Applied Mathematics, Computer Science and Statistics, Ghent University, Gent, Belgium
e-mail: hans.demeyer@ugent.be

F. Marta L. Di Lascio
Faculty of Economics and Management, Free University of Bozen-Bolzano, Bolzano, Italy
e-mail: marta.dilascio@unibz.it

Manuel Díaz Carrillo
Departamento de Análisis Matemático, Universidad de Granada, Granada, Spain
e-mail: madiaz@ugr.es

Fabrizio Durante
Dipartimento di Scienze dell'Economia, Università del Salento, Lecce, Italy
e-mail: fabrizio.durante@unisalento.it

Arturo Erdely
Facultad de Estudios Superiores Acatlán, Universidad Nacional Autónoma de México, Naucalpan, Mexico

e-mail: arturo.erdely@comunidad.unam.mx

Juan Fernández Sánchez
Grupo de Investigación de Análisis Matemático, Universidad de Almería, Almería, Spain
e-mail: juanfernandez@ual.es

Christian Genest
Department of Mathematics and Statistics, McGill University, Montréal, Québec, Canada
e-mail: christian.genest@mcgill.ca

Piotr Jaworski
Institute of Mathematics, University of Warsaw, Warsaw, Poland
e-mail: P.Jaworski@mimuw.edu.pl

Harry Joe
Departament of Statistics, University of British Columbia, Vancouver, Canada
e-mail: Harry.Joe@ubc.ca

Erich Peter Klement
Department of Knowledge-Based Mathematical Systems, Johannes Kepler University, Linz, Austria
e-mail: ep.klement@jku.at

Anna Kolesárová
Department of Information Engineering, Automation and Mathematics, Faculty of Chemical and Food Technology, Slovak University of Technology, Bratislava, Slovakia
e-mail: anna.kolesarova@stuba.sk

Gaspar Mayor
Department of Mathematics and Computer Science, Universitat de les Illes Balears, Palma de Mallorca, Spain
e-mail: gmayor@uib.es

Radko Mesiar
Department of Mathematics and Descriptive Geometry, Faculty of Civil Engineering, Slovak University of Technology, Bratislava, Slovakia
e-mail: radko.mesiar@stuba.sk

Sabrina Mulinacci
Department of Statistics, University of Bologna, Bologna, Italy
e-mail: sabrina.mulinacci@unibo.it

Johanna G. Nešlehová
Department of Mathematics and Statistics, McGill University, Montréal, Québec, Canada
e-mail: johanna.neslehova@mcgill.ca

Roberta Pappadà
Department of Economics, Business, Mathematics and Statistics “Bruno de Finetti”, University of Trieste, Trieste, Italy
e-mail: rpappada@units.it

José Juan Quesada-Molina
Department of Applied Mathematics, University of Granada, Granada, Spain
e-mail: jquesada@ugr.es

Ludger Rüschendorf
University of Freiburg, Freiburg, Germany
e-mail: ruschen@stochastik.uni-freiburg.de

Susanne Saminger-Platz
Department of Knowledge-Based Mathematical Systems, Johannes Kepler University, Linz, Austria
e-mail: susanne.saminger-platz@jku.at

Carlo Sempi
Dipartimento di Matematica e Fisica “Ennio De Giorgi”, Università del Salento, Lecce, Italy
e-mail: carlo.sempi@unisalento.it

Jaume Suñer
Departament of Mathematics and Computer Science, Universitat de les Illes Balears, Palma de Mallorca, Spain
e-mail: jaume.sunyer@uib.es

Joan Torrens
Departament of Mathematics and Computer Science, Universitat de les Illes Balears, Palma de Mallorca, Spain
e-mail: jts224@uib.es

Wolfgang Trutschnig
Department for Mathematics, University Salzburg, Salzburg, Austria
e-mail: wolfgang@trutschnig.net

Manuel Úbeda-Flores
Department of Mathematics, University of Almería, Almería, Spain
e-mail: mubeda@ual.es

Chapter 1
Constructions of copulas under prescribed sections

Enrique de Amo, Manuel Díaz Carrillo, and Juan Fernández Sánchez

Abstract The main problems related to copula sections are reviewed and various methods for constructing copulas that preserve some partial information, are presented. Here, the scope of interest extends from seminal work on the existence and construction of copulas with given diagonal sections to the most recent research on copulas with given diagonal and opposite diagonal sections. Also a survey is given on the state of the art on related domains such as the construction of special families of copulas and generalizations of the copula concept.

1.1 Introduction

Since copulas were introduced by A. Sklar in 1959 as the solution to a question proposed by M. Fréchet, this special type of distribution function has had an increasing role in probability and statistics. But the interest in copulas is not reduced, whatsoever, to these topics. Copula theory has also revealed to be of importance in fields such as probabilistic metric spaces, fuzzy set theory or aggregation theory.

One of the problems initially studied in copula theory was that of constructing copulas given some partial information, for instance, copulas with given section(s). Much attention was focused on the construction of copulas with given diagonal section and on copulas having quadratic functions as vertical or horizontal sections.

Enrique de Amo
Departamento de Matemáticas, Universidad de Almería, Almería, Spain
e-mail: edeamo@ual.es

Manuel Díaz Carrillo
Departamento de Análisis Matemático, Universidad de Granada, Granada, Spain
e-mail: madiaz@ugr.es

Juan Fernández Sánchez
Grupo de Investigación de Análisis Matemático, Universidad de Almería, Almería, Spain
e-mail: juanfernandez@ual.es

M. Úbeda Flores et al. (eds.), *Copulas and Dependence Models with Applications*,
DOI 10.1007/978-3-319-64221-5_1

Since then, an ever growing variety of problems has been studied: copulas with vertical and horizontal sections, compatibility between them, diagonal and opposite diagonal and their compatibility, multidimensional generalizations, etc.

Here we give a survey of the main problems related to copula sections and of the results and methods that are used to construct copulas that preserve some partial information.

1.2 Preliminaries

Let us recall the definitions we use throughout this chapter. By $\mathbb{I}$ we denote the unit closed interval $[0,1]$, $\mathbb{Z}^+$ is the set of positive integers, by id we denote the identity function.

Definition 1.1. A *two-dimensional copula* is a function $C:\mathbb{I}^2 \to \mathbb{I}$ with the following properties:

C.1 (Grounded) For every $u,v \in \mathbb{I}$, $C(u,0) = 0 = C(0,v)$.
C.2 (With Uniform margins) For every $u,v \in \mathbb{I}$, $C(u,1) = u$, $C(1,v) = v$.
C.3 (2-increasing) For every $u_1,u_2,v_1,v_2 \in \mathbb{I}^2$ such that $u_1 \le u_2$ and $v_1 \le v_2$,

$$C(u_2,v_2) - C(u_2,v_1) - C(u_1,v_2) + C(u_1,v_1) \ge 0.$$

Alternatively, we can say that a copula is a bidimensional distribution function concentred on $\mathbb{I}^2$ whose marginal distribution functions are uniformly distributed. Therefore, each copula C induces a probability measure μ_C on $\mathbb{I}^2$ in the following way. We start defining the C-volume of rectangles $[a,b]\times[c,d]$ via the formula

$$V_C([a,b]\times[c,d]) = C(b,d) - C(b,c) - C(a,d) + C(a,c).$$

V_C can be extended to the σ-algebra $\mathscr{B}\left(\mathbb{I}^2\right)$ of Borel subsets to a corresponding μ_C measure. We denote by λ the standard Lebesgue measure on the σ-algebra of Borel sets. The *support of a copula C* is the complement of the union of all open subsets of $\mathbb{I}^2$ with μ_C-measure equal to zero, and we denote it by $\mathrm{Supp}(C)$.

We denote by $\mathscr{C}^2$ the class of all two-dimensional copulas. For each $n \ge 2$ we can also generalise the concept of copula:

Definition 1.2. An *n-dimensional copula* is a function $C:\mathbb{I}^n \to \mathbb{I}$ with the following properties:

C.1' (Grounded) $C(u_1,\dots,u_{i-1},0,u_{i+1},\dots,u_n) = 0$ for all $i = 1,\dots,n$.
C.2' (With Uniform univariate margins) $C(1,\dots,1,u_i,1,\dots,1) = u_i$ for all $i = 1,\dots,n$.
C.3' (n-increasing) For each rectangle $[\mathrm{a},\mathrm{b}] := \mathop{\times}\limits_{i=1}^{n} [a_i,b_i] \subset \mathbb{I}^n$, with $a_i \le b_i$ for all $i = 1,\dots,n$,

$$V_C([\mathrm{a},\mathrm{b}]) := \sum_{j_1=1}^{2} \cdots \sum_{j_n=1}^{2} (-1)^{j_1+\cdots+j_n} C\left(u_{1_{j_1}}, \ldots, u_{n_{j_n}}\right) \geq 0$$

where $u_{i_1} = a_i$ and $u_{i_2} = b_i$ for all $i = 1, \ldots, n$.

We denote by $\mathscr{C}^n$ the set of all *n-copulas*.

For any integer $n \geq 2$, an n-dimensional copula is the restriction to the unit n-cube $\mathbb{I}^n$ of a multivariate cumulative distribution function whose margins are uniform on $\mathbb{I}$. They were introduced by Sklar in 1959 (see [54]), as the answer to a question posed by M. Fréchet, and they allow to represent a joint distribution of random variables as a function of marginal distributions. Precisely, Sklar enunciated the result that follows:

Theorem 1.1 (Sklar). *If F is the joint distribution function of n random variables $X_1, \ldots, X_n$, and $F_1, \ldots, F_n$ are the distribution functions of $X_1, \ldots, X_n$, resp., then there exists an n-dimensional copula C such that*

$$F(x_1, \ldots, x_n) = C(F_1(x_1), \ldots, F_n(x_n))$$

for all $(x_1, \ldots, x_n) \in \mathbb{R}^n$. Such C is uniquely determined on $\mathrm{Ran}(F_1) \times \cdots \times \mathrm{Ran}(F_n)$. *Furthermore, the copula C is uniquely determined when the margins $F_1, \ldots, F_n$ are continuous. (Observe that* $\mathrm{x} = (x_1, \ldots, x_n) \in \mathbb{R}^n$, *but* $\mathrm{u} = (F_1(x_1), \ldots, F_n(x_n)) \in \mathbb{I}^n$.*)*

The first proof of this theorem (in the bidimensional case) was published in 1974 by Schweizer and Sklar [50]. For a constructive proof of Sklar's Theorem see [3, 5, 57].

Example 1.1. The function M, given by $M(\mathrm{u}) := \min\{u_1, \ldots, u_n\}$, for all $\mathrm{u} := (u_1, \ldots, u_n)$
$\in \mathbb{I}^n$, is a copula for all n. It is called the *comonotonicity copula.* The function W is given by $W(\mathrm{u}) := \max\{u_1 + \cdots + u_n - n + 1, 0\}$, for all $\mathrm{u} := (u_1, \ldots, u_n) \in \mathbb{I}^n$ is a copula only if $n = 2$. In this case it is called the *countermonotonicity copula*. But (*Fréchet-Hoeffding bounds inequality*), for every n-copula C,

$$W(\mathrm{u}) \leq C(\mathrm{u}) \leq M(\mathrm{u}), \text{ for all } \mathrm{u} := (u_1, \ldots, u_n) \in \mathbb{I}^n;$$

hence, M and W are respectively called the *upper* and *lower Fréchet-Hoeffding bounds*. Moreover:

$$W(\mathrm{u}) = \inf\{C(\mathrm{u}) : C \in \mathscr{C}^n\}, \quad M(\mathrm{u}) = \sup\{C(\mathrm{u}) : C \in \mathscr{C}^n\}.$$

The two following lemmas state very important properties for copulas:

Lemma 1.1. *Copulas are* 1-Lipschitz functions*: for all $C \in \mathscr{C}^n$, and* x *and* y *in* $\mathbb{I}^n$, *we have*

$$|C(\mathrm{x}) - C(\mathrm{y})| \leq \sum_{i=1}^{n} |x_i - y_i| =: \|\mathrm{x} - \mathrm{y}\|_1.$$

Lemma 1.2. *The class $\mathscr{C}^n$ of the n-copulas is a convex set. This set is also compact w.r.t. the metric of the supremum.*

One first schema of constructing copulas is given by the so-called ordinal sums:

Definition 1.3. Let $\mathrm{N} \subset \mathbb{Z}^+$ and $\{]a_i, b_i[\subset \mathbb{I} : i \in \mathrm{N}\}$ be a family of non-empty and two-by-two non-overlapping open intervals. Let $\{C_i : i \in \mathrm{N}\}$ be a family of copulas. The *ordinal sum of* $\{C_i : i \in \mathrm{N}\}$ *w.r.t.* $\{]a_i, b_i[\subset \mathbb{I} : i \in \mathrm{N}\}$ is defined, for each $\mathrm{u} \in \mathbb{I}^n$, by

$$C(\mathrm{u}) := \sum_{i \in \mathrm{N}} (b_i - a_i) C_i \left(\frac{u_1 - a_i}{b_i - a_i}, \ldots, \frac{u_n - a_i}{b_i - a_i} \right) \\ + \lambda \left([0, \min\{u_1, \ldots, u_n\}] \setminus \cup_{i \in \mathrm{N}}]a_i, b_i[\right),$$

where λ is the Lebesgue measure.

The ordinal sum above is, in fact, a copula (see [21, Th.3.8.2]). We denote it by $C = \{\langle]a_i, b_i[, C_i \rangle : i \in \mathrm{N}\}$.

1.3 Quasi-copulas

Quasi-copulas were introduced in [1] (see [52]) in order to characterize operations on univariate distribution functions that can or cannot be derived from corresponding operations on random variables (defined on the same probability space). The original definition was as follows:

Definition 1.4. A *quasi-copula* is a function $Q : \mathbb{I}^2 \to \mathbb{I}$ such that for every *track* B in $\mathbb{I}^2$ (i.e., B can be described as $\{(\alpha_1(t), \alpha_2(t)) : t \in \mathbb{I}\}$ for some continuous and nondecreasing functions α_i with $\alpha_i(0) = 1 - \alpha_i(1) = 0$ $i \in \{1, 2\}$), there exists a copula C_B such that $Q(x, y) = C_\mathrm{B}(x, y)$ with $(x, y) \in \mathrm{B}$.

First, we should clarify whether there exist quasi-copulas that are not copulas. As this is the case, it is necesary a characterization of quasi-copulas. The answer is based on the following result.

Proposition 1.1. *Let $x, y, q : \mathbb{I} \to \mathbb{I}$ be three continuous functions satisfying:*

i. *x, y and q are non-decreasing;*
ii. $0 \le q(t_2) - q(t_1) \le |x(t_2) - x(t_1)| + |y(t_2) - y(t_1)|$, *for all* $0 \le t_1 \le t_2 \le 1$;
iii. $\max(0, x(t) + y(t) - 1) \le q(t) \le \min(x(t), y(t))$, *for all* $t \in \mathbb{I}$.

Then there exists a copula C such that $C(x(t), y(t)) = q(t)$ for all $t \in \mathbb{I}$.

Corollary 1.1 ([30]). *A function $Q : \mathbb{I}^2 \to \mathbb{I}$ is a quasi-copula if and only if*

i. $Q(0, x) = Q(x, 0) = 0$ *and* $Q(1, x) = Q(x, 1) = x$ *for all* $x \in \mathbb{I}$;

ii. Q is a non-decreasing function in each of its arguments;
iii. Q is 1-Lipschitz: $|Q(u,v) - Q(x,y)| \leq |u-x| + |v-y|$*, for all* $u,v,x,y \in \mathbb{I}$.

Example 1.2 ([41, Exercise 2.11]). Let

$$K(x,y) := \begin{cases} \min\{x,y,1/3,x+y-2/3\}, & 2/3 \leq x+y \leq 4/3 \\ M(x,y), & \text{otherwise.} \end{cases}$$

This function K is a quasi-copula by the corollary above. But it is not a copula: $V_K\left([1/3,2/3]^2\right) < 0$.

1.4 Diagonal sections

The *diagonal section* of a copula C is the function $\delta_C : \mathbb{I} \to \mathbb{I}$ defined by $\delta_C(t) := C(t,t)$ for all $t \in \mathbb{I}$. Note that δ_C has the following properties:

D.1 $\delta_C(1) = 1$,
D.2 $\delta_C(t) \leq t$ for all $t \in \mathbb{I}$,
D.3 δ_C is increasing,
D.4 $|\delta_C(x) - \delta_C(y)| \leq 2|x-y|$ for all $x,y \in \mathbb{I}$.

A function δ satisfying conditions D.1 to D.4 is named a *diagonal function*. The set of all diagonal functions is denoted by $\mathscr{D}_2$. A natural question arises: for a given $\delta \in \mathscr{D}_2$, does there exist a copula C such that $\delta_C = \delta$? The answer is "yes", and it is a consequence of Corollary 1.1, because when $x(t) = y(t) = t$, conditions a., b., and c. in the statement are equivalent to conditions D.1 to D.4.

Theorem 1.2. *If* $\delta \in \mathscr{D}_2$*, then there exists a copula* C *such that* $\delta = \delta_C$.

Theorem 1.2 only states the existence of a such copula C, but does not provide any qualitative information on it. One can ask, for instance, whether C is necessarily a singular, or an absolutely continuous copula? In the following, we will recall various constructions methods and study the properties of the copulas we generate. Given a diagonal function δ, we write by $\mathscr{C}_\delta$ the set of all copulas C such that $\delta_C = \delta$.

1.4.1 Diagonal and Hairpin copulas

In the investigations about extreme points in the class of copulas, the concepts of *hairpin sets* and *two-dimensional hairpin copulas* have been introduced. Given an increasing homeomorphism $g : \mathbb{I} \to \mathbb{I}$ fulfilling $g(x) < x$ for every $x \in \,]0,1[$ the (compact) set $\Gamma^\star(g) := \mathrm{Gr}(g) \cup \mathrm{Gr}\left(g^{-1}\right)$ (where $\mathrm{Gr}(f) := \{(t,f(t) : t \in \mathbb{I})\}$) will be called a *two-dimensional hairpin set*. The class of all these homeomorphisms will be denoted by $\mathscr{G}$. For $g \in \mathscr{G}$, the inverse g^{-1} is strictly increasing and fulfills $g^{-1}(x) > x$ for every $x \in \,]0,1[$. We will write g^j (resp., g^{-j}) for the j-times composition of g (resp., g^{-1}) with itself for every $j \in \mathbb{Z}^+$ and set $g^0 := \mathrm{id}_{\mathbb{I}}$. Following [42, 53] (see [51] as well) we say that C is a *two-dimensional hairpin copula* if $\mathrm{Supp}(C) \subseteq \Gamma^\star(g)$ for some $g \in \mathscr{G}$.

Theorem 1.3 ([51, 53]). *For every $g \in \mathscr{G}$ there exists at most one copula $C \in \mathscr{C}^2$ with* $\mathrm{Supp}(C) \subseteq \Gamma^\star(g)$. *If such a copula exists it is necessarily symmetric and for every $x \in \,]0,1[$,*

$$f(x) := \mu_C\left(\{(t,g(t)) : 0 \le t \le x\}\right) = \sum_{n=1}^{\infty} (-1)^{n+1} g^n(x).$$

The function f so defined is named as *mass spreader*. Using notation above, we have $\delta_C = 2f$.

Theorem 1.4 ([14, 53]). *If for a given $g \in \mathscr{G}$ there exists one copula $C \in \mathscr{C}^2$ with* $\mathrm{Supp}(C) \subseteq \Gamma^\star(g)$, *then:*

i. The set O_x, defined by $O_x = \bigcup_{n\in\mathbb{Z}^+} \left[g^{2n+2}(x), g^{2n+1}(x)\right]$, fulfills $\lambda(O_x) = 1/2$.
ii. The diagonal δ_C fulfills

$$2g(x) = \delta_C(x) + \delta_C(g(x)) \tag{1.1}$$

for every $x \in \mathbb{I}$.

As a consequence of Theorem 1.3, we have the next result.

Corollary 1.2. *Hairpin copulas are extreme points in $\mathscr{C}^2$.*

To the best of our knowledge, the first time that the formula of a copula C with mass concentrated in $\Gamma^\star(g)$ appears in explicit form is in the proof of Theorem 2.4 in [38]. Nevertheless, it was already implicit in [51] and [53]. Its expression is given in terms of the mass spreader function f, but we will use the diagonal function δ_C.

Theorem 1.5. *Suppose that for a given $g \in \mathscr{G}$ there is one copula $C \in \mathscr{C}^2$ with* $\mathrm{Supp}(C) \subseteq \Gamma^\star(g)$. *Then*

$$C(x,y) = \min\left\{x, y, \frac{\delta_C(x) + \delta_C(y)}{2}\right\}$$

In the light of this theorem, it is reasonable to study the functions

$$C_\delta^{\mathbf{FN}}(x,y) := \min\left\{x, y, \frac{\delta(x) + \delta(y)}{2}\right\},$$

with $\delta \in \mathscr{D}_2$.

This study was made by Fredricks and Nelsen, and the main result is the following.

Theorem 1.6 ([28, 42]). *If $\delta \in \mathscr{D}_2$, then $C_\delta^{\mathbf{FN}}$ is a copula. Moreover, for each symmetric copula $C \in \mathscr{C}_\delta$, it follows that $C \leq C_\delta^{\mathbf{FN}}$.*

We call δ a *strict diagonal* if $\delta(x) < x$ for all $x \in]0,1[$, and both δ and $\tilde{\delta} : \mathbb{I} \to \mathbb{I}$, defined by

$$\tilde{\delta}(x) = 2x - \delta(x),$$

are strictly increasing. As a consequence, $C_\delta^{\mathbf{FN}}$ is a hairpin copula if and only if δ is a strict diagonal.

Copulas in the form $C_\delta^{\mathbf{FN}}$ are called Fredricks-Nelsen copulas, or diagonal copulas. The theorem above again states that for a given diagonal function δ, there exists at least one copula C such that $\delta_C = \delta$. Moreover, it provides a method to find an example of this kind of copulas.

Example 1.3. Let us consider the strict diagonal $\delta(x) = x^2$ for $x \in \mathbb{I}$. Then $C_\delta^{\mathbf{FN}}$ is a hairpin copula and the corresponding homeomorphism g is given by $g(x) = 1 - \sqrt{1-x^2}$, i.e. $\Gamma^\star(g)$ is the union of two quarter-circles (see [42]).

The question whether for a given homeomorphism g there exists a copula that concentrates it mass in $\Gamma^\star(g)$, must be answered negatively: taking a singular homeomorphism g (that is, $g' = 0$ in a set of measure 1), it is impossible to find such a copula. Indeed, g been singular, there exists a set N_g satisfying $\lambda(\mathrm{N}_g) = 0$ and $\lambda(g(\mathrm{N}_g)) = 1$. Moreover, even for absolutely continuous or convex sections g, the answer remains negative. This is made explicit by the following theorem.

Theorem 1.7 ([51]). *If $a > 1$ and $g(x) = x^a$, then there does not exist $C \in \mathscr{C}^2$ with* $\mathrm{Supp}(C) \subseteq \Gamma^\star(g)$.

The following example shows that hairpin copulas may be properly generalized by shuffles of M in the sense of [20] (also see [55]), so in particular (mutually) completely dependent copulas. Necessary and sufficient conditions for a shuffle of M to be a diagonal copula can be found in [27] and in [42]. In the former article, the following result was proved:

Theorem 1.8 ([27]). *Suppose that δ is a two-dimensional diagonal. Then the diagonal copula $C_\delta^{\mathbf{FN}}$ is a generalized shuffle of M whose support is contained in the graph of a λ-preserving bijection $S : \mathbb{I} \to \mathbb{I}$ fulfilling $S \circ S = \mathrm{id}_{\mathbb{I}}$ if and only if for almost every $x \in \mathbb{I}$ either $\delta'(x) \in \{0,2\}$ or $\delta(x) = x$ holds.*

Example 1.4. Consider the partition $\mathfrak{Z} = \{I_1, J_1, I_2, J_2, \ldots\}$ of $\mathbb{I}$, whereby

$$I_n = \left[1 - 1/2^n, 1 - 1/2^{n+1}\right] \quad \text{and} \quad J_n = \left[1/2^{n+1}, 1/2^n\right]$$

for every $n \in \mathbb{Z}^+$, and the ordinal sum of the copula $C_{\delta_W}^{\mathbf{FN}}$ with respect to $\mathfrak{Z}$, whereby δ_W denotes the diagonal of W. It follows immediately from the construction that $C_{\delta_W}^{\mathbf{FN}}$ is a generalized shuffle of M.

A multidimensional generalization of these results can be found in [14] and [31]. More precisely, we say that an $n-$dimensional diagonal function is a function $\delta : \mathbb{I} \to \mathbb{I}$ satisfying D.1 to D.3 and

D.4' $|\delta(x) - \delta(y)| \leq n|x-y|$ for all $x, y \in \mathbb{I}$.

Theorem 1.9 ([14, 31]). *For a given n-dimensional diagonal function δ, the function*

$$C(u_1, \ldots, u_n) := \frac{1}{n} \sum_{i=1}^{n} \min\left\{ f(u_{\tau^i(1)}), \ldots, f(u_{\tau^i(n-1)}),\, \delta(u_{\tau^i(n)}) \right\} \tag{1.2}$$

for all $(u_1, \ldots, u_n) \in \mathbb{I}^n$, where τ is the permutation in $\{1, \ldots, n\}$ given by $\tau^i(k) := (k+i)$ mod n and $f(x) := \frac{nx-\delta(x)}{n-1}$ for every $x \in \mathbb{I}$, is a copula satisfying $\delta_C = \delta$.

1.4.1.1 Absolutely continuous copulas

All copulas $C_\delta^{\mathbf{FN}}$ are singular. The characterisation of the diagonal functions δ for which there exists an absolutely continuous copula C with diagonal section $\delta = \delta_C$, is a problem that was studied and solved by Durante and Jaworski in [15]. The method is based upon the modification on the copulas $C_\delta^{\mathbf{FN}}$. Denote by $\mathbb{I}(\delta)$ the set of points t in the unit interval such that δ has a derivative $\delta'(t)$. Setting $\sup\{\delta'(t) : t \in \mathbb{I}(\delta)\} = \frac{2}{1+\varepsilon_\delta}$, the next statement is true:

Theorem 1.10 ([15]). *For every $\delta \in \mathscr{D}_2$, and $\alpha \in \left[\frac{1}{2} - \varepsilon_\delta, \frac{1}{2} + \varepsilon_\delta\right]$, the function defined by*

$$K_{\delta,\alpha}(u,v) := \min\{u, v, \alpha\delta(u) + (1-\alpha)\,\delta(v)\},$$

is a copula with diagonal section equal to δ. Moreover, if $\varepsilon_\delta > 0$ then the function

$$C_\delta(u,v) := \int_{\frac{1}{2}-\varepsilon_\delta}^{\frac{1}{2}+\varepsilon_\delta} K_{\delta,\alpha}(u,v)\mathrm{d}\alpha,$$

is an absolutely continuous copula with diagonal section equal to δ.

With the aid of this theorem we can give a characterization of diagonal sections of absolutely continuous copulas with full support.

Theorem 1.11 ([13, 15]).

a. A diagonal function $\delta \in \mathscr{D}_2$ is the diagonal of an absolutely continuous copula if and only if $\lambda(\{t : \delta(t) = t\}) = 0$.
b. Let $\delta \in \mathscr{D}_2$. The following statements are equivalent:

i. δ is the diagonal of an absolutely continuous copula with full support;
ii. $\delta(t) < t$ for every $t \in]0,1[$ and there is no interval J included in $\mathbb{I}$ such that either $\delta' = 0$ on J or $\delta' = 2$ on J.

In [4] appears an alternative method yielding statement a. above; and in [31] is given a multidimensional generalization of this statement.

1.4.2 Bertino copulas

Given a diagonal δ, the *Bertino copula* $C_\delta^{\mathbf{Ber}}$ is defined by (see [29] and the seminal paper [6]):

$$C_\delta^{\mathbf{Ber}}(x,y) := M(x,y) - \min\left\{\hat{\delta}(t) : t \in [\min\{x,y\}, \max\{x,y\}]\right\}, \qquad (1.3)$$

where $\hat{\delta}(t) := t - \delta(t)$. It is well know that $C_\delta^{\mathbf{Ber}}$ is the minimal element in $\mathscr{C}_\delta$ (see [29]). We define the two following functions $l, u : \mathbb{I} \to \mathbb{I}$:

$$\begin{aligned} u(x) &: = \max\left\{y \geq x : \hat{\delta}(t) \geq \hat{\delta}(x) \text{ for all } t \in [x,y]\right\} \\ l(x) &: = \min\left\{y \leq x : \hat{\delta}(t) \geq \hat{\delta}(x) \text{ for all } t \in [y,x]\right\}. \end{aligned} \qquad (1.4)$$

Theorem 1.12. *The support of the Bertino copula* $C_\delta^{\mathbf{Ber}}$ *is contained in the union of the diagonal and the closure of the graph of the measurable function* $S : \mathbb{I} \to \mathbb{I}$, *defined by*

$$S(x) = \begin{cases} u(x), \text{ if } w_\delta(x) > 0 \\ l(x), \text{ if } w_\delta(x) \leq 0 \end{cases} \qquad (1.5)$$

where w_δ *is (a versions of) the* derivative of δ.

Therefore, $C_\delta^{\mathbf{Ber}}$ *is a singular copula.*

Note that the copulas $C_\delta^{\mathbf{Ber}}$ and $C_\delta^{\mathbf{FN}}$ only coincide in case $\delta = \delta_M$ (see [11]).

1.5 Semilinear and conic copulas

We have seen that for any diagonal function δ there exists a copula C with diagonal section $\delta_C = \delta$. Quite a lot of research has been focused on the compatibility of diagonal functions δ with the diagonal sections of copulas that are generated by very special constructions methods, such as the methods based on linear interpolation. In the next subsections we consecutively consider copulas of which certain sections parallel to the borders of the unit square or sections that join one of the points $(0,1)$ and $(1,0)$ to points on the diagonal of the unit square, are linear functions.

1.5.1 Semilinear copulas

We call a function $L : \mathbb{I}^2 \longrightarrow \mathbb{I}$ a *lower semilinear function* if it satisfies the boundary conditions C.1 and C.2 of a copula and if for all $x \in]0,1]$, the mappings

$$h_x : [0,x] \to \mathbb{I}, h_x(t) := L(t,x),$$
$$v_x : [0,x] \to \mathbb{I}, v_x(t) := L(x,t),$$

are linear. Analogously, a function U is called un *upper semilinear* if for all $x \in [0,1[$, the mappings

$$h_x : [x,1] \to \mathbb{I}, h_x(t) := U(t,x),$$
$$v_x : [x,1] \to \mathbb{I}, v_x(t) := U(x,t),$$

are linear.

It is easily verified that given a diagonal function $\delta \in \mathscr{D}_2$, the function $L_\delta : \mathbb{I}^2 \to \mathbb{I}$ defined by

$$L_\delta(x,y) = \begin{cases} \frac{y}{x}\delta_C(x), \ y \leq x \\ \frac{x}{y}\delta_C(y), \ \text{otherwise} \end{cases}$$

where the convention $\frac{0}{0} := 0$ is adopted, is the unique lower semilinear function with diagonal section δ_{L_δ} equal to δ.

The upper semilinear function U_δ with diagonal section δ is defined in an analogous way. Note that every upper semilinear copula U_δ with diagonal section δ is equivalently given by

$$U_\delta(x,y) = x + y - 1 + L_\delta(1-x, 1-y)$$

where L_δ is the lower semilinear copula determined by the diagonal section $\widehat{\delta}(t) = 2t - 1 + \delta(1-t)$. This property allows to restrict the further analysis to just one of these semilinear functions.

The characterization of the diagonal functions that generate lower semilinear copulas is as follows:

Theorem 1.13 ([16]). *Given a diagonal function δ, the lower semilinear function L_δ with diagonal section δ is a copula if and only if the functions $\varphi_\delta, \eta_\delta :]0,1] \to \mathbb{I}$ defined by $\varphi_\delta(x) := \frac{\delta(x)}{x}$, $\eta_\delta(x) := \frac{\delta(x)}{x^2}$ are non-decreasing and non-increasing, respectively.*

Subsequently, the concept of semilinear copulas has been extended in various ways (see [8, 24, 25, 32, 35, 36]).

1.5.2 Biconic copulas

We call a function $F_\delta : \mathbb{I}^2 \to \mathbb{I}$ a *biconic* function if it satisfies the boundary conditions C.1 and C.2 of a copula and if it is linear on segments that join a point on the diagonal of the unit square to one of the corners points $(1,0)$ and $(0,1)$.

It is easily verified that given a diagonal function $\delta \in \mathscr{D}_2$, the function $F_\delta : \mathbb{I}^2 \to \mathbb{I}$ defined by

$$F_\delta(x,y) := \begin{cases} \delta\left(\dfrac{y}{y+1-x}\right)(y+1-x), & x \geq y \\ \delta\left(\dfrac{x}{x+1-y}\right)(x+1-y), & x < y, \end{cases} \tag{1.6}$$

where the convention $\delta\left(\frac{0}{0}\right) := 0$ is adopted. In a closed form,

$$F_\delta(x,y) = \delta\left(\tfrac{x\wedge y}{x\wedge y+1-x\vee y}\right)(x\wedge y+1-x\vee y).$$

The characterization of the diagonal functions generating biconic copulas is as follows:

Theorem 1.14 ([12, 33]). *Given a diagonal function δ, the biconic function F_δ with diagonal section δ is a copula if and only if δ is convex.*

Further studies on these copulas and their generalizations can be found in [26, 34, 37].

1.6 Subdiagonal sections

The study of diagonal sections was generalised to sub-diagonal sections by Quesada-Molina et al. in [46]. We recall basic notions and properties:

a) Given a copula C and $x_0 \in]0,1[$, the *sub-diagonal section* $\delta_{x_0,C}$ of C at x_0 is the function $\delta_{x_0,C} : [0,1-x_0] \to [0,1-x_0]$ defined by $\delta_{x_0,C}(t) = C(x_0+t,t)$.

b) Given $x_0 \in]0,1[$, a *sub-diagonal function* δ_{x_0} is a function $[0,1-x_0] \to [0,1-x_0]$ with the following three properties:

$$\begin{cases} \text{SD.1} \quad \delta_{x_0}(1-x_0) = 1-x_0, \\ \text{SD.2} \quad 0 \leq \delta_{x_0}(t) \leq t \text{ for every } t \in [0,1-x_0], \\ \text{SD.3} \quad 0 \leq \delta_{x_0}(t) - \delta_{x_0}(t') \leq 2(t-t'), \text{ for every } t \text{ and } t' \in [0,1-x_0] \text{ with } t' \leq t. \end{cases}$$

Remark 1.1. If (X,Y) is a pair of random variables distributed according to the copula C, the function $\frac{\delta_{x_0}}{1-x_0}$ is the restriction to $\mathbb{I}$ of the conditional distribution function of $\max(X-x_0,Y)$, given $Y \leq 1-x_0$.

For a fixed sub-diagonal function δ_{x_0}, we denote by $\mathscr{C}_{\delta_{x_0}}$ (resp. $\mathscr{C}^{ac}_{\delta_{x_0}}$) the class of all copulas (resp. absolutely continuous copulas) whose sub-diagonal section is δ_{x_0}.

For the next result we need some of notation:

$$\begin{aligned} m_{x_0}(x,y) &:= \min(x-x_0,y), \\ M_{x_0}(x,y) &:= \max(x-x_0,y), \\ h_{x_0}(x,y) &:= \min\left\{\widehat{\delta}_{x_0}(t) : t \in \left[m_{x_0}(x,y), M_{x_0}(x,y)\right]\right\}. \end{aligned}$$

Theorem 1.15 ([46]). *For every $x_0 \in]0,1[$ and every sub-diagonal δ_{x_0}, there exist a copula $C \in \mathscr{C}_{\delta_{x_0}}$. Moreover, it is possible to find an element in $\mathscr{C}_{\delta_{x_0}}$ being symmetric.*

The set $\mathscr{C}_{\delta_{x_0}}$ has a minimun which is given by the formula

$$B_{\delta_{x_0}}(x,y) := \begin{cases} m_{x_0}(x,y) - h_{x_0(x,y)}, & \textit{if } (x,y) \in [x_0,1] \times [0,1-x_0] \\ W(x,y), & \textit{otherwise.} \end{cases}$$

Theorem 1.16 ([4]). *Let $x_0 \in]0,1[$ and δ_{x_0} be a sub-diagonal function. Then the set $C^{ac}_{\delta_{x_0}}$ is non-empty.*

1.7 Diagonal and opposite diagonal sections

The *opposite diagonal section* ω_C of a copula C is a function $\omega_C \colon \mathbb{I} \to \mathbb{I}$ defined by $\omega_C(t) = C(t, 1-t)$. An *opposite diagonal function* is a function $\omega \colon \mathbb{I} \to \mathbb{I}$ satisfying the following conditions:

$$\begin{cases} \text{W.1} \quad \omega(t) \leq \min(t, 1-t), \text{ for all } t \in \mathbb{I}, \\ \text{W.2} \quad |\omega(v) - \omega(u)| \leq |v-u| \text{ for all } u,v \in \mathbb{I}. \end{cases}$$

Note that there exits a bijection between $\mathscr{D}_2$ and the set of opposite diagonals which is given by $\delta \to \omega_\delta$, with $\omega_\delta(t) := t - \delta(t)$. Also, many of the properties of diagonal functions can be transferred to the opposite diagonal functions because the copula $C'(x,y) := x - C(x, 1-y)$ satisfies $\omega_{C'} = \omega_{\delta_C}$ (see [10]).

We can ask for necessary and sufficient conditions for a diagonal function δ and an opposite diagonal function ω to be compatible, that is, there exists a copula C such that $\delta_C = \delta$ and $\omega_C = \omega$. The first steps in this direction were made in [9], where with the help of cross-copulas the authors have put forward compatibility conditions between δ and ω.

We consider the following assumptions. Let δ be a diagonal function and ω be an opposite diagonal function satisfying:

$$\begin{aligned} &\text{a) } \forall t \in [0,1/2],\ 0 \leq \omega(t) - \delta(t) \text{ and } 0 \leq \omega(1-t) - \delta(t), \\ &\text{b) } \forall t \in [1/2,1],\ \delta(t) - \omega(t) \leq 2t-1 \text{ and } \delta(t) - \omega(1-t) \leq 2t-1, \\ &\text{c) } \forall t,t' \in [0,1/2], t < t' \quad \Rightarrow \\ &\quad \delta(t) + \delta(1-t) - \omega(t) - \omega(1-t) \geq \delta(t') + \delta(1-t') - \omega(t') - \omega(1-t'). \end{aligned} \tag{1.7}$$

From (1.7.a) and (1.7.b), it follows that $\delta\left(\frac{1}{2}\right) = \omega\left(\frac{1}{2}\right)$. Taking this equality into account and putting $t' = \frac{1}{2}$ in (1.7.c), we have that $\delta(t) + \delta(1-t) - \omega(t) - \omega(1-t) \geq 0$, for all $t \in [0,1/2]$.

The solution to this problem was obtained through its reformulation as a linear programming problem. The latter could be realized with the help of an algebraic tool known as Farkas' Lemma (see [49]).

Theorem 1.17 ([2]). *Let δ be a diagonal function and ω be an opposite diagonal function for which (1.7) holds. Then there exists a copula C with diagonal section and opposite diagonal section equal to δ and ω, respectively.*

This allowed to solve two problems posed earlier by Klement and Kolesárová in [39].

Theorem 1.18. *Let C be a copula, and let δ and ω be its diagonal and opposite diagonal sections, respectively. Then C is the unique copula with these sections if and only if*

$$\begin{cases} \delta(x) = \omega(x) \\ \delta(x) = \delta(1-x) - 1 + 2x \\ \omega(x) = \omega(1-x) \end{cases} \tag{1.8}$$

for all $x \in [0, 1/2]$.

Theorem 1.19. *Let δ be a diagonal function and ω be an opposite diagonal function. Then there exists a unique copula having δ and ω as diagonal and opposite diagonal sections, respectively, if and only if δ and ω satisfy conditions (1.7) and (1.8).*

These ideas can be applied to compatibility problems between other types of sections that generalise either diagonal sections or opposite diagonal sections, as can be found in [23].

1.8 Copulas with given quadratic sections

The notation and results we present in this section are those introduced in [45].

A function $(x,y) \to C(x,y)$ is said to have quadratic sections in x iff its intersection with $y = y_0$ is quadratic in x for each $y_0 \in \mathbb{I}$; that is equivalent to say that C has quadratic sections in x if and only if it can be expressed in the form

$$C(x,y) = a(y)x^2 + b(y)x + c(y), \quad (x,y) \in \mathbb{I}^2, \tag{1.9}$$

where a, b, and c are real-valued functions defined on the unit interval $\mathbb{I}$. Function C can be characterised as a copula in terms of these a, b, and c by:

Theorem 1.20 ([45]). *Let a, b, and c are real-valued functions defined on $\mathbb{I}$, and $C : \mathbb{I}^2 \to \mathbb{R}$ be the function defined by the formula (1.9). Then these two statements are equivalent:*

i. C is a copula.
ii. The following conditions are satisfied:

- $b(y) = y - a(y)$ *and* $c(x) = 0$ *for all* $y \in \mathbb{I}$;
- $a(0) = a(1) = 0$;

- *a is 1-Lipschitz; i.e., $|a(s) - a(t)| \leq |s - t|$ for all $s, t \in \mathbb{I}$.*

In such a case, C is absolutely continuous.

Copulas with quadratic sections in y can be obtained by exchanging the variables: C is a copula with quadratic sections in y if and only if its transpose $C^t(x,y) = C(y,x)$ is a copula with quadratic sections in x.

We can write, for the sake of simplicity,

$$C(x,y) = xy + a(y)x(1-x), \quad (x,y) \in \mathbb{I}^2, \tag{1.10}$$

and the characterization of C as a copula is given by

Corollary 1.3. *The function C in equation (1.10) is a copula if and only if a satisfies the three following conditions:*

a. $y \to a(y)$ *is an absolutely continuous function;*
b. $|a'(y)| \leq 1$ *almost everywhere in* $\mathbb{I}$*; and*
c. $|a'(y)| \leq \min\{y, 1-y\}$ *for all* $y \in \mathbb{I}$.

In such a case, C is an absolutely continuous copula.

The next result, a direct consequence of the preceding corollary, is of practical interest rather than the theoretical one for functions C given as above:

Corollary 1.4. *Let a be a continuous function defined on $\mathbb{I}$ such that $a(0) = a(1) = 0$. Suppose that a is a differentiable function with continuous derivative a' satisfying $|a'(y)| \leq 1$. Then C is an absolutely continuous copula.*

Other methods for obtaining copulas in other ways using quadratic constructions either in one variable or in the multivariate case can be found in [43, 47, 48, 56].

1.9 Rectangular Patchwork and Cross Copulas

We conclude with this section which is devoted to the study of the compatibility between horizontal and vertical sections of a copula with the upper and lower bounds of this copula.

Let C be a copula. The horizontal b-section of C is the function $h_{C,b} : \mathbb{I} \to [0,b]$ given by $h_{C,b}(t) := C(t,b)$ and the vertical a-section of C is the function $v_{C,a} : \mathbb{I} \to [0,a]$ given by $v_{C,a}(t) := C(a,t)$.

Let a and b be fixed in $]0,1[$. We denote by $\mathscr{H}_b$ the set of increasing and 1-Lipschitz functions $h : \mathbb{I} \to \mathbb{I}$ such that $h(0) = 0$ and $h(1) = b$ with $\max\{0, t+b-1\} \leq h(t) \leq \min\{t, b\}$. Analogously, we denote by $\mathscr{V}_a$ the set of increasing and 1-Lipschitz functions $v : \mathbb{I} \to \mathbb{I}$ such that $v(0) = 0$ and $v(1) = a$ with $\max\{0, t+a-1\} \leq v(t) \leq \min\{a, t\}$. These functions will be called as horizontal b-sections and vetical a-sections, respectively.

It is easy to prove that for each $b \in]0,1[$ and $C \in \mathscr{C}^2$ then $h_{C,b} \in \mathscr{H}_b$. The same is true for vertical sections. By application of Proposition 1.1, we obtain the next result:

Theorem 1.21 ([40]). *For each $h \in \mathscr{H}_b$ (resp., $v \in \mathscr{V}_a$), there exists a copula C such that $h_{C,b} = h$ (resp. $v_{C,a} = v$).*

Actually, we can say more: there exists infinitely many copulas satisfying that property. To see this property and other results related with maxima and minima in the set $\mathscr{H}_b$, we need a new tool which we describe in what follows.

Rectangular patchworks appear in [19] as an unification of different ideas that were implicit in the literature (see [7, 16, 18, 22, 39, 40, 44, 46]). We introduce some notation to describe them.

Let $a_1, a_2, b_1, b_2, c_1, c_2$ be in $\mathbb{I}$ with $a_1 < a_2, b_1 < b_2$ and $c_1 \leq c_2$. Given a function $F : [a_1, a_2] \times [b_1, b_2] \to [c_1, c_2]$, the margins of F are the functions $h_{b_1}^F, h_{b_2}^F, v_{a_1}^F$ and $v_{a_2}^F$ defined by

$$\begin{aligned} h_{b_1}^F, h_{b_2}^F &: [a_1, a_2] \to [c_1, c_2], \; h_{b_1}^F(x) := F(x, b_1); \; h_{b_2}^F(x) := F(x, b_2), \\ v_{a_1}^F, v_{a_2}^F &: [b_1, b_2] \to [c_1, c_2], \; v_{a_1}^F(y) := F(a_1, y); \; v_{a_2}^F(y) := F(a_2, y). \end{aligned}$$

Theorem 1.22 ([19]). *Be given a family $\{C_i\}_{i \in \mathscr{I}}$ of copulas and a family of rectangles $\left\{\mathrm{R}_i = \left[a_1^i, a_2^i\right] \times \left[b_1^i, b_2^i\right]\right\}_{i \in \mathscr{I}}$ in $\mathbb{I}^2$ such that $\mathrm{R}_i \cap \mathrm{R}_j \subseteq \partial \mathrm{R}_i \cap \partial \mathrm{R}_j$, for every $i = j$. Let C be a copula and put $\lambda_i = V_C(\mathrm{R}_i)$. Let $\widetilde{C} : \mathbb{I}^2 \to \mathbb{I}$ be defined, for every $x, y \in \mathbb{I}$, by*

$$\widetilde{C}(x,y) = \begin{cases} \lambda_i C_i \left(\frac{V_C\left(\left[a_1^i, x\right] \times \left[b_1^i, b_2^i\right]\right)}{\lambda_i}, \frac{V_C\left(\left[a_1^i, a_2^i\right] \times \left[b_1^i, y\right]\right)}{\lambda_i} \right) & (x,y) \in \mathrm{R}_i, \lambda_i \neq 0 \\ \quad + h_{b_1^i}^C(x) + v_{a_1^i}^C(y) - h_{b_1^i}^C(a_1^i), & \\ C(x,y), & \text{otherwise} \end{cases} \tag{1.11}$$

Then $\widetilde{C}$ is a copula.

We use the notation $\widetilde{C} = (\langle \mathrm{R}_i, C_i \rangle)_{i \in \mathscr{I}}^C$ for indicating the *rectangular patchwork* of $(\langle \mathrm{R}_i, C_i \rangle)_{i \in \mathscr{I}}$ into the copula C.

In case of a finite number of rectangles whose union fills the unit square it is unnecessary to know the copula C since $\widetilde{C}$ is completely determined by the horizontal and vertical sections $h_{b_1^i}^C$, $h_{b_2^i}^C$, $v_{a_1^i}^C$ and $v_{a_2^i}^C$, and the copulas C_i,. Therefore, we can state the following result:

Corollary 1.5. *Let $0 = a_0 < a_1 < \cdots < a_n = 1$, $0 = b_0 < b_1 < \cdots < b_m = 1$, $\{v_i \in \mathscr{V}_{a_i} : i \in \{0, 1, \ldots, n\}\}$, and $\left\{h_j \in \mathscr{H}_{b_j} : j \in \{0, 1, \ldots, m\}\right\}$ satisfying the following compatibility conditions $v_i(b_j) = h_j(a_i)$ and also the 2-increasingness:*

a. $h_{j+1}(x) + h_j(a_i) - h_j(x) - h_{j+1}(a_i)$ non decreasing when $x \in [a_i, a_{i+1}]$, $\forall i \in \{0, 1, \ldots, n-1\}$, $\forall j \in \{1, \ldots, m\}$.

b. $v_{i+1}(x)+v_i(b_j)-v_i(x)-v_{i+1}(b_j)\geq 0$ *non decreasing when* $x\in\left[b_j,b_{j+1}\right], \forall j\in \{0,1,\ldots,m-1\}, \forall i\in\{1,\ldots,n\}.$

Let $\left\{C_{i,j}\in\mathscr{C}^2 : j\in\{1,\ldots,m\}, i\in\{1,\ldots,n\}\right\}$ *a family of copulas and set* $\lambda_{ij}=h_{j+1}(a_{i+1})+h_j(a_i)-h_j(a_{i+1})-h_{j+1}(a_i)$*. Let* $C:\mathbb{I}^2\to\mathbb{I}$ *be defined, for every* $(x,y)\in\mathbb{I}^2$*, by*

$$C(x,y)=\lambda_{ij}C_{ij}\left(\frac{h_{j+1}(x)+h_j(a_i)-h_j(x)-h_{j+1}(a_i)}{\lambda_{ij}},\frac{v_{i+1}(x)+v_i(b_j)-v_i(x)-v_{i+1}(b_j)}{\lambda_{ij}}\right)+h_j(x)+v_i^C(y)-h_j(a_i)$$

Then C is a copula with horizontal b_j*-sections and vertical* a_i*-sections which coincide with* h_j *and* v_i*, respectively.*

The copula C obtained above is represented by $\left\langle\left\{h_j\right\},\left\{v_i\right\},\left\{C_{i,j}\right\}\right\rangle$, for the sake of simplicity. Observe that $h_0=v_0=0$ and $h_1=v_1=\mathrm{id}$, and therefore it is unnecessary to include them in the above notation.

With the aid of this result we can describe the compatibility between elements of $\mathscr{H}_b$ and $\mathscr{V}_a$. Given $h\in\mathscr{H}_b$ and $v\in\mathscr{V}_a$, we represent by $\mathscr{C}_{h,v}$ the set of copulas $\left\{C\in\mathscr{C}^2 : h_b^C=h, v_a^C=v\right\}$. These are the Cross Copulas cited in [17]. Setting $C=\langle\{h_b\},\{v_a\},\{C_{11},C_{12},C_{21},C_{22}\}\rangle$, without restrictions on the copulas C_{ij} it follows:

Theorem 1.23 ([17]). *Let h and v be given in* $\mathscr{H}_b$ *and* $\mathscr{V}_a$*, respectively, with* $h(a)=v(b)=c$*, and* $\max\{a+b-1,0\}<c<\min\{a,b\}$*. Then,* $\mathscr{C}_{h,v}$ *is a non empty set.*

The conditions of Theorem 1.23 can be weakened to $\max\{a+b-1,0\}\leq c\leq\min\{a,b\}$. Although condition "$a,b\in\,]0,1[$" is introduced, it is possible to consider a or b to be zero. In this case, the results obtained for the case of horizontal b-sections or vertical a-sections are retrieved(see [40]).

Taking copulas C_{ij} to be equal to M and W we obtain:

Theorem 1.24 ([17]). *Let h be in* $\mathscr{H}_b$ *and v be in* $\mathscr{V}_a$ *with* $h(a)=v(b)=c$*. Then the set* $\mathscr{C}_{h,v}$ *has maximal a element that is* $\langle\{h_b\},\{v_a\},\{M,M,M,M\}\rangle$ *and minimal element given by* $\langle\{h_b\},\{v_a\},\{W,W,W,W\}\rangle$*.*

Acknowledgements This work was supported by the Ministerio de Economía y Competitividad (Spain) under Research Project No. MTM2014-60594-P and partially supported by FEDER.

References

1. Alsina, C.; Nelsen, R.B. and Schweizer, B.: *On the characterization of a class of binary operations on distribution functions*. Statist. Probab. Lett. **17** (2) (1993) 85-89
2. Amo, E. de, De Meyer, H., Díaz Carrillo, M. and Fernández Sánchez, J.: *Characterizations of copulas with given diagonal and opposite diagonal sections*, Fuzzy Sets and Systems **284** (2016) 63-77

3. Amo, E. de, Díaz Carrillo, M. Durante, F. and Fernández Sánchez, J.: *Extension of subcopulas*, J. Math. Anal. Appl. (2017), http://dx.doi.org/10.1016/j.jmaa.2017.02.061
4. Amo, E. de, Díaz Carrillo, M. and Fernández Sánchez, J.: *Absolutely continuous copulas with given sub-diagonal sections*, Fuzzy Sets and Systems **228** (2013) 105-113
5. Amo, E. de, Díaz Carrillo, M. and Fernández Sánchez, J.: *Characterizations of all copulas with non-continuous random variables*, Fuzzy Sets and Systems **191** (2012) 103-112
6. Bertino, S.: *Sulla dissomiglianza tra mutabili cicliche*, (Italian) Metron 35 (1977), no. 1-2, 53-88 (1979).
7. De Baets, B. and De Meyer, H.: *Orthogonal grid constructions of copulas*, IEEE Trans. Fuzzy Syst. **15** (6) (2007) 1053-1062
8. De Baets, B., De Meyer, H. and Mesiar, R.: *Asymmetric semilinear copulas*, Kybernetika (Prague) **43** (2) (2007) 221-233
9. De Baets, B., De Meyer, H. and Úbeda-Flores, M.: *Constructing copulas with given diagonal and opposite diagonal sections*, Commun. Stat. Theory Methods **40** (5) (2011) 828-843
10. De Baets, B., De Meyer, H. and Úbeda-Flores, M.: *Opposite diagonal sections of quasi-copulas and copulas*, Int. J. Uncertain. Fuzziness Knowl.-Based Syst. **17** (4) (2009) 481-490
11. Durante, F. and Fernández Sánchez, J.: *On the classes of copulas and quasi-copulas with a given diagonal section*, Int. J. Uncertain. Fuzziness Knowl.-Based Syst. **19** (1) (2011) 1-10
12. Durante, F. and Fernández-Sánchez, J.: *A note on biconic copulas*, Kybernetika (Prague) **47** (4) (2011) 532-540
13. Durante, F., Fernández Sánchez, J. and Pappadà, R.: *Copulas, diagonals, and tail dependence*, Fuzzy Sets and Systems **264** (2015) 22-41
14. Durante, F., Fernández Sánchez, J. and Trustchnig, W.: *Multivariate copulas with hairpin support*, J. Multivar. Anal. **130** (2014) 323-324
15. Durante, F. and Jaworski, P.: *Absolutely continuous copulas with given diagonal sections*, Commun. Statist. Theory Methods **37** (18) (2008) 2924-2942
16. Durante, F., Kolesárová, A., Mesiar, R. and Sempi, C.: *Semilinear copulas*, Fuzzy Sets and Systems **159** (2008) 63-76
17. Durante, F., Kolesárová, A., Mesiar, R. and Sempi, C.: *Copulas with given values on a horizontal and a vertical section*, Kybernetika (Prague) **43** (2) (2007) 209-220
18. Durante, F., Mesiar, R., and Sempi, C.: *On a family of copulas constructed from the diagonal section*, Soft. Comput. **10** (6) (2006) 490-494
19. Durante, F., Saminger-Platz, S., and Sarkoci, P.: *Rectangular patchwork for bivariate copulas and tail dependence*, Comm. Statist. Theory Methods **38** (13-15) (2009) 2515-2527
20. Durante, F., Sarkoci, P., and Sempi, C.: *Shuffles of copulas*, J. Math. Anal. Appl. **352** (2) (2009) 914-921
21. Durante, F. and Sempi, C.: *Principles of Copula Theory*, (2016) CRC/Chapman & Hall, Boca Raton, FL
22. Erdely, A. and González-Barrios, J.M.: *On the construction of families of absolutely continuous copulas with given restrictions*, Commun. Stat. Theory Methods **35** (4-6) (2006) 649-659
23. Fernández Sánchez, J. and Úbeda Flores, M.: *Copulas with given track and opposite track sections*, Fuzzy Sets and Systems, (accepted for publication)
24. Fernández Sánchez, J. and Úbeda-Flores, M.: *On copulas that generalize semilinear copulas*, Kybernetika (Prague) **48** (5) (2012) 968-976
25. Fernández Sánchez, J. and Úbeda-Flores, M.: *Semi-polynomial copulas*, J. Nonparametr. Stat. **26** (1) (2014) 129-140
26. Fernández Sánchez, J. and Ubeda Flores, M. *The distribution of the probability mass of biconic copulas*, Information Sciences, **329** (2016) 177-183
27. Fernández Sánchez, J. and Trutschnig, W.: *Some members of the class of (quasi-)copulas with given diagonal from the Markov kernel perspective*, Comm. Statist. Theory Methods **45** (5) (2016) 1508-1526
28. Fredricks, G.A. and Nelsen, R.B.: *Copulas constructed from diagonal sections*, in: J. Beneš and J. Štěpán (Eds.) Distributions with given Marginals and Moment Problems (Prague 1996), Kluver Acad. Publ., Dordrecht, 1997, 129-136

29. Fredricks, G.A. and Nelsen, R.B.: *The Bertino family of copulas*, Distributions with given marginals and statistical modelling, Kluwer Acad. Publ., Dordrecht, (2002) 81-91
30. Genest, C., Quesada-Molina, J.J., Rodríguez-Lallena, J.A. and Sempi, C.: *A characterization of quasi-copulas*, J. Multivariate Anal. **69** (1999) 193-205
31. Jaworski, P.: *On copulas and their diagonals*, Inform. Sci. **179** (1) (2009) 2863-2871
32. Jwaid, T., De Baets, B. and De Meyer, H.: *Orbital semilinear copulas*, Kybernetika (Prague) **45** (6) (2009) 1012-1029
33. Jwaid, T., De Baets, B. and De Meyer, H.: *Biconic aggregation functions*, Inform. Sci. **187** (2012) 129-150
34. Jwaid, T., De Baets, B. and De Meyer, H.: *Ortholinear and paralinear semi-copulas*, Fuzzy Sets and Systems **252** (2014) 76-98
35. Jwaid, J., De Meyer, H. and De Baets, B.: *Semiquadratic copulas based on horizontal and vertical interpolation*, Fuzzy Sets and Systems **264** (2015) 3-21
36. Jwaid, J., De Meyer, H. and De Baets, B.: *Lower semiquadratic copulas with a given diagonal section*, J. Stat. Plan. Inf. **143** (8) (2013) 1355-1370
37. Jwaid, J., De Meyer, H., Mesiar, R. and De Baets, B.: *The role of generalized convexity in conic copula constructions*, J. Math. Anal. Appl. **425** (2015) 864-885
38. Kamiński, A., Mikusiński, P., Sherwood, H. and Taylor, M.D.: *Doubly Stochastic Measures, Topology, and Latticework Hairpin*, J. Math. Anal. Appl. **152** (1990) 252-268
39. Klement, E.P. and Kolesárová, A.: *Intervals of 1-Lipschitz aggregation operators, quasi-copulas and copulas with given affine section*, Monaths. Maths. **152** (2) (2007) 151-167
40. Klement, E.P., Kolesárová, A., Mesiar, R. and Sempi, C.: *Copulas constructed from horizontal sections*, Commun. Stat. Theory Methods **36** (13-16) (2007) 2901-2911
41. Nelsen, R.B.: *Introduction to Copulas*, (2nd ed.) (2006) Springer Series in Statistics Springer, New York
42. Nelsen, R.B. and Fredricks, G.A.: *Diagonal Copulas*, in: J. Beneš and J. Štěpán (Eds.) Distributions with given Marginals and Moment Problems (Prague 1996), Kluver Acad. Publ., Dordrecht (1997) 121-128
43. Nelsen, R.B., Quesada-Molina, J.J. and Rodríguez-Lallena, J.A.: *Bivariate copulas with cubic sections*, J. Nonparametr. Statist. **7** (3) (1997) 205-220
44. Nelsen, R.B., Quesada-Molina, J.J., Rodríguez-Lallena, J.A. and Úbeda-Flores, M.: *On the construction of copulas and quasi-copulas with given diagonal sections*, Insurance Math. Econ. **42** (2) (2008) 473-483
45. Quesada-Molina, J.J. and Rodríguez-Lallena, J.A.: *Bivariate copulas with cuadratic sections*, J. Nonparametr. Stat. **5** (4) (1995) 323-337
46. Quesada-Molina, J.J., Saminger-Platz, S. and Sempi, C.: *Quasi-copulas with a given sub-diagonal sections*, Nonlinear Anal. **69** (12) (2008) 4654-4673
47. Rodríguez-Lallena, J.A.: *A class of copulas with piecewise linear horizontal sections*, J. Stat. Plan. Infer. **139** (11) (2009) 3908-3920
48. Rodríguez-Lallena, J.A. and Úbeda Flores, M.: *Multivariate copulas with quadratic sections in one variable*, Metrika **72** (2009) 331-349
49. Schrijver, A.: *Theory of Linear and Integer Programming*, (1998) John Wiley & Sons, New York
50. Schweizer, B. and Sklar, A., *Operations on distribution functions not derivable from operations on random variables*, Stud. Math. **52** (1974) 43-52
51. Seethoff, T.L. and Shiflett, R.C.: *Doubly stochastic measures with prescribed support*, Z. Wahrscheinlichkeitstheorie und Verw. Gebiete **41** (4) 1977/78 283-288
52. Sempi, C.: *Quasi-copulas: a brief survey*. Appeared in this monograph
53. Sherwood, H. and Taylor, M.D.: *Doubly Stochastic Measures with Hairpin Support*, Prob. Th. Rel. Fields **78** (1988) 617-626
54. Sklar, A., *Fonctions de répartition à n dimensions et leurs marges*, Publ. Inst. Statist. Univ. Paris **8** (1959) 229-231
55. Trutschnig, W. and Fernández Sánchez, J.: *Some results on shuffles of two-dimensional copulas*, J. Statist. Plann. Inference **143** (2) (2013) 251-260

56. Úbeda Flores, M.: *Multivariate copulas with given cubic sections in one variable*, J. Nonparametr. Stat. **20** (1) (2008) 91-98
57. Úbeda Flores, M. and Fernández Sánchez, J.: Sklar's theorem: The cornerstone of the Theory of Copulas. Appeared in this monograph

Chapter 2
The Gumbel-Marshall-Olkin distribution

Umberto Cherubini and Sabrina Mulinacci

Abstract In this paper we introduce a generalization of the Marshall-Olkin distribution that allows for some dependence among the shock arrival times while it preserves exponentially distributed observed lifetimes: these features make the resulting distribution well suited for credit risk applications. The main result of the paper is that the only Archimedean dependence structure consistent with these requirements is the Gumbel one.

2.1 Introduction

During the financial crisis, much of the literature devoted to credit risk has naturally turned to look at the Marshall-Olkin model (see Marshall and Olkin, 1967) as an important tool to represent and study systemic crises. From the point of view of credit risk analysis, in fact, the Marshall-Olkin model has many advantages and one main shortcoming.

The main advantages make the model perfectly well-suited to address a systemic crisis. First, there are unobserved shocks, some of which are unique to each and every individual and some others which are common to a subset of individuals. Second, these common shocks allow for simultaneous default of the elements in the cluster defined as the subset of individuals exposed to the same common factor. This feature cannot be captured by the typical multivariate approach that has been applied to credit risk, that is multivariate distributions obtained using absolutely continuous copulas. This set of copula functions, and the corresponding multivari-

Umberto Cherubini
Department of Statistics, University of Bologna, Via delle Belle Arti 41, 40126 Bologna, Italy
e-mail: umberto.cherubini@unibo.it

Sabrina Mulinacci
Department of Statistics, University of Bologna, Via delle Belle Arti 41, 40126 Bologna, Italy
e-mail: sabrina.mulinacci@unibo.it

M. Úbeda Flores et al. (eds.), *Copulas and Dependence Models with Applications*,
DOI 10.1007/978-3-319-64221-5_2

ate distributions, structurally lead to underestimate the joint default of a number of individuals in a short period of time, which is exactly what happens in a systemic crisis. The third beauty of the Marshall-Olkin model is to preserve the exponential marginal distribution for the observed default times (see also the seminal paper of Gumbel, 1960). This is particularly relevant for credit risk models in which the constant intensity structure of the default event is still considered an important working assumption that simplifies very much the calibration of the model. Indeed, while for univariate models it may be worthwhile to model the intensity of default as a process, this makes the model quite difficult to address and calibrate in a multivariate setting, so that marginal exponential distributions are the standard choice in most of the multivariate analyses of credit risk. The few exceptions to this standard, such as to jointly model self-exciting shocks to the intensities, are not able yet to deal with the singularity issue, that is the simultaneous default of a number of issuers.

As for the flaw of the Marshall-Olkin model, the main issue has to do with the fact that the unobserved shocks are independent. This assumption is not realistic for credit risk applications, particularly with respect to banking crises, in which any idiosyncratic shock can turn into a crisis of the banking system as a whole. Assuming a dependence structure among the unobserved components has important effects on the analysis of dependence among observed default events, both in terms of measurement and interpretation. As for measurement, it is well known that in the Marshall-Olkin model the dependence structure arises from the sensitivity of each obligor to a systemic shock. Actually, it can be easily proved that the Kendall's tau between the systemic shock and the default time of an individual is simply the ratio between the intensity of the systemic shock and the marginal default intensity. It is clear that this dependence relationship should be made stronger by the assumption that the unobserved components are also dependent in the first place. As for the interpretation of dependence, it is clear that the dependence due to sensitivity to the systemic shock is very different from dependence due to the fact that the systemic shock could be actually triggered by the obligor.

A question is how to extend the Marshall-Olkin model keeping all its advantages, including marginal exponential distributions of default times, while introducing a dependence structure among the unobserved components. For this purpose, in this paper we introduce a multivariate model to relax the assumption of independent shocks arrival times, whose dependence structure is assumed of Archimedean type, while preserving the main features of the Marshall-Olkin distribution: first, the intensity of each arrival time of a shock is identified by a nonnegative parameter; second, the observed lifetimes are exponentially distributed. The main result will be that the only Archimedean dependence structure consistent with this extension of the Marshall-Olkin structure is given by a Gumbel copula. For this reason we call this distribution Gumbel-Marshall-Olkin.

The plan of the paper is as follows. In Section 2 we review the main literature connected to this paper. In section 3 we will lay out the main assumptions of the model. In section 4 we will derive the main result of the paper. Section 5 concludes.

2.2 Literature review

There is a wide literature concerning extensions of the Multivariate Marshall-Olkin distribution. Extensions of the Marshall-Olkin distribution have been considered in Li (2009), where the so-called scale-mixtures of the Marshall-Olkin distributions are introduced: such distributions are obtained as a scale-mixture of a Marshall-Olkin distribution. The approach of obtaining generalizations of distributions reinforcing the dependence through a scale mixing technique is well-established in literature and its application for the construction of extensions of the Marshall-Olkin distribution is investigated in Mulinacci (2015). Scale-mixture of the Marshall-Olkin distribution have also been considered in Mai et al. (2013) and in Bernhart et al. (2013) in the exchangeable case which is applied to the CDO pricing problem. In all these cases the resulting dependence structure of the shocks arrival times is of Archimedean type when the generator is the Laplace transform of a positive random variable, since this kind of copulas are obtained through a mixing technique. A more general model with a dependence structure of Archimedean type with general generator is considered in Mulinacci (2017) while the approach of combining Archimedean copulas with extreme value copulas is extended in Charpentier et al. (2014). The multivariate distribution studied in this paper is actually a special case in that general setting.

Concerning credit risk applications, the Marshall-Olkin distribution and its extensions have become extremely relevant in the aftermath of the crisis, because of their singularity property according to which events of simultaneous defaults have positive probability mass. The use of common shocks in the general case of credit risk has been applied in Giesecke (2003) and Lindskog and McNeil (2003) while pricing and hedging applications to credit derivatives are addressed in Elouerkhaoui (2007) and Mai and Scherer (2009). This singularity feature is particularly important in the analysis of banking crises, in which simultaneous defaults are a practical matter, and are the main reason of worries for the regulatory bodies. In this line of research, Baglioni and Cherubini (2013a) applied a standard Marshall-Olkin model to estimate the actuarial value of the liability to be faced by public finance, due to a systemic banking crisis. Baglioni and Cherubini (2013b) extend the research to investigate the relationship between credit risk of the banking system and sovereign credit risk. Cherubini and Mulinacci (2016) extend the analysis to address the issue of dependence among idiosyncratic and systemic shocks, in a model in which every cluster is characterized by a single common shock.

The results presented in this paper are quite natural generalizations of those showed in Cherubini and Mulinacci (2016) where the only shocks considered are those causing the end of a single lifetime and that causing the simultaneous end of all the lifetimes considered: in that case the setting is simplified because of the application considered and the estimation tractability of the model. Here, instead, the Gumbel-Marshall-Olkin distribution is derived in full generality for what concerns the structure of the unobserved shocks. As for the literature on Marshall-Olkin extensions, the distributional model considered in this paper is a particular case of the extensions of the Marshall-Olkin distribution obtained through the mixing technique mentioned above, where it can be characterized as the unique one satisfying the requirements that preserve the Marshall-Olkin structure.

2.3 Model assumptions

Let us consider the classical construction of Marshall-Olkin distributions (see Marshall and Olkin, 1967).

Let $(\Omega,\mathscr{F},\mathbb{P})$ be a probability space. If $d\geq 2$, we consider the class of subsets of $\{1,\ldots,d\}$ given by $\mathscr{P}=\{S\subset\{1,\ldots,d\}:S\neq\emptyset\}$ and $\mathbf{X}=\{X_S\}_{S\in\mathscr{P}}$ a collection of independent and exponentially distributed random variables with intensity $\lambda_S\geq 0$ representing the arrival times of some unobservable shocks, that is, if $\mathbf{x}=(x_S)_{S\in\mathscr{P}}$, $x_S\geq 0$ for all $S\in\mathscr{P}$

$$\bar{F}_{\mathbf{X}}(\mathbf{x})=\exp\left(-\sum_{S\in\mathscr{P}}\lambda_S x_S\right).$$

We denote with T_j, $j=1,\ldots,d$, the j-th observable lifetime whose end is caused by the first arrival time among those shocks X_S with $j\in S$, that is

$$T_j=\min_{S:j\in S}X_S.$$

In this construction X_S represents the arrival time of a shock causing the simultaneous end of all lifetimes T_j for which $j\in S$: the case $\lambda_S=0$ is allowed in order to take into account the case in which the occurrence of the simultaneous default of a specific subset of the defaultable entities has probability zero.

The survival distribution of $\mathbf{T}=(T_1,\ldots,T_d)$ defines the so called *Multivariate Marshall-Olkin distribution* whose survival version, for $\mathbf{t}=(t_1,\ldots,t_d)$ with $t_j\geq 0$ for all $j=1,\ldots,d$ is given by

$$\bar{F}(\mathbf{t})=\exp\left(-\sum_{S\in\mathscr{P}}\lambda_S\max_{j\in S}\{t_j\}\right).$$

We assume

$$\bar{\lambda}_j=\sum_{S:j\in S}\lambda_S>0,\text{ for }j=1,\ldots,d \tag{2.1}$$

and the corresponding survival marginal distributions are

$$\bar{F}_{T_j}(t_j)=\exp(-\bar{\lambda}_j t_j)$$

while the corresponding survival copula (see Li (2008)) is

$$\hat{C}(\mathbf{u})=\prod_{S\in\mathscr{P}}\min_{j\in S}u_j^{\alpha_{S,j}}$$

where

$$\alpha_{S,j} = \lambda_S / \bar{\lambda}_j, \text{ with } j \in S. \tag{2.2}$$

In this paper we consider a generalization of the Marshall-Olkin model assuming a more general distribution for the unobserved shock arrival times **X**. The model is built starting from two basic assumptions.

Assumption 2.3.1. *The joint survival distribution of the shocks arrival times* **X** *is, for* $\mathbf{x} = (x_S)_{S\in\mathscr{P}}$, $x_S \geq 0$ *for all* $S \in \mathscr{P}$,

$$\bar{F}(\mathbf{x}) = \psi\left(\sum_{S\in\mathscr{P}} \lambda_S H(x_S)\right)$$

with $\lambda_S \geq 0$ *and* $\sum_{S:j\in S} \lambda_S > 0$, *for* $j = 1,\dots,d$ *(as in (2.1)) and where*

- $\psi : [0,+\infty) \to [0,1]$ *is the Laplace transform of a positive random variable*
- $H : [0,+\infty) \to [0,+\infty)$ *is strictly increasing with* $H(0) = 0$ *and* $\lim_{x\to+\infty} H(x) = +\infty$.

Notice that in case $\psi(z) = e^{-z}$ and $H(x) = x$ we recover the Marshall-Olkin setting.

As a consequence of Assumption 2.3.1 we have that the survival marginal distributions of the occurrence times X_S, for $S \in \mathscr{P}$, of the unobserved shocks are of same type

$$\bar{F}_S(x) = \psi(\lambda_S H(x))$$

differing only for the intensity parameters λ_S. Notice that since ψ is the Laplace transform of a positive random variable, this is always a survival distribution function for $x \geq 0$.

As for the survival dependence structure induced by Assumption 2.3.1 we trivially recognize that it is of Archimedean type, being

$$\hat{C}(\mathbf{u}) = \psi\left(\sum_{S\in\mathscr{P}} \psi^{-1}(u_S)\right), \mathbf{u} = (u_S)_{S\in\mathscr{P}}, u_S \in [0,1].$$

It is well known that the requirement for ψ to be a Laplace transform is restrictive, since Archimedean copulas are defined for more general classes of generators ψ (see McNeil and Nešlehová, 2009).

While with Assumption 2.3.1 we allow to deal with more general distributions than that underlying the Marshall-Olkin construction, as for the marginal distribution of the observed lifetimes, we still require that they are of exponential type that is:

Assumption 2.3.2. *Default times are marginally exponentially distributed*

$$\bar{F}_{T_j}(x) = \exp(-\mu_j x), \; j = 1,\dots,d$$

where

$$\mu_j = f(\lambda_S : j \in S)$$

denotes the intensity parameter. We also assume that f is a strictly positive differentiable function, strictly increasing with respect to each argument.

2.4 The Gumbel-Marshall-Olkin distribution

Under Assumption 2.3.1 we have that the joint survival distribution function of the random vector $\mathbf{T} = (T_1, \ldots, T_d)$ is

$$\bar{F}_{\mathbf{T}}(\mathbf{t}) = \psi\left(\sum_{S\in\mathscr{P}} \lambda_S H\left(\max_{j\in S}\{t_j\}\right)\right) \tag{2.3}$$

for $\mathbf{t} = (t_1, \ldots, t_d) \in [0, +\infty)^d$. As a consequence, the marginal survival distribution functions are

$$\bar{F}_{T_j}(t) = \psi\left(\bar{\lambda}_j H(t)\right), t \in [0, +\infty) \tag{2.4}$$

where $\bar{\lambda}_j$ is defined in (2.1) and, applying Sklar's theorem, the associated survival copula is

$$\hat{C}(u_1, \ldots, u_d) = \psi\left(\sum_{S\in\mathscr{P}} \max_{j\in S}\left\{\alpha_{S,j}\, \psi^{-1}(u_j)\right\}\right) \tag{2.5}$$

where the $\alpha_{S,j}$'s are defined in (2.2).

In next Proposition, that is the main result of the paper, we show that, in our setting, the requirement of exponentially distributed observed lifetimes is equivalent to restrict to an Archimedean generator of Gumbel type. This is in line with the classical result in Genest and Rivest (1989) according to which the Gumbel one is the only Archimedean copula which is also extreme value. Of course it is the requirement that the observed lifetimes are exponentially distributed that leads to a copula of the extreme value class, according to a very well known result (see section 6.2 in Joe, 1997).

Proposition 2.1. *Under Assumption 2.3.1, Assumption 2.3.2 is equivalent to $\psi(x) = e^{-x^{\frac{1}{\theta}}}$ and $H(x) = \beta x^\theta$, for $\theta \geq 1$ and $\beta > 0$.*

Proof. By (2.4), Assumption 2.3.2 is equivalent to

$$H(x) = \frac{\psi^{-1}\left(e^{-f(\lambda_S : j\in S)x}\right)}{\bar{\lambda}_j}, \text{ for } j = 1, \ldots d. \tag{2.6}$$

Since H is given independently of every possible set of parameters λ_S, $S \in \mathscr{P}$, setting $\psi^{-1} = \phi$, necessarily

$$\frac{\partial}{\partial \lambda_S}\left(\frac{\phi\left(e^{-f(\lambda_S:j\in S)x}\right)}{\bar{\lambda}_j}\right)=0,\ \forall x\geq 0 \text{ and } \forall \lambda_S \text{ such that } \sum_{S:j\in S}\lambda_S>0.$$

It follows, $\forall \lambda_S$, such that $\sum_{S:j\in S}\lambda_S>0$ and $\forall x\geq 0$,

$$-\phi'\left(e^{-f(\lambda_S:j\in S)x}\right)e^{-f(\lambda_S:j\in S)x}\frac{\partial}{\partial \lambda_S}f(\lambda_S:j\in S)x\bar{\lambda}_j-\phi\left(e^{-f(\lambda_S:j\in S)x}\right)=0.$$

Setting $z=e^{-f(\lambda_S:j\in S)x}$,

$$\frac{\phi'(z)}{\phi(z)}z\log z=\frac{f(\lambda_S:j\in S)}{\frac{\partial}{\partial \lambda_S}f(\lambda_S:j\in S)\bar{\lambda}_j}$$

for all λ_S such that $\sum_{S:j\in S}\lambda_S>0$ and for all $z\in(0,1)$. It follows that there exists a constant $\theta>0$ such that

$$\frac{\phi'(z)}{\phi(z)}z\log z=\theta \qquad \text{and} \qquad \frac{f(\lambda_S:j\in S)}{\frac{\partial}{\partial \lambda_S}f(\lambda_S:j\in S)\bar{\lambda}_j}=\theta$$

from which

$$\phi(z)=\gamma(-\log z)^{\theta}$$

where $\gamma>0$ and

$$f(\lambda_S:j\in S)=\delta\,\bar{\lambda}_j^{\frac{1}{\theta}}$$

with $\delta>0$. From (2.6) it follows that

$$H(x)=\gamma\delta^{\theta}x^{\theta}.$$

Since it is a known fact that Archimedean generators having proportional inverse, generate the same copula function, setting $\gamma=1$, we recover $\psi(x)=e^{-x^{\frac{1}{\theta}}}$, which is an Archimedean generator for $\theta\geq 1$, and $H(x)=\delta^{\theta}x^{\theta}$.

As a consequence of the above result, under Assumptions 2.3.1 and 2.3.2, the joint survival distribution function of the unobservable shocks arrival times $\mathbf{X}$ is of type

$$\bar{F}(\mathbf{x})=e^{-\left[\sum_{S\in\mathscr{P}}\lambda_S\delta^{\theta}x_S^{\theta}\right]^{\frac{1}{\theta}}}$$

and the corresponding marginal survival distributions are of exponential type, $\bar{F}_{X_S}(x)=e^{-(\delta^{\theta}\lambda_S)^{\frac{1}{\theta}}x}$ with $\theta\geq 1$ and $\delta>0$.

Since different δ's translate in differently proportional parameters λ_S, we normalize the setting assuming $\delta=1$.

This result is consistent with the finding of Corollary 3.4 in Mai and Scherer (2014), according to which the Gumbel is the only Archimedean copula consistent with exponential distributions of unobserved and observed lifetimes. Beyond this we recover the unique multivariate distribution of the observed lifetimes and the specific shape of marginal intensities.

We notice that, under Assumptions 2.3.1 and 2.3.2, we necessarily recover the same survival distribution presented in Example 5.6 in Mulinacci (2015) which is a particular specification of the Mixed Generalized Marshall-Olkin distributions there introduced and studied.

The resulting distributional features of the obtained distributions are summarized in the following Proposition.

Proposition 2.2. *Under Assumptions 2.3.1 and 2.3.2, the joint survival distribution function of the random vector* $\mathbf{T} = (T_1, T_2, \ldots, T_d)$ *is, for* $\mathbf{t} = (t_1, \ldots, t_d) \in [0, +\infty)^d$, *of type (see (2.3))*

$$\bar{F}_{\mathbf{T}}(t) = \exp\left\{ -\left(\sum_{S\in\mathscr{P}} \lambda_S \max_{j\in S}\{t_j^\theta\} \right)^{\frac{1}{\theta}} \right\}$$

where $\theta \geq 1$. *The marginal survival functions are*

$$\bar{F}_{T_j}(t) = \exp\left(-\left(\sum_{S:j\in S} \lambda_S \right)^{\frac{1}{\theta}} t \right)$$

while the associated survival copula function is (see (2.5))

$$\hat{C}(\mathbf{u}) = \exp\left\{ -\left(\sum_{S\in\mathscr{P}} \max_{j\in S} \left\{ \alpha_{S,j} (-\log u_j)^\theta \right\} \right)^{\frac{1}{\theta}} \right\}$$

where $\alpha_{S,j} \in [0,1]$ *is defined in (2.2).*

As mentioned in Section 2, the above survival copula function is a particular case among the multidimensional generalizations of the copulas studied in Mulinacci (2017) and of the copulas studied in Charpentier et al. (2014). Actually, from the copula structure it is immediate to derive the effect of the dependence among the unobserved components on the dependence of the observed default times. In particular, let us consider the pairs of observed default times (T_i, T_k), and define

$$\hat{\lambda}_{ik} = \sum_{S:i,k\in S} \lambda_S, \qquad \hat{\lambda}_{i(k)} = \sum_{S:i\in S, k\notin S} \lambda_S \qquad \text{and} \qquad \hat{\lambda}_{k(i)} = \sum_{S:k\in S, i\notin S} \lambda_S.$$

Then, the joint survival distribution of (T_i, T_k) is defined for $(t_i, t_k) \in [0, +\infty)^2$ as

$$\bar{F}_{ik}(t_i,t_k) = \exp\left\{-\left(\hat{\lambda}_{ik}\max\{t_i^\theta,t_k^\theta\} + \hat{\lambda}_{i(k)}t_i^\theta + \hat{\lambda}_{k(i)}t_k^\theta\right)^{\frac{1}{\theta}}\right\}.$$

and the associated bivariate copula function is

$$\begin{aligned}\hat{C}_{ik}(u_i,u_k) = &\exp\big\{-\big(\max\{\alpha_i^{ik}(-\log u_i)^\theta, \alpha_k^{ik}(-\log u_k)^\theta\}\\ &+(1-\alpha_i^{ik})(-\log u_i)^\theta + (1-\alpha_k^{ik})(-\log u_k)^\theta\big)^{\frac{1}{\theta}}\big\}\end{aligned}$$

where $\alpha_i^{ik} = \frac{\hat{\lambda}_{ik}}{\bar{\lambda}_i}$ and $\alpha_k^{ik} = \frac{\hat{\lambda}_{ik}}{\bar{\lambda}_k}$.

According to the results shown in Capéraà et al. (2000) and in Mulinacci (2017) we have that the corresponding pairwise Kendall's tau is

$$\tau_{i,k} = \frac{\theta-1}{\theta} + \frac{\tau_{i,k}^{MO}}{\theta}$$

where

$$\tau_{i,k}^{MO} = \frac{\alpha_i^{ik}\alpha_k^{ik}}{\alpha_i^{ik}+\alpha_k^{ik}-\alpha_i^{ik}\alpha_k^{ik}}$$

is the Kendall's tau of the Marshall-Olkin copula.

Notice that in the standard Marshall-Olkin case the dependence among the default times is only defined in terms of the α parameters of the copula function, that represent the measures, in the $[0,1]$ interval, of the sensitivity of each default times to the common components. In this model, the default time of every individual that is not sensitive to any common shock is independent of the default time of any other individual. This is no longer true in the Gumbel-Marshall-Olkin model, since in this case even individuals whose lifetime is insensitive to common shocks play a role in the dependence structure of the default times. This is due to the dependence structure of the unobserved components, so that even the default time of individuals that are insensitive to common shocks may play the role of triggers of the other components, and then in turn may affect the default time of the other individuals in the system. So, the Archimedean dependence has the effect of introducing a floor of positive dependence, raising the general level of dependence in the system. Ignoring the dependence structure of the unobserved components would then naturally lead to mis-specification of the sensitivities of each individual to the common shocks.

While this approach is purely theoretical, in practical applications the model would run into a typical curse of dimensionality problem, with exploding number of unobserved shocks and parameters. Two different routes are suggested to address the problem. The first assumes an exchangeable model with same intensities for each subset of shocks affecting the same number of observed lifetimes (see Hering and Mai, 2012). The other route is to restrict the set of shocks to a smaller number of common shocks and to identify clusters of observable lifetimes exposed to the same set of shocks: in this case the estimation can be conducted relying on the bivariate dependence structures of the observed variables (see Cherubini and Mulinacci, 2016 and Mazo et al., 2015).

2.5 Conclusion

In this paper we provide an extension of the Marshall-Olkin multivariate distribution that could allow for Archimedean dependence of the occurrence times of the unobserved components while keeping the main advantages of the Marshall-Olkin model and preserving exponential marginal default probabilities. The model is derived in full generality for what concerns the number of shocks reaching the system. The main result is that the only extension satisfying these requirements imposes a specific type of dependence among the unobserved components, that is that of a Gumbel copula. For this reason, we call this multivariate distribution the Gumbel-Marshall-Olkin distribution. Like in the Marshall-Olkin distributon, the marginal distribution of the observed default times is exponential, but differently from that, the marginal intensities are no longer linear functions of the unobserved intensities. Moreover, the dependence structure among the observed components is higher than in the Marshall-Olkin model because of the assumption of positive dependence among the unobserved components. Nevertheless, the dependence among the observed default times is again determined by the sensitivities to the common shocks, like in the Marshall-Olkin model, increased by the parameter representing the Archimedean dependence of the unobserved components.

References

1. A. Baglioni, U. Cherubini (2013a), Marking-to-market government guarantees to financial systems: theory and evidence for Europe, Journal of Internationsl Money and Finance, 32, 990-1007
2. A. Baglioni, U. Cherubini (2013b), *Within* and *between* systemic country risk: theory and evidence from the sovereign crisis in Europe, Journal of Economic Dynamics and Control, 37, 1581-1597
3. G. Bernhart, M. Escobar Anel, J.F. Mai, M. Scherer (2013): Default models based on scale mixtures of Marshall-Olkin Copulas: properties and applications. Metrika, 76(2), 179-203.
4. P. Capéraà, A.-L. Fougères, C. Genest (2000): Bivariate distributions with given extreme value attractor. J. of Multivariate Analysis, 72(1), 30-49
5. A. Charpentier, A.-L. Fougeres, C. Genest, J.G. Nešlehová (2014): Multivariate Archimax copulas. Journal of Multivariate Analysis, 126, 118-136
6. U. Cherubini, S. Mulinacci (2016): Marking to market credit derivatives on simultaneous credit events, pre-print
7. Y. Elouerkhaoui (2007): Pricing and hedging in a dynamic credit model. Int. J. of Theoretical and Applied Finance, 10, 703-731
8. C. Genest, L.-P. Rivest (1989): A characterization of Gumbel's family of extreme value distributions. Statistics & Probability Letters, 8, 207-211
9. K. Giesecke (2003): A simple exponential model for dependent defaults. Journal of Fixed Income, 13(3), 74-83
10. E.J. Gumbel (1960): Bivariate exponential distributions. Journal of the American Statistical Association, 55(292), 698-707
11. C. Hering, J.-F. Mai (2012): Moment-based estimation of extendible Marshall-Olkin copulas. Metrika, 75, 601-620
12. H. Joe, (1997): Multivariate Models and Multivariate Dependence Concepts. Chapman & Hall, London

13. H. Li (2008): Tail dependence comparison of survival Marshall-Olkin copulas. Methodology and Computing in Applied Probability 10(1), 3954.
14. H. Li (2009): Orthant tail dependence of multivariate extreme value distributions. J. of Multivariate Analysis, 100(1), 243-256.
15. F. Lindskog, A.J. McNeil (2003): Common Poisson shock models: applications to insurance and credit risk modeling. Astin Bull., 33, 209-238
16. J.F. Mai, M. Scherer (2009): A tractable multivariate default model based on a stochastic time change. Int. J. of Theoretical and Applied Finance, 12(2), 227-249
17. J.F. Mai, M. Scherer, R. Zagst (2013): CIID frailty models and implied copulas. In: Copulae in Mathematical and Quantitative Finance, Lecture Notes in Statistics 213, Springer Verlag , 201-230
18. F. Mai, M. Scherer (2014): Characterization of extendible distributions with exponential minima via processes that are infinitely divisible with respect to time. Extremes, 17(1), 77-95
19. A. W. Marshall, I. Olkin (1967): A multivariate exponential distribution. J. Amer. Statist. Ass., 62(317), 30-44.
20. G. Mazo, S. Girard, F. Forbes (2015): Weighted least-squares inference for multivariate copulas based on dependence coefficients, ESAIM: Probability and Statistics, 19, 746-765.
21. A. J. McNeil, J.G. Nešlehová (2009): Multivariate Archimedean copulas, d-monotone functions and L1-norm symmetric distributions. The Annals of Statistics, 37, 3059-3097.
22. S. Mulinacci (2015): Marshall-Olkin Machinery and Power Mixing: The Mixed Generalized Marshall-Olkin Distribution. In U. Cherubini, F. Durante, S. Mulinacci (eds.) Marshall-Olkin Distributions-Advances in Theory and Applications. Springer Proceedings in Mathematics and Statistics 141. Springer International Publishing Switzerland (2015), 65-86
23. S. Mulinacci (2017): Archimedean-based Marshall-Olkin Distributions and Related Dependence Structures, Methodology and Computing in Applied Probability, DOI: 10.1007/s11009-016-9539-y
24. R.B. Nelsen (2006): An Introduction to Copulas, Second Edition, Springer, New York

Chapter 3
A look at copulas in a curved mirror

Bernard De Baets and Hans De Meyer

Abstract We extend the seminal work of Roger Nelsen on symmetry-related properties and the degree of asymmetry of copulas, by reattributing the role the diagonal plays as axis of symmetry to a continuous strictly increasing curve in the unit square. First, we make explicit the geometrical notion of symmetry of a function on the unit square with respect to a curve. Next, we provide a measure for quantifying to what extent a quasi-copula or copula can be regarded asymmetric with respect to a given curve. Finally, we derive a lower and upper bound on the degree of asymmetry a quasi-copula can possess with respect to a given curve and show that each bound is sharp within the class of copulas.

3.1 Introduction

It was in the winter of 2003 when Roger Nelsen, who was invited speaker at the 24th Linz Seminar on Fuzzy Systems (Austria), introduced us in the world of copulas. Although we had heard of copulas shortly before and even had used them in our research on preference models, we were at that time not at all aware of the rich potential of these operations. Roger's enthusiastic and clear teaching style aroused our interest in the topic and encouraged us to start digging into his famous book. Soon we realized that almost everything known on copulas could be found in it and that any fresh idea on the subject could eventually be traced back to some germ hidden in it. Looking back, our first meeting with Roger in Linz was one of those

Bernard De Baets
KERMIT, Department of Mathematical Modelling, Statistics and Bioinformatics, Ghent University, Coupure links 653, B-9000 Gent, Belgium, e-mail: bernard.debaets@ugent.be

Hans De Meyer
Department of Applied Mathematics, Computer Science and Statistics, Ghent University, Krijgslaan 281 S9, B-9000 Gent, Belgium, e-mail: hans.demeyer@ugent.be

M. Úbeda Flores et al. (eds.), *Copulas and Dependence Models with Applications*,
DOI 10.1007/978-3-319-64221-5_3

lucky circumstances that have left a significant mark on the development of our later research work and for which we still remain very grateful.

As a tribute to the inspiring personality of Roger Nelsen, and at the occasion of his 75th birthday, we want to expose in the present chapter some of our findings on a generalization of the concept of degree of symmetry (or asymmetry) of quasi-copulas and copulas, a concept that is strongly entangled with the probabilistic concept of exchangeability of random variables and that, in geometric terms, is related to diagonal reflection in the unit square (i.e., with the diagonal of the unit square acting as a mirror), whence also indirectly related to the concept of diagonal copulas. Clearly, any one of these viewpoints and insights can, in some way or another, be traced back to the seminal work of Roger Nelsen.

More in detail, we will investigate in the present contribution what are the changes induced by bending the diagonal into a curvilinear mirror, in other words, when we look at copulas and quasi-copulas through a curvilinear looking glass. First, we will define what can be understood as symmetry with respect to a curve in the unit square. Next, we will define a degree of asymmetry of a copula, and more generally of a binary aggregation function. Finally, we will investigate what is the minimum and maximum degree of asymmetry of quasi-copulas and copulas.

To make apparent the differences with the classical approach, we first recall some definitions and properties of copulas and quasi-copulas that will be relevant to the present exposition.

Definition 3.1. A two-dimensional quasi-copula is a function $Q : [0,1]^2 \to [0,1]$ with the following properties [10, 11, 12]:

(i) Q has absorbing element 0 and neutral element 1, i.e. for every $u, v \in [0,1]$,

$$Q(u,0) = Q(0,v) = 0\,, \qquad Q(u,1) = u \quad \text{and} \quad Q(1,v) = v\,;$$

(ii) Q is increasing , i.e. for every $u_1, u_2, v_1, v_2 \in [0,1]$ such that $u_1 \leq u_2$ and $v_1 \leq v_2$,

$$Q(u_2, v_2) \geq Q(u_1, v_1)\,;$$

(iii) Q is 1-Lipschitz continuous, i.e. for every $u_1, u_2, v_1, v_2 \in [0,1]$,

$$|Q(u_2,v_2) - Q(u_1,v_1)| \leq |u_2 - u_1| + |v_2 - v_1|\,.$$

Definition 3.2. A two-dimensional copula is a function $C : [0,1]^2 \to [0,1]$ with the following properties [1, 12, 15]:

(i) C has absorbing element 0 and neutral element 1, i.e. for every $u, v \in [0,1]$,

$$C(u,0) = C(0,v) = 0\,, \qquad C(u,1) = u \quad \text{and} \quad C(1,v) = v\,;$$

(ii) C is 2-increasing, i.e. for every $u_1, u_{2,1}, v_2 \in [0,1]$ such that $u_1 \leq u_2$ and $v_1 \leq v_2$,

$$C(u_2,v_2) - C(u_2,v_1) - C(u_1,v_2) + C(u_1,v_1) \geq 0\,.$$

While every copula is a quasi-copula, there exist proper quasi-copulas that are not copulas. The copulas M and W, defined by $M(u,v) = \min(u,v)$ and $W(u,v) = \max(u+v-1,0)$ are, respectively, the greatest and the smallest quasi-copula, i.e. for any quasi-copula Q, it holds that $W \leq Q \leq M$. The product copula Π, defined by $\Pi(u,v) = uv$, is also known as the independence copula. The set of all quasi-copulas (resp. copulas) will be denoted by $\mathscr{Q}$ (resp. $\mathscr{C}$).

The copula C of an ordered pair of continuous real-valued random variables (r.v.'s) X and Y, defined on the same probability space and such that F_X and F_y are the univariate distribution functions of X and Y, respectively, and $F_{X,Y}$ and $F_{Y,X}$ are the joint distribution functions of (X,Y) and (Y,X), respectively, is the joint distribution function of the random vector $(F_X(X),F_Y(Y))$, i.e. $C(u,v) = F_{X,Y}(F_X(u),F_Y(v))$. The r.v.'s X and Y are said to be exchangeable if $F_X = F_y$ and $F_{X,Y} = F_{Y,X}$. Two identically distributed r.v.'s are exchangeable if and only if their copula C is symmetric, i.e. $C(u,v) = C(v,u)$ for every $(u,v) \in [0,1]^2$.

The diagonal section of a quasi-copula Q is the function $\delta_Q : [0,1] \to [0,1]$ defined by $\delta_Q(t) = Q(t,t)$. Diagonal sections of copulas have the following probabilistic interpretation. If X and Y are identically distributed r.v.'s with distribution function F and copula C, then the distribution function of the r.v. $\max(X,Y)$ is $\delta_C(F(t))$.

The diagonal section of a quasi-copula has the following properties.

Proposition 3.1. *If δ is the diagonal section of a quasi-copula, then*

(i) $\delta(1) = 1$ *;*
(ii) *δ is increasing and 2-Lipschitz continuous, i.e. for any $t_1,t_2 \in [0,1]$ such that $t_1 \leq t_2$,*

$$0 \leq \delta(t_2) - \delta(t_1) \leq 2(t_2 - t_1);$$

(iii) $\delta \leq id$ (the identity function on $[0,1]$), i.e. for every $t \in [0,1]$,

$$\delta(t) \leq t\,.$$

We call any function $\delta : [0,1] \to [0,1]$ that satisfies (i)–(iii) of Proposition 3.1 a diagonal function. Properties (i)–(iii) characterize the diagonal section of a copula.

Proposition 3.2. *For any diagonal function δ there exists at least one copula whose diagonal section is δ.*

This result was proven firstly by Fredricks and Nelsen in [8], who have shown that given a diagonal function δ, the function $K_\delta : [0,1]^2 \to [0,1]$ defined by

$$K_\delta(u,v) = \min\left(u, v, \frac{\delta(u)+\delta(v)}{2}\right),$$

is a copula whose diagonal section is δ. K_δ is called the diagonal copula with diagonal section δ; it is, moreover, the greatest symmetric copula with diagonal section δ. Other structural properties of diagonal copulas have been reported in [17], alternative constructions of copulas with given diagonal section have been investigated in [4, 18], while in [5] necessary and sufficient conditions on a diagonal function have been formulated that ensure the existence of an absolutely continuous copula having this diagonal function as diagonal section. Also mentioned in [15] is the smallest copula with prescribed diagonal section δ, known as the Bertino copula B_δ with diagonal section δ, and given by [3, 9]

$$B_\delta(u,v) = \begin{cases} u - \min\limits_{u \le t \le v} (t - \delta(t)) & \text{, if } u \le v\,, \\ v - \min\limits_{v \le t \le u} (t - \delta(t)) & \text{, if } v \le u\,. \end{cases} \tag{3.1}$$

Note that B_δ is symmetric.

Since exchangeability of random variables is, in general, not a desired property in statistical modelling, there have been many efforts to construct families of non-symmetric copulas. Among the wide variety of techniques for constructing such copulas, it is worth mentioning the technique of diagonal splicing of symmetric copulas with common diagonal section [18]. Together with the development of methods to construct asymmetric copulas, the need arose for defining a measure that quantifies the degree of asymmetry, or in probabilistic terms, the degree of non-exchangeability of a pair of identically distributed r.v.'s. Among various ways for measuring this asymmetry [6, 7], by far the most interesting one is the measure μ_∞, given by:

$$\mu_\infty(Q) = 3\left(\max_{[u,v]\in[0,1]^2} \mid Q(u,v) - Q(v,u) \mid \right),$$

where Q can be any quasi-copula. For this measure, the question how non-symmetric a copula can be, was first raised by Klement and Mesiar [13] and also treated by Nelsen [16] who characterized the copulas that achieve this maximum degree of asymmetry, a result that was subsequently refined in [2]. The scale factor 3 in the definition of μ_∞ ensures that this measure takes values in $[0,1]$.

3.2 Curvilinear sections of a (quasi-)copula

As a first step in the proposed generalization, we have to define the kind of curves (curved mirrors) we want to consider. Let G_ϕ denote the curve in the unit square from (0,0) to (1,1) that is functionally described by an automorphism ϕ of $[0,1]$, i.e. G_ϕ is the graph of a function ϕ having the following properties:

(i) $\phi(0) = 0$ and $\phi(1) = 1$;
(ii) ϕ is continuous and strictly increasing on $[0,1]$.

Clearly, if $\phi = \text{id}$, G_{id} is the diagonal of the unit square, while if $\phi \le \text{id}$ (resp. $\phi \ge \text{id}$), G_ϕ is a curve below (resp. above) the diagonal. We denote by $\mathscr{A}$ the set of all automorphisms ϕ on $[0,1]$ and by $\mathscr{A}_l$ (resp. $\mathscr{A}_u$) the subset of all automorphisms such that $\phi \le id$ (resp. $\phi \ge \text{id}$).

Let the curve G_ϕ be the graph of a given automorphism ϕ. The curvilinear section of a quasi-copula Q, or, shortly, the G_ϕ-section of Q, is the function

$g_{Q,\phi} : [0,1] \to [0,1]$ that is the restriction of Q to G_ϕ, i.e. the function defined by $g_{Q,\phi}(t) = Q(t,\phi(t))$. For example, if $\phi(t) = t^a$ (with $a > 0$), then $g_{M,\phi}(t) = t^a$ and $g_{\Pi(t),\phi} = t^{a+1}$. Note that if $\phi = \mathrm{id}$, then $g_{Q,\mathrm{id}}$ is the diagonal section δ_Q of Q.

We can attribute the following probabilistic meaning to the G_ϕ-section of a copula. If X and Y are identically distributed r.v.'s with distribution function F, assumed to be a bijection, and with copula C, then the distribution function of the r.v. $\max(X,\psi(Y))$, with $\psi : \mathbb{R} \to \mathbb{R}$ a bijection, is $g_{C,\phi}(F(t))$, where $\phi = F\psi^{-1}F^{-1}$. In particular, if X and Y are uniformly distributed r.v.'s on $[0,1]$ with copula C, then the distribution function of the r.v. $\max(X,\phi(Y))$ is $g_{C,\phi^{-1}}(t)$.

The G_ϕ-section of a quasi-copula has the following properties.

Proposition 3.3. *If g is the G_ϕ-section of a quasi-copula for some $\phi \in \mathscr{A}$, then:*

(i) $\max(0,t+\phi(t)-1) \leq g(t) \leq \min(t,\phi(t))$ *for any $t \in [0,1]$;*
(ii) $0 \leq g(t') - g(t) \leq t' - t + \phi(t') - \phi(t)$ *for any $t \leq t' \in [0,1]$.*

These properties immediately follow from the fact that any quasi-copula is 1-Lipschitz continuous, and is bounded from below by W and from above by M.

Given a curve $G_\phi \in \mathscr{A}$, any function $g : [0,1] \to [0,1]$ that has properties (i) and (ii) of Proposition 3.3 will be called a G_ϕ-function. The set of all G_ϕ-functions will be denoted by $\mathscr{S}_\phi$. The following proposition provides a characterization of the curvilinear section of a (quasi-)copula.

Proposition 3.4. *Let $\phi \in \mathscr{A}$. A function $g : [0,1] \to [0,1]$ is a G_ϕ-section of a copula if and only if $g \in \mathscr{S}_\phi$.*

Proof. If g is a G_ϕ-section of a copula Q, then obviously $g \in \mathscr{S}_\phi$. Conversely, consider a function $g \in \mathscr{S}_\phi$. In [19], it is shown that for an increasing compact set S and a quasi-copula Q, the function $C_{S,Q} : [0,1]^2 \to [0,1]$ defined by

$$C_{S,Q}(u,v) = \max(0, u+v-1, \max_{(a,b)\in S} (Q(a,b) - \max(a-u,0) - \max(b-v,0)))$$

is a copula. Moreover, for any copula C that coincides with Q on the subset S, it holds that $C_{S,Q} \leq C$. Since ϕ is increasing and continuous, it follows that its graph is an increasing compact set. Denoting $a = z$, $b = \phi(z)$ and $C_{S,Q} = C_g$, we obtain

$$C_g(u,v) = \max(0, u+v-1, \max_{z\in[0,1]} (g(z) - \max(z-u,0) - \max(\phi(z)-v,0)))\,.$$

Suppose that $v \leq \phi(u)$, then $C_g(u,v)$ can be rewritten as

$$\begin{aligned} C_g(u,v) = \max(0, u+v-1, \quad & \max_{z\in[0,\phi^{-1}(v)]} g(z), \max_{z\in[\phi^{-1}(v),u]} (g(z)-\phi(z)+v), \\ & \max_{z\in[u,1]} (g(z)-\phi(z)-z+u+v)))\,. \end{aligned}$$

Property (ii) of Proposition 3.3 expresses that g is increasing and that $g(z)-\phi(z)-z$ is decreasing, whereas property (i) implies that $g(u)-\phi(u)+v \geq u+v-1$ for any $(u,v) \in [0,1]^2$. Hence, C_g can be further simplified to the equivalent form

$$C_g(u,v) = \max(g(\phi^{-1}(v)), \max_{z\in[\phi^{-1}(v),u]} (g(z)-\phi(z)+v), g(u)-\phi(u)+v)\,.$$

Note that $g(z)-\phi(z)+v = g(\phi^{-1}(v))$ when $z=\phi^{-1}(v)$, while $g(z)-\phi(z)+v = g(u)-\phi(u)+v$ when $z=u$. Hence,

$$C_g(u,v) = v - \min_{z\in[\phi^{-1}(v),u]} (\phi(z)-g(z))\,.$$

Similarly, we can prove that $C_g(u,v) = u - \min_{z\in[u,\phi^{-1}(v)]} (z-g(z))$ when $v > \phi(u)$. Combining the above, we finally obtain

$$C_g(u,v) = \begin{cases} v - \min\{\phi(z)-g(z) \mid z \in [\phi^{-1}(v),u]\} & \text{, if } v \le \phi(u)\,, \\ u - \min\{z-g(z) \mid z \in [u,\phi^{-1}(v)]\} & \text{, otherwise.} \end{cases} \tag{3.2}$$

Moreover, C_g is the smallest copula with G_ϕ-section g. □

If ϕ=id, then C_g in (3.2) is the Bertino copula, whence it is justified to call (3.2) the Bertino copula with G_ϕ-section g. Note that in contrast to the proof of Nelsen in the classical case of given diagonal section δ, we have not constructed a generalization of the copula K_δ. This is not surprising, as the obvious generalization of K_δ in case of a given G_ϕ-section g is the bivariate function

$$K_{g,\phi}(u,v) = \min\left(u, v, \frac{g(u)+g(\phi^{-1}(v))}{2}\right),$$

and this function is not necessarily a copula, not even a quasi-copula. Indeed, it is not difficult to find G_ϕ-functions g such that in some point (u,v) where $K_{g,\phi}(u,v) < M(u,v)$, the 1-Lipschitz continuity condition is violated.

3.3 Symmetry of a function with respect to a curve

The notion of symmetry of a $[0,1]^2 \to [0,1]$ function w.r.t. a curve G_ϕ, which is the graphical representation of a function $\phi \in \mathscr{A}$, has been investigated before [14]. It generalizes the classical concept of symmetry in the following manner. Any point $(u,v) \in [0,1]^2$ determines together with the points $(\phi^{-1}(v),v)$, $(u,\phi(u))$ and $(\phi^{-1}(v),\phi(u))$, a rectangle in the unit square (see Figure 3.1). Note that the points (u,v) and $(\phi^{-1}(v),\phi(u)$ determine the same rectangle, while the two other points lie on the curve G_ϕ.

Definition 3.3. Let $\phi \in \mathscr{A}$. A function $f : [0,1]^2 \to [0,1]$ is called G_ϕ-symmetric if for every $(u,v) \in [0,1]^2$,

$$f(u,v) = f(\phi^{-1}(v),\phi(u))\,.$$

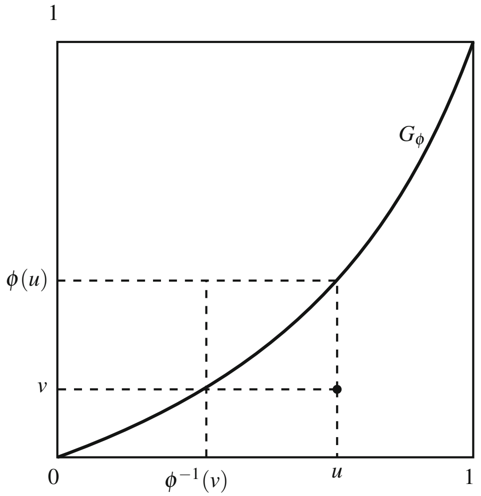


Fig. 3.1: Illustration of the unique rectangle that is determined by the curve G_ϕ ($\phi \in \mathscr{A}_l$) and the point $(u,v) \in [0,1]^2$.

Example 3.1.

1. Consider the function $\phi : [0,1] \to [0,1]$ defined by $\phi(t) = t^2$ and the function $f : [0,1]^2 \to [0,1]$ defined by $f(u,v) = u^2 v$. One easily verifies that f is G_ϕ-symmetric.
2. Consider the function $\phi : [0,1] \to [0,1]$ defined by $\phi(t) = \sqrt{t}$ and the function $f : [0,1]^2 \to [0,1]$ defined by $f(u,v) = \min(uv, u^2\sqrt{v})$. Again, one easily verifies that f is G_ϕ-symmetric. $\triangleleft$

When applied to a copula, the geometric notion of G_ϕ-symmetry can be related to the probabilistic notion of ϕ-exchangeability in the following sense. Let X and Y be uniformly distributed r.v.'s on $[0,1]$ with copula C and $\phi \in \mathscr{A}$ a given automorphism. The r.v.'s X and $\phi(Y)$ are exchangeable r.v.'s if and only if $\text{Prob}(X \leq a, \phi(Y) \leq b) = \text{Prob}(X \leq b, \phi(Y) \leq a)$ for any $a,b \in [0,1]$, hence, if and only if $C(a,\phi^{-1}(b)) = C(b,\phi^{-1}(a))$ for any $a,b \in [0,1]$. This last condition is equivalent with the property of G_ϕ-symmetry of C. Obviously, X and $\phi(Y)$ can only be exchangeable r.v.'s if $\phi = \text{id}$. To see this, it suffices to put $b = 1$ hereabove, which leads to the condition $a = \phi(a)$ for all $a \in [0,1]$.

Next, we want to measure, for a given curve G_ϕ, the extent to which a function $f : [0,1]^2 \to [0,1]$ is not G_ϕ-symmetric, in other words, we want a measure of G_ϕ-asymmetry. Inspired by the measure μ_∞ that quantifies the degree of asymmetry w.r.t. the diagonal, we propose the following extension.

Definition 3.4. Given an automorphism ϕ, the degree of G_ϕ-asymmetry $\mu_\phi(f)$ of a function $f : [0,1]^2 \to [0,1]$ is defined by

$$\mu_\phi(f) = \sup_{(u,v)\in[0,1]^2} |f(u,v) - f(\phi^{-1}(v), \phi(u))| . \tag{3.3}$$

When μ_ϕ is applied to a copula C, it may be called a measure of indistinguishability between r.v.'s X and $\phi(Y)$ where X and Y are uniformly distributed r.v.'s on $[0,1]$ with copula C. Note that $\mu_{\mathrm{id}}(C) = \frac{1}{3}\mu_\infty(C)$. It follows that in the class of copulas $\mathscr{C}$, μ_{id} takes values from 0 (for any symmetric copula) to 1/3 (for the most asymmetric copulas). In the following two sections we investigate, again in the class $\mathscr{C}$, what are the minimum and maximum values of μ_ϕ, given an arbitrary but fixed $\phi \in \mathscr{A}$.

3.4 Minimum degree of asymmetry with respect to a curve

In this section, we investigate the minimum degree of asymmetry of a quasi-copula Q with respect to a curve G_ϕ ($\phi \in \mathscr{A}$) according to the measure μ_ϕ. Let us introduce the following notation

$$\lambda_\phi = \max_{t\in[0,1]} |t - \phi(t)| .$$

This quantity, which only depends on ϕ, turns out to provide a lower bound for μ_ϕ.

Proposition 3.5. *For any $\phi \in \mathscr{A}$ and every $Q \in \mathscr{Q}$,*

$$\mu_\phi(Q) \geq \lambda_\phi .$$

Proof. One easily verifies that

$$\mu_\phi(Q) \geq \sup_{u\in[0,1]} |Q(u,1) - Q(1,\phi(u))| = \max_{u\in[0,1]} |u - \phi(u)| = \lambda_\phi . \qquad \square$$

Note that $0 \leq \lambda_\phi < 1$ for any $\phi \in \mathscr{A}$. Consider, for instance, the function $\phi_n(t) = t^n$, with $n \in \mathbb{N}$. Clearly $\lambda_{\phi_1} = 0$, while λ_{ϕ_n} gets closer to 1 when n increases.

Inspired by Proposition 3.5, given $\phi \in \mathscr{A}$, we say that a quasi-copula has minimum degree of asymmetry w.r.t. to the curve G_ϕ if $\mu_\phi(Q) = \lambda_\phi$.

From here onward we will consider mainly automorphisms $\phi \in \mathscr{A}_l \cup \mathscr{A}_u$, in other words, curves G_ϕ that lie entirely below or entirely above the diagonal. Moreover, we can even restrict the discussion to automorphisms $\phi \in \mathscr{A}_l$; indeed, the situation of $\phi \in \mathscr{A}_u$ is equivalent to the situation of $\phi^{-1} \in \mathscr{A}_l$ together with the interchange of the coordinates u and v.

Lemma 3.1. *Let $\phi \in \mathscr{A}_l$ and $Q \in \mathscr{Q}$. Then for every $(u,v) \in [0,1]^2$,*

$$|Q(u,\phi(v)) - Q(\phi(u),v)| \leq \max(u - \phi(u), v - \phi(v)) .$$

Proof. Using $\phi \leq \mathrm{id}$, and the increasingness and 1-Lipschitz continuity of Q, it follows that

$$Q(u,\phi(v)) - Q(\phi(u),v) \leq Q(u,v) - Q(\phi(u),v) \leq u - \phi(u)$$

and

$$Q(\phi(u),v) - Q(u,\phi(v)) \leq Q(u,v) - Q(u,\phi(v)) \leq v - \phi(v)\,,$$

which implies that $|Q(u,\phi(v)) - Q(\phi(u),v)| \leq \max(u-\phi(u), v-\phi(v))$. □

The following proposition identifies some interesting quasi-copulas that have minimum degree of G_ϕ-asymmetry.

Proposition 3.6. *Let $\phi \in \mathscr{A}_l$. For any symmetric $Q \in \mathscr{Q}$, it holds that*

$$\mu_\phi(Q) = \lambda_\phi\,.$$

Proof. For all $Q \in \mathscr{Q}$ it holds that

$$\mu_\phi(Q) \geq \sup_{u\in[0,1]} |Q(u,1) - Q(1,\phi(u))| = \max_{u\in[0,1]} (u-\phi(u)) = \lambda_\phi\,.$$

Using z as a shorthand notation for $\phi^{-1}(v)$, it follows that

$$\mu_\phi(Q) = \sup_{(u,v)\in[0,1]^2} |Q(u,v) - Q(\phi^{-1}(v),\phi(u))| = \sup_{(u,z)\in[0,1]^2} |Q(u,\phi(z)) - Q(z,\phi(u))|\,.$$

Since Q is symmetric, $\mu_\phi(Q)$ can be rewritten in the form

$$\mu_\phi(Q) = \sup_{(u,z)\in[0,1]^2} |Q(u,\phi(z)) - Q(\phi(u),z)|\,.$$

Using Lemma 3.1 and Proposition 3.5, it follows that

$$\mu_\phi(Q) \leq \sup_{(u,z)\in[0,1]^2} \max(u-\phi(u), z-\phi(z)) = \sup_{u\in[0,1]} (u-\phi(u)) = \lambda_\phi \leq \mu_\phi(Q)\,.$$

Therefore, $\mu_\phi(Q) = \lambda_\phi$. □

The following proposition illustrates that there are also non-symmetric quasi-copulas that have minimum degree of asymmetry w.r.t. a given curve G_ϕ.

Proposition 3.7. *Let $\phi \in \mathscr{A}_l$. For any G_ϕ-function g, the Bertino-copula C_g with G_ϕ-section g, given by (3.2), has minimum degree of G_ϕ-asymmetry.*

Proof. For every point (u,v) such that $v \leq \phi(u)$, we obtain

$$\begin{aligned} &C_g(\phi^{-1}(v),\phi(u)) - C_g(u,v) \\ &\quad = \phi^{-1}(v) - v - \min_{z\in[\phi^{-1}(v),u]} (z-g(z)) + \min_{z\in[\phi^{-1}(v),u]} (\phi(z)-g(z))\,. \end{aligned}$$

Let z^* denote a point where the function $z-g(z)$ attains its maximum value in $[\phi^{-1}(v),u]$, then

$$C_g(\phi^{-1}(v),\phi(u)) - C_g(u,v) \leq \phi^{-1}(v) - v - z^* + \phi(z^*)\,.$$

Since both $0 \leq \phi^{-1}(v) - v \leq \lambda_\phi$ and $z^* - \phi(z^*) \leq \lambda_\phi$, it follows that

$$C_g(\phi^{-1}(v), \phi(u)) - C_g(u,v) \leq \lambda_\phi \,.$$

Similarly, we can prove that

$$C_g(\phi^{-1}(v), \phi(u)) - C_g(u,v) \geq -\lambda_\phi \,,$$

which completes the proof. □

3.5 Maximum degree of asymmetry with respect to a curve

In this section, we investigate the maximum degree of G_ϕ-asymmetry a quasi-copula Q can have in case $\phi \in \mathscr{A}_l$. In fact, we are looking for an upper bound of $|Q(u,v) - Q(\phi^{-1}(v), \phi(u))|$ valid for every $(u,v) \in [0,1]^2$ and any $Q \in \mathscr{Q}$. To avoid working with absolute values, we split the problem into two subproblems:

1. Find an upper bound on $Q(u,v) - Q(\phi^{-1}(v), \phi(u))$ for every $(u,v) \in [0,1]^2$ such that $v \geq \phi(u)$ and every $Q \in \mathscr{Q}$ such that $Q(u,v) \geq Q(\phi^{-1}(v), \phi(u))$;
2. Find an upper bound on $Q(u,v) - Q(\phi^{-1}(v), \phi(u))$ for every $(u,v) \in [0,1]^2$ such that $v \leq \phi(u)$ and every $Q \in \mathscr{Q}$ such that $Q(u,v) \geq Q(\phi^{-1}(v), \phi(u))$.

The required upper bound is then the maximum of the upper bounds of subproblems 1 and 2, respectively. Note that (u,v) is always the point where Q attains a value greater than it does in the opposite corner point $(\phi^{-1}(v), \phi(u))$.

In both subproblems, the upper bound follows from two conditions, one related to the 1-Lipschitz continuity of Q, the other to the bounding inequalities $W \leq Q \leq M$. Making this explicit, we obtain

1. If $v \geq \phi(u)$ and $Q(u,v) \geq Q(\phi^{-1}(v), \phi(u))$, it must hold that

$$\begin{aligned} &Q(u,v) - Q(\phi^{-1}(v), \phi(u)) \\ &\quad \leq \min[\max(\phi^{-1}(v) - u, v - \phi(u)), \min(u,v) - \max(\phi^{-1}(v) + \phi(u) - 1, 0)] \,. \end{aligned}$$

2. If $v \leq \phi(u)$ and $Q(u,v) \geq Q(\phi^{-1}(v), \phi(u))$, it must hold that

$$\begin{aligned} &Q(u,v) - Q(\phi^{-1}(v), \phi(u)) \\ &\quad \leq \min[\max(u - \phi^{-1}(v), \phi(u) - v), v - \max(\phi^{-1}(v) + \phi(u) - 1, 0)] \,. \end{aligned}$$

This is the starting point for proving the following proposition.

Proposition 3.8. *For any $\phi \in \mathscr{A}_l$, it holds for every $Q \in \mathscr{Q}$ that*

$$\mu_\phi(Q) \leq \max(p_1, p_2) \,,$$

with p_1 the unique fixed point in $[0,1]$ of the function f_1 given by

$$f_1(t) = \phi(1-\phi(t)) - \phi(t)\,,$$

and p_2 *the unique fixed point in* $[0,1]$ *of the function* f_2 *given by*

$$f_2(t) = 1 - t - \phi(t)\,.$$

Proof. Note that since ϕ is increasing, f_1 and f_2 are decreasing. Since these functions are also continuous, they have exactly one fixed point. Also $2p_2 = 1-\phi(p_2)$, whence $p_2 \leq 1/2$. We analyze the subproblem of $v \geq \phi(u)$. In this case, the upper bound for a fixed (u,v) is

$$\begin{aligned}
&\min[\max(\phi^{-1}(v)-u, v-\phi(u)), \min(u,v) - \max(\phi^{-1}(v)+\phi(u)-1,0)] \\
&= \max[\min(\phi^{-1}(v)-u, \min(u,v)-\max(\phi^{-1}(v)+\phi(u)-1,0)), \\
&\qquad \min(v-\phi(u), \min(u,v)-\max(\phi^{-1}(v)+\phi(u)-1,0))]\,.
\end{aligned}$$

We need to compute

$$\begin{aligned}
&\max_{v\geq\phi(u)} \max[\min(\phi^{-1}(v)-u, \min(u,v)-\max(\phi^{-1}(v)+\phi(u)-1,0)), \\
&\qquad \min(v-\phi(u), \min(u,v)-\max(\phi^{-1}(v)+\phi(u)-1,0))] \\
&= \max\Bigg(\max_{v\geq\phi(u)} \min[\phi^{-1}(v)-u, \min(u,v)-\max(\phi^{-1}(v)+\phi(u)-1,0)], \\
&\qquad \max_{v\geq\phi(u)} \min[\phi^{-1}(v)-u, \min(u,v)-\max(\phi^{-1}(v)+\phi(u)-1,0)]\Bigg)\,.
\end{aligned}$$

Let us denote by S_ϕ the graph of the function $\phi(1-\phi(t))$. Note that the part of the curve S_ϕ above (resp. below) G_ϕ is the mirror image with respect to G_ϕ of the part of the opposite diagonal below (resp. above) G_ϕ. Indeed, let $(t, 1-t)$ be the coordinates of a point on the opposite diagonal, then the mirrored point w.r.t. G_ϕ is the point with coordinates $(\phi^{-1}(1-t), \phi(t))$ which clearly lies on S_ϕ.

We use the curve S_ϕ and the diagonal D to separate the part of the unit square above the curve G_ϕ into four subdomains labeled I to IV (see Figure 3.2). The optimization problem hereabove can then be split into eight optimization problems, two on each of the subdomains I–IV. For example, on subdomain I, delimited by the left border of the unit square, the diagonal D and the curve S_ϕ, we first consider the optimization problem

$$\max_{(u,v)\in I} \min(\phi^{-1}(v)-u, u)\,.$$

The maximum is attained in the intersection point of the curve F_ϕ, which is the graph of the function $\phi(2t)$, and the curve S_ϕ, hence in the point (x,y) where x is the unique solution of the equation $\phi(2t) = \phi(1-\phi(t))$, or equivalently of the equation $2t = 1-\phi(t)$, and $y = \phi(2x)$. Clearly, $(x,y) = (p_2, \phi(2p_2))$. Note that p_2 is also the abscissa of the intersection point of the curve G_ϕ and the line with equation $v = 1 - 2u$ (see Figure 3.2). The value of the maximum is p_2.

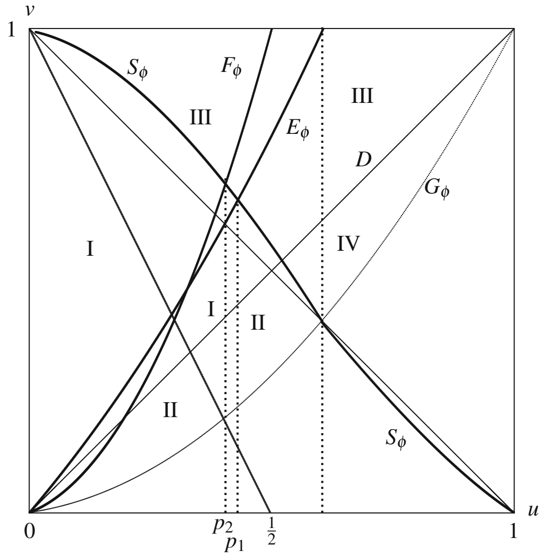


Fig. 3.2: Different curves in the unit square that delimit subdomains I-IV and determine the position of the points from which the upper bound of the degree of G_ϕ-asymmetry is computed. Shown is a situation where $p_1 > p_2$.

The second optimization problem on subdomain I is

$$\max_{(u,v)\in I} \min(v - \phi(u), u)\,.$$

Now the maximum is attained in the intersection point of the curve E_ϕ, which is the graph of the function $\phi(t) + t$, and the curve S_ϕ, hence in the point (x, y) where x is the unique solution of the equation $\phi(t) + t = \phi(1 - \phi(t))$ and $y = \phi(x) + x$. Clearly, $(x, y) = (p_1, \phi(p_1) + p_1)$. The value of the maximum is p_1. It follows that

$$\max_{(u,v)\in I} \big(Q(u,v) - Q(\phi^{-1}(v), \phi(u))\big) \leq \max(p_1, p_2)\,.$$

It is straightforward to analyze in a similar way the optimization problems on the domains II-IV, and after that, to analyze the problem on the part of the unit square below the curve G_ϕ. The outcome of all these lengthy computations is nonetheless surprisingly simple: the maximum that we obtained so far gets not surpassed. Hence, $\mu_\phi(Q) \leq \max(p_1, p_2)$. □

Depending on the curve G_ϕ, the maximum can be either p_1 or p_2. This is illustrated in the following example.

Example 3.2. Let $\phi \in \mathscr{A}_l$ be the piecewise linear function with graph G_ϕ composed of a segment from $(0,0)$ to (a,b) and a segment from (a,b) to $(1,1)$, with $a > b$, i.e.,

$$\phi(t) = \begin{cases} \dfrac{b}{a}t & , \text{ if } t \leq a, \\ 1 - \dfrac{1-b}{1-a}(1-t) & , \text{ if } t > a. \end{cases}$$

One can easily verify that $p_1 > p_2$ if $a+b < 1$, $p_1 = p_2$ if $a+b = 1$, and $p_1 < p_2$ if $a+b > 1$. ◁

The following proposition yields a sufficient condition on ϕ ensuring that $p_1 \geq p_2$.

Proposition 3.9. *If $\phi \in \mathscr{A}_l$ satisfies the inequality*

$$\phi(1-\phi(t)) \geq 1-t \tag{3.4}$$

for all $t \in [0,1/2]$, then $p_1 \geq p_2$.

Proof. Suppose $p_1 < p_2$. Since ϕ is strictly increasing, it holds that $\phi(p_1) \leq \phi(p_2)$. It follows that

$$p_2 = 1 - p_2 - \phi(p_2) \leq \phi(1-\phi(p_2)) - \phi(p_2) > \phi(1-\phi(p_1)) - \phi(p_1) = p_1\,,$$

which is a contradiction. □

Example 3.3. Consider the function $\phi_n : [0,1] \to [0,1]$ defined by $\phi_n(t) = t^n$, with $n \in \mathbb{N}$. Clearly, ϕ_n satisfies (3.4) for any $n \in \mathbb{N}$. In Table 1, the values of λ_{ϕ_n} and p_1 are listed for some values of n. Note that the length of the interval $[\lambda_{\phi_n}, p_1]$ decreases when n grows. ◁

Table 3.1: The values of λ_{ϕ_n} and p_1 for the case $\phi_n(t) = t^n$ and for different values of n.

n	1	2	4	8	10	∞
λ_{ϕ_n}	0	1/4	0.472470	0.650123	0.696837	1
p_1	1/3	0.445042	0.560689	0.671621	0.704910	1

Example 3.4. Consider the function $\phi_\theta : [0,1] \to [0,1]$ defined by $\phi_\theta(t) = \frac{\theta t}{1-(1-\theta)t}$, with $\theta \in [0,1]$. Clearly, ϕ_θ satisfies (3.4) for any $\theta \in [0,1]$. One easily verifies that

$$p_1 = \frac{3-\sqrt{1+8\theta}}{4(1-\theta)} \quad \text{and} \quad \lambda_{\phi_\theta} = \frac{1-\sqrt{\theta}}{1+\sqrt{\theta}}\,.$$

◁

The upper bound $\max(p_1, p_2)$ is sharp. In both situations, whether $p_1 \geq p_2$ or $p_2 > p_1$, a shuffle of M can be found for which μ_ϕ attains the upper bound, namely if $p_1 \geq p_2$:

$$C_1(u,v) = \min(u,v,\max(u-1+\phi(p_1),0)+\max(v-\phi(p_1),0))\,;$$

if $p_2 > p_1$:

$$\begin{aligned} C_2(u,v) &= \min(u,v,\max(u-1+\phi(p_2),0) \\ &\quad + \min(\max(v-\phi(p_2),0),\max(u-p_2,0)+\max(v+p_2-\phi(2p_2),0)))\,. \end{aligned}$$

The support of these shuffles is shown in Figure 3.

Since shuffles of M are copulas, the upper bound is also a sharp upper bound on the degree of G_ϕ-asymmetry in the class of copulas.

Proposition 3.10. *Given $\phi \in \mathscr{A}$. For any copula $C \in \mathscr{C}$, it holds that*

$$\lambda_\phi \leq \mu_\phi(C) \leq \max(p_1, p_2)\,,$$

with p_1 and p_2 the fixed points in $[0,1]$ of the functions $\phi(1-\phi(t))-\phi(t)$ and $1-t-\phi(t)$, respectively.

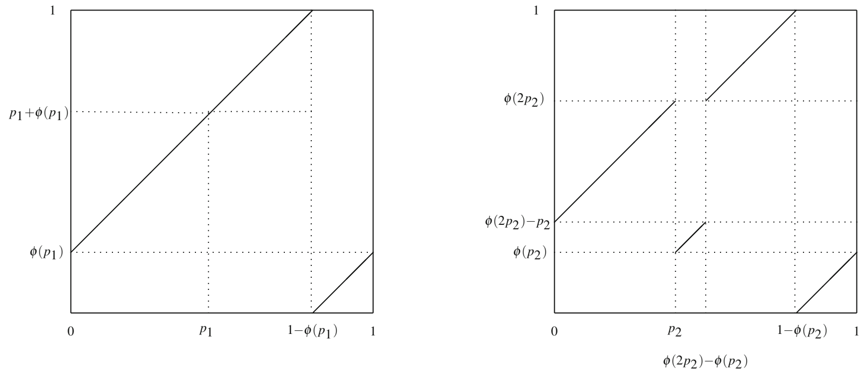


Fig. 3.3: The support of shuffle C_1 (left figure) in the case $p_1 \geq p_2$ and the support of shuffle C_2 (right figure) in the case $p_2 > p_1$.

3.6 Conclusions

Our investigation of a generalisation of the degree of asymmetry of copulas with respect to a curve has revealed two major differences with respect to the classical situation where the curve is the diagonal: there do not exist symmetric copulas (in other words, the minimum degree of asymmetry is strictly positive), and the maximum degree of asymmetry follows from the relative position of the fixed points of two functions that depend on the given curve.

References

1. Alsina, C., Frank, M.J., Schweizer, B.: Associative Functions: Triangular Norms and Copulas. World Scientific, Singapore (2006)
2. Beliakov, G., De Baets, B., De Meyer, H., Nelsen, R.B., Úbeda Flores, M.: Best possible bounds on the set of copulas with given degree of non-exchangeability. Journal of Mathematical Analysis and Applications **417**, 451–468 (2014)
3. Bertino, S.: Sulla dissomiglianza tra mutabili cicliche. Metron **35**, 53–88 (2011)
4. De Baets, B., De Meyer, H., Úbeda Flores, M.: Constructing copulas with given diagonal and opposite diagonal sections. Communications in Statistics - Theory and Methods **40**, 828–843 (2011)
5. Durante, F., Jaworski, P.: Absolutely continuous copulas with given diagonal sections. Communications in Statistics - Theory and Methods **37**, 2924–2942 (2008)
6. Durante, F., Klement, E., Sempi, C., Úbeda Flores, M.: Measures of non-exchangeability for bivariate random vectors. Statistical Papers **151**, 687–699 (2010)
7. Durante, F., Perrone, E.: Asymmetric copulas and their application in design of experiments. In: S. Saminger-Platz, R. Mesiar (eds.) On Logical, Algebraic, and Probabilistic Aspects of Fuzzy Set Theory, Studies in Fuzziness and Soft Computing 336, pp. 157–172. Springer Intern. Publ. Switzerland (2016)
8. Fredricks, G.A., Nelsen, R.B.: Copulas constructed from diagonal sections. In: V. Beneš, J. Štěpán (eds.) Distributions with given marginals and moment problems, pp. 129–136. Kluwer Acad. Publ., Dordrecht (1997)
9. Fredricks, G.A., Nelsen, R.B.: The Bertino family of copulas. In: C.M. Cuadras, J. Fortiana, J. Rodríguez Lallena (eds.) Distributions with given Marginals and Statistical Modelling, pp. 81–91. Kluwer Acad. Publ., Dordrecht (2002)
10. Genest, C., Quesada Molina, J.J., Rodríguez Lallena, J.A., Sempi, C.: A characterization of quasi-copulas. Journal of Multivariate Analysis **69**, 193–205 (1999)
11. Janssens, S., De Baets, B., De Meyer, H.: Bell-type inequalities for quasi-copulas. Fuzzy Sets and Systems **148**, 263–278 (2004)
12. Klement, E., Kolesárová, A.: Extension to copulas and quasi-copulas as special 1-Lipschitz aggregation operators. Kybernetika **41**, 329–348 (2005)
13. Klement, E., Mesiar, R.: How non-symmetric can a copula be? Commentationes Mathematicae Universitatis Carolinae **47**, 141–148 (2006)
14. Maes, K.C., De Baets, B.: Orthosymmetrical monotone functions. Bulletin of the Belgian Mathematical Society - Simon Stevin **14**, 99–116 (2007)
15. Nelsen, R.B.: An Introduction to Copulas, second edn. Springer Series in Statistics. Springer, New York (2006)
16. Nelsen, R.B.: Extremes of nonexchangeability. Statistical papers **48**, 329–336 (2007)
17. Nelsen, R.B., Fredricks, G.: Diagonal copulas. In: V. Beneš, J. Štěpán (eds.) Distributions with given Marginals and Moment Problems, pp. 121–128. Kluwer Academic Publishers, Dordrecht (1997)
18. Nelsen, R.B., Quesada-Molina, J.J., Rodríguez-Lallena, J.A., Úbeda Flores, M.: On the construction of copulas and quasi-copulas with given diagonal sections. Insurance: Mathematics and Economics **42**, 473–483 (2008)
19. Tankov, P.: Improved Fréchet bounds and model-free pricing of multi-asset options. Journal of Applied Probability **48**, 389–403 (2011)

Chapter 4
Copula–based clustering methods

F. Marta L. Di Lascio, Fabrizio Durante, and Roberta Pappadà

Abstract We review some recent clustering methods based on copulas. Specifically, in the dissimilarity–based clustering framework, we describe and compare methods based on concordance or tail-dependence concept. An illustration is hence provided by using a time series dataset formed by the constituent data of the S&P 500 observed during the financial crisis of 2007-2008. Next, in the likelihood–based clustering framework, we present and discuss a clustering algorithm based on copula and called CoClust. Here, an application to the gene expression profiles of human tumour cell lines is provided to describe the methodology. Finally, a comparison between the two different approaches is performed through a case study on environmental data.

4.1 Introduction

Cluster analysis, also called *data segmentation*, is related to the idea of grouping or segmenting a collection of objects (observations or variables) into subsets or "clusters", such that those within each cluster are more closely related to one another (according to a specific criterion) than objects assigned to different clusters. Cluster analysis is sometimes used to form descriptive statistics to ascertain whether or not

F. Marta L. Di Lascio
Faculty of Economics and Management, Free University of Bozen-Bolzano, Italy
e-mail: marta.dilascio@unibz.it

Fabrizio Durante
Dipartimento di Scienze dell'Economia, Università del Salento, Lecce, Italy
e-mail: fabrizio.durante@unisalento.it

Roberta Pappadà
Department of Economics, Business, Mathematics and Statistics "Bruno de Finetti", University of Trieste, Italy
e-mail: rpappada@units.it

M. Úbeda Flores et al. (eds.), *Copulas and Dependence Models with Applications*,
DOI 10.1007/978-3-319-64221-5_4

the data consists of a set of distinct groups, each group representing objects with specific properties.

Generally, any cluster analysis starts with a $(n \times p)$*–data matrix* $\mathbf{X}$

$$\mathbf{X} = \begin{bmatrix} x_{11} & \dots & x_{1j} & \dots & x_{1j'} & \dots & x_{1p} \\ \vdots & \dots & \vdots & \dots & \vdots & \dots & \vdots \\ x_{i1} & \dots & x_{ij} & \dots & x_{ij'} & \dots & x_{ip} \\ \vdots & \dots & \vdots & \dots & \vdots & \dots & \vdots \\ x_{i'1} & \dots & x_{i'j} & \dots & x_{i'j'} & \dots & x_{i'p} \\ \vdots & \dots & \vdots & \dots & \vdots & \dots & \vdots \\ x_{n1} & \dots & x_{nj} & \dots & x_{nj'} & \dots & x_{np} \end{bmatrix} \tag{4.1}$$

and aims at dividing the n rows or the p columns, says r elements, into K non-empty subsets $\mathscr{C}_1, \dots, \mathscr{C}_K$, such that:

- each $\mathscr{C}_k$ $(k = 1, \dots, K)$ contains (at least) one element;
- $|\mathscr{C}_1| + \cdots + |\mathscr{C}_K| \leq r$, i.e. the sum of the cardinalities of all subsets is bounded by r. Notice that, we allow some elements to be not assigned to any cluster.

Model–based clustering methods assume that the data matrix $\mathbf{X}$ is generated according to a specific (stochastic) data generating process (DGP hereafter). Here, inspired by Sklar's decomposition of a multivariate probability law [50], we assume that the data matrix $\mathbf{X}$ is generated by a K-dimensional copula such that each of the K subsets of elements is a realization of a (continuous) random variable X_k (with $k = 1, \dots, K$), and the joint distribution function of $(X_1, \dots, X_k, \dots, X_K)$ can be expressed as $C(F_1, \dots, F_k, \dots, F_K)$, where $F_1, \dots, F_k, \dots, F_K$ are the cumulative distribution functions of $(X_1, \dots, X_k, \dots, X_K)$, while C is the linking copula, which is unique since the marginal distributions are continuous (see, e.g., [24, 43]). In practice, K is often assumed to be equal to the number of columns (respectively, rows) and the elements of each column (respectively, row) are realizations from a given univariate model, linked to all other columns via $C(\cdot)$. Hence, copula–based clustering analysis is a sort of sub-class of model–based clustering methods that uses the copula information (in a parametric as well as non-parametric form) to derive the specific criterion that determines the clustering composition. As such, copula–based clustering methods are *rank invariant*, i.e., two data matrices produce the same cluster composition if one matrix is obtained from the other one by a monotone increasing transformations of some of its columns/rows (depending on the context)..

Here, we aim at revisiting some clustering procedures which use a suitably defined dissimilarity measure based either on concordance or tail dependence and a clustering algorithm based on the likelihood of the copula, called CoClust, which has been introduced in [13] and further developed and implemented in [14, 15].

4.2 Dissimilarity–based clustering methods

Central to the goals of cluster analysis is the notion of the degree of *similarity* (or proximity) between the individual objects being clustered, which can only come from subject matter considerations. Dissimilarity–based clustering methods provide this information in a direct way by replacing the $(n \times p)$ data matrix $\mathbf{X}$ in eq. (4.1) with a $(p \times p)$ matrix $\Delta = (\Delta_{jj'})$ with the following properties:

- for every $j, j' = 1, \dots, p$, $\Delta_{jj'} \geq 0$;
- for every $j, j' = 1, \dots, p$, $\Delta_{jj'} = \Delta_{j'j}$;
- for every $j = 1, \dots, p$, $\Delta_{jj} = 0$.

Basically, each entry $\Delta_{jj'}$ represents the dissimilarity among the columns $\mathbf{x}_j$ and $\mathbf{x}_{j'}$ in the original data matrix so that the value $\Delta_{jj'}$ is small (close to zero) when $\mathbf{x}_j$ and $\mathbf{x}_{j'}$ are "near" to each other and becomes large when $\mathbf{x}_j$ and $\mathbf{x}_{j'}$ are very different. Notice that, in general, the triangle inequality does not hold for the elements of Δ.

Traditionally, dissimilarity matrices have focused on the use of Pearson correlation (see, for instance, [34, 37]). However, clustering methods based on linear measures of correlation are inadequate to capture association when the involved random variables are not (jointly) elliptical (see, for instance, [39]).

In a copula–based approach, the dissimilarity matrix usually depends on the copula $C_{jj'}$ that can be associated with the column data $\mathbf{x}_j$ and $\mathbf{x}_{j'}$. In particular, it is often assumed that, if $C_{jj'}$ is equal to the comonotonicity copula M, which represents perfect positive dependence among $\mathbf{x}_j$ and $\mathbf{x}_{j'}$, then $\Delta_{jj'} = 0$ (the converse implication may not be true). For an alternative approach see [35].

Clearly, $C_{jj'}$ is unknown and could be replaced by one of its estimates in a parametric/non-parametric setting.

Once a dissimilarity matrix has been obtained, a cluster analysis can be performed by following (at least) two different approaches:

- Apply any classical dissimilarity–based technique (for instance, an agglomerative hierarchical algorithm) by using as input the matrix Δ.
- Perform a non–metric Multidimensional Scaling (MDS) in order to obtain a representation of the dissimilarity matrix into $\mathbf{y}_1, \dots, \mathbf{y}_p$ points in $\mathbb{R}^q$. Then, one may use any classical distance–based algorithm (for instance, K–means algorithm) by using as inputs $\mathbf{y}_1, \dots, \mathbf{y}_p$.

Clearly, both the choice of dissimilarity and the subsequent choice of the clustering algorithm can influence the final clustering composition. Here, we focus our attention on the construction of suitable copula–based dissimilarities and we review some recent methods that can be founded in the literature (for another approach, see also [25]).

Dissimilarity based on measures of concordance.
These methods are derived from the fact that concordance measures can be expressed in terms of the copula associated with the involved random variables (see, e.g., [27, 42, 49, 53]). Specifically, let κ be a concordance measure taking values

in $[-1,1]$. Let $g\colon [-1,1] \to [0,+\infty]$ be a decreasing function such that $g(1)=0$. Then we can define a suitable dissimilarity between the random variables X and Y via the formula

$$diss(X,Y) = g(\kappa(X,Y)).$$

In fact, if X and Y are comonotonic, then $\kappa(X,Y)=1$ and, hence, $diss(X,Y)=0$. Analogously, the other properties of a dissimilarity can be proved using the properties of a measure of concordance.
A popular choice in this context is to adopt $g_1(\kappa)=\sqrt{1-\kappa^2}$, where κ is equal to Kendall's τ or Spearman's ρ. Actually, such a g also defines a pseudo–matrix; see, for instance, [6, 17]. Such methods are also used in vine copula model selection (see. e.g., [8, 16]).
Another popular choice is to use $g_2(\kappa)=\sqrt{1-\kappa}$, where κ is the Spearman's rank correlation coefficient ρ (compare with [3]).
The main difference between g_1 and g_2 is that, while g_1 does not distinguish between comonotonicity and countermonotonicity, g_2 does.

Dissimilarity based on tail–dependence.
Instead of constructing a dissimilarity based on the (global) association between two random variables, one may consider specific features of the joint distribution function, like the tail behavior.
This approach was introduced in [9] and further developed in [10, 11, 12] and [23]. It is based on the *tail dependence coefficients* (either lower, λ_L, or upper, λ_U) associated with a copula C, usually defined as:

$$\lambda_L = \lim_{t\to 0^+} \frac{\delta_C(t)}{t} \quad \text{and} \quad \lambda_U = \lim_{t\to 1^-} \frac{1-2t+\delta_C(t)}{1-t},$$

provided that $\delta_C(t)=C(t,t)$ is the diagonal section of the copula C and the above limits exist.
In this setting, a common approach is to define a dissimilarity between the random variables X and Y via the formula

$$diss(X,Y) = -log(\lambda_L) \quad \text{or} \quad diss(X,Y) = -log(\lambda_U).$$

Thus, if X and Y are comonotonic, then $\lambda_L=\lambda_U=1$ and $diss(X,Y)=0$, while $diss(X,Y)=+\infty$ when X and Y are asymptotically independent, i.e. $\lambda_L=\lambda_U=0$.
Now, while tail dependence coefficients give an asymptotic approximation of the tail behavior of the copula, it might be also of interest to consider the tail behavior at some (finite) points nears the corners of the unit square, as stressed in [18] (see also [52]). To this end, one can consider the so–called *tail concentration function*, defined as the function $q_C\colon (0,1)\to[0,1]$ given by

$$q_C(t) = \frac{\delta_C(t)}{t}\cdot \mathbf{1}_{(0,0.5]} + \frac{1-2t+\delta_C(t)}{1-t}\cdot \mathbf{1}_{(0.5,1)},$$

where $\mathbf{1}_S$ is the indicator function of the set S. See, for instance, [55] and [44, 45]. In [18] suitable distances between q_C and the tail concentration function of the comonotone copula are used to define a dissimilarity.

Another analogous way consists of replacing the tail dependence coefficient by a local concordance measure that focuses on the tail regions of the domain of the joint distribution. In [22], for instance, the conditional Spearman's ρ of (X,Y), given that both variables are below their α–quantile for small values of $\alpha \in (0,1)$, is used and can be interpreted as the degree of dependence between two profit/loss random variables X and Y, provided that both are taking on small values.

Dissimilarity based on risk measures.

For particular purposes, various other dissimilarity measures can be defined using special functionals on the class of copulas, which can be interpreted in terms of risk measures associated with a given system. Examples of such situations are described below.

- In [21], the Kendall distribution function associated with a copula is used to derive a dissimilarity measure. This choice is mainly motivated by the use of Kendall hazard scenarios in defining risky regions in hydrology and environmental sciences (see, for instance, [48]).
- In [19], a contagion index is constructed in order to detect anomalies in the correlation among financial time series during normal and crisis periods. Such an index is then transformed to perform a (fuzzy) clustering of different stock market data.

4.2.1 Illustration: S&P 500 dataset

Here, we illustrate the two above methods based on measures of concordance and tail dependence by means of a time series dataset formed by the constituent data of the S&P 500 in a time window covering the financial crisis of 2007-2008. The end-day prices of all 505 constituents of the S&P 500 are available in the `R` package `qrmdata` (see [31]). As in [32] we consider $T = 756$ daily records from 2007-01-01 to 2009-12-31 on $p = 465$ constituents (which have maximally 20% missing data). Moreover, the data are classified according to the Global Industry Classification Standard (GICS) sector information (the number of companies in each sector is in parenthesis): Consumer discretionary (78), Consumer staples (33), Energy (36), Financials (85), Health care (51), Industrials (63), Information technology (60), Materials (25), Telecommunications services (5), and Utilities (29).

Following the standard approach in financial time series framework (see, e.g., [44]), we fit a suitable marginal model to each of the 755 log-returns from the portfolio of 465 constituents. In particular, the ARMA(1,1)-GARCH(1,1) model with innovations following a Student-t distribution is adopted and the corresponding standardized residuals are extracted. Moreover, the joint distribution function of the standardized residuals $(\varepsilon_{t,1},\dots,\varepsilon_{t,p})$, $t \in \{1,\dots,T\}$, can be expressed as

$$(\varepsilon_{t,1},\dots,\varepsilon_{t,p}) \sim C(F_1,\dots,F_p), \tag{4.2}$$

for a copula C and continuous marginal distributions F_j, $j \in \{1,\dots,p\}$, both assumed to be time–invariant. In a non–parametric approach, the *pseudo-observations* are hence computed via the empirical distribution functions $\hat{F}_{T,1},\dots,\hat{F}_{T,p}$ as

$$\tilde{U}_{t,j} = \frac{T}{T+1}\hat{F}_{T,j}(\widehat{\varepsilon}_{t,j}) = \frac{R_{t,j}}{T+1}, \tag{4.3}$$

for $j \in \{1,\dots,p\}$, where $R_{t,j}$ denotes the rank of $\varepsilon_{t,j}$ among $\varepsilon_{t,1},\dots,\varepsilon_{t,p}$. Now, the multivariate vector of pseudo–observations $(\tilde{U}_{t,1},\dots,\tilde{U}_{t,p})$ takes values on $[0,1]^p$, and can be used to calculate various bivariate measures of association related with the data.

Specifically, we calculate the $p \times p$ matrix of empirical Spearman's rho coefficients $\rho_T^{jj'}$, for $j, j' \in \{1,\dots,p\}$, and use it to determine a dissimilarity measure between the portfolio constituents via the formula

$$diss(\tilde{U}_{t,j},\tilde{U}_{t,j'}) = \sqrt{2(1-\rho_T^{jj'})}. \tag{4.4}$$

Here, the multiplier 2 is used to highlight the differences in the dissimilarity values.

Once the dissimilarity matrix has been computed we apply an agglomerative hierarchical clustering (method= complete linkage). The group composition determined by the hierarchical algorithm is hence compared with the groupings induced by sectors (as given by GICS). That is, we are assuming that the desired number, K, of clusters is equal to the number of sectors ($K = 10$). We adopt the well-known *Adjusted Rand Index* (ARI) [33], which is bounded above by 1 and takes the value 0 when the index equals its expected value, as a measure of agreement between the external criteria and clustering results. The obtained ARI value equals 0.31, which suggests that the constituents in each sector are only partly assigned to the same group, as one could expect.

It is interesting to observe that the difference between GICS sectors and groups determined by statistical analysis becomes more relevant when considering a dissimilarity matrix taking into account the tail dependence among the variables. To this end, let us use in (4.4) the *lower conditional Spearman's* rank-correlation coefficient, defined by conditioning on the pseudo–observations belonging to T_α, defined as $T_\alpha :\, = \{(u,v) \in [0,1]^2 : u \le \alpha, v \le \alpha\}$, where $\alpha \in (0,0.5]$. Then, we repeat the hierarchical clustering procedure for a specific choice of α. For instance, we can set $\alpha \in \{0.25, 0.10\}$, meaning that we are considering those pseudo–observations in the lower left area $[0,0.25]^2$ and $[0,0.10]^2$ of the copula domain, respectively.

The agreement between the agglomerative hierarchical clustering (with complete linkage) and the benchmark classification by sectors is around 9.4%, in terms of ARI, when considering $\alpha = 0.25$; while it gives a value of 2.8%, when the threshold is set to $\alpha = 0.10$. Such findings seem to suggest that the natural grouping of the companies by sectors reflects neither the degree of concordance between the data

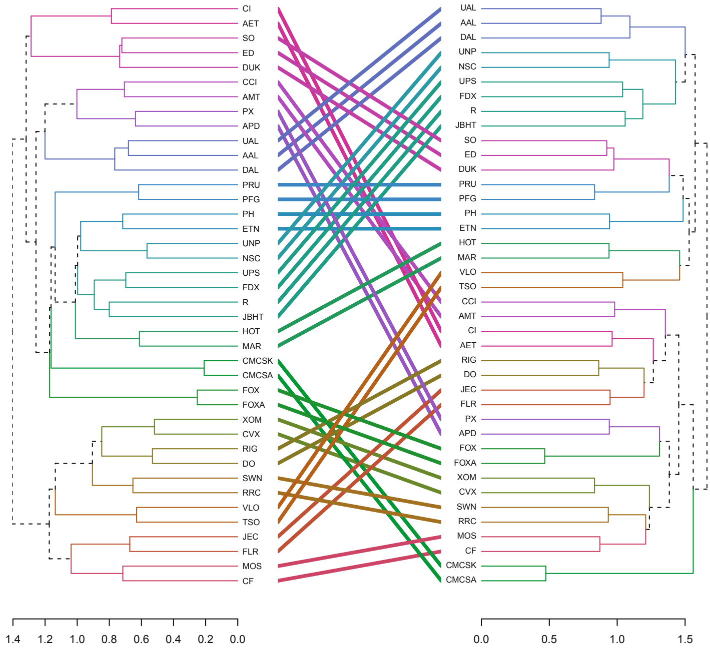


Fig. 4.1: Side by side dendrogram subtrees for S&P 500 constituents, based on agglomerative hierarchical clustering according to Spearman's ρ (left) and lower conditional Spearman's ρ for $\alpha = 0.25$ (right). The subtrees are displayed so that they include only a subset of 40 constituents over 465.

constituents, as measured by Spearman's ρ, nor their tail dependence. Fig. 4.1 can help us see which patterns are somewhat preserved between the two dendrograms based on Spearman's ρ and lower conditional Spearman's ρ, for $\alpha = 0.25$, respectively, by visualizing the subtrees of the two dendrograms related to the variables grouped together in both clustering solutions.

Such findings are in agreement with previous studies emphasising the need of considering specific tools to understand rank-invariant dependencies in risky regions (see, for instance, [10, 20, 22] among others).

4.3 Likelihood–based clustering methods

In this Section we describe and review the methodological framework of cluster analysis based on the likelihood function of a probability model. In this approach to clustering, the population of interest $\mathbf{X}$ given in eq. (4.1) is supposed to consist of K

different subpopulations

$$\mathbf{X} = \begin{bmatrix} x_{11} & \dots & x_{1j} & \dots & x_{1j'} & \dots & x_{1p} \\ \vdots & \dots & \vdots & \dots & \vdots & \dots & \vdots \\ x_{i1} & \dots & x_{ij} & \dots & x_{ij'} & \dots & x_{ip} \\ \vdots & \dots & \vdots & \dots & \vdots & \dots & \vdots \\ x_{i'1} & \dots & x_{i'j} & \dots & x_{i'j'} & \dots & x_{i'p} \\ \vdots & \dots & \vdots & \dots & \vdots & \dots & \vdots \\ x_{n1} & \dots & x_{nj} & \dots & x_{nj'} & \dots & x_{np} \end{bmatrix} = \begin{bmatrix} \mathbf{x}_1 \\ \vdots \\ \mathbf{x}_i \\ \vdots \\ \mathbf{x}_{i'} \\ \vdots \\ \mathbf{x}_n \end{bmatrix} \begin{matrix} \left.\right\} \to \text{1-st subpopulation} \\ \vdots \\ \left.\right\} \to k\text{-th subpopulation} \\ \vdots \\ \left.\right\} \to K\text{-th subpopulation} \end{matrix} \qquad (4.5)$$

and the density of a p-dimensional observation $\mathbf{x}_i$ from the k-th subpopulation is supposed to be $f_k(\mathbf{x}_i;\theta)$ for some unknown (vector of) parameter(s) θ. Given observations $(\mathbf{x}_1,\dots,\mathbf{x}_n)$, we let $\gamma_1,\dots,\gamma_n$ where γ_i denotes the identifying labels, where $\gamma_i = k$ if $\mathbf{x}_i$ comes from the k-th subpopulation. In the maximum likelihood–based clustering procedure, θ and γ are chosen so as to maximize the likelihood $L(\theta,\gamma) = \prod_{i=1}^{n} f_{\gamma_i}(\mathbf{x}_i,\theta)$. In this approach, the generic k-th cluster is defined in terms of a probability density function of the form $f_k(\cdot)$.

In [41, 2] (see also [38, 4]) a clustering strategy based on multivariate normal mixture models with covariances parameterized by eigenvalue decomposition has been developed. For a recent review of mixture model–based clustering we refer to [40]. Here, it is sufficient to recall that the whole population is a mixture model with K components and the likelihood becomes

$$L(\theta_1 \dots, \theta_K; \pi_1, \dots, \pi_K) = \prod_{i=1}^{n} \sum_{k=1}^{K} f_{\pi_k}(\mathbf{x}_i, \theta_k)$$

where π_k is the label for the generic k-th cluster. In this framework, a cluster k is defined in terms of a component within an appropriate finite mixture model. The term "appropriate" here means that the mixture model is tailored to the data under consideration. Gaussian mixture model–based clustering with the EM algorithm for maximum likelihood estimation has been extensively studied and the method has been extended by [26] to select the parameterization of the model as well as the number of clusters simultaneously using the Bayesian Information Criterion (BIC). Nowadays, many deviations from the Gaussian model have been taken into consideration (see, e.g., [40]).

The likelihood–based clustering approach has been used with copula models in several different approaches and contexts (see [13, 1, 5, 36]). [5] developed a network clustering technique based on the likelihood of the copula that finds a partition of objects such that the ones belonging to the same cluster show a dependence structure. The approach in [13] has been then further developed in [15], where we focus our attention. The CoClust assumes that the data are generated by a K-dimensional copula whose arguments $F_1,\dots,F_k,\dots,F_K$ are the probability-integral transforms of the density functions that generate the clusters; thus, in this approach, each cluster is defined in terms of a (marginal) univariate density probability function. It is important to stress that here the definitions of "cluster" and "clustering" are somehow different from the classical ones. Indeed, even though the assumptions of a probability model to generate a cluster and a multivariate probability model to generate the whole final clustering are largely accepted, two main differences can be found: i) a cluster is a set of independent observations, i.e. independent and identically distributed realizations of a univariate model and ii) in the final clustering the interest focuses on the among-group relationship, which is the multivariate dependence, and not on the within-group relationship, which is the independence.

The starting point of the CoClust algorithm is the standard $n \times p$ data matrix in eq. (4.1) in which n objects have to be grouped in K groups and there are p segmentation variables. The algorithm can be applied to the row or column data matrix according to the purpose of the analysis. Here, the CoClust is described as applied to the rows of the data matrix. The basic idea behind the CoClust consists in a forward procedure that allocates a K-plet of row data matrix, i.e. a p-dimensional vector for each cluster, at a time and the allocation of each K-plet of rows is performed on the basis of the log-likelihood of the copula fit. This likelihood is computed by using the K-plets already allocated and the one candidate to the allocation, say $\mathbf{x}_{i'} = (x_{i'1}, \dots, x_{i'j}, \dots, x_{i'p})$, by varying the permutations of observations in $\mathbf{x}'_i$ in order to find, if it exists, the combination that maximizes the copula fit.

The main steps of the algorithm to cluster n row data matrix are described:

1. for $k = K_{min}, \dots, K_{max}$, where K_{min} and K_{max} are, respectively, the minimum and the maximum number of clusters to be tried, such that $2 \le K_{min} \le K_{max} \le n$,

   a. select a subset of n_k k-plets of rows in the data matrix in eq. (4.1) on the basis of the following multivariate measure of association based on pairwise Spearman's ρ correlation coefficient:

   $$H(\Lambda_2|\Lambda_1) = \max_{i' \in \Lambda_2} \left\{ \psi_{i \in \Lambda_1} \left(\rho(\mathbf{x}_i, \mathbf{x}_{i'}) \right) \right\} \tag{4.6}$$

   where Λ is a set of row index vectors such that $\Lambda = \Lambda_1 \cup \Lambda_2$, where Λ_1 is the subset of vectors already selected to compose a k-plet and Λ_2 is the set of remaining candidates to complete it, and ψ is an average aggregation function (see [29]), for instance, the mean, the median or the maximum;

   b. fit on the n_k k-plets of rows the copula model through the maximum pseudo-likelihood estimation [28, 54] that estimates each univariate margins through the empirical cumulative distribution functions $\hat{F}_k(X_k)$ with $k = 1, \dots, K$, and, then, the maximum likelihood is used to estimate the copula parameter as follows

   $$\hat{\theta} = \arg\max_{\theta} \sum_{i=1}^{n_k} \log c \left\{ \hat{F}_1(X_{1i}), \dots, \hat{F}_k(X_{ki}); \theta \right\}; \tag{4.7}$$

2. select the subset of n_k k-plets of rows, say n_K K-plets, that maximizes the log–likelihood of the copula; hence, the number of clusters K, that is the dimension of the copula, is automatically chosen;
3. select a K-plet of rows (among the remaining ones) on the basis of the eq. (4.6) and estimate $K!$ copulas by using the observations already clustered and a permutation of the candidate to the allocation;
4. allocate the permutation of the selected K-plet to the clustering by assigning each row to the corresponding cluster only if it increases the log-likelihood of the copula fit, otherwise drop the entire K-plet of rows;
5. repeat steps 3. and 4. until all the observations are evaluated (either allocated or discarded).

The main purpose of the CoClust is to identify dependent groups in such a way that the complex dependence among observations can be uncovered. Hence, at the end of the procedure we obtain a clustering of K dependent clusters each one contained at maximum $(n/K)p$ independent observations.

As described above, the CoClust algorithm selects automatically the number of clusters K on the basis of the log-likelihood of the copula estimated on the subsets of k-plets allocated until a step predefined by the user. On the contrary, the selection of the copula model is not automatic, but an information criterion should be employed. The Bayesian information criterion (BIC, from now on) has the following expression for a K-dimensional copula model m with s independent parameters:

$$\mathrm{BIC}_{K,m} = -2\log \Pi_{i=1}^{n} c_m \left\{\hat{F}_1(X_{1i}),\ldots,\hat{F}_k(X_{ki}),\ldots,\hat{F}_K(X_{Ki});\hat{\theta}\right\} + s\log\left((n/K)p\right) \tag{4.8}$$

where $\hat{\theta}$ is as in eq. (4.7) with the summation over the number of allocated observations, which equals maximum $(n/K)p$ (i.e. n/K p-dimensional vectors). According to [46], we select the copula model that minimizes the BIC (for possible alternative approaches, see also [30]).

The CoClust has been implemented in an R package available on the CRAN [14]. The main function `CoClust` makes it possible to perform the cluster analysis by employing Elliptical and Archimedean copula models and setting the set of numbers of clusters to be tried, the number n_k of observations to be used for the selection of K and the kind of function Ψ in eq. (4.6).

4.3.1 Illustration: NCI60 data

The National Cancer Institute's (NCI) Developmental Therapeutics Program (DTP) has carried out intensive studies of 60 cancer cell lines derived from tumours from a variety of tissues and organs [47]. Here, the interest focuses on the NCI60 data set which contains the gene expression profiles of 60 human tumour cell lines derived from patients with leukaemia (LEUK), melanoma (MELAN), non-small colon lung (NSCLC), colon (COLON), central nervous system (CNS), ovarian (OVAR), renal (RENAL), breast (BREAST) and prostate (PROSTATE) cancers. This panel of cell lines have been subjected to several different DNA microarray studies using different array technologies, like spotted cDNA [47] and Affimetrix [51]. The data set employed here is available in the R package `made4` and contains subsets from one cDNA spotted which has been pre-processed as described by [7].

The NCI60 data set contains 144 gene expression (log-ratio measurements) rows and 60 cell line columns. We apply the CoClust to the cell lines with the main purpose of evaluating the capability of the CoClust to recognize the organ specific of each tumour type. Since [47] found that the gene expression patterns of cell lines derived from non-small lung carcinoma, breast and prostate tumours are heterogeneous, we apply the CoClust to the 41 cell lines derived from central nervous system, colon, leukaemia, melanoma, ovarian and renal tissues. For the distribution of the cell lines per organ type see Tab. 4.1. The CoClust has been applied by varying the number of clusters from $K_{min} = 2$ to $K_{max} = 8$ and copula models among all the copulas in the Elliptical and Archimedean families. As for the t-Student copula, we evaluated the clustering by using both $\nu = 2$ and $\nu = 4$ degrees of freedom. Moreover, the value of n_k has been set to the minimum integer value n/K_{max}, which is 5.

Table 4.1: Number of tumours for each kind of organ.

Organ type	CNS	COLON	LEUK	MELAN	OVAR	RENAL
N. of tissues	6	7	6	8	6	8

The copula model and the number of clusters of the final clustering have been selected on the basis of the BIC as explained in Section 4.3. The selected copula model is a 7-dimensional t-Student with $\nu = 4$ degrees of freedom, estimated dependence parameter $\hat{\theta} = 0.651$ (which is equal to $\rho = 0.633$) and log-likelihood of the fit equals 1911.04. The obtained clustering is shown in Tab. 4.2. On the one hand, the CoClust has been able to almost perfectly recognize LEUK, MELAN, COLON and RENAL tumour types, so gene expressions of cell lines with common presumptive tissues of origin are strongly dependent to each other. On the other hand, OVAR and CNS tumour types are not well-identified and six cell lines has been dropped out of the clustering; specifically, they are 1 MELAN, 2 CNS, and 3 OVAR cell lines. It could be possible that CNS and OVAR belong to a DGP with a dependence structure different from the t-Student one.

4.4 Empirical comparison of the two approaches

In order to briefly illustrate a possible application of both the proposed methodologies and its comparison we present here a case study on environmental data. The data were collected by “Ufficio Idrografico” of the province of Bozen-Bolzano and

Table 4.2: Identification of cancer types through CoClust algorithm. 7-dimensional t-Student copula model with $\nu = 4$ and $\hat{\theta} = 0.651$.

$\mathscr{C}_1$	$\mathscr{C}_2$	$\mathscr{C}_3$	$\mathscr{C}_4$	$\mathscr{C}_5$	$\mathscr{C}_6$	$\mathscr{C}_7$
LEUK	LEUK	COLON	LEUK	LEUK	LEUK	LEUK
MELAN	MELAN	MELAN	MELAN	MELAN	MELAN	MELAN
RENAL	RENAL	CNS	CNS	CNS	OVAR	CNS
COLON	COLON	COLON	COLON	COLON	OVAR	COLON
RENAL	RENAL	RENAL	RENAL	OVAR	RENAL	RENAL

are available online. They are related to daily rainfall measurements recorded at 18 gauge stations spread across the province of Bolzano-Bozen in the North-Eastern Italy, from 1961 to 2010. This results in a set of $p = 18$ time series originally formed by $T = 18262$ observations. Tab. 4.3 reports the available information on the analysed rainfall records.

Table 4.3: Summary of the rainfall measurement stations.

Code	Station	Longitude	Latitude	Height (mt)
0220	S.VALENTINO ALLA MUTA	10.5277	46.7745	1520
0310	TUBRE	10.4775	46.6503	1119
2090	PLATA	11.1783	46.8225	1147
3140	FLERES	11.3477	46.9639	1246
3260	VIPITENO-CONVENTO	11.4295	46.8978	948
8320	BOLZANO	11.3127	46.4976	254
9150	SESTO	12.3477	46.7035	1310
0250	MONTE MARIA	10.5213	46.7057	1310
0480	MAZIA	10.6175	46.6943	1570
1580	VERNAGO	10.8493	46.7357	1700
2170	S.LEONARDO PASSIARIA	11.2471	46.8091	644
2670	PAVICOLO	11.1093	46.6278	1400
3450	RIDANNA	11.3068	46.9091	1350
4450	S.MADDALENA IN CASIES	12.2427	46.8353	1398
6650	FUNDRES	11.7029	46.8872	1159
8570	BRONZOLO	11.3111	46.4065	226
8730	REDAGNO	11.3968	46.3465	1562
4100	ANTERIVO	11.3678	46.2773	1209

From these time series, in order to focus on extreme observations, annual maxima at each spatial location are extracted resulting in a 50×18 matrix of time series observations $\tilde{X}_1^i, \dots, \tilde{X}_p^i$, $i \in \{1, \dots, 50\}$, summarized by Fig. 4.2.

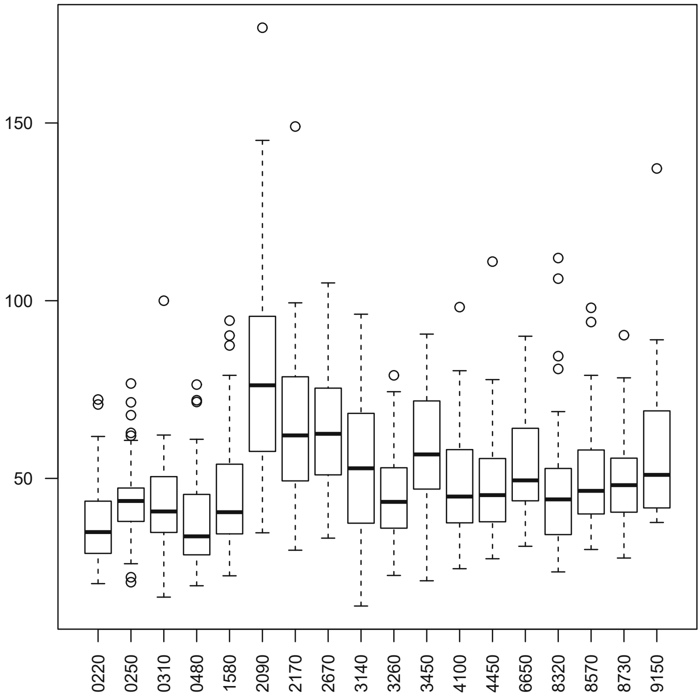


Fig. 4.2: Boxplot of annual maxima at each station from 1961 to 2010. The station codes are as Tab. 4.3. On the y-axis the amount of rainfall is measured in millimeters.

4.4.1 Clustering of rainfall data based on tail dependence

As shown in Section 4.2, the dependence between the time series of annual maxima can be studied in terms of a copula–based measure of dependence estimated from the pseudo-observations in $[0,1]^p$, $(\tilde{U}_{i,1},\ldots,\tilde{U}_{i,p})$, $i \in \{1,\ldots,50\}$. Being interested in exploring the tail behaviour of the data, we estimate the upper tail dependence coefficient λ_U of the pair $(\tilde{U}_{i,j},\tilde{U}_{i,j'})$, for all $j,j' \in \{1,\ldots,p\}$, $j<j'$. Specifically, the function `fitLambda` of the `copula` package computes non-parametric estimators of the matrix of tail dependence coefficients. By setting `method="t"` we fit a t-Student copula and consider the implied tail-dependence coefficient (see [39]). In the sequel, we will denote by $\hat{\lambda}^U_{jj'}$ the pairwise coefficients of upper tail dependence.

Such matrix can be transformed through a monotonic function in order to obtain a dissimilarity measure between the 18 time series. As previously discussed, we can consider the matrix $\Delta = (\Delta_{ij})$ with elements

$$\Delta_{jj'} = -\log(\hat{\lambda}^U_{jj'}). \tag{4.9}$$

The result of hierarchical agglomerative clustering is presented in Fig. 4.3 (here, the complete linkage method is adopted). In the literature, a wide variety of methods

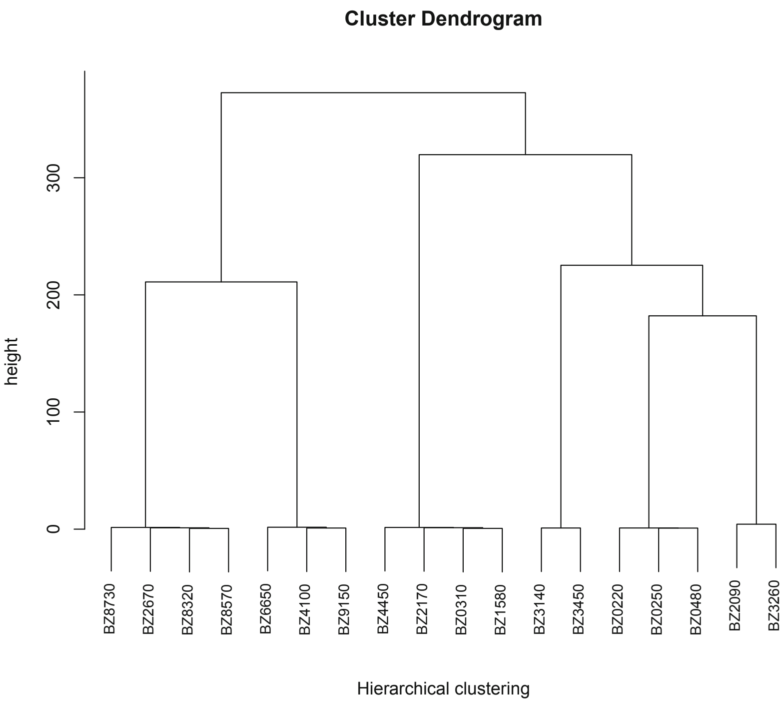


Fig. 4.3: Dendrogram for the 18 rainfall measurement stations listed in Tab. 4.3 based on the complete linkage method.

Table 4.4: Hierarchical clustering of rainfall measurement stations.

$\mathscr{C}_1$	BZ0220	BZ0250	BZ0480	BZ2090	BZ3140	BZ3260	BZ3450
$\mathscr{C}_2$	BZ0310	BZ1580	BZ2170	BZ4450			
$\mathscr{C}_3$	BZ2670	BZ4100	BZ6650	BZ8320	BZ8570	BZ8730	BZ9150

have been proposed to find the optimal number of clusters in a partitioning of a data set. Here, we select $K = 3$ in order to have a clustering comparable with the one obtained through the CoClust algorithm and shown in the next Section.

4.4.2 Clustering of rainfall data based on the CoClust algorithm

In this Section, we apply the CoClust algorithm to the precipitation data described in Section 4.4. As in Section 4.3, the algorithm has been applied by varying the copula models among all the copulas in the Elliptical and Archimedean families (the t-Student copula has been used with two different levels of degree of freedom $\nu = 2$ and $\nu = 4$), whereas the number of clusters has been automatically selected in the range from $K_{min} = 2$ to $K_{max} = 6$ by using $n_k = 3$ observations. The copula model and the number of clusters of the final clustering have been selected on the basis of the BIC as explained in Section 4.3.

The obtained clustering is shown in Tab. 4.5. The selected model is a 5-dimensional Gaussian copula with $\hat{\theta} = 0.555$ and log-likelihood equals 142.9247. Note that the three stations: BZ0310, BZ1580, and BZ3140, have been dropped out of the clustering. The geographical parameter appears to be a quite valid criterion for evaluating

Table 4.5: CoClust results on the rainfall measurement stations.

$\mathscr{C}_1$	$\mathscr{C}_2$	$\mathscr{C}_3$	$\mathscr{C}_4$	$\mathscr{C}_5$
BZ0250	BZ0220	BZ0480	BZ8320	BZ6650
BZ2670	BZ3260	BZ3450	BZ2170	BZ2090
BZ8730	BZ4100	BZ4450	BZ8570	BZ9150

the resulting clustering. Indeed, the second row contains stations of Alta Val d'Isarco and Burgraviato, which are really close to each other, and the third row contains stations located in the south and east part of the land. Only the first row does not reflect the geographical criterion because there are three stations in Val Venosta, which is in the north-west of Alto-Adige, and two stations in the center of the region.

4.4.3 Comparing the two clusterings of rainfall data

In order to compare the two clusterings shown in the previous Sections, we should find similarities and differences between the rows of tables 4.4 and 4.5. To summarize them, Figure 4.4 shows and compares the two obtained clustering solutions: for a given row i and column j in the graphical matrix, a black point in the corresponding cell means that stations i and j belong to the same cluster according to the clustering based on tail dependence, while an empty square means that stations i and j are dependent to each other and belong to different clusters in the method based on the log-likelihood of the copula fit; the presence of both the point and the square means that the two stations have been identified as "similar/associated/dependent" by both the clustering procedures. The figure shows a certain agreement between the two findings, particularly for the clusters in the lower and upper corner of the matrix. In the middle of the matrix, on the contrary, there is a non-overlapping area that confirm the differences between the two approaches.

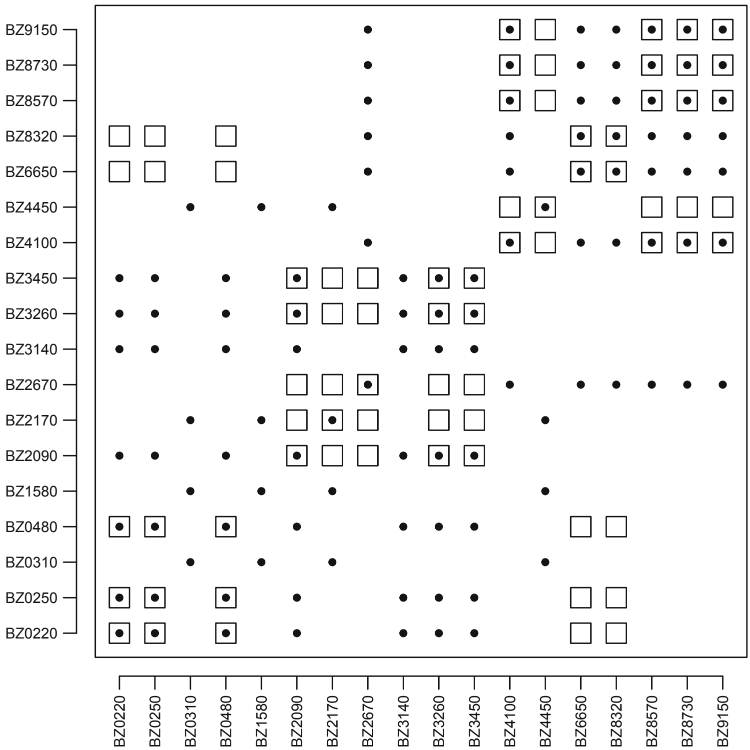


Fig. 4.4: Comparison between tail–based and likelihood–based clustering results.

4.5 Conclusions

Two different approaches to clustering based on copula and dependence measures have been discussed. Firstly, dissimilarity measures based on concordance or tail dependence have been presented and used to perform a cluster analysis of financial time series. Secondly, a copula-based clustering algorithm that uses the log-likelihood of the copula fit has been described and an application to gene expression profiles has been provided. Both the methods take into consideration the dependence between variables but in a different way. The former aims at creating groups of variables characterized by a higher global association or a similar tail behavior. The CoClust, on the contrary, attempts to discover clusters of variables dependent to each other, whereas each cluster is a set of independent variables, so the interest focuses on the among-group relationship. A case study from environmental data has been performed in order to compare the two approaches in practice.

Acknowledgements On the occasion of his 75th birthday, we would like to thank Roger Nelsen for writing his famous book on copulas. During the preparation of our PhD thesis (and later on) we have spent plenty of hours reading and studying "An Introduction to Copulas". We are indebted to him for having provided a valuable reference book for our research.

The first author acknowledge the support of Free University of Bozen-Bolzano, Faculty of Economics and Management, via the project "Correlation, Clustering, and Copulas: Detecting association in multivariate data" (COCCO). The second author acknowledge the support of Free University of Bozen-Bolzano, Faculty of Economics and Management, via the project "Aggregation functions for Innovation and Data Analysis" (AIDA). The second author also acknowledges the support of the grant MTM2014-60594-P (partially supported by FEDER) from the Spanish "Ministerio de Economía y Competitividad".

References

1. Arakelian, V., Karlis, D.: Clustering dependencies via mixtures of copulas. Comm. Statist. Simulation Comput. **43**(7), 1644–1661 (2014)
2. Banfield, J.D., Raftery, A.E.: Model-based gaussian and non-gaussian clustering. Biometrics **49**, 803–821 (1993)
3. Bonanno, G., Caldarelli, G., Lillo, F., Miccichè, S., Vandewalle, N., Mantegna, R.: Networks of equities in financial markets. Eur. Phys. J. B **38**(2), 363–371 (2004)
4. Celeux, G., Govaert, G.: Gaussian parsimonious clustering models. Pattern Recognition **28**, 781–793 (1995)
5. Chessa, A., Crimaldi, I., Riccaboni, M., Trapin, L.: Cluster analysis of weighted bipartite networks: A new copula-based approach. PLoS ONE **9**(10), e109,507 (2014)
6. Côté, M.P., Genest, C.: A copula–based risk aggregation model. Canad. J. Statist. **43**(1), 60–81 (2015)
7. Culhane, A.C., Perrière, G., Higgins, D.G.: Cross-platform comparison and visualisation of gene expression data using co-inertia analysis. BMC Bioinformatics **21**, 4–59 (2003)
8. Czado, C., Schepsmeier, U., Min, A.: Maximum likelihood estimation of mixed C-vines with application to exchange rates. Stat. Model. **12**(3), 229–255 (2012)
9. De Luca, G., Zuccolotto, P.: A tail dependence-based dissimilarity measure for financial time series clustering. Adv. Data Anal. Classif. **5**(4), 323–340 (2011)
10. De Luca, G., Zuccolotto, P.: Time series clustering on lower tail dependence for portfolio selection. In: M. Corazza, C. Pizzi (eds.) Mathematical and Statistical Methods for Actuarial Sciences and Finance, pp. 131–140. Springer, Berlin (2014)
11. De Luca, G., Zuccolotto, P.: Dynamic tail dependence clustering of financial time series. Statistical Papers, in press (2017)
12. De Luca, G., Zuccolotto, P.: A double clustering algorithm for financial time series based on extreme events. Stat. Risk Model., in press (2017)
13. Di Lascio, F.M.L., Giannerini, S.: A copula-based algorithm for discovering patterns of dependent observations. Journal of Classification **29**(1), 50–75 (2012)
14. Di Lascio, F.M.L., Giannerini, S.: CoClust. R package version 0.3-1 (2015). URL http://CRAN.R-project.org/package=CoClust
15. Di Lascio, F.M.L., Giannerini, S.: Clustering dependent observations with copula functions. Statistical Papers pp. 1–17 (2016). DOI 10.1007/s00362-016-0822-3
16. Dißmann, J., Brechmann, E.C., Czado, C., Kurowicka, D.: Selecting and estimating regular vine copulae and application to financial returns. Comput. Statist. Data Anal. **59**, 52–69 (2013)

17. van Dongen, S., Enright, A.J.: Metric distances derived from cosine similarity and Pearson and Spearman correlations. ArXiv e-prints (2012). URL http://arxiv.org/abs/1208.3145
18. Durante, F., Fernández-Sánchez, J., Pappadà, R.: Copulas, diagonals and tail dependence. Fuzzy Sets and Systems **264**, 22–41 (2015)
19. Durante, F., Foscolo, E.: An analysis of the dependence among financial markets by spatial contagion. Int. J. Intell. Syst. **28**(4), 319–331 (2013)
20. Durante, F., Foscolo, E., Jaworski, P., Wang, H.: A spatial contagion measure for financial time series. Expert Syst. Appl. **41**(8), 4023–4034 (2014)
21. Durante, F., Pappadà, R.: Cluster analysis of time series via Kendall distribution. In: P. Grzegorzewski, M. Gagolewski, O. Hryniewicz, M. Gil (eds.) Strengthening Links Between Data Analysis and Soft Computing, *Advances in Intelligent Systems and Computing*, vol. 315, pp. 209–216. Springer International Publishing (2015)
22. Durante, F., Pappadà, R., Torelli, N.: Clustering of financial time series in risky scenarios. Adv. Data Anal. Classif. **8**, 359–376 (2014)
23. Durante, F., Pappadà, R., Torelli, N.: Clustering of time series via non–parametric tail dependence estimation. Statist. Papers **56**(3), 701–721 (2015)
24. Durante, F., Sempi, C.: Principles of copula theory. CRC Press, Boca Raton, FL (2016)
25. D'Urso, P., Disegna, M., Durante, F.: Copula-based fuzzy clustering of time series. In: F. Mola, C. Conversano (eds.) Book of Abstracts of the 10th Scientific Meeting of the Classification and Data Analysis Group, p. 4. CUEC, Cagliari (2015)
26. Fraley, C., Raftery, A.E.: How many clusters? which clustering method? answers via model–based cluster analysis. The Computer Journal **41**(8), 578–588 (1998)
27. Fuchs, S.: Copula–Induced Measures of Concordance. Depend. Model. **4**(1), 205–214 (2016)
28. Genest, C., Ghoudi, K., Rivest, L.P.: A semiparametric estimation procedure of dependence parameters in multivariate families of distributions. Biometrika **82**, 543–552 (1995)
29. Grabisch, M., Marichal, J.L., Mesiar, R., Pap, E.: Aggregation Functions. Encyclopedia of Mathematics and its Applications (No. 127). Cambridge University Press, New York (2009)
30. Grønneberg, S., Hjort, N.L.: The copula information criteria. Scand. J. Stat. **41**(2), 436–459 (2014)
31. Hofert, M., Hornik, K.: qrmdata: Data Sets for Quantitative Risk Management Practice (2016). URL https://CRAN.R-project.org/package=qrmdata. R package version 2016-01-03-1
32. Hofert, M., Oldford, W.: Visualizing dependence in high-dimensional data: An application to S&P 500 constituent data. ArXiv e-prints (2016). URL http://arxiv.org/abs/1609.09429
33. Hubert, L., Arabie, P.: Comparing partitions. J. Classif. **2**, 193–218 (1985)
34. Kaufman, L., Rousseeuw, P.: Finding groups in data. Wiley Series in Probability and Mathematical Statistics: Applied Probability and Statistics. John Wiley & Sons Inc., New York (1990)
35. Kojadinovic, I.: Hierarchical clustering of continuous variables based on the empirical copula process and permutation linkages. Comput. Statist. Data Anal. **54**(1), 90–108 (2010)
36. Kosmidis, I., Karlis, D.: Model-based clustering using copulas with applications. Stat. Comput. **26**(5), 1079–1099 (2016)
37. Mantegna, R.N.: Hierarchical structure in financial markets. Euro. Phys. J. B **11**(1), 193–197 (1999)
38. McLachlan, G.J., Basford, K.E.: Mixture Models: Inference and Applications to Clustering. New York: Marcel Dekker (1988)
39. McNeil, A.J., Frey, R., Embrechts, P.: Quantitative risk management. Concepts, techniques and tools. Princeton Series in Finance. Princeton University Press, Princeton, NJ (2005)
40. McNicholas, P.D.: Model-based clustering. Journal of Classification **33**(3), 331–373 (2016)
41. Murtagh, F., Raftery, A.E.: Fitting straight lines to point patterns. Pattern Recognition **17**, 479–483 (1984)
42. Nelsen, R.B.: Concordance and copulas: a survey. In: C.M. Cuadras, J. Fortiana, J.A. Rodríguez Lallena (eds.) Distributions with given marginals and Statistical Modelling, pp. 169–178. Kluwer, Dordrecht (2003)
43. Nelsen, R.B.: An Introduction to Copulas, second edn. Springer Series in Statistics. Springer, New York (2006)

44. Patton, A.J.: A review of copula models for economic time series. J. Multivariate Anal. **110**, 4–18 (2012)
45. Patton, A.J.: Copula methods for forecasting multivariate time series. In: G. Elliott, A. Timmermann (eds.) Handbook of Economic Forecasting, vol. 2, pp. 899–960. Elsevier, Oxford (2013)
46. Raftery, A., Nema, D.: Variable selection for model-based clustering. Journal of the American Statistical Association **101**(473), 168–178 (2006)
47. Ross, D.T., Scherf, U., Eisen, M.B., Perou, C.M., Rees, C., Spellman, P., Iyer, V., Jeffrey, S.S., Van de Rijn, M., Waltham, M., Pergamenschikov, A., Lee, J.C., Lashkari, D., Shalon, D., Myers, T.G., Weinstein, J.N., Botstein, D., Brown, P.O.: Systematic variation in gene expression patterns in human cancer cell lines. Nature Genetics **24**, 227–235 (2000)
48. Salvadori, G., Durante, F., De Michele, C., Bernardi, M., Petrella, L.: A multivariate copula-based framework for dealing with hazard scenarios and failure probabilities. Water Resour. Res. **52**(5), 3701–3721 (2016)
49. Scarsini, M.: On measures of concordance. Stochastica **8**(3), 201–218 (1984)
50. Sklar, A.: Fonctions de répartition à *n* dimensions et leurs marges. Publ. Inst. Statist. Univ. Paris **8**, 229–231 (1959)
51. Staunton, J.E., Slonim, D.K., Coller, H.A., Tamayo, P., Angelo, M.J., Park, J., Scherf, U., Lee, J.K., Reinhold, W.O., Weinstein, J.N., J.P., M., Lander, E., Golub, T.R.: Chemosensitivity prediction by transcriptional profiling. Proceedings of the National Academy of Sciences of the USA **98**, 10,787–10,792 (2001)
52. Sweeting, P., Fotiou, F.: Calculating and communicating tail association and the risk of extreme loss. British Actuarial Journal **18**, 13–72 (2013)
53. Taylor, M.D.: Multivariate measures of concordance for copulas and their marginals. Depend. Model. **4**(1), 224–236 (2016)
54. Tsukahara, H.: Semiparametric estimation in copula models. The Canadian Journal of Statistics **33**(3), 357–375 (2005)
55. Venter, G.G.: Tails of copulas. In: Proceedings ASTIN Washington, USA, pp. 68–113 (2001)

Chapter 5
Copula-based piecewise regression

Arturo Erdely

Abstract Most common parametric families of copulas are totally ordered, and in many cases they are also positively or negatively regression dependent and therefore they lead to monotone regression functions, which makes them not suitable for dependence relationships that imply or suggest a non-monotone regression function. A gluing copula approach is proposed to decompose the underlying copula into totally ordered copulas that once combined may lead to a non-monotone regression function.

5.1 Introduction

Given a bivariate random vector (X,Y) with joint probability distribution function $F_{X,Y}(x,y) = \mathbb{P}(X \leq x, Y \leq y)$ it is possible to assess uncertainty about one of the random variables conditioning on certain values of the other, for example through the univariate conditional probability distribution of Y given $X = x$, that is $F_{Y|X}(y\,|\,x) = \mathbb{P}(Y \leq y\,|\,X = x)$. As a point estimate for a future value of Y given $X = x$ we may calculate central tendency measures with $F_{Y|X}$ such as the mean (whenever it exists) or the median (which always exists in the continuous case) which will depend on the conditioning value x and therefore such point estimates depending on x may be denoted by $\mu(x)$ and are called *regression function* for Y given $X = x$.

As a consequence of *Sklar's Theorem* [9] for continuous random variables there exists a unique copula C such that the joint probability distribution function $F_{X,Y}(x,y) = C(F_X(x), F_Y(y))$ where $F_X(x) = \mathbb{P}(X \leq x)$ and $F_Y(y) = \mathbb{P}(Y \leq y)$ are the marginal probability distribution functions of X and Y, respectively. As explained in [5], the conditional distribution of Y given $X = x$ can be obtained by

Arturo Erdely
Universidad Nacional Autónoma de México, Facultad de Estudios Superiores Acatlán, e-mail: arturo.erdely@comunidad.unam.mx

M. Úbeda Flores et al. (eds.), *Copulas and Dependence Models with Applications*,
DOI 10.1007/978-3-319-64221-5_5

$$F_{Y|X}(y\,|\,x) = \frac{\partial C(u,v)}{\partial u}\bigg|_{u=F_X(x)\,,\,v=F_Y(y)} \tag{5.1}$$

and therefore to find the median regression function for Y given $X = x$ whenever $F_{Y|X}$ is a continuous distribution function, we proceed as follows:

Algorithm 1

1. Set $\partial C(u,v)/\partial u = 1/2$;
2. solve for the regression function of $V = F_Y(Y)$ given $U = F_X(X) = u$, and obtain $v = \psi(u)$;
3. replace u by $F_X(x)$ and v for $F_Y(y)$;
4. solve for the regression function of Y given $X = x$:

$$y = \mu(x) = F_Y^{(-1)}(\psi(F_X(x))). \tag{5.2}$$

It is worth to notice that since F_X and F_Y only explain the individual (marginal) probabilistic behavior of the continuous random variables X and Y, respectively, then the information about their dependence for regression purposes is contained in ψ. A survey of copula-based regression models may be found in [3] and estimation/inference procedures for such purpose in [6].

5.2 Piecewise monotone regression

In [1] it is argued that when the regression function is non-monotone, copula-based regression estimates do not reproduce the qualitative features of the regression function under commonly used parametric copula families. This occurs because very often such parametric copulas lead to monotone regression functions, but in case there is evidence that the underlying regression function is non-monotone a *piecewise regression* approach may be applied in order to break up a non-monotone relationship into a piecewise monotonic one, and then seek for the best copula fit for each piece.

Piecewise (or segmented) monotone regression for Y given $X = x$ is defined by partitioning the support of X into a finite number of intervals such that restricted to each one it is possible to obtain a monotone regression function. For example, instead of (5.1) we may obtain something like

$$F_{Y|X}(y\,|\,x) = \begin{cases} \frac{\partial}{\partial u} C_1(u,v)\Big|_{u=F_{X|X\leq b}(x)\,,\,v=F_Y(y)}, & \text{if } x \leq b, \\ \frac{\partial}{\partial u} C_2(u,v)\Big|_{u=F_{X|X>b}(x)\,,\,v=F_Y(y)}, & \text{if } x > b, \end{cases} \tag{5.3}$$

with C_1 and C_2 two different copulas, b is called a *break-point* for explanatory variable X, and where $F_{X|X\leq b}$ and $F_{X|X>b}$ are the conditional distribution functions of X given $X \leq b$ and $X > b$, respectively. This may be justified in terms of the *gluing copula* technique [8] as explained in [2] for the particular case of vertical section gluing and bivariate copulas. Specifically, given two bivariate copulas C_1 and C_2, and a fixed value $0 < \theta < 1$ (gluing point), we may scale C_1 to $[0,\theta] \times [0,1]$ and C_2 to $[\theta,1] \times [0,1]$ and *glue* them into a single copula

$$C_{1,2,\theta}(u,v) = \begin{cases} \theta C_1(\frac{u}{\theta},v), & 0 \leq u \leq \theta, \\ (1-\theta)C_2(\frac{u-\theta}{1-\theta},v) + \theta v, & \theta \leq u \leq 1. \end{cases} \tag{5.4}$$

Then

$$\frac{\partial}{\partial u} C_{1,2,\theta}(u,v) = \begin{cases} \frac{\partial}{\partial u} C_1(\frac{u}{\theta},v), & 0 \leq u \leq \theta, \\ \frac{\partial}{\partial u} C_2(\frac{u-\theta}{1-\theta},v), & \theta \leq u \leq 1, \end{cases} \tag{5.5}$$

and by (5.1)

$$\begin{aligned} F_{Y|X}(y\,|\,x) &= \left.\frac{\partial C_{1,2,\theta}(u,v)}{\partial u}\right|_{u=F_X(x)\,,\,v=F_Y(y)} \\ &= \begin{cases} \frac{\partial}{\partial u} C_1(\frac{F_X(x)}{\theta},F_Y(y)), & 0 \leq F_X(x) \leq \theta, \\ \frac{\partial}{\partial u} C_2(\frac{F_X(x)-\theta}{1-\theta},F_Y(y)), & \theta \leq F_X(x) \leq 1, \end{cases} \\ &= \begin{cases} \left.\frac{\partial}{\partial u} C_1(u,v)\right|_{u=F_{X|X\leq b}(x)\,,\,v=F_Y(y)}, & x \leq F_X^{(-1)}(\theta) = b, \\ \left.\frac{\partial}{\partial u} C_2(u,v)\right|_{u=F_{X|X> b}(x)\,,\,v=F_Y(y)}, & x > F_X^{(-1)}(\theta) = b, \end{cases} \end{aligned} \tag{5.6}$$

since $F_{X|X\leq b}(x) = \mathbb{P}(X \leq x\,|\,X \leq b) = \mathbb{P}(X \leq x)/\mathbb{P}(X \leq b) = F_X(x)/\theta$ and $F_{X|X>b}(x) = \mathbb{P}(b < X \leq x)/\mathbb{P}(X > b) = (F_X(x)-\theta)/(1-\theta)$. The result obtained in (5.6) leads to a regression function of the form

$$\mu(x) = \begin{cases} \mu_1(x), & \text{if } x \leq b, \\ \mu_2(x), & \text{if } x > b, \end{cases} \tag{5.7}$$

where, for example, if $\mu_1(x)$ is an increasing function and $\mu_2(x)$ a decreasing one, then $\mu(x)$ is non-monotone.

Example 5.1. From example 3.3 in [5] if a probability mass $0 < \theta < 1$ is uniformly distributed on the line segment joining $(0,0)$ to $(\theta,1)$, and a probability mass $1-\theta$ is uniformly distributed on the line segment joining $(\theta,1)$ to $(1,0)$, see Fig. 5.1, the underlying copula for a random vector (X,Y) of continuous Uniform$(0,1)$ random variables with such non-monotone dependence is given by

$$C_\theta(u,v) = \begin{cases} u, & 0 \leq u \leq \theta v \leq \theta, \\ \theta v, & 0 \leq \theta v < u < 1-(1-\theta)v, \\ u+v-1, & \theta \leq 1-(1-\theta)v \leq u \leq 1. \end{cases} \tag{5.8}$$

By construction we have that $\mathbb{P}(Y = \frac{x}{\theta}\,|\,X = x) = 1$ whenever $0 \leq x \leq \theta$ and $\mathbb{P}(Y = \frac{1-x}{1-\theta}\,|\,X = x) = 1$ whenever $\theta \leq x \leq 1$, which implies that the regression function of Y given $X = x$ is

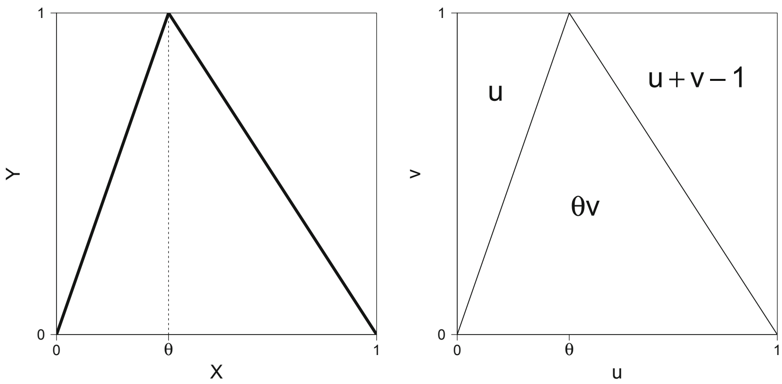


Fig. 5.1: Example 5.1. Left: (X,Y) dependence. Right: underlying copula (5.8).

$$\mu(x) = \begin{cases} \frac{x}{\theta}, & 0 \leq x \leq \theta, \\ \frac{1-x}{1-\theta}, & \theta \leq x \leq 1, \end{cases} \tag{5.9}$$

clearly a non-monotone function: linearly increasing for $0 \leq x \leq \theta$ and linearly decreasing for $\theta \leq x \leq 1$, which suggests in this case that the underlying dependence might be split by means of the gluing copula technique in terms of two copulas, with θ as gluing point. Indeed, let $C_1(u,v) = \min\{u,v\}$ (the Fréchet-Hoeffding upper bound that represents the case when one variable is almost surely an increasing function of the other) and $C_2(u,v) = \max\{u+v-1,0\}$ (the Fréchet-Hoeffding lower bound that represents the case when one variable is almost surely a decreasing function of the other), then applying (5.4) it is straightforward to verify that the resulting gluing copula $C_{1,2,\theta}$ is equal to (5.8).

Therefore, the same regression function obtained in (5.9) could be obtained in two pieces: the first one in terms of the random vector (X_1,Y) with underlying copula C_1 and where the distribution of X_1 is the conditional distribution of X given $X \leq \theta$, which turns to be uniform$(0,\theta)$, and the second one in terms of the random vector (X_2,Y) with underlying copula C_2 and where the distribution of X_2 is the conditional distribution of X given $X > \theta$, which turns to be uniform$(\theta,1)$. Applying (5.1) to the first piece we obtain the following:

$$\begin{aligned} F_{Y|X_1}(y\,|\,x) &= \frac{\partial}{\partial u} C_1(u,v)\Big|_{u=\frac{x}{\theta},\,v=y} \\ &= \begin{cases} 1, & \text{if } y \geq \frac{x}{\theta}, \\ 0, & \text{if } y < \frac{x}{\theta} \end{cases} \end{aligned} \tag{5.10}$$

from which we get $\mu_1(x) = \frac{x}{\theta}$ whenever $0 \leq x \leq \theta$, and similarly from $F_{Y|X_2}(y\,|\,x)$ we obtain $\mu_2(x) = \frac{1-x}{1-\theta}$ whenever $\theta \leq x \leq 1$, as expected. ■

For simplicity's sake, the case for a single break-point has been analyzed, but the analogous idea may be applied for finitely many break-points. For each interval I induced in the support of the explanatory variable, the conditional distribution of Y given $X = x$ is obtained by

$$F_{Y|X}(y\,|\,x) = \frac{\partial}{\partial u}C_I(u,v)\Big|_{u=F_{X|X\in I}(x)\,,\,v=F_Y(y)} \tag{5.11}$$

and with it the regression function $\mu(x)$ for $x \in I$ may be calculated.

5.3 Dependence and regression

In this section the concepts of quadrant and regression dependence by [4] are recalled.

Definition 5.1. A bivariate random vector (X,Y) or its joint distribution function $F_{X,Y}$ is *positively quadrant dependent* and abbreviated as $\mathrm{PQD}(X,Y)$ if

$$\mathbb{P}(X \le x, Y \le y) \ge \mathbb{P}(X \le x)\mathbb{P}(Y \le y)\,, \qquad \text{for all } x \text{ and } y, \tag{5.12}$$

and *negatively quadrant dependent* $\mathrm{NQD}(X,Y)$ if (5.12) holds with the inequality sign reversed.

In the particular case where both X and Y are continuous random variables with underlying copula C, as an immediate consequence of *Sklar's Theorem* [9] we have that $\mathrm{PQD}(X,Y)$ is equivalent to $C(u,v) \ge uv$ for all u,v in $[0,1]$, and $\mathrm{NQD}(X,Y)$ with this last inequality sign reversed. From [5] we have the following:

Definition 5.2. If C_1 and C_2 are copulas, we say that C_1 is *smaller than* C_2 (or C_2 *is larger than* C_1), and write $C_1 \prec C_2$ (or $C_2 \succ C_1$) if $C_1(u,v) \le C_2(u,v)$ for all u,v in $[0,1]$.

This point-wise partial ordering of the set of copulas is called *concordance ordering*. It is a partial order rather than a total order because not every pair of copulas is comparable. However, there are families of copulas that are totally ordered. We will call a totally ordered parametric family $\{C_\theta\}$ of copulas *positively ordered* if $C_\alpha \prec C_\beta$ whenever $\alpha \le \beta$; and *negatively ordered* if $C_\alpha \succ C_\beta$ whenever $\alpha \le \beta$. Many of well known one-parameter families of copulas are totally ordered and include $\Pi(u,v) = uv$, and hence have subfamilies of PQD and NQD copulas.

As mentioned in [5] one form to calculate *Spearman's concordance measure* is

$$\rho_C = 12\iint_{[0,1]^2}\big[C(u,v) - uv\big]\,dudv = 12\iint_{[0,1]^2} C(u,v)\,dudv - 3\,, \tag{5.13}$$

and hence $\rho_C/12$ can be interpreted as a measure of "average" quadrant dependence (both positive and negative) for continuous random variables whose copula is C. Closely related to (5.13) is the L_1 distance between C and the (sometimes called) independence copula $\Pi(u,v) = uv$ known as *Schweizer-Wolff's dependence measure* [7] defined as

$$\sigma_C = 12 \iint_{[0,1]^2} \left|C(u,v) - uv\right| dudv. \tag{5.14}$$

Two main differences (among others) are that $-1 \leq \rho_C \leq 1$ in contrast to $0 \leq \sigma_C \leq 1$, and that $\sigma_C = 0$ if and only X and Y are independent (that is $C = \Pi$) while $\rho_C = 0$ does not necessarily imply independence. Moreover, as explained in [5]:

> Of course, it is immediate that if X and Y are PQD, then $\sigma_{X,Y} = \rho_{X,Y}$; and that if X and Y are NQD, then $\sigma_{X,Y} = -\rho_{X,Y}$. Hence for many of the totally ordered families of copulas presented in earlier chapters (e.g., Plackett, Farlie-Gumbel-Morgenstern, and many families of Archimedean copulas), $\sigma_{X,Y} = |\rho_{X,Y}|$. But for random variables X and Y that are neither PQD nor NQD, i.e., random variables whose copulas are neither larger nor smaller than Π, σ is often a better measure than ρ [...]

Definition 5.3. A random variable Y is *positively regression dependent* on a random variable X and abbreviated as PRD$(Y|X)$ if

$$F_{Y|X}(y\,|\,x) = \mathbb{P}(Y \leq y\,|\,X = x) \quad \text{is non-increasing in } x, \tag{5.15}$$

and *negatively regression dependent* NRD$(Y|X)$ if (5.15) is non-decreasing in x.

From theorems 5.2.4 and 5.2.12 in [5] or from Lemma 4 in [4] we have the following:

Corollary 5.1. *Given (X,Y) a bivariate random vector:*

a) If PRD$(Y|X)$ then PQD(X,Y).
b) If NRD$(Y|X)$ then NQD(X,Y).

By arguments explained in [5] the reverse implications in Corollary 5.1 do not necessarily hold.

Corollary 5.2. *If (X,Y) are continuous random variables with underlying copula C then:*

a) PRD$(Y|X)$ if and only if for any v in $[0,1]$ and for almost all u, $\partial C(u,v)/\partial u$ is non-increasing in u;
b) NRD$(Y|X)$ if and only if for any v in $[0,1]$ and for almost all u, $\partial C(u,v)/\partial u$ is non-decreasing in u.

In case the conditional expectation exists it is possible to obtain a *mean regression function*

$$\mu(x) = \mathbb{E}(Y\,|\,X = x) = \int_0^\infty [1 - F_{Y|X}(y\,|\,x)]\,dy - \int_{-\infty}^0 F_{Y|X}(y\,|\,x)\,dy, \tag{5.16}$$

and in case $F_{Y|X}(y\,|\,x)$ is a continuous function of y then it is possible to obtain a *median regression function*

$$\mu(x) = \text{median}(Y\,|\,X = x) = F_{Y|X}^{(-1)}(0.5\,|\,x)\,. \tag{5.17}$$

Proposition 5.1. *Let $\mu(x)$ be a mean or median regression function:*

a) If PRD$(Y|X)$ then $\mu(x)$ is non-decreasing.
b) If NRD$(Y|X)$ then $\mu(x)$ is non-increasing.

Proof. If PRD$(Y|X)$ then for all $x_1 < x_2$

$$-F_{Y|X}(y\,|\,x_1) \leq -F_{Y|X}(y\,|\,x_2)\,, \tag{5.18}$$

$$1 - F_{Y|X}(y\,|\,x_1) \leq 1 - F_{Y|X}(y\,|\,x_2)\,. \tag{5.19}$$

Integration of (5.18) on $]-\infty,0]$ and of (5.19) on $[0,\infty[$, and adding the results according to the inequalities it is obtained $\mu(x_1) = \mathbb{E}(Y\,|\,X = x_1) \leq \mathbb{E}(Y\,|\,X = x_2) = \mu(x_2)$, as required. Now from (5.18) we have $F_{Y|X}(y\,|\,x_1) \geq F_{Y|X}(y\,|\,x_2)$, and since $F_{Y|X}(y\,|\,x)$ is non-decreasing in y for any x so is $F_{Y|X}^{(-1)}(u\,|\,x)$ as a function of u and therefore $F_{Y|X}^{(-1)}(u\,|\,x_1) \leq F_{Y|X}^{(-1)}(u\,|\,x_2)$, hence $\mu(x_1) = \text{median}(Y\,|\,X = x_1) = F_{Y|X}^{(-1)}(0.5\,|\,x_1) \leq F_{Y|X}^{(-1)}(0.5\,|\,x_2) = \text{median}(Y\,|\,X = x_2) = \mu(x_2)$, as required. ∎

But the reverse implications in this last proposition do not necessarily hold, as it can be easily verified by similar arguments.

Example 5.2. Continuing with Example 5.1, applying formulas (5.13) and (5.14) it is straightforward to verify that Spearman's $\rho_\theta = 2\theta - 1$ and Schweizer-Wolff's $\sigma_\theta = \theta^2 + (\theta - 1)^2$, and since $0 < \theta < 1$ then $|\rho_\theta| < \sigma_\theta$ and therefore neither we have PQD nor NQD, and neither PRD nor NRD. Moreover, if $\theta = \frac{1}{2}$ then $\rho_{1/2} = 0$ but this does not imply independence since $\sigma_{1/2} = \frac{1}{2}$ (its minimum possible value, by the way). See Fig. 5.2 (left). ∎

5.4 Change-point detection

The ideas explained in the previous sections may be useful in tackling the concerns raised by [1] when the dependence relationship between random variables implies a non-monotone regression function, considering that the most common families of parametric copulas lead to monotone regression functions, and a possible solution might be to break up such dependence into pieces such that within each one the dependence implies a piecewise monotone regression function, and possibly one of the common families of parametric copulas may have an acceptable fit for each piece. In pursuing this objective, when dealing with data from which the dependence has to be estimated, a methodology to find break-point candidates, that is *change-point detection,* becomes necessary.

Definition 5.4. The *diagonal section* of a copula C is a function $\delta_C : [0,1] \to [0,1]$ given by $\delta_C(t) = C(t,t)$.

Since every copula C is bounded by the Fréchet-Hoeffding bounds $\max\{u+v-1,0\} \leq C(u,v) \leq \min\{u,v\}$ then $\max\{2t-1,0\} \leq \delta_C(t) \leq t$. If $C = \Pi$ (independence copula) then $\delta_\Pi(t) = t^2$. If (X,Y) is a random vector of continuous random variables with underlying copula C and $\mathrm{PQD}(X,Y)$ or $\mathrm{NQD}(X,Y)$ then $C(u,v) \geq uv$ or $C(u,v) \leq uv$, respectively, for all (u,v) in $[0,1]^2$, and therefore $\delta_C(t) \geq t^2$ or $\delta_C(t) \leq t^2$, respectively, for all t in $[0,1]$. Hence, if there exist t_1, t_2 in $[0,1]$ such that $\delta_C(t_1) < t_1^2$ and $\delta_C(t_2) > t_2^2$ then neither $\mathrm{PQD}(X,Y)$ nor $\mathrm{NQD}(X,Y)$, and consequently this would imply that neither $\mathrm{PRD}(Y|X)$ nor $\mathrm{NRD}(Y|X)$. In case of this last scenario, this would not necessarily imply that a mean or median regression function $\mu(x)$ is non-monotone since Proposition 5.1 is a one-way implication, but at least raises the question and leads to propose and analyze break-point candidates. The following result is straightforward:

Proposition 5.2. *Let C_1 and C_2 be two copulas such that $C_1(u,v) \geq uv$ and $C_2(u,v) \leq uv$ for all $(u,v) \in [0,1]^2$, and let $0 < \theta < 1$. Then the diagonal section of the resulting gluing copula $C_{1,2,\theta}$ as in (5.4) satisfies*

$$\delta_{1,2,\theta}(t) \begin{cases} \geq t^2, & \text{if } 0 \leq t \leq \theta, \\ = \theta^2, & \text{if } t = \theta, \\ \leq t^2, & \text{if } \theta \leq t \leq 1. \end{cases} \tag{5.20}$$

Since the diagonal section δ_C of any copula C is a continuous function, see [5], we may choose and analyze as possible break-point candidates those where crossings between δ_C and δ_Π take place.

Example 5.3. Continuing with Example 5.1, from formula (5.8) the corresponding diagonal section is:

$$\delta_\theta(t) = C_\theta(t,t) = \begin{cases} \theta t\,, & t \leq \frac{1}{2-\theta}\,, \\ 2t-1\,, & t \geq \frac{1}{2-\theta}\,. \end{cases} \tag{5.21}$$

If $0 < t \leq \frac{1}{2-\theta}$ then $\delta_\theta(t) \geq t^2$ if and only if $t \leq \theta$. If $\frac{1}{2-\theta} \leq t < 1$ then $\delta_\theta(t) \leq t^2$. Since $0 < \theta < 1$ then $\theta < \frac{1}{2-\theta}$ and therefore we conclude that $\delta_\theta(t) \geq t^2$ if and only if $t \leq \theta$, and $\delta_\theta(t) \leq t^2$ if and only if $t \geq \theta$. Hence, we would propose $t = \theta$ as break-point candidate, as expected. See Fig. 5.2 (right). ■

Example 5.4. This is one of the examples used in [1] to raise concerns about the use of copulas when the dependence relationship between random variables implies

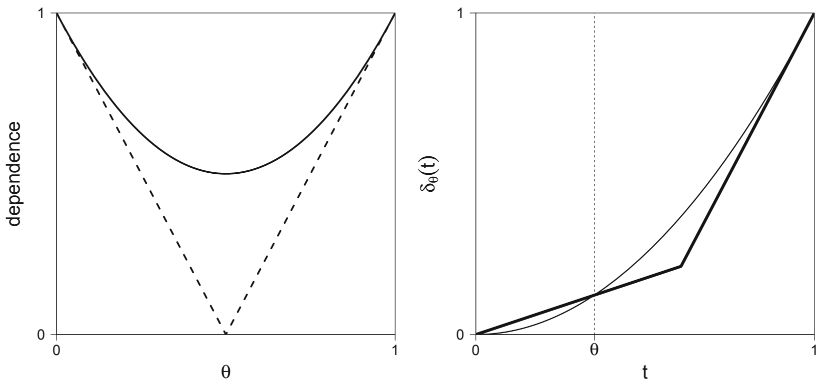


Fig. 5.2: Left: $|\rho_\theta|$ (dashed line) and σ_θ (continuous line) in Example 5.2. Right: δ_θ (thick line) and δ_Π (thin line) in Example 5.3.

a non-monotone regression function. Let ε be a Normal$(0,1)$ random variable, a constant $k^2 = 0.01$, and X a Uniform$(0,1)$ random variable independent from ε. Now define the random variable:

$$Y = (X - 0.5)^2 + k\varepsilon\,. \tag{5.22}$$

Then the conditional distribution of Y given $X = x$ is Normal$\big((x-0.5)^2, k^2\big)$ and therefore the corresponding mean regression function is given by:

$$\mu(x) = \mathbb{E}(Y\,|\,X=x) = (x-0.5)^2, \quad 0 \le x \le 1, \tag{5.23}$$

clearly a non-monotone regression function (decreasing when $x \le 0.5$, increasing when $x \ge 0.5$). Since the joint probability density of (X,Y) is given by $f_{X,Y}(x,y) = f_X(x)f_{Y|X}(y\,|\,x)$ then:

$$\begin{aligned} F_{X,Y}(x,y) &= \int_{-\infty}^{x} f_X(r)\int_{-\infty}^{y} f_{Y|X}(s\,|\,x)\,ds\,dr = \int_{-\infty}^{x} f_X(r)F_{Y|X}(y\,|\,r)\,dr \\ &= \begin{cases} 0, & \text{if } x \le 0, \\ \displaystyle\int_0^x \Phi\Big(\frac{y-(r-0.5)^2}{k}\Big)\,dr, & \text{if } 0 < x < 1, \\ \displaystyle\int_0^1 \Phi\Big(\frac{y-(r-0.5)^2}{k}\Big)\,dr, & \text{if } x \ge 1, \end{cases} \end{aligned} \tag{5.24}$$

where Φ is the distribution function for a Normal$(0,1)$ random variable. From (5.24) it is possible obtain the following expression for the marginal distribution function of Y :

$$F_Y(y) = F_{X,Y}(+\infty,y) = \int_0^1 \Phi\Big(\frac{y-(r-0.5)^2}{k}\Big)\,dr. \tag{5.25}$$

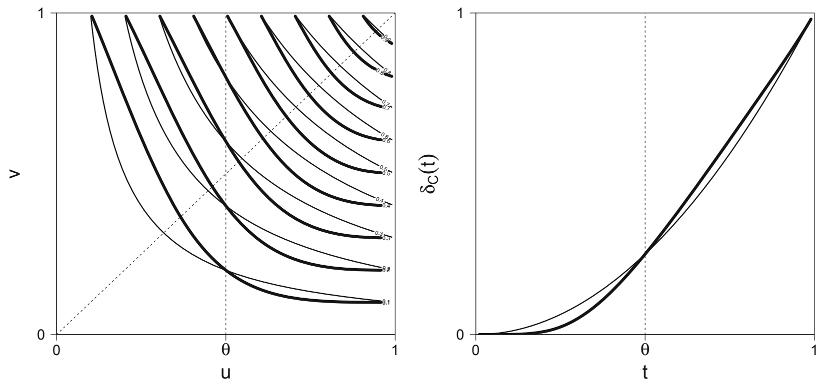


Fig. 5.3: Example 5.4. Left: Level curves (thick style) of copula (5.26) versus level curves (thin style) of product (or independence) copula. Right: Diagonal section (thick style) of copula (5.26) versus diagonal section (thin style) of product (or independence) copula.

Hence, by Sklar's Corollary 2.3.7 in [5] it is possible to obtain the following expression for the underlying copula of (X,Y) :

$$C(u,v) = F_{X,Y}\big(F_X^{(-1)}(u), F_Y^{(-1)}(v)\big) = \int_0^u \Phi\Big(\frac{F_Y^{-1}(v) - (r-0.5)^2}{k}\Big)\, dr, \quad (5.26)$$

and consequently the diagonal section of such copula is given by:

$$\delta_C(t) = C(t,t) = \int_0^t \Phi\Big(\frac{F_Y^{-1}(t) - (r-0.5)^2}{k}\Big)\, dr. \quad (5.27)$$

In Fig. 5.3 (left) we may notice crossings between copula (5.26) level curves (thick style) and the product (or independence) copula $\Pi(u,v) = uv$ level curves (thin style), with the following interpretation: thick curve below thin curve implies $C(u,v) \geq \Pi(u,v)$ and thick curve above thin curve implies $C(u,v) \leq \Pi(u,v)$. In Fig. 5.3 (right) the graph of the diagonal section (5.27) is compared to the graph of the diagonal section of Π from where we get as gluing point candidate $u = \theta = 1/2$.

Then we proceed to a *gluing copula decomposition* by means of (5.4) where $C_{1,2,\theta} = C$. For $0 \leq u \leq \theta$ we get $\theta C_1(\frac{u}{\theta}, v) = C(u,v)$, and if we let $u_* = \frac{u}{\theta} \in [0,1]$ then:

$$C_1(u_*, v) = \frac{1}{\theta} C(\theta u_*, v) = 2\int_0^{u_*/2} \Phi\Big(\frac{F_Y^{-1}(v) - (r-0.5)^2}{k}\Big)\, dr, \quad (5.28)$$

and therefore:

$$\frac{\partial}{\partial u_*} C_1(u_*, v) = \Phi\Big(\frac{F_Y^{-1}(v) - 0.25(1-u_*)^2}{k}\Big), \quad (5.29)$$

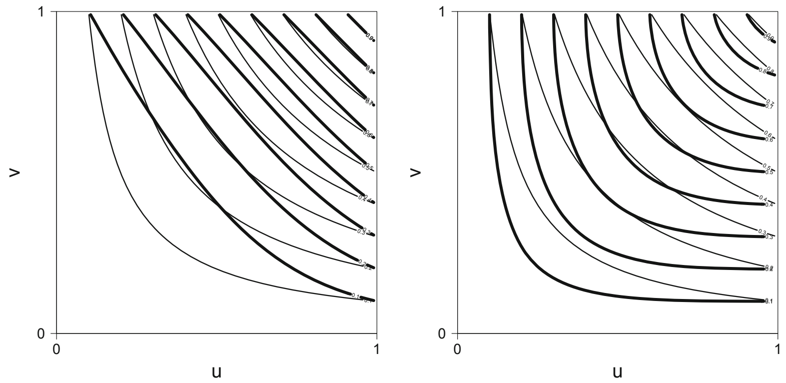


Fig. 5.4: Example 5.4. Left: Level curves (thick style) of copula (5.28) versus level curves (thin style) of product (or independence) copula. Right: Level curves (thick style) of copula (5.4) versus level curves (thin style) of product (or independence) copula.

where clearly (5.29) is a non-decreasing function of u_* which by Corollary 5.2 implies NRD for copula C_1, and consequently NQD by Corollary 5.1. Also, by Proposition 5.1 we get that a regression function $\mu_1(x)$ based on C_1 will lead to a non-increasing function of x. See Fig. 5.4 (left) for the level curves of C_1 (thick style) versus the level curves (thin lines) of $\Pi(u,v) = uv$, where all the level curves of C_1 are above the corresponding ones to Π implying that $C_1(u,v) \leq \Pi(u,v)$, as expected.

Similarly, for $\theta \leq u \leq 1$ we get $(1-\theta)C_2(\frac{u-\theta}{1-\theta}, v) + \theta v = C(u,v)$ and if we let $u_* = \frac{u-\theta}{1-\theta} \in [0,1]$ then:

$$
\begin{aligned}
C_2(u_*, v) &= \frac{C((1-\theta)u_* + \theta, v) - \theta v}{1-\theta} \\
&= 2\int_0^{(u_*+1)/2} \Phi\Big(\frac{F_Y^{-1}(v) - (r-0.5)^2}{k}\Big)\, dr - v\,, \qquad (5.30)
\end{aligned}
$$

and therefore:

$$
\frac{\partial}{\partial u_*} C_2(u_*, v) = \Phi\Big(\frac{F_Y^{-1}(v) - 0.25u_*^2)}{k}\Big), \qquad (5.31)
$$

where clearly (5.31) is a non-increasing function of u_* which by Corollary 5.2 implies PRD for copula C_2, and consequently PQD by Corollary 5.1. Also, by Proposition 5.1 we get that a regression function $\mu_2(x)$ based on C_2 will lead to a non-decreasing function of x. See Fig. 5.4 (right) for the level curves of C_2 (thick style) versus the level curves (thin lines) of $\Pi(u,v) = uv$, where all the level curves of C_2 are below the corresponding ones to Π implying that $C_2(u,v) \geq \Pi(u,v)$, as expected.

In summary, the dependence between X and Y induced by (5.22), which by construction has a regression function $\mu(x)$ that is non-monotone, has an underlying copula C given by (5.26) with a diagonal section δ_C given by (5.27) that gives as gluing point candidate $\theta = 1/2$, leading to a gluing copula decomposition as in (5.4) where C_1 is NQD and NRD and therefore leads to a non-increasing regression function $\mu_1(x)$, and where C_2 is PQD and PRD and therefore leads to a non-decreasing regression function $\mu_2(x)$, that is:

$$\mu(x) = \begin{cases} \mu_1(x) \downarrow, & u = F_X(x) = x \leq \theta = 1/2, \\ \mu_2(x) \uparrow, & u = F_X(x) = x \geq \theta = 1/2. \end{cases} \tag{5.32}$$

In this example it was possible to obtain a gluing copula decomposition as in (5.4) of the underlying copula C into C_1 and C_2 being these last two copulas NQD and PQD, respectively, and therefore candidates to be approximated by well known totally ordered families of copulas. ■

5.5 Final remarks

If (X,Y) is a bivariate random vector of continuous random variables with an underlying copula C such that $|\rho_C| < \sigma_C$ then C is neither PQD nor NQD and therefore neither PRD nor NRD. Many of well known parametric families of copulas are totally ordered (that is, PQD and/or NQD) and in such case they have to be discarded as admissible copulas for (X,Y). To face this challenge, in the present work it has been proposed a *gluing copula decomposition* of C into totally ordered copulas that combined may lead to a non-monotone regression function.

Acknowledgements The present work was partially supported by project IN115817 from Programa de Apoyo a Proyectos de Investigación e Innovación Tecnológica (PAPIIT) at Universidad Nacional Autónoma de México.

References

1. Dette, H., Van Hecke, R., Volgushev, S.: Some comments on copula-based regression. J. Amer. Statist. Assoc. **109**(507), 1319–1324 (2014)
2. Erdely, A., Diaz-Viera, M.: Nonparametric and semiparametric bivariate modeling of petrophysical porosity-permeability dependence from well log data. In: P. Jaworski, F. Durante, W. Härdle, T. Rychlik (eds.) Copula Theory and Its Applications, pp. 267–278. Springer, Berlin (2010)
3. Kolev, N., Paiva, D.: Copula-based regression models: A survey. J. Statist. Plan. Inference **139**, 3847–3856 (2009)
4. Lehmann, E.: Some concepts of dependence. Ann. Math. Statist. **37**, 1137–1153 (1966)
5. Nelsen, R.B.: An Introduction to Copulas, second edn. Springer Series in Statistics. Springer, New York (2006)

6. Noh, H., El Ghouch, A., Bouezmarni, T.: Copula-based regression estimation and inference. J. Amer. Statist. Assoc. **108**(502), 676–688 (2013)
7. Schweizer, B., Wolff, E.: On nonparametric measures of dependence for random variables. Ann. Statist. **9**, 879–885 (1981)
8. Siburg, K., Stoimenov, P.: Gluing copulas. Commun. Statist. Theory Methods **37**(19), 3124–3134 (2008)
9. Sklar, A.: Fonctions de répartition à *n* dimensions et leurs marges. Publ. Inst. Statist. Univ. Paris **8**, 229–231 (1959)

Chapter 6
When Gumbel met Galambos

Christian Genest and Johanna G. Nešlehová

Abstract The well known bivariate Gumbel and Galambos copulas have analytical forms whose similarity is intriguing. As explored here, several deep connections indeed exist between these two parametric families of copulas in any dimension.

6.1 Introduction

Emil Gumbel (1891–1966) and János Galambos (1940–) are two early contributors to extreme-value theory who wrote influential books on the subject [10, 11, 16]. While they probably never met, they are frequently cited together in the statistical literature on extremes, such as [1, 5], through families of multivariate dependence structures that bear their name. In the context of copula theory, which Roger Nelsen promoted and developed through research and the two editions of his book [28, 30], the Gumbel and Galambos parametric families of copulas are well known.

The bivariate Gumbel copula with parameter $\rho \in (0,1)$ is defined, for all $u_1, u_2 \in (0,1)$, by

$$\mathrm{GU}_\rho(u_1,u_2) = \exp[-\{|\ln(u_1)|^\rho + |\ln(u_2)|^\rho\}^{1/\rho}]. \tag{6.1}$$

The limiting cases $\rho = 0$ and $\rho = 1$ correspond respectively to the Fréchet–Hoeffding upper bound M and the independence copula Π, where for all u_1, u_2, $M(u_1,u_2) = \min(u_1,u_2)$ and $\Pi(u_1,u_2) = u_1 u_2$.

As for the bivariate Galambos copula with parameter $\rho \in (0,\infty)$, it is defined, for all $u_1, u_2 \in (0,1)$, by

The original version of this chapter was revised. An erratum to this chapter can be found at https://doi.org/10.1007/978-3-319-64221-5_16

Christian Genest
Department of Mathematics and Statistics, McGill University, 805, rue Sherbrooke ouest, Montréal (Québec) Canada H3A 0B9 e-mail: christian.genest@mcgill.ca

Johanna G. Nešlehová
Department of Mathematics and Statistics, McGill University, 805, rue Sherbrooke ouest, Montréal (Québec) Canada H3A 0B9 e-mail: johanna.neslehova@mcgill.ca

M. Úbeda Flores et al. (eds.), *Copulas and Dependence Models with Applications*,
DOI 10.1007/978-3-319-64221-5_6

$$\mathrm{GA}_\rho(u_1,u_2) = \frac{u_1u_2}{\exp[-\{|\ln(u_1)|^{-1/\rho}+|\ln(u_2)|^{-1/\rho}\}^{-\rho}]}. \tag{6.2}$$

The limiting cases $\rho = 0$ and $\rho = \infty$ correspond to the Fréchet–Hoeffding upper bound M and the independence copula Π, respectively.

Beyond the fact that both families belong to the class of extreme-value copulas, they have expressions that seem tantalizingly similar. The purpose of this contribution to Roger Nelsen's Festschrift is to unveil some of the deep connections that exist between these two families of bivariate copulas and the following d-variate extensions thereof, defined for arbitrary $u_1,\ldots,u_d \in (0,1)$, by

$$\mathrm{GU}_\rho(u_1,\ldots,u_d) = \exp\Big[-\Big\{\sum_{i=1}^{d}|\ln(u_i)|^\rho\Big\}^{1/\rho}\Big],$$

$$\mathrm{GA}_\rho(u_1,\ldots,u_d) = \exp\Big[\sum_{A\subseteq\{1,\ldots,d\},\,A\neq\emptyset}(-1)^{|A|}\Big\{\sum_{i\in A}|\ln(u_i)|^{-1/\rho}\Big\}^{-\rho}\Big].$$

In Section 6.2 we explore the Gumbel family from the perspective of Archimedean copulas. In Section 6.3 we embed the Galambos family in the reciprocal Archimedean class, newly introduced in [13]. Finally, Section 6.4 highlights the similarities between the two families when viewed as members of the extreme-value copula class.

6.2 Gumbel copulas as Archimedean copulas

Copulas of the form (6.1) were introduced implicitly by Gumbel, almost as an afterthought, at the end of a paper studying new bivariate exponential distributions [17]. These copulas are easily seen to be Archimedean [28, 30, 36] and they are in fact the only copulas that are both extreme-value and Archimedean [14].

In arbitrary dimension $d \geq 2$, a copula is called Archimedean if it can be expressed, for all $u_1,\ldots,u_d \in [0,1]$, in the form

$$C_\psi(u_1,\ldots,u_d) = \psi\{\psi^{-1}(u_1)+\cdots+\psi^{-1}(u_d)\}, \tag{6.3}$$

where $\psi : [0,\infty) \to [0,1]$ is d-monotone on $(0,\infty)$ and such that $\psi(0) = 1$ and $\psi(x) \to 0$ as $x \to \infty$; see [24, 26]. By convention, $\psi^{-1}(0) = \inf\{x \in \mathbb{R} : \psi(x) = 0\}$. Following [38], a function f is said to be d-monotone on $(0,\infty)$ if for all $k \in \{0,\ldots,d-2\}$, the kth derivative of f exists everywhere on $(0,\infty)$ and satisfies $(-1)^k f^{(k)} \geq 0$, and if $(-1)^{d-2} f^{(d-2)}$ is non-increasing and convex.

The Gumbel copula is of the form (6.3) in every dimension $d \geq 2$. Its generator is defined, for all $x \in [0,\infty)$, by $\psi_\rho(x) = \exp(-x^\rho)$. This generator is completely monotone and has been identified as the Laplace transform of a variable T_ρ whose distribution γ_ρ is positive stable with parameter $\rho \in (0,1)$ and unit scaling [8]. The copula GU_ρ may thus be regarded as the dependence structure of a multiplicative hazard model [25, 31]. It was actually introduced as such by Hougaard [21], independently of Gumbel's work. Specifically, if $X_1,\ldots,X_d$ are mutually independent unit exponential random variables that are independent of the frailty T_ρ, GU_ρ is the survival copula of the random vector $(X_1,\ldots,X_d)/T_\rho$. This fact can be used, e.g., for sampling [25].

The Gumbel generator ψ_ρ is also the survival function of the Weibull distribution with shape parameter ρ and unit scaling. By Sklar's Theorem for survival functions, the mapping defined, for all $x_1,\ldots,x_d \in [0,\infty)$, by

$$\mathrm{GU}_\rho\{\psi_\rho(x_1),\ldots,\psi_\rho(x_d)\} = \psi_\rho(x_1+\cdots+x_d)$$

is thus a d-variate survival function with Weibull margins. Because this survival function depends on $x_1,\ldots,x_d$ only through their sum, it is Schur-constant.

As recognized by Nelsen [29], the survival copula of a Schur-constant survival function is necessarily Archimedean. The converse is also true, as shown in [26]. In fact, distributions on $[0,\infty)^d$ with Schur-constant survival function coincide with the class of ℓ_1-norm symmetric distributions. In particular, GU_ρ is the survival copula of the random vector $R_\rho \times (S_1,\ldots,S_d)$, where $(S_1,\ldots,S_d)$ is uniformly distributed on the unit simplex $\mathbb{S}_d = \{(w_1,\ldots,w_d) \in [0,1] : w_1+\cdots+w_d = 1\}$ and independent of the strictly positive radial random variable R_ρ. Put differently, the Gumbel copula is the dependence structure of a resource sharing model, in which a random resource R_ρ is distributed equitably among $d \geq 2$ agents [27].

The distribution function of R_ρ is given through the inverse of the Williamson d-transform of ψ_ρ; see [26]. That is, for all $r \in (0,\infty)$,

$$\Pr(R_\rho \leq r) = 1 - \sum_{k=0}^{d-1} \frac{(-1)^k}{k!}\, r^k \psi_\rho^{(k)}(r).$$

The derivatives of ψ_ρ seem cumbersome to compute at first sight. Thanks to [20], however, it is known that, for all $d \in \mathbb{N}$ and $x \in (0,\infty)$,

$$(-1)^d \psi_\rho^{(d)}(x) = \frac{\psi_\rho(x)}{x^d} \sum_{k=1}^{d} x^{k\rho} a_{kd}(\rho),$$

where for all $k \in \{1,\ldots,d\}$, $a_{dk}(\rho)$ is a constant given in terms of the Stirling numbers s and S of the first and second kind, respectively, viz.

$$a_{dk}(\rho) = (-1)^{d-k} \sum_{i=k}^{d} \rho^i s(d,i) S(i,k) = \frac{d!}{k!} \sum_{i=1}^{k} \binom{k}{i} \binom{\rho i}{d} (-1)^{d-i}.$$

A simple calculation shows that the density of R_ρ is given, for all $r \in (0,\infty)$, by

$$g_\rho(r) = \frac{(-1)^d}{(d-1)!}\, r^{d-1} \psi_\rho^{(d)}(r) = \sum_{k=1}^{d} m_{dk}(\rho) \left\{ \frac{\rho}{(k-1)!}\, r^{\rho k-1} e^{-r^\rho} \right\}, \tag{6.4}$$

where, for each $k \in \{1,\ldots,d\}$, $m_{dk}(\rho) = a_{dk}(\rho)(k-1)!/\{(d-1)!\rho\}$. In (6.4), the term in curly brackets is the density of $Q^{1/\rho}$, where Q is a Gamma random variable with shape parameter $k \in \{1,\dots,d\}$ and unit scaling, called a generalized Gamma in [37]. As the weights $m_{d1}(\rho),\dots,m_{dd}(\rho)$ are positive and add up to 1 for all $\rho \in (0,1)$, g_ρ can be recognized as a d-fold mixture of generalized Gamma random variables. This leads to an easy implementation of the alternative sampling algorithm for Archimedean copulas described in [26].

6.3 Galambos copulas as reciprocal Archimedean copulas

Formula (6.2) and its d-variate extension were obtained by Galambos [9] as the dependence structure of the extremal attractor of Mardia's multivariate survival Pareto distribution. Just as Gumbel copulas are the only extreme-value copulas that are Archimedean, Galambos copulas are the only extreme-value copulas that are reciprocal Archimedean in the sense of [13]. In the latter paper, a copula C is said to be reciprocal Archimedean if it can be expressed, for all $u_1,\dots,u_d \in [0,1]$, in the form

$$C_F(u_1,\dots,u_d) = \prod_{A\in\mathscr{P}_{d,o}} F\Big\{\sum_{k\in A} F^{-1}(u_k)\Big\}\Big/ \prod_{A\in\mathscr{P}_{d,e}} F\Big\{\sum_{k\in A} F^{-1}(u_k)\Big\}, \tag{6.5}$$

where $\mathscr{P}_{d,o} = \{A \subseteq \{1,\dots,d\} : A \neq \emptyset \text{ and } |A| \text{ is odd}\}$, $\mathscr{P}_{d,e} = \{A \subseteq \{1,\dots,d\} : A \neq \emptyset \text{ and } |A| \text{ is even}\}$, and the reciprocal Archimedean generator F is a continuous univariate distribution with support $[0,x_F]$, where $0 < x_F \le \infty$. In the case $d = 2$, formula (6.5) reduces, for all $u_1,u_2 \in [0,1]$, to the "reciprocal Archimedean" form

$$C_F(u_1,u_2) = \frac{u_1u_2}{F\{F^{-1}(u_1)+F^{-1}(u_2)\}},$$

which justifies in part the terminology. The copula GA_ρ is indeed of this form with generator defined, for all $x \in [0,\infty)$, by $F_\rho(x) = \exp(-x^{-\rho})$, i.e., its generator is the Fréchet distribution function with parameter $\rho \in (0,\infty)$.

By Sklar's Theorem, the distribution function given, for all $x_1,\dots,x_d \in [0,\infty)$, by

$$H_F(x_1,\dots,x_d) = C_F\{F(x_1),\dots,F(x_d)\} = \prod_{A\in\mathscr{P}_{d,o}} F\Big(\sum_{k\in A} x_k\Big)\Big/ \prod_{A\in\mathscr{P}_{d,e}} F\Big(\sum_{k\in A} x_k\Big) \tag{6.6}$$

has margins F and a reciprocal Archimedean copula with generator F. Clearly, the survival function corresponding to H_F in (6.6) is not Schur-convex. However, H_F admits an alternative expression in which Schur-convexity plays a prominent role.

As shown in [13], Eq. (6.5) defines a copula if and only if $\Lambda = -\ln(F)$ is d-monotone on $(0,\infty)$. The authors further prove that H_F is max-infinitely divisible (max-id), i.e., H_F^p is a *bona fide* d-variate distribution function for any $p > 0$. Moreover, one has, for all $x_1,\dots,x_d \in [0,\infty)$,

$$H_F(x_1,\dots,x_d) = \exp\big\{-\mu_F([-\infty,x_1]\times\cdots\times[-\infty,x_d])^{\complement}\big\},$$

where the complement is taken with respect to the set $E_d = [0,\infty]^d \setminus \{(0,\dots,0)\}$ and μ_F is a measure on E_d such that, for all $k \in \{1,\dots,d\}$, $\mu_F\{(x_1,\dots,x_d) \in E_d : x_k = 0\} = 0$, and for all $(x_1,\dots,x_d) \in E_d$,

$$\mu_F((x_1,\infty] \times \cdots \times (x_d,\infty]) = \Lambda(x_1 + \cdots + x_d).$$

In other words, the generalized survival function induced by μ_F is Schur-convex.

Alternatively, the exponent measure μ_F is ℓ_1-norm symmetric. It is further proved in [13] that the class of max-id distributions whose exponent measure is ℓ_1-norm symmetric has a stochastic representation which parallels that of ℓ_1-norm symmetric distributions. Indeed, given a distribution of the form (6.6), define first the radial measure ν_F on $(0,\infty]$ in such a way that, for all $r \in (0,\infty)$,

$$\nu_F(r,\infty] = \sum_{k=0}^{d-2} \frac{(-1)^k \Lambda^{(k)}(r)}{k!} r^k + \frac{(-1)^{d-1}\Lambda_+^{(d-1)}(r)}{(d-1)!} r^{d-1},$$

where $\Lambda_+^{(d-1)}$ denotes the right-hand derivative of $\Lambda^{(d-2)}$. Next introduce a Poisson random measure $\zeta = \sum_k \delta(r_k, s_{k1},\dots,s_{kd})$ on $(0,\infty] \times \mathbb{S}_d$ with mean measure $\nu_F \times \sigma_d$, where δ denotes a Dirac point measure and σ_d is the uniform probability measure on $\mathbb{S}_d$. Then H_F is the distribution function of the random vector

$$Y = (Y_1,\dots,Y_d) = \max_k \{r_k \times (s_{k1},\dots,s_{kd})\} \vee (0,\dots,0),$$

where the maximum is taken component-wise and $a \vee b = \max(a,b)$ for any vectors $a,b \in \mathbb{R}^d$. As discussed in [13], this representation leads, among other things, to simulation algorithms for copulas of the form (6.5).

For the Galambos copula with parameter $\rho \in (0,\infty)$, one has $F_\rho(x) = \exp(-x^{-\rho})$ for all $x \in [0,\infty)$, i.e., $\Lambda_\rho(x) = -\ln\{F_\rho(x)\} = x^{-\rho}$. This function is completely monotone and, for all $k \in \mathbb{N}$, its kth derivative is given, for all $x \in (0,\infty)$, by

$$\Lambda_\rho^{(k)}(x) = (-1)^k x^{-\rho-k} \prod_{j=0}^{k-1} (\rho + j).$$

Furthermore, the corresponding radial measure ν_{F_ρ} is such that, for all $r \in (0,\infty)$,

$$\nu_{F_\rho}(r,\infty] = \sum_{k=0}^{d-1} \frac{(-1)^k}{k!} r^k (-1)^k x^{-\rho-k} \prod_{j=0}^{k-1} (\rho + j) = r^{-\rho} \frac{\Gamma(\rho+d)}{\Gamma(d)\Gamma(\rho+1)},$$

where Γ is the Gamma function. Moreover, Λ_ρ is the Laplace transform of a measure γ_ρ on $[0,\infty)$ with $\gamma_\rho[0,t] = t^\rho/\Gamma(1+\rho)$ for all $t \in [0,\infty)$, as for all $x \in (0,\infty)$,

$$\int_0^\infty e^{-tx} \gamma_\rho(t) = \frac{\rho}{\Gamma(1+\rho)} \int_0^\infty e^{-tx} t^{\rho-1} dt = x^{-\rho} = \Lambda_\rho(x).$$

Now let $\mathscr{E}_1$ denote the unit exponential distribution on $(0,\infty)$ and consider a Poisson random measure $\xi = \sum_k \delta(x_{k1},\dots,x_{kd},t_k)$ on $(0,\infty)^d \times [0,\infty)$ with mean measure $\mathcal{E}_1 \times \cdots \times \mathcal{E}_1 \times \gamma_\rho$. The function H_{F_ρ} given in (6.6) is the distribution function of

$$(Z_1,\ldots,Z_d) = \max_k \{(x_{k1},\ldots,x_{kd})/t_k\} \vee (0,\ldots,0),$$

as established in [13]. This stochastic representation parallels the interpretation of Archimedean copulas with completely monotone generators as the dependence structures of multiplicative frailty models. The measure γ_ρ plays the role of the frailty distribution. However, the crucial difference is that $\gamma_\rho([0,\infty))$ is infinite, so that γ_ρ cannot be scaled to be a probability measure.

6.4 Gumbel and Galambos brought together

Being extreme-value, the Gumbel and Galambos copulas can both be written, for all $u_1,\ldots,u_d \in (0,1)$, in the form

$$C_\ell(u_1,\ldots,u_d) = \exp[-\ell\{-\ln(u_1),\ldots,-\ln(u_d)\}], \tag{6.7}$$

in terms of a function $\ell : [0,\infty)^d \to [0,\infty)$ called the stable tail dependence function [12, 15, 19] whose analytical characterization is given in [3, 34].

The stable tail dependence function of the Gumbel copula with parameter $\rho \in (0,1)$ is given, for all $x_1,\ldots,x_d \in [0,\infty)$, by

$$\ell_\rho^{\mathrm{GU}}(x_1,\ldots,x_d) = \left(x_1^{1/\rho} + \cdots + x_d^{1/\rho}\right)^\rho.$$

In contrast, the stable tail dependence function of the Galambos copula with parameter $\rho \in (0,\infty)$ is given, for all $x_1,\ldots,x_d \in [0,\infty)$, by

$$\ell_\rho^{\mathrm{GA}}(x_1,\ldots,x_d) = \sum_{A\subset\{1,\ldots,d\},\, A\neq\emptyset} (-1)^{|A|+1} \Big(\sum_{i\in A} x_i^{-1/\rho}\Big)^\rho.$$

As already mentioned, members of the Gumbel and Galambos families are the only extreme-value copulas that are Archimedean and reciprocal Archimedean, respectively. It is also known that Gumbel copulas are the only extreme-value attractors of the Archimedean class [4, 14, 23]. Proposition 6.1 below establishes a similar result for the Galambos family within the class of reciprocal Archimedean copulas.

Recall first that an arbitrary copula C lies in the maximum domain of attraction of an extreme-value copula C_0 if, for all $u_1,\ldots,u_d \in [0,1]$,

$$\lim_{n\to\infty} C^n(u_1^{1/n},\ldots,u_d^{1/n}) = C_0(u_1,\ldots,u_d).$$

Proposition 6.1. *Let C_F be a d-variate reciprocal Archimedean copula with generator F such that $1-F$ is regularly varying with index $-\rho$ for $\rho \in (0,\infty)$, i.e., such that for any $x \in (0,\infty)$, $\{1-F(xt)\}/\{1-F(t)\} \to x^{-\rho}$ as $t \to \infty$. Then C_F is in the maximum domain of attraction of the Galambos copula* GA_ρ.

Proof. Recall from [33] that if $1-F$ is regularly varying with index $-\rho$, there exists a sequence (a_n) of positive constants such that, for any $x \in (0,\infty)$, $F^n(a_n x) \to e^{-x^{-\rho}}$ as $n \to \infty$. Thus H_F given in Eq. (6.6) satisfies, for all $x_1,\ldots,x_d \in (0,\infty)$,

$$\lim_{n\to\infty} H_F^n(a_n x_1,\ldots,a_n x_d) = \lim_{n\to\infty} \prod_{A\in\mathscr{P}_{d,o}} F^n\Big(a_n \sum_{k\in A} x_k\Big) \Big/ \prod_{A\in\mathscr{P}_{d,e}} F^n\Big(a_n \sum_{k\in A} x_k\Big)$$
$$= \prod_{A\in\mathscr{P}_{d,o}} \exp\{-\big(\sum_{k\in A} x_k\big)^{-\rho}\} \Big/ \prod_{A\in\mathscr{P}_{d,e}} \exp\{-\big(\sum_{k\in A} x_k\big)^{-\rho}\}.$$

The last expression is easily recognized as the distribution function H_{F_ρ}, where for all $x \in (0,\infty)$, $F_\rho(x) = \exp(-x^{-\rho})$ is the Galambos generator. Consequently, the copula of H_F is in the maximum domain of attraction of GA_ρ, as claimed. □

Regular variation of $1-F$ required in Proposition 6.1 parallels the condition needed for an Archimedean copula C_ψ to be in the domain of attraction of GU_ρ, namely that $1-\psi(1/\cdot)$ is regularly varying with index $-\rho \in [-1,0)$. The limiting case $\rho = 1$ refers to the independence copula Π.

In the remainder of this section, we show that the Gumbel and Galambos families can in fact be embedded in a single parametric class of extreme-value copulas. To this end, note that the stable tail dependence function in (6.7) can also be conveniently characterized in terms of a probability measure υ_d on $\mathbb{S}_d$, called the spectral distribution, with the property that, for all $k \in \{1,\ldots,d\}$,

$$\int_{\mathbb{S}_d} w_k \,\mathrm{d}\upsilon_d(w_1,\ldots,w_d) = \frac{1}{d}.$$

Following [11, 32], one then has, for all $x_1,\ldots,x_d \in [0,\infty)$,

$$\ell(x_1,\ldots,x_d) = d\int_{\mathbb{S}_d} \max(w_1x_1,\ldots,w_dx_d)\mathrm{d}\upsilon_d(w_1,\ldots,w_d).$$

In the special case of the Gumbel copula, a simple calculation shows that the density of the spectral distribution υ_d is given, for all $(w_1,\ldots,w_d) \in \mathbb{S}_d$, by

$$h_\rho^{\mathrm{GU}}(w_1,\ldots,w_d) = \frac{\Gamma(d-\rho)}{d\rho^{d-1}\Gamma(1-\rho)} \Big(\prod_{k=1}^d w_k\Big)^{-1/\rho-1} \Big(\sum_{k=1}^d w_k^{-1/\rho}\Big)^{\rho-d}, \tag{6.8}$$

which is the spectral density of Gumbel's multivariate logistic model; see [6, 18]. For the Galambos copula, the density of υ_d is given, for all $(w_1,\ldots,w_d) \in \mathbb{S}_d$, by

$$h_\rho^{\mathrm{GA}}(w_1,\ldots,w_d) = \frac{\Gamma(d+\rho)}{d\rho^{d-1}\Gamma(1+\rho)} \Big(\prod_{k=1}^d w_k\Big)^{1/\rho-1} \Big(\sum_{k=1}^d w_k^{1/\rho}\Big)^{-\rho-d}, \tag{6.9}$$

which is the spectral density of the negative logistic model; see, e.g., [6, 22].

Comparing (6.8) and (6.9), one can see that the formulas are again strikingly similar. In fact, it turns out that they can be embedded in one and the same family of spectral densities as follows. For all $(w_1,\ldots,w_d)\in\mathbb{S}_d$ and $\rho\in(-\infty,1)\setminus\{0\}$, let

$$h^{\mathrm{G}}_{\rho}(w_1,\ldots,w_d)=\frac{\Gamma(d-\rho)}{d|\rho|^{d-1}\Gamma(1-\rho)}\Big(\prod_{k=1}^{d}w_k\Big)^{-1/\rho-1}\Big(\sum_{k=1}^{d}w_k^{-1/\rho}\Big)^{\rho-d}. \tag{6.10}$$

Then, for all $(w_1,\ldots,w_d)\in\mathbb{S}_d$, one finds

$$h^{\mathrm{G}}_{\rho}(w_1,\ldots,w_d)=\begin{cases}h^{\mathrm{GU}}_{\rho}(w_1,\ldots,w_d) & \text{if } \rho\in(0,1),\\ h^{\mathrm{GA}}_{-\rho}(w_1,\ldots,w_d) & \text{if } \rho\in(-\infty,0).\end{cases}$$

The family of stable tail dependence functions with spectral density (6.10) is a special case of the scaled extremal Dirichlet model derived in [2]. As shown therein, this model has a simple stochastic representation which is stated and proved below in the specific case of h^{G}_{ρ}.

Proposition 6.2. *For arbitrary $\rho\in(-\infty,1)\setminus\{0\}$, the stable tail dependence function ℓ^{G}_{ρ} with spectral density* (6.10) *is given, for all $x_1,\ldots,x_d\in[0,\infty)$, by*

$$\begin{aligned}\ell^{\mathrm{G}}_{\rho}(x_1,\ldots,x_d)&=\frac{\Gamma(d-\rho)}{\Gamma(d)\Gamma(1-\rho)}\,\mathrm{E}\{\max(x_1S_1^{-\rho},\ldots,x_dS_d^{-\rho})\}\\&=\frac{1}{\Gamma(1-\rho)}\,\mathrm{E}\{\max(x_1X_1^{-\rho},\ldots,x_dX_d^{-\rho})\},\end{aligned}$$

where $(S_1,\ldots,S_d)$ is uniformly distributed on the unit simplex $\mathbb{S}_d$ and $X_1,\ldots,X_d$ are mutually independent unit exponential random variables.

Proof. As is well known, $X=X_1+\cdots+X_d$ is independent of $(X_1/X,\ldots,X_d/X)$ and the latter has the same distribution as $(S_1,\ldots,S_d)$. Moreover, X is Gamma with shape parameter d and unit scaling. Therefore, $\mathrm{E}(X^{-\rho})=\Gamma(d-\rho)/\Gamma(d)$ and thus

$$\mathrm{E}\{\max(x_1X_1^{-\rho},\ldots,x_dX_d^{-\rho})\}=\frac{\Gamma(d-\rho)}{\Gamma(d)}\,\mathrm{E}\{\max(x_1S_1^{-\rho},\ldots,x_dS_d^{-\rho})\}.$$

Furthermore, the left-hand side can be computed to be

$$\frac{1}{\Gamma(1-\rho)}\int_0^{\infty}\cdots\int_0^{\infty}\max(x_1t_1^{-\rho},\ldots,x_dt_d^{-\rho})\Big(\prod_{k=1}^{d}e^{-t_k}\Big)\mathrm{d}t_1\cdots\mathrm{d}t_d. \tag{6.11}$$

Now set $q=t_1^{-\rho}+\cdots+t_d^{-\rho}$ and $w_k=t_k^{-\rho}/q$ for all $k\in\{1,\ldots,d-1\}$. Let also $w_d=1-(w_1+\cdots+w_{d-1})$ so that, for all $k\in\{1,\ldots,d\}$, $t_k=(qw_k)^{-1/\rho}$. As the absolute value of the Jacobian of this transformation is

$$|J|=\frac{1}{|\rho|^d}q^{-d/\rho-1}\Big(\prod_{k=1}^{d}w_k\Big)^{-1/\rho-1},$$

the expression (6.11) can be rewritten as

$$\frac{1}{|\rho|^d \Gamma(1-\rho)} \int_{\mathbb{S}_d} \max(x_1 w_1, \ldots, x_d w_d) \Big(\prod_{k=1}^d w_k \Big)^{-1/\rho - 1} \\ \times \int_0^\infty q^{-d/\rho} e^{-q^{-1/\rho} \sum_{k=1}^d w_k^{-1/\rho}} \mathrm{d}q \mathrm{d}w_1 \cdots \mathrm{d}w_{d-1}.$$

Equation (6.10) now follows from the fact that

$$\int_0^\infty q^{-d/\rho} e^{-q^{-1/\rho} \sum_{k=1}^d w_k^{-1/\rho}} \mathrm{d}q = |\rho| \Gamma(d-\rho) \Big(\sum_{k=1}^d w_k^{-1/\rho} \Big)^{\rho - d}.$$

This concludes the argument. □

Further note that, for all $x_1, \ldots, x_d \in [0, \infty)$,

$$\lim_{\rho \to 0} \ell_\rho^{\mathrm{G}}(x_1, \ldots, x_d) = \max(x_1, \ldots, x_d),$$

which is the stable tail dependence function of the Fréchet–Hoeffding upper bound [28, 30], i.e., the copula M of d comonotonic uniform random variables. The spectral distribution of M is degenerate; it places mass 1 at $(1/d, \ldots, 1/d)$.

From Proposition 6.2 one sees that, for any $x_1, \ldots, x_d \in [0, \infty)$ and any $\rho \in (0,1)$,

$$\ell_\rho^{\mathrm{GU}}(x_1, \ldots, x_d) = \frac{1}{\Gamma(1-\rho)} \mathrm{E}\{\max(x_1 X_1^{-\rho}, \ldots, x_d X_d^{-\rho})\} \tag{6.12}$$

while for any $\rho \in (0, \infty)$,

$$\ell_\rho^{\mathrm{GA}}(x_1, \ldots, x_d) = \frac{1}{\Gamma(1+\rho)} \mathrm{E}\{\max(x_1 X_1^{\rho}, \ldots, x_d X_d^{\rho})\}. \tag{6.13}$$

It is interesting to note that when X is unit exponential and $\rho > 0$, $\mathrm{E}(X^\rho) = \Gamma(1+\rho)$ is always finite, while $\mathrm{E}(X^{-\rho})$ is finite and equal to $\Gamma(1-\rho)$ only when $\rho < 1$. Thus the parameter space of the Gumbel copula can only be extended to $[0,1]$ but not beyond; the case $\rho = 1$ corresponds to independence.

Beyond highlighting the kinship between the Gumbel and Galambos families of copulas, formulas (6.12)–(6.13) lead to a unified simulation algorithm for these two dependence structures. This procedure, adapted from [7, 35], is presented in a broader context in [2]. Gumbel and Galambos are thereby united, at last.

Acknowledgements This work was completed while the authors were visiting Université Claude-Bernard, Lyon 1, whose hospitality is gratefully acknowledged. Partial funding in support of this work was provided by the Canada Research Chairs Program, the Natural Sciences and Engineering Research Council of Canada (2015–06801, 2016–04720), the Canadian Statistical Sciences Institute, and the Fonds de recherche du Québec – Nature et technologies (2015–PR–183236).

References

1. Beirlant, J., Goegebeur, Y., Teugels, J., Segers, J.: Statistics of Extremes. Wiley, Chichester (2004)
2. Belzile, L.R., Nešlehová, J.G.: Extremal attractors of Liouville copulas (2017). Submitted
3. Charpentier, A., Fougères, A.L., Genest, C., Nešlehová, J.G.: Multivariate Archimax copulas. J. Multivariate Anal. **126**, 118–136 (2014)
4. Charpentier, A., Segers, J.: Tails of multivariate Archimedean copulas. J. Multivariate Anal. **100**, 1521–1537 (2009)
5. Coles, S.: An Introduction to Statistical Modeling of Extreme Values. Springer, London (2001)
6. Coles, S.G., Tawn, J.A.: Modelling extreme multivariate events. J. Roy. Statist. Soc. Ser. B **53**, 377–392 (1991)
7. Dombry, C., Engelke, S., Oesting, M.: Exact simulation of max-stable processes. Biometrika **103**, 303–317 (2016)
8. Feller, W.J.: An Introduction to Probability Theory and its Applications. Vol. II. Second edition. Wiley, New York (1971)
9. Galambos, J.: Order statistics of samples from multivariate distributions. J. Amer. Statist. Assoc. **70**, 674–680 (1975)
10. Galambos, J.: The Asymptotic Theory of Extreme Order Statistics. Wiley, New York (1978)
11. Galambos, J.: The Asymptotic Theory of Extreme Order Statistics, 2nd edn. Robert E. Krieger, Melbourne, FL (1987)
12. Genest, C., Nešlehová, J.: Copula modeling for extremes. In: A.H. El-Shaarawi, W.W. Piegorsch (eds.) Encyclopedia of Environmetrics, 2nd edn. Wiley, Chichester (2012)
13. Genest, C., Nešlehová, J.G., Rivest, L.P.: The class of multivariate max-id copulas with ℓ_1-norm symmetric exponent measure (2016). Submitted
14. Genest, C., Rivest, L.P.: A characterization of Gumbel's family of extreme value distributions. Statist. Probab. Lett. **8**, 207–211 (1989)
15. Gudendorf, G., Segers, J.: Extreme-value copulas. In: Copula Theory and its Applications, *Lect. Notes Stat. Proc.*, vol. 198, pp. 127–145. Springer, Heidelberg (2010)
16. Gumbel, E.J.: Statistics of Extremes. Columbia University Press, New York (1958)
17. Gumbel, E.J.: Bivariate exponential distributions. J. Amer. Statist. Assoc. **55**, 698–707 (1960)
18. Gumbel, E.J.: Distributions des valeurs extrêmes en plusieurs dimensions. Publ. Inst. Statist. Univ. Paris **9**, 171–173 (1960)
19. de Haan, L., Resnick, S.I.: Limit theory for multivariate sample extremes. Z. Wahrscheinlichkeitstheorie und Verw. Gebiete **40**, 317–337 (1977)
20. Hofert, M., Mächler, M., McNeil, A.J.: Likelihood inference for Archimedean copulas in high dimensions under known margins. J. Multivariate Anal. **110**, 133–150 (2012)
21. Hougaard, P.: A class of multivariate failure time distributions. Biometrika **73**, 671–678 (1986)
22. Joe, H.: Families of min-stable multivariate exponential and multivariate extreme value distributions. Statist. Probab. Lett. **9**, 75–81 (1990)
23. Larsson, M., Nešlehová, J.: Extremal behavior of Archimedean copulas. Adv. Appl. Probab. **43**, 195–216 (2011)
24. Malov, S.: On finite-dimensional Archimedean copulas. In: N. Balakrishnan, I. Ibragimov, V. Nevzorov (eds.) Asymptotic Methods in Probability and Statistics With Applications, pp. 19–35. Birkhäuser, Basel (2001)
25. Marshall, A.W., Olkin, I.: Families of multivariate distributions. J. Amer. Statist. Assoc. **83**, 834–841 (1988)
26. McNeil, A.J., Nešlehová, J.: Multivariate Archimedean copulas, d-monotone functions and ℓ_1-norm symmetric distributions. Ann. Statist. **37**, 3059–3097 (2009)
27. McNeil, A.J., Nešlehová, J.: From Archimedean to Liouville copulas. J. Multivariate Anal. **101**, 1772–1790 (2010)
28. Nelsen, R.B.: An Introduction to Copulas. Springer, New York (1999)
29. Nelsen, R.B.: Some properties of Schur-constant survival models and their copulas. Braz. J. Probab. Stat. **19**, 179–190 (2005)

30. Nelsen, R.B.: An Introduction to Copulas, 2nd edn. Springer, New York (2006)
31. Oakes, D.: Bivariate survival models induced by frailties. J. Amer. Statist. Assoc. **84**, 487–493 (1989)
32. Pickands III, J.: Multivariate extreme value distributions. In: Proceedings of the 43rd session of the International Statistical Institute, Vol. 2 (Buenos Aires, 1981), vol. 49, pp. 859–878, 894–902 (1981)
33. Resnick, S.I.: Extreme Values, Regular Variation and Point Processes. Springer, New York (2008)
34. Ressel, P.: Homogeneous distributions—and a spectral representation of classical mean values and stable tail dependence functions. J. Multivariate Anal. **117**, 246–256 (2013)
35. Schlather, M.: Models for stationary max-stable random fields. Extremes **5**, 33–44 (2002)
36. Schweizer, B., Sklar, A.: Probabilistic Metric Spaces. North-Holland, New York (1983)
37. Stacy, E.W.: A generalization of the Gamma distribution. Ann. Math. Statist. **33**, 1187–1192 (1962)
38. Williamson, R.E.: Multiply monotone functions and their Laplace transforms. Duke Math. J. **23**, 189–207 (1956)

Chapter 7
On the Conditional Value-at-Risk (CoVaR) in copula setting

Piotr Jaworski

Abstract This survey is dedicated to an application of copula methodology to systemic risk management. We deal with the modified Conditional Value-at-Risk (CoVaR) for various families of copulas. We study to what extent the tail behaviour of the copula determines the limiting performance of the CoVaR when the conditioning event is becoming more extreme.

7.1 Introduction

The theory of copulas provides a useful tool for modeling dependence in risk management. In insurance and finance, as well as in other applications like hydrology, dependence of extreme events is particularly important, hence there is a need for the detailed study of the tail behaviour of copulas.

To fix the notation, in this paper we will base on the Profit/Loss (P/L) approach as for example in [12, 1, 2, 22].

We will study random variables X and Y, which are modelling for example: welfare of the financial institutions, financial positions, gains of the investments, or rates of returns of stock prices and indices. So generally

"The higher value of X the better".

We recall that Value-at-Risk, at a given significance level $\alpha \in (0,1)$, of a P/L random variable X, is defined as follows ([12]):

$$VaR_\alpha(X) = \inf\{v \in \mathbb{R} : \mathbb{P}(X + v < 0) \leq \alpha\}.$$

Piotr Jaworski
Institute of Mathematics, University of Warsaw, Poland
e-mail: P.Jaworski@mimuw.edu.pl

M. Úbeda Flores et al. (eds.), *Copulas and Dependence Models with Applications*,
DOI 10.1007/978-3-319-64221-5_7

The above can be expressed in terms of quantiles. Namely Value-at-Risk at a level α is equal to minus upper α quantile of X or lower $1-\alpha$ quantile of the loss $-X$

$$VaR_{\alpha}(X) = -Q^{+}_{\alpha}(X) = Q^{-}_{1-\alpha}(-X).$$

We recall that for a given random variable X and a given level $\kappa \in (0,1)$ the set of κ quantiles is a closed interval $Q_{\kappa}(X)$, which might be reduced to a point. The end points of the interval $Q_{\kappa}(X)$ are referred to as upper and lower quantiles.

$$[Q^{-}_{\kappa}(X), Q^{+}_{\kappa}(X)] = Q_{\kappa}(X) = \{x : \mathbb{P}(X \leq x) \geq \kappa \text{ and } \mathbb{P}(X \geq x) \geq 1-\kappa\}.$$

To switch to the alternative Loss/Profit (L/P) approach (applied for example in [32, 15, 3]) when random variables are modelling losses of the financial investments, actuarial risks or high water levels in hydrology, it is is enough to change the sign of the variables

$$L = -X,$$

and remember that, by a convention, the subscript is changed. The significance level α is replaced by the confidence level $c = 1-\alpha$

$$VaR_{c}(L) = VaR_{\alpha}(X).$$

Now assume, that we are measuring the risk given some stress event. For example, we want to determine how big bailout would be necessary to keep a financial institution Y solvent with probability at least $1-\beta$ when a financial institution X would perform badly. Conditional Value-at-Risk (CoVaR) introduced in 2008 by Adrian and Brunnermeier ([1]) and its later modifications proved to be very useful tools for measuring (quantifying) such phenomena.

Let X and Y be random variables modelling positions. CoVaR is defined as VaR of Y conditioned by X. In more details:

$$CoVaR(Y|X) = VaR_{\beta}(Y|X \in E),$$

where a Borel subset of the real line E is modeling some adverse event concerning X. Most often E consists of one point (a threshold) or is a half-line bounded by a threshold.

As we see, to deal with CoVaR, one has to model the dependence between Y and X. This can be achieved by means of copulas. In this survey we recapitulate the results stated in [32, 2, 22] putting a special attention to the asymptotic dependence and asymptotic independence between X and Y. Since we are following a P/L approach, we are interested in the shape of the copula close to the left border of the unit square, especially around the origin.

Adrian and Brunnermeier ([1]) applied the construction with E consisting of one point. Such approach has a certain drawback, pointed out for example by Mainik and Schaanning in [32], which is due to the fact that the standard CoVaR is not compatible with the concordance ordering. Hence it is "breaking" the paradigm:

more dependence, more systemic risk (see also [22]). To avoid this inconsistency a modified definition of CoVaR was introduced in 2013 and 2014 by Girardi and A.T. Ergün ([13]) and by Mainik and Schaanning ([32]), both in L/P setting.
The modified Conditional-VaR at a level (α,β), which is a main objective of this survey, is defined as VaR at level β of Y under the condition that $X \leq -VaR_\alpha(X)$.

Definition 7.1.

$$CoVaR^{\leq}_{\alpha,\beta}(Y|X) = VaR_\beta(Y \mid X \leq -VaR_\alpha(X)),$$

which can be expressed in terms of quantiles as:

$$CoVaR^{\leq}_{\alpha,\beta}(Y|X) = -Q^{+}_{\beta}(Y \mid X \leq Q^{+}_{\alpha}(X)).$$

The paper is organized in the following way:
In section 7.2 we recapitulate the basic facts about copulas, their geometric transformations and tail expansions.
Section 7.3 contains the main "general" results. We recall how to express modified CoVaR in terms of copulas. Following [2] we study the threshold $w_*(\alpha,\beta,C)$ such that $CoVaR^{\leq}(Y) = VaR_{w_*}(Y)$. We discuss the compatibility of modified CoVaR with concordance ordering of copulas i.e. with the strength of dependence between the conditioned and conditioning variable and provide the rough bounds for the thresholds. Next we deal with copulas with non-trivial tail expansions at one of points (0,0) or (0,1) or both of them. We show that for such copulas the first order limiting properties of w_* are fully determined by the tail dependence functions. In the last part of the section we deal with copulas that are differentiable and close to the independent copula.
We illustrate the above results in last section. We study the implied threshold w_* for several families of copulas.

7.2 Copulas

7.2.1 Basic notation

To fix the notation we collect here some basic facts about copulas. For more details the reader is referred to the monograph of Roger Nelsen [35] or other publications on the subject like [26, 8, 6, 23, 24, 10, 34].

We recall that the function

$$C : [0,1]^2 \longrightarrow [0,1]$$

is called a copula if the following three properties hold:

(c1) $\forall u_1,u_2 \in [0,1] \quad C(u_1,0) = 0, \quad C(0,u_2) = 0;$

(c2) $\forall u_1, u_2 \in [0,1] \quad C(u_1, 1) = u_1, \quad C(1, u_2) = u_2$;
(c3) $\forall u_1, u_2, v_1, v_2 \in [0,1], \quad u_1 \leq v_1, u_2 \leq v_2$

$$C(v_1, v_2) - C(u_1, v_2) - C(v_1, u_2) + C(u_1, u_2) \geq 0.$$

We can alternatively characterize copulas in a more probabilistic way. Namely, a function C is a copula if and only if there exist random variables U, V, which are uniformly distributed on $[0,1]$, such that C is a restriction to the unit square $[0,1]^2$ of their joint distribution function. Random variables U and V are called the representers of the copula C.

Note that when random variables X and Y have continuous distribution functions F_X and F_Y, then the random variables

$$F_X(X), F_Y(Y)$$

are representers of the copula C of the pair (X, Y) (compare [18] Pr.1).

One more premise to use copulas to model systemic risk follows from the fact that the copulas are true measures of interdependence between random phenomena. Namely they do not depend on the scale in which these phenomena are quantified.

Indeed, if C is a copula of a random pair

$$X = (X_1, X_2),$$

and the functions h_1, h_2 are defined and strictly increasing on the supports of X_1, X_2, then C is also a copula of the transformed random pair

$$Y = (h_1(X_1), h_2(X_2)).$$

7.2.2 Geometrical transformations of copulas

We recall that there exist eight linear isometric transformations of the unit square $[0,1]^2$: two mirror reflections with respect to the diagonals, two mirror reflections with respect to bisectors, one point reflection, two rotations $\pm\pi/2$ (90 and 270 degrees) and identity. They induce the transformations of copulas. Namely let random variables (U_1, U_2) be representers of a copula C and

$$\sigma : [0.1]^2 \longrightarrow [0,1]^2$$

be an isometry, then random variables V_1, V_2 given by

$$(V_1, V_2) = \sigma(U_1, U_2)$$

are uniformly distributed on the unit interval $[0,1]$. The copula C_σ of the pair V_1, V_2 is called a reflection or respectively rotation of the copula C.

The copulas obtained by the point reflection are better known under the name "survival copulas"

$$C^{surv}(u,v) = u + v - 1 + C(1-u, 1-v).$$

Note that in this case

$$V_1 = 1 - U_1, \ V_2 = 1 - U_2.$$

Survival copulas are a useful tool when one is switching from P/L to L/P setting. Indeed when C is a copula of gains X and Y, then the survival copula C^{surv} is a copula of losses $L_1 = -X$ and $L_2 = -Y$.

In the following we will study the tail behaviour of copulas at the left side of its domain. Therefore we will be checking the tail dependence not only at the origin (0,0) but also at the vertex (0,1). We will denote by $\widehat{C}$ the reflection of a copula C with respect to the horizontal line $v = 0.5$.

$$\widehat{C}(u,v) = u - C(u, 1-v).$$

In this case

$$V_1 = U_1, \quad V_2 = 1 - U_2.$$

7.2.3 Copulas admitting tail expansions

In risk management one has to deal with extreme events and the interdependencies between them. This leads to the study of the tail behaviour of a copula, i.e. of the possible approximations of a copula close to the vertices of the unit square. Since applying a proper geometric transformation one may map any vertex of the unit square $[0,1]^2$ to the selected one, we restrict ourselves to the vertex $(0,0)$ (the origin).

Definition 7.2. We say that a copula C has a tail expansion at the vertex $(0,0)$ of the unit square if the limit

$$\lim_{t \to 0^+} \frac{C(tx, ty)}{t}$$

exists for all nonnegative x, y.

The function

$$L : [0,\infty]^2 \longrightarrow [0,\infty), \quad L(x,y) = \lim_{t \to 0} \frac{C(tx,ty)}{t},$$

is called the tail dependence function or the leading term of the tail expansion. The second naming follows from the fact, proved in [19]: if L exists then we have a decomposition of a copula

$$C(u,v) = L(u,v) + R(u,v)(u+v),$$

where R is bounded and

$$\lim_{(u,v)\to(0,0)} R(u,v) = 0.$$

The above can be applied to other vertices as well. It is enough to reflect the copula. For example for the vertex (0,1) we get

$$\widehat{L}(x,y) = \lim_{t\to 0} \frac{\widehat{C}(tx,ty)}{t} = \lim_{t\to 0} \frac{tx - C(tx, 1-ty)}{t}.$$

Note that $L(1,1)$ is equal to the lower tail dependence coefficient λ_L.

We recall the basic properties of the tail dependence functions (for details see [17, 18, 19, 20, 21, 28, 4, 5, 27, 30]).

Lemma 7.1. *[see [18, 19]] The tail dependence function induced by a copula C,*

$$L(\mathbf{u}) = \lim_{t\to 0^+} \frac{C(t\mathbf{u})}{t}, \quad \mathbf{u} \in [0,+\infty)^2,$$

is

1. *homogeneous of degree 1,*
2. *2-nondecreasing and nondecreasing with respect to every variable,*
3. *nonnegative and bounded by the smaller coordinate of* $\mathbf{u}$*:*

$$0 \le L(\mathbf{u}) \le \min(u_1, u_2).$$

4. *Lipschitz with Lipschitz constant 1:*

$$|L(\mathbf{v}) - L(\mathbf{u})| \le |v_1 - u_1| + |v_2 - u_2|.$$

5. *concave:*

$$\forall \lambda_1, \lambda_2 \ge 0, \lambda_1 + \lambda_2 = 1 \quad L(\lambda_1 \mathbf{u} + \lambda_2 \mathbf{v}) \ge \lambda_1 L(\mathbf{u}) + \lambda_2 L(\mathbf{v}).$$

Due to homogeneity, the leading term L is uniquely described by vertical sections like

$$l(t) = L(1,t).$$

Theorem 7.1. *Let*

$$l : [0,\infty) \to [0,1], \quad l(0) = 0,$$

be a nondecreasing, concave function, such that $l(t) \le t$*. Then the function*

$$L : [0,+\infty)^2 \longrightarrow [0,+\infty), \quad L(x,y) = \begin{cases} xl\left(\frac{y}{x}\right) & \text{for } x > 0, \\ 0 & \text{for } x = 0. \end{cases}$$

is a leading term of some copula.

The function $l(t) = L(1,t)$ will be called a generator of the leading term. The proof of theorem 7.1 follows from the examples in last section (see [18], section 2.1, and [17] for detailed calculations).

7.3 Modified CoVaR by copulas

We follow the P/L approach from [2]. For the L/P setting the reader is referred to [32] Theorem 3.1.b.

Let $C(u,v)$ be a copula of random variables X and Y having continuous distribution functions F_X and F_Y, then

$$\mathbb{P}(Y \le y \mid X \le Q_\alpha^+(X)) = \frac{\mathbb{P}(Y \le y \wedge X \le Q_\alpha^+(X))}{\alpha} = \frac{C(\alpha, F_Y(y))}{\alpha}.$$

Therefore

$$CoVaR_{\alpha,\beta}^{\le}(Y|X) = VaR_{w_*}(Y),$$

where $w_* = w_*(\alpha,\beta,C)$ is the largest solution of the equation

$$C(\alpha, w_*) = \alpha\beta.$$

Note that:

$$w_* = -CoVaR_{\alpha,\beta}^{\le}(F_Y(Y)|F_X(X)).$$

Furthermore as was observed in [32] Theorem 3.4 for the L/P setting, modified CoVaR is compatible with the concordance ordering of copulas. The same is valid for the P/L setting ([22]). In more details:

Theorem 7.2. *Let $C_i(u,v)$, $i = 1,2$, be a copula of a random pair (X_i, Y_i) having continuous marginal distribution functions F_{X_i} and F_{Y_i} and $\alpha,\beta \in (0,1]$ some fixed thresholds. If*

$$\forall (u,v) \in [0,1]^2 \quad C_1(u,v) \le C_2(u,v),$$

then

$$w_*(\alpha,\beta,C_1) \ge w_*(\alpha,\beta,C_2). \tag{7.1}$$

If, furthermore,

$$\forall t \in (-\infty,+\infty) \quad F_{Y_1}(t) \ge F_{Y_2}(t),$$

then

$$CoVaR_{\alpha,\beta}^{\le}(Y_1|X_1) \le CoVaR_{\alpha,\beta}^{\le}(Y_2|X_2). \tag{7.2}$$

Theorem 7.2 implies the rough bounds for the threshold w_*. Due to Fréchet-Hoeffding bounds (see [35, 8]) we have

$$M(u,v) \ge C(u,v) \ge W(u,v),$$

where $M(u,v) = \min(u,v)$ is the comonotonicity copula and $W(u,v) = (u+v-1)^+$ the countermonotonicity one. Since $w_*(\alpha,\beta,M) = \alpha\beta$ and $w_*(\alpha,\beta,W) = 1-\alpha+\alpha\beta$ ([2] §3.1), we get:

Corollary 7.1. *Let C be any bivariate copula. Then*

$$\alpha\beta \leq w_*(\alpha,\beta,C) \leq 1-\alpha(1-\beta).$$

If furthermore we assume that copula C is PQD (positively quadrant dependent) i.e. C dominates the independence copula $\Pi(u,v) = uv$

$$\forall (u,v) \in [0,1]^2 \quad C(u,v) \geq uv = \Pi(u,v),$$

we may improve the upper bound. Indeed, since $w_*(\alpha,\beta,\Pi) = \beta$, we get:

Corollary 7.2. *Let C be a PQD copula then*

$$\alpha\beta \leq w_*(\alpha,\beta,C) \leq \beta.$$

So, if random variables X and Y with continuous distribution functions are more likely to be large together or to be small together compared with a pair of independent random variables (X',Y') where $X' \stackrel{d}{=} X$ and $Y' \stackrel{d}{=} Y$ then

$$CoVaR^{\leq}_{\alpha,\beta}(Y|X) \geq VaR_\beta(Y).$$

Indeed, a positively quadrant dependent random pair (X,Y) admits a positively quadrant dependent copula (see [35] section 5.2.1).

7.3.1 Copulas with non-trivial tail expansions

In this section we deal with copulas having non-trivial tail expansions at the left vertices of the unit square. We show that the generators of the leading terms describe the limiting properties $w_*(\alpha,\beta)$ for respectively small and large (close to 1) values of β. Most of the copulas used to model dependencies between extremes have non-trivial tail expansion at one vertex. But there are significant exceptions. For example the t-Student copula has non-trivial tail expansions at all four vertices of the unit square.

Theorem 7.3. *([22]) Let the copula C have a nonzero tail dependence function L*

$$\lim_{t\to 0}\frac{C(tu,tv)}{t} = L(u,v) = ul\left(\frac{v}{u}\right).$$

Then for $\beta < l(\infty)$

$$\lim_{\alpha\to 0} w_*(\alpha,\beta,C) = 0,$$

$$\lim_{\alpha \to 0} \frac{w_*(\alpha, \beta, C)}{\alpha} = l^{-1}(\beta).$$

Proof. For consistency we recall the proof. We have to solve the equation

$$C(\alpha, w_*) = \alpha\beta. \tag{7.3}$$

First we show that for sufficiently small α, $w_*(\alpha, \beta, C)$ is bounded by some linear function of α. Indeed. We choose β_1 from the interval $(\beta, l(+\infty))$. We obtain

$$\lim_{\alpha \to 0} \frac{C(\alpha, \alpha l^{-1}(\beta_1))}{\alpha} = L(1, l^{-1}(\beta_1)) = \beta_1 > \beta.$$

So, for α smaller than sufficiently small α_1

$$C(\alpha, \alpha l^{-1}(\beta_1)) > \alpha\beta.$$

Since C is continuous and monotonic in the second variable, we get that for $\alpha \in (0, \alpha_1)$ the solution w_* of (7.3) is between 0 and $\alpha l^{-1}(\beta_1)$. Hence,

$$\lim_{\alpha \to 0} w_*(\alpha, \beta, C) = 0.$$

To show the second equality we decompose C

$$C(u, v) = L(u, v) + R(u, v)(u + v).$$

As it was shown in [19], R is bounded and has a limit at zero

$$\lim_{(u,v) \to (0,0)} R(u, v) = 0,$$

i.e. for any two sequences of numbers u_n and v_n from the unit interval, which are tending to 0 when $n \to \infty$, the sequence $R_n = R(u_n, v_n)$ is tending to 0 as well.
We rewrite equation (7.3)

$$L(\alpha, w_*) + R(\alpha, w_*)(\alpha + w_*) = \alpha\beta.$$

We divide both sides by α.

$$l\left(\frac{w_*}{\alpha}\right) = \beta - R(\alpha, w_*)\left(1 + \frac{w_*}{\alpha}\right).$$

Hence,

$$\frac{w_*}{\alpha} = l^{-1}\left(\beta - R(\alpha, w_*)\left(1 + \frac{w_*}{\alpha}\right)\right).$$

Since for $\alpha < \alpha_1$ $w_* < \alpha l^{-1}(\beta_1)$, we get

$$\lim_{\alpha \to 0} R(\alpha, w_*)\left(1 + \frac{w_*}{\alpha}\right) = 0.$$

Since l^{-1} is continuous, we obtain

$$\lim_{\alpha\to 0}\frac{w_*(\alpha,\beta,C)}{\alpha}=l^{-1}(\beta).$$

The similar results are valid for the vertex (0,1).

Theorem 7.4. *Let the copula C have a nonzero tail dependence function $\widehat{L}$ at the vertex (0,1)*

$$\lim_{t\to 0}\frac{tu-C(tu,1-tv)}{t}=\widehat{L}(u,v)=u\hat{l}\left(\frac{v}{u}\right).$$

Then for $\beta>1-\hat{l}(\infty)$

$$\lim_{\alpha\to 0}w_*(\alpha,\beta,C)=1,$$

$$\lim_{\alpha\to 0}\frac{w_*(\alpha,\beta,C)-1}{\alpha}=\hat{l}^{-1}(1-\beta).$$

Proof. Let

$$\widehat{C}(u,v)=u-C(u,1-v)$$

be the reflected copula. Since $w_*=w_*(\alpha,\beta,C)$ is a solution of the equation

$$C(\alpha,w_*)=\alpha\beta,$$

we get

$$\widehat{C}(\alpha,1-w_*)=\alpha-C(\alpha,w_*)=\alpha(1-\beta).$$

Hence,

$$w_*(\alpha,\beta,C)=1-w_*(\alpha,1-\beta,\widehat{C}).$$

From theorem 7.3 applied for $\widehat{C}$ and $1-\beta$ we get for $1-\beta<\hat{l}(+\infty)$

$$w_*(\alpha,\beta,C)=1-\alpha\hat{l}^{-1}(1-\beta)+o(\alpha).$$

7.3.2 Auxiliary results

In the next theorem we show that regularity conditions like differentiability of the copula and non-vanishing density may simplify the study of the threshold w_*. We shall deal with copulas that are three times continuously differentiable on the neighbourhood of the segment $\{0\}\times(0,1)$. In more details, we assume that there exists a function C^+ defined on $\mathbb{R}^2$ such that :
1. Copula C is a restriction of C^+,

$$C(u,v)=C^+(u,v)\quad\text{for all }\ (u,v)\in[0,1]^2.$$

2. The extension C^+ is three times continuously differentiable on some neighbourhood of the segment $\{0\}\times(0,1)$.

Theorem 7.5. *Let the copula C admits an extension C^+ which is three times continuously differentiable on some convex neighbourhood U of the segment $\{0\} \times (0,1)$ and its mixed derivative is positive on $\{0\} \times (0,1)$. Moreover, let the function $\varphi : [0,1] \to [0,1]$ given by*

$$\varphi(v) = \frac{\partial C(0^+, v)}{\partial u}$$

be continuous.

Then, for all $\beta \in (0,1)$,

$$\lim_{\alpha \to 0} w_*(\alpha, \beta, C) = \varphi^{-1}(\beta).$$

and

$$\lim_{\alpha \to 0} \frac{w_*(\alpha,\beta,C) - \varphi^{-1}(\beta)}{\alpha} = -\frac{\frac{\partial^2 C^+(0,\varphi^{-1}(\beta))}{\partial u^2}}{\frac{\partial^2 C^+(0,\varphi^{-1}(\beta))}{\partial u \partial v}}.$$

Proof. Since the mixed derivative of C^+ is positive along $\{0\} \times (0,1)$, $\varphi(v)$ is continuous and strictly increasing on $[0,1]$. Hence there exists exactly one $v_0 \in (0,1)$ such that

$$\frac{\partial C^+(0, v_0)}{\partial u} = \varphi(v_0) = \beta.$$

Equivalently one may put

$$v_0 = \varphi^{-1}(\beta).$$

We shall study a new function

$$\Psi(u,v) = C^+(u,v) - u\frac{\partial C^+(0,v_0)}{\partial u}.$$

Note that Ψ and its gradient are vanishing at the point $(0, v_0)$. Hence there exists functions $b_{1,1}$, $b_{1,2}$ and $b_{2,2}$ differentiable on U, such that

$$\Psi(u,v) = b_{1,1}(u,v)u^2 + b_{1,2}(u,v)u(v - v_0) + b_{2,2}(u,v)(v - v_0)^2$$

and

$$b_{2,2}(0,v_0) = 0, \;\; b_{1,2}(0,v_0) = \frac{\partial^2 C^+(0,v_0)}{\partial u \partial v}, \;\;\; b_{1,1}(0,v_0) = \frac{\partial^2 C^+(0,v_0)}{\partial u^2}.$$

For details see [16] Ch. 6 §1.1. Furthermore due to assumption about positivity of a mixed derivative, $b_{1,2}(0,v_0)$ is positive. Since $\frac{\partial C^+(0,v_0)}{\partial u} = \beta$ we have to solve an equation

$$\Psi(\alpha, w_*) = 0.$$

We substitute $w_*(\alpha) = v_0 + \alpha v_1(\alpha)$ and divide by α^2. We get

$$b_{1,1}(\alpha, v_0 + \alpha v_1) + b_{1,2}(\alpha, v_0 + \alpha v_1)v_1 + b_{2,2}(\alpha, v_0 + \alpha v_1)v_1^2 = 0.$$

Since the left side of the equation is vanishing at the point $\alpha = 0, v_1 = -\frac{b_{1,1}(0,v_0)}{b_{1,2}(0,v_0)}$ and the derivative of the left side with respect to v_1 is positive at this point, the implicit function theorem implies the existence of the function $v_1(\alpha)$ which solves the above equation. Obviously

$$v_1(0) = -\frac{b_{1,1}(0,v_0)}{b_{1,2}(0,v_0)}.$$

Next theorem shows what happens when we drop the assumption of the existence and continuity of the right-sided derivate with respect to u at $u = 0$.

Theorem 7.6. *If for every* $v \in (0,1]$

$$\liminf_{\alpha\to 0} \frac{C(\alpha,v)}{\alpha} \geq l_* > 0,$$

then for $\beta < l_*$

$$\lim_{\alpha\to 0} w_*(\alpha,\beta,C) = 0.$$

Proof. Let us assume that there exists a sequence α_k, $\lim_{k\to\infty} \alpha_k = 0$, such that for fixed $\beta < l_*$

$$\lim_{k\to\infty} w_*(\alpha_k,\beta) = v_* > 0.$$

Then for sufficiently large k

$$w_*(\alpha_k,\beta) > \frac{v_*}{2}.$$

Hence for such k

$$\beta = \frac{C(\alpha_k, w_*(\alpha_k,\beta))}{\alpha_k} \geq \frac{C\left(\alpha_k, \frac{v_*}{2}\right)}{\alpha_k}.$$

However

$$\liminf_{\alpha\to 0} \frac{C\left(\alpha_k, \frac{v_*}{2}\right)}{\alpha_k} \geq l_* > \beta$$

and a contradiction holds.

7.4 Examples

7.4.1 Survival extreme value copulas

Let $l : [0,+\infty] \to [0,1]$ be a concave, nondecreasing function, such that $l(t) \leq t$, then the function

$$C_l : [0,1]^2 \longrightarrow [0,1],$$
$$C_l(u,v) = u + v - 1 + (1-u)(1-v)\exp\left(-\ln(1-u)l\left(\frac{\ln(1-v)}{\ln(1-u)}\right)\right)$$

is a copula. Indeed, the survival copula is given by

$$C_l^{surv}(u,v) = uv\exp\left(-\ln(u)l\left(\frac{\ln(v)}{\ln(u)}\right)\right).$$

It belongs to the well-known class of extreme value copulas (see for example [35] §3.3.4, [8] §6.6. or [14]). When we put

$$l(t) = 1+t-(1+t^\theta)^{1/\theta}, \theta \geq 1,$$

we get the Gumbel family of copulas.

Copula C_l has a non-trivial tail expansion at the origin (compare [18] §5.4)

$$L(x,y) = xl\left(\frac{y}{x}\right)$$

and trivial tail expansion at vertex (0,1). We get a following equation for w_*.

$$\alpha + w_* - 1 + (1-\alpha)(1-w_*)\exp\left(-\ln(1-\alpha)l\left(\frac{\ln(1-w_*)}{\ln(1-\alpha)}\right)\right) = \alpha\beta.$$

We solve it with respect to l.

$$l\left(\frac{\ln(1-w_*)}{\ln(1-\alpha)}\right) = \frac{\ln(1+\alpha\beta-\alpha-w_*)-\ln(1-\alpha)-\ln(1-w_*)}{-\ln(1-\alpha)} = \beta + O(\alpha).$$

Next we inverse l and assume that $\beta < l(+\infty)$. Theorem 7.3 implies that $w_* = O(\alpha)$ and we get

$$\begin{aligned} w_* &= 1-\exp\left(\ln(1-\alpha)l^{-1}\left(\frac{\ln(1+\alpha\beta-\alpha-w_*)-\ln(1-\alpha)-\ln(1-w_*)}{-\ln(1-\alpha)}\right)\right) \\ &= \alpha l^{-1}(\beta) + O(\alpha^2). \end{aligned}$$

To get a better approximation of w_* one may apply the following recurrence:

Lemma 7.2. *([22]) Let the functions* $w_1, w_2 : [0,1] \longrightarrow [0,1]$ *fulfill the equation*

$$w_2(\alpha) = 1-\exp\left(\ln(1-\alpha)l^{-1}\left(\frac{\ln(1+\alpha\beta-\alpha-w_1(\alpha))-\ln(1-\alpha)-\ln(1-w_1(\alpha))}{-\ln(1-\alpha)}\right)\right).$$

If $w_*(\alpha) - w_1(\alpha) = O(\alpha^k)$*,* $k \geq 2$*, then* $w_*(\alpha) - w_2(\alpha) = O(\alpha^{k+1})$.

7.4.2 Survival conic copulas

Let $l : [0,+\infty] \to [0,1]$ be a concave, nondecreasing function, such that $l(t) \le t$, then the function

$$C_l : [0,1]^2 \longrightarrow [0,1], \quad C_l(u,v) = \max\left(ul\left(\frac{v}{u}\right), u+v-1\right)$$

is a copula with a tail dependence function

$$L(x,y) = xl\left(\frac{y}{x}\right).$$

At the vertex (0,1), we get

$$\widehat{L}(x,y) = \min((1-l(+\infty))x, y),$$

with generator

$$\hat{l}(t) = \min(1-l(+\infty), t).$$

Hence,

$$\hat{l}(+\infty) = 1-l(+\infty).$$

Copulas of this form were used in [18] to prove the existence of copulas with given lower and upper tail dependence functions. The survival copulas are given by

$$C_l^{surv} : [0,1]^2 \longrightarrow [0,1],$$
$$C_l^{surv}(u,v) = \max\left(1+(1-u)\left(l\left(\frac{1-v}{1-u}\right) - \frac{1-v}{1-u} - 1\right), 0\right).$$

They are known under the name ”conic copulas” (see [29, 11]).

To get the threshold w_* we have to solve an equation

$$\max\left(l\left(\frac{w}{\alpha}\right), 1+\frac{w}{\alpha} - \frac{1}{\alpha}\right) = \beta.$$

We get

$$w_* = \begin{cases} \min(\alpha l^{-1}(\beta), 1-\alpha(1-\beta)) & \text{for } \beta < l(\infty), \\ 1-\alpha(1-\beta) & \text{for } \beta \ge l(\infty), \end{cases}$$

Hence for $\beta < l(\infty)$ and $\alpha \le \frac{1}{1+l^{-1}(\beta)-\beta}$

$$w_*(\alpha,\beta,C_l) = \alpha l^{-1}(\beta).$$

If $l(+\infty) < 1$ we have additionally an illustration to theorem 7.4.
For $\beta \ge l(+\infty) = 1-\hat{l}(+\infty)$

$$w_*(\alpha,\beta,C_l) = 1-\alpha(1-\beta) = 1-\alpha\hat{l}^{-1}(1-\beta).$$

7.4.3 LTI copulas

Let $f : [0,+\infty] \to [0,1]$ be a surjective, concave and nondecreasing function and g its right inverse $(f(g(y)) = y)$. Then the function

$$C_f : [0,1]^2 \longrightarrow [0,1], \quad C_f(x,y) = \begin{cases} 0 & \text{for } x = 0, \\ xf\left(\frac{g(y)}{x}\right) & \text{for } x > 0. \end{cases}$$

is a copula introduced and considered in [7, 9]. It belongs to the class of copulas that are invariant under left truncation. For a suitable generator f, the popular Clayton copulas belong to this class. Namely

$$f(t) = (1+t^{-\theta})^{-1/\theta}, \quad \theta > 0.$$

Furthermore (see [9] proposition 4.1) the leading term of C_f equals

$$L(x,y) = xf\left(g'(0^+)\frac{y}{x}\right).$$

Note that since g is a convex increasing function its right sided derivative at 0 exists and is nonnegative. Furthermore L is nonzero if and only if $g'(0^+) > 0$. Then the generator $l(t)$ equals $f(g'(0^+)t)$ and

$$l(+\infty) = \lim_{t\to+\infty} l(t) = 1.$$

We have to solve an equation

$$f\left(\frac{g(w)}{\alpha}\right) = \beta.$$

We get ([2])

$$w_* = f(\alpha g(\beta)),$$

$$\lim_{\alpha\to 0} w_* = f(0) = 0, \quad \lim_{\alpha\to 0} \frac{\partial w_*}{\partial \alpha} = \lim_{\alpha\to 0} \frac{w_*}{\alpha} = f'(0^+)g(\beta).$$

Note that since f is concave, nondecreasing its derivative may be finite or infinite. In the first case C_f has a non-trivial leading term and

$$w_* = \alpha f'(0^+)g(\beta) + o(\alpha).$$

7.4.4 Fréchet copulas

To illustrate the fact that copula may have two non-trivial tails along the $x = 0$ side of the unit square we consider the subclass of copulas from the Fréchet family that consists of the convex combinations of the comonotonicity copula M and the countermonotonicity copula W. Copulas of this type are also known as X–copulas, since their support is contained in the two diagonals of the unit square (see [8]). We put for all $\lambda \in (0,1)$

$$C_{\lambda}^{\mathbf{Fre}}(u,v) = \lambda \min(u,v) + (1-\lambda)\max(u+v-1,0). \tag{7.4}$$

We get

$$L(x,y) = \lambda \min(x,y), \quad l(t) = \lambda \min(1,t).$$
$$\widehat{L}(x,y) = (1-\lambda)\min(x,y), \quad \hat{l}(t) = (1-\lambda)\min(1,t).$$

Note that since

$$l(+\infty) = \lim_{t\to+\infty} l(t) = \lambda \quad \text{and} \quad \hat{l}(+\infty) = \lim_{t\to+\infty} \hat{l}(t) = 1-\lambda,$$

we get for a convex combination of M and W

$$l(+\infty) + \hat{l}(+\infty) = 1.$$

Let us consider the case $\alpha < 0.5$ (which is the case of practical interest in the calculation of these risky quantities). As for the calculation of w_* (in the case $\alpha < 0.5$), easy calculations yield

$$w_* = \begin{cases} \frac{\alpha\beta}{\lambda}, & \text{for } \beta < \lambda, \\ 1 - \frac{\alpha(1-\beta)}{1-\lambda}, & \text{for } \beta \geq \lambda. \end{cases}$$

For a pair of random variables with copula $C_{\lambda}^{\mathbf{Fre}}$, we get

$$\mathrm{CoVaR}^{\leq}_{\alpha,\beta}(Y \mid X) = \begin{cases} \mathrm{VaR}_{\frac{\alpha\beta}{\lambda}}(Y), & \text{for } \beta < \lambda, \\ \mathrm{VaR}_{1-\frac{\alpha(1-\beta)}{1-\lambda}}(Y), & \text{for } \beta \geq \lambda. \end{cases}$$

7.4.5 t-Student copulas

The construction of t-Student copulas is based on the Sklar's theorem.

$$C_t(u,v;\nu,r) = \Phi_{t(\nu,R)}\left(\Phi_{t(\nu,1)}^{-1}(u), \Phi_{t(\nu,1)}^{-1}(v)\right),$$

$$R = \begin{pmatrix} 1 & \rho \\ \rho & 1 \end{pmatrix}, \quad \rho \in (-1,1), \quad \nu >= 1,$$

where $\Phi_{t(\nu,R)}$ and $\Phi_{t(\nu,1)}$ are distribution functions of bivariate and univariate t-Student's probability law.

The t-Student copula has a non-trivial tail expansions at all four corners of the unit square. For the origin (0,0) we have ([36], Thm.2.3)

$$\begin{aligned} L(x,y) = & x\Phi_{t(\nu+1,1)}\left(\frac{\sqrt{\nu+1}}{\sqrt{1-\rho^2}}\left(-\left(\frac{x}{y}\right)^{\frac{1}{\nu}}+\rho\right)\right) \\ & + y\Phi_{t(\nu+1,1)}\left(\frac{\sqrt{\nu+1}}{\sqrt{1-\rho^2}}\left(-\left(\frac{y}{x}\right)^{\frac{1}{\nu}}+\rho\right)\right) \end{aligned}$$

with a generator

$$l(t) = \Phi_{t(\nu+1,1)}\left(\frac{\sqrt{\nu+1}}{\sqrt{1-\rho^2}}\left(-t^{-\frac{1}{\nu}}+\rho\right)\right) + t\Phi_{t(\nu+1,1)}\left(\frac{\sqrt{\nu+1}}{\sqrt{1-\rho^2}}\left(-t^{\frac{1}{\nu}}+\rho\right)\right).$$

Since the reflected t-Student copula is a t-Student copula with $-\rho$ we get at vertex (0,1)

$$\begin{aligned} \widehat{L}(x,y) = & x\Phi_{t(\nu+1,1)}\left(\frac{\sqrt{\nu+1}}{\sqrt{1-\rho^2}}\left(-\left(\frac{x}{y}\right)^{\frac{1}{\nu}}-\rho\right)\right) \\ & + y\Phi_{t(\nu+1,1)}\left(\frac{\sqrt{\nu+1}}{\sqrt{1-\rho^2}}\left(-\left(\frac{y}{x}\right)^{\frac{1}{\nu}}-\rho\right)\right) \end{aligned}$$

with a generator

$$\hat{l}(t) = \Phi_{t(\nu+1,1)}\left(\frac{\sqrt{\nu+1}}{\sqrt{1-\rho^2}}\left(-t^{-\frac{1}{\nu}}-\rho\right)\right) + t\Phi_{t(\nu+1,1)}\left(\frac{\sqrt{\nu+1}}{\sqrt{1-\rho^2}}\left(-t^{\frac{1}{\nu}}-\rho\right)\right).$$

Note that

$$l(+\infty) = \Phi_{t(\nu+1,1)}\left(\frac{\sqrt{\nu+1}}{\sqrt{1-\rho^2}}\rho\right), \quad \hat{l}(+\infty) = \Phi_{t(\nu+1,1)}\left(-\frac{\sqrt{\nu+1}}{\sqrt{1-\rho^2}}\rho\right).$$

Hence similarly as in the previous examples

$$l(+\infty) + \hat{l}(+\infty) = 1.$$

Theorems 7.3 and 7.4 imply that:
For $\beta < l(+\infty)$

$$w_*(\alpha,\beta,C_t) = \alpha l^{-1}(\beta) + o(\alpha).$$

For $\beta > l(+\infty)$

$$w_*(\alpha,\beta,C_t) = 1 - \alpha\hat{l}^{-1}(1-\beta) + o(\alpha).$$

7.4.6 FGM copulas

Now, consider the FGM family of copulas (see [8]), which is given for $(u,v) \in [0,1]^2$ and $\theta \in [-1,1]$ by

$$C_\theta^{\mathbf{FGM}}(u,v) = uv\,(1+\theta\,(1-u)\,(1-v))\,. \tag{7.5}$$

Note that FGM copulas are restrictions of polynomials. Hence they are illustrating theorem 7.5. As for the calculation, we can consider that w_* is the solution of the equation

$$v(1+\theta(1-\alpha)(1-v)) = \beta,$$

which gives

$$w_* = \frac{2\beta}{1+\theta\,(1-\alpha)+\sqrt{\Delta}}, \quad \text{with} \quad \Delta = (1+\theta\,(1-\alpha))^2 - 4\beta\theta\,(1-\alpha)\,.$$

In the independence case (i.e. $\theta = 0$), we have $\mathrm{CoVaR}^{\leq}_{\alpha,\beta}(Y \mid X) = \mathrm{VaR}_\beta(Y)$, as expected. It is of interest to note that

$$\lim_{\alpha\to 0^+} w_* = \frac{2\beta}{1+\theta+\sqrt{1+2\theta(1-2\beta)+\theta^2}}.$$

7.4.7 Marshall–Olkin copulas

Let $(u,v) \in [0,1]^2$, and let $a,b \in (0,1)$ be the copula parameters, then the Marshall–Olkin copula (see, e.g., [8]) is defined as

$$C_{a,b}^{\mathbf{MO}}(u,v) = \begin{cases} u^{1-a}v, & \text{for } u^a \geq v^b, \\ uv^{1-b}, & \text{for } u^a < v^b. \end{cases}$$

Note that theorem 7.5 applies to the Marshall-Olkin family of copulas. Indeed, we construct the extension C^+ in a following way:

$$C^+(u,v) = \begin{cases} u^{1-a}v, & \text{for } u^a \geq v^b \geq 0, \\ uv^{1-b}, & \text{for } 0 \leq u^a < v^b, \\ uv^{1-b}, & \text{for } u < 0 \leq v, \\ 0, & \text{for } v < 0. \end{cases}$$

For the calculation of w_*, it can be checked that

$$w_* = \begin{cases} \beta\alpha^a, & \text{for } \beta \in (0, \alpha^{(1-b)a/b}), \\ \beta^{1/(1-b)}, & \text{for } \beta \in (\alpha^{(1-b)a/b}, 1). \end{cases}$$

In particular, we have

$$\lim_{\alpha\to 0^+} w_* = \beta^{\frac{1}{1-b}}.$$

7.4.8 Gaussian copulas

The construction of Gaussian copulas similarly as t-Student copulas is based on the Sklar's theorem.

$$C_{Ga}(u,v;r) = \Phi_{N(0,R)}\left(\Phi^{-1}_{N(0,1)}(u), \Phi^{-1}_{N(0,1)}(v)\right), \; R = \begin{pmatrix} 1 & r \\ r & 1 \end{pmatrix}, \quad r \in (-1,1).$$

where $\Phi_{N(0,R)}$ and $\Phi_{N(0,1)}$ are distribution functions of bivariate and univariate Gaussian probability law. Note that Gaussian copulas have trivial tail expansions at all four vertices of the unit square (compare [18] Pr. 3). But with positive r they fulfill assumptions of theorem 7.6. Indeed

$$\frac{\partial C_{Ga}(u,v;r)}{\partial u} = \Phi_{N(0,1)}\left(\frac{\Phi^{-1}_{N(0,1)}(v) - r\Phi^{-1}_{N(0,1)}(u)}{\sqrt{1-r^2}}\right).$$

At the limit $u = 0$ we get for $v > 0$ and $r > 0$

$$\frac{\partial C_{Ga}(0^+,v;r)}{\partial u} = 1.$$

Hence,

$$\lim_{\alpha\to 0} \frac{C_{Ga}(\alpha,v;r)}{\alpha} = \frac{\partial C_{Ga}(0^+,v;r)}{\partial u} = 1.$$

By a direct calculation we get:

$$\lim_{\alpha\to 0} w_* = \begin{cases} 0 & r > 0, \\ \beta & r = 0, \\ 1 & r < 0, \end{cases}$$

$$\lim_{\alpha\to 0} \frac{\partial w_*}{\partial \alpha} = \begin{cases} \infty & r > 0, \\ 0 & r = 0, \\ -\infty & r < 0, \end{cases}$$

7.4.9 Archimedean copulas

We recall that a bivariate copula C is called Archimedean (see [35, 33]) if there exist generators ψ and φ such that

$$C_\varphi(u,v) = \psi(\varphi(u)+\varphi(v)).$$

The generators are convex decreasing functions

$$\psi : [0,\infty] \to [0,1], \qquad \varphi : [0,1] \to [0,\infty],$$

such that

$$\psi(0) = 1, \quad \varphi(1) = 0, \quad \text{and} \quad \forall t \in [0,1] \quad \psi(\varphi(t)) = t.$$

We observe that

$$w_* = \psi(\varphi(\alpha\beta) - \varphi(\alpha)). \tag{7.6}$$

If φ is not a strict generator, i.e. $\varphi(0) < +\infty$, then

$$\lim_{\alpha\to 0^+} w_* = \psi(\varphi(0)-\varphi(0)) = \psi(0) = 1. \tag{7.7}$$

which is similar to what we can get from the countermonotonicity copula.

If φ is strict, i.e. $\varphi(0) = +\infty$, and regularly varying at 0 with a negative index, i.e. there exists $d > 0$ such that, for every $x > 0$, $\lim_{t\to 0^+} \frac{\varphi(tx)}{\varphi(t)} = x^{-d}$ (compare [5, 18]), then copula C_φ has a non-trivial tail expansion at the origin

$$L(x,y) = x\left(1+\left(\frac{y}{x}\right)^{-d}\right)^{-\frac{1}{d}} = \left(x^{-d}+y^{-d}\right)^{-\frac{1}{d}},$$

with the generator

$$l(t) = \left(1+t^{-d}\right)^{-\frac{1}{d}}, \quad l(+\infty) = 1,$$

and a trivial one at the vertex (0,1). Due to theorem 7.3 we get in such a case

$$w_*(\alpha,\beta,C_\varphi) = \alpha\left(\beta^{-d}-1\right)^{-\frac{1}{d}} + o(\alpha).$$

7.4.9.1 Clayton copulas

Consider the Clayton copula C_θ given, for $\theta > 0$, by

$$C_\theta^{\mathbf{Cl}}(u,v) = \left(u^{-\theta}+v^{-\theta}-1\right)^{-\frac{1}{\theta}}.$$

Moreover, it can be calculated that

$$w_* = \alpha\beta(1+\alpha^\theta\beta^\theta-\beta^\theta)^{-1/\theta} = \alpha\beta(1-\beta^\theta)^{-1/\theta} + O(\alpha^\theta)$$

which goes almost linearly to 0 as $\alpha \to 0^+$, since the generator of these copulas is regularly varying at 0 with index $-\theta$.

7.4.9.2 Frank copulas

Consider the Frank copula C_θ given, for $\theta \in (-\infty, 0) \cup (0, +\infty)$, by

$$C_\theta^{\mathbf{Fr}}(u,v) = -\frac{1}{\theta}\ln\left(1 + \frac{(\exp(-\theta u) - 1)(\exp(-\theta v) - 1)}{\exp(-\theta) - 1}\right).$$

Then

$$\begin{aligned} w_* &= -\frac{1}{\theta}\ln\left\{1 + \frac{(\exp(-\theta\alpha\beta) - 1)(\exp(-\theta) - 1)}{\exp(-\theta\alpha) - 1}\right\} \\ &= -\frac{1}{\theta}\ln\{1 - \beta(1 - \exp(-\theta))\} + O(\alpha). \end{aligned}$$

Note that although the generator of the Frank copula is regularly varying at 0, the variation index is 0 and w_* need not to tend to 0 as in the case of the Clayton family. On the other side Frank copula is a restriction of an analytic function and theorem 7.5 is applicable.

7.4.9.3 Gumbel copulas

Consider the Gumbel copula C_θ given, for $\theta \in (1, +\infty)$, by

$$C_\theta^{\mathbf{Gu}}(u,v) = \exp\left\{-\left((-\ln(u))^\theta + (-\ln(v))^\theta\right)^{\frac{1}{\theta}}\right\}. \tag{7.8}$$

Then a closed form expression for w_* can be calculated as follows

$$w_* = \exp\left\{-\left((-\ln(\alpha\beta))^\theta - (-\ln(\alpha))^\theta\right)^{\frac{1}{\theta}}\right\}. \tag{7.9}$$

Subsequently

$$\lim_{\alpha\to 0^+} w_* = 0 \quad \text{with} \quad \lim_{\alpha\to 0^+} \frac{w_*}{\alpha} = \infty.$$

Note that similarly as in the case of a Frank copulas, the generator of a Gumbel copula is regularly varying at 0 with index 0. But in this case for $v > 0$

$$\lim_{\alpha\to 0} \frac{C_\theta^{\mathbf{Gu}}(\alpha, v)}{\alpha} = 1$$

and theorem 7.6 is applicable.

Acknowledgements The author acknowledges the support from National Science Centre, Poland, via project 2015/17/B/HS4/00911.

References

1. Adrian, T., Brunnermeier, M. K.: CoVaR, *The American Economic Review* **106.7**, 1705-1741 (2016).
2. Bernardi, M., Durante, F., Jaworski, P.: CoVaR of families of copulas. *Statistics and Probability Letters* **120**, 8-17 (2017).
3. Bernardi, M., Durante, F., Jaworski, P., Petrella, L., Salvadori, G.: Conditional Risk based on multivariate Hazard Scenarios. Submitted.
4. Charpentier, A., Juri,A.: Limiting dependence structures for tail events, with applications to credit derivatives. *J. Appl. Probab.*, **43**(2), 563–586 (2006).
5. Charpentier, A., Segers,J.: Tails of multivariate Archimedean copulas. *Journal of Multivariate Analysis* **100**, 1521–1537 (2009).
6. Cherubini, U., Luciano, E., Vecchiato,W .: *Copula Methods in Finance*, John Wiley & Sons Ltd 2004.
7. Durante, F., Jaworski, P.: Invariant dependence structure under univariate truncation, *Statistics* **46**, 263-277 (2012).
8. Durante, F., Sempi, C.: *Principles of copula theory*, CRC Press, 2016.
9. Durante, F., Jaworski, P., Mesiar, R.: Invariant dependence structures and Archimedean copulas. *Statistics and Probability Letters* **81**, 1995-2003 (2011).
10. Embrechts, P.: Copulas: a personal view. *Journal of Risk and Insurance*, **76**, 639–650 (2009).
11. Fernández-Sánchez, J., Úbeda-Flores, M.: The distribution of the probability mass of conic copulas. *Fuzzy Sets and Systems* **284**, 138–145 (2016).
12. Föllmer, H., Schied, A.: *Stochastic Finance. An Introduction in Discrete Time*, de Gruyter, 2nd Edition, 2004.
13. Girardi, G., Ergün, T. A.: Systemic risk measurement: Multivariate GARCH estimation of CoVar, Journal of Banking & Finance 37, (2013) 3169-3180.
14. Gudendorf, G., Segers, J.: Extreme-Value Copulas In: Jaworski, P., Durante,F., Härdle, W. K. and Rychlik, T. editors. *Copula Theory and Its Applications, Proceedings of the Workshop Held in Warsaw, 25-26 September 2009*, Lecture Notes in Statistics - Proceedings 198. Springer 2010, pp. 126-145.
15. Hakwa, B., Jäger-Ambrozewicz, M., Rüdiger, B.: Analysing systemic risk contribution using a closed formula for conditional Value at Risk through copula. *Commun. Stoch. Anal.*, 9(1):131–158, 2015.
16. Hirsch, M.: *Differential Topology*, Springer-Verlag 1976.
17. Jaworski, P.: Asymptotics of bivariate copulas (in Polish), Matematyka Stosowana 4 (2003) 78-89.
18. Jaworski, P.: On uniform tail expansions of bivariate copulas, *Applicationes Mathematicae*, **31**, 397–415 (2004).
19. Jaworski, P.: On uniform tail expansions of multivariate copulas and wide convergence of measures, *Applicationes Mathematicae*, **33**, 159-184 (2006).
20. Jaworski, P.: Tail Behaviour of Copulas, In: Jaworski, P., Durante, F., Härdle, W. K. and Rychlik, T. editors. *Copula Theory and Its Applications, Proceedings of the Workshop Held in Warsaw, 25-26 September 2009*, Lecture Notes in Statistics - Proceedings 198. Springer 2010, pp. 161-186.
21. Jaworski, P.: The limiting properties of copulas under univariate conditioning, In: Copulae in Mathematical and Quantitative Finance, P. Jaworski, F. Durante, W. K. Härdle (Eds.), Springer 2013, pp.129–163.
22. Jaworski, P.: On the Conditional Value at Risk (CoVaR) for tail-dependent copulas, *Dependence Modeling* **5**, 1–19 (2017).
23. Jaworski, P., Durante, F., Härdle, W. K., Rychlik, T., editors. *Copula Theory and its Applications*. Lecture Notes in Statistics - Proceedings 198, Springer 2010.
24. Jaworski, P., Durante, F., Härdle, W. K., editors. *Copulae in Mathematical and Quantitative Finance*. Lecture Notes in Statistics - Proceedings 213, Springer 20103

25. Joe, H.: *Multivariate models and dependence concepts*, volume 73 of *Monographs on Statistics and Applied Probability*. Chapman & Hall, London, 1997.
26. Joe, H.: *Dependence Modeling with Copulas.* Chapman & Hall/CRC, London, 2014.
27. Joe, H., Li, H., Nikoloulopoulos, A.: Tail Dependence Functions and Vine Copulas, *J. Multivariate Anal.* **101**, 252-270 (2010).
28. Juri, A., Wüthrich, M. V.: Copula convergence theorems for tail events. *Insurance Math. Econom.*, **30**, 405–420 (2002).
29. Jwaid, T., De Baets, B., Kalická, J., Mesiar, R.: Conic aggregation functions, Fuzzy sets and systems **167**, 3-20 (2011)
30. Li, H., Wu, P.: Extremal dependence of copulas: A tail density approach, Journal of Multivariate Analysis, **114**, 99–111 (2013).
31. Łojasiewicz, S.: *An introduction to the theory of real functions*. A Wiley-Interscience Publication. John Wiley & Sons Ltd., Chichester, third edition,(1988).
32. Mainik G., Schaanning E.: On dependence consistency of CoVaR and some other systemic risk measures. *Stat. Risk Model.*, **31**, 49–77 (2014).
33. McNeil ,A. J., Nešlehová, J.: Multivariate Archimedean copulas, d-monotone functions and ℓ_1-norm symmetric distributions. *Ann. Statist.*, 37(5B):3059–3097 (2009).
34. McNeil, A. J., Frey, R., Embrechts, P.. *Quantitative risk management. Concepts, techniques and tools*. Princeton Series in Finance. Princeton University Press, Princeton, NJ, 2005.
35. Nelsen, R. B.: *An introduction to copulas*. Springer Series in Statistics. Springer, New York, second edition, 2006.
36. Nikoloulopoulos, A., Joe, H., Li, H.: Extreme value properties of multivariate t copulas, Extremes **12** (2009) 129–148.

Chapter 8
Parametric copula families for statistical models

Harry Joe

Abstract For data analysis with copulas, one tries to match features seen in data to properties of parametric copula models. Relevant tail asymmetry and dependence properties and measures are summarized. New parametric bivariate copula families in the class of [Durante and Jaworski, 2012] are presented, and some of their dependence and asymmetry properties are determined.

8.1 Introduction

For high-dimensional applications, vine and factor copula models are most useful for their flexible and parsimonious dependence structures. Because these classes of copulas are usually constructed from a sequence of parametric bivariate copulas, I will mention (a) how to match bivariate copula families to asymmetry features seen in bivariate scatterplots and (b) further research needed for constructing 2-parameter and 3-parameter bivariate copula families to cover different tail properties.

I will illustrate construction and analysis of new bivariate parametric copula families with examples satisfying the class in [Durante and Jaworski, 2012]; members of this class have a generator that is either a concave cumulative distribution function (cdf) or a concave survival function. Some links to bivariate Archimedean copulas are mentioned.

Section 8.2 has background on fitting copula models in practice. Section 8.3 has concepts used to assess asymmetries in copulas. Section 8.4 summarizes some useful parametric copula families to use within vines. Section 8.5 has some new parametric bivariate copula families in the class in [Durante and Jaworski, 2012], and theory from Section 8.3 is used to obtained some of their properties. For the remainder of this chapter, if no specific reference for a concept or topic, the general reference is [Joe, 2014].

Harry Joe
University of British Columbia, e-mail: Harry.Joe@ubc.ca

M. Úbeda Flores et al. (eds.), *Copulas and Dependence Models with Applications*,
DOI 10.1007/978-3-319-64221-5_8

8.2 Fitting copula models in practice

Sklar's theorem implies that any d-variate distribution F can be written as a copula C with univariate margins $F_1,\dots,F_d$, as arguments:

$$F_{1\cdots d}(y_1,\dots,y_d) = C(F_1(y_1),\dots,F_d(y_d)).$$

The copula is unique if $F_1,\dots,F_d$ are continuous.

For statistical practice with parametric models, one often considers parametric families for the univariate margins and for the copula, for example,

$$F_{1\cdots d}(y_1,\dots,y_d;\theta,\eta_1,\dots,\eta_d) = C\big(F_1(y_1;\eta_1),\dots,F_d(y_d;\eta_d);\theta\big),$$

where η_j is a parameter vector for the jth margin F_j and θ is a parameter vector for the copula.

The flexible construction of vine copulas is based on a sequential mixing of conditional distributions, so it is indicated below how Sklar's theorem applies for this class.

Let S be a subset of indices from $\{1,\dots,d\}$ with cardinality between 1 and $d-2$ inclusive. Let j,k be two distinct indices that are not in S. Consider the Fréchet class of multivariate distributions with given margins $F_{\{j\}\cup S}$ and $F_{\{k\}\cup S}$ for $(Y_j,\mathbf{Y}_S)$ and $(Y_k,\mathbf{Y}_S)$ respectively, where $\mathbf{Y}_S = (Y_i : i \in S)$. Let $F_{j|S}(\cdot|\mathbf{y}_S)$ and $F_{k|S}(y_k|\mathbf{y}_S)$ be conditional distributions given $\mathbf{Y}_S = \mathbf{y}_S$ for $\mathbf{y}_S$ in the domain of $\mathbf{Y}_S$. If $F_{jk|S}(\cdot|\mathbf{y}_S)$ is known for all $\mathbf{y}_S$, then

$$\begin{aligned} F_{\{j,k\}\cup S}(y_j,y_k,\mathbf{y}_S) &= \int_{(-\infty,\dots,-\infty)}^{\mathbf{y}_S} F_{jk|S}(y_j,y_k|\mathbf{z}_S)\,\mathrm{d}F_S(\mathbf{z}_S) \\ &= \int_{(-\infty,\dots,-\infty)}^{\mathbf{y}_S} C_{jk;S}\big(F_{j|S}(y_j|\mathbf{z}_S)F_{k|S}(y_k|\mathbf{z}_S);\mathbf{z}_S\big)\,\mathrm{d}F_S(\mathbf{z}_S), \end{aligned}$$

where $C_{jk;S}(\cdot;\mathbf{z}_S)$ is a bivariate copula for $F_{jk|S}(\cdot|\mathbf{z}_S)$ with univariate conditional margins $F_{j|S}(\cdot|\mathbf{z}_S)$ and $F_{k|S}(\cdot|\mathbf{z}_S)$.

For statistical modeling, this result is usually applied with $C_{jk;S}(\cdot;\theta_{jk;S})$ being in a parametric family, not depending on $\mathbf{z}_S$, and considered as a copula for $F_{j|S}(\cdot|\mathbf{z}_S)$ and $F_{k|S}(\cdot|\mathbf{z}_S)$, where these conditional distributions have parameters from an earlier stage of the (vine) sequence. When $C_{jk;S}(\cdot;\theta_{jk;S})$ doesn't depend on $\mathbf{z}_S$; this is called the *simplifying assumption* for vines. Note that if all of the above copulas amd distributions are multivariate Gaussian, then the simplifying assumption does hold when Sklar's theorem is applied for $F_{jk|S}$. That is, vine copulas wih the simplifying assumption are an extension of Gaussian copulas after the multivariate Gaussian distribution has been parametrized in terms of $d-1$ correlations and $(d-1)(d-2)/2$ partial correlations that are algebraically independent.

For the dependence structure covered by the copula model, the dimension of θ is of the order $O(d)$ to $O(d^2)$ to cover a range of flexible dependence. If dependence in the variables can be explained by common or group-specific latent factors, then I find that factor copula models (which generalize classical Gaussian factor models) are good-fitting models. If dependence cannot be explained by latent factors, then vine pair-copula constructions (continuous, discrete and mixed) are the most flexible. The vine pair-copula construction is based on a sequence of bivariate copulas; in tree 1, bivariate copulas are applied to univariate margins, and in trees 2 to $d-1$, they apply to conditional univariate margins where the conditioning is based on marginal multivariate distributions that exist from previous trees. Generally for applications of the vine pair-copula construction, one puts pairs of variables with strongest dependence in tree 1 and assumes the simplifying assumption for higher-order trees. If the dimension is large, one would used a truncated vine ([Brechmann et al., 2012]) as a parsimonious dependence model; in a truncated vine, trees at some level m and higher have independence copulas on all edges to represent some conditional independence (given lower order trees).

In summary, vine, truncated vine and factor copulas cover flexible and parsimonious dependence structures, with classical Gaussian models as special cases. But the flexibility depends on having a sufficiently large number of different bivariate parametric copula families. Do there exists a sufficient number of bivariate parametric copulas families to cover different contour shapes that one might see in bivariate scatterplots? This is discussed further below, after some review of how to match features seen in scatterplots to the contours of copula densities.

For d continuous variables, the first step before the copula is to fit univariate margins and then transform to $N(0,1)$ (normal score transform) or use ranks to convert to normal scores. Bivariate scatterplots of normal scores are used to check for deviation from elliptical shape that would be expected if a multivariate Gaussian model were valid.

Candidate bivariate families would depend on the deviations from elliptical, strength of tail dependence (sharpness or roundness in corners relative to ellipse) and tail asymmetries (joint lower versus joint upper, and lower right versus upper left corners). “Tailweights” and measures of tail asymmetry are discussed in Section 8.3, and discussion of bivariate parametric copula families to satisfy different features is given in Section 8.4.

For univariate models for unimodal densities/histograms, as the sample size gets larger, one needs parametric families with two to four parameters to handle asymmetries and tail behavior. Not surprisingly, this is also the case for bivariate copula families for tree 1 of a vine copula model. With stronger dependence, different parametric families are more easily distinguished based on likelihood inference. For weaker dependence, the impact of different families is less, and hence the choice of bivariate copula families has less of an effect in higher-order trees of a vine when the conditional dependence is weaker.

Suppose the original data are $(y_{i1},\dots,y_{id})$ for $i = 1,\dots,n$, and they have been transformed to normal scores. Preliminary data analysis involves looking at bivariate normal scores scatterplots of the pairs with strongest dependence and detecting possible joint tail asymmetries and tail dependence relative to Gaussian. If some variables are discrete, one can compare empirical probability mass functions with fits from discretized multivariate Gaussian distributions or Gaussian copulas.

For factor copulas, results in [Krupskii and Joe, 2013] link strength of dependence in joint tails for bivariate copulas that link observed variables and latent variables to those for two observed variables. The bivariate scatterplots and tail-weighted dependence measures can help in choice of parametric copula families in the vine rooted at one or more latent variables.

For vine models without latent variables, one approach ([Dissmann et al., 2013], [Panagiotelis et al., 2017]) is to assign pairs with strongest dependence to tree 1. For continuous variables, after tree 1, one can compute pseudo-observations of the form

$$v_{ij;k} = \widehat{F}_{j|k}(y_{ij}|y_{ik}) = C_{j|k}(F_j(y_{ij};\hat{\eta}_j)|F_k(y_{ik};\hat{\eta}_k);\hat{\theta}_{jk}),$$
$$v_{ik;j} = \widehat{F}_{k|j}(y_{ik}|y_{ij}) = C_{k|j}(F_j(y_{ik};\hat{\eta}_k)|F_k(y_{ij};\hat{\eta}_j);\hat{\theta}_{jk}),$$

where variables j,k form an edge in tree 1 with fitted copula $C_{jk}(\cdot;\hat{\theta}_{jk})$ and corresponding conditional distributions $C_{j|k}(u_j|u_k;\hat{\theta}_{jk})$ and $C_{k|j}u_k|u_j;\hat{\theta}_{jk})$. If (j_1,k) and (j_2,k) are edges in tree 1, one can look at normal score plots of $(v_{ij_1;k}, v_{ij_2;k})$ for $i=1,\ldots,n$ to see the shape and choose some appropriate copula familes for the tree 2 edge $(j_1|k, j_2|k)$. This could be done for any potential edge for tree 2.

After deciding on edges for tree 2 and fitting bivariate parametric copula families. If $C_{j_1,j_2;k}(\cdot;\hat{\theta}_{j_1,j_2;k})$ is the fitted copula to $(v_{ij_1;k}, v_{ij_2;k})$, then it leads to pseudo-observations $v_{ij_1;j_2k} = \widehat{F}_{j_1|j_2k}(y_{ij_1}|y_{ij_2}, y_{ik})$ and $v_{ij_2;j_1k} = \widehat{F}_{j_2|j_1k}(y_{ij_2}|y_{ij_1}, y_{ik})$ for tree 3. This procedure iterates to the trees at higher levels.

For a small dimension d, it is useful for one to go through the above procedure manually one tree at a time. For a large dimension d, one can automate the procedure; summary statistics to diagnose tail dependence and tail asymmetries can be used to narrow the choices of bivariate copula families for each edge. See Sections 2.15, 2.17 and 5.12.1 of [Joe, 2014] for tail-weighted measures of dependence and measures of tail asymmetries.

8.3 Tail asymmetry and strength of dependence in tails

To assess the strength of dependence in a joint lower or upper tail of a copula, the concept of a *tail order* is analogous to the tailweight for the tails of a univariate density. By comparing the tail orders of the joint lower and joint upper tails, one can get a sense of tail asymmetry.

For copulas, [Hua and Joe, 2011] define that tail orders based on the rates that $C(u\mathbf{1}_d)$ and $\overline{C}((1-u)\mathbf{1}_d)$ go to 0 as $u \to 0^+$. If $C(u\mathbf{1}_d) \sim \ell_L(u)u^{\kappa_L}$ as $u \to 0^+$, where $\ell_L(u)$ is a slowly varying function, then the lower tail order is defined as κ_L. Similarly if $\overline{C}((1-u)\mathbf{1}_d) \sim \ell_U(u)u^{\kappa_U}$ as $u \to 0^+$, where $\ell_U(u)$ is a slowly varying function, then the upper tail order is defined as κ_U. If $1 < \kappa_L < d$ (or $1 < \kappa_U < d$) then $\lambda_L = 0$ (or $\lambda_U = 0$ respectively), and this is termed *intermediate tail dependence* in [Hua and Joe, 2011]. A smaller value of the tail order corresponds to more dependence in the joint tail (more probability in the corner). Then, the strongest dependence in the tail corresponds to $\kappa_L = 1$ or $\kappa_U = 1$. When $\kappa_L = d$ or $\kappa_U = d$ with a constant for the slowly varying function ℓ_L or ℓ_U, then this is referred to as tail orthant independence (or tail quadrant independence for $d = 2$). The tail order can be greater than d for negative dependence.

The tail order, as well as the tail dependence coefficient, are defined as limits. For data, tail-weighted dependence measures can be used to assess strength of tail dependence relative to Gaussian. For example, the lower and upper semi-correlations of normal scores are defined for bivariate copulas, and these have sample counterparts for data — say $\hat{\rho}^-_{jk,L}$ and $\hat{\rho}^+_{jk,U}$ for variables j,k. For a bivariate Gaussian copula with correlation ρ, the lower and upper semi-correlations are both $\zeta(\rho)$ as given in equation (2.59) in [Joe, 2014]. An indicator of stronger lower tail dependence when $\hat{\rho}^-_{jk,L} > \zeta(\hat{\rho}_{jk,N})$, where $\hat{\rho}_{jk,N}$ is the correlation of normal scores for variables j,k.

Other measures of tail-weighted dependence are in [Krupskii and Joe, 2015]. See also [Krupskii, 2017], and references therein, for (i) measures of reflection asymmetry (to summarize the probability in the joint upper tail has relative to that in the joint lower tail), and (ii) measures of bivariate permutation asymmetry (to summarize the tail probability in the lower right corner relative to that in the upper left corner).

A bivariate copula is reflection symmetric if $C(u,v) = \widehat{C}(u,v)$ for all $(u,v) \in [0,1]^2$, where $\widehat{C}(u,v) = u+v-1+C(1-u,1-v)$. If $(U,V) \sim C$ so that the reflection $(1-U,1-V) \sim \widehat{C}$, then reflection symmetry implies that $(U,V) \stackrel{d}{=} (1-U,1-V)$. A bivariate copula is permutation symmetric if $C(u,v) = C(v,u)$ for all $(u,v) \in [0,1]^2$. If $(U,V) \sim C$, then permutation symmetry implies that $(U,V) \stackrel{d}{=} (V,U)$.

For copula families, these corner probabilities can also be analyzed the appropriate tail orders; examples are shown in Section 8.5 when analyzing some new parametric copula families for the two tail asymmetries: reflection and permutation.

8.4 Parametric bivariate families

This section summarizes some bivariate parametric copula that could be useful within vines. Based on the tail asymmetries mentioned in Section 8.3, indications are given for what might be lacking.

Simple bivariate parametric copula families have monotonic dependence between that two variables, so that the common parametric families are unsuitable for non-monotonic dependence. For the latter, one might consider non-parametric methods (see [Nagler and Czado, 2016]). But for monotonic dependence, it is simpler to use parametric models to avoid the "curse of dimensionality" of non-parametric methods for high-dimensional data.

There are many 1-parameter bivariate Archimedean families in [Nelsen, 2006] and there are several 2-parameter bivariate Archimedean and non-Archimedean families in [Joe, 2014]. They can cover a variety of tail behavior in the joint lower and upper tails, but they assume permutation symmetry. There is a need for simple

3-parameter and 4-parameter bivariate copula families that can add some permutation asymmetry.

Some of the useful copula families with permutation symmetry include the following.

- 2-parameter BB1 Archimedean copula (8.3) based on $\psi(s;\theta,\delta) = (1+s^{1/\delta})^{-1/\theta}$ for $\theta > 0$ and $\delta > 1$. This has asymmetric lower/upper tail dependence. For data, sometimes the reflected BB1 copula provides a better fit.
- 2-parameter Archimedean copula (8.3) based on the integrated Mittag-Leffler Laplace transform $\psi(s;\theta,\delta) = 1 - F_B(s^{1/\delta}/(1+s^{1/\delta});\delta,\theta^{-1})$ where $\theta > 0$, $\delta > 1$ and $F_B(\cdot;a,b)$ is the cdf of the Beta(a,b) random variable. This has upper tail dependence ($\kappa_U = 1$) and lower intermediate tail dependence ($1 < \kappa_L < 2$) and the reflected copula has lower tail dependence and upper intermediate tail dependence.
- 2-parameter BB8 copula: this has asymmetric tails with $\kappa_L = \kappa_U = 2$.

There are several 1-parameter and 2-parameter Archimedean copulas with $(\kappa_L, \kappa_U) = (1,2)$ or $(2,1)$ but these are quite tail asymmetric and often provide poorer fits to data. There is no Archimedean copulas in the above references with one κ equal to 2 and the other in the interval $(1,2)$.

For bivariate copulas with flexible permutation asymmetry, some useful families are the following.

- The 3-parameter bivariate skew-normal copula based on the bivariate skew-normal distribution in [Azzalini and Dalla Valle, 1996].
- The 4-parameter bivariate skew-t copula based on the bivariate skew-t distribution in [Azzalini and Capitanio, 2003].
- A permutation symmetric bivariate copula $C(u,v)$ with an added parameter to attain permutation asymmetry. For example, with $0 \le \alpha \le 1$, define $C^*(u,v) = u^\alpha C(u^{1-\alpha},v)$ or $C^*(u,v) = v^\alpha C(u,v^{1-\alpha})$.

For bivariate/multivariate skew-normal and skew-t distributions, see [Azzalini, 2013], [Yoshiba, 2015] shows that these can be implemented for faster maximum likelihood estimation if monotone interpolation is used for the univariate skew-normal and skew-t inverse cdfs.

For a given parametric copula family, there are associated families based on reflected $U(0,1)$ random variables; these are useful to reorient the tail asymmetry or to get negative dependence. Suppose $(U,V) \sim C$ for a bivariate copula C. The copula of $(1-U,1-V)$ is denoted as $\widehat{C}(u,v) = u+v-1+C(1-u,1-v)$, and is called the *survival or reflected copula*. The copula of $(1-U,V)$ is $v - C(1-u,v)$ and is called the *1-reflected copula*, and the copula of $(U,1-V)$ is $u - C(u,1-v)$ and is called the *2-reflected copula*. This terminology is used in [Panagiotelis et al., 2017]. These copulas are sometimes referred to as rotated copulas, but they are not rotations of the (U,V); they are rotations of the contours of the copula density when $C(u,v) = C(v,u)$. The operations to get 1-reflected and 2-reflected copulas can change a family with positively dependence to one with negative dependence. The operation to get a survival copula can change the direction of skewness relative to joint lower/upper tails.

8.5 Durante-Jaworski class of bivariate copulas

In this section, some properties are obtained for the class of bivariate copulas in [Durante and Jaworski, 2012]. Based on these properties, some parametric copulas families that interpolate for independence to comonotonicity or countermonotonicity are derived. Then their tail properties are obtained. The latter parts of this section show steps that one can go through to determine properties and potential usefulness of new constructions of copula families.

The Durante-Jaworski (DJ) class generally have permutation asymmetry and tail properties not in Archimedean families, but some links are shown between the DJ class and the class of bivariate Archimedean copulas due to the two classes having generators with common monotonicity properties.

In the remainder of this chapter, the superscript $\leftarrow$ to denotes functional inverse of a monotonically increasing or decreasing function.

Let $\varphi : [0,\infty) \to [0,1]$ be a function satisfying (i) φ is concave increasing with $\varphi(0) = 0$, $\varphi(\infty) = 1$, or (ii) φ is convex decreasing with $\varphi(0) = 1$, $\varphi(\infty) = 0$. Through univariate truncation, [Durante and Jaworski, 2012] show that

$$C_\varphi(u,v) = u\varphi[\varphi^{\leftarrow}(v)/u], \quad 0 \le u \le 1,\ 0 \le v \le 1, \tag{8.1}$$

is a bivariate copula cdf. The generator φ in (8.1) is a concave cdf on $[0,\infty)$ or a convex survival function on $[0,\infty)$. Some basic properties are given next as they are not stated together previously — P3 is in [Durante and Jaworski, 2012]; P4–P5 are in Section 7.6.1 of [Jaworski, 2013]; and P6 is Proposition 2.1 in [Durante et al., 2011].

P1. Invariance to scale changes. If $\varphi(z)$ is a convex survival function or concave cdf on $[0,\infty)$, then so is $\varphi_2(z) = \varphi(\alpha z)$ for $\alpha > 0$. Note that, with $\varphi_2^{\leftarrow}(p) = \alpha^{-1}\varphi^{\leftarrow}(p)$,

$$C_{\varphi_2}(u,v) = u\varphi_2[u^{-1}\varphi_2^{\leftarrow}(v)] = u\varphi_2[u^{-1}\alpha^{-1}\varphi^{\leftarrow}(v)] = u^{-1}\varphi[u^{-1}\varphi^{\leftarrow}(v)].$$

P2. Power change: If $\varphi(z)$ is a convex survival function or concave cdf on $[0,\infty)$, then so is $\varphi_2(z) = \varphi(z^\alpha)$ for $0 < \alpha \le 1$. Note that, with $\varphi_2^{\leftarrow}(p) = [\varphi^{\leftarrow}(p)]^{1/\alpha}$,

$$C_{\varphi_2}(u,v) = u\varphi_2[u^{-1}\varphi_2^{\leftarrow}(v)] = u\varphi_2[u^{-1}\{\varphi^{\leftarrow}(v)\}^{1/\alpha}] = u\varphi[u^{-\alpha}\varphi^{\leftarrow}(v)].$$

P3. From cdf to survival function, and vice versa. If φ is a concave cdf and $\varphi_2 = 1-\varphi$ is a convex survival function (and vice versa). Since $\varphi_2^{\leftarrow}(p) = \varphi^{\leftarrow}(1-p)$,

$$C_{\varphi_2}(u,v) = u\varphi_2[u^{-1}\varphi^{\leftarrow}(1-v)] = u\{1-\varphi[u^{-1}\varphi^{\leftarrow}(1-v)]\} = u - C_\varphi(u,1-v).$$

Hence $C_{\varphi_2}(u,v)$ is the 2-reflection of $C_\varphi(u,v)$.

P4. Comonotonicity: $\varphi(z) = \min\{z,1\}$, for $z > 0$, is the cdf of $U(0,1)$ and it is concave. Since $\varphi^{\leftarrow}(p) = p$, this leads to:

$$C_\varphi(u,v) = u\min\{u^{-1}\varphi^{\leftarrow}(v),1\} = u\min\{u^{-1}v,1\} = \min\{u,v\} = C^+(u,v).$$

Hence the comonotonicity copula C^+ is within the DJ class.

P5. Countermonotonicity: From the above (items 3 and 4), this obtains with the $U(0,1)$ survival function $\varphi(z) = \max\{1-z,0\}$ and $\varphi^{\leftarrow}(p) = 1-p$, and

$$C_\varphi(u,v) = u\max\{1-u^{-1}(1-v),0\} = \max\{u-1+v,0\} = C^-(u,v).$$

P6. Positive quadrant dependence (PQD) holds if φ is a concave cdf and negative quadrant dependence (NQD) holds if φ is a convex survival function. That is, $C_\varphi(u,v) \ge uv$ in $[0,1]^2$ if φ is a concave cdf and $C_\varphi(u,v) \le uv$ in $[0,1]^2$ if φ is a convex survival function.
Proof. If φ is a concave cdf, then φ is increasing and

$$u\varphi(u^{-1}\varphi^{\leftarrow}(v)) \ge uv \iff \varphi(u^{-1}\varphi^{\leftarrow}(v)) \ge v \iff u^{-1}\varphi^{\leftarrow}(v)) \ge \varphi^{\leftarrow}(v)).$$

If φ is a convex survival function, φ is decreasing and

$$u\varphi(u^{-1}\varphi^{\leftarrow}(v)) \le uv \iff \varphi(u^{-1}\varphi^{\leftarrow}(v)) \le v \iff u^{-1}\varphi^{\leftarrow}(v)) \ge \varphi^{\leftarrow}(v)).$$

Other properties are stated as propositions. These include stochastic representation, concordance ordering and a formula for Kendall's τ that is a one-dimensional integral similar to that for bivariate Archimedean copulas. The stochastic representation in the next proposition is different from that in [Durante et al., 2011], but is equivalent to that in Theorem 2.2 of [Di Lascio et al., 2016].

Proposition 8.1. *(Stochastic representation). Let φ be a concave cdf on $[0,\infty)$. and let $F_Y(y) = \varphi(y)$ be the cdf of Y. Let Z be defined so that its conditional distribution given $Y = y$ is: $\overline{F}_{Z|Y}(z|y) = f_Y(z)/f_Y(y)$, for $z \ge y$; this is well defined because f_Y is decreasing. Then $Y/Z \sim U(0,1)$ and the copula of Y/Z and Y is C_φ.*

Proof: For $0 < u < 1$,

$$\begin{aligned}\Pr(Y/Z \le u) &= \Pr(Z \ge Y/u) = \int_0^\infty \Pr(Z > y/u | Y = y)\, f_Y(y)\,\mathrm{d}y \\ &= \int_0^\infty \frac{f_Y(y/u)}{f_Y(y)} f_Y(y)\,\mathrm{d}y = u\int_0^\infty f_Y(y/u)\,u^{-1}\mathrm{d}y = u.\end{aligned}$$

For the joint distribution of Y/Z and Y,

$$\begin{aligned}\Pr(Y/Z \le u, Y \le y) &= \int_0^y \Pr(Z > s/u | Y = s)\, f_Y(s)\,\mathrm{d}s = u\int_0^y f_Y(s/u)\,u^{-1}\mathrm{d}s \\ &= u\int_0^{y/u} f_Y(t)\,\mathrm{d}t = uF_Y(y/u) = u\varphi(u^{-1}y).\end{aligned}$$

Then the copula is $u\varphi[u^{-1}\varphi^{\leftarrow}(v)]$, for $u,v \in [0,1]$. □

Proposition 8.2. *(Concordance ordering).*

(a) Let φ_1, φ_2 be concave cdfs on $[0,\infty)$. If $\varphi_2^{\leftarrow} \circ \varphi_1$ is anti-starshaped with respect to the origin, then $C_{\varphi_1}(u,v) \le C_{\varphi_2}(u,v)$ for all $(u,v) \in [0,1]^2$.
(b) Let φ_1, φ_2 be convex survival functions on $[0,\infty)$. If $\varphi_2^{\leftarrow} \circ \varphi_1$ is anti-starshaped with respect to the origin, then $C_{\varphi_1}(u,v) \ge C_{\varphi_2}(u,v)$ for all $(u,v) \in [0,1]^2$.

Proof: (a)

$$\begin{aligned}
& u\varphi_1[u^{-1}\varphi_1^{\leftarrow}(v)] \le u\varphi_2[u^{-1}\varphi_2^{\leftarrow}(v)], \quad \forall u,v \in (0,1),\\
\iff & u^{-1}\varphi_1^{\leftarrow}(v)] \le \varphi_1^{\leftarrow} \circ \varphi_2[u^{-1}\varphi_2^{\leftarrow}(v)], \quad \forall u,v \in (0,1),\\
\iff & u^{-1}y \le \varphi_1^{\leftarrow} \circ \varphi_2[u^{-1}\varphi_2^{\leftarrow} \circ \varphi_1(y)], \quad \forall u \in (0,1), y>0,\\
\iff & b(u^{-1}y) \le u^{-1}b(y), \quad \forall u \in (0,1), y>0,
\end{aligned}$$

where $b = \varphi_2^{\leftarrow} \circ \varphi_1 : [0,\infty) \to [0,\infty)$. Since $\varphi_j(0) = 0$ for a continuous concave cdf, the condition is the same as b anti-starshaped with respect to the origin. Let $y^* = u^{-1}y > y$. The above implies $b(u^{-1}y)/[u^{-1}y] \le b(y)/y$ or $b(y^*)/y^* \le b(y)/y$ for $y^* > y > 0$, which is the definition of anti-starshapedness.

The proof of (b) is similar, or can make use of Property P3. □

Proposition 8.3. *(Kendall's τ). Let $\tau(C_\varphi)$ be the Kendall tau value for the copula C_φ in* (8.1). *Then*

$$\tau(C_\varphi) = \begin{cases} 2\int_0^\infty x[\varphi'(x)]^2 \mathrm{d}x, & \varphi \text{ concave cdf},\\ -2\int_0^\infty x[\varphi'(x)]^2 \mathrm{d}x, & \varphi \text{ convex survival}. \end{cases} \tag{8.2}$$

Proof: We use the formula in [Fredricks and Nelsen, 2007] and Section 2.12.1 of [Joe, 2014]. With conditional distributions $C_{1|2}(u|v)$ and $C_{2|1}(v|u)$, then $\tau = 1 - 4\int_{[0,1]^2} C_{2|1}(v|u)C_{1|2}(u|v)\,\mathrm{d}u\,\mathrm{d}v$. For C_φ, the conditional distributions are (i) $C_{1|2}(u|v) = \varphi'[u^{-1}\varphi^{\leftarrow}(v)]/\varphi'[\varphi^{\leftarrow}(v)]$, and (ii) $C_{2|1}(v|u) = \gamma_\varphi(u^{-1}\varphi^{\leftarrow}(v))$, where $\gamma_\varphi(t) = \varphi(t) - t\varphi'(t)$ for $t \ge 0$. First consider φ being a concave cdf. Let $y = \varphi^{\leftarrow}(v)$, $x = u^{-1}\varphi^{\leftarrow}(v) = u^{-1}y$, with $x \ge y$. Then $u = y/x$, $v = \varphi(y)$, $\mathrm{d}v = \varphi'(y)\,\mathrm{d}y$, $\mathrm{d}u = -yx^{-2}\mathrm{d}x$, and

$$\begin{aligned}
\int_{[0,1]^2} C_{2|1}(v|u)C_{1|2}(u|v)\,\mathrm{d}u\,\mathrm{d}v &= \int_{y=0}^\infty \int_{x=y}^\infty [\varphi(x) - x\varphi'(x)]\varphi'(x)yx^{-2}\mathrm{d}x\,\mathrm{d}y\\
&= \int_{x=0}^\infty [\varphi(x) - x\varphi'(x)]\varphi'(x)x^{-2}\int_{y=0}^x y\,\mathrm{d}y\,\mathrm{d}x\\
&= \tfrac{1}{2}\int_0^\infty [\varphi(x) - x\varphi'(x)]\varphi'(x)\mathrm{d}x = \tfrac{1}{2}\Big[\tfrac{1}{2} - \int_0^\infty x[\varphi'(x)]^2\mathrm{d}x\Big].
\end{aligned}$$

Hence for a concave cdf φ, $\tau(C_\varphi) = 2\int_0^\infty x[\varphi'(x)]^2 \mathrm{d}x$.

Next, consider φ being a convex survival function. Starting with the same steps: let $y = \varphi^{\leftarrow}(v)$, $x = u^{-1}\varphi^{\leftarrow}(v) = u^{-1}y$, with $x \ge y$. Then $\varphi' \le 0$, $\mathrm{d}v = \varphi'(y)\,\mathrm{d}y$, $\mathrm{d}u = -yx^{-2}\mathrm{d}x$,

$$\begin{aligned}&\int_{[0,1]^2} C_{2|1}(v|u)\,C_{1|2}(u|v)\,\mathrm{d}u\,\mathrm{d}v\\&= -\int_{y=0}^{\infty}\int_{x=y}^{\infty}[\varphi(x)-x\varphi'(x)]\varphi'(x)yx^{-2}\mathrm{d}x\,\mathrm{d}y\\&= -\tfrac{1}{2}\int_0^{\infty}[\varphi(x)-x\varphi'(x)]\varphi'(x)\mathrm{d}x = -\tfrac{1}{2}[-\tfrac{1}{2}-\int_0^{\infty}x[\varphi'(x)]^2\mathrm{d}x.\end{aligned}$$

Hence, for a convex survival function φ, $\tau(C_\varphi) = -2\int_0^\infty x[\varphi'(x)]^2\mathrm{d}x$. □

Some additional tail properties are given in Subsection 8.5.2. Other tail properties and extensions of the DJ class are given in [Coia, 2017].

8.5.1 Connections to Archimedean copulas

In this subsection, Propositions 8.2 and 8.3 are used to show connections between the DJ class with positive dependence and the class of bivariate Archimedean copulas. These classes have the same generators as shown below. See also Section 3 of [Durante et al., 2011] for other similarities of the two classes.

For a bivariate Archimedean copula

$$C_{A,\psi}(u,v) = \psi(\psi^{\leftarrow}(u)+\psi^{\leftarrow}(v)), \tag{8.3}$$

the generator $\psi : [0,\infty) \to [0,1]$ is a Laplace transform (LT) or a Williamson 2-transform. That is, ψ is a convex survival function. Hence $\varphi = 1-\psi$ is a concave cdf and can be used as a generator for the DJ class to get positive dependence, and $\varphi = \psi$ as a convex survival function can be used for the DJ class to get negative dependence. Note that this representation of Archimedean copulas, as given in [Joe, 1997] and [McNeil and Nešlehová, 2009], is needed to show the connection, and not the form of generator in [Nelsen, 2006].

Note the following similarities/contrasts.

1. Let ψ be a decreasing convex survival function on $[0,\infty)$. For an Archimedean copula, Kendall's τ is

   $$\tau_A(C_{A,\psi}) = 1-4\int_0^{\infty} s[\psi'(s)]^2\,ds,$$

   as given in Theorem 4.3 of [Joe, 1997]. For a DJ copula, Kendall's τ is (from Proposition 8.3) given in (8.2).
2. For two generators ψ_1,ψ_2, the concordance ordering for Archimedean copulas $C_{\psi_1} \prec_c C_{\psi_2}$ holds iff $\psi_2^{\leftarrow}\circ\psi_1$ is superadditive; starshapedness is a sufficient condition for superadditivity. See Theorem 4.1 of [Joe, 1997]. For two generators φ_1,φ_2 that are concave cdfs, the concordance ordering for DJ copulas $C_{\varphi_1} \prec_c C_{\varphi_2}$ holds iff $\varphi_2^{\leftarrow}\circ\varphi_1$ is anti-starshaped (from Proposition 8.2). Hence with a family of Williamson 2-transforms or corresponding cdfs that are con-

cave and satisfy starshapedness or anti-starshapedness, the orderings of concordance for Archimedean copulas and DJ copulas are in opposite directions.

Let $\varphi = 1-\psi$ be a cdf where ψ is a Williamson 2-transform, then $\int_0^\infty z[\varphi'(z)]^2\,dz = \int_0^\infty s[\psi'(s)]^2\,ds$, and

$$\begin{aligned}\tau_{DJ}(C_\varphi) &= \tfrac{1}{2}[1-\tau_A(C_{A,\psi})],\\ \tau_{DJ}(C_\psi) &= -\tfrac{1}{2}[1-\tau_A(C_{A,\psi})].\end{aligned}$$

Suppose $\psi(\cdot;\theta) = 1-\varphi(\cdot;\theta)$ is a 1-parameter family such that the corresponding Archimedean copulas are increasing in concordance.

- If $\tau_A(C_{A,\psi(\cdot;\theta)})$ goes from 0 to 1 (independence $C^\perp$ to comonotonicity C^+ as θ increases, then $\tau_{DJ}(C_{\varphi(\cdot;\theta)})$ goes from $\frac{1}{2}$ to 0 as θ increases and $\tau_{DJ}(C_{\psi(\cdot;\theta)})$ goes from $-\frac{1}{2}$ to 0 as θ increases.
- If $\tau_A(C_{A,\psi(\cdot;\theta)})$ goes from -1 to 1 (countermonotonicity C^- to comonotonicity C^+ as θ increases, then $\tau_{DJ}(C_{\varphi(\cdot;\theta)})$ goes from 1 to 0 as θ increases and $\tau_{DJ}(C_{\psi(\cdot;\theta)})$ goes from -1 to 0 as θ increases.

Therefore, from families of parametric Archimedean copulas that range from C^- to C^+, we can get parametric DJ copulas that range from $C^\perp$ to C^+ with the same family of generators. For bivariate Archimedean copulas, a sufficient condition (Corollary 8.23 of [Joe, 2014]) for PQD is that $-\log\psi$ is concave (in addition to ψ decreasing convex) and a a sufficient condition for NQD is that $-\log\psi$ is convex.

With $\varphi = 1-\psi$ being a concave cdf, the following hold.

- If ψ is a decreasing convex survival function and $-\log\psi$ is concave (or ψ is a survival function with decreasing failure rate; see [Barlow and Proschan, 1981], then $\tau_A(C_{A,\psi}) \geq 0$ and $\tau_{DJ}(C_\varphi) \in [0,\frac{1}{2}]$.
- If $\psi(s) = 1-e^{-s}$ so that the Archimedean copula is the independence copula, then $\tau_A(C_{A,\psi}) = 0$ and $\tau_{DJ}(C_\varphi) = \frac{1}{2}$.
- If ψ is a decreasing convex survival function and $-\log\psi$ is convex (equivalent to ψ being a survival function with increasing failure rate), then $\tau_A(C_{A,\psi}) \leq 0$ and $\tau_{DJ}(C_\varphi) \in [\frac{1}{2},1]$.

Note that any generator of Archimedean copulas that is a Laplace transform satisfies concavity for $-\log\psi$ so that $\tau_{DJ}(C_\varphi) \in [0,\frac{1}{2}]$ when $\varphi = 1-\psi$. To get a Kendall τ value exceeding $\frac{1}{2}$, one can consider survival functions ψ that are decreasing convex and have increasing failure rate (and hence not Laplace transforms). Examples of parametric copula families in the DJ class using known families of Archimedean generators are given next.

1. Positive stable LTs or Weibull distributions. With $\varphi(z;\alpha) = 1-\exp\{-z^\alpha\}$ for $0<\alpha\leq 1$, one has $\varphi^{\leftarrow}(p;\alpha) = [-\log(1-p)]^{1/\alpha}$ and

$$C_\varphi(u,v;\alpha) = u[1-(1-v)^{1/u^\alpha}], \quad 0\leq u,v\leq 1.$$

Kendall's τ goes from 0 to $\frac{1}{2}$ as α goes from 0 to 1.

2. Gamma LTs and extension to Williamson 2-transforms, or Pareto distributions and extension to Beta distributions (combined as generalized Pareto distributions). With notation $x_+ = \max\{0,x\}$, let $\varphi(z;\xi) = 1-(1+\xi z)_+^{-1/\xi}$ and $\varphi^{\leftarrow}(p;\xi) = [(1-p^{-\xi}-1]/\xi$ for $-1<\xi<\infty$ (with the exponential distribution in the limit as $\xi \to 0$). This leads to:

$$C_\varphi(u,v;\xi) = u\{1-(1+u^{-1}[(1-v)^{-\xi}-1])_+^{-1/\xi}\}, \quad 0 \le u,v \le 1.$$

Kendall's τ goes from 1 to 0 as ξ goes from -1 to ∞. When $-1<\xi<0$, the Archimedean copulas based on $1-\varphi$ or the DJ copulas based on φ do not have support on all of $[0,1]^2$; the support of the latter is on the set $\{(u,v):(1-u) \le (1-v)^{-\xi}\}$. This includes the corner point $(1,0)$ but not the corner point $(0,1)$.
3. Logarithmic series LTs and extension to Williamson 2-transforms. Let $\varphi(z;\theta) = 1+\theta^{-1}\log[1-(1-e^{-\theta})e^{-z}]$, for $-\infty<\theta<\infty$, with $\varphi^{\leftarrow}(p;\theta) = -\log[(1-e^{-\theta(1-p)})/(1-e^{-\theta})]$ for $0<p<1$. The DJ copula family becomes:

$$C_\varphi(u,v;\theta) = u\Big\{1+\theta^{-1}\log\Big[1-(1-e^{-\theta})\Big(\frac{1-e^{-\theta(1-v)}}{1-e^{-\theta}}\Big)^{1/u}\Big]\Big\}, \quad 0 \le u,v \le 1.$$

This goes from C^+ to $C^\perp$ as θ increases from $-\infty$ to ∞. The support of the copula is all of $[0,1]^2$.
4. Integrated positive stable LTs and extension to Williamson 2-transforms; cdfs are generalized gamma. For the Archimedean copula family, see Section 4.11 of Joe (2014). Let $F_\Gamma(\cdot;\eta)$ be the Gamma$(\eta,1)$ cdf with $\eta>0$. Let $\varphi(z;\theta) = F_\Gamma(z^{1/\theta};\theta)$, for $0<\theta<\infty$, with $\varphi^{\leftarrow}(p;\theta) = [F_\Gamma^{\leftarrow}(p;\theta)]^\theta$ for $0<p<1$. The DJ copula family becomes:

$$C_\varphi(u,v;\theta) = uF_\Gamma(u^{-1/\theta}F_\Gamma^{\leftarrow}(v;\theta);\theta), \quad 0 \le u,v \le 1.$$

This goes from C^+ to $C^\perp$ as θ increases from 0 to ∞. The support of the copula is all of $[0,1]^2$.

One can also go in the opposite direction from the DJ class to a family of Archimedean copulas. [Durante and Jaworski, 2012] show in their Theorem 4.1 that the bivariate Mardia-Takahasi-Clayton-Cook-Johnson (MTCJ) copula also in DJ class with $\varphi(z;\alpha) = (1+z^{-\alpha})^{-1/\alpha}$ with $\alpha>0$ (independence copula as $\alpha \to 0^+$ and comonotonicity copula C^+ as $\alpha \to \infty$). Since the MTCJ family is ordered by concordance, $\varphi^{\leftarrow}(\varphi(z;\alpha_1);\alpha_2)$ is anti-starshaped for $0<\alpha_1<\alpha_2$.

With $\psi(s;\alpha) = 1-(1+s^{-\alpha})^{-1/\alpha}$, one gets a family of Archimedean copulas:

$$C_{A,\psi}(u,v;\alpha) = 1-\big\{1+\big[((1-u)^{-\alpha}-1)^{-1/\alpha}+((1-v)^{-\alpha}-1)^{-1/\alpha}\big]^{-\alpha}\big\}^{-1/\alpha}, \tag{8.4}$$

for $\alpha>0$. Note that (8.4) decreases in concordance from C^+ to C^- as α goes from 0 to ∞, and its Kendall's tau value is $\tau = (2-\alpha)/(2+\alpha)$. For $\alpha=2$, $\tau=0$ but the copula is not the independence copula; the copula can be greater than or less than uv, depending on (u,v). Also $\psi'(s;\alpha) = -(1+s^\alpha)^{-1/\alpha-1}<0$, $\psi''(s;\alpha) =$

$(1+\alpha)(1+s^{\alpha})^{-1/\alpha-2}s^{\alpha-1} > 0$ and further derivatives alternative in sign only if $0 < \alpha \le 1$. That is, $\psi(s;\alpha) = 1-(1+s^{-\alpha})^{-1/\alpha}$ is a LT only for $0 < \alpha \le 1$.

The copula (8.4) has some tail properties not common for Archimedean copulas. (i) The lower tail dependence parameter is $\lambda_L = 2^{-\alpha}$ and there is (weak) lower tail dependence even for $\alpha \ge 2$ with $\tau \le 0$. (ii) The upper tail order is $\kappa_U = \alpha+1$. Hence there is intermediate upper tail dependence ($\kappa_U < 2$) for $0 < \alpha < 1$, and negative tail quadrant dependence ($\kappa_U > 2$) for $\alpha > 1$. (iii) Using Theorem 8.38 in [Joe, 2014], the tail order in the $(0,1)$ and $(1,0)$ corners behaves like that of $-\psi'(\psi^{\leftarrow}(u))\cdot\psi^{\leftarrow}(1-u)$ as $u \to 0^+$, leading to $O(u^{2+1/\alpha})$. (iv) The combination of these tail properties is especially clear in contour plots of the density $c(\Phi(y_1),\Phi(y_2))\,\phi(y_1)\,\phi(y_2)$ for α near 2 (overall negative dependence with joint lower tail dependence); here Φ,ϕ are the standard normal cdf and density respectively.

8.5.2 Asymmetries, tail properties and orders

In this subsection, tail properties of (8.1) are studied using the tail order in all four corners. Proposition 4.1 of [Durante et al., 2011] has results for tail dependence only. The results are for cases of positive dependence with concave cdfs, and the corresponding results for convex survival functions follow with 2-reflection.

The tail orders in the different corners mainly depend on the behavior of $\varphi(z)$ as $z \to 0^+$ and as $z \to \infty$. Next, the form of tails are given for the Pareto, Beta, Weibull, logseries-mixture and generalized gamma cdfs, in Section 8.5.1.

- Pareto: $\varphi(z;\alpha) = 1-(1+z)^{-\alpha}$ for $\alpha > 0$, with $\varphi^{\leftarrow}(p;\alpha) = (1-p)^{-1/\alpha}-1$. As $z \to \infty$, $\varphi(z;\alpha) \sim 1-z^{-\alpha}$ and as $z \to 0$, $\varphi(z;\alpha) \sim \alpha z$. As $p = 1-q \to 1$, $\varphi^{\leftarrow}(p;\alpha) \sim q^{-1/\alpha}$ and as $p \to 0$, $\varphi^{\leftarrow}(p;\alpha) \sim p/\alpha$.
- Beta$(1,\gamma)$: $\varphi(z;\gamma) = 1-(1-z)^{\gamma}$ for $\gamma > 1$, with $\varphi^{\leftarrow}(p;\gamma) = 1-(1-p)^{1/\gamma}$. As $z \to 0$, $\varphi(z;\gamma) \sim \gamma z$. As $p = 1-q \to 1$, $\varphi^{\leftarrow}(p;\gamma) \sim 1-q^{1/\gamma}$. As $p \to 0$, $\varphi^{\leftarrow}(p;\gamma) \sim \gamma^{-1}p$.
- Weibull: $\varphi(z;\alpha) = 1-\exp\{-z^{\alpha}\}$ for $0 < \alpha \le 1$, with $\varphi^{\leftarrow}(p;\alpha) = [-\log(1-p)]^{1/\alpha}$. As $z \to 0$, $\varphi(z;\alpha) \sim z^{\alpha}$. As $p = 1-q \to 1$, $\varphi^{\leftarrow}(p;\alpha) \sim [-\log q]^{1/\alpha}$. As $p \to 0$, $\varphi^{\leftarrow}(p;\alpha) \sim p^{1/\alpha}$.
- logseries-mixture: $\varphi(z;\theta) = 1+\theta^{-1}\log[1-(1-e^{-\theta})e^{-z}]$, for $-\infty < \theta < \infty$. As $z \to \infty$, $\varphi(z;\theta) \sim 1-m(\theta)e^{-z}$ with $m(\theta) = \theta^{-1}(1-e^{-\theta}) > 0$. As $p \to 1$, $\varphi^{\leftarrow}(p;\theta) \sim -\log[(1-p)/m(\theta)]$.
  As $z \to 0$, $\varphi(z;\theta) \sim 1+\theta^{-1}\log[1-(1-e^{-\theta})(1-z)] \sim 1+\theta^{-1}\log[e^{-\theta}(1-z+e^{\theta}z)] \sim \theta^{-1}(e^{\theta}-1)z$. As $p \to 0$, $\varphi^{\leftarrow}(p;\theta) \sim p\theta/(e^{\theta}-1)$.
- generalized gamma: $\varphi(z;\theta) = F_{\Gamma}(z^{1/\theta};\theta)$ for $\theta > 0$. As $z \to \infty$, $\varphi(z;\theta) \sim 1-ze^{-z^{1/\theta}}$. As $p \to 1$, $\varphi^{\leftarrow}(p;\theta) \sim [-\log(1-p)]^{\theta}$.
  As $z \to 0$, $\varphi(z;\theta) \sim z/\Gamma(\theta+1)$. As $p \to 0$, $\varphi^{\leftarrow}(p;\theta) \sim p\Gamma(\theta+1)$.

Note that ψ increasing concave implies that $\psi^{\leftarrow}$ is increasing convex. Also, $(-\log(1-p))^{\zeta}$ is convex in large p for any $\zeta > 0$.

Next, tail properties in all four corners of the copula (8.1) are obtained assuming the tails of φ are like one of the above cases.

Lower left corner

The corner probability is $C_\varphi(u,u) = u\varphi(u^{-1}\varphi^{\leftarrow}(u))$. As $u \to 0^+$, this depends mainly on the behavior of $\varphi^{\leftarrow}$ near 0.
(a) If $\varphi^{\leftarrow}(p) \sim kp$ as $p \to 0^+$ with $k > 0$, then $C_\varphi(u,u) \sim u\varphi(k)$ as $u \to 0^+$ and there is lower tail dependence. This applies for the Pareto, logseries-mixture and generalized gamma cdfs.
(b) If $\varphi^{\leftarrow}(p) \sim kp^\zeta$ as $p \to 0^+$ with $k > 0$ and $\zeta > 1$, then $C_\varphi(u,u) \sim u\varphi(ku^{\zeta-1})$ as $u \to 0^+$ and there is no lower tail dependence with $C_\varphi(u,u)/u \to \psi(0) = 0$. For the Weibull cdf for φ, one gets $C_\varphi(u,u) \sim u \cdot (u^{1/\alpha-1})^\alpha = u^{2-\alpha}$ implying intermediate lower tail dependence with tail order $\kappa_L = 2-\alpha \in (0,1)$ for $0 < \alpha < 1$.

Upper right corner

If $(U,V) \sim C$, the upper corner probability is $\widehat{C}_\varphi(u,u) = 2u - 1 + (1-u)\varphi((1-u)^{-1}\varphi^{\leftarrow}(1-u))$ as $u \to 0^+$.
(a) If φ has support on the positive real line, then $\varphi^{\leftarrow}(1-u)$ dominates $1-u$ as $u \to 0$, and $\varphi[(1-u)^{-1}\varphi^{\leftarrow}(1-u)] \sim \varphi[\varphi^{\leftarrow}(1-u)] = 1-u$. Hence $\widehat{C}_\varphi(u,u) \sim 2u - 1 + (1-u)^2 = u^2$, as $u \to 0^+$, with upper tail quadrant independence.
(b) For a cdf φ with finite upper support point, such as generalized Pareto with negative parameter or equivalently, Beta$(1,\gamma)$ with $\gamma > 1$. $\varphi[(1-u)^{-1}\varphi^{\leftarrow}(1-u)] \sim \varphi[(1-u)^{-1}(1-u^{1/\gamma})] \sim \varphi(1-u^{1/\gamma} + o(u)) \sim 1-u$, and $\widehat{C}_\varphi(u,u) \sim u^2$.
(c) For a cdf φ with finite upper support point, such that $\varphi'[\varphi^{\leftarrow}(1)] > 0$, then $\widehat{C}_\varphi(u,u) = O(u)$ using l'Hopital's rule and there is tail dependence (Proposition 4.1 of [Durante et al., 2011]).

Lower right corner

If $(U,V) \sim C$, for 1-reflection, the copula of $(1-U,V)$ is $v - C(1-u,v)$. For (8.1), the corner probability is $u - C_\varphi(1-u,u) = u - (1-u)\varphi[(1-u)^{-1}\varphi^{\leftarrow}(u)]$ as $u \to 0^+$. It depends mainly on $\varphi^{\leftarrow}(p)$ for p near 0. If $\varphi^{\leftarrow}(p) \sim kp^\zeta$ as $p \to 0^+$ with $k > 0$ and $\zeta \geq 1$, then $u - C_\varphi(1-u,u) \sim u - (1-u)\varphi(ku^\zeta) \sim u - (1-u)\varphi(\varphi^{\leftarrow}(u)) = u^2$ and this tail behaves like quadrant independence.

Upper left corner

If $(U,V) \sim C$, for 2-reflection, the copula of $(U,1-V)$ is $u - C(u,1-v)$. For (8.1), the corner probability is $u - C_\varphi(u,1-u) = u - u\varphi[(u^{-1}\varphi^{\leftarrow}(1-u)]$ as $u \to 0^+$. It depends mainly on $\varphi^{\leftarrow}(p)$ for p near 1.
(a) If $\varphi^{\leftarrow}(1-q) \sim kq^{-\zeta} \sim q$ as $q \to 0^+$ with $k > 0$ and $\zeta > 0$, then $u - C_\varphi(u,1-u) \sim u - u\varphi(u^{-1}ku^{-\zeta}) \sim u - u(1-u^{1+1/\zeta}) \sim u^{2+1/\zeta}$.
(b) If $\varphi^{\leftarrow}(1-q) \sim k(-\log q)^\zeta \sim q$ as $q \to 0^+$ with $k > 0$ and $\zeta > 0$, then $u - C_\varphi(u,1-u) \sim u - u\varphi(u^{-1}k(-\log u)^\zeta) \sim u - u[1 - \exp\{-u^{-1/\zeta}(-\log u)\} \sim u\exp\{-u^{-1/\zeta}(-\log u)\}$. This goes to zero faster than rate u^2 and indicates dependence weaker than quadrant independence.

Conclusions about asymmetries

For positive dependence, as $u \to 0^+$, generally for

- the lower left corner, $C_\varphi(u,u)$ can be $O(u)$ to $O(u^2)$,
- for the upper right corner, $2u-1+C_\varphi(1-u,1-u) = O(u^2)$,
- for the lower right corner, $u-C_\varphi(1-u,u) = O(u^2)$,
- for the upper left corner, $u-C_\varphi(u,1-u) = o(u^{2+\zeta})$ with $\zeta \geq 0$; also this corner can have zero probability (as $u \to 0^+$) if φ has a finite upper end point of support.

Hence the general pattern is skewness to joint lower tail relative to joint upper tail (reflection asymetry skewed to joint lower), and permutation asymmetry skewed to lower right corner.

8.6 Closing remarks

It is shown how to analyze some tail properties of new parametric families of copulas. Several new parametric families of bivariate copulas interpolating $C^\perp$ to C^+ or C^- to C^+ have been constructed, but it remains to be seen if they are useful for applications within vines. The DJ class (8.1) is derived based on univariate truncation. Copulas based on truncation tend to have one joint (upper or lower) orthant that is tail orthant independent, so they can be quite tail asymmetric if the opposite orthant has tail dependence. In data applications, it is rare to see such deviations in tail asymmetry relative to Gaussian copulas. However, perhaps the DJ class can be extended to a larger class that covers more flexibility in tail asymmetries.

Acknowledgements Research has been supported from NSERC Discovery Grant 8698 and a Collaborative Research Team grant for the project: *Copula Dependence Modelling: Theory and Applications* of the Canadian Statistical Sciences Institute.

References

[Azzalini, 2013] Azzalini, A. (2013). *The Skew-normal and Related Families*. Cambridge University Press, Cambridge.

[Azzalini and Capitanio, 2003] Azzalini, A. and Capitanio, A. (2003). Distributions generated by perturbation of symmetry with emphasis on a multivariate skew t-distribution. *Journal of the Royal Statistical Society B*, 65(2):367–389.

[Azzalini and Dalla Valle, 1996] Azzalini, A. and Dalla Valle, A. (1996). The multivariate skew-normal distribution. *Biometrika*, 83(4):715–726.

[Barlow and Proschan, 1981] Barlow, R. E. and Proschan, F. (1981). *Statistical Theory of Reliability and Life Testing*. To Begin With, Silver Spring, MD.

[Brechmann et al., 2012] Brechmann, E. C., Czado, C., and Aas, K. (2012). Truncated regular vines in high dimensions with application to financial data. *Canadian Journal of Statistics*, 40(1):68–85.

[Coia, 2017] Coia, V. (2017). *Forecasting of Nonlinear Extreme Quantiles using Copula Models*. PhD thesis, University of British Columbia.

[Di Lascio et al., 2016] Di Lascio, F. M., Durante, F., and Jaworski, P. (2016). Truncation invariant copulas and a testing procedure. *Journal of Statistical Computation and Simulation*, 86(12):2362–2378.

[Dissmann et al., 2013] Dissmann, J., Brechmann, E. C., Czado, C., and Kurowicka, D. (2013). Selecting and estimating regular vine copulae and application to financial returns. *Computational Statistics and Data Analysis*, 59:52–69.

[Durante and Jaworski, 2012] Durante, F. and Jaworski, P. (2012). Invariant dependence structure under univariate truncation. *Statistics*, 46(2):263–277.

[Durante et al., 2011] Durante, F., Jaworski, P., and Mesiar, R. (2011). Invariant dependence structures and Archimedean copulas. *Statistics & Probability Letters*, 81(12):1995–2003.

[Fredricks and Nelsen, 2007] Fredricks, G. A. and Nelsen, R. B. (2007). On the relationship between Spearman's rho and Kendall's tau for pairs of continuous random variables. *Journal of Statistical Planning and Inference*, 137(7):2143–2150.

[Hua and Joe, 2011] Hua, L. and Joe, H. (2011). Tail order and intermediate tail dependence of multivariate copulas. *Journal of Multivariate Analysis*, 102(10):1454–1471.

[Jaworski, 2013] Jaworski, P. (2013). The limiting properties of copulas under univariate conditioning. In Jaworski, P., Durante, F., and Härdle, W. K., editors, *Copulae in Mathematical and Quantitative Finance*, pages 129–163. Springer, Heidelberg.

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

[Joe, 2014] Joe, H. (2014). *Dependence Modeling with Copulas*. Chapman & Hall/CRC, Boca Raton, FL.

[Krupskii, 2017] Krupskii, P. (2017). Copula-based measures of reflection and permutation asymmetry and statistical tests. *Statistical Papers*, in press.

[Krupskii and Joe, 2013] Krupskii, P. and Joe, H. (2013). Factor copula models for multivariate data. *Journal of Multivariate Analysis*, 120:85–101.

[Krupskii and Joe, 2015] Krupskii, P. and Joe, H. (2015). Tail-weighted measures of dependence. *Journal of Applied Statistics*, 42:614–629.

[McNeil and Nešlehová, 2009] McNeil, A. J. and Nešlehová, J. (2009). Multivariate Archimedean copulas, d-monotone functions and l_1-norm symmetric distributions. *Annals of Statistics*, 37(5B):3059–3097.

[Nagler and Czado, 2016] Nagler, T. and Czado, C. (2016). Evading the curse of dimensionality in nonparametric density estimation with simplified vine copulas. *Journal of Multivariate Analysis*, 151:69–89.

[Nelsen, 2006] Nelsen, R. B. (2006). *An Introduction to Copulas*. Springer, New York, second edition.

[Panagiotelis et al., 2017] Panagiotelis, A., Czado, C., Joe, H., and Jacob, S. (2017). Model selection for discrete regular vine copulas. *Computational Statistics and Data Analysis*, 106:138–152.

[Yoshiba, 2015] Yoshiba, T. (2015). Maximum likelihood estimation of skew t-copulas, with its applications to stock returns. Technical report, ISM Research Memorandum No. 1195.

Chapter 9
Copula constructions using ultramodularity

Erich Peter Klement, Anna Kolesárová, Radko Mesiar, and Susanne Saminger-Platz

To Professor Roger B. Nelsen
on the occasion of his 75th birthday

Abstract We discuss some copula constructions by means of ultramodular bivariate copulas. In general, the ultramodularity of a real function is a stronger version of both its convexity and its supermodularity (the latter property being always satisfied in the case of a bivariate copula). In a statistical sense, ultramodular bivariate copulas are related to random vectors whose components are mutually stochastically decreasing with respect to each other. Analytically speaking, an ultramodular bivariate copula is characterized by the convexity of all of its horizontal and vertical sections. Among other results, we give a sufficient condition for the additive generators of Archimedean ultramodular bivariate copulas, and we propose two constructions for bivariate copulas: the first one being based on ultramodular aggregation functions, and the other one showing the special role of ultramodularity and Schur concavity for a product-like composition of bivariate copulas being again a bivariate copula.

Erich Peter Klement
Department of Knowledge-Based Mathematical Systems, Johannes Kepler University, Linz, Austria; e-mail: ep.klement@jku.at

Anna Kolesárová
Department of Information Engineering, Automation and Mathematics, Faculty of Chemical and Food Technology, Slovak University of Technology, Bratislava, Slovakia; e-mail: anna.kolesarova@stuba.sk

Radko Mesiar
Department of Mathematics and Descriptive Geometry, Faculty of Civil Engineering, Slovak University of Technology, Bratislava, Slovakia; e-mail: radko.mesiar@stuba.sk

Susanne Saminger-Platz
Department of Knowledge-Based Mathematical Systems, Johannes Kepler University, Linz, Austria; e-mail: susanne.saminger-platz@jku.at

M. Úbeda Flores et al. (eds.), *Copulas and Dependence Models with Applications*,
DOI 10.1007/978-3-319-64221-5_9

9.1 Introduction

We work in the framework of aggregation functions [22] with special properties, in particular with boundary conditions (extending the Boolean conjunction) and supermodularity (i.e., with bivariate copulas [17, 26, 53, 69]). When constructing bivariate copulas, it is therefore necessary to guarantee these two properties which can be treated rather independently.

This survey is organized as follows. In the following section, the necessary preliminaries from the theory of aggregation functions, quasi-copulas and copulas are given, including ultramodularity and Schur concavity. Ultramodular bivariate copulas are discussed in Section 9.3.

In Section 9.4 we exploit the ultramodularity of copulas using the important fact that in a composition of functions the supermodularity of the inner function is preserved just when the outer function is ultramodular [44]. We discuss some variants of this approach and present a number of examples, some of them leading to bivariate copulas which are well-known from the literature.

Section 9.5 is devoted to the so-called D-product of a bivariate copula and its dual, where the bivariate copula D is ultramodular and Schur concave on the upper left triangle of the unit square [31, Theorem 2.3]. These properties of D are shown to be sufficient for the D-product of an arbitrary copula and its dual to be a bivariate copula, and several examples and counterexamples in this context are given.

9.2 Preliminaries

For elements $(x_1, x_2, \dots, x_n)$ of the n-dimensional Euclidean space $\mathbb{R}^n$ we shall also write $\mathbf{x}$, whatever is more convenient in the given context. In one case (when discussing the proof of Theorem 9.5.2 in Section 9.5) we will interpret elements of $\mathbb{R}^2$ as points in the Euclidian plane and denote them by capital letters. The order on $\mathbb{R}^n$ is induced by the usual linear order in its coordinates.

To simplify some of the formulas to come, we shall also use the infix notations $\wedge$ and $\vee$ for the lattice operations meet and join in $\mathbb{R}^n$.

If $A\colon [0,1]^2 \to [0,1]$ is a binary function, we shall sometimes work with its *diagonal section* $\delta_A\colon [0,1] \to [0,1]$, its *horizontal section* $h_{A,\alpha}\colon [0,1] \to [0,1]$ and with its *vertical section* $v_{A,\alpha}\colon [0,1] \to [0,1]$ at level $\alpha \in [0,1]$ given by, respectively,

$$\delta_A(x) = A(x,x), \qquad h_{A,\alpha}(x) = A(x,\alpha), \qquad v_{A,\alpha}(x) = A(\alpha,x).$$

We often require a function $f\colon \Omega \to \mathbb{R}$, where Ω is a non-empty subset of $\mathbb{R}^n$, to be 1-*Lipschitz*, i.e., for all $\mathbf{x}, \mathbf{y} \in \Omega$ we have $|f(\mathbf{x}) - f(\mathbf{y})| \leq \sum_{i=1}^n |x_i - y_i|$.

Three particular functions will play a special role in our considerations: the *Fréchet-Hoeffding lower bound* $W\colon [0,1]^n \to [0,1]$, the *product* $\Pi\colon [0,1]^n \to [0,1]$, and the *Fréchet-Hoeffding upper bound* $M\colon [0,1]^n \to [0,1]$ given by, respectively, $W(\mathbf{x}) = \max\left(\sum_{i=1}^n x_i - (n-1), 0\right)$, $\Pi(\mathbf{x}) = \prod_{i=1}^n x_i$, and $M(\mathbf{x}) = \min(x_1, x_2, \dots, x_n)$.

9.2.1 Aggregation functions, quasi-copulas, copulas

An *(n-ary) aggregation function* [22] is a function $A\colon [0,1]^n \to [0,1]$ which is monotone non-decreasing in each component and which satisfies $A(0,0,\dots,0) = 0$ and $A(1,1,\dots,1) = 1$.

Given a binary 1-Lipschitz aggregation function $A\colon [0,1]^2 \to [0,1]$, its *dual* [22] $A^*\colon [0,1]^2 \to [0,1]$ is defined by $A^*(x,y) = x+y-A(x,y)$.

Each n-ary 1-Lipschitz aggregation function A satisfies $W \le A \le W^*$, where the dual function $W^*\colon [0,1]^n \to [0,1]$ of the Fréchet-Hoeffding lower bound W is given by $W^*(\mathbf{x}) = \min\bigl(\sum_{i=1}^n x_i, 1\bigr)$.

A (*bivariate*) *quasi-copula* (see [2, 20]) $Q\colon [0,1]^2 \to [0,1]$ is a 1-Lipschitz aggregation function which satisfies $Q(0,x) = Q(x,0) = 0$ and $Q(1,x) = Q(x,1) = x$ for all $x \in [0,1]$.

Observe that a 1-Lipschitz aggregation function $A\colon [0,1]^2 \to [0,1]$ is a quasi-copula if and only if $A(0,1) = A(1,0) = 0$ (see [37]) or, equivalently, if and only if it is bounded from above by the Fréchet-Hoeffding upper bound M, i.e., $A \le M$.

A (*bivariate*) *copula* $C\colon [0,1]^2 \to [0,1]$ (see [69, 53]) is a function which satisfies $C(0,x) = C(x,0) = 0$ and $C(1,x) = C(x,1) = x$ for all $x \in [0,1]$ and which is 2-*increasing*, i.e., for all $x_1,x_2,y_1,y_2 \in [0,1]$ with $x_1 \le x_2$ and $y_1 \le y_2$ for the C-volume V_C of the rectangle $[x_1,x_2] \times [y_1,y_2]$ we have

$$V_C([x_1,x_2] \times [y_1,y_2]) = C(x_1,y_1) - C(x_1,y_2) + C(x_2,y_2) - C(x_2,y_1) \ge 0. \tag{9.1}$$

Obviously, each bivariate copula is a quasi-copula but not vice versa. Each quasi-copula Q satisfies $W \le Q \le M$, and so the same inequalities hold for each copula.

A bivariate copula $C\colon [0,1]^2 \to [0,1]$ is called *Archimedean* if there is a continuous, strictly decreasing convex function $t\colon [0,1] \to [0,\infty]$ with $t(1) = 0$ such that (see [51]) for all $(x,y) \in [0,1]^2$

$$C(x,y) = t^{-1}(\min(t(x)+t(y),t(0))). \tag{9.2}$$

The function t is called an *additive generator* of C, and it is unique up to a positive multiplicative constant. Note that a bivariate Archimedean copula is necessarily associative and satisfies $C(x,x) < x$ for all $x \in \,]0,1[$. For many more facts about quasi-copulas and copulas see [17, 26, 53].

If $(C_i)_{i\in I}$ is a family of copulas and $(]a_i,e_i[)_{i\in I}$ a family of non-empty, pairwise disjoint open subintervals of $[0,1]$, the *ordinal sum* of the *summands* $(\langle a_i,e_i,C_i\rangle)_{i\in I}$ (which we shall call here *M-ordinal sum*) is well-known from the literature [18, 35, 42, 62]), and it is based on a result in the theory of abstract semigroups [9]. In this construction, the “gaps” between the squares $[a_i,e_i]^2$ are filled by the upper Fréchet-Hoeffding bound M, and the result is always a copula. Another ordinal sum construction of copulas is based on the lower Fréchet-Hoeffding bound W, and it was considered more recently in [10, 14, 32, 50].

Let $(C_i)_{i\in I}$ be a family of copulas and $(]a_i,e_i[)_{i\in I}$ be a family of non-empty, pairwise disjoint open subintervals of $[0,1]$.

(i) The copula $C = M\text{-}(\langle a_i, e_i, C_i \rangle)_{i \in I}$ defined by

$$C(x,y) = \begin{cases} a_i + (e_i - a_i) C_i\left(\frac{x-a_i}{e_i-a_i}, \frac{y-a_i}{e_i-a_i}\right) & \text{if } (x,y) \in [a_i, e_i]^2, \\ M(x,y) & \text{otherwise,} \end{cases} \tag{9.3}$$

is called the *M-ordinal sum* of the *summands* $(\langle a_i, e_i, C_i \rangle)_{i \in I}$.

(ii) The copula $C = W\text{-}(\langle a_i, e_i, C_i \rangle)_{i \in I}$ defined by

$$C(x,y) = \begin{cases} a_i + (e_i - a_i) C_i\left(\frac{x-a_i}{e_i-a_i}, \frac{y-1+e_i}{e_i-a_i}\right) & \text{if } (x,y) \in [a_i, e_i[\times [1-e_i, 1-a_i[, \\ W(x,y) & \text{otherwise,} \end{cases} \tag{9.4}$$

is called the *W-ordinal sum* of the *summands* $(\langle a_i, e_i, C_i \rangle)_{i \in I}$.

More recently, also other constructions in this spirit were proposed, using patchwork [10, 13, 14] and gluing techniques [68], or ordinal sums based on the product Π [38, 48, 60].

9.2.2 Modular, supermodular, and ultramodular functions

Let Ω be a convex sublattice of $\mathbb{R}^n$. A function $f\colon \Omega \to [0,1]$ is called *convex* if, for all $\mathbf{x}, \mathbf{y} \in \Omega$ and for all $\lambda \in [0,1]$, we have $f\big(\lambda \mathbf{x} + (1-\lambda)\mathbf{y}\big) \le \lambda f(\mathbf{x}) + (1-\lambda) f(\mathbf{y})$.

A function $f\colon \Omega \to [0,1]$ is called

[M] *modular* if, for all $\mathbf{x}, \mathbf{y} \in \Omega$, we have $f\big(\mathbf{x} \vee \mathbf{y}\big) + f\big(\mathbf{x} \wedge \mathbf{y}\big) = f(\mathbf{x}) + f(\mathbf{y})$;

[SM] *supermodular* [6, 33, 44] if, for all $\mathbf{x}, \mathbf{y} \in \Omega$,

$$f\big(\mathbf{x} \vee \mathbf{y}\big) + f\big(\mathbf{x} \wedge \mathbf{y}\big) \ge f(\mathbf{x}) + f(\mathbf{y});$$

[UM] *ultramodular* [44] if, for all $\mathbf{x}, \mathbf{y} \in \Omega$ with $\mathbf{x} \le \mathbf{y}$ and for all $\mathbf{h} \in \mathbb{R}^n$ with $\mathbf{h} \ge \mathbf{0}$ and $\mathbf{x} + \mathbf{h}, \mathbf{y} + \mathbf{h} \in \Omega$,

$$f(\mathbf{x} + \mathbf{h}) - f(\mathbf{x}) \le f(\mathbf{y} + \mathbf{h}) - f(\mathbf{y}).$$

Each modular function is ultramodular, and ultramodularity implies supermodularity [44]. As a consequence of [25, 28, 30, 72], the ultramodular functions form a proper subset of the set of all convex functions. Ultramodular functions were also called *Wright convex* [57] by some authors (mainly in mathematical analysis where they first appeared in [72] and originally just were called convex functions).

In the case of an n-dimensional real domain, ultramodularity can be seen as a version of convexity. Assuming some mild regularity, the set of ultramodular functions equals the intersection of the set of all supermodular functions and the set of all functions which are convex in each variable. Ultramodular functions are used in economics, in particular in game theory when dealing with convex measure games [3], but they also have applications in multicriteria decision support systems [5].

In statistics, ultramodular functions play an important role in modelling stochastic orders and positive dependence among random vectors (see [52, 64]), and they are known there also as *directionally convex* functions. For more details about ultramodular real functions we recommend [44].

Trivially, each additive function $f\colon [0,1] \to \mathbb{R}$, i.e., for all $x,y \in [0,1]$ we have $f(x+y) = f(x)+f(y)$, is an ultramodular function. However, it is linear (and, therefore, convex) only if we require additional properties, such as continuity, monotonicity or boundedness. Otherwise, an additive function $f\colon [0,1] \to \mathbb{R}$ can be quite pathological (for instance, its graph may be a dense subset of $[0,1] \times \mathbb{R}$).

For a function of one variable, ultramodularity and convexity are equivalent. If $n > 1$, ultramodularity and convexity of a function $f\colon \mathbb{R}^n \to \mathbb{R}$ are independent notions. There are, on the one hand, convex functions which are not ultramodular (e.g., $f(\mathbf{x}) = \|\mathbf{x}\|$) and, on the other hand, ultramodular functions which are not convex (e.g., $f(\mathbf{x}) = \prod_{i=1}^{n} x_i$).

In special cases, there are several additional characterizations of and some interesting relations between these properties.

A function of one variable defined on an interval which is twice differentiable is a convex function if and only if its second derivative is non-negative. A continuous, twice differentiable function of several variables defined on a convex set is a convex function if and only if its Hessian matrix is positive semidefinite on the interior of the convex set.

A monotone non-decreasing function $f\colon [0,1]^2 \to [0,1]$ is supermodular if and only if for all $x, x_*, y, y_* \in [0,1]$ with $x \le x_*$ and $y \le y_*$

$$f(x_*,y_*) - f(x_*,y) - f(x,y_*) + f(x,y) \ge 0. \tag{9.5}$$

Binary aggregation functions satisfying condition (9.5) are also called 2-*increasing* [53] or of *moderate growth* [29]. Mainly in economics, ultramodular functions $f\colon \Omega \to [0,1]$ are said to have *non-decreasing increments* [7].

Observe that a function $C\colon [0,1]^2 \to [0,1]$ is a bivariate copula if and only if it is a supermodular quasi-copula.

In [44] several equivalent conditions for a function $f\colon \Omega \to [0,1]$ to be ultramodular are given:

[UM1] for all $\mathbf{x},\mathbf{y},\mathbf{z} \in \Omega$ with $\mathbf{y},\mathbf{z} \ge \mathbf{0}$ and $\mathbf{x}+\mathbf{y}+\mathbf{z} \in \Omega$

$$f(\mathbf{x}+\mathbf{z}) - f(\mathbf{x}) \le f(\mathbf{x}+\mathbf{y}+\mathbf{z}) - f(\mathbf{x}+\mathbf{y});$$

[UM2] for all $\mathbf{x},\mathbf{y},\mathbf{z},\mathbf{w} \in \Omega$ with $\mathbf{x} \le \mathbf{y} \le \mathbf{w}$ and $\mathbf{x}+\mathbf{w} = \mathbf{y}+\mathbf{z}$

$$f(\mathbf{y}) + f(\mathbf{z}) \le f(\mathbf{w}) + f(\mathbf{x});$$

[UM3] for all $\mathbf{x} \in \Omega$ and all $\mathbf{h},\mathbf{k} \in \mathbb{R}^n$ with $\mathbf{h} \ge \mathbf{0}$ and $(|k_1|,|k_2|,\ldots,|k_n|) \le \mathbf{h}$ and $\mathbf{x}+\mathbf{h}, \mathbf{x}-\mathbf{h}, \mathbf{x}+\mathbf{k}, \mathbf{x}-\mathbf{k} \in \Omega$

$$f(\mathbf{x}+\mathbf{k}) + f(\mathbf{x}-\mathbf{k}) \le f(\mathbf{x}+\mathbf{h}) + f(\mathbf{x}-\mathbf{h}).$$

The vectors $\mathbf{x}, \mathbf{y}, \mathbf{z}, \mathbf{w} \in \Omega$ in [UM2] can be viewed as the vertices of a parallelogram centered at $\frac{1}{2}(\mathbf{x}+\mathbf{w})$ (see Figure 9.1).

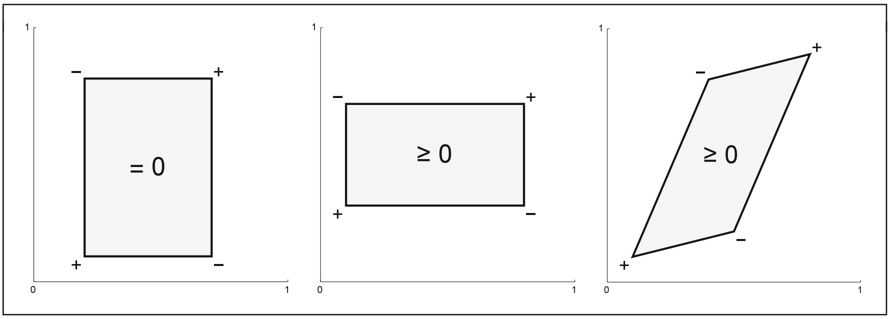


Fig. 9.1: Modularity (left), supermodularity (center), and ultramodularity (right) condition for a binary function defined on the unit square.

The following characterization of supermodular functions $f\colon [0,1]^n \to [0,1]$ is due to [6, 27]:

Proposition 9.2.1. *An n-ary function* $f\colon [0,1]^n \to [0,1]$ *is supermodular if and only if each of its two-dimensional sections is supermodular, i.e., for each* $\mathbf{x} \in [0,1]^n$ *and all* $i, j \in \{1,2,\ldots,n\}$ *with* $i \neq j$*, the function* $f_{\mathbf{x},i,j}\colon [0,1]^2 \to [0,1]$ *given by* $f_{\mathbf{x},i,j}(u,v) = f(\mathbf{y})$*, where* $y_i = u$*,* $y_j = v$ *and* $y_k = x_k$ *for* $k \in \{1,2,\ldots,n\} \setminus \{i,j\}$*, is supermodular.*

The following result (Corollary 4.1 of [44]) states the exact relationship between ultramodular and supermodular functions $f\colon [0,1]^n \to [0,1]$:

Proposition 9.2.2. *A function* $f\colon [0,1]^n \to [0,1]$ *is ultramodular if and only if f is supermodular and if each of its one-dimensional sections is convex, i.e., if for each* $\mathbf{x} \in [0,1]^n$ *and each* $i \in \{1,\ldots,n\}$ *the function* $f_{\mathbf{x},i}\colon [0,1] \to [0,1]$ *which is given by* $f_{\mathbf{x},i}(u) = f(\mathbf{y})$*, where* $y_i = u$ *and* $y_j = x_j$ *whenever* $j \neq i$*, is convex.*

For an n-ary aggregation function $A\colon [0,1]^n \to [0,1]$ the following are equivalent as a consequence of Propositions 9.2.1 and 9.2.2:

(i) A is ultramodular;
(ii) each two-dimensional section of A is ultramodular;
(iii) each two-dimensional section of A is supermodular and each one-dimensional section of A is convex.

For $n = 2$, the ultramodularity [UM] of an aggregation function $A\colon [0,1]^2 \to [0,1]$ is equivalent to A being *P-increasing* (see [16]), i.e., to

$$A(u_1,v_1) + A(u_4,v_4) \geq \max\big(A(u_2,v_2) + A(u_3,v_3), A(u_3,v_2) + A(u_2,v_3)\big)$$

for all $u_1,u_2,u_3,u_4,v_1,v_2,v_3,v_4 \in [0,1]$ which satisfy $u_1 \leq u_2 \wedge u_3 \leq u_2 \vee u_3 \leq u_4$, $v_1 \leq v_2 \wedge v_3 \leq v_2 \vee v_3 \leq v_4$, $u_1 + u_4 \geq u_2 + u_3$, and $v_1 + v_4 \geq v_2 + v_3$.

9.2.3 Ultramodular aggregation functions

In general, the class of ultramodular aggregation functions is closed under composition (see Theorem 3.2 in [33]), i.e., ultramodular aggregation functions form a clone [39]. Therefore, any n-ary extension of a bivariate ultramodular aggregation function is again ultramodular:

Proposition 9.2.3. *Let $A\colon [0,1]^2 \to [0,1]$ be an ultramodular bivariate aggregation function and define for $n \geq 2$ its n-ary extension $A^{[n]}\colon [0,1]^n \to [0,1]$ inductively by*

$$A^{[2]} = A \quad \textit{and} \quad A^{[n+1]}(x_1,x_2,\ldots,x_{n+1}) = A\big(A^{[n]}(x_1,x_2,\ldots,x_n),x_{n+1}\big). \tag{9.6}$$

Then $A^{[n]}$ is an n-ary ultramodular aggregation function for each $n \geq 2$.

Proof. The function $A^{[3]}\colon [0,1]^3 \to [0,1]$ given by $A^{[3]}(x_1,x_2,x_3) = A(A(x_1,x_2),x_3)$ is ultramodular because of the ultramodularity of A and the ultramodularity of the two functions $B_1\colon [0,1]^3 \to [0,1]$, $B_1(x_1,x_2,x_3) = A(x_1,x_2)$ and $B_2\colon [0,1]^3 \to [0,1]$, $B_2(x_1,x_2,x_3) = x_3$. The rest follows by induction. □

If $A\colon [0,1]^2 \to [0,1]$ is an ultramodular bivariate aggregation function which is not associative then we can define another trivariate extension $B^{[3]}\colon [0,1]^3 \to [0,1]$ by $B^{[3]}(x_1,x_2,x_3) = A(x_1,A(x_2,x_3))$ which may be different from $A^{[3]}$ given in (9.6), but which is also ultramodular.

The following result is a consequence of Theorems 3.1 and 3.2 in [33], and it generalizes [12, Theorem 5.2]:

Theorem 9.2.4. *Let $A\colon [0,1]^n \to [0,1]$ be an ultramodular n-ary aggregation function and let $B_1,B_2,\ldots,B_n\colon [0,1]^k \to [0,1]$ be supermodular, monotone non-decreasing functions of k variables such that $A\big(B_1(\mathbf{0}),B_2(\mathbf{0}),\ldots,B_n(\mathbf{0})\big) = 0$ and $A\big(B_1(\mathbf{1}),B_2(\mathbf{1}),\ldots,B_n(\mathbf{1})\big) = 1$. Then the composite function $D\colon [0,1]^k \to [0,1]$ defined by $D(\mathbf{x}) = A\big(B_1(\mathbf{x}),B_2(\mathbf{x}),\ldots,B_n(\mathbf{x})\big)$ is a supermodular k-ary aggregation function.*

9.2.4 Schur concave functions and copulas

Another property of real functions which will play a crucial role in the constructions in Section 9.5 is the Schur concavity. The concept of Schur convex functions (and Schur concave functions as their negations) was presented in [61] as a variant of convexity and concavity of real functions, respectively (see also [58]). For example, each symmetric convex function is Schur convex (and each symmetric concave function is Schur concave). An example of a Schur convex function is the maximum. The minimum and the product (the latter only in the case of strictly positive factors) are Schur concave, as well as all elementary symmetric functions (again only if all components are strictly positive) [66, 67].

Schur convexity is a special type of monotonicity which reverses majorization. Majorization [46], a preorder on vectors of real numbers, and inequalities related to it [65] are used when comparing income inequality, and they have applications also in physics, chemistry, political science, engineering, and economics [47]. Looking at bivariate copulas, one sees that each Schur concave copula is necessarily symmetric, and that each associative copula is Schur concave [15]. Also other Schur convex (concave) functions have interpretations in stochastics and/or aggregation [22]: the variance and the standard deviation turn out to be Schur convex, whereas the Shannon entropy, the Rényi entropy and the Gini coefficient are Schur concave [4, 21, 40, 55, 56].

If $\mathbf{x} = (x_1, x_2, \dots, x_n) \in \mathbb{R}^n$ is a vector then we denote by $\mathbf{x}^{\downarrow} = (x_1^{\downarrow}, x_2^{\downarrow}, \dots, x_n^{\downarrow}) \in \mathbb{R}^n$ the vector having the same components as $\mathbf{x}$, but sorted in descending order.

We say that $\mathbf{y} = (y_1, y_2, \dots, y_n) \in \mathbb{R}^n$ *majorizes* $\mathbf{x} = (x_1, x_2, \dots, x_n) \in \mathbb{R}^n$ (in symbols $\mathbf{y} \succ \mathbf{x}$) if we have $\sum_{i=1}^n x_i = \sum_{i=1}^n y_i$ and $\sum_{i=k}^n x_i^{\downarrow} \leq \sum_{i=k}^n y_i^{\downarrow}$ for all $k \in \{1, 2, \dots, n\}$.

The relation $\succ$ on $\mathbb{R}^n$ is reflexive and transitive, i.e., a preorder, but not antisymmetric and, therefore, not a partial order: from $\mathbf{y} \succ \mathbf{x}$ and $\mathbf{x} \succ \mathbf{y}$ it only follows that $\mathbf{x}$ and $\mathbf{y}$ have the same components, but not necessarily in the same order.

If Ω is a subset of $\mathbb{R}^n$ then a function $f \colon \Omega \to \mathbb{R}$ is called *Schur convex* if for all $\mathbf{x}, \mathbf{y} \in \Omega$ with $\mathbf{y} \succ \mathbf{x}$ we have $f(\mathbf{y}) \geq f(\mathbf{x})$, and it is called *Schur concave* if its negation $-f$ is Schur convex.

When considering the special scenario $n = 2$ and $\Omega = [0,1]^2$, the conditions for Schur concavity become somewhat simpler:

In this case, a function $f \colon [0,1]^2 \to [0,1]$ is Schur concave if and only if, for all $(x,y), (u,v) \in [0,1]^2$ satisying $x + y = u + v$ and $\min(x,y) \leq \min(u,v)$, we have $f(x,y) \leq f(u,v)$.

Equivalently, a function $f \colon [0,1]^2 \to [0,1]$ is Schur concave if and only if we have $f(x,y) \leq f(\lambda \cdot x + (1-\lambda) \cdot y, (1-\lambda) \cdot x + \lambda \cdot y)$ for all $(x,y) \in [0,1]^2$ and all $\lambda \in [0,1]$.

We will often require the Schur concavity of a copula $C \colon [0,1]^2 \to [0,1]$ on the upper left triangle $\Delta = \{(x,y) \in [0,1]^2 \mid x \leq y\}$ of the unit square only, which means that $C(x,y) \leq C(x+\varepsilon, y-\varepsilon)$ for all $(x,y) \in \Delta$ and for all $\varepsilon > 0$ with $(x+\varepsilon, y-\varepsilon) \in \Delta$.

Here are some well-known facts about Schur concave copulas: first of all, each associative copula is Schur concave. As a consequence, the three basic copulas W, Π and M and also all Archimedean copulas are Schur concave. Trivially, each convex combination of Schur concave copulas is also Schur concave (implying that, for example, each member of the two families of Fréchet [19] and Mardia [43] copulas as discussed in [53, Exercise 2.4] is Schur concave).

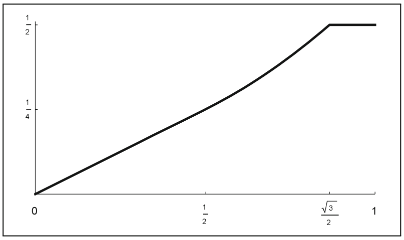
Moreover, each Schur concave copula is symmetric, and each symmetric copula which is Schur concave on the upper left triangle Δ is Schur concave.

In general, however, the Schur concavity of a copula on the upper left triangle Δ does not imply its symmetry. A counterexample is given by the copula C defined by

$$C(x,y) = \begin{cases} \Pi(x,y) & \text{if } x \leq y, \\ \min\left(\frac{x^2+y^2}{2}, y\right) & \text{otherwise.} \end{cases}$$

Since C coincides with Π on the upper left triangle Δ it is Schur concave on Δ. On the other hand, C is not ultramodular: for example, the horizontal section $h_{C,\frac{1}{2}}\colon [0,1] \to \left[0,\frac{1}{2}\right]$ is not a convex function (see figure below).

$$h_{C,\frac{1}{2}}(x) = \begin{cases} \frac{x}{2} & \text{if } x \in \left[0,\frac{1}{2}\right], \\ \frac{4x^2+1}{8} & \text{if } x \in \left[\frac{1}{2},\frac{\sqrt{3}}{2}\right], \\ \frac{1}{2} & \text{otherwise.} \end{cases}$$



9.3 Ultramodular copulas

Ultramodular bivariate copulas, characterized by the convexity of all of their horizontal and vertical sections [44, Corollary 4.1] (see also [34, Proposition 2.3]), were studied recently in [33, 34]. An ultramodular copula describes the dependence structure of stochastically decreasing random vectors (see [53]), and thus each ultramodular copula is said to be *negative quadrant dependent*.

The set $\mathscr{C}_u$ of all ultramodular copulas is closed under convex combinations and pointwise limits, i.e., it is a compact and convex subset of the set of all binary functions from $[0,1]^2$ to $[0,1]$.

The smallest element of $\mathscr{C}_u$ is the Fréchet-Hoeffding lower bound W, and its greatest element is the product copula Π (observe that each section of Π is a linear function i.e., the greatest convex function satisfying the boundary conditions of copulas).

Moreover, each survival copula [17, 53] of an ultramodular copula is ultramodular, and a W-ordinal sum $W\text{-}(\langle a_i, e_i, C_i\rangle)_{i\in I}$ of bivariate copulas as given by (9.4) is ultramodular if and only if each copula C_i is ultramodular [33].

No copula which is greater than Π (for example, the Fréchet-Hoeffding upper bound M) can be ultramodular. Also no copula C with an idempotent element different from 0 and 1, i.e., $C(a,a) = a$ for some $a \in\,]0,1[$, can be ultramodular since this would imply $C(x,a) = \min(x,a)$ for all $x \in [0,1]$, i.e., the existence of some non-convex horizontal section. As a consequence, a non-trivial M-ordinal sum of copulas [9, 18, 35, 42, 53, 62, 63]) can never be ultramodular.

Because of [53, 63], each associative copula is an M-ordinal sum of Archimedean copulas. Therefore, each ultramodular associative copula C is a trivial ordinal sum of Archimedean copulas, i.e., C itself must be Archimedean. The set of all ultramodular Archimedean copulas is a compact (but not convex) subset of $\mathscr{C}_u$ (see [36]).

Some of the one-parameter families of Archimedean copulas listed in [53, Table 4.1] contain ultramodular copulas: the complete family of copulas given in [53, (4.2.7)] is ultramodular (in the framework of triangular norms it is also known as family of Sugeno-Weber t-norms [35, 70, 71]), as well as "half" of the family (with parameters in $[-1,0]$) of Ali-Mikhail-Haq copulas [1] (also known as family of Hamacher t-norms [35, 23, 24]) given in [53, (4.2.3)] and "half" of the family (with parameters in $[-\infty,0]$) of Frank copulas [18] given in [53, (4.2.5)].

In [59, Proposition 3.2] (compare also Proposition 9.2.3) it was shown that, for each ultramodular bivariate copula $C\colon [0,1]^2 \to [0,1]$ and for each $n \geq 2$, the function $C^{[n]}\colon [0,1]^n \to [0,1]$ defined as in (9.6) is an n-ary ultramodular quasi-copula. Even if the bivariate copula C is associative, the quasi-copula $C^{[n]}$ is not a copula in general: the Fréchet-Hoeffding lower bound is a well-known counterexample.

The diagonal section $\delta_C\colon [0,1] \to [0,1]$ of an ultramodular bivariate copula C is strictly increasing on the preimage $\delta_C^{\leftarrow}(]0,1]) = \{x \in [0,1] \mid \delta_C(x) \in]0,1]\}$:

Lemma 9.3.1. *If $C\colon [0,1]^2 \to [0,1]$ is an ultramodular bivariate copula then there are no numbers $x_1, x_2 \in]0,1[$ with $x_1 < x_2$ such that $0 < \delta_C(x_1) = \delta_C(x_2)$.*

Proof. Assume, to the contrary, that such numbers exist and assume without loss of generality $x_1 = \inf\{x \in]0,1[\mid \delta_C(x) = \delta_C(x_2)\}$. Choose some $\varepsilon \in \left]0, \frac{x_2-x_1}{2}\right]$ and consider $\mathbf{x}, \mathbf{y}, \mathbf{z} \in [0,1]^2$ given by $\mathbf{x} = (x_1-\varepsilon, x_1-\varepsilon)$, $\mathbf{y} = (\varepsilon, 2\varepsilon)$, and $\mathbf{z} = (2\varepsilon, \varepsilon)$. Then we obtain, due to the monotonicity of C, $C(\mathbf{x}) = \delta_C(x_1-\varepsilon) < \delta_C(x_2)$ and $C(\mathbf{x}+\mathbf{y}) = C(\mathbf{x}+\mathbf{z}) = C(\mathbf{x}+\mathbf{y}+\mathbf{z}) = \delta_C(x_2)$, implying that

$$C(\mathbf{x}+\mathbf{y}+\mathbf{z}) - C(\mathbf{x}+\mathbf{y}) - C(\mathbf{x}+\mathbf{z}) + C(\mathbf{x}) = \delta_C(x_1-\varepsilon) - \delta_C(x_2) < 0,$$

in contradiction to the ultramodularity of C. □

If we want to see whether an Archimedean copula is ultramodular, i.e., has convex horizontal and vertical sections, its symmetry (as a consequence of (9.2)) and boundary conditions tell us that it suffices to check the convexity of all horizontal sections for $a \in]0,1[$.

The following characterization of the ultramodularity of an Archimedean bivariate copula $C\colon [0,1]^2 \to [0,1]$ with a two times differentiable additive generator $t\colon [0,1] \to [0,\infty]$ was given in [34, Theorem 3.1] (see also [8]): C is ultramodular if and only if $1/t'$ is a convex function.

9.4 A rather general construction method

Based on some earlier results of [11, Theorem 2] and [34, Theorem 4.1], the following was shown in [31, Theorem 2.3]:

Theorem 9.4.1. *Let $C, D_1, D_2\colon [0,1]^2 \to [0,1]$ be bivariate copulas and suppose that C is ultramodular. Then, for all monotone non-decreasing continuous functions $f_1, f_2, g_1, g_2\colon [0,1] \to [0,1]$ such that $C(f_1(x), f_2(x)) = C(g_1(x), g_2(x)) = x$ for each $x \in [0,1]$, also the function $E\colon [0,1]^2 \to [0,1]$ given by*

$$E(x,y) = C\big(D_1(f_1(x), g_1(y)), D_2(f_2(x), g_2(y))\big) \tag{9.7}$$

is a bivariate copula.

As a consequence of Proposition 9.2.3 and Theorem 9.2.4, there is an extension of Theorem 9.4.1, generalizing a result shown for the product copula in [41]:

Corollary 9.4.2. *Let* $Q\colon [0,1]^n \to [0,1]$ *be an ultramodular n-ary quasi-copula and suppose that* $D_1, D_2, \ldots, D_n\colon [0,1]^2 \to [0,1]$ *are bivariate copulas. If the functions* $f_1, f_2, \ldots, f_n, g_1, g_2, \ldots, g_n\colon [0,1] \to [0,1]$ *are monotone non-decreasing and continuous and satisfy* $Q(f_1(x), f_2(x), \ldots, f_n(x)) = Q(g_1(x), g_2(x), \ldots, g_n(x)) = x$ *for each* $x \in [0,1]$*, then also the function* $E\colon [0,1]^2 \to [0,1]$ *given by*

$$E(x,y) = Q\big(D_1(f_1(x), g_1(y)), D_2(f_2(x), g_2(y)), \ldots, D_n(f_n(x), g_n(y))\big)$$

is a bivariate copula.

Example 9.4.3. Consider the n-ary Fréchet-Hoeffding lower bound $W\colon [0,1]^n \to [0,1]$ which is an ultramodular quasi-copula. Note that for all monotone non-decreasing continuous functions $f_1, f_2, \ldots, f_n\colon [0,1] \to [0,1]$ and for each $x \in [0,1]$ we have $W\big(f_1(x), f_2(x), \ldots, f_n(x)\big) = x$ if and only if $\sum_{i=1}^n f_i = \mathrm{id}_{[0,1]} + n - 1$, where the identity function $\mathrm{id}_{[0,1]}\colon [0,1] \to [0,1]$ is given by $\mathrm{id}_{[0,1]}(x) = x$. Defining for each $i \in \{1, 2, \ldots, n\}$ the function $\varphi_i\colon [0,1] \to [0,1]$ by $\varphi_i = f_i - f_i(0)$ and putting $\xi_i = \varphi_i(1) \in [0,1]$, we obviously get $\sum_{i=1}^n \varphi_i = \mathrm{id}_{[0,1]}$ and $\sum_{i=1}^n \xi_i = 1$. Note that functions $\varphi_1, \varphi_2, \ldots, \varphi_n\colon [0,1] \to [0,1]$ with these properties are necessarily 1-Lipschitz, and they exist if and only if there is a bivariate copula C such that, for each $i \in \{1, 2, \ldots, n\}$ and for all $x \in [0,1]$,

$$\varphi_i(x) = C\big(x, \xi_1 + \xi_2 + \cdots + \xi_i\big) - C\big(x, \xi_1 + \xi_2 + \cdots + \xi_{i-1}\big), \tag{9.8}$$

where $\xi_0 = 0$ by convention (compare the *ordered modular averages* discussed in [49]). If now, for each $i \in \{1, 2, \ldots, n\}$, we choose $\varphi_i(x) = \xi_i x$ for each $x \in [0,1]$ (this corresponds to the case $C = \Pi$ in (9.8)), i.e., $f_i(x) = 1 - \xi_i + \xi_i x$, and if, for some $(\eta_1, \eta_2, \ldots, \eta_n) \in [0,1]^n$ satisfying $\sum_{i=1}^n \eta_i = 1$, we define the functions $g_1, g_2, \ldots, g_n\colon [0,1] \to [0,1]$ by $g_i(x) = 1 - \eta_i + \eta_i x$, for $D_1 = D_2 = \cdots = D_n = \Pi$ we obtain

$$\begin{aligned} W\Big(D_1\big(f_1(x), g_1(y)\big), D_2\big(f_2(x), g_2(y)\big), \ldots, D_n\big(f_n(x), g_n(y)\big)\Big) \\ = \max\Big(\sum_{i=1}^n (1 - \xi_i + \xi_i x)(1 - \eta_i + \eta_i y) - (n-1), 0\Big) \qquad (9.9) \\ = \max\Big(\Big(\sum_{i=1}^n \xi_i \eta_i\Big) xy + \Big(1 - \sum_{i=1}^n \xi_i \eta_i\Big)(x + y - 1), 0\Big). \end{aligned}$$

Observe that the copulas defined by (9.9) are all members of the family of bivariate copulas given by (4.2.7) in [53, Table 4.1], where the parameter $\theta \in [0,1]$ used there equals $\sum_{i=1}^n \xi_i \eta_i$ (note that, in the framework of *triangular norms* [35], this family is known as family of *Sugeno-Weber t-norms*).

If the bivariate copula C in Theorem 9.4.1 fails to be ultramodular, then the function constructed via (9.7) may not even be a quasi-copula.

Example 9.4.4. Put $C = M$ and $D_1 = D_2 = \Pi$ and recall that M is not ultramodular. Define the functions $f_1, f_2\colon [0,1] \to [0,1]$ by

$$f_1(x) = \begin{cases} x & \text{if } x \in \left[0, \frac{1}{2}\right], \\ 4x - \frac{3}{2} & \text{if } x \in \left]\frac{1}{2}, \frac{5}{8}\right], \\ 1 & \text{otherwise}, \end{cases} \qquad f_2(x) = \max(x, \tfrac{1}{2}),$$

and put $g_1 = f_1$ and $g_2 = f_2$. Clearly, $M(f_1(x), f_2(x)) = M(g_1(x), g_2(x)) = x$ for each $x \in [0,1]$. For $E\colon [0,1]^2 \to [0,1]$ given by $E(x,y) = M\big(f_1(x) \cdot g_1(y), f_2(x) \cdot g_2(y)\big)$ we have $\left|E\left(\frac{5}{8}, \frac{3}{8}\right) - E\left(\frac{1}{2}, \frac{3}{8}\right)\right| > \left|\frac{5}{8} - \frac{1}{2}\right| + \left|\frac{3}{8} - \frac{3}{8}\right|$, i.e., E is not 1-Lipschitz and, therefore, not even a quasi-copula.

For $\alpha \in [0,1]$ define the function $p_\alpha\colon [0,1] \to [0,1]$ given by $p_\alpha(x) = \max(x, \alpha)$. Obviously, $p_0 = \mathrm{id}_{[0,1]}$ and $p_1(x) = 1$ for all $x \in [0,1]$, and for each ultramodular copula C and for each $x \in [0,1]$ we have $C(p_0(x), p_1(x)) = C(p_1(x), p_0(x)) = x$.

Example 9.4.5. Define the functions $f_1, f_2\colon [0,1] \to [0,1]$ by $f_1(x) = \min(2x, 1)$ and $f_2(x) = \max(x, 1-x)$, respectively. Note that f_1 is not 1-Lipschitz and f_2 is not monotone, but we have $W(f_1(x), f_2(x)) = x$ for all $x \in [0,1]$. Applying the construction of Theorem 9.4.1 to the case $C = W$, $D_1 = D_2$, $g_1 = p_0$, and $g_2 = p_1$ we obtain

$D_1 = W$: $W\big(W(f_1(x), y), W(f_2(x), 1)\big) = W(x,y)$;
$D_1 = \Pi$: $W\big(\Pi(f_1(x), y), \Pi(f_2(x), 1)\big) = W\text{-}(\langle 0, \frac{1}{2}, \Pi\rangle)(x,y)$ which is a W-ordinal sum and, therefore, an ultramodular copula;
$D_1 = M$: $W\big(M(f_1(x), y), M(f_2(x), 1)\big) = \max(\max(2x + y - 1, 0) - x, 0)$ which is not monotone and, therefore, not even an aggregation function.

If C is an ultramodular copula we shall write $\mathscr{I}_C$ for the set of all pairs (f_1, f_2) of continuous, monotone non-decreasing functions $f_1, f_2\colon [0,1] \to [0,1]$ satisfying $C(f_1(x), f_2(x)) = x$ for each $x \in [0,1]$.

Let us fix some ultramodular bivariate copula $C\colon [0,1]^2 \to [0,1]$. In order to exploit Theorem 9.4.1 for the construction of (possibly non-symmetric) copulas by means of C it is advantageous to have at hand a reasonable sample of functions f_1 and f_2 such that $(f_1, f_2) \in \mathscr{I}_C$.

To start with, recall the diagonal section $\delta_C\colon [0,1] \to [0,1]$ of the ultramodular copula C, put $d_C = \sup\{x \in [0,1] \mid \delta_C(x) = 0\}$ and observe that the restriction $\delta_C \restriction [d_C, 1]$ of δ_C to $[d_C, 1]$ is a strictly increasing bijection because of Lemma 9.3.1. Defining $f\colon [0,1] \to [0,1]$ as its inverse, i.e., $f(x) = \big(\delta_C \restriction [d_C, 1]\big)^{-1}(x)$, we obviously get $C(f(x), f(x)) = \delta_C \circ f(x)$ for each $x \in [0,1]$, i.e., $(f, f) \in \mathscr{I}_C$.

Proposition 9.4.6. *For each ultramodular bivariate copula $C\colon [0,1]^2 \to [0,1]$ and for each $\alpha \in \,]0,1[$ there exist two uniquely determined monotone non-decreasing, continuous functions $r_\alpha^C, q_\alpha^C\colon [0,1] \to [0,1]$ such that $\{(p_\alpha, r_\alpha^C), (q_\alpha^C, p_\alpha)\} \subseteq \mathscr{I}_C$.*

Proof. Because of the ultramodularity of C, for each $\alpha \in \,]0,1[$ both the horizontal section $h_{C,\alpha}\colon [0,1] \to [0,\alpha]$ and the vertical section $v_{C,\alpha}\colon [0,1] \to [0,\alpha]$ of C given by $h_{C,\alpha}(x) = C(x,\alpha)$ and $v_{C,\alpha}(x) = C(\alpha, x)$, respectively, are convex, monotone non-decreasing, and 1-Lipschitz, and they satisfy $h_{C,\alpha}(0) = v_{C,\alpha}(0) = 0$ and $h_{C,\alpha}(1) = v_{C,\alpha}(1) = \alpha$. If we denote $a_{C,\alpha} = \sup\{x \in [0,1] \mid h_{C,\alpha}(x) = 0\}$ and $b_{C,\alpha} = \sup\{x \in [0,1] \mid v_{C,\alpha}(x) = 0\}$, then the restrictions $h_{C,\alpha} \upharpoonright [a_{C,\alpha},1]$ and $v_{C,\alpha} \upharpoonright [b_{C,\alpha},1]$ of $h_{C,\alpha}$ to $[a_{C,\alpha},1]$ and $v_{C,\alpha}$ to $[b_{C,\alpha},1]$, respectively, are strictly increasing bijections. Now, for each $\alpha \in [0,1]$ we can define the two continuous functions $r^C_\alpha, q^C_\alpha \colon [0,1] \to [0,1]$ by

$$r^C_\alpha(x) = \begin{cases} p_{1-\alpha}(x) & \text{if } \alpha \in \{0,1\}, \\ \left(v_{C,\alpha} \upharpoonright [b_{C,\alpha},1]\right)^{-1}(x) & \text{if } \alpha \in]0,1[\text{ and } x \in [0,\alpha[, \\ 1 & \text{otherwise,} \end{cases}$$

$$q^C_\alpha(x) = \begin{cases} p_{1-\alpha}(x) & \text{if } \alpha \in \{0,1\}, \\ \left(h_{C,\alpha} \upharpoonright [a_{C,\alpha},1]\right)^{-1}(x) & \text{if } \alpha \in]0,1[\text{ and } x \in [0,\alpha[, \\ 1 & \text{otherwise,} \end{cases}$$

and we have $\{(p_\alpha, r^C_\alpha), (q^C_\alpha, p_\alpha)\} \subseteq \mathscr{I}_C$. □

Clearly, if the bivariate copula $C\colon [0,1]^2 \to [0,1]$ is symmetric and the function $r^C_\alpha \colon [0,1] \to [0,1]$ exists then we have $q^C_\alpha = r^C_\alpha$.

Example 9.4.7. Consider the convex combination $C_\lambda = \lambda \Pi + (1-\lambda)W$ of the two extremal ultramodular bivariate copulas Π and W with $\lambda \in [0,1]$. Looking at its vertical sections, we have $v_{C_\lambda,\alpha}(x) = \lambda \alpha x + (1-\lambda)\max(\alpha + x - 1, 0)$, which in the case $\alpha \in]0,1[$ implies $b_{C_\lambda,\alpha} = \sup\{x \in [0,1] \mid v_{C_\lambda,\alpha}(x) = 0\} = 0$. This means that the vertical section $v_{C_\lambda,\alpha}$ itself is a bijection and that $r^{C_\lambda}_\alpha$ equals its inverse for $\alpha \in]0,1[$, i.e., $r^{C_\lambda}_\alpha$ is given by

$$r^{C_\lambda}_\alpha(x) = \begin{cases} p_{1-\alpha}(x) & \text{if } \alpha \in \{0,1\}, \\ \frac{x}{\lambda\alpha} & \text{if } \alpha \in]0,1[\text{ and } x \in [0, \lambda\alpha(1-\alpha)], \\ 1 - \frac{\alpha - x}{1-\lambda+\lambda\alpha} & \text{if } \alpha \in]0,1[\text{ and } x \in]\lambda\alpha(1-\alpha), \alpha[, \\ 1 & \text{otherwise.} \end{cases}$$

Since C_λ is ultramodular and symmetric we may put $q^{C_\lambda}_\alpha = r^{C_\lambda}_\alpha$ for each $\alpha \in [0,1]$.

The approach in Theorem 9.4.1 can also be used to generate a copula which is a Π-horizontal ordinal sum [48, 60] (see also the gluing of copulas [68]).

Example 9.4.8. For $\lambda \in]0,1[$ consider, as in Example 9.4.7, the convex combination $C_\lambda = \lambda\Pi + (1-\lambda)W$ of the ultramodular bivariate copulas Π and W and the function $r^{C_\lambda}_\alpha \colon [0,1] \to [0,1]$. Then, for each $\alpha \in]0,1[$ the function $C\colon [0,1]^2 \to [0,1]$ given by

$$C(x,y) = C_\lambda\left(\Pi(x, p_\alpha(y)), W\left(1, r^{C_\lambda}_\alpha(y)\right)\right) = C_\lambda\left(x p_\alpha(y), r^{C_\lambda}_\alpha(y)\right)$$

is a bivariate copula (visualized in Figure 9.2 left and center). It is not difficult to see that C is indeed a Π-horizontal ordinal sum: keeping the notations of [60] and putting $\kappa = \frac{\alpha\lambda}{\alpha\lambda - \lambda + 1}$, we obtain $C = \Pi_h$-$(\langle \lambda\alpha(1-\alpha), \alpha, C_\kappa \rangle)$.

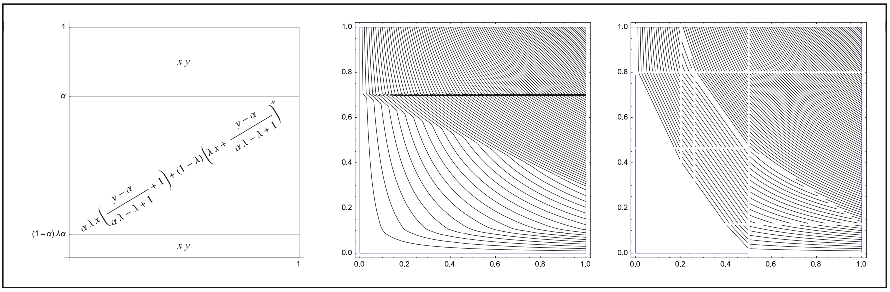


Fig. 9.2: The Π-horizontal sum (left and center) in Example 9.4.8, and the copula in Example 9.4.9

Example 9.4.9. For $\alpha, \beta, \lambda \in [0,1]$ consider, as in Example 9.4.7, the convex combination $C_\lambda = \lambda \Pi + (1-\lambda)W$ of the two ultramodular bivariate copulas Π and W and the functions $r_\alpha^{C_\lambda}, r_\beta^{C_\lambda} \colon [0,1] \to [0,1]$. Then, for each $\alpha, \beta, \lambda \in \,]0,1[$ the function $C \colon [0,1]^2 \to [0,1]$ given by

$$C(x,y) = C_\lambda\Big(\Pi(p_\alpha(x), p_\beta(y)), W\Big(r_\alpha^{C_\lambda}(x), r_\beta^{C_\lambda}(y)\Big)\Big)$$

is a bivariate copula (visualized in Figure 9.2 right). Observe that C coincides with Π on the set $([\alpha,1]\times[\beta,1]) \cup ([\alpha,1]\times[0,\lambda\beta(1-\beta)]) \cup ([0,\lambda\alpha(1-\alpha)]\times[\beta,1])$, and that C vanishes (coinciding there with W) on a set containing the triangle determined by the vertices $(0,0)$, $(0,\lambda\beta(1-\beta))$, and $(\lambda\alpha(1-\alpha),0)$.

Example 9.4.10. Let $(]a_i,b_i[)_{i\in I}$ be a family of non-empty, pairwise disjoint open subintervals of $[0,1]$ and $(C_i)_{i\in I}$ be a family of ultramodular bivariate copulas, and let C be their W-ordinal sum, i.e., $C = W\text{-}(\langle a_i,b_i,C_i\rangle)_{i\in I}$. If $\alpha \in \,]0,1[$ then for the continuous function $r_\alpha^C \colon [0,1] \to [0,1]$ which satisfies $C\big(p_\alpha(x), r_\alpha^C(x)\big) = x$ for each $x \in [0,1]$ we have to distinguish two cases, depending on the position of α with respect to the family of open intervals $(]a_i,b_i[)_{i\in I}$:

(i) if there is an $i_* \in I$ such that $\lambda \in \,]a_{i_*}, b_{i_*}[$ then $r_\alpha^C \colon [0,1] \to [0,1]$ is given by

$$r_\alpha^C(x) = \begin{cases} (1-b_{i_*}) + (b_{i_*}-a_{i_*})\, r^{C_{i_*}}_{\frac{\alpha-a_{i_*}}{b_{i_*}-a_{i_*}}}\left(\frac{x}{b_{i_*}-a_{i_*}}\right) & \text{if } x \in [0, \alpha - a_{i_*}], \\ 1-\alpha+x & \text{if } x \in \,]\alpha - a_{i_*}, \alpha[, \\ 1 & \text{otherwise;} \end{cases}$$

(ii) if $\lambda \notin \bigcup_{i\in I}]a_i,b_i[$ then $r_\alpha^C \colon [0,1] \to [0,1]$ is given by $r_\alpha^C(x) = \min(1-\alpha+x, 1)$, and we have $r_\alpha^C = r_\alpha^W$.

There is also a counterpart of Theorem 9.4.1 for (ultramodular) quasi-copulas. We formulate it here for the bivariate case only (see [31, Corollary 5.2]):

Corollary 9.4.11. *Let $R, Q_1, Q_2\colon [0,1]^2 \to [0,1]$ be bivariate quasi-copulas and assume that R is ultramodular. If $f_1, f_2, g_1, g_2\colon [0,1] \to [0,1]$ are monotone non-decreasing continuous functions satisfying $R(f_1(x), f_2(x)) = R(g_1(x), g_2(x)) = x$ for each $x \in [0,1]$, then the function $L\colon [0,1]^2 \to [0,1]$ given by*

$$L(x,y) = R\big(Q_1(f_1(x), g_1(y)), Q_2(f_2(x), g_2(y))\big) \tag{9.10}$$

is a quasi-copula.

9.5 A product-like construction

Now we summarize some results from [31] demonstrating the role of (partial) ultramodularity and Schur concavity when constructing copulas.

If $D, A\colon [0,1]^2 \to [0,1]$ are binary aggregation functions and if A is 1-Lipschitz then the function $D(A, A^*)\colon [0,1]^2 \to [0,1]$ given by

$$D(A,A^*)(x,y) = D(A(x,y), A^*(x,y)) \tag{9.11}$$

is called the *D-product of A and its dual A^**.

When we start with two bivariate quasi-copulas $D, Q\colon [0,1]^2 \to [0,1]$ then it was shown in Proposition 5.3 in [31] that the D-product $D(Q, Q^*)$ of Q and its dual Q^* is always a bivariate quasi-copula.

Example 9.5.1. If $D, C\colon [0,1]^2 \to [0,1]$ are bivariate copulas and if one of them equals one of the three basic copulas W, Π and M, then the D-product of C and its dual C^* quite often (but not always) is also a copula [31].

(i) For each copula C we trivially get $W(C, C^*) = W$ and $M(C, C^*) = C$, and for each copula D we have $D(W, W^*) = W$.

(ii) In [38] it was shown that, for each copula C, also $\Pi(C, C^*)$ is a copula.

(iii) Let D be a copula. Then $D(M, M^*)$ is a copula if and only if, for the copula $C_{\delta_D} : [0,1]^2 \to [0,1]$ given by

$$C_{\delta_D}(x,y) = \min\big(x, y, \tfrac{D(x,x)+D(y,y)}{2}\big)$$

and for the restriction $D\restriction_\Delta$ of D to Δ we have $D\restriction_\Delta \leq C_{\delta_D}\restriction_\Delta$. Note that the copula C_{δ_D} is the *diagonal copula* (see [54]) whose diagonal section coincides with the diagonal section δ_D of D.

(iv) For a symmetric copula D we obtain $D(M, M^*) = D$. Consider the (non-symmetric) copula $D = W\text{-}(\langle 0, 0.5, \Pi\rangle)$ given by

$$D(x,y) = \begin{cases} \Pi(x, 2y-1) & \text{if } (x,y) \in [0,0.5] \times [0.5,1], \\ W(x,y) & \text{otherwise,} \end{cases}$$

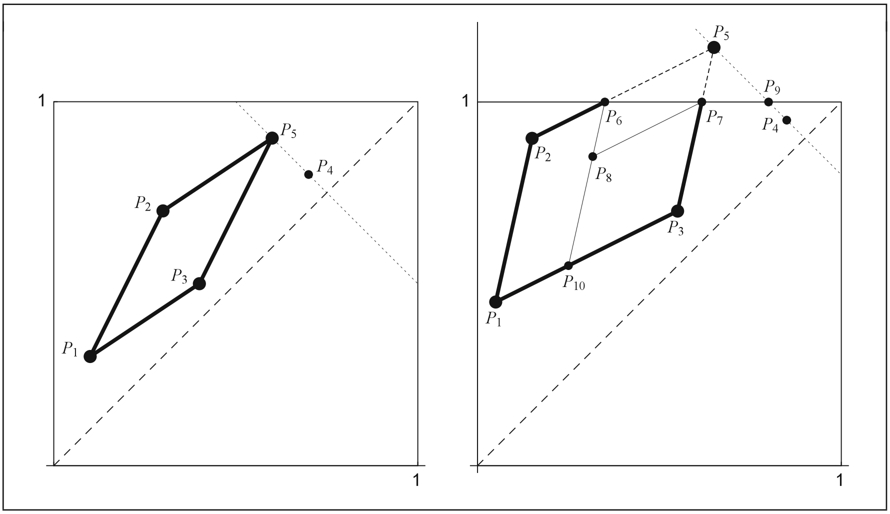


Fig. 9.3: The two cases in the proof of Theorem 9.5.2

and notice that it satisfies $D \restriction_\Delta \leq C_{\delta_D} \restriction_\Delta$. Then $D(M,M^*)$ is a copula different from D: indeed, $D(M,M^*) = W\text{-}(\langle 0,0.5,\Pi\rangle, \langle 0.5,1,\Pi\rangle) \neq D$.

(v) Consider the (non-symmetric) W-ordinal sum $D = W\text{-}(\langle 0,0.8,M\rangle)$ given by

$$D(x,y) = \begin{cases} M(x,y-0.2) & \text{if } (x,y) \in [0,0.8] \times [0.2,1], \\ W(x,y) & \text{otherwise.} \end{cases}$$

Then the $D(M,M^*)$-volume $V_{D(M,M^*)}$ of, e.g., the square $[0.4,0.6]^2 \subseteq [0,1]^2$ is negative, i.e., $D(M,M^*)$ is not a copula.

The following main result of [31] (there it is Theorem 3.1) shows that, for a copula D, the validity of the ultramodularity [UM] and the Schur concavity on the upper left triangle $\Delta \subseteq [0,1]^2$ are sufficient conditions for the D-product of an arbitrary copula C and its dual C^* to be a copula. Note that, although the upper Fréchet-Hoeffding bound M is not ultramodular on $[0,1]^2$, it is ultramodular on the upper left triangle Δ, so Theorem 9.5.2 will apply to $D = M$ as well.

Theorem 9.5.2. *Let D be a copula which is ultramodular and Schur concave on the upper left triangle $\Delta = \{(x,y) \in [0,1]^2 \mid x \leq y\}$. Then, for each copula C, the D-product $D(C,C^*)$ of C and its dual C^* is a copula.*

The proof of Theorem 9.5.2 as given in [31] turns out to be quite elaborate, so we present here a sketch of it highlighting its key points. Since the boundary conditions are trivial, the main issue is to show that $D(C,C^*)$ is 2-increasing, i.e., that inequality (9.1) holds for arbitrary, but fixed, numbers $x_1,x_2,y_1,y_2 \in [0,1]$ with $x_1 \leq x_2$ and $y_1 \leq y_2$. Choosing the points P_1,P_2,P_3 and P_4 in $[0,1]^2$ according to

$$P_1 = (C(x_1,y_1),C^*(x_1,y_1)), \qquad P_2 = (C(x_2,y_1),C^*(x_2,y_1)),$$
$$P_3 = (C(x_1,y_2),C^*(x_1,y_2)), \qquad P_4 = (C(x_2,y_2),C^*(x_2,y_2)),$$

it is not difficult to see that they all belong to the upper left triangle Δ. Putting

$$\begin{aligned} P_5 &= P_2 + P_3 - P_1 \\ &= (C(x_2,y_1)+C(x_1,y_2)-C(x_1,y_1),C^*(x_2,y_1)+C^*(x_1,y_2)-C^*(x_1,y_1)), \end{aligned}$$

i.e., extending the triangle with vertices P_1, P_2 and P_3 to a parallelogram, there are two possible cases.

If $P_5 \in \Delta$ (as in Figure 9.3 left) then the Schur concavity of D on Δ implies $D(C,C^*)(x_2,y_2) = D(P_4) \geq D(P_5)$, and the ultramodularity of D on Δ then leads to $D(P_1) - D(P_3) + D(P_5) - D(P_2) \geq 0$, which means that $D(C,C^*)$ satisfies inequality (9.1) in this case and, therefore, is a copula.

If $P_5 \notin \Delta$ (as in Figure 9.3 right) then we consider the points P_6–P_{10} as in the figure and observe that, again because of the Schur concavity of D on Δ,

$$\begin{aligned} &D(C,C^*)(x_1,y_1) - D(C,C^*)(x_1,y_2) + D(C,C^*)(x_2,y_2) - D(C,C^*)(x_2,y_1) \\ &\quad = D(P_1) - D(P_3) + D(P_4) - D(P_2) \\ &\quad = D(P_1) + D(P_6) - D(P_2) - D(P_{10}) && \text{(a)} \\ &\qquad + D(P_{10}) + D(P_7) - D(P_3) - D(P_8) && \text{(b)} \\ &\qquad + D(P_9) - D(P_7) - D(P_6) + D(P_8) && \text{(c)} \\ &\qquad + D(P_4) - D(P_9). && \text{(d)} \end{aligned}$$

Now it is not too difficult to see that the expressions (a) and (b) are both non-negative because of the ultramodularity of D on Δ, that expression (c) is non-negative because of the Schur concavity of D on Δ, and that the non-negativity of expression (d) is essentially due to $D \geq W$. Summarizing, $D(C,C^*)$ satisfies inequality (9.1) also in this case and, therefore, is a copula.

Example 9.5.3. The following examples (see Example 4.1 in [31]) illustrate the importance of the hypotheses in Theorem 9.5.2:

(i) Consider the W-ordinal sum $D = W\text{-}(\langle 0, 0.5, \Pi \rangle)$ as in Example 9.5.1(iv) and the M-ordinal sum $C = M\text{-}(\langle \frac{1}{4}, \frac{13}{24}, W \rangle)$ given by

$$C(x,y) = \begin{cases} \max(x+y-\frac{13}{24},\frac{1}{4}) & \text{if } (x,y) \in \left[\frac{1}{4},\frac{13}{24}\right]^2, \\ M(x,y) & \text{otherwise.} \end{cases}$$

Observe that D is ultramodular, but not Schur concave on the upper left triangle Δ. Then the $D(C,C^*)$-volume $V_{D(C,C^*)}$ of the square $[0.375,0.5]^2 \subseteq [0,1]^2$ is negative, i.e., $D(C,C^*)$ is not a copula.

(ii) Consider the M-ordinal sum $D = M\text{-}(\langle 0.5, 1, W \rangle)$ given by

$$D(x,y)=\begin{cases}\max(x+y-1,0.5) & \text{if } (x,y)\in[0.5,1]^2,\\ M(x,y) & \text{otherwise.}\end{cases}$$

Observe that D is Schur concave, but not ultramodular on the upper left triangle Δ. Then the $D(\Pi,\Pi^*)$-volume $V_{D(\Pi,\Pi^*)}$ of the square $\left[\frac{2}{3},\frac{3}{4}\right]^2\subseteq[0,1]^2$ is negative, i.e., $D(\Pi,\Pi^*)$ is not a copula.

(iii) Theorem 9.5.2 cannot be modified replacing the dual copula C^* by the co-copula $\overline{C}\colon [0,1]^2\to[0,1]$ given by $\overline{C}(x,y)=1-C(1-x,1-y)$. Indeed, for the M-ordinal sum $C=M\text{-}(\langle 0,0.5,\Pi\rangle)$ given by

$$C(x,y)=\begin{cases}2\Pi(x,y) & \text{if } (x,y)\in[0,0.5]^2,\\ M(x,y) & \text{otherwise}\end{cases}$$

the $\Pi(C,\overline{C})$-volume $V_{\Pi(C,\overline{C})}$ of $[0.3,0.4]^2$ is negative, i.e., $\Pi(C,\overline{C})$ is not a copula.

It is remarkable (see Proposition 3.2 in [31]) that the construction in Theorem 9.5.2 preserves the ultramodularity and the Schur concavity on Δ of the copulas involved. More precisely, if $C,D\colon [0,1]^2\to[0,1]$ are bivariate copulas which are ultramodular and Schur concave on the upper left triangle Δ, then also the copula $D(C,C^*)$ is ultramodular and Schur concave on Δ.

It also turns out (Theorem 3.3 in [31]) that the ultramodularity of D on Δ is a necessary condition in Theorem 9.5.2: if $D\colon [0,1]^2\to[0,1]$ is a bivariate copula such that for each bivariate copula $C\colon [0,1]^2\to[0,1]$ also the function $D(C,C^*)$ is a copula, then D must be ultramodular on the upper left triangle Δ.

The construction in Theorem 9.5.2 allows us also to construct sequences of copulas converging to the Fréchet-Hoeffding lower bound W (see [31]):

Let $D,C\colon [0,1]^2\to[0,1]$ be bivariate copulas and assume that D is ultramodular and Schur concave on Δ and different from the Fréchet-Hoeffding upper bound M. Then we necessarily have $D(x,y)<x$ for all $(x,y)\in\Delta\cap\,]0,1[^2$ and $D(C,C^*)<C$. Therefore, we can inductively define a sequence of bivariate copulas $(C_n)_{n\in\mathbb{N}}$ putting $C_1=C$, and $C_{n+1}=D(C_n,(C_n)^*)$ for each $n\in\mathbb{N}$, and we obtain $\lim\limits_{n\to\infty}C_n=W$.

Keep the notations of the previous paragraph and define the trivariate function $f\colon \big\{(x,y,z)\in[0,1]^3 \,\big|\, x+y-z\in[0,1]\big\}\to[0,1]$ by $f(x,y,z)=D(z,x+y-z)$, and put $z_n=C_n(x,y)$, where (x,y) is an arbitrary but fixed point in $[0,1]^2$. Then we have $z_1=C(x,y)$ and $z_{n+1}=f(x,y,z_n)$ for each $n\in\mathbb{N}$, implying $\lim\limits_{n\to\infty}z_n=W(x,y)$.

Taking into account the flipping method for constructing new bivariate copulas [53] which transforms ultramodular bivariate copulas into copulas with concave horizontal and vertical sections [45], Theorem 9.5.2 can be modified as follows (Corollary 3.5 in [31]):

Let $E\colon [0,1]^2\to[0,1]$ be a bivariate copula which is concave on the horizontal and vertical sections which are contained in the lower left triangle Δ_* determined by the vertices $(0,0)$, $(0,1)$ and $(1,0)$, and assume that E is 1-Lipschitz with respect to the Chebyshev norm on the affine sections which are parallel to the main diagonal of

$[0,1]^2$ and are contained in Δ_*, i.e., $E(x+\varepsilon, y+\varepsilon) - E(x,y) \leq \varepsilon$ for all $x,y \geq 0$ and all $\varepsilon > 0$ with $x+y+2\varepsilon \leq 1$. Then, for each bivariate copula $C\colon [0,1]^2 \to [0,1]$, the function $E_C\colon [0,1]^2 \to [0,1]$ given by $E_C(x,y) = C(x,y) - E(C(x,y), 1 - C^*(x,y))$ is a bivariate copula.

Since the product copula Π is invariant under flipping, this construction and the one in Theorem 9.5.2 yield the same result in the case $D = E = \Pi$, i.e., $\Pi(C, C^*) = \Pi_C$ for each bivariate copula $C\colon [0,1]^2 \to [0,1]$. For the extremal cases $E = W$ (which vanishes on Δ_*) and $E = M$ we obtain $W_C = C$ and $M_C = W$ for each copula C.

9.6 Concluding remarks

We have presented two constructions of bivariate copulas involving ultramodular copulas.

The first approach extends the result of [44] that, in a composition of functions, the ultramodularity of the outer function preserves the supermodularity of the inner function.

In [38] it was shown that the product of a bivariate copula and its dual is always a copula. In our second construction the product was replaced by a copula D which is ultramodular and Schur concave on the upper left triangle Δ of the unit square [31]. The ultramodularity on Δ is necessary, but it is still an open question whether or to which extent the Schur concavity on Δ can be weakened in order to obtain a condition which is necessary and sufficient.

Most of the constructions in this survey led to bivariate copulas. The next challenge will be to obtain copulas of higher dimensions, a task which will require some strengthening of the ultramodularity.

Acknowledgements This work was supported by the "Technologie-Transfer-Förderung" of the Upper Austrian Government (Wi-2014-200710/13-Kx/Kai), and by grant APVV-14-0013.

References

1. Ali, M.M., Mikhail, N.N., Haq, M.S.: A class of bivariate distributions including the bivariate logistic. J. Multivariate Anal. **8**, 405–412 (1978)
2. Alsina, C., Nelsen, R.B., Schweizer, B.: On the characterization of a class of binary operations on distribution functions. Statist. Probab. Lett. **17**, 85–89 (1993)
3. Aumann, R.J., Shapley, L.S.: Values of Non-Atomic Games. Princeton University Press, Princeton (1974)
4. Barlow, R.E., Spizzichino, F.: Schur-concave survival functions and survival analysis. J. Comput. Appl. Math. **46**, 437–447 (1993)
5. Beccacece, F., Borgonovo, E.: Functional ANOVA, ultramodularity and monotonicity: Applications in multiattribute utility theory. European J. Oper. Res. **210**, 326–335 (2011)

6. Block, H.W., Griffith, W.S., Savits, T.H.: L-superadditive structure functions. Adv. in Appl. Probab. **21**, 919–929 (1989)
7. Brunk, H.D.: Integral inequalities for functions with nondecreasing increments. Pacific J. Math. **14**, 783–793 (1964)
8. Capéraà, P., Genest, C.: Spearman's ρ is larger than Kendall's τ for positively dependent random variables. J. Nonparametr. Statist. **2**, 183–194 (1993)
9. Clifford, A.H.: Naturally totally ordered commutative semigroups. Amer. J. Math. **76**, 631–646 (1954)
10. De Baets, B., De Meyer, H.: Orthogonal grid constructions of copulas. IEEE Trans. Fuzzy Systems **15**, 1053–1062 (2007)
11. Durante, F.: Construction of non-exchangeable bivariate distribution functions. Statist. Papers **50**, 383–391 (2009)
12. Durante, F., Mesiar, R., Papini, P.L., Sempi, C.: 2-increasing binary aggregation operators. Inform. Sci. **177**, 111–129 (2007)
13. Durante, F., Saminger-Platz, S., Sarkoci, P.: On representations of 2-increasing binary aggregation functions. Inform. Sci. **178**, 4534–4541 (2008)
14. Durante, F., Saminger-Platz, S., Sarkoci, P.: Rectangular patchwork for bivariate copulas and tail dependence. Comm. Statist. Theory Methods **38**, 2515–2527 (2009)
15. Durante, F., Sempi, C.: Copulæ and Schur-concavity. Int. Math. J. **3**, 893–905 (2003)
16. Durante, F., Sempi, C.: On the characterization of a class of binary operations on bivariate distribution functions. Publ. Math. Debrecen **69**, 47–63 (2006)
17. Durante, F., Sempi, C.: Principles of Copula Theory. CRC Press, Boca Raton (2015)
18. Frank, M.J.: On the simultaneous associativity of $F(x,y)$ and $x+y-F(x,y)$. Aequationes Math. **19**, 194–226 (1979)
19. Fréchet, M.: Remarques au sujet de la note précédente. C. R. Acad. Sci. Paris **246**, 2719–2720 (1958)
20. Genest, C., Quesada-Molina, J.J., Rodríguez-Lallena, J.A., Sempi, C.: A characterization of quasi-copulas. J. Multivariate Anal. **69**, 193–205 (1999)
21. Gong, W.M., Sun, H., Chu, Y.M.: The Schur convexity for the generalized Muirhead mean. J. Math. Inequal. **8**, 855–862 (2014)
22. Grabisch, M., Marichal, J.L., Mesiar, R., Pap, E.: Aggregation Functions. Cambridge University Press, Cambridge (2009)
23. Hamacher, H.: Über logische Aggregationen nicht-binär explizierter Entscheidungskriterien. Rita G. Fischer Verlag, Frankfurt (1978)
24. Hamacher, H.: Über logische Verknüpfungen unscharfer Aussagen und deren zugehörige Bewertungsfunktionen. In: R. Trappl, G.J. Klir, L. Ricciardi (eds.) Progress in Cybernetics and Systems Research III, pp. 276–288. Hemisphere Publ. Comp., New York (1978)
25. Jensen, J.L.W.V.: Sur les fonctions convexes et les inégalités entre les valeurs moyennes. Acta Math. **30**, 175–193 (1906)
26. Joe, H.: Multivariate Models and Dependence Concepts. Chapman & Hall, London (1997)
27. Kemperman, J.H.B.: On the FKG-inequality for measures on a partially ordered space. Nederl. Akad. Wetensch. Proc. Ser. A **39**, 313–331 (1977)
28. Kenyon, H.: Note on convex functions. Amer. Math. Monthly **63**, 107 (1956)
29. Kimberling, C.: On a class of associative functions. Publ. Math. Debrecen **20**, 21–39 (1973)
30. Klee, V.L.: Solution of a problem of E. M. Wright on convex functions. Amer. Math. Monthly **63**, 106–107 (1956)
31. Klement, E.P., Kolesárová, A., Mesiar, R., Saminger-Platz, S.: On the role of ultramodularity and Schur concavity in the construction of binary copulas. J. Math. Inequal. **11**, 361–381 (2017)
32. Klement, E.P., Kolesárová, A., Mesiar, R., Sempi, C.: Copulas constructed from horizontal sections. Comm. Statist. Theory Methods **36**, 2901–2911 (2007)
33. Klement, E.P., Manzi, M., Mesiar, R.: Ultramodular aggregation functions. Inform. Sci. **181**, 4101–4111 (2011)
34. Klement, E.P., Manzi, M., Mesiar, R.: Ultramodularity and copulas. Rocky Mountain J. Math. **44**, 189–202 (2014)

35. Klement, E.P., Mesiar, R., Pap, E.: Triangular Norms. Kluwer Academic Publishers, Dordrecht (2000)
36. Klement, E.P., Mesiar, R., Pap, E.: Uniform approximation of associative copulas by strict and non-strict copulas. Illinois J. Math. **45**, 1393–1400 (2001)
37. Kolesárová, A.: 1-Lipschitz aggregation operators and quasi-copulas. Kybernetika **39**, 615–629 (2003)
38. Kolesárová, A., Mesiar, R., Kalická, J.: On a new construction of 1-Lipschitz aggregation functions, quasi-copulas and copulas. Fuzzy Sets and Systems **226**, 19–31 (2013)
39. Lau, D.: Function Algebras on Finite Sets. A Basic Course on Many-Valued Logic and Clone Theory. Springer, Berlin (2006)
40. Li, D.M., Shi, H.N.: Schur convexity and Schur-geometrically concavity of generalized exponent mean. J. Math. Inequal. **3**, 217–225 (2009)
41. Liebscher, E.: Construction of asymmetric multivariate copulas. J. Multivariate Anal. **99**, 2234–2250 (2008)
42. Ling, C.M.: Representation of associative functions. Publ. Math. Debrecen **12**, 189–212 (1965)
43. Mardia, K.V.: Families of Bivariate Distributions. Charles Griffin & Co., London (1970)
44. Marinacci, M., Montrucchio, L.: Ultramodular functions. Math. Oper. Res. **30**, 311–332 (2005)
45. Marinacci, M., Montrucchio, L.: On concavity and supermodularity. J. Math. Anal. Appl. **344**, 642–654 (2008)
46. Marshall, A.W., Olkin, I.: Majorization in multivariate distributions. Ann. Statist. **2**, 1189–1200 (1974)
47. Marshall, A.W., Olkin, I., Arnold, B.C.: Inequalities: Theory of Majorization and Its Applications, second edn. Springer Series in Statistics. Springer, New York (2011)
48. Mesiar, R., Jágr, V., Juráňová, M., Komorníková, M.: Univariate conditioning of copulas. Kybernetika **44**, 807–816 (2008)
49. Mesiar, R., Mesiarová-Zemánková, A.: The ordered modular averages. IEEE Trans. Fuzzy Systems **19**, 42–50 (2011)
50. Mesiar, R., Szolgay, J.: W-ordinal sums of copulas and quasi-copulas. In: Proceedings MAGIA & UWPM 2004, pp. 78–83. Slovak University of Technology, Publishing House of STU, Bratislava (2004)
51. Moynihan, R.: On τ_T semigroups of probability distribution functions II. Aequationes Math. **17**, 19–40 (1978)
52. Müller, A., Scarsini, M.: Stochastic comparison of random vectors with a common copula. Math. Oper. Res. **26**, 723–740 (2001)
53. Nelsen, R.B.: An Introduction to Copulas, second edn. Springer, New York (2006)
54. Nelsen, R.B., Fredricks, G.A.: Diagonal copulas. In: V. Beneš, J. Štěpán (eds.) Distributions with given Marginals and Moment Problems, pp. 121–128. Kluwer Academic Publishers, Dordrecht (1997)
55. Nevius, S.E., Proschan, F., Sethuraman, J.: Schur functions in statistics. II. Stochastic majorization. Ann. Statist. **5**, 263–273 (1977)
56. Proschan, F., Sethuraman, J.: Schur functions in statistics. I. The preservation theorem. Ann. Statist. **5**, 256–262 (1977)
57. Roberts, A.W., Varberg, D.E.: Convex functions. Academic Press, New York (1973)
58. Rovenţa, I.: A note on Schur-concave functions. J. Inequal. Appl. **2012**, 9 pages (2012)
59. Saminger-Platz, S., Arias-García, J.J., Mesiar, R., Klement, E.P.: Characterizations of bivariate conic, extreme value, and Archimax copulas. Depend. Model. **5**, 45–58 (2017)
60. Saminger-Platz, S., Dibala, M., Klement, E.P., Mesiar, R.: Ordinal sums of binary conjunctive operations based on the product. Publ. Math. Debrecen **91** (2017, in press)
61. Schur, I.: Über eine Klasse von Mittelbildungen mit Anwendungen auf die Determinantentheorie. Sitzungsber. Berl. Math. Ges. **22**, 9–20 (1923)
62. Schweizer, B., Sklar, A.: Associative functions and abstract semigroups. Publ. Math. Debrecen **10**, 69–81 (1963)

63. Schweizer, B., Sklar, A.: Probabilistic Metric Spaces. Dover Publications, Mineola (2006)
64. Shaked, M., Shanthikumar, J.G.: Stochastic orders. Springer, New York (2007)
65. Shi, H.N.: Schur-convex functions related to Hadamard-type inequalities. J. Math. Inequal. **1**, 127–136 (2007)
66. Shi, H.N., Zhang, J.: Schur-convexity, Schur geometric and Schur harmonic convexities of dual form of a class symmetric functions. J. Math. Inequal. **8**, 349–358 (2014)
67. Shi, H.N., Zhang, J., Gu, C.: New proofs of Schur-concavity for a class of symmetric functions. J. Inequal. Appl. pp. 2012:12, 4 (2012)
68. Siburg, K.F., Stoimenov, P.A.: Gluing copulas. Comm. Statist. Theory Methods **37**, 3124–3134 (2008)
69. Sklar, A.: Fonctions de répartition à *n* dimensions et leurs marges. Publ. Inst. Statist. Univ. Paris **8**, 229–231 (1959)
70. Sugeno, M.: Theory of fuzzy integrals and its applications. Ph.D. thesis, Tokyo Institute of Technology (1974)
71. Weber, S.: A general concept of fuzzy connectives, negations and implications based on t-norms and t-conorms. Fuzzy Sets and Systems **11**, 115–134 (1983)
72. Wright, E.M.: An inequality for convex functions. Amer. Math. Monthly **61**, 620–622 (1954)

Chapter 10
Operations on Finite Settings: from Triangular Norms to Copulas

Gaspar Mayor, Jaume Suñer, and Joan Torrens

Abstract Operations defined on finite chains, usually known as discrete operations, have become a topic of increasing interest because of their applications in many fields. In this paper, we recall two of the main families of discrete operations: discrete t-norms with applications in fuzzy logic, approximate reasoning and computing with words, and discrete copulas with applications in probability, statistics and economy. For both kinds of operations, we emphasize their main properties, their structure and characterization. We also devote some part of the work to highlight the importance of these discrete operations and their impact in the mentioned fields.

10.1 Introduction

Although we already superficially knew copulas and their applications, it was after meeting Professor Roger Nelsen that we really care and devote part of our investigation to the field of copulas. Thus, our production in this topic is partially due to the influence of Professor Nelsen in our research and it is for this reason that we are very glad to have been invited to collaborate in this book devoted to him.

Taking into account that part of our field of expertise lies in the topic of *discrete operations*, understood as operations defined on finite chains, the introduction and study of copula-like operations defined on a finite chain L was one of the first topics related to copulas that we dealt with. The importance of discrete operations lies in the fact that they allow us to avoid numerical interpretations of linguistic variables. Although the most usual scale in fuzzy logic is the real unit interval $[0,1]$, only a finite number of values is used in most applications. Moreover, expert reasoning is usually carried out through linguistic terms that are always increasingly ordered

Gaspar Mayor, Jaume Suñer, Joan Torrens
Department of Mathematics and Computer Science, Universitat de les Illes Balears, 07122 Palma de Mallorca, Spain
e-mail: {gmayor,jaume.sunyer,jts224}@uib.es

M. Úbeda Flores et al. (eds.), *Copulas and Dependence Models with Applications*,
DOI 10.1007/978-3-319-64221-5_10

forming a finite chain. Thus, operations defined on this discrete setting allow us to directly manage these linguistic terms, with the consequent applications in fuzzy logic, approximate reasoning and computing with words.

The current paper has been thought as a naive survey on discrete operations, paying special attention to the seminal work that was the starting point of the theory, that is, the paper where discrete t-norms and t-conorms were introduced, and the paper that was devoted to the introduction and study of discrete copulas. However, we also compile several references where other kind of discrete operators can be found, as well as many works derived from these two aforementioned papers in which their importance and their impact was pointed out. We will do it devoting one section to each one of the two types of discrete operators: t-norms and copulas.

10.2 Discrete t-norms

All through this section the domain of all discrete binary operations will be the finite chain $L = \{0, 1, \ldots, n\}, n \geq 1$.

Definition 10.1. A mapping $F : L \times L \to L$ is said to be a *discrete aggregation function if it is increasing in each argument and such that* $F(0,0) = 0$, $F(n,n) = n$.

This section is devoted to a special class of aggregation functions, the so-called t-norms.

Definition 10.2. A *discrete triangular norm* (briefly *discrete t-norm*) on L is a function $T : L \times L \to L$ such that for all $x, y, z \in L$ the following axioms are satisfied:

(T1) $T(x,y) = T(y,x)$ (Commutativity)
(T2) $T(T(x,y),z) = T(x,T(y,z))$ (Associativity)
(T3) $T(x,y) \leq T(x,z)$ whenever $y \leq z$ (Monotonicity)
(T4) $T(x,1) = x$ (Boundary conditions)

Basic examples of discrete t-norms on L are the *minimum* T_M and the *drastic* T_D, which are given by:

$$T_M(x,y) = \min(x,y) \qquad \text{and} \qquad T_D(x,y) = \begin{cases} 0 & \text{if } \max(x,y) \neq n \\ \min(x,y) & \text{if } \max(x,y) = n \end{cases}$$

With similar arguments as in the case $L = [0,1]$, basic properties of any discrete t-norm on L can be proved. We collect some of them as follows.

Proposition 10.1. *Let T be a discrete t-norm on L. Then:*

- $T_D \leq T \leq T_M$*, that is T_D and T_M are the smallest and the largest t-norms respectively (where $\leq$ denotes the pointwise ordering).*
- $T(x,0) = T(0,x) = 0$ *for all $x \in L$ (that is, 0 is an annihilator).*
- T_M *is the only t-norm for which all elements in L are idempotent (that is, $T_M(x,x) = x$ for all $x \in L$).*

Some additional properties can be required to a discrete t-norm depending on the context where they have to be applied. Some of them are listed in the following definition stated for any commutative aggregation function F in general[1].

Definition 10.3. Let $F : L \times L \to L$ be a commutative discrete aggregation function. We say that F

a) is *divisible* if for all $x, y \in L$ with $x \leq y$, there is $z \in L$ such that $x = F(y, z)$,
b) is *smooth* if for all $x, y \in L$ with $x \leq n-1$, $F(x+1, y) - F(x, y) \leq 1$,
c) satisfies the *1-Lipschitz condition* if for all $x, y, z \in L$ such that $z \geq x$, $F(z, y) - F(x, y) \leq z - x$,
d) satisfies the *"intermediate value theorem"* if the following condition holds: let $F(x, y) = z$ and $F(x, y') = z'$ with $z < z'$; then for all $z'' \in [z, z']$, there is some $y'' \in [y, y']$ such that $F(x, y'') = z''$.

All these conditions are in fact equivalent when we deal with discrete t-norms as follows.

Proposition 10.2. *Given any discrete t-norm T on L, the following statements are equivalent:*

- *T is divisible.*
- *T is smooth.*
- *T satisfies the 1-Lipschitz condition.*
- *T satisfies the "intermediate value theorem".*

Another important property for discrete t-norms is given in the following definition.

Definition 10.4. Let $T : L \times L \to L$ be a discrete t-norm. We say that T is *Archimedean* if it satisfies the following condition:

$$\text{For all } x, y \in L \setminus \{0, n\} \text{ there is } m \in \mathbb{N} \text{ such that } x_T^{(m)} < y \qquad (10.1)$$

where $x_T^{(m)}$ is defined in the usual way by induction,

$$x_T^{(m)} = \begin{cases} x, & \text{if } m = 1, \\ T(x_T^{(m-1)}, x), & \text{if } m \geq 2. \end{cases}$$

For any discrete t-norm, 0 and n are always idempotent elements, i.e., $T(0,0) = 0$, $T(n,n) = n$. They are usually called *trivial idempotent elements*. In view of this fact, we give the following definition.

Definition 10.5. A discrete t-norm T on L is *idempotent-free* if the only idempotent elements of T are 0 and n.

[1] Moreover, all these properties can be extended also to the non-commutative case just by requiring the conditions in both variables instead of requiring them only in one variable.

For discrete t-norms we have the following equivalence, which in the case of the real interval $[0,1]$, only holds for continuous t-norms.

Proposition 10.3. *Let T be a t-norm on L. Then T is Archimedean if and only if it is idempotent-free.*

We recall now a method for constructing a new t-norm from a family of given t-norms. It is based on results concerning ordinal sums of semigroups (see [6]), that can be adapted to the case of t-norms on $[0,1]$ (see [13]) and also in our discrete setting. In fact, we only need here a simple version of this method. First let us point out the following notation.

Remark 10.1. Given two elements $i, j \in L$ with $i < j$ we will denote by $[i, j]$ the subset of L given by

$$[i, j] = \{k \in L \mid i \leq k \leq j\}.$$

Note that such interval is again a finite chain of $j - i + 1$ elements and so we can consider also t-norms on this chain $[i, j]$.

Thus, given $r \in \{0, 1, \ldots, n-1\}$, let $a_0, a_1, \ldots, a_r, a_{r+1} \in L$ such that $0 = a_0 < a_1 < \cdots < a_r < a_{r+1} = n$. Let us take $L_i = [a_i, a_{i+1}]$ with $i \in J = \{0, 1, \ldots, r\}$, a family of intervals of L, and for each $i \in J$, let T_i be a discrete t-norm defined on $[a_i, a_{i+1}]$. We can construct a new discrete t-norm on L as follows:

$$T(x,y) = \begin{cases} T_i(x,y) & \text{if } (x,y) \in [a_i, a_{i+1}]^2 \\ \min(x,y) & \text{otherwise.} \end{cases} \tag{10.2}$$

Definition 10.6. Let $([a_i, a_{i+1}], T_i)$, $i \in J$, be a family of intervals of L and discrete t-norms on these intervals like above. Then the discrete t-norm T on L defined by (10.2) is called the *ordinal sum* of $([a_i, a_{i+1}], T_i)$, $i \in J$, and we will denote it by $T =< ([a_i, a_{i+1}], T_i)_{i \in J} >$.

Remark 10.2. The set of idempotent elements of $T =< ([a_i, a_{i+1}], T_i)_{i \in J} >$ contains $0, n$ and all the other end-points of the intervals $[a_i, a_{i+1}]$, $i \in J$. Note that the equality between these two sets occurs if and only if the set of idempotent elements of each t-norm T_i is $\{a_i, a_{i+1}\}$, that is, each T_i is idempotent-free.

The structure of an ordinal sum like in the definition above can be viewed in Figure 1.

Proposition 10.4. *Let $T =< ([a_i, a_{i+1}], T_i)_{i \in J} >$ be an ordinal sum of discrete t-norms. Then T is divisible if and only if T_i is divisible for all $i \in J$.*

In this general but simple assumption, we can establish a representation theorem of divisible t-norms on L by means of ordinal sums.

Theorem 10.1. *Let T be a discrete t-norm on L with the set of its idempotent elements given by $\{0 = a_0, a_1, \ldots, a_r, a_{r+1} = n\}$ and let $J = \{0, 1, \ldots, r\}$. Then the following are equivalent:*

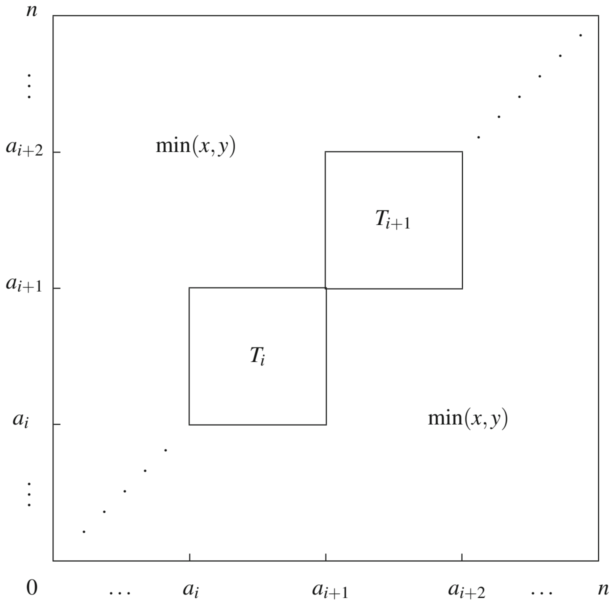


Figure 1. An ordinal sum of t-norms T_i defined on intervals $[a_i, a_{i+1}]$.

i) T is divisible
ii) T is an ordinal sum $T =< ([a_i, a_{i+1}], T_i)_{i \in J} >$ *where each* T_i *is a divisible, idempotent-free t-norm on* $[a_i, a_{i+1}]$.

With respect to the Archimedean property, we have the following results.

Proposition 10.5. *Let T be a discrete t-norm on L. Then T is Archimedean if and only if*

$$T(x,y) \neq \min(x,y) \qquad \textit{for all} \quad x,y \in L \setminus \{0,n\}$$

Proposition 10.6. *The only Archimedean divisible discrete t-norm on L is the Łukasiewicz t-norm*

$$T_L(x,y) = \max(0, x+y-n) \qquad \textit{for all} \quad x,y \in L$$

Now, we can state the following characterization of the class of divisible discrete t-norms as ordinal sums of Łukasiewicz t-norms.

Theorem 10.2. *A discrete t-norm T on* $L = \{0,1,...,n\}$ *is divisible if and only if there exists a natural number r with* $0 \leq r \leq n-1$ *and a subset of L,* $I = \{0 = a_0 < a_1 < \ldots < a_r < a_{r+1} = n\}$ *such that T is given by:*

$$T(x,y) = \begin{cases} \max(a_i, x+y-a_{i+1}) & \textit{if } (x,y) \in [a_i, a_{i+1}]^2 \textit{ for some } 0 \leq i \leq r \\ \min(x,y) & \textit{otherwise} \end{cases} \tag{10.3}$$

Remark 10.3. Let us denote by T^I the discrete t-norm given by equation (10.3). I is the set of idempotent elements of T. Note that

- In case $r=0$, that is $I=\{0,n\}$, we obtain $T^I=T_L$.
- In case $r=n-1$, that is $I=L$, we obtain $T^I=T_M$.
- It is easily deduced from Theorem 10.2 that $T_L \leq T^I \leq T_M$ for any divisible discrete t-norm T^I.

Example 10.1. Consider $n=7, L=\{0,1,2,3,4,5,6,7\}$ and $I=\{0,4,7\}$. According to the previous results, there is a unique divisible discrete t-norm on L with I as the set of its idempotent elements. It is given by

$$T^I(x,y)=\begin{cases}\max(0,x+y-4) & \text{if } (x,y)\in[0,4]^2\\ \max(4,x+y-7) & \text{if } (x,y)\in[4,7]^2\\ \min(x,y) & \text{otherwise}\end{cases}$$

and it can be viewed explicitly in Table 1.

T	**0**	**1**	**2**	**3**	**4**	**5**	**6**	**7**
0	**0**	0	0	0	0	0	0	0
1	0	0	0	0	1	1	1	1
2	0	0	0	1	2	2	2	2
3	0	0	1	2	3	3	3	3
4	0	1	2	3	**4**	4	4	4
5	0	1	2	3	4	4	4	5
6	0	1	2	3	4	4	5	6
7	0	1	2	3	4	5	6	**7**

Table 1. The only divisible t-norm on $L=\{0,1,2,3,4,5,6,7\}$ with idempotent elements 0, 4, 7.

Let us denote by $\mathcal{T}_{DIV}(L)$ the set of all divisible discrete t-norms on L. From the given classifications of divisible discrete t-norms, we obtain the following Corollary.

Corollary 10.1. *The correspondence $\psi:\mathcal{T}_{DIV}(L)\to\mathcal{P}(L\setminus\{0,n\})$ between the set of all divisible discrete t-norms on L and the power set of $\{1,2,\ldots,n-1\}$, defined by $\psi(T)=I\setminus\{0,n\}$ (the set of non-trivial idempotent elements of T), is a bijection. Thus there are exactly 2^{n-1} divisible discrete t-norms on L.*

In table 2 we can view, depending on n, the number of discrete t-norms and the number of divisible discrete t-norms on $L=\{0,1,\ldots,n\}$.

The results presented here are only the most important ones related to discrete t-norms, but many other details and facts can be found in [28]. To finish this section let us highlight the usefulness and impact of the results recalled just before. This study

n	$\mid \mathscr{T}(L) \mid$	$\mid \mathscr{T}_{Div}(L) \mid$
1	1	1
2	2	2
3	6	4
4	22	8
5	94	16
6	451	32
7	2386	64

Table 2. Number of discrete t-norms and divisible discrete t-norms on $L = \{0, 1, ..., n\}$.

on discrete t-norms was firstly published in [27] and, due to its impact and interest, it was rewritten and completed in several aspects leading to the book chapter in [28]. Some relevant aspects to be stressed are as follows:

- As we have already commented in the introduction, the importance of discrete operations lies in the fact that they allow us to avoid numerical interpretations of linguistic variables, with the consequent applications in fuzzy logic and approximate reasoning. In this sense, not only t-norms were studied on finite chains, but also other logical operators like t-conorms [28], uninorms [5, 17] and also some different kinds of implication functions like residual and material implication derived from discrete t-norms and t-conorms [19], QL and D-implications [20], or residual implications derived from discrete uninorms [16].
- The one-to-one correspondence between divisible discrete t-norms and finite BL-chains relates this work with BL-algebras and related structures, as well as with formal fuzzy logic. For instance, this relation has been recently used to provide an equational characterization of any divisible t-norm, see [7] and the references therein.
- This work was pioneer in the study of aggregation functions on finite scales or discrete aggregation functions and depicted the starting point of many later papers devoted to this topic, which is the basis of many linguistic approaches for decision making, consensus processes and linguistic preference relations. Thus, discrete t-norms have been later used to characterize many new classes of discrete aggregation functions [9, 15, 21, 22] and, in particular, uninorms and null-norms [17], non-commutative versions of them [8, 18], idempotent uninorms [5], weighted means [10], and also copulas and quasi-copulas [1, 2, 11, 12, 25, 26] (see also next section).
- Moreover, discrete t-norms have been more deeply studied leading to additional papers related to this topic. In particular, the research of additive generators for discrete t-norms led to some interesting papers proving for instance that all di-

visible discrete t-norms, but not all discrete t-norms, are additively generated [14, 23, 24]. On the other hand, extensions of discrete t-norms to more general settings have been investigated like extensions to multisets [3] or to discrete fuzzy numbers [4].

10.3 Discrete copulas

Definition 10.7. A *(discrete) copula C on $L=\{0,1,\ldots,n\}, n\geq 1$* is a binary operation on L, i.e., $C: L\times L \longrightarrow L$, such that the following axioms are satisfied:

(C1) $C(i,0)=C(0,i)=0\ \forall i\in L$
(C2) $C(i,n)=C(n,i)=i\ \forall i\in L$
(C3) $C(i,j)+C(i',j')\geq C(i,j')+C(i',j)$ whenever $i\leq i', j\leq j'$
(2-increasing condition)

An operation $C: D\times D' \longrightarrow L$, with $\{0,n\}\subset D, D'\subset L$ satisfying $(C1)-(C3)$ of the definition above is called a *discrete subcopula on L*. Both the concepts of discrete copula and discrete subcopula are the discrete versions of copula and subcopula on [0,1].

Remark 10.4. Let us consider I_n the subset of $[0,1]$ given by $I_n=\{0,1/n,\ldots,(n-1)/n,1\}$.

i) Note that any operation $C: L\times L\to L$ is a discrete copula on L if and only if the operation $C': I_n\times I_n\to I_n$ given by $C'(a,b)=1/n\cdot C(na,nb)$ for all $a,b\in I_n$ is a discrete copula on I_n. Moreover, note that in this case the operation C' can be also understood as a subcopula on $[0,1]$ with domain $DomC'=I_n\times I_n$ and range $RanC'=I_n$.
Reciprocally, if C' is a subcopula on $[0,1]$ with domain $DomC'=I_n\times I_n$ and range $RanC'=I_n$, then $C(a,b)=n\cdot C'(a/n,b/n)$ for all $a,b\in L$ is a discrete copula on L.
ii) Similarly, there is a close relation between discrete copulas and empirical copulas as follows. Consider $P=\{(x_1,y_1),\ldots,(x_n,y_n)\}$ a set of n ordered real pairs where $x_i\neq x_j,\ y_i\neq y_j$, for $i\neq j,\ \ i,j: 1,\ldots,n$. Consider $x_{(1)}\leq\cdots\leq x_{(n)}$ and $y_{(1)}\leq\cdots\leq y_{(n)}$ the increasing reordering of the collections $x_1,\ldots,x_n$ and $y_1,\ldots,y_n$ respectively. The *empirical* discrete copula defined from P is the discrete copula, say C_P, defined on $L=\{0,1,\ldots,n\}$ by: $\forall (i,j)\in L^2$,

$$C_P(i,j)=\begin{cases}0 & \text{if } i=0 \text{ or } j=0\\ \left|\{(x_k,y_k)\in P: x_k\leq x_{(i)}, y_k\leq y_{(j)}\}\right| & \text{otherwise}\end{cases}$$

Reciprocally, any discrete copula on L is the empirical discrete copula of some set P.

Observe that if we apply the construcion given in the previous item to C_P, we obtain the empirical copula in the sense of [30]: For all $i, j \in L$,

$$C_n\left(\frac{i}{n}, \frac{j}{n}\right) = 1/n \cdot C(i,j) = \begin{cases} 0 & \text{if } i = 0 \text{ or } j = 0 \\ \dfrac{\left|\{(x_k, y_k) \in P : x_k \leq x_{(i)}, y_k \leq y_{(j)}\}\right|}{n} & \text{otherwise} \end{cases}$$

Example 10.2. It is clear that discrete copulas need not to be commutative nor associative and so they are far to coincide with discrete t-norms. However, as in the case of $[0,1]$ there exist intersections between both families of discrete operations. For instance we have the following examples.

(i) $C_M(x,y) = \min(x,y)$ and $C_L(x,y) = \max(0, x+y-n)$ are discrete copulas. We call them the minimum and the Łukasiewicz copula respectively (in statistics and probability fields, C_L is commonly known as the Fréchet copula). It is already known that these copulas are also (divisible) discrete t-norms on L.

(ii) C defined on $L = \{0,1,2,3\}$ by

$$C(x,y) = \begin{cases} C_L(x,y) = \max(0, x+y-3) & \text{if } (x,y) \neq (1,2) \\ 1 & \text{if } (x,y) = (1,2) \end{cases}$$

is a discrete copula which is not a t-norm.

Next we summarize some basic properties of discrete copulas that can be proved in the same way as for copulas defined on [0,1].

Proposition 10.7. *Let C be a discrete copula on L. The following statements are valid:*

(i) $C_L \leq C \leq C_M$.
(ii) C is non-decreasing in each variable.
(iii) C satisfies the Lipschitz condition.
(iv) The only discrete copula C that satisfies $C(x,x) = x\ \forall x \in L$ is $C = C_M$.
(v) The only discrete copula C that satisfies $C(x, n-x) = 0$ (or $C(n-x, x) = 0$) for all $x \in L$ is $C = C_L$.

Similarly to the case of discrete t-norms, a discrete copula C is said to be *divisible* when for all $x, y \in L$ with $x \leq y$, there are $z, z' \in L$ such that $x = C(y,z) = C(z',y)$. The second part of the next result is a consequence of the structure of divisible discrete t-norms given in Theorem 10.2.

Proposition 10.8. *All discrete copulas on L are divisible and all divisible discrete t-norms are (associative) discrete copulas.*

An explicit and useful characterization of discrete copulas can be obtained using permutation matrices. Recall that an $n \times n$ permutation matrix A is an $n \times n$ matrix (a_{ij}) such that there exists a permutation σ_A of $\{1, 2, \ldots, n\}$ such that

$$a_{ij} = \begin{cases} 1 & \text{if } i = \sigma_A(j) \\ 0 & \text{otherwise} \end{cases}$$

Note that this is equivalent to saying that in each row and each column of A all entries are equal to 0 except one which is 1.

Proposition 10.9. *A binary operation C on L is a discrete copula if and only if there exists an $n \times n$ permutation matrix $A = (a_{ij})$ such that $\forall (x,y) \in L^2$,*

$$C(x,y) = \begin{cases} 0 & \text{if } x = 0 \text{ or } y = 0 \\ \sum\limits_{\substack{i \le x \\ j \le y}} a_{ij} & \text{otherwise} \end{cases} \tag{10.4}$$

A discrete copula C on L given through an $n \times n$ permutation matrix A by equation (10.4) will be denoted by C_A. Thus, the previous proposition says that for any discrete copula C on L, there exists an $n \times n$ permutation matrix A such that $C = C_A$. From this fact, we easily obtain the following corollaries.

Corollary 10.2. *The correspondence $A \longrightarrow C_A$ defined by equation (10.4) is a bijection between the set of $n \times n$ permutation matrices and the set of discrete copulas on L. Thus, there are $n!$ discrete copulas on L.*

Corollary 10.3. *The set of discrete copulas on L can be equipped in a natural way with a group structure defining the product of two discrete copulas by $C_A \cdot C_B = C_{A \cdot B}$. This group is obviously isomorphic to the symmetric group S_n of all permutations on $\{1, \ldots, n\}$ and has the copula minimum as neutral element.*

Corollary 10.4. *A discrete copula is commutative if and only if its associated permutation matrix is symmetric.*

Example 10.3.

$i)$ The identity matrix I_n is the associated $n \times n$ permutation matrix to the copula C_M. In this case the associated permutation is the identity.

$ii)$ The corresponding matrix to the Łukasiewicz copula C_L is the $n \times n$ permutation matrix $A = (a_{ij})$ given by:

$$a_{ij} = \begin{cases} 1, & \text{if } i + j = n + 1, \\ 0, & \text{otherwise,} \end{cases}$$

and for this reason we call A the $n \times n$ *Łukasiewicz matrix*. Note that in this case, the associated permutation σ_A is given by:

$$\sigma_A = (n, n-1, \ldots, 1) = \prod_{i:1 \le i < n+1-i} (i, n+1-i)$$

where (i, j) denotes the transposition which interchanges i and j, and $\prod$ means the composition of the transpositions $(i, n+1-i)$.

Given $A_1,\ldots,A_k, k \geq 1$, where each A_i is the $n_i \times n_i$ Łukasiewicz permutation matrix, we can construct a new $n \times n$ $(n = n_1 + \cdots + n_k)$ permutation matrix $A = A_1 \oplus \cdots \oplus A_k$ in the following form:

$$A = \begin{pmatrix} A_1 & \cdots & 0 \\ \cdots & \cdots & \cdots \\ 0 & \cdots & A_k \end{pmatrix}$$

We call A the *ordinal sum* of $A_1,\ldots,A_k$. Of course, A is symmetric.

Example 10.4. Let C_A be the discrete copula on $L = \{0,1,2,3,4,5,6,7,8,9\}$ given by the matrix:

$$A = \begin{pmatrix} 0&0&1&0&0&0&0&0&0 \\ 0&1&0&0&0&0&0&0&0 \\ 1&0&0&0&0&0&0&0&0 \\ 0&0&0&1&0&0&0&0&0 \\ 0&0&0&0&0&1&0&0&0 \\ 0&0&0&0&1&0&0&0&0 \\ 0&0&0&0&0&0&0&0&1 \\ 0&0&0&0&0&0&0&1&0 \\ 0&0&0&0&0&0&1&0&0 \end{pmatrix}$$

A is an ordinal sum of four Łukasiewicz matrices: $A = A_1 \oplus A_2 \oplus A_3 \oplus A_4$, where

$$A_1 = \begin{pmatrix} 0&0&1 \\ 0&1&0 \\ 1&0&0 \end{pmatrix}, \qquad A_2 = (1), \qquad A_3 = \begin{pmatrix} 0&1 \\ 1&0 \end{pmatrix}, \qquad A_4 = \begin{pmatrix} 0&0&1 \\ 0&1&0 \\ 1&0&0 \end{pmatrix},$$

$\sigma_A = (3,2,1,4,6,5,9,8,7)$ is the permutation associated to A.

The next two propositions give results on associative copulas.

Proposition 10.10. *The following items hold.*

i) The only associative discrete copula on L with 0 *and n as the only idempotent elements is the Łukasiewicz copula.*

ii) A discrete copula is associative if and only if its associated matrix is an ordinal sum of Łukasiewicz matrices.

iii) If a discrete copula is associative, then it is commutative.

iv) The class of associative discrete copulas coincides with the class of divisible discrete t-norms.

Proposition 10.11. *Any discrete copula is a product of associative discrete copulas. In other words, for each copula C there exist divisible discrete t-norms $T_1,\ldots,T_r$ such that $C = T_1 \cdot \cdots \cdot T_r$.*

Note that these are only the main results about discrete copulas, but more details can be found in [25], including for instance different ways to extend discrete copulas to copulas on $[0,1]$: the bilinear extension that gives a monomorphism from the group of discrete copulas into a semigroup of copulas on $[0,1]$, and 2^n different extensions of any discrete copula to n-regular shuffles of M.

To end this section we would like to highlight again the importance and impact of the results recalled here as we have done in the previous section with discrete t-norms. What is named here “discrete copula” corresponds to those with minimal range (that is, taking values in the same finite set I_n). They are important because of their relationship with discrete bivariate distribution functions with uniform discrete univariate marginals. When the marginals are uniformly distributed with step $1/n$, the joint distribution can be represented by a discrete copula C that gives information on the dependence structure of the involved random variables. In this sense, the most immediate sequel of this work (that was published in [25]) was the discrete version of the well known Sklar Theorem proving that the joint distribution function of two discrete marginals with range included in I_n can be described through a discrete copula C.

From these two works, some other papers related to discrete copulas have appeared generalizing the results presented here. For instance,

- Discrete copulas in general, that is, with range in the unit interval $[0,1]$ were studied in [11]. In this paper, these general operations are called discrete copulas, whereas the discrete operations presented here are referred as “irreducible discrete copulas”. It is proved that there is an annalogous relation between discrete copulas and bistochastic matrices and that the set of all discrete copulas is simply the convex closure of the set of all irreducible discrete copulas.
- An extremely related paper is [12] where irreducible discrete quasi-copulas (that is, those with minimal range, of the form $Q : I_n \times I_n \to I_n$) and some of their properties are investigated. Ordinal sum structures are analyzed and the problem of finding irreducible quasi-copulas with some given diagonal section is solved, characterizing the cases when such extension is unique (except for the case of copulas). However, the following two open problems arise:
  - Counting the exact number of irreduccible discrete quasi-copulas existing on I_n, and
  - Characterizing all diagonal sections δ for which there exists a unique copula C having δ as its diagonal.
- These two open problems were solved in [1] precisely by generalyzing the study of discrete copulas given here to discrete quasi-copulas. Thus, in a similar way as in the current section, a one-to-one correspondence between irreducible discrete quasi-copulas and Alternating Sign Matrices (ASM’s for short) is proved. This result allows to find the number of irreducible discrete quasi-copulas as the same number of ASM’s (see [31]), that is:

$$\prod_{k=0}^{n-1} \frac{(3k+1)!}{(n+k)!} = \prod_{k=1}^{n} \frac{(3k-2)!}{(2n-k)!}.$$

  Similarly, it is proved that the set of discrete quasi-copulas is bijective with a class of matrices that generalize both, ASM’s and bistochastic matrices, called there

Generalized Bistochastic Matrices (GBM's for short). Again, it is also proved that the set of all discrete quasi-copulas is the convex closure of the set of all irreducible discrete quasi-copulas. Moreover, by using the structure of ASM's it is easy to see when, given a diagonal section δ, there exists a unique irreducible discrete quasi-copula with diagonal δ retriving the results in [12], but solving also the second open problem related to copulas.

- Finally, note that all previous works deal with discrete copulas and quasi-copulas defined on $I_n \times I_n$. Generalizations to non-square grids of the unit square, that is to domains of the form $I_n \times I_m$ with $n \neq m$, can be found in [2] and [29].

Acknowledgements The authors want to dedicate this work to professor Roger Nelsen on the occasion of his 75th birthday.

The last two authors have been partially supported by the Spanish Grant TIN2016-75404-P, AEI/FEDER, UE.

References

1. I. Aguiló, J. Suñer, J. Torrens, Matrix representation of discrete quasi-copulas, *Fuzzy Sets and Systems* 159 (2008) 1658–1672.
2. I. Aguiló, J. Suñer, J. Torrens, Matrix representation of copulas and quasi-copulas defined on non-square grids of the unit square, *Fuzzy Sets and Systems* 161 (2010) 254–268.
3. J. Casasnovas, G. Mayor, Discrete t-norms and operations on extended multisets, *Fuzzy Sets and Systems* 159 (2008) 1165–1177.
4. J. Casasnovas, J.V. Riera, Extension of discrete t-norms and t-conorms to discrete fuzzy numbers, *Fuzzy Sets and Systems* 167(1) (2011) 65–81.
5. B. De Baets, J. Fodor, D. Ruiz-Aguilera, J. Torrens, Idempotent uninorms on finite ordinal scales, *International Journal of Uncertainty, Fuzziness and Knowledge-Based Systems* 17 (2009) 1-14.
6. A.H. Clifford (1954), Naturally totally ordered commutative semigroups, *Amer. J. Math.*, 76, pp. 631-646.
7. F. Esteva, A. García-Cerdaña, L. Godo, Smooth Finite t-norms and Their Equational Axiomatization, in T. Calvo, J. Torrens (Eds), "Fuzzy Logic and Information Fusion. To commemorate the 70th Birthday of Professor Gaspar Mayor", *Studies in Fuzziness and Soft Computing*, 339 (2016) pp. 11–24.
8. J.C. Fodor, Smooth associative operations on finite ordinal scales, *IEEE Transactions on Fuzzy Systems* 8 (2000) 791-795.
9. L. Godo, V. Torra, On aggregation operators for ordinal qualitative information, *IEEE Transactions on Fuzzy Systems* 8 (2000) 143-154.
10. A. Kolesarova, G. Mayor, R. Mesiar, Weighted ordinal means, *Information Sciences* 177 (2007) 3822–3830.
11. A. Kolesárová, R. Mesiar, J. Mordelová, C. Sempi, Discrete copulas, *IEEE Transactions on Fuzzy Systems* 14 (2006) 698–705.
12. A. Kolesárová, J. Mordelová, Quasi-copulas and copulas on a discrete scale, *Soft Computing* 10 (2006) 495–501.
13. E. P. Klement, R. Mesiar and E. Pap (2000), *Triangular norms*, Kluwer Academic Publishers. Dordrecht.
14. J. Martín, G. Mayor, J. Monreal, The problem of the additive generation of finitely-valued t-conorms, *Mathware and Soft Computing* 16 (2009) 17–27.

15. J. Martín, G. Mayor, J. Torrens, On locally internal monotonic operations, *Fuzzy Sets and Systems* 137 (2003) 27–42.
16. M. Mas, G. Mayor, M. Monserrat, J. Torrens, Residual Implications from Discrete Uninorms. A Characterization, in L. Magdalena, J.L. Verdegay F. Esteva (Eds)"Enric Trillas: A Passion for Fuzzy Sets", *Studies in Fuzziness and Soft Computing*, 322 (2015) pp. 27–40.
17. M. Mas, G. Mayor, J. Torrens, t-Operators and uninorms on a finite totally ordered set, *International Journal of Intelligent Systems* 14 (1999) 909-922.
18. M. Mas, M. Monserrat, J. Torrens, On left and right uninorms on a finite chain, *Fuzzy Sets and Systems* 146 (2004) 3-17.
19. M. Mas, M. Monserrat, J. Torrens, S-implications and R-implications on a finite chain, *Kybernetica* 40 (2004) 3–20.
20. M. Mas, M. Monserrat, J. Torrens, On two types of discrete implications, *International Journal of Approximate Resoning* 40 (2005) 262–279.
21. M. Mas, M. Monserrat, J. Torrens: Smooth aggregation functions on finite scales. In E. Hüllermeier, T. Kruse, F. Hoffmann, (eds.) Proceedings of IPMU'2010, LNAI 6178, pp. 398–407, Dortmund (2010)
22. M. Mas, M. Monserrat, J. Torrens, Kernel aggregation functions on finite scales. Constructions from their marginals, *Fuzzy Sets and Systems* 241 (2014) 27–40.
23. G. Mayor, J. Monreal, Additive generators of discrete conjunctive aggregation operations, *IEEE Transactions on Fuzzy Systems* 15 (6) (2007) 1046–1052.
24. G. Mayor, J. Monreal, On some classes of additive generators, *Fuzzy Sets and Systems* 264 (2015) 110–120
25. G. Mayor, J. Suñer, J. Torrens, Copula-like operations on finite settings, *IEEE Transactions on Fuzzy Systems* 13 (2005) 468–477.
26. G. Mayor, J. Suñer, J. Torrens, Sklar's theorem on finite settings, *IEEE Transactions on Fuzzy Systems* 15 (2007) 410–416.
27. G. Mayor, J. Torrens, On a class of operators for expert systems, *International Journal of Intelligent Systems* 8 (1993) 771–778.
28. G. Mayor, J. Torrens, Triangular norms on discrete settings, in E.P. Klement and R. Mesiar (eds.), *Logical, Algebraic, Analytic, and Probabilistic Aspects of Triangular Norms*, Elsevier, 2005, pp. 189–230.
29. J. Mordelová, A. Kolesárová, Some results on discrete copulas, in Proceedings of *EUSFLAT*-2007, pp. 145–150.
30. R. B. Nelsen, *An Introduction to Copulas*, New York: Springer-Verlag, 1999.
31. D.P. Robbins, H. Rumsey, Determinants and alternating-sign matrices, *Advances in Mathematics* 62 (1986) 169–184.

Chapter 11
My meetings with Roger B. Nelsen

José Juan Quesada-Molina

On the occasion of the 75th anniversary of Roger B. Nelsen

Abstract In this paper, we will review our many meetings with Roger B. Nelsen, from August 1986 to the present day. A commented summary of the main results of Roger B. Nelsen and other colleagues work with the author is presented. Most of the mathematical results described here were obtained in the theory of copulas and quasi-copulas.

11.1 Introduction and first meeting

I met Roger B. Nelsen for the first time in August 1986, in Mount Holyoke College (South Hadley, Massachusetts, USA) at the XXIV International Symposium on Functional Equations, organized by Berthold Schweizer. Since September 1985, I had been visiting the University of Massachusetts in Amherst, to work on research under the advice and expertise of Berthold Schweizer, after my Ph.D. thesis at the University of Granada, in Spain. I must say that I met such a great group of people at that conference in Mount Holyoke: good mathematicians and good friends afterwards. I had had the opportunity to meet some of them before that conference: Claudi Alsina, in Barcelona (Spain); Jerry Frank, in Amherst (Massachusetts, USA); Abe Sklar, in Chicago (Illinois, USA); and Howard Sherwood and Michael D. Taylor, in Orlando (Florida, USA). I also met many others for the first time at that conference: Carlo Sempi, Robert Moynihan, János Aczél, Edward F. Wolff, Robert Tardiff, Richard Rice,… and Roger B. Nelsen. I will never be able to thank my good friend and mentor, Professor Berthold Schweizer, enough, for his hospitality

José Juan Quesada-Molina
Department of Applied Mathematics, University of Granada, 18071 Granada (Spain)
e-mail: jquesada@ugr.es

M. Úbeda Flores et al. (eds.), *Copulas and Dependence Models with Applications*,
DOI 10.1007/978-3-319-64221-5_11

in Amherst, for having faith in me, for teaching me so many valuable things that I could never have learned from books, and for introducing me to such a great group of friends and colleagues.

That first meeting with Roger B. Nelsen in Mount Holyoke College in August 1986 was the first of many meetings over the last thirty years, where friendship and mathematics have been shared with other colleagues: Berthold Schweizer, Carlo Sempi, José Antonio Rodríguez-Lallena, Manuel Úbeda-Flores, Fabrizio Durante, and Juan Fernández-Sánchez.

In the following sections of this paper, the main results of the work of Roger B. Nelsen and the above mentioned colleagues with the author are presented sequentially. Most of the mathematical results described here were obtained in the theory of copulas and quasi-copulas.

11.2 Second meeting: Via the Università di Lecce (Italy)

In September 1990, I visited Carlo Sempi for the first time at the Università di Lecce, in Italy. Lecce is a beautiful city in the south of Italy, in a region called Puglia. I spent a month there working with Carlo in a wonderful environment. It was a pleasure for me to be in Lecce. Carlo provided me with everything for a wonderful stay, and as a result, it was a delightful and fruitful visit; the beginning of a sequence of visits to Lecce, and the start of a great friendship with Carlo and his family.

During that visit to Lecce, Carlo and I were working on the problem of the derivability of operations on distribution functions. I remember that, back in Granada a few months later, I wrote Roger an e-mail, and I explained to him what Carlo and I had been researching. He told me that something very similar was the subject of ongoing research with Bert Schweizer. Actually, we were using different notation, but the same concepts with different names. So, we (Bert, Roger, Carlo and I) decided to join forces, and write a joint paper on the subject. As a result, we wrote [27], and we presented it at the Symposium on Distributions with Fixed Marginals, Doubly Stochastic Measures, and Markov Operators held in Seattle (Washington, USA), in August 1993. In that paper, we characterized the operations on distribution functions that are both derivable from functions on random variables defined on a common probability space and induced pointwise by functions from $[0,1]^n$ into $[0,1]$. We specified the class of functions on random variables from which the operations are derived, and showed that it includes all order statistics. We also described the n-place functions from which these operations are induced pointwise. Finally, we also showed, by way of illustration, that mixtures, which are induced pointwise, are not derivable. I remember this as an interesting and very technical paper, and we truly enjoyed working on it.

That was my second meeting with Roger B. Nelsen, not personally, but through Carlo Sempi and the Università di Lecce.

11.3 Third meeting: Lewis & Clark College (Portland, OR, USA)

In January-February 1995, I visited Roger B. Nelsen at Lewis & Clark College in Portland (Oregon, USA) for three weeks. I hold wonderful memories from that visit. During my stay at Lewis & Clark College, Roger and I, together with José Antonio Rodríguez-Lallena, from Almería (Spain), started working on the problem of constructing copulas whose horizontal or vertical sections are cubic polynomials. Before continuing, I must state the obvious: the concept of copula was introduced by Abe Sklar in 1959 [33]. It was a very clever idea from Abe Sklar: A copula is a function linking the joint distribution function of a random vector to their marginal distribution functions; and, therefore, the copula captures the whole dependence structure among the variables of the random vector. This is why Berthold Schweizer and Edward F. Wolff used copulas to define and study new nonparametric measures of dependence for random variables (see [32]). Back to the work with Roger at Lewis & Clark, the idea of looking for such a type of copulas was, in a sense, a continuation of previous work (see [28]) that José Antonio Rodríguez-Lallena and I developed two years before, about bivariate copulas with quadratic sections. One of the main results of that work, with Roger and José Antonio, as a method to construct bivariate copulas with cubic sections, follows ([16], Theorem 2.4):

Let C be a function from $[0,1]^2$ into $\mathbb{I} := [0,1]$, given by

$$C(u,v) = uv + u(1-u)[\alpha(v)(1-u) + \beta(v)u],$$

where α and β are two functions from $\mathbb{I}$ to $\mathbb{R}$, satisfying $\alpha(0) = \alpha(1) = \beta(0) = \beta(1) = 0$. It is easy to see that C has cubic sections in u. It was shown in [16] that C is a copula if, and only if,

(i) $\alpha(v)$ and $\beta(v)$ are absolutely continuous, and
(ii) for almost every v in $\mathbb{I}$, either

$$-1 \le \alpha'(v) \le 2 \text{ and } -2 \le \beta'(v) \le 1$$

or

$$[\alpha'(v)]^2 - \alpha'(v)\beta'(v) + [\beta'(v)]^2 - 3\alpha'(v) + 3\beta'(v) \le 0.$$

Moreover, C is absolutely continuous.

In [16], we also studied dependence properties, measures of association, and concepts of symmetry for these copulas with cubic cross-sections. We explored examples of copulas with cubic sections, which extend some well-known families of bivariate copulas, such as the iterated Farlie-Gumbel-Morgenstern, Kimeldorf and Sampson, Lin, and Sarmanov families of copulas, and which provide second-order approximations to the Frank and Plackett families of copulas.

My third meeting with Roger B. Nelsen, this time at Lewis & Clark College, was fruitful and very enjoyable, and with a beautiful snow storm during my stay.

11.4 Fourth meeting: University of Massachusetts (Amherst, MA, USA)

In 1995, Berthold Schweizer had retired from the Department of Mathematics and Statistics of the University of Massachusetts in Amherst. He spent most of his academic career at UMass. He was advised by Karl Menger, and gained his Ph.D. degree at the Illinois Institute of Tecnhology in Chicago in 1956.

A good group of people, friends and colleagues, were invited to Amherst in November 1995, to celebrate the retirement of Professor Berthold Schweizer. We all met at UMass in Amherst, and held a symposium in honor of Bert Schweizer, and another colleague from the Department of Mathematics and Statistics, Haskell Cohen, who was also retiring. Carlo Sempi was there, as well as Jerry Frank, Abe Sklar, Robert Moynihan, János Aczél, Thomas Riedel, Howard Sherwood, Michael D. Taylor, Claudi Alsina,... and Roger B. Nelsen. It was there, during a break at that symposium, when some of us, after talking about copulas, suggested the idea of writing an introductory-level monograph on this subject to Roger, which would be the first book on copulas, his very well known "*An Introduction to Copulas*" [15]. Before Roger's book on copulas, Berthold Schweizer and Abe Sklar devoted Chapter 6 of their "*Probabilistic Metric Spaces*" book [31] to copulas, Bert Schweizer himself with his work in [30], and just as Harry Joe devoted Chapter 5 of his "*Multivariate Models and Dependence Concepts*" book [9]. After Roger's book on copulas, several books on the applications of the theory of copulas to different fields have appeared in recent years, for instance the books by Gianfausto Salvadori *et al.* [29] Umberto Cherubini *et al.* [3], Jan-Frederik Mai and Matthias Scherer [14], and the book "*Dependence Modeling with Copulas*" [10] by Harry Joe. Recently, Fabrizio Durante and Carlo Sempi have written a beautiful book entirely devoted to copulas, "*Principles of Copula Theory*" [4], which will be extremely useful for researchers in this field.

Once again, it was such a joy for me to be in Amherst with all those friends and colleagues, ten years after my stay there in 1985-86. The symposium was truly enjoyable, as was the celebration on the occasion of the retirement of Bert Schweizer.

This was my fourth meeting with Roger B. Nelsen, and, at that time, he was aware of our invitation for him to visit the Universities of Granada and Almería in Spain. Just a few years later, Roger visited our universities, and we commenced our engaging, continuous, and fruitful collaboration, that has never ended.

11.5 Fifth, sixth,... and many other meetings: Granada, Almería, Barcelona (Spain)

After my meeting with Roger in Amherst in November 1995, to celebrate the retirement of Professor Berthold Schweizer, he has visited Granada and Almería many times. We have thoroughly enjoyed working on different research problems in the theory of copulas and quasi-copulas all these years. We have shared wonderful moments of mathematics and friendship in Granada and Almería, working together with José Antonio Rodríguez-Lallena and Manuel Úbeda Flores, and also with Fabrizio Durante and Juan Fernández-Sánchez.

The research work of these years has been presented in several papers (see [6, 7, 17, 18, 19, 20, 21, 22, 23, 24, 25, 26]). In what follows, the main results of these papers will be described.

It is well known that the fundamental best-possible bounds inequality for bivariate distribution functions with given margins is the Fréchet-Hoeffding inequality [8]. Namely, if H denotes the joint distribution function of random variables X and Y whose margins are F and G, respectively, then

$$\max\{0, F(x)+G(y)-1\} \leq H(x,y) \leq \min\{F(x), G(y)\}$$

for all x, y in $[-\infty, \infty]$. In [17], we studied the problem of finding bounds on distribution functions with given margins when a value of the population version of a measure of association, such as Kendall's tau or Spearman's rho, is also given. Let $\tau(C)$ denote the Kendall's tau of a copula C, and let $\mathbf{T}_t$ denote the set of copulas with a common value $t \in [-1,1]$ of Kendall's tau, i.e.,

$$\mathbf{T}_t = \{C \mid C \in \mathscr{C}, \tau(C) = t\},$$

where $\mathscr{C}$ denotes the set of all copulas. Let $\underline{T}_t$ and $\overline{T}_t$ denote the pointwise infimum and supremum, respectively, of $\mathbf{T}_t$ for t in $[-1,1]$, i.e., for each (u,v) in $\mathbb{I}^2$,

$$\underline{T}_t(u,v) = \inf\{C(u,v) \mid C \in \mathbf{T}_t\} \quad \text{and} \quad \overline{T}_t(u,v) = \sup\{C(u,v) \mid C \in \mathbf{T}_t\}.$$

Then,

$$\underline{T}_t(u,v) = \max\left(0, u+v-1, \frac{1}{2}\left[(u+v) - \sqrt{(u-v)^2+1-t}\right]\right)$$

and

$$\underline{T}_t(u,v) = \min\left(u, v, \frac{1}{2}\left[(u+v-1) - \sqrt{(u+v-1)^2+1+t}\right]\right),$$

which are copulas themselves. Hence, if X and Y are continuous random variables with joint distribution function H and respective marginal distribution functions F and G, and such that $\tau(X,Y) = t$, then the best-possible bounds for H are

$$\underline{T}_t(F(x), G(y)) \leq H(x,y) \leq \overline{T}_t(F(x), G(y))$$

for all (x,y) in $[-\infty, \infty]^2$.

A similar result was obtained for the set $\mathbf{P}_t = \{C \mid C \in \mathscr{C}, \rho(C) = t\}$, where $\rho(C)$ denotes the Spearman's rho of the copula C ([17], Theorem 4).

Now, let X and Y be continuous random variables with distribution functions F and G, respectively, and let H_1 and H_2 be bivariate distribution functions whose univariate margins are F and G, respectively. In [18] (see also [19]), we showed that the distribution function of the random variable $H_1(X,Y)$, given that the joint distribution function of X and Y is H_2, depends only on the copulas C_1 and C_2 associated with H_1 and H_2, respectively. In particular, like the (one-dimensional) probability integral transform, it is independent of the marginal distribution functions F and G. However, unlike the probability integral transform, it is rarely uniformly distributed on $\mathbb{I}$.

In July 2000, a great group of people went to Barcelona (Spain) for the "Distributions with given marginal and statistical modelling" meeting. Among them, Claudi Alsina, Barry C. Arnold, Gregory A. Fredricks, Christian Genest, Harry Joe, Samuel Kotz, Ingram Olkin, Ludger Rüschendorf, Carmen Ruiz-Rivas, Allan R. Sampson, Carlo Sempi, Abe Sklar, Michael D. Taylor, Manuel Úbeda-Flores,... and Roger B. Nelsen. And, of course, the organizers and editors of the book of papers presented at the meeting: Carles M. Cuadras, Josep Fortiana and José Antonio Rodríguez-Lallena. It was a delightful meeting in the beautiful city of Barcelona, where we missed Berthold Schweizer, who could not attend.

We presented two papers at the meeting. One of them devoted to the study of a class of multivariate quasi-copulas, the Archimedean n-quasi-copulas, which have a wide range of applications. So, we studied basic properties of multivariate Archimedean quasi-copulas—that are also properties of Archimedean n-copulas—of particular interest for proper n-quasi-copulas. By the way, the notion of quasi-copula, that generalizes the concept of copula, was introduced by Claudi Alsina, Roger B. Nelsen and Berthold Schweizer in 1993 (see [1]), in order to show that a certain class of operations on univariate distribution functions is not derivable from corresponding operations on random variables defined on the same probability space. This concept was also used in [27], by Roger B. Nelsen, Berthold Schweizer, Carlo Sempi and I, to characterize, in a given class of operations on distribution functions, those that do derive from corresponding operations on random variables. A few years later, in 1999, the quasi-copula concept was characterized in simpler operational terms by Christian Genest, José Antonio Rodríguez-Lallena, Carlo Sempi and I (see [11]).

In the second paper at the meeting of Barcelona [21], we presented a new simple characterization and some new properties of quasi-copulas, all of them concerning the mass distribution of these functions. For instance, we showed that for any quasi-copula Q, and any rectangle $R = [u_1,u_2] \times [v_1,v_2]$ in $\mathbb{I}^2$, the $Q-$volume of R, given by

$$V_Q(R) = Q(u_2,v_2) - Q(u_2,v_1) - Q(u_1,v_2) + Q(u_1,v_1),$$

satisfies that $-1/3 \le V_Q(R) \le 1$. Moreover, $V_Q(R) = 1$ if, and only if, $R = \mathbb{I}^2$, and $V_Q(R) = -1/3$ implies that $R = [1/3,2/3]^2$. We also showed that any quasi-copula (and hence any copula) can be approximated arbitrarily closely by a quasi-copula with as much negative mass as desired. Namely, for any $\varepsilon > 0$, $M > 0$, and any quasi-copula $\overline{Q}$, there exists a quasi-copula Q and a set S in $\mathbb{I}^2$ such that:

(i) $\mu_Q(S) < -M$;
(ii) $\left|Q(u,v) - \overline{Q}(u,v)\right| < \varepsilon$ for all u, v in $\mathbb{I}$,

where $\mu_Q(S) = \sum_i V_Q(R_i)$, whenever $S = \cup_i R_i$.

In [22], we studied various properties of Kendall distribution functions and their consequences. If X and Y are continuous random variables with joint distribution function H, then the *Kendall distribution function* of (X,Y) is the distribution function of the random variable $H(X,Y)$. This function was called a "decomposition of Kendall's tau" by Christian Genest and Louis-Paul Rivest in [10], and by Philippe Capéraà, Anne-Laure Fougères and Christian Genest (see [2]), when defining a stochastic ordering based on this function. Furthermore, it appears in [18] as a bivariate probability integral transform. We used the Bertino family of copulas [9] to show that every distribution function satisfying the properties of a Kendall distribution function is the Kendall distribution function of some pair of random variables. We also examined the equivalence relation on the set of copulas induced by Kendall distribution functions, and, finally we studied empirical Kendall distribution functions and their relationships to the ordinary sample version of Kendall's tau. A few years later, we proved (see [25]) that any Kendall distribution function is the Kendall distribution function of some associative copula. Moreover, this fact allowed us to show that each equivalence class of the relation "to have the same Kendall distribution function as" contains a unique associative copula.

In [23], we used quasi-copulas to express the pointwise best-possible bounds on nonempty sets of distribution functions, copulas or quasi-copulas, and showed that every set $\mathscr{S}$ of (quasi-)copulas has the smallest upper bound and the greatest lower bound in the set of quasi-copulas (in the sense of pointwisely ordered functions). These bounds do not necessarily belong to the set $\mathscr{S}$, nor they are necessarily copulas whenever $\mathscr{S}$ does not contain any quasi-copula. We also presented an application to sets of copulas with some common property, such as a common diagonal section or common values at quartiles.

In [24], we studied a method, which we called a copula (or quasi-copula) *diagonal splice*, for creating new functions by joining portions of two copulas (or quasi-copulas) with a common diagonal section. The diagonal splice of two quasi-copulas is always a quasi-copula, and we found a necessary and sufficient condition for the diagonal splice of two copulas to be a copula. An alternative approach to our method was independently proposed by Fabrizio Durante, Anna Kolesárová, Radko Mesiar and Carlo Sempi in [5]. We showed that, in particular, such a method can be used to construct absolutely continuous asymmetric copulas with a prescribed diagonal section. In [21], we also introduced the concept of a simple diagonal and showed that many of the most commonly used copulas have simple diagonal sections. We also showed that an important subclass of such diagonals is the set of convex diagonals. We found an elementary way to construct asymmetric copulas with simple diagonal sections and, as an application, we obtained the best-possible upper bound for the set of copulas with a given simple diagonal. I remember that we particularly enjoyed working on this paper in Granada and Almería.

In [26], we studied the relationship between multivariate quasi-copulas and measures that they may or may not induce on $\mathbb{I}^n$. We studied the mass distribution of the pointwise best possible lower bound for the set of n-quasi-copulas for $n \geq 3$; and, as a consequence, we showed that not every n-quasi-copula induces a signed measure on $\mathbb{I}^n$. That was a quite technical research work that Roger Nelsen, José Antonio Rodríguez-Lallena, Manuel Úbeda-Flores and I completed in 2010.

In one of the recent meetings with Roger in Granada, we have investigated, together with Fabrizio Durante and Manuel Úbeda-Flores (see [6]), whether pairwise dependence properties related to all the bivariate margins of a trivariate copula imply the corresponding trivariate dependence property. What we have found is that, in general, information about the pairwise dependence is not sufficient to infer some aspects of global dependence. In other words, that stochastic dependence has so many facets, that they cannot be recovered from its lower-dimensional projections. For instance, pairwise independence does not imply mutual independence; positive quadrant dependence (PQD) of all bivariate margins does not imply positive lower orthant dependence (PLOD) or positive upper orthant dependence (PUOD) for the corresponding 3-copula; pairwise exchangeability does not imply mutual exchangeability.

Roger's most recent visit to Granada took place in May 2014. During that meeting with Roger, together with Juan Fernández-Sánchez and Manuel Úbeda-Flores, we were interested in the study of some properties of multivariate lower and upper tail dependence coefficients. As a result, in [7] we have presented those properties, and have used copulas to analyze the relationship between pairwise tail independence, mutual tail independence, and extremal independence.

11.6 Conclusions and next meetings

There is a popular saying that: "Time flies". I think most of us are in agreement on the truth in this. Looking back to 1986, when I met Roger B. Nelsen for the first time in Mount Holyoke College, I realize how fast these last thirty years have flown by. Now, we look forward to the coming years, with further meetings, and more good moments to share our friendship and mathematics.

Finally, I wish to express my sincere gratitude to Roger B. Nelsen, and to all my friends and colleagues, for their friendship, and for everything that they have taught me over the years.

Acknowledgements I thank Manuel Úbeda-Flores for inviting me to write this paper, for his comments and suggestions, which have improved the original manuscript, and for his assistance and help in the typesetting into Latex. I also thank Catherine L. Way for her valuable comments and suggestions, which have improved the readability of the paper.

This contribution has been supported by the Ministerio de Economía y Competitividad (Spain) under research project MTM2014-60594-P and partially supported by FEDER.

References

1. C. Alsina, R.B. Nelsen, B. Schweizer, On the characterization of a class of binary operations on distribution functions, Statist. Probab. Lett. 17, 85–89 (1993).
2. P. Capéraà, A.-L. Fougères and C. Genest. A stochastic ordering based on a decomposition of Kendall's tau, in *Distributions with Given Marginals and Moment Problems*, ed. by V. Beneš, J. Štěpán (Kluwer Academic Publishers, Dordrecht, 1997) pp. 81–92.
3. U. Cherubini, E. Luciano, W. Vecchiato, *Copula Methods in Finance* (Wiley Finance Series, John Wiley and Sons Ltd., Chichester, 2004).
4. F. Durante, C. Sempi, *Principles of Copula Theory* (Chapman & Hall/CRC, Boca Raton, FL, 2016).
5. F. Durante, A. Kolesárová, R. Mesiar, C. Sempi, Copulas with given diagonal sections: Novel constructions and applications, Int. J. Uncertainty Fuzziness Knowledge Based Syst. 15, 397–410 (2007).
6. F. Durante, R.B. Nelsen, J.J. Quesada-Molina, M. Úbeda-Flores, Pairwise and global dependence in trivariate copula models, in *Information Processing and Management of Uncertainty in Knowledge-Based Systems*, ed. by A. Laurent, O. Strauss, B. Bouchon-Meunier, R.R. Yager (Springer International Publishing Switzerland, Cham, 2014) pp. 243–251.
7. J. Fernández-Sánchez, R.B. Nelsen, J.J. Quesada-Molina, M. Úbeda-Flores, Independence results for multivariate tail dependence coefficients, Statist. Probab. Lett. 284, 129–137 (2016).
8. M. Fréchet, Sur les tableaux de corrélation dont les marges sont données, Ann. Univ. Lyon 9, Sect. A, 53–77 (1951).
9. G.A. Fredricks, R.B. Nelsen, The Bertino family of copulas, in *Distributions with Given Marginals and Statistical Modelling*, ed. by C.M. Cuadras, J. Fortiana, J.A. Rodríguez-Lallena (Kluwer Academic Publishers, Dordrecht, 2002) pp. 81–91.
10. C. Genest, L.-P. Rivest, Statistical inference procedures for bivariate Archimedean copulas. J. Amer. Statist. Assoc. 88, 1034–1043 (1993).
11. C. Genest, J.J. Quesada-Molina, J.A. Rodríguez-Lallena, C. Sempi, A characterization of quasi-copulas, J. Multivar. Anal. 69, 193–205 (1999).
12. H. Joe, *Multivariate Models and Dependence Concepts* (Monographs on Statistics and Applied Probability, vol. 73, Chapman & Hall, London, 1997).
13. H. Joe, *Dependence Modeling with Copulas*, (Chapman & Hall/CRC, London, 2014).
14. J.-F. Mai, M. Scherer, *Simulating Copulas: Stochastic Models, Sampling Algorithms and Applications*, (Imperial College Press, London, 2012).
15. R.B. Nelsen, *An Introduction to Copulas*, (Springer, New York, 1999), 2nd edn. (2006).
16. R.B. Nelsen, J.J. Quesada-Molina, J.A. Rodríguez-Lallena, Bivariate copulas with cubic sections, J. Nonparametr. Statist. 7, 205–220 (1997).
17. R.B. Nelsen, J.J. Quesada-Molina, J.A. Rodríguez-Lallena, M. Úbeda-Flores, Bounds on bivariate distribution functions with given margins and measures of association, Comm. Statist. Theor. Meth. 30, 1155–1162 (2001).
18. R.B. Nelsen, J.J. Quesada-Molina, J.A. Rodríguez-Lallena, M. Úbeda-Flores, Distribution functions of copulas: a class of bivariate probability integral transforms, Statist. Probab. Lett. 54, 277–282 (2001).
19. R.B. Nelsen, J.J. Quesada-Molina, J.A. Rodríguez-Lallena, M. Úbeda-Flores, Funciones de distribución univariantes de funciones de distribución bivariantes, in *Actas del Encuentro de Matemáticos Andaluces*, ed. by E. Briales Morales, Alfonso Carriazo Rubio, Tomás Chacón

Rebollo, Pedro Real Jurado, Alvaro Romero Jiménez (Universidad de Sevilla, Sevilla, 2001) pp. 681–690.
20. R.B. Nelsen, J.J. Quesada-Molina, J.A. Rodríguez-Lallena, M. Úbeda-Flores, Multivariate Archimedean quasi-copulas, in *Distributions with Given Marginals and Statistical Modelling*, ed. by C.M. Cuadras, J. Fortiana, J.A. Rodríguez-Lallena (Kluwer Academic Publishers, Dordrecht, 2002) pp. 179–186.
21. R.B. Nelsen, J.J. Quesada-Molina, J.A. Rodríguez-Lallena, M. Úbeda-Flores, Some new properties of quasi-copulas, in *Distributions with Given Marginals and Statistical Modelling*, ed. by C.M. Cuadras, J. Fortiana, J.A. Rodríguez-Lallena (Kluwer Academic Publishers, Dordrecht, 2002) pp. 187-194.
22. R.B. Nelsen, J.J. Quesada-Molina, J.A. Rodríguez-Lallena, M. Úbeda-Flores, Kendall distribution functions, Statist. Probab. Lett. 65, 263–268 (2003).
23. R.B. Nelsen, J.J. Quesada-Molina, J.A. Rodríguez-Lallena, M. Úbeda-Flores, Best-possible bounds on sets of bivariate distribution functions, J. Multivar. Anal. 90, 348-358 (2004).
24. R.B. Nelsen, J.J. Quesada-Molina, J.A. Rodríguez-Lallena, M. Úbeda-Flores, On the construction of copulas and quasi-copulas with given diagonal sections, Insur. Math. Econ. 42, 473–483 (2008).
25. R.B. Nelsen, J.J. Quesada-Molina, J.A. Rodríguez-Lallena, M. Úbeda-Flores, Kendall distribution functions and associative copulas Fuzzy Sets Syst. 160, 52–57 (2009).
26. R.B. Nelsen, J.J. Quesada-Molina, J.A. Rodríguez-Lallena, M. Úbeda-Flores, Quasi-copulas and signed measures, Fuzzy Sets Syst. 161, 2328–2336 (2010).
27. R.B. Nelsen, J.J. Quesada-Molina, B. Schweizer, C. Sempi, On the derivability of some operations on distribution functions, in *Distributions with Fixed Marginals and Related Topics*, ed. by L. Rüschendorf, B. Schweizer, M.D. Taylor (IMS Lecture Notes-Monograph Series, vol. 28, Hayward, CA, 1996) pp. 233–243.
28. J.J. Quesada-Molina, J.A. Rodríguez-Lallena, Bivariate copulas with quadratic sections, J. Nonparametr. Statist. 5, 323–337 (1995).
29. G. Salvadori, C. De Michele, N.T. Kottegoda, R. Rosso, *Extremes in nature. An approach using copulas* (Springer, Dordrecht, 2007).
30. B. Schweizer, Thirty years of copulas, in *Advances in Probability Distributions with Given Marginals*, ed. by G. DallAglio, S. Kotz, G. Salinetti (Kluwer Academic Publishers, Dordrecht, 1991) pp. 13–50.
31. B. Schweizer, A. Sklar, *Probabilistic Metric Spaces* (Elsevier, New York, 1983), reprint (Dover Publications, New York, 2005).
32. B. Schweizer, E.F. Wolff, On nonparametric measures of dependence for random variables, Ann. Statist. 9, 879–885 (1981).
33. A. Sklar, Fonctions de répartition à n dimensions et leurs marges, Publ. Inst. Statist. Univ. Paris 8, 229–231 (1959).

Chapter 12
Improved Hoeffding–Fréchet bounds and applications to VaR estimates

Ludger Rüschendorf

Abstract The classical Fréchet bounds determine upper and lower bounds for the distribution function F of a random vector X, when the marginal df's F_i are fixed. As consequence these bounds imply also upper and lower bounds for the expectation $E\varphi(X)$ of a certain class of functions $\varphi(X)$. The classical examples are the Hoeffding bounds for the expectation of the product EX_1X_2 of two random variables. In this paper we review and partially elaborate on several developments of improved Hoeffding–Fréchet bounds which assume some restriction on the dependence structure additional to the information on the marginals. We describe applications of the results to obtain improved VaR bounds for the joint portfolio of risk vectors. We consider in particular improved VaR bounds in the case where information of the joint distribution function resp. on the copula is available on some subsets and the case where higher order marginal information is available.

12.1 Hoeffding–Fréchet bounds

The classical Fréchet bounds are one of the most prominent results in stochastic ordering. They can be stated in the following form. For an n-dimensional df F holds: $F \in \mathscr{F}(F_1,\dots,F_n)$ – the Fréchet class of n-dimensional df's with marginals $F_1,\dots,F_n$ – if and only if

$$F_- \le F \le F_+, \tag{12.1}$$

where $F_+(x) := \min_{1\le i\le n} F_i(x_i)$ and $F_-(x) := \max\left\{0, \sum_{i=1}^n F_i(x_i) - (n-1)\right\}$ are the upper and lower Fréchet bounds. While F_+ is in general a df and thus $F_+ \in \mathscr{F}(F_1,\dots,F_n)$, it holds that $F_- \in \mathscr{F}(F_1,\dots,F_n)$ only for $n=2$ and for rare cases when $n \ge 3$. These cases were characterized in [Dall'Aglio (1972)]. F_+ is denoted

Ludger Rüschendorf
University of Freiburg, Eckerstraße 1, 79104 Freiburg, Germany
e-mail: ruschen@stochastik.uni-freiburg.de

M. Úbeda Flores et al. (eds.), *Copulas and Dependence Models with Applications*,
DOI 10.1007/978-3-319-64221-5_12

the comonotonic distribution, F_- is called in case $n = 2$ the antimonotonic distribution.

The inequalities in (12.1) imply some integral inequalities. Let for a real function $\varphi = \varphi(x_1, \dots, x_n)$ of n variables

$$\begin{aligned} M(\varphi) &:= \sup\Big\{ \int \varphi dP; P \in M(P_1, \dots, P_n) \Big\} \\ \text{and} \quad m(\varphi) &:= \inf\Big\{ \int \varphi dP; P \in M(P_1, \dots, P_n) \Big\} \end{aligned} \tag{12.2}$$

denote the generalized Hoeffding–Fréchet functionals, where P_i are probability measures on some measurable spaces $(E_i, \mathscr{A}_i)$. Since Hoeffding's paper from 1940 belongs to the earliest papers on the bounds in (12.1) we call bounds of this type in the following invariably Fréchet bounds or Hoeffding–Fréchet bounds. In particular this includes the case where $(E_i, \mathscr{A}_i) = (\mathbb{R}^1, \mathscr{B}^1)$ and P_i have distribution functions F_i.

A basic consequence of the Fréchet bounds in (12.1) is the following result on the supermodular ordering of distributions. Define for $P, Q \in M(P_1, \dots, P_n)$ – i. e. P and Q have marginals $P_1, \dots, P_n$ –

$$P \leq_{\text{sm}} Q \text{ if } \int f dP \leq \int f dQ \tag{12.3}$$

for all supermodular functions $f \in \mathscr{F}_{\text{sm}}$ such that the integrals exist. Let $\leq_{\text{uo}}$ denote the upper orthant ordering, i. e. $P \leq_{\text{uo}} Q$ if $P([a, \infty)) \leq Q([a, \infty))$ for all $a \in \mathbb{R}^n$. Then the following result holds

Theorem 12.1 (Supermodular order). *Let $P, Q \in M^1(\mathbb{R}^n, \mathscr{B}^n)$, then*

a) In case $n = 2$ it holds:

$$P \leq_{\text{sm}} Q \Leftrightarrow P \leq_{\text{uo}} Q \tag{12.4}$$

b) ***Lorentz Theorem****: For any $P \in M(P_1, \dots, P_n)$ holds*

$$P \leq_{\text{sm}} P_+, \tag{12.5}$$

where $P_+ \sim F_+$ is the comonotonic probability measure with marginals P_i.

The characterization of the supermodular ordering by the upper orthant order $\leq_{\text{uo}}$ in a) is due to [Cambanis et al. (1976)]. It generalizes in particular the classical Hoeffding bounds for the expectation of the product of two random variables. Let $X \sim F$, $Y \sim G$, and $U \sim U(0,1)$, then

$$EF^{-1}(U)G^{-1}(1-U) \leq EXY \leq EF^{-1}(U)G^{-1}(U). \tag{12.6}$$

Part b) was proved in [Tchen (1980)] by discrete approximation and reduction to the [Lorentz (1953)] inequalities. In Rueschendorf[1](1979,1983) the problem to

[1] In the further text Rüschendorf is abbreviated as Rü.

determine the generalized Hoeffding–Fréchet functional was identified with a rearrangement problem for functions. The Lorentz Theorem was reduced to the Lorentz inequality for functions.

There is also an analogue of the Fréchet bounds in (12.1) for the survival functions $\overline{F}_i(x_i) = P(X_i \ge x_i)$ and $\overline{F}(x_1,\dots,x_n) = P(X_i \ge x_i, 1 \le i \le n)$

$$\overline{F}^-(x) := \Big(\sum_{i=1}^n \overline{F}_i(x_i) - (n-1)\Big)_+ \le \overline{F}(x_1,\dots,x_n) \le \min_{1\le n} \overline{F}_i(x_i) := \overline{F}^+(x). \quad (12.7)$$

This version of the Fréchet bounds leads to an ordering result for the class of Δ-monotone functions (also called n-increasing) by means of a partial integration formula

Theorem 12.2 (Δ-monotone ordering). *Let $f : \mathbb{R}^n \to \mathbb{R}^1$ be Δ-monotone and assume that for any $1 \le i \le n$, $\lim_{x_i\to-\infty} f(x_1,\dots,x_i,\dots,x_n) = 0$, $\forall x_j$, $j \ne i$,*

a) For any $F \in \mathscr{F}(F_1,\dots,F_n)$ holds

$$\int \overline{F}^-(x)df(x) \le \int f dF \le \int \overline{F}^+(x)df(x). \quad (12.8)$$

b) If $F \in \mathscr{F}(F_1,\dots,F_n)$ and $\overline{G},\overline{H}$ are decreasing functions with

$$\overline{F}^- \le \overline{G} \le \overline{F} \le \overline{H} \le \overline{F}^+, \quad (12.9)$$

then

$$\int \overline{G}(x)df(x) \le \int f dF \le \int \overline{H}(x)df(x). \quad (12.10)$$

For part a) see Rü (2004, Theorem 5.5). Part b) is a direct consequence of the proof of Part a) and the assumptions in Part b).

Remark 12.1. a) *Convex ordering of risk portfolios.* Taking functions of the form $\varphi(x) = \Psi(\sum_{i=1}^n x_i)$, Ψ convex, we obtain that $\varphi \in \mathscr{F}_{\text{sm}}$ and the Lorentz Theorem implies that for any random vector $X = (X_1,\dots,X_n)$, $X_i \sim F_i$ it holds that

$$\sum_{i=1}^n X_i \le_{\text{cx}} \sum_{i=1}^n X_i^c, \quad (12.11)$$

where $X^c = (X_i^c)$ is a comonotonic vector with comonotonic distribution P_+; a result due to [Meilijson and Nadas (1979)]. In risk theory this result implies that X^c is the worst case dependence structure for the joint portfolio for any convex risk measure ρ. Further recent results on the convex ordering of risk portfolios as well as to VaR bounds of joint portfolios are described in [Puccetti and Wang (2015)].

b) *Parameter free price bounds under dependence constraints.* Part b) in Theorem 12.2 implies that positive or negative dependence constraints on the survival function $\overline{F}$ in terms of the upper orthant ordering $\le_{\text{uo}}$ imply directly improved parameter free bounds for the value of options defined by Δ-monotone functions.

Dependence restrictions as in (12.8) depend typically only on the copulas. They are closely related to modelling of dependence structures and to various specific constructions and bounds on copulas. A wealth of relevant material on these kind of theory and models is given in the by now classical book of [Nelsen (2006)] as well as in the recent book on copulas of [Durante and Sempi (2016)].

In the following Section 12.2 we discuss various forms of dual representations which are available to deal with the Hoeffding–Fréchet functional for general aggregation functions φ and also allow to include dependence information.

In Section 12.3 we consider the case where additional information on the dependence structure is available on some part of the domain. Finally in Section 12.4 we discuss additional information on the dependence structure by including second order marginal information. We discuss in particular applications to the problem of establishing VaR-bounds for the joint portfolio.

12.2 Dual representation of Hoeffding–Fréchet bounds

The most relevant and general information on the generalized Hoeffding–Fréchet functional is given by the dual representation of these functionals. The basic duality theorem states under some general conditions on φ equality of $M(\varphi)$ with a dual functional $U(\varphi)$. For detailed conditions see Rü (1991a,2007):
Duality Theorem:

$$M(\varphi) = U(\varphi) := \inf\Big\{\sum_{i=1}^n \int f_i dP_i; \sum_{i=1}^n f_i(x_i) \geq \varphi(x_1,\dots,x_n)\Big\}. \tag{12.12}$$

Similarly,

$$m(\varphi) = I(\varphi) := \sup\Big\{\sum_{i=1}^n \int f_i dP_i; \sum_{i=1}^n f_i \leq \varphi\Big\}. \tag{12.13}$$

Remark 12.2. Some history: The duality result was proved in Rü (1979,1981) and Gaffke and Rü (1981) including existence of solutions for the case where φ is bounded continuous. For the case of bounded measurable functions it was shown in these papers that replacing the σ-additive measures by finitely additive measures with marginals P_i and defining

$$\widetilde{M}(\varphi) := \sup\Big\{\int \varphi d\mu; \mu \in \mathrm{ba}(P_1,\dots,P_n)\Big\}$$

one gets

$$\widetilde{M}(\varphi) = U(\varphi). \tag{12.14}$$

This is a consequence of the Hahn–Banach separation theorem combined with Riesz representation theorem. Under suitable regularity on the spaces and on φ one obtains that $\widetilde{M}(\varphi) = M(\varphi)$. These duality results were then extended to some general classes of functions based on continuity properties of the functionals U, I and on the Choquet capacity theorem in [Kellerer (1984)].

For a survey of these developments, see Rü (1991b,2007) or Rachev and Rü (1998). It should be noted that in case $n = 2$ the duality result in (12.12) and (12.13) was the first instance of the Kantorovich duality theorem for mass transportation for *general* functionals φ. Kantorovich (1942) had established his duality result in the case where $\varphi(x_1, x_2)$ is a metric on a compact space.

As consequence of the duality theorem some basic inequalities and bounds were obtained, as for example the following result (see Rü (1981)).

Define for $A \in \mathscr{A}_1 \otimes \cdots \otimes \mathscr{A}_n$, $M(A) := M(\mathbb{1}_A)$, then:

Theorem 12.3 (Sharpness of Fréchet bounds). *For any* $A_i \in \mathscr{A}_i$, $1 \le i \le n$ *holds*

$$M(A_1 \times \cdots \times A_n) = \min\{P_i(A_i); 1 \le i \le n\} \tag{12.15}$$

$$\text{and} \quad m(A_1 \times \cdots \times A_n) = \min\left(\sum_{i=1}^n P_i(A_i) - (n-1)\right)_+ . \tag{12.16}$$

In particular (12.16) was the first proof of the sharpness of the lower Fréchet bounds in (12.1). An alternative proof was given in Sklar (1998) (see Theorem 2.10.13 in [Nelsen (2006)]).

A consequence of the duality theorem (or alternatively of Strassens's Theorem) is also a formula for the maximal and minimal value of the distribution function of the sum in case $n = 2$ due to Makarov (1981) and Rü (1982):

$$\begin{aligned} M(s) &:= \sup\{P(X_1 + X_2 \le x); X_1 \sim F_i\} \\ &= \inf_{x\in\mathbb{R}} (F_1(x) + F_2(t-x)) =: F_1 \wedge F_2(t) \end{aligned} \tag{12.17}$$

$$\begin{aligned} m(s) &:= \inf\{P(X_1 + X_2 < s); X_i \sim F_i\} \\ &= 1 - \sup_{x\in\mathbb{R}} (\overline{F}_1(x) + \overline{F}_2(t-x) := 1 - \overline{F}_1 \vee \overline{F}_2(t) \end{aligned} \tag{12.18}$$

This implies by inversion sharp bounds for the Value at Risk (VaR) in case $n = 2$.

Embrechts and Puccetti (2006a, 2006b) relaxed the dual representation by restricting to admissible piecewise linear functions of a simple form. This way they establish the following dual bounds:

$$M(s) \le D(s) = \inf_{u\in\overline{U}(s)} \min\left\{ \frac{\sum_{i=1}^n \int_{u_i}^{s-\sum_{j\neq i} u_j} \overline{F}_i(t)dt}{s - \sum_{i=1}^n u_i}, 1 \right\} \tag{12.19}$$

$$\text{and } m(s) \ge d(s) = \sup_{u\in\underline{U}(s)} \max\left\{ \frac{\sum_{i=1}^n \int_{u_i}^{s-\sum_{j\neq i} u_j} \overline{F}_i(t)dt}{s - \sum_{i=1}^n u_i} - d + 1, 0 \right\} \tag{12.20}$$

where $\overline{U}(s) = \left\{u \in \mathbb{R}^n; \sum_{i=1}^n u_i < s\right\}$ and $\underline{U}(s) = \left\{u \in \mathbb{R}^n; \sum_{i=1}^n u_1 > s\right\}$.

In the homogeneous case these dual bounds simplify strongly and can be shown under some mixing conditions to be sharp bounds (see Puccetti and Rü (2012)).

The method of proving the duality theorem described above is flexible enough to be able to handle also additional constraints. Some examples of additional constraints have been considered in Ramachandran and Rü (1997,1999,2002) who considered e. g. upper or lower bounds on the marginals, restrictions on the domain of the admissible measures or upper local bounds on the class of admissible measures. In a similar way also other types of constraints can be dealt with this approach. Consider for example an additional positive upper orthant dependence assumption and define

$$M_{\mathrm{PUOD}}(P_1,\dots,P_n) = \{P \in M(P_1,\dots,P_n) : P \text{ is positive upper orthant dependent}\}. \tag{12.21}$$

Here P is called positive upper orthant dependent (PUOD) if $P^\perp \le_{\mathrm{uo}} P$, where $P^\perp = \otimes_{i=1}^n P_i$.

Let $\mathscr{F}_\Delta$ denote the cone of Δ-monotone functions, then for $P \in M(P_1,\dots,P_n)$ holds:

$$P \text{ is PUOD iff } P^\perp \le_{\mathscr{F}_\Delta} P \text{ i. e. } \int f dP^\perp \le \int f dP \text{ for all } f \in \mathscr{F}_\Delta. \tag{12.22}$$

For a duality statement we consider as above the modified Hoeffding–Fréchet problem with finitely additive measures

$$\widetilde{M}_{\mathrm{PUOD}}(P_1,\dots,P_n) = \{\mu \in \mathrm{ba}(P_1,\dots,P_n); \mu \text{ is PUOD}\} \tag{12.23}$$

and for bounded measurable φ

$$\widetilde{M}_{\mathrm{PUOD}}(\varphi) = \sup\{\varphi d\mu; \mu \in \widetilde{M}_{\mathrm{PUOD}}(P_1,\dots,P_n)\}. \tag{12.24}$$

There are several possible classes of dual problems. Define

$$\mathscr{H}_1 = \Big\{((f_i),g); g \in \mathscr{F}_\Delta, f_i \in \mathscr{L}^1(P_i), \sum_{i=1}^n f_i - g \ge \varphi\Big\}.$$

Theorem 12.4 (Dual representation with PUOD constraints). *For any bounded measurable function φ holds:*

$$\widetilde{M}_{\mathrm{PUOD}}(\varphi) = I_1(\varphi) := \inf\Big\{\sum_{i=1}^n \int f_i dP_i - \int g d \otimes_{i=1}^n P_i; ((f_i),g) \in \mathscr{H}_1\Big\} \tag{12.25}$$

Proof. For any $\mu \in \widetilde{M}_{\mathrm{PUOD}}(P_1,\dots,P_n)$ and any $g \in \mathscr{F}_\Delta$ holds by (12.22)

$$\int g d\mu \ge \int g d \otimes_{i=1}^n P_i.$$

This implies for any $((f_i),g) \in \mathscr{H}_1$

$$\sum_{i=1}^n \int f_i dP_i - \int g d\otimes_{i=1}^n P_i \geq \sum_{i=1}^n \int f_i dP_i - \int g d\mu$$
$$= \int \Big(\sum_{i=1}^n f_i - g\Big) d\mu \geq \int \varphi d\mu.$$

As consequence we get

$$\widetilde{M}_{\text{PUOD}}(\varphi) \leq I_1(\varphi). \tag{12.26}$$

By Riesz representation theorem any continuous linear functional T on $B(\mathbb{R}^n, \mathscr{B}^n)$ with $Tf_i = \int f_i dP_i$ for $f_i = f_i(x_i) \in B(\mathbb{R}^1, \mathscr{B}^1)$ can be identified with an element $\widetilde{\mu} \in \widetilde{M}(P_1, \ldots, P_n)$. Further it holds that

$$\widetilde{\mu} \in \widetilde{M}_{\text{PUOD}}(P_1, \ldots, P_n) \Leftrightarrow \widetilde{\mu} \leq I_1.$$

Therefore, the Hahn–Banach separation theorem implies that for any $\varphi \in B(\mathbb{R}^n, \mathscr{B}^n)$

$$\widetilde{M}_{\text{PUOD}}(\varphi) = I_1(\varphi). \qquad \square$$

Remark 12.3.

a) *Existence and extensions:* The separation theorem also implies the existence of a solution $\widetilde{\mu} \in \text{ba}(P_1, \ldots, P_n)$. Restricting to the class of continuous bounded functions C_b on $\mathbb{R}^n$ we obtain by Riesz representation theorem

$$M_{\text{PUOD}}(\varphi) = \widetilde{M}_{\text{PUOD}}(\varphi) = I_1(\varphi), \;\; \varphi \in C_b. \tag{12.27}$$

This duality result can be further extended to more general classes of functions by suitable continuity properties of the functionals M_{PUOD} and I_1 as in the simple marginal case.

b) *Reduced dual problem:* The duality statements in (12.25) and (12.27) give an exact upper bound for $\int \varphi dP$, $P \in M_{\text{PUOD}}(P_1, \ldots, P_n)$ which however as in the simple marginal case in general is not easy to evaluate. For the dual functional we can restrict to a more simple generator of $\mathscr{V}_\Delta$ by restricting to g of the form

$$g(x) = \sum_{i=1}^m \alpha_i 1_{[a_i, \infty)}(x) \quad \text{with } \alpha_i \geq 0.$$

Then the dual problem reduces to optimize

$$\sum \int f_i dP_i - \sum_{i=1}^m \alpha_i \overline{F^\perp}(a_i) \tag{12.28}$$

over all admissible duals of this form, where $\overline{F^\perp}(x) = \prod_{i=1}^n P_i([x, \infty))$ is the survival function. In particular, any admissible dual choice yields by (12.25) and (12.26) to an upper bound.

As second example we consider the case where additional to the marginals also the covariances $\sigma_{ij} = \text{Cov}(X_i, X_j) = EX_iX_j - a_ia_j$, $a_i = EX_i$ are specified. In a similar way as above one gets the dual representation for the class $M_\Sigma(P_1,\dots,P_n)$ of measures P with marginals P_i and correlation matrix $\Sigma = (\sigma_{ij})$.

Theorem 12.5 (Fixed correlations). *For any $\varphi \in B(\mathbb{R}^n, \mathscr{B}^n)$ holds*

$$\widetilde{M}_\Sigma(\varphi) = I_2(\varphi) := \inf\Big\{ \sum_{i=1}^n \int f_i dP_i + \sum_{(i,j)} \alpha_{ij} s_{ij}; \varphi \le \sum_{i=1}^n f_i(x_i) + \Sigma \alpha_{ij} x_i x_j \Big\} \quad (12.29)$$

Remark 12.4.

a) Similarly as above for $\varphi \in C_b$, $\widetilde{M}_\Sigma(\varphi) = M_\Sigma(\varphi)$ and the duality can be extended to more general classes of functions φ.

b) If we consider as in [Bernard et al. (2015)] $\varphi = \mathbb{1}_{\{\sum_{i=1}^n x_i \ge t\}}$ and assume that it is known that additional to the marginals P_i, also it is known that $\mathrm{VaR}(S_n) \le \sigma^2$, then the dual in (12.29) simplifies strongly to the form

$$\begin{aligned} I_{2,\sigma^2}(\varphi) = \inf\Big\{ &\sum_{i=1}^n \int f_i dP_i + \alpha(\sigma^2 - \mu^2); \\ &\varphi(x) \le \sum_{i=1}^n f_i(x_i) + \alpha\Big[\Big(\sum_{i=1}^n x_i\Big)^2 - \mu^2\Big], \alpha \ge 0, f_i \in L^1(P_i)\Big\} \end{aligned} \quad (12.30)$$

In [Bernard et al. (2015)] good upper bounds for this case were given. In contrast formula (12.30) gives theoretically sharp upper bounds.

c) Model independent price bounds: In a similar way the above sketched method also applies to various other types of constraints. For robust model independent price bounds in recent years dual representations with martingale constraints have been developed (see e. g. [Acciaio et al. (2013)] and [Beiglböck et al. (2013)]). This kind of constraints is due to the fact, that reasonable pricing measures have the martingale property. Also this type of constraints can be dealt with by the above described method.

12.3 Improved Hoeffding–Fréchet bounds – distributional information on domains

Motivated by the problem to determine good bounds for the Value at Risk (VaR) of the joint portfolio there has been a lot of recent papers to improve the Fréchet bounds in (12.1) by including additional dependence information and as consequence to obtain improved bounds for the tail risk $P(\sum_{i=1}^n X_i \ge t)$ or on $\mathrm{VaR}_\alpha(\sum_{i=1}^n X_i)$.

For a random vector X with $X_i \sim F_i$ and distribution function $F_X = F$ we can pose positive dependence restrictions in the form that $F \le G$ or that the survival function $\overline{F} \ge \overline{G}$ for some increasing (resp. decreasing) function G ($\overline{G}$), where G and $\overline{G}$ are bounded above and below by the Fréchet bounds in (12.1) resp. (12.7). Defining $\mathscr{A}(s) := \{u = (u_1,\dots,u_n) \in \mathbb{R}^n : \sum_{i=1}^n u_i = s\}$ and $\bigwedge G(s) =$

$\inf_{u\in\mathscr{A}(s)} G(u)$, $\bigvee\overline{G}(s) = \sup_{u\in\mathscr{A}(s)} \overline{G}(u)$, the following improved standard bounds on the joint portfolio have been given in several similar forms in the literature, see [Williamson and Downs(1990), Embrechts et al. (2003)], Rü (2005,2013), Embrechts and Puccetti (2006b,2010), and Puccetti and Rü (2012).

Theorem 12.6 (Improved standard risk bounds under positive dependence restriction). *Let X,F,G, and $\overline{G}$ be as introduced above.*

a) If $G \le F$, then

$$P\Big(\sum_{i=1}^n X_i \le s\Big) \ge \bigvee G(s) \tag{12.31}$$

b) If $\overline{G} \le \overline{F}$, then

$$P\Big(\sum_{i=1}^n X_i < s\Big) \le 1 - \bigvee \overline{G}(s). \tag{12.32}$$

In the case where $G = F_-$ resp. $\overline{G} = \overline{F}_-$ include no further dependence information these bounds are called standard bounds. By inversion we obtain the standard bounds for VaR which depend only on the lower Fréchet copula bound $W(u) = (\sum_{i=1}^n u_i - (n-1))_+$. We denote the corresponding standard VaR bound by VaR^W.

In particular (12.31) and (12.32) give upper and lower bounds for the distribution function and thus also for the tail risk of the sum if the risk vector X is positive quadrant dependent (i. e. PUOD and PLOD).

To establish bounds for the df as in (12.31) or in (12.32) it is of course sufficient to have bounds for the copula $C = C_X$. An elaboration on the method induced by Theorem 12.6 to VaR bounds has been given in [Embrechts et al. (2013), Embrechts et al. (2014)]. Also several alternative ways to include dependence information in order to obtain improved VaR bounds for the sum have been discussed in the recent literature. For example, Bernard et al. (2016a,2016b) derive improved risk bounds based on additional variance or moment information. Positive and negative dependence restrictions as in (12.31) and (12.32) based on independence and positive dependence information in subgroups were considered in [Bignozzi et al. (2015)] and [Puccetti et al. (2015)]. Structural information by partially specified risk factor models was investigated in [Bernard et al. (2016b)]. A survey of these developments is given in Rü (2016).

Assuming that for a distribution function F with marginals F_i it is known that $F \le G$ and/or that $F \ge G$ on some subset $S \subset \mathbb{R}^n$ one obtains the following improved Hoeffding–Fréchet bounds which were given independently in [Puccetti et al. (2016)] and in [Lux and Papapantoleon (2016)].

Theorem 12.7 (Improved Hoeffding–Fréchet bounds). *Let $G : \mathbb{R}^n \to \mathbb{R}$ be an increasing function with $F_- \le G \le F_+$ and define*

$$F^*(x) = \min\Big(\min_{1\le i\le n} F_i(x_i), \inf_{y\in S}\Big\{G(y) + \sum_{i=1}^n (F_i(x_i) - F_i(y_i))_+\Big\}\Big)$$

$$F_*(x) = \max\Big(0, \sum_{i=1}^n F_i(x_i) - (n-1), \sup_{y\in S}\Big\{G(y) - \sum_{i=1}^n (F_i(y_i) - F_i(x_i))_+\Big\}\Big)$$

Then for $F \in \mathscr{F}(F_1,\dots,F_n)$ holds

i) If $F(y) \le G(y)$ for all $y \in S$, then $F(x) \le F^*(x)$ for all $x \in \mathbb{R}^n$.
ii) If $F(y) \ge G(y)$ for all $y \in S$, then $F(x) \ge F_*(x)$ for all $x \in \mathbb{R}^n$.
iii) If $F(y) = G(y)$ for all $y \in S$, then $F_*(x) \le F(x) \le F^*(x)$ for all $x \in \mathbb{R}^n$.

Remark 12.5. In the case $n = 2$ the improved Hoeffding–Fréchet bounds in Theorem 12.7 are due to Rachev and Rü (1994). They were restated in the case of uniform marginals i. e. for copulas in [Tankov (2011)] for the case of equality constraints. In this paper also a sharpness result for increasing sets S and an application to model free pricing bounds for multi-asset options is given. In the case that S is singleton and $n = 2$ these bounds and their sharpness were shown in [Nelsen et al. (2004), Theorem 3.2.2].

In [Nelsen (2006)] several constructions are given for copulas with given sections. An interesting construction are f. e. the diagonal copulas with prescribed diagonal section. They are however different from the upper or lower bounds F^* resp. F_* as specified in Theorem 12.7. For $n \ge 2$ and S being a singleton the improved bounds were given in [Rodríguez-Lallena and Úbeda-Flores (2004)] and in [Sadooghi-Alvandi et al. (2013)] for finite sets S.

Extensions of the sharpness result are given in [Bernard et al. (2012)]. The paper of [Bernard et al. (2013)] discusses as application the case where S is the central part of the distribution.

As corollary in the case that $F_i \sim U(0,1)$, $1 \le i \le n$ Theorem 12.7 implies the following improved bounds for the copula of a risk vector.

Corollary 12.8 (Improved copula bounds)**.** *Let $S \subset [0,1]^n$ and let Q be a componentwise increasing function on $[0,1]^n$ such that $W(u) \le Q(u) \le M(u)$, $u \in [0,1]^n$. Define the bounds $A^{S,Q}, B^{S,Q} : [0,1]^n \to [0,1]$ as*

$$A^{S,Q}(u) = \min\Big(M(u), \inf_{a\in S}\Big\{Q(a) + \sum_{i=1}^n (u_i - a_i)_+\Big\}\Big)$$

$$B^{S,Q}(u) = \max\Big(W(u), \sup_{a\in S}\Big\{Q(a) - \sum_{i=1}^n (a_i - u_i)_+\Big\}\Big)$$

Then for an n-dimensional copula C, it holds

i) If $C(u) \le Q(u)$ for all $u \in S$, then $C(u) \le A^{S,Q}(u)$ for all $u \in [0,1]^n$.
ii) If $C(u) \ge Q(u)$ for all $u \in S$, then $C(u) \ge B^{S,Q}(u)$ for all $u \in [0,1]^n$.
iii) If $C(u) = Q(u)$ for all $u \in S$, then $B^{S,Q}(u) \le C(u) \le A^{S,Q}(u)$ for all $u \in [0,1]^n$.

It is shown in [Puccetti et al. (2016)] in several relevant examples that the improved Hoeffding–Fréchet bounds in Theorem 12.7 may lead to strongly improved VaR bounds for the joint portfolio based on the method of improved standard risk bounds in Theorem 12.6.

Some examples showing the effect of the improved Hoeffding–Fréchet bounds are discussed in [Puccetti et al. (2015)]. The following example shows related results in a graphical way.

Example 12.1.

a) *Positive dependence in the tails.* We consider the case $n = 2$ with $F_1 = F_2 = Pareto(2)$ and assume that the copula Q of the risk vector is comonotonic on the tail area $S = [0.9, 1]^2$, i. e. for a copula vector $(U_1, U_2) \sim Q$ holds

$$P(U_1 \geq u_1, U_2 \geq u_2) = \min(1 - u_1, 1 - u_2), u_i \geq 0.9.$$

This models a case where in extreme situations a strong form of positive dependence arises. As consequence of this strong positive dependence in the tails we obtain from Corollary 12.8 and Theorem 12.6 a remarkable reduction of the improved VaR bounds $\mathrm{VaR}_\alpha^{B^{S,Q}}$ for moderate and in particular for high quantile levels α (see Figure 12.1). In fact in this example the standard bounds are know to be sharp bounds.

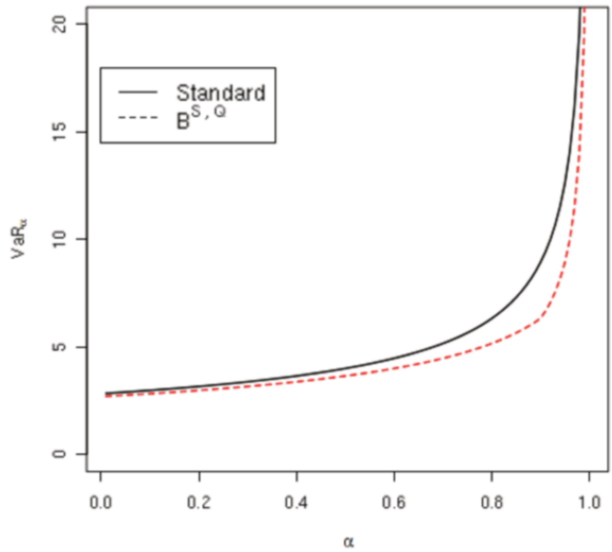


Fig. 12.1: Comparison of $\mathrm{VaR}_\alpha^{B^{S,Q}}$ and the standard bound VaR_α^{S} for $n = 2$ with $Pareto(2)$ marginals.

Based on Corollary 12.8 a similar effect also holds in the case that $n \geq 2$. The assumption of comonotonicity in the tails is a strong assumption.
In [Puccetti et al. (2016)] it is shown that some related effects are obtained when replacing the strong positive dependence assumption in the tail in Corollary 12.8 by a weaker assumption of the form $\overline{Q}(u) = 1 - Q(u) \geq G_\vartheta(u)$, $u \in S$, where G_ϑ is a parametric class of Gumbel or of Gaussian copulas.

b) *Independent subgroups with positive internal dependence.* In this example we modify the model assumption investigated in [Bignozzi et al. (2015)]. We consider the case that the risks are split into k independent subgroups I_j.
[Bignozzi et al. (2015)] allow any kind of dependence within these subgroups. In comparison we assume that the risks within the subgroups are strongly positive dependent (comonotonic) in the tails, i. e., similar as in Example 12.1 a) on $[0.9,1]^{n_i}$, where $n_i = |I_i|$.
As concrete example we consider the case where $n = 20$, with $k = 1, 10, 20$ subgroups, where the subgroup sizes are equal to $\frac{20}{k}$. We further assume that $F_i = Pareto(2) = F$, $1 \leq i \leq n$. As consequence of Theorem 12.6 and Corollary 12.8 we obtain

$$P\Big(\sum_{i=1}^n X_i \leq s\Big) \geq B_k^{S,Q}\Big(F\Big(\frac{s}{n}\Big),\dots,F\Big(\frac{s}{n}\Big)\Big)$$
$$= \max\Big(nF\Big(\frac{s}{n}\Big) - (n-1), \max_{a\in S}\Big\{Q(a) - \sum_{i=1}^n \Big(a_i - F\Big(\frac{s}{n}\Big)\Big)_+\Big\}\Big),$$

where $S = [0.9,1]^n$ and $Q(a) := \prod_{j=1}^k \min_{i\in I_j} a_i$. The corresponding VaR bounds $\mathrm{VaR}_\alpha^{B_k^{S,Q}}$ are obtained by inversion and are given in Figure 12.2. In that paper also the improved standard bounds are compared with the (sharp) bounds with marginal information only. For strong enough positive dependence the improved standard bounds are sharper than the dual bounds.
The results obtained can be expected. The worst bound is the standard bound. The best bound is obtained for the case $k = 1$ of general comonotonicity in the tails. The case of 10 independent subgroups with positive tail dependence leads to a considerable reduction.
As in Example 12.1 a) [Puccetti et al. (2016)] describe similar effects in this example when replacing the comonotonicity assumption inside the groups by weaker Gumbel type specification in the tails.

12.4 Higher order marginal information; comparison of various VaR bounds for the joint portfolio

If higher order marginal distributions of the risk vector X are known then it is possible to improve the Hoeffding–Fréchet bounds and as consequence of (12.31), (12.32), (12.19), and (12.20) one gets improved standard bounds for the VaR. In

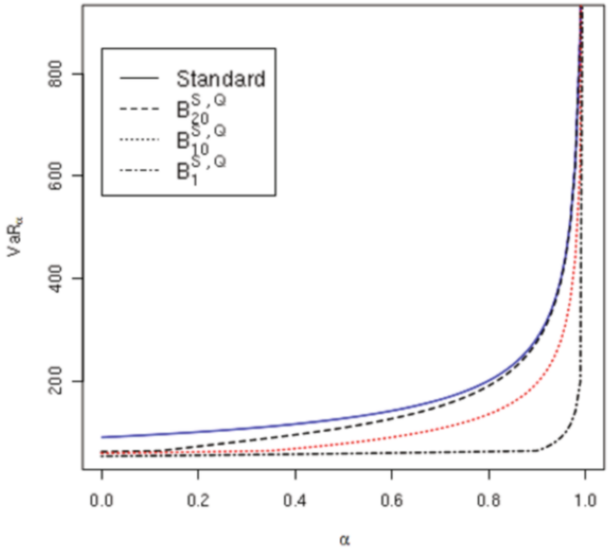


Fig. 12.2: Comparison of $\mathrm{VaR}_\alpha^{B_k^{S,Q}}$ for $k = 1, 10, 20$ and standard bound VaR_α^S, $n = 20$, $F_i = Pareto(2)$.

this section we consider the case where two dimensional marginal distributions are known. Alternative dual bounds with higher order marginals called 'reduced bounds' have been discussed in [Embrechts and Puccetti (2006a)], Puccetti and Rü (2012), and in [Embrechts et al. (2013)]. As a result it was found in these papers that the additional information of higher dimensional marginals may lead to considerably improved upper VaR bounds, when the joint marginals are not 'too close' to the upper Hoeffding–Fréchet bounds.

One obtains improved Hoeffding–Fréchet bounds for the distribution function (resp. for the copula) by means of Bonferroni-type bounds (see Prop. 6 in Rü (1991a)).

Proposition 12.1 (Bonferroni-type bounds). *Let C be an n-dimensional copula with bivariate marginals $C_{i,j}$ for $i \neq j$. Then*

$$\min_{i \neq j} C_{ij} \geq C \geq W_B \geq W_A \geq W, \tag{12.33}$$

where $W(u) = (\sum_{i=1}^n u_i - (n-1))_+$ is the Hoeffding–Fréchet lower bound,

$$W_A(u) = \Big(\sum_{i=1}^n u_i - (n-1) + \frac{2}{n}\sum_{i<j}(1-u_i-u_j+C_{i,j}(u_i,u_j))\Big)_+ \qquad (12.34)$$

$$\textit{and} \quad W_B(u) = \Big(\sum_{i=1}^n u_i - (n-1) + \sup_\tau \sum_{(i,j)\in\tau}(1-u_i-u_j+C_{i,j}(u_i,u_j))\Big)_+, \qquad (12.35)$$

the sup being taken over all spanning trees of the complete graph induced by $\{1,\dots,n\}$.

The bound W_B is a consequence of the Bonferroni inequality from [Hunter (1976)] (see Prop. 6 in Rü (1991a)). It improves the bound W_A arising from a Bonferroni bound of [Hunter (1976)] and [Worsley(1982)]. As consequence of (12.31) and (12.32) these bounds imply improved bounds for the tail-risk and the VaR of the joint portfolio $\sum_{i=1}^n X_i$, where (X_i,X_j) have copulas $C_{i,j}$. Let

$$\begin{aligned} &\mathrm{VaR}_\alpha^W = W(F_1,\dots,F_n)^{-1}(\alpha), \quad \mathrm{VaR}_\alpha^{W_A} = W_A(F_1,\dots,F_n)^{-1}(\alpha) \\ \text{and} \quad &\mathrm{VaR}_\alpha^{W_B} = W_B(F_1,\dots,F_n)^{-1}(\alpha) \end{aligned} \qquad (12.36)$$

denote the upper α-quantiles of W, W_A, W_B with marginals $F_1,\dots,F_n$. Then we obtain as consequence of (12.33)

$$\mathrm{VaR}_\alpha(S) \le \mathrm{VaR}_\alpha^{W_B} \le \mathrm{VaR}_\alpha^{W_A} \le \mathrm{VaR}_\alpha^W. \qquad (12.37)$$

The upper bound $\mathrm{VaR}_\alpha^{W_A}$ has been investigated in [Liu and Chan (2011)]. In contrast to their statement this bound is not the 'best possible upper bound' for $\mathrm{VaR}_\alpha(S)$. As their numerical results indicate the bound $\mathrm{VaR}_\alpha^{W_A}$ improves on the dual bound, which is based solely on marginal information, only for high confidence levels α and for highly positive correlated two-dimensional marginals. Correspondingly it was seen in [Embrechts et al. (2013)] that strong improvements of lower bounds are obtained, when the two-dimensional marginals are independent.

In the following examples we compare the Bonferroni bounds $\mathrm{VaR}_\alpha^{W_A}$ and $\mathrm{VaR}_\alpha^{W_B}$ with each other and with the standard bounds VaR_α^W as well as with the dual bound VaR_α^D arising from (12.19) for various dependence levels on the bivariate marginals.

By (12.31) we have

$$P\Big(\sum_{i=1}^n X_i \le t\Big) \ge \sup_{u\in\mathscr{U}(t)} C_L(F_1(u_1),\dots,F_n(u_n)), \qquad (12.38)$$

where C_L is either W or is one of the (improved) bounds W_A, W_B. For $u = \left(\frac{t}{n},\dots,\frac{t}{n}\right)$ we get the lower bound

$$P\Big(\sum_{i=1}^n X_i \le t\Big) \ge C_L\Big(F_1\Big(\frac{t}{n}\Big),\dots,F_n\Big(\frac{t}{n}\Big)\Big). \qquad (12.39)$$

In general the improvements of the Fréchet bounds as in (12.33) can be considerable. The improved standard bounds in (12.38) are not easy to determine in general in explicit form. In several cases however conditions are easy to state which allow to determine them explicitly. In general we obtain the strongest improvement of the upper bound VaR_α^W if the two-dimensional copulas are comonotonic.

We next state for some cases explicit solutions to (12.38). If $C_L = W$ and $F_1,\dots,F_n$ have decreasing densities and $u^* \in U(t)$ satisfies $F_1(u_1^*) = \cdots = F_n(u_n^*)$ then $u^* = (u_1^*,\dots,u_n^*)$ is uniquely determined and u^* is a solution to (12.38). If $F_1 = \cdots = F_n$ has a decreasing density, then $\left(\frac{t}{n},\dots,\frac{t}{n}\right)$ is a solution to (12.38) and thus the bound in (12.39) coincides with that in (12.38).

More generally let

$$A = \{(F_1(u_1),\dots,F_n(u_n));\ \ u = (u_i) \in \mathscr{U}(t)\}$$

and assume that $(F_i(u_i^*))$ is a smallest element of A w.r.t. the increasing Schur convex order $\preceq_S$, then

$$\sup_{u\in\mathscr{U}(t)} W(F_1(u_1),\dots,F_n(u_n)) = W(F_1(u_1^*),\dots,F_n(u_n^*)). \tag{12.40}$$

Similarly, assuming that W_A resp W_B are Schur concave, i. e. decreasing w.r.t. the increasing Schur convex order $\preceq_S$ we obtain

$$\sup_{u\in\mathscr{U}(t)} W_A(F_1(u_1),\dots,F_n(u_n)) = W_A(F_1(u_1^*),\dots,F_n(u_n^*)) \tag{12.41}$$

$$\text{resp.}\quad \sup_{u\in\mathscr{U}(t)} W_B(F_1(u_1),\dots,F_n(u_n)) = W_B(F_1(u_1^*),\dots,F_n(u_n^*)). \tag{12.42}$$

Sufficient conditions for Schur concavity of W_A and W_B can be inferred from Chapters 3 and 4 in [Marshall and Olkin (1979)]. For example, in the homogeneous case $C_{ij} = C_2$ for all i, j, if C_2 is concave and symmetric or more generally is Schur concave, then W_A and W_B are Schur concave.

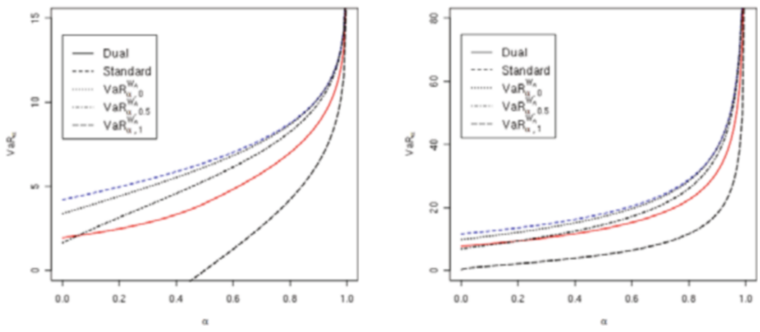
In the following we use the vector u^* with identical components $(F_1(u_1^*),\dots,F_n(u_n^*))$ as above as a proxy for comparison of the upper bounds in (12.40)–(12.42). In particular in the case $F_1 = \cdots = F_n = F$ we use the vector $\left(F\left(\frac{t}{n}\right),\dots,F\left(\frac{t}{n}\right)\right)$. In contrast to statements in [Liu and Chan (2011)] this choice will not give the exact bounds in (12.40) and (12.41) (and also in (12.42)) in general.

In the following examples we consider the homogeneous case where $F_i = F$ and where $C_{i,j} = C_2$ for all $i < j$. We concentrate on the approximate bounds based on u^*.

Comparison of VaR^{W_A}, standard bounds, and dual bound

In the first example we compare the standard bound, i. e. the VaR bound induced by W, the VaR bound induced by W_A and the dual bound D, which gives the optimal bound with only marginal information in this example.

Let $n = 5$ and let X_i be standard normal resp. log-normal distributed, $1 \le i \le 5$. Let C_2 be a Gauß-copula with correlations $\rho = 0,\ 0.5,\ 1$. Figure 12.3 compares the $\mathrm{VaR}_{\alpha,\rho}^{W_A}$ upper bounds with the dual bound VaR_α^D in dependence on α and ρ for both distributions. Note that using the proxies the bounds $\mathrm{VaR}^{W_A}_{\alpha,\rho}$ and $\mathrm{VaR}^{W_B}_{\alpha,\rho}$ coincide in this case.



((a)) ((b))

Fig. 12.3: Comparison $\mathrm{VaR}^{W_A}_{\alpha,\rho}$, standard bound and dual bound, $n = 5$, $\rho = 0,\ 0.5,\ 1$, in case of Gaußcopula in (a) and log normal copula in (b).

Figure 12.3 (a) shows that the $\mathrm{VaR}^{W_A}_{\alpha,\rho}$ bound improves with increasing correlation. In particular the case $\rho = 1$ (comonotonicity) for the two dimensional marginals gives better upper bounds than the case $\rho = 0$ (independence). This kind of dependence on ρ can also be seen directly from the definition of W_A in (12.34). Further one finds as expected, that for any ρ the $\mathrm{VaR}^{W_A}_{\alpha,\rho}$ bound using information on two-dimensional marginals is an improvement on the standard bound based on marginal information only.

The dual bound $\mathrm{VaR}^{D}_{\alpha}$ is a strong improvement over the standard bound, both being based on marginal information only. It is known that the dual bound is optimal in this example. This example shows that the technique of standard bounds does not work well in higher dimensions.

From Figure 12.3 and Table 12.1 one sees that the dual bound $\mathrm{VaR}^{D}_{\alpha}$ is even an improvement over the bounds $\mathrm{VaR}^{W_A}_{\alpha,\rho}$ when $\rho < 0.9$ and $\alpha \geq 0.9$, i.e. the information on two-dimensional marginal information does not lead to an improved upper bound in these cases, when using the method of improved standard bounds.

α	VaR^S_α	VaR^D_α	$\mathrm{VaR}^{W_A}_{\alpha,0}$	$\mathrm{VaR}^{W_A}_{\alpha,0.5}$	$\mathrm{VaR}^{W_A}_{\alpha,0.9}$	$\mathrm{VaR}^{W_A}_{\alpha,1}$
0.9	10.268	8.773	10.234	9.943	8.764	6.407
0.95	11.631	10.311	11.616	11.415	10.425	8.224
0.99	14.390	13.322	14.388	14.297	13.589	11.631

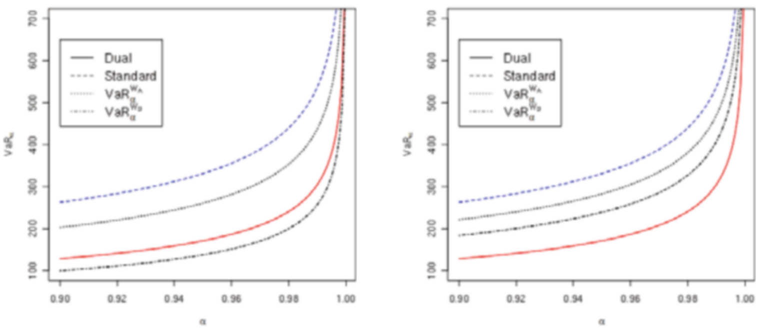
Table 12.1: Comparison of VaR^S, VaR^D and VaR^{W_A}

In Figure 12.3 (b) we see that in the case of log-normal distributions with heavy tails we obtain a similar picture of the relation between these VaR bounds.

While in this example the bounds VaR^{W_A} and VaR^{W_B} coincide when using the proxies, in the following example we show that in inhomogeneous cases the difference can be quite big so that VaR^{W_B} is a strong improvement over VaR^{W_A}.

Comparison of VaR^{W_A} and VaR^{W_B}
We consider the case $n = 20$ where the marginals X_i are log-normal distributed. We assume that $C_{i,j}(u_i, u_j)$ is a t-copula with three degrees of freedom and correlation ρ. The risks X_i are divided into two groups of equal size 10. Within the groups the rv's are pairwise comonotone, i. e. $\rho = \rho_1 = 1$ and between the groups the rv's are pairwise independent, i. e. $\rho = \rho_2 = 0$.

In this case the sup in (12.35) is attained by the tree which uses only once the correlation $\rho_2 = 0$. On the other hand $\mathrm{VaR}_\alpha^{W_A}$ can be seen as an average over all starwise trees which also contains trees which use several times the low correlation connections with $\rho_2 = 0$. This construction makes the difference between both bounds in a particular way big. We find in Figure 12.4 (a) that in this case $\mathrm{VaR}_\alpha^{W_B}$ is strongly improved compared to the VaR bound $\mathrm{VaR}_\alpha^{W_A}$. For example we obtain $\mathrm{VaR}_{0.9}^{W_B} = 99.5875$ which is about 50 % better than $\mathrm{VaR}_{0.9}^{W_A} = 202.6817$. The difference between the bounds is increasing in α. For $\alpha = 0.99$ we have for example $\mathrm{VaR}_{0.99}^{W_B} = 257.1075$ an improvement of 59 % over $\mathrm{VaR}_{0.99}^{W_A} = 437.2221$. $\mathrm{VaR}_\alpha^{W_B}$ improves over the dual bound VaR_α^D whereas $\mathrm{VaR}_\alpha^{W_A}$ is worse than the dual bound.



((a)) $\rho_1 = 1$, $\rho_2 = 0$ ((b)) $\rho_1 = 0.9$, $\rho_2 = 0.1$

Fig. 12.4: Comparison of $\mathrm{VaR}_\alpha^{W_A}$, $\mathrm{VaR}_\alpha^{W_B}$, VaR_α^D, inhomogeneous case, $C_{i,j}$ t-copula

In Figure 12.4 (b) we see that under slightly weaker differences for the correlations with $\rho_1 = 0.9$ and $\rho_2 = 0.1$ the dual bound VaR_α^D is better than the Bonferroni bounds $\mathrm{VaR}_\alpha^{W_A}$ and $\mathrm{VaR}_\alpha^{W_B}$ indicating again a weakness of the method of improved standard bounds. While the Fréchet bounds for the df's improve considerably by inclusion of two dimensional marginals, the corresponding VaR bounds for the aggregated sums only improve in certain cases which exhibit strong enough positive dependence.

Remark 12.6 (Reduced bounds versus Bonferroni-type bounds). The reduced bounds in [Embrechts et al. (2013)] consider the case with the weaker assumption that only the non-overlapping distributions F_{12} of $(X_1,X_2),\dots,F_{2n-1,2n}$ of (X_{2n-1},X_{2n}) are known and give a non-sharp reduction to the one-dimensional case for $Y_1 = X_1+X_2,\dots,Y_n = X_{2n-1}+X_{2n}$, which can be handled by the dual bounds. From the examples in the present and in the related papers on improved bounds, it seems reasonable to expect that the reduced bounds may be well better than the Bonferroni bounds if the positive dependence on the 2-dimensional marginals is not strong. But in grouped examples like in Example 12.1 b) or in related structured examples the Bonferroni bounds are better able to make use of this structure by the choice of a suitable dependence tree τ consisting of strongly positive dependent components. So in case that information on more than the serial two-dinensional marginals are available it seems that the Bonferroni-type bounds in combination with the improved standard bounds are better than the reduced bounds.

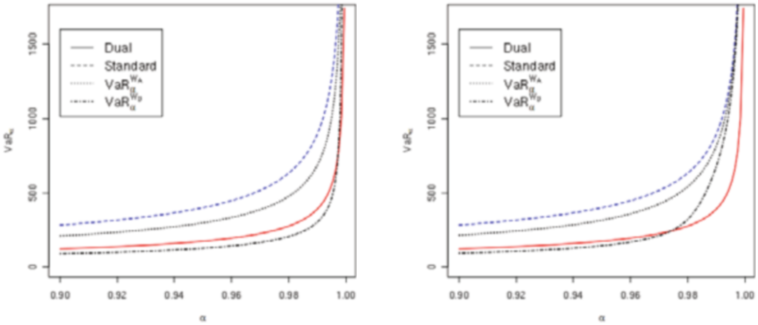
In the following example we compare the bounds for a set of heavy tailed marginal distributions and a different set of bivariate copulas.

Comparison of VaR bounds for bivariate Clayton copulas
We assume that $n = 20$ and X_i are Pareto-distributed, i. e. $F_i(x) = 1-x^{-2}$, $x \geq 1$. We assume that $C_{i,j}(u_i,u_j)$ is a Clayton copula with parameter ϑ. Note that for $\vartheta \to \infty$ the Clayton copula approaches comonotonicity while for $\vartheta \to 0$ it approaches independence. As in the third example we consider the case that the risks are divided into two groups. Within the groups the risks are approximatively comonotone (strongly positive dependent), i. e. the Clayton parameter $\vartheta = \vartheta_1$ is big. Between the groups the risks are approximatively independent, i. e. the Clayton parameter $\vartheta = \vartheta_2$ is small. This construction allows us to investigate the behaviour of the various VaR bounds in dependence of the dependence parameter ϑ of the copulas.

In Figure 12.5 and Table 12.2 we consider the choice $\vartheta_1 = 10\,000$, $\vartheta_2 = 0.1$ in 5(a) and $\vartheta = 1\,000$, $\vartheta_2 = 1$ in 5(b). As in the case of log-normal distributions we find that the Bonferroni bound $\mathrm{VaR}_\alpha^{W_B}$ is significantly better than $\mathrm{VaR}_\alpha^{W_A}$ and in particular improves the standard bound VaR_α^{S}.

In case $\vartheta_1 = 10\,000$ and $\vartheta_2 = 0.1$ the dual bound VaR_α^{D} improves on the Bonferroni bound $\mathrm{VaR}_\alpha^{W_B}$ for $\alpha \geq 0.9975 = \alpha_0$. Experience of further examples shows that this turning point moves to smaller values of α, the smaller the dependence parameter ϑ_1 gets. For example, for $\vartheta_1 = 1\,000$ and $\vartheta_2 = 1$ the turning point is $\alpha_0 = 0.975$. For $\alpha > \alpha_0$ the dual bounds are better than the Bonferroni bounds if the model is in enough distance to the comonotonic case.

As general conclusion of the examples in this section we obtain that the Bonferroni bound VaR_α^{B} and the dual bound VaR_α^{D} improve upon the standard bound



((a)) $\vartheta_1 = 10\,000$, $\vartheta_2 = 0.1$ ((b)) $\vartheta_1 = 1\,000$, $\vartheta_2 = 1$

Fig. 12.5: Comparison of VaR bounds, $n = 20$, Pareto-marginals, bivariate Clayton copula

			$\vartheta_1 = 10\,000$, $\vartheta_2 = 0.1$		$\vartheta_1 = 1\,000$, $\vartheta_2 = 1$	
α	VaR^S_α	VaR^D_α	$\mathrm{VaR}^{W_A}_\alpha$	$\mathrm{VaR}^{W_B}_\alpha$	$\mathrm{VaR}^{W_A}_\alpha$	$\mathrm{VaR}^{W_B}_\alpha$
0.9	282.842	123.288	209.452	88.717	214.864	93.168
0.99	894.427	389.871	684.720	301.371	813.773	676.727
0.999	2 828.427	1 232.883	2 574.672	2 141.456	2 797.193	2 764.304

Table 12.2: Comparison of VaR bounds, $n = 20$, Pareto-marginals, Clayton copulas with parameter ϑ_1 and ϑ_2 for $\alpha \geq 0.9$

VaR^W_α. $\mathrm{VaR}^{W_B}_\alpha$ also improves generally on $\mathrm{VaR}^{W_A}_\alpha$. The Bonferroni bound VaR^{W_B} improves for high degree of positive dependence on the dual bound VaR^D but for weaker forms of positive dependence the dual bound may be preferable. It should be noted however that the dual bound is typically only calculable for small dimensions for inhomogeneous cases. In these cases however the rearrangement algorithm (RA) can be applied to yield sharp marginal bounds. In our applications we used proxies for the calculation of the Bonferroni bounds. These were shown above to be sharp under some conditions.

Acknowledgements The simulations in Sections 12.3 and 12.4 were produced by Dennis Manko in the frame of his master thesis.

References

[Acciaio et al. (2013)] B. Acciaio, M. Beiglböck, F. Penkner, and W. Schachermayer. A model free version of the fundamental theorem of asset pricing and the super replication theorem. *Mathematical Finance*, 23(2):233–251, 2013.

[Beiglböck et al. (2013)] M. Beiglböck, P. Henry-Labordère, and F. Penkner. Model-independent bounds for option prices – a mass transport approach. *Finance and Stochastics*, 17(3):477–501, 2013.

[Bernard and Vanduffel (2015)] C. Bernard and S. Vanduffel. A new approach to assessing model risk in high dimensions. *J. Bank. Financ.*, 58:166–178, 2015.

[Bernard et al. (2012)] C. Bernard, X. Jiang, and S. Vanduffel. A note on 'Improved Fréchet bounds and model-free pricing of multi-asset options' by Tankov (2011). *Journal of Applied Probability*, 49(3):866–875, 2012.

[Bernard et al. (2013)] C. Bernard, Y. Liu, N. MacGillivray, and J. Zhang. Bounds on capital requirements for bivariate risk with given marginals and partial information on the dependence. *Depend. Model.*, 1:37–53, 2013.

[Bernard et al. (2015)] C. Bernard, L. Rüschendorf, and S. Vanduffel. Value-at-risk bounds with variance constraints. *Journal of Risk Insurance*, 2015. doi: 10.2139/ssrn.2572508.

[Bernard et al. (2016a)] C. Bernard, L. Rüschendorf, S. Vanduffel, and R. Wang. Risk bounds for factor models. *Social Science Research Network*, 2016a.

[Bernard et al. (2016b)] C. Bernard, L. Rüschendorf, S. Vanduffel, and J. Yao. How robust is the value-at-risk of credit risk portfolios? *Eur. J. Finance*, 2016b. doi: 10.1080/1351847X.2015.1104370.

[Bignozzi et al. (2015)] V. Bignozzi, G. Puccetti, and L. Rüschendorf. Reducing model risk via positive and negative dependence assumptions. *Insurance: Mathematics and Economics*, 61: 17–26, 2015.

[Cambanis et al. (1976)] S. Cambanis, G. Simons, and W. Stout. Inequalities for $Ek(X,Y)$ when the marginals are fixed. *Z. Wahrscheinlichkeitstheorie Verw. Geb.*, 36:285–294, 1976.

[Dall'Aglio (1972)] G. Dall'Aglio. Fréchet classes and compatibility of distribution functions. In *Symposia Mathematica*, volume 9, pages 131–150. Amer. Math. Soc., Providence, R.I., 1972.

[Denuit et al. (1999)] M. Denuit, C. Genest, and E. Marceau. Stochastic bounds on sums of dependent risks. *Insurance: Mathematics and Economics*, 25:85–104, 1999.

[Durante and Sempi (2016)] F. Durante and C. Sempi. *Principles of Copula Theory.* Boca Raton, FL: CRC Press, 2016.

[Embrechts and Puccetti (2006a)] P. Embrechts and G. Puccetti. Aggregating risk capital, with an application to operational risk. *Geneva Risk Insur. Rev.*, 31(2):71–90, 2006a.

[Embrechts and Puccetti (2006b)] P. Embrechts and G. Puccetti. Bounds for functions of dependent risks. *Finance and Stochastics*, 10:341–352, 2006b.

[Embrechts and Puccetti (2010)] P. Embrechts and G. Puccetti. Bounds for the sum of dependent risks having overlapping marginals. *Journal of Multivariate Analysis*, 101:177–190, 2010.

[Embrechts et al. (2003)] P. Embrechts, A. Höing, and A. Juri. Using copulae to bound the value-at-risk for functions of dependent risks. *Finance and Stochastics*, 7:145–167, 2003.

[Embrechts et al. (2013)] P. Embrechts, G. Puccetti, and L. Rüschendorf. Model uncertainty and VaR aggregation. *J. Bank Financ.*, 58(8):2750–2764, 2013.

[Embrechts et al. (2014)] P. Embrechts, G. Puccetti, L. Rüschendorf, R. Wang, and A. Beleraj. An academic response to Basel 3.5. *Risks*, 2(1):25–48, 2014.

[Frank et al. (1987)] M. J. Frank, R. B. Nelsen, and B. Schweizer. Best-possible bounds for the distribution of a sum – a problem of Kolmogorov. *Probability Theory and Related Fields*, 74: 199–211, 1987.

[Gaffke and Rüschendorf (1981)] N. Gaffke and L. Rüschendorf. On a class of extremal problems in statistics. *Math. Operationsforschung Statist.*, 12:123–135, 1981.

[Hoeffding (1940)] W. Hoeffding. *Maßstabinvariante Korrelationstheorie*. Ph.D.Thesis, Schr. Math. Inst. u. Inst. Angew. Math. Univ. Berlin 5, 181–233 (1940) and Berlin: Dissertation, 1940.

[Hunter (1976)] D. Hunter. An upper bound for the probability of a union. *Journal of Applied Probability*, 13:597–603, 1976.

[Kantorovich (1942)] L. V. Kantorovich. On the transfer of masses. *Dokl. Akad. Nauk USSR*, 37: 7–8, 1942.

[Kellerer (1984)] H. G. Kellerer. Duality theorems for marginal problems. *Z. Wahrscheinlichkeitstheorie Verw. Geb.*, 67:399–432, 1984.

[Kounias (1968)] E. G. Kounias. Bounds for the probability of a union, with applications. *Annals of Mathematical Statistics*, 39:2154–2158, 1968.

[Liu and Chan (2011)] H. Liu and C.-H. Chan. Best possible upper bound on VaR for dependent portfolio risk. *ICIC Express Letters*, 5(5):1795–1800, 2011.

[Lorentz (1953)] G. G. Lorentz. An inequality for rearrangements. *American Math. Monthly*, 60: 176–179, 1953.

[Lux and Papapantoleon (2016)] T. Lux and A. Papapantoleon. Improved Fréchet–Hoeffding bounds on d-copulas and applications in model-free finance. Preprint, 2016.

[Makarov (1981)] G. D. Makarov. Estimates for the distribution function of a sum of two random variables when the marginal distributions are fixed. *Theory of Probability and Its Applications*, 26:803–806, 1981.

[Manko (2015)] D. Manko. *Verteilungsschranken fr Portfolios von abhngigen Risiken*. Masterarbeit, Universitt Freiburg, 2015.

[Marshall and Olkin (1979)] A. W. Marshall and I. Olkin. *Inequalities: Theory of Majorization and Its Applications.*, volume 143 of *Mathematics in Science and Engineering*. New York etc.: Academic Press, 1979.

[Meilijson and Nadas (1979)] I. Meilijson and A. Nadas. Convex majorization with an application to the length of critical paths. *Journal of Applied Probability*, 16:671–677, 1979.

[Mesfioui and Quessy(2005)] M. Mesfioui and J.-F. Quessy. Bounds on the value-at-risk for the sum of possibly dependent risks. *Insurance: Mathematics and Economics*, 37(1):135–151, 2005.

[Nelsen (2006)] R. B. Nelsen. *An Introduction to Copulas. Properties and Applications*, volume 139 of *Lecture Notes in Statistics*. Springer, 1st edition 1999, 2nd edition, 2006.

[Nelsen et al. (2004)] R. B. Nelsen, J. J. Quesada Molina, J. A. Rodríguez Lallena, and M. Úbeda Flores. Best-possible bounds on sets of bivariate distribution functions. *Journal of Multivariate Analysis*, 90(2):348–358, 2004.

[Puccetti and Rüschendorf (2012)] G. Puccetti and L. Rüschendorf. Bounds for joint portfolios of dependent risks. *Statistics & Risk Modeling*, 29(2):107–131, 2012.

[Puccetti and Rüschendorf (2013)] G. Puccetti and L. Rüschendorf. Sharp bounds for sums of dependent risks. *Journal of Applied Probability*, 50(1):42–53, 2013.

[Puccetti and Wang (2015)] G. Puccetti and R. Wang. Extremal dependence concepts. *Statistical Science*, 30:485–517, 2015.

[Puccetti et al. (2013)] G. Puccetti, B. Wang, and R. Wang. Complete mixability and asymptotic equivalence of worst-possible VaR and ES estimates. *Insurance: Mathematics and Economics*, 53(3):821–828, 2013.

[Puccetti et al. (2015)] G. Puccetti, L. Rüschendorf, D. Small, and S. Vanduffel. Reduction of Value-at-Risk bounds via independent and variance information. *Scand. Actuar. J., (to appear)*, 2015. doi: 10.1080/03461248.2015.1119717.

[Puccetti et al. (2016)] G. Puccetti, L. Rüschendorf, and D. Manko. VaR bounds for joint portfolios with dependence constraints. *Dependence Modelling*, 4:368–381, 2016.

[Rachev and Rüschendorf (1994)] S. T. Rachev and L. Rüschendorf. Solution of some transportation problems with relaxed and additional constraints. *SIAM Control and Optimization*, 32: 673–689, 1994.

[Rachev and Rüschendorf (1998a)] S. T. Rachev and L. Rüschendorf. *Mass Transportation Problems. Vol. I: Theory.* Probability and its Applications. Springer, 1998a.

[Rachev and Rüschendorf (1998b)] S. T. Rachev and L. Rüschendorf. *Mass Transportation Problems.Vol. II: Applications.* Probability and its Applications. Springer, 1998b.

[Ramachandran and Rüschendorf (1997)] D. Ramachandran and Rüschendorf. Duality theorems for assignments with upper bounds. In V. Beneš et al., editors, *Distributions with Given Marginals and Moment Problems. Proceedings of the 1996 Conference, Prague, Czech Republic*, pages 283–290. Kluwer Academic Publishers, 1997.

[Ramachandran and Rüschendorf (1999)] D. Ramachandran and L. Rüschendorf. An extension of the nonatomic assignment problem. *Current Trends in Economics. Studies in Economic Theory*, 8:405–412, 1999.

[Ramachandran and Rüschendorf (2002)] D. Ramachandran and L. Rüschendorf. Assignment models for constrained marginals and restricted markets. In C. M. Cuadras, J. Fortiana, and J. A. Rodriguez-Lallena, editors, *Distributions with Given Marginals and Statistical Modelling*, pages 195–210. Kluwer Academic Publishing, 2002.

[Rodríguez-Lallena and Úbeda-Flores (2004)] J. A. Rodríguez-Lallena and M. Úbeda-Flores. Best-possible bounds on sets of multivariate distribution functions. *Commun. Stat., Theory Methods*, 33(4):805–820, 2004.

[Rüschendorf (1979)] L. Rüschendorf. *Vergleich von Zufallsvariablen bzgl. integralinduzierter Halbordnungen.* Habilitationsschrift, Mathematisch-Naturwissenschaftliche Fakultät der Rheinisch-Westfälischen Technischen Hochschule Aachen. 133 S., 1979.

[Rüschendorf (1980)] L. Rüschendorf. Inequalities for the expectation of Δ-monotone functions. *Z. Wahrscheinlichkeitstheorie Verw. Geb.*, 54:341–354, 1980.

[Rüschendorf (1981)] L. Rüschendorf. Sharpness of Fréchet bounds. *Z. Wahrscheinlichkeitstheorie Verw. Geb.*, 57:293–302, 1981.

[Rüschendorf (1982)] L. Rüschendorf. Random variables with maximum sums. *Advances in Applied Probability*, 14:623–632, 1982.

[Rüschendorf (1983)] L. Rüschendorf. Solution of a statistical optimization problem by rearrangement methods. *Metrika*, 30:55–62, 1983.

[Rüschendorf(1991a)] L. Rüschendorf. Bounds for distributions with multivariate marginals. In K. Mosler and M. Scarsini, editors, *Stochastic Orders and Decision Under Risk*, volume 19 of *IMS Lecture Notes*, pages 285–310. 1991a.

[Rüschendorf (1991b)] L. Rüschendorf. Fréchet bounds and their applications. In G. Dall'Aglio, S. Kotz, and G. Salinetti, editors, *Advances in Probability Distributions with Given Marginals*, volume 67 of *Mathematics and Its Applications*, pages 151–188. Springer, 1991b.

[Rüschendorf (2004)] L. Rüschendorf. Comparison of multivariate risks and positive dependence. *Journal of Applied Probability*, 41:391–406, 2004.

[Rüschendorf (2005)] L. Rüschendorf. Stochastic ordering of risks, influence of dependence and a.s. constructions. In N. Balakrishnan, I. G. Bairamov, and O. L. Gebizlioglu, editors, *Advances on Models, Characterizations and Applications*, pages 19–56. Chapman & Hall/CRC Press, 2005.

[Rüschendorf (2007)] L. Rüschendorf. Monge–Kantorovich transportation problem and optimal couplings. *Jahresbericht der DMV*, 109:113–137, 2007.

[Rüschendorf (2013)] L. Rüschendorf. *Mathematical Risk Analysis. Dependence, Risk Bounds, Optimal Allocations and Portfolios.* Berlin: Springer, 2013.

[Rüschendorf (2017)] L. Rüschendorf. Risk bounds and partial dependence information. In D. Ferger, W. Gonzáles Manteiga, T. Schmidt, and J.-L. Wang, editors, *From Statistics to Mathematical Finance, (to appear)*. Springer, 2017.

[Sadooghi-Alvandi et al. (2013)] S. Sadooghi-Alvandi, Z. Shishebor, and H. Mardani-Fard. Sharp bounds on a class of copulas with known values at several points. *Commun. Stat., Theory Methods*, 42(12):2215–2228, 2013.

[Tankov (2011)] P. Tankov. Improved Fréchet bounds and model-free pricing of multi-asset options. *Journal of Applied Probability*, 48(2):389–403, 2011.

[Tchen (1980)] A. H. Tchen. Inequalities for distributions with given marginals. *Annals of Probability*, 8:814–827, 1980.

[Wang and Wang (2011)] B. Wang and R. Wang. The complete mixability and convex minimization problems with monotone marginal densities. *Journal of Multivariate Analysis*, 102(10): 1344–1360, 2011.

[Wang et al. (2013)] R. Wang, L. Peng, and J. Yang. Bounds for the sum of dependent risks and worst value-at-risk with monotone marginal densities. *Finance and Stochastics*, 17(2):395–417, 2013.

[Williamson and Downs(1990)] R. C. Williamson and T. Downs. Probabilistic arithmetic. I. Numerical methods for calculating convolutions and dependency bounds. *International Journal of Approximate Reasoning*, 4(2):89–158, 1990.

[Worsley(1982)] K. J. Worsley. An improved Bonferroni inequality and applications. *Biometrika*, 69(2):297–302, 1982.

Chapter 13
Quasi–copulas: a brief survey

Carlo Sempi

Abstract This article is a survey of quasi–copulas. The *genesis*, the history and the properties of quasi–copulas are presented, with special emphasis on their characterisation and on recent results. The presentation will also highlight the key part played by Roger Nelsen in all stages of the development of the theory.

13.1 Preliminaries and notation

The reader is supposed to have some familiarity with the definition and the main properties of copulas, as they can be derived from Nelsen's classical book ([22] for the first edition and [24] for the second one) or from the more recent [10]. Here only the notation is recalled that will be used in the sequel.

The extended reals will be denoted by $\overline{\mathbb{R}} = \mathbb{R} \cup \{-\infty, +\infty\}$ and the unit interval $[0,1]$ by $\mathbb{I}$. A distribution function(=df) is a function $F : \overline{\mathbb{R}} \to \mathbb{I}$ that is increasing (in the weak sense), right–continuous on $\mathbb{R}$ and such that $F(-\infty) = 0$ and $F(+\infty) = 1$. Given a probability space $(\Omega, \mathscr{F}, \mathbb{P})$, a random variable(=rv) is a measurable function from Ω into $\overline{\mathbb{R}}$; its df F_X is defined by $F_X(t) := \mathbb{P}(X \leq t)$. An almost surely constant rv $X = a$ has a df given by the unit step function with jump at a, i.e.

$$\varepsilon_a(t) := \begin{cases} 0, & t < a\,, \\ 1, & t \geq a\,. \end{cases}$$

The set of df's will be denoted by $\mathscr{D}$.

Carlo Sempi
Dipartimento di Matematica e Fisica "Ennio De Giorgi", Università del Salento, Lecce, 73100 Italy
e-mail: carlo.sempi@unisalento.it

M. Úbeda Flores et al. (eds.), *Copulas and Dependence Models with Applications*,
DOI 10.1007/978-3-319-64221-5_13

13.2 Quasi–copulas make their appearance

In a probabilistic setting, quasi–copulas were first introduced in 1993 in a paper by Alsina, Nelsen and Schweizer [1] in the bivariate case. In the same year, but published in 1996, the definition was extended to the general case $d \geq 3$ by Nelsen, Quesada Molina, Schweizer and Sempi [29].

The definition of a quasi–copula requires the notion of a track, which is here recalled in the d–dimensional setting ($d \geq 2$).

Definition 13.1. A *track* B in the unit square $\mathbb{I}^d$ is a subset of $\mathbb{I}^d$ that can be written in the form

$$B := \{(F_1(t), \dots, F_d(t)) : t \in \mathbb{I}\}, \tag{13.1}$$

where F_1, $\dots$, F_d are continuous distribution functions such that $F_j(0) = 0$ and $F_j(1) = 1$ for every j in $\{1, \dots, d\}$. ♢

Definition 13.2. A *d-quasi-copula* is a function $Q : \mathbb{I}^d \to \mathbb{I}$ such that for every track B in $\mathbb{I}^d$ there exists a d–copula C_B that coincides with Q on B, namely such that, for every point $\mathbf{u} \in B$, $Q(\mathbf{u}) = C_B(\mathbf{u})$. ♢

It follows immediately from this definition that every copula $C \in \mathscr{C}_d$ is a quasi–copula, since it suffices to take $C_B = C$ for every track B. However, once we recall a characterisation of quasi–copulas, it will be seen that quasi–copulas exist that are not copulas; therefore, if $\mathscr{Q}_d$ denotes the set of d–quasi–copulas one has the strict inclusion $\mathscr{C}_d \subset \mathscr{Q}_d$. A quasi-copula that is not a copula will be called *proper*.

But why were quasi–copulas introduced? It had been known at least since the publication of the paper by Schweizer and Sklar [34] that, while an operation on random variables defined on the same probability space $(\Omega, \mathscr{F}, \mathbb{P})$ induces a corresponding operation on their distribution functions, not all operations on distribution functions may be derived from a corresponding operation on random variables defined on the *same* probability space. Below we provide a formal definition of derivability in this sense (see [29]).

Definition 13.3. An operation φ on $\mathscr{D}$ is said to be *derivable* from a function on random variables if there exists a Borel–measurable function $V : \mathbb{R}^d \to \mathbb{R}$ that satisfies the following condition: for every collection of d dfs $F_1, \dots, F_d$ in $\mathscr{D}$ there exist a common probability space $(\Omega, \mathscr{F}, \mathbb{P})$ and d random variables $X_1, \dots, X_d$ on $(\Omega, \mathscr{F}, \mathbb{P})$ such that, for every $j \in \{1, \dots, d\}$, F_j is the df of X_j, $F_{X_j} = F_j$ and such that $\varphi(F_1, \dots, F_d)$ is the df of $V(X_1, \dots, X_d)$. ♢

It is easily seen that the convex combination of two dfs is not derivable from a binary operation on random variables defined on the same probability space (see [2]).

Example 13.1. Assume, if possible, that the convex combination $\varphi = \alpha F + (1 - \alpha) G$ of two dfs F and G with parameter $\alpha \in]0, 1[$ is derivable from a suitable function V on corresponding rvs.

For all $a,b \in \mathbb{R}$, $a \neq b$, let F and G be the unit step functions at a and b, respectively. Then F and G are, respectively, the dfs of the rvs X and Y, which are defined on a common probability space and are almost surely equal, respectively, to a and b. Hence $V(X,Y)$ is a rv defined on the same probability space and almost surely equal to $V(a,b)$. Thus the distribution function of $V(X,Y)$ is $\varepsilon_{V(a,b)}$, the unit step function at $V(a,b)$. But because φ is derivable from V, the distribution function of $V(X,Y)$ must be equal to $\alpha F + (1-\alpha)G$, which is not a step function with a single jump, and, hence, cannot be the df of $V(X,Y)$. ■

As a consequence, it is of interest to look for, and to characterise, classes of operations on dfs that are indeed derivable in the sense of Definition 13.3. The first class to be studied under this framework was that of the pointwise induced operations.

Definition 13.4. A d–operation φ on $\mathscr{D}$ is said to be *induced pointwise* by a function $\psi : \mathbb{I}^d \to \mathbb{I}$ if,

$$\varphi(F_1,\dots,F_d)(t) = \psi(F_1(t),\dots,F_d(t))$$

for every choice of $F_1,\dots,F_d$ in $\mathscr{D}$ and for every $t \in \mathbb{R}$. ◊

A typical example of operation on $\mathscr{D}$ that is induced pointwise is the convex combination of dfs, namely a mixture in the language of Statistics, which can be derived by the function $\psi(\mathbf{u}) = \sum_{j=1}^{d} \alpha_j u_j$, where $\alpha_j \geq 0$ for every index j and $\sum_{j=1}^{d} \alpha_j = 1$.

Of course, not every operation on $\mathscr{D}$ is induced pointwise. Consider, for instance, the *convolution* of two dfs F and G defined, for every $x \in \mathbb{R}$, by

$$F \otimes G(x) = \int_{\mathbb{R}} F(x-t)\, \mathrm{d}G(t)\,.$$

Now, since the value of the convolution of two dfs F and G at the point x generally depends on more than just the values of F and G at x, convolution is not induced pointwise by any two–place function. However, as is well known, the convolution is interpretable in terms of rvs, since $F \otimes G$ can be seen as the df of the sum of independent rvs X and Y such that $X \sim F$ and $Y \sim G$.

The class of operations that are both induced pointwise and derivable from operations on random variables turns out to be quite small, as the following theorem shows.

Theorem 13.1. *Let φ be a binary operation on $\mathscr{D}$ that is both induced pointwise by a two-place function $\psi : \mathbb{I}^2 \to \mathbb{I}$ and derivable from a function V on random variables defined on a common probability space. Then precisely one of the following holds:*

(a) $V = \max$ *and ψ is a quasi-copula;*
(b) $V = \min$ *and $\psi(x,y) = x + y - Q(x,y)$, where Q is a quasi-copula;*
(c) *V and φ are trivial in the sense that, for all x and y in $\mathbb{R}$ and for all a and b in $\mathbb{I}$, either $V(x,y) = x$ and $\psi(a,b) = a$ or $V(x,y) = y$ and $\psi(a,b) = b$.*

Proof. Since $\varphi(F,G)$, F and G are dfs, the function ψ is right-continuous and increasing in each place; moreover, $\psi(0,0)=0$ and $\psi(1,1)=1$. Consider $F=\varepsilon_x$ and $G=\varepsilon_y$ the unit step functions at x and y, respectively, with $x,y\in\mathbb{R}$. Then F and G are the dfs of two random variables X and Y that may be defined on a common probability space $(\Omega,\mathscr{F},\mathbb{P})$ and which are respectively equal to x and y almost surely. It follows that $V(X,Y)$ is also a random variable on $(\Omega,\mathscr{F},\mathbb{P})$ and that it takes the value $V(x,y)$ almost surely; hence, its df is the step function $\varepsilon_{V(x,y)}$, which, since φ is derivable from V, must be $\varphi(\varepsilon_x,\varepsilon_y)$. Thus, for every $t\in\mathbb{R}$, one has

$$\begin{aligned}\varepsilon_{V(x,y)}(t) &= \varphi(\varepsilon_x,\varepsilon_y)(t)=\psi(\varepsilon_x(t),\varepsilon_y(t))\\ &=\begin{cases}\psi(1,0)\,\varepsilon_x(t)+(1-\psi(1,0))\,\varepsilon_y(t)\,, & x\le y\,,\\ (1-\psi(0,1))\,\varepsilon_x(t)+\psi(0,1)\,\varepsilon_y(t)\,, & x\ge y\,.\end{cases}\end{aligned}$$

But $\varepsilon_{V(x,y)}$ takes only the values 0 and 1, so that one of the following four cases holds:

(a) $\psi(1,0)=0$, $\psi(0,1)=0$ and $V(x,y)=\max\{x,y\}$;
(b) $\psi(1,0)=1$, $\psi(0,1)=1$ and $V(x,y)=\min\{x,y\}$;
(c_1) $\psi(1,0)=1$, $\psi(0,1)=0$ and $V(x,y)=x$;
(c_2) $\psi(1,0)=0$, $\psi(0,1)=1$ and $V(x,y)=y$.

Now let $B=\{(\alpha(t),\beta(t)):t\in\mathbb{I}\}$ be a track and let F and G be given by

$$F(t)=\begin{cases}0\,, & t<0\,,\\ \alpha(t)\,, & t\in\mathbb{I}\,,\\ 1\,, & t>1\,,\end{cases}\qquad G(t)=\begin{cases}0\,, & t<0\,,\\ \beta(t)\,, & t\in\mathbb{I}\,,\\ 1\,, & t>1\,.\end{cases}$$

Let X, Y and $V(X,Y)$ be rvs having dfs F, G and $\varphi(F,G)$, respectively, and let H and C be the joint df and the (uniquely defined) copula of X and Y. In case (a), one has $V(X,Y)=\max\{X,Y\}$, so that $F_{V(X,Y)}$, the df of $V(X,Y)$, is given by $F_{V(X,Y)}(t)=H(t,t)$; then

$$\begin{aligned}\psi(F(t),G(t)) &= \varphi(F,G)(t)=F_{V(X,Y)}(t)=H(t,t)\\ &=C(F(t),G(t))\,.\end{aligned}$$

Thus ψ agrees with a copula on the track B, which is arbitrary; therefore ψ is a quasi-copula.

In case (b), one has $V(X,Y)=\min\{X,Y\}$, so that $F_{V(X,Y)}(t)=F(t)+G(t)-H(t,t)$. Therefore

$$\begin{aligned}\psi(F(t),G(t)) &= \varphi(F,G)(t)=F_{V(X,Y)}(t)\\ &=F(t)+G(t)-C(F(t),G(t))\,,\end{aligned}$$

which is the assertion in this case.

In case (c_1) one has $V(X,Y) = X$, whence $\psi(F(t),G(t)) = F_{V(X,Y)}(t) = F(t)$; thus $\psi(u,v) = u$; the remaining case (c_2) is completely analogous. □

Theorem 13.1 is a special case, for $d = 2$, of a more general result holding for every $d \geq 2$ (see [29], especially Theorem 2.5 and Corollary 2.6).

A quasi-copula satisfies the same general bounds as a copula.

Theorem 13.2. *For every d-quasi-copula Q and for every point* $\mathbf{u} \in \mathbb{I}^d$

$$W_d(\mathbf{u}) \leq Q(\mathbf{u}) \leq M_d(\mathbf{u}). \tag{13.2}$$

Proof. Given a point $\mathbf{u} \in \mathbb{I}^d$, consider any track B going through $\mathbf{u}$ and a copula C_B that coincides with Q on B; then

$$W_d(\mathbf{u}) \leq C_B(\mathbf{u}) \leq M_d(\mathbf{u}).$$

Since $Q(\mathbf{u}) = C_B(\mathbf{u})$, the inequalities (13.2) have been proved. □

The lower and upper bounds provided by (13.2) are, obviously, the best possible in $\mathscr{Q}_d$.

It easily follows from (13.2) that every quasi–copula $Q \in \mathscr{Q}_2$ satisfies

$$\max\{u,v\} \leq u+v-Q(u,v) \leq \min\{u+v,1\} \tag{13.3}$$

for all $(u,v) \in \mathbb{I}^2$. Since neither the arithmetic mean $\psi(u,v) = \alpha u + (1-\alpha)v$, $\alpha \in]0,1[$, nor the geometric mean $\psi(u,v) = \sqrt{uv}$ satisfies (13.3), it follows at once, as a consequence of Theorem 13.1, that neither mixtures nor the geometric mean are derivable.

13.3 Characterisation of quasi–copulas

Since it was hard to check whether a function $Q : \mathbb{I}^d \to \mathbb{I}$ is a proper quasi–copula[1], research was soon started in order to characterise quasi–copulas in an operationally convenient manner. The necessary conditions that a quasi–copula satisfies are given in the next theorem.

Theorem 13.3. *A d-quasi-copula Q satisfies the following properties:*

(a) *for every* $j \in \{1,\dots,d\}$, $Q(1,\dots,1,u_j,1,\dots,1) = u_j$;
(b) *Q is increasing in each of its arguments, viz. for every* $j \in \{1,\dots,d\}$ *and for every point* $(u_1,\dots,u_{j-1},u_{j+1},\dots,u_d)$ *in* $\mathbb{I}^{d-1}$, *the function* $t \mapsto Q(\mathbf{u}_j(t))$, *where* $\mathbf{u}_j(t) := (u_1,\dots,u_{j-1},t,u_{j+1},\dots,u_d)$, *is increasing;*

[1] It is easily checked whether Q is a copula, so that the identification problem remains only for proper quasi–copulas.

(c) *Q satisfies the following Lipschitz condition: if* $\mathbf{u}$ *and* $\mathbf{v}$ *are in* $\mathbb{I}^d$*, then*

$$|Q(\mathbf{v}) - Q(\mathbf{u})| \leq \sum_{j=1}^{d} |v_j - u_j| . \tag{13.4}$$

Properties (a) and (b) of Theorem 13.3 together imply $Q(\mathbf{u}) = 0$ when at least one of the components of $\mathbf{u}$ vanishes, viz. when $\min\{u_1, \dots, u_d\} = 0$.

But not only are the conditions of Theorem 13.3 necessary; they are also sufficient, so that they completely characterise quasi–copulas. This is stated in the next theorem.

Theorem 13.4. *A function $Q : \mathbb{I}^d \to \mathbb{I}$ is a d-quasi-copula if, and only if, it satisfies the following conditions:*

(a) *for every $j \in \{1, \dots, d\}$, $Q(1, \dots, 1, u_j, 1, \dots, 1) = u_j$;*
(b) *Q is increasing in each of its arguments;*
(c) *Q satisfies Lipschitz condition* (13.4).

The characterisation of Theorem 13.4 is now frequently used as the standard definition of a quasi–copula. Using it one immediately proves that W_d is a d-quasi-copula, and, since it is already known that it is not a d-copula if $d \geq 3$, it is a proper d-quasi-copula for $d \geq 3$.

The proof of Theorem 13.4 was given by Genest *et al.* ([17]) in the case $d = 2$ and by Cuculescu and Theodorescu ([4]) for $d \geq 3$; the proofs provided by these papers are not similar. At the time of publication, however, it was not clear to the authors of the first one of them how the proof given in the appendix could be generalised to characterise quasi–copulas in higher dimension. There was an attempt by the same authors to prove the general case along the same lines of [17]; they even submitted a paper with the general "proof" but promptly withdrew it when they found a counterexample to one of the crucial steps. Then Cuculescu and Theoderescu characterised quasi–copulas using completely different ideas and techniques. It is worth giving at least a glimpse of their complicated proof [2].

The idea of the proof is to approximate a given, but arbitrary, track by a polygonal; more specifically, given a d-quasi-copula Q and a track, one can consider n points along this track and the polygonal having those points as the endpoints of the segments constituting it. For n sufficiently large, the polygonal approximates the given track; moreover, one can find a d–copula agreeing with Q at those points. To this end one has recourse to the concept of system.

Definition 13.5. A *system* will be a finite sequence of points in $\mathbb{I}^d$

$$\mathbf{0} = \mathbf{a}^{(0)} \leq \mathbf{a}^{(1)} \leq \cdots \leq \mathbf{a}^{(n)} = \mathbf{1},$$

together with a finite sequence, of the same length n, of numbers in $\mathbb{I}$, $0 = q_0 \leq q_1 \leq \cdots \leq q_n = 1$, such that

[2] The reader is referred to [10] for the proofs of the results reported in this section.

$$\max\{0, a_1^{(j)} + a_2^{(j)} + \cdots + a_d^{(j)} - d + 1\} \le q_j \le \min\{a_1^{(j)}, a_2^{(j)}, \ldots, a_d^{(j)}\}$$

and

$$q_{j+1} - q_j \le \left\| \mathbf{a}^{(j+1)} - \mathbf{a}^{(j)} \right\|_1 = \sum_{i=1}^{d} \left(a_i^{(j+1)} - a_i^{(j)} \right) .$$

Such a system will be denoted by

$$\Sigma = \left(\mathbf{a}^{(0)}, \mathbf{a}^{(1)}, \ldots, \mathbf{a}^{(n)}; q_0, q_1, \ldots, q_n \right) ,$$

and it will be said to be *increasing* if $\mathbf{a}^{(j)} < \mathbf{a}^{(j+1)}$ for every $j = 0, 1, \ldots, n-1$. Finally a system will be said to be *simple* if

- it is increasing;
- it is of even length;
- $q_{j+1} = q_j$ for every even index j, while $q_{j+1} - q_j = \|\mathbf{a}^{(j+1)} - \mathbf{a}^{(j)}\|_1$ for every odd index.

The *track of a system* is the polygonal line that consists of all the segments S_j with endpoints $\mathbf{a}^{(j)}$ and $\mathbf{a}^{(j+1)}$ $(j = 1, \ldots, n-1)$. The *function of a system* is the piecewise linear function f defined by $f(\mathbf{a}^{(j)}) := q_j$ and linear on S_j. One speaks of a d-copula C as the *solution* for a system when $C(\mathbf{a}^{(j)}) = q_j$ for $j = 0, 1, \ldots, n$. ◇

The proof consists in showing that every system has a solution and is achieved in a series of steps.

Lemma 13.1. *Let* Σ *be an increasing system. If* k *is such that* $2 \le k \le n-2$*, assume the following:*

(a) *the system* Σ_k *obtained from* Σ *by deleting the elements with indices* $1, \ldots, k-1$*, has a solution* C_k*;*
(b) *there exists a measure* μ *on the Borel sets of* $[\mathbf{0}, \mathbf{a}^{(k)}]$ *such that* $\mu([\mathbf{0}, \mathbf{a}^{(j)}]) = q_j$ *for* $j \le k$ *and such that its i-th margin* μ_i *satisfies*

$$\mu_i \le \mathbf{1}_{[0, a_i^{(k)}]} \cdot \lambda = \lambda \Big|_{[0, a_i^{(k)}]} .$$

Then Σ *has a solution.*

Corollary 13.1. *Let* Σ *be as in Lemma 13.1 and assume that a solution exists for the system* Σ_k' *obtained from* Σ *by deleting the elements with index* $j \ge k+1$*. Then condition* (b) *of Lemma 13.1 holds.*

Lemma 13.2. *Let* Σ *be a simple system*

$$\Sigma = \left(\mathbf{a}^{(0)}, \mathbf{a}^{(1)}, \ldots, \mathbf{a}^{(n)}; q_0, q_1, \ldots, q_n \right) .$$

Then there exists a measure μ *on the Borel sets of* $[\mathbf{0}, \mathbf{a}^{(2)}]$ *such that*

$$\mu([\mathbf{0},\mathbf{a}^{(1)}]) = q_1 \quad \textit{and} \quad \mu([\mathbf{0},\mathbf{a}^{(2)}]) = q_2\,, \tag{13.5}$$

and such that for every one of its margins μ_i*, one has* $\mu_i \le \lambda$ *.*

Two operations are introduced on a system Σ; each of them yields a new system as a result.

Straightening: this operation consists in deleting some elements of a system Σ; specifically, one deletes the elements of index j if either one of the following conditions holds:

- $q_{j-1} = q_j = q_{j+1}$;
- $\left\|\mathbf{a}^{(j)} - \mathbf{a}^{(j-1)}\right\|_1 = q_j - q_{j-1}$ and $\left\|\mathbf{a}^{(j+1)} - \mathbf{a}^{(j)}\right\|_1 = q_{j+1} - q_j$.

Smashing: Consider all the segments of endpoints $\mathbf{a}^{(j)}$ and $\mathbf{a}^{(j+1)}$ such that

$$q_j < q_{j+1} < q_j + \left\|\mathbf{a}^{(j+1)} - \mathbf{a}^{(j)}\right\|_1\,,$$

and on each such segment insert a new point $\mathbf{b}^{(j)}$ determined by the condition

$$\left\|\mathbf{a}^{(j+1)} - \mathbf{b}^{(j)}\right\|_1 = q_{j+1} - q_j\,;$$

the function of Σ then takes the value q_j at $\mathbf{b}^{(j)}$ and is linear on the segment of endpoints $\mathbf{b}^{(j)}$ and $\mathbf{a}^{(j+1)}$.

Lemma 13.3. *If* C *is a solution for a system* Σ*, it is also a solution for the system obtained from* Σ *by either smashing or straightening it, or both.*

Lemma 13.4. *Let*

$$\Sigma = \left(\mathbf{a}^{(0)},\mathbf{a}^{(1)},\mathbf{a}^{(2)},\mathbf{a}^{(3)},\mathbf{1};0,0,q,q,1\right)$$

be a simple system of length 4*; then the system*

$$\Sigma' = \left(\mathbf{a}^{(0)},\mathbf{a}^{(2)},\mathbf{a}^{(3)},\mathbf{1};0,q,q,1\right)$$

obtained from Σ *by deleting the elements with index* 1 *has a solution.*

Lemmata 13.2 and 13.4 together yield

Corollary 13.2. *Every simple system of length* 4 *has a solution.*

The final auxiliary result is provided by the following lemma.

Lemma 13.5. *Every system* Σ *has a solution.*

At this point one is ready for the final step in the proof.

Proof (of Theorem 13.4). Because of Theorem 13.3, it is already known that a quasi-copula Q satisfies the properties of the statement.

Conversely, let a function $Q : \mathbb{I}^d \to \mathbb{I}$ be given that satisfies properties (a), (b) and (c), and let

$$B = \{(F_1(t), \dots, F_d(t)) : t \in \mathbb{I}\}$$

be a track in $\mathbb{I}^d$. For $n = 2^m$ consider the points in $\mathbb{I}^d$

$$\mathbf{u}^{(j)} = \left(F_1\left(\frac{j}{2^m}\right), \dots, F_d\left(\frac{j}{2^m}\right)\right),$$

and set $q_j := Q\left(\mathbf{u}^{(j)}\right)$ $(j = 1, 2, \dots, 2^m)$. As a consequence of Lemma 13.5, for every $m \in \mathbb{N}$, there exists a copula C_m that coincides with Q at each of these points, viz., $C_m(\mathbf{u}^{(j)}) = Q(\mathbf{u}^{(j)})$.

Since $\mathscr{C}_d$ is compact, the sequence (C_m) of copulas obtained in this fashion contains a subsequence $\left(C_{m(k)}\right)_{k\in\mathbb{N}}$ that converges to a copula C. This limiting copula coincides with Q on B. To see this, fix $t \in \mathbb{I}$ and for each $k \in \mathbb{N}$, let $j = j(m(k))$ be the largest integer smaller than, or, at most, equal to, $2^{m(k)} - 1$ for which

$$\left| t - \frac{j}{2^{m(k)}} \right| \le \frac{1}{2^{m(k)}}.$$

Because of the continuity of the functions F_i $(i = 1, \dots, d)$ it is then possible to take $m(k)$ large enough to ensure that

$$\left| F_i(t) - F_i\left(\frac{j}{2^m}\right) \right| < \frac{\varepsilon}{3d}$$

simultaneously for every index $(i = 1, \dots, d)$ and for a given $\varepsilon > 0$. Since $(C_{m(k)})$ converges to C on $\mathbb{I}^d$, for k sufficiently large, one has

$$\left| C_{m(k)}(F_1(t), \dots, F_d(t)) - C(F_1(t), \dots, F_d(t)) \right| < \frac{\varepsilon}{3}.$$

Keeping in mind that $C_{m(k)}(\mathbf{u}^{(j)}) = Q(\mathbf{u}^{(j)})$ $(j = 1, \dots, 2^{m(k)})$ and that both Q and $C_{m(k)}$ satisfy the Lipschitz condition, then one has

$$\begin{aligned}
|Q(F_1(t), \dots, F_d(t)) - C(F_1(t), \dots, F_d(t))| \\
&\le \left| Q(F_1(t), \dots, F_d(t)) - Q(\mathbf{u}^{(j)}) \right| \\
&\quad + \left| C_{m(k)}(\mathbf{u}^{(j)}) - C_{m(k)}(F_1(t), \dots, F_d(t)) \right| \\
&\quad + \left| C_{m(k)}(F_1(t), \dots, F_d(t)) - C(F_1(t), \dots, F_d(t)) \right| \\
&< d\,\frac{\varepsilon}{3d} + d\,\frac{\varepsilon}{3d} + \frac{\varepsilon}{3} = \varepsilon.
\end{aligned}$$

Because of the arbitrariness of ε, Q coincides with the copula C on the track B and is, therefore, a d-quasi-copula. □

It is this author's belief, and his hope, that the proof just outlined will in the future be replaced by a simpler and more transparent one.

It is worth mentioning that a different characterisation is possible (see [32]).

Theorem 13.5. *For a function $Q : \mathbb{I}^d \to \mathbb{I}$ the following statements are equivalent:*

(a) *Q is a d–quasi–copula;*
(b) *Q satisfies property* (a) *of Theorem* 13.3*, and the functions*

$$t \mapsto Q(\mathbf{u}_j(t)) \qquad (j = 1, \dots, d) \tag{13.6}$$

are absolutely continuous for every choice of $\mathbf{u} = (u_1, \dots, u_d)$ in $\mathbb{I}^{d-1}$ and

$$0 \le \partial_j Q(\mathbf{u}_j(t)) \le 1 \qquad (j = 1, \dots, d) \tag{13.7}$$

for almost every $t \in \mathbb{I}$ and for all $(u_1, \dots, u_d) \in \mathbb{I}^{d-1}$.

13.4 Archimedeanity

In order to speak of Archimedean quasi–copulas it is convenient to compare them with the analogous concept for copulas. To this end the well known concept of generator will be needed, since Archimedean copulas are parametrised via a function of a single variable, which is defined below.

Definition 13.6. A function $\varphi : [0, +\infty[\to \mathbb{I}$ is said to be a *generator* if it is continuous, decreasing and $\varphi(0) = 1$, $\lim_{t \to +\infty} \varphi(t) = 0$ and is strictly decreasing on $[0, t_0]$, where $t_0 := \inf\{t > 0 : \varphi(t) = 0\}$. ◊

The *pseudo–inverse* of the generator φ is defined by

$$\varphi^{(-1)}(t) := \begin{cases} \varphi^{-1}(t), & t \in]0, 1], \\ t_0, & t = 0. \end{cases} \tag{13.8}$$

Notice that $\varphi^{(-1)}(\varphi(t)) = \min\{t, t_0\}$ for every $t \ge 0$.

Definition 13.7. A d–copula C with $d \ge 2$ is said to be *Archimedean* if a generator φ exists such that C may be represented in the form

$$C(\mathbf{u}) = \varphi\left(\varphi^{(-1)}(u_1) + \cdots + \varphi^{(-1)}(u_d)\right), \tag{13.9}$$

for every $\mathbf{u} \in \mathbb{I}^d$. ◊

McNeil and Nešlehová [21] have characterised the generators for which the function defined by (13.9) is a d–copula.

Theorem 13.6. *Let $\varphi : [0,+\infty[\to \mathbb{I}$ be a generator. Then the following statements are equivalent:*

(a) *φ is d-monotone on $[0,+\infty[$;*
(b) *the function $C : \mathbb{I}^d \to \mathbb{I}$ defined by* (13.9) *is a d-copula.*

A function $f :]a,b[\to \mathbb{R}$ is called *d–monotone* in $]a,b[$, where $-\infty \leq a < b \leq +\infty$ and $d \geq 2$ if

- it is differentiable up to order $d-2$;
- for every $x \in]a,b[$, its derivatives satisfy

$$(-1)^k f^{(k)}(x) \geq 0$$

  for $k = 0, 1, \dots, d-2$;
- $(-1)^{d-2} f^{(d-2)}$ is decreasing and convex in $]a,b[$.

Moreover, if f has derivatives of every order in $]a,b[$ and if

$$(-1)^k f^{(k)}(x) \geq 0,$$

for every $x \in]a,b[$ and for every $k \in \mathbb{Z}_+$, then f is said to be *completely monotone*. Let $I \subseteq \mathbb{R}$ be an interval. A function $f : I \to \mathbb{R}$ is said to be *d–monotone* (respectively, *completely monotone*) on I, with $d \in \mathbb{N}$, if it is continuous on I and if its restriction to the interior I° of I is d-monotone (respectively, completely monotone).

For $d = 1$, f is said to be (1-)monotone on $]a,b[$ if it is decreasing and positive on this interval, while in the case $d = 2$, a 2-monotone function f is simply a decreasing and convex function.

Kimberling [19] proved the following

Corollary 13.3. *Let $\varphi : [0,+\infty[\to \mathbb{I}$ be a generator. Then the following statements are equivalent:*

(a) *φ is completely monotone on $[0,+\infty[$;*
(b) *the function $C_\varphi : \mathbb{I}^d \to \mathbb{I}$ defined by* (13.9) *is a d-copula for every $d \geq 2$.*

Nelsen *et al.* [26] completely characterised Archimedean quasi–copulas.

Theorem 13.7. *For $d \geq 2$ and for a given generator φ, the following statements are equivalent:*

(a) *the generator φ is convex;*
(b) *the function $Q_\varphi : \mathbb{I}^d \to \mathbb{I}$ defined by*

$$Q_\varphi(\mathbf{u}) = \varphi\left(\varphi^{(-1)}(u_1) + \cdots + \varphi^{(-1)}(u_d)\right)$$

is a d–quasi–copula.

As a consequence of this latter results no proper Archimedean quasi–copula exists when $d = 2$, while for $d > 2$ there are proper Archimedean quasi–copulas, the easiest example being, for instance for $d = 3$, the lower Fréchet–Hoeffding bound W_3, which is a proper 3–quasi–copula; in fact its generator $\varphi(t) = \max\{1-t, 0\}$ is convex but not differantiable at $t = 1$, and, hence, not 3–monotone. Another example is given in [26].

13.5 Discrete quasi–copulas

The construction of discrete quasi–copulas was introduced in [30], for the case $d = 2$, although the construction applies in principle to any dimension d. Here, only the case $d = 2$ will be briefly discussed. Let $I_{n,m}$ be the set of points of $\mathbb{I}^2$ given by

$$I_{n,m} := \left\{ \left(\frac{i}{n}, \frac{j}{m}\right) : i \in \{0, 1, \dots, n\}, j \in \{0, 1, \dots, m\} \right\},$$

where n and m are natural numbers with $n, m \geq 2$.

Definition 13.8. A *discrete quasi–copula* on $I_{n,m}$ is a function $Q_{n,m} : I_{n,m} \to \mathbb{I}$ that satisfies the following conditions

(a) $Q_{n,m}\left(\frac{i}{n}, 0\right) = Q_{n,m}\left(0, \frac{j}{m}\right) = 0, \quad i \in \{0, \dots, n\},\ j \in \{0, \dots, m\};$
(b) $Q_{n,m}\left(\frac{i}{n}, 1\right) = \frac{i}{n}, \quad Q_{n,m}\left(1, \frac{j}{m}\right) = \frac{j}{m} \qquad i \in \{0, \dots, n\},\ j \in \{0, \dots, m\};$
(c) $0 \leq Q_{n,m}\left(\frac{i}{n}, \frac{j}{m}\right) - Q_{n,m}\left(\frac{i-1}{n}, \frac{j}{m}\right) \leq \frac{1}{n} \quad i \in \{1, \dots, n\},\ j \in \{1, \dots, m\};$
(d) $0 \leq Q_{n,m}\left(\frac{i}{n}, \frac{j}{m}\right) - Q_{n,m}\left(\frac{i}{n}, \frac{j-1}{m}\right) \leq \frac{1}{m} \quad i \in \{1, \dots, n\},\ j \in \{1, \dots, m\}.$

The discrete quasi–copula $Q_{n,m}$ is said to be proper if at least a pair (i, j) exists such that the inequality

$$Q_{n,m}\left(\frac{i}{n}, \frac{j}{m}\right) + Q_{n,m}\left(\frac{i-1}{n}, \frac{j-1}{m}\right) < Q_{n,m}\left(\frac{i-1}{n}, \frac{j}{m}\right) + Q_{n,m}\left(\frac{i}{n}, \frac{j-1}{m}\right) \tag{13.10}$$

is fullfilled. ◊

Theorem 13.8. *The restriction to $I_{n,m}$ of a 2–quasi–copula Q is a discrete quasi–copula. Moreover, if Q is a copula, then it satisfies eq.* (13.10) *where the inequality sign $<$ is replaced by $\geq$.*

Given a quasi–copula Q on $\mathbb{I}^2$, the discrete quasi–copula $Q_{n,m}$ defined by

$$Q_{n,m}\left(\frac{i}{n}, \frac{j}{m}\right) := Q\left(\frac{i}{n}, \frac{j}{m}\right)$$

will be called the *discretisation of order* $n \times m$ of Q. When $m = n$ it will be denoted by Q_n and called the discretisation of order n.

Theorem 13.9. *Let Q be a 2–quasi–copula and Q_n its discretisation of order n; then*

$$\lim_{n\to+\infty} Q_n\left(\frac{[nu]}{n}, \frac{[nv]}{n}\right) = Q(u,v),$$

where $[x]$ denotes the integral part of x.

In the opposite direction one has

Theorem 13.10. *Let $k \geq 2$ be a natural number and let $(Q_{k^m})_{m\in\mathbb{N},m\geq 2}$ be a sequence of discrete quasi–copulas satisfying the consistency condition*

$$Q_{k^m}\left(\frac{i}{k^{m-1}}, \frac{j}{k^{m-1}}\right) = Q_{k^{m-1}}\left(\frac{i}{k^{m-1}}, \frac{j}{k^{m-1}}\right) \quad i,j \in \{0,\dots,k^{m-1}\}.$$

Then the limit of the sequence

$$\left(Q_{k^m}\left(\frac{[k^m u]}{k^m}, \frac{[k^m v]}{k^m}\right)\right)$$

exists at every point $(u,v) \in \mathbb{I}^2$ and the function $Q : \mathbb{I}^2 \to \mathbb{I}$ defined by

$$Q(u,v) : \lim_{m\to+\infty} Q_{k^m}\left(\frac{[k^m u]}{k^m}, \frac{[k^m v]}{k^m}\right)$$

is a quasi–copula.

Given a discrete (not necessarily proper) quasi–copula Q_n it is possible to construct a whole set of quasi–copulas having Q_n as their discretisation of order n (see [30, Section 4]).

A modification of the transformation matrices studied in [16] can be used in order to construct quasi–copulas ([13]); this will be useful in proving some of the results of sections 13.8 and 13.9.

A matrix $T = (t_{ij})_{i,j=1}^m \in [-1/3,1]^{m\times m}$ will be said to be a *quasi–transformation matrix* if

(a) $\sum_{i,j=1}^m t_{i,j} = 1$;
(b) $\sum_{j=1}^m t_{i,j} > 0$ for every $i \in \{1,\dots,m\}$;
(c) $\sum_{i=1}^m t_{i,j} > 0$ for every $j \in \{1,\dots,m\}$;
(d) the sum of the entries in any submatrix of T that contains at least one entry from the first or the last row or column of T is nonnegative.

A quasi–transformation matrix T is said to be *proper* if at least two indices i and j exist such that $t_{i,j} < 0$.

Given a quasi–transformation matrix T, let p_i (respectively, q_j) denote the sum of the entries of the first i columns (respectively, of first j rows) T with $i, j \in \{1, \dots, m\}$; set $p_0 = q_0 := 0$. Consider the m^2 rectangles $R_{i,j} := [p_{i-1}, p_i] \times [q_{j-1}, q_j]$; they have disjoint interiors and their union is the whole unit square $\mathbb{I}^2$. Thus, a quasi–transformation matrix induces an operator W_T on $\mathscr{Q}_2$ defined for $Q \in \mathscr{Q}_2$ and $(u,v) \in R_{i_0,j_0}$ by

$$W_T(Q)(u,v) := \sum_{i<i_0, j<j_0} t_{i,j} + \frac{u - p_{i_0-1}}{p_{i_0} - p_{i_0-1}} \sum_{j<j_0} t_{i_0,j} + \frac{v - q_{j_0-1}}{q_{j_0} - q_{j_0-1}} \sum_{i<i_0} t_{i,j_0} + t_{i_0,j_0}\, Q\left(\frac{u - p_{i_0-1}}{p_{i_0} - p_{i_0-1}}, \frac{v - q_{j_0-1}}{q_{j_0} - q_{j_0-1}}\right). \tag{13.11}$$

In [13] it is proved that

(a) $W_T(Q)$ is a quasi–copula for every quasi–copula Q, so that $W_T(\mathscr{Q}_2) \subseteq \mathscr{Q}_2$;
(b) W_T is a contraction on $(\mathscr{Q}_2, d_\infty)$;
(c) if T is proper then $W_T(Q)$ is also proper for every $Q \in \mathscr{Q}_2$, so that one has $W_T(\mathscr{Q}_2) \subseteq \mathscr{Q}_2 \setminus \mathscr{C}_2$, when T is proper;
(d) $W_T^n(Q)$ is a proper quasi–copula for every $n \in \mathbb{N}$, if, and only if, Q is proper;
(e) a unique quasi–copula Q_T exists such that $W_T(Q_T) = Q_T$;
(f) for every $Q \in \mathscr{Q}_2$, $\lim_{n\to+\infty} W_T^n(Q) = Q_T$.

13.6 The space of quasi–copulas and its lattice structure

While quasi–copulas were originally introduced in characterising operations on distribution functions that are induced pointwise, they play a considerable role in studying the structure of the set of copulas, and, in time, this aspect of quasi–copulas has become their most important one; to it is devoted the present section.

It is easily checked that $\mathscr{Q}_d$, the space of d-quasi-copulas, is a closed subset of the space $(\Xi(\mathbb{I}^d), d_\infty)$ of continuous functions on $\mathbb{I}^d$ endowed with the topology of uniform convergence.

Theorem 13.11. *The set $\mathscr{Q}_d$ of d-quasi-copulas is a compact and convex subset of $(\Xi(\mathbb{I}^d), d_\infty)$.*

We recall a few definitions from lattice theory (see, e.g., [5]). Given a partial ordered set (=poset) $(P, \leq)$, and two elements x and y in P, $x \vee y$ denotes their *join*, namely their least upper bound when it exists, while $x \wedge y$ denotes their *meet*, namely their greatest lower bound, when it exists. If S is a subset of P, $\vee S$ and $\wedge S$ are defined in the obvious way.

Now one can consider the classical pointwise order among functions in $\mathscr{Q}_d$, namely, for $Q_1, Q_2 \in \mathscr{Q}_d$, $Q_1 \leq Q_2$ if, and only if, $Q_1(\mathbf{u}) \leq Q_2(\mathbf{u})$ for every $\mathbf{u} \in \mathbb{I}^d$. In particular, one may define

$$Q_1 \vee Q_2 := \inf\{Q \in \mathscr{Q}_d : Q_1 \leq Q, Q_2 \leq Q\},$$
$$Q_1 \wedge Q_2 := \sup\{Q \in \mathscr{Q}_d : Q \leq Q_1, Q \leq Q_2\}.$$

When the join, or meet, is found in a particular poset P, one writes $\vee_P S$ and $\wedge_P S$. Given two posets A and B, A is said to be *join–dense* (respectively, *meet–dense*) in B if for every $D \in B$, there exists a set $S \subseteq A$, such that $D = \vee_B S$ (respectively, $D = \wedge_B S$). A poset $P \neq \emptyset$ is said to be a *lattice* if for all x and y in P, both $x \vee y$ and $x \wedge y$ are in P; and P is a *complete lattice* if both $\vee S$ and $\wedge S$ are in P for every subset S of P. If $\varphi : P \to L$ is an order-preserving injection of a poset $(P, \prec)$ into a complete lattice $(L, \prec_1)$, then L is said to be a *completion* of P; in particular, if φ maps P onto L, then it is an *order–isomorphism*.

Definition 13.9. A complete lattice $(L, \prec_1)$ is said to be the *Dedekind–MacNeille completion* of a poset $(P, \prec)$ (also referred to as the *normal completion* or the *completion by cuts*) if $(P, \prec)$ is both join-dense and meet-dense in $(L, \prec_1)$. ◊

Theorem 13.2 states that the upper and lower bounds in $\mathscr{Q}_d$ coincide with the Hoeffding–Fréchet bounds for copulas; however, more can be said. The relevant result is the following theorem (see [27] in the case $d = 2$ and [32] for $d > 2$).

Theorem 13.12. *The set $\mathscr{Q}_d$ of d-quasi-copulas is a complete lattice.*

Since the set $\mathscr{C}_d$ of d-copulas is included in $\mathscr{Q}_d$, one immediately has the following corollary.

Corollary 13.4. *Both the join and the meet of every set of d-copulas are d-quasi-copulas.*

However, proper subsets of $\mathscr{Q}_d$ may not be closed under supremum and infimum operations; two important cases in point are given in the next result.

Theorem 13.13. *For every $d \geq 2$, neither the set $\mathscr{C}_d$ of copulas nor the set $\mathscr{Q}_d \setminus \mathscr{C}_d$ of proper quasi–copulas is a lattice.*

In examining the question of the Dedekind–MacNeille completion of $\mathscr{C}_d$ one has to distinguish the two cases $d = 2$ and $d > 2$. As a preliminary, a few lemmata will be needed.

In the next result, which refers to bivariate copulas, sharper bounds, both upper and lower, are given related to the class of all 2-copulas taking a specified value at a point in the interior of the unit square.

Lemma 13.6. *Let the 2-copula C take the value θ at the point $(a,b) \in\,]0,1[^2$, i.e. $C(a,b) = \theta$, where θ belongs to the interval $[\max\{a+b-1, 0\}, \min\{a,b\}]$. Then, for every (u,v) in $\mathbb{I}^2$,*

$$C_L(u,v) \leq C(u,v) \leq C_U(u,v), \tag{13.12}$$

where C_L and C_U are defined by

$$C_L(u,v) := \begin{cases} \max\{0, u-a+v-b+\theta\}, & (u,v) \in [0,a] \times [0,b], \\ \max\{0, u+v-1, u-a+\theta\}, & (u,v) \in [0,a] \times [b,1], \\ \max\{0, u+v-1, v-b+\theta\}, & (u,v) \in [a,1] \times [0,b], \\ \max\{\theta, u+v-1\}, & (u,v) \in [a,1] \times [b,1], \end{cases}$$

and

$$C_U(u,v) = \begin{cases} \min\{u, v, \theta\}, & (u,v) \in [0,a] \times [0,b], \\ \min\{u, v-b+\theta\}, & (u,v) \in [0,a] \times [b,1], \\ \min\{u-a+\theta, v\}, & (u,v) \in [a,1] \times [0,b], \\ \min\{u, v, u-a+v-b+\theta\}, & (u,v) \in [a,1] \times [b,1], \end{cases}$$

respectively. The bounds in (13.12) *are the best possible.*

As usual set $x^+ := \max\{x, 0\}$.

Lemma 13.7. *For $(a,b) \in]0,1[^2$ and for $\theta \in [W_2(a,b), M_2(a,b)]$ let $S_{a,b,\theta}$ denote the set of quasi–copulas that take the value θ at (a,b),*

$$S_{a,b,\theta} := \{Q \in \mathscr{Q}_2 : Q(a,b) = \theta\}.$$

Then both $\vee S_{a,b,\theta}$ and $\wedge S_{a,b,\theta}$ are copulas and

$$\bigvee S_{a,b,\theta}(u,v) = \min\{M_2(u,v), \theta + (u-a)^+ + (v-b)^+\},$$
$$\bigwedge S_{a,b,\theta}(u,v) = \max\{W_2(u,v), \theta - (a-u)^+ - (b-v)^+\}.$$

Lemma 13.8. *For every quasi-copula $Q \in \mathscr{Q}_2$, one has $Q = \vee S_1(Q)$, where $S_1(Q) := \{C \in \mathscr{C}_2 : C \leq Q\}$, and $Q = \wedge S_2(Q)$, where $S_2(Q) := \{C \in \mathscr{C}_2 : C \geq Q\}$.*

In particular, it follows that 2-quasi-copulas can be characterised in terms of copulas. The following result is now obvious.

Lemma 13.9. *The set of bivariate copulas $\mathscr{C}_2$ is both join-dense and meet-dense in $\mathscr{Q}_2$.*

Theorem 13.14. *The complete lattice $\mathscr{Q}_2$ of 2–quasi–copulas is order–isomorphic to the Dedekind–MacNeille completion of the poset $\mathscr{C}_2$ of bivariate copulas.*

The sets $\mathscr{C}_2$ of bivariate copulas and $\mathscr{C}_d$ with $d > 2$ of multivariate copulas differ with respect to their Dedekind–MacNeille completion: in fact, while $\mathscr{Q}_2$ is the Dedekind–MacNeille completion of $\mathscr{C}_2$, $\mathscr{Q}_d$ is not the Dedekind–MacNeille completion of $\mathscr{C}_d$ for $d > 2$ (see [25] and [12]).

Theorem 13.15. *For $d > 2$, the complete lattice $\mathscr{Q}_d$ of d–quasi–copulas is not order–isomorphic to the Dedekind–MacNeille completion of the poset $\mathscr{C}_d$ of d–copulas.*

Proof. It is enough to consider that W_d is a proper quasi-copula and that, because of Theorem 13.2, W_d cannot be the upper bound of any set of d–copulas. □

13.7 Best bounds for quasi–copulas with a given diagonal

It follows from the results of the previous section that quasi–copulas are likely to appear whenever bounds for specific sets of copulas are considered. We provide one further example: consider the set of 2–copulas with a given diagonal. We recall that a function $\delta : \mathbb{I} \to \mathbb{I}$ is the diagonal section of a 2–copula C if, and only if, it satisfies the following properties: (a) $\delta(0) = 0$ and $\delta(1) = 1$, (b) $\delta(t) \le t$ for every $t \in \mathbb{I}$, (c) δ is increasing and (d) δ satisfies the Lipschitz condition $|\delta(t') - \delta(t)| \le 2\,|t' - t|$ for all t and t' in $\mathbb{I}$.

Theorem 13.16. *For every diagonal δ the function $A_\delta : \mathbb{I}^2 \to \mathbb{I}$ defined by*

$$A_\delta(u,v) := \min\Big\{u, v, \max\{u,v\} - \max\{\hat{\delta}(t) : t \in [u \wedge v, u \vee v]\}\Big\} \tag{13.13}$$

$$= \begin{cases} \min\left\{u, v - \max_{t\in[u,v]}\{t - \delta(t)\}\right\}, & u < v, \\ \min\left\{v, u - \max_{t\in[v,u]}\{t - \delta(t)\}\right\}, & v \le u, \end{cases}$$

is a symmetric 2–quasi–copula having diagonal equal to δ.

While A_δ is in general a proper quasi–copula, Fernández Sánchez and Trutschnig [14] have recently given necessary and sufficient conditions on the graph of δ that ensure that it is a copula (see also [35] and [27, 20]).

Once A_δ has been introduced, the following holds

Theorem 13.17. *For every diagonal δ and for every quasi-copula Q having δ as its diagonal, viz. $\delta(t) = Q(t,t)$ for every $t \in \mathbb{I}$, one has*

$$C_\delta^{\mathbf{Ber}} \le Q \le A_\delta\,, \tag{13.14}$$

where $C_\delta^{\mathbf{Ber}}$ is the Bertino copula defined by

$$C_\delta^{\mathbf{Ber}}(u,v) := \min\{u,v\} - \min\{\hat{\delta}(t) : t \in [u \wedge v, u \vee v]\}$$

$$= \begin{cases} u - \min_{t\in[u,v]}\{t - \delta(t)\}, & u \le v, \\ v - \min_{t\in[v,u]}\{t - \delta(t)\}, & v < u, \end{cases}$$

The Bertino's copula was introduced by Fredricks and Nelsen [15] following a paper by Bertino [3].

13.8 Mass distribution

A d–copula C generates a d–fold stochastic measure on the Borel sets of $\mathbb{I}^d$: if $R = \mathbb{I}^d \cap \prod_{j=1}^d [a_j, b_j[$ is a rectangle in $\mathbb{I}^d$, then one defines

$$\mu_C(R) := V(R) = \sum_{\mathbf{v}\in \mathrm{ver}(]\mathbf{a},\mathbf{b}])} \mathrm{sign}(\mathbf{v})\, C(\mathbf{v}), \tag{13.15}$$

where

$$\mathrm{sign}(\mathbf{v}) = \begin{cases} 1, & \text{if } v_j = a_j \text{ for an even number of indices,} \\ -1, & \text{if } v_j = a_j \text{ for an odd number of indices,} \end{cases}$$

Through the usual techniques of measure theory the above definition of μ_C is then extended to the set $\mathscr{B}(\mathbb{I}^d)$ of Borel sets (see, e.g., [10]). This measure is d–fold stochastic in the sense that, for every $A \in \mathscr{B}(\mathbb{I})$ and for every $j \in \{1, \dots, d\}$,

$$\mu(\underbrace{\mathbb{I} \times \cdots \times \mathbb{I}}_{j-1} \times A \times \mathbb{I} \times \cdots \times \mathbb{I}) = \lambda(A), \tag{13.16}$$

where λ denotes the (restriction to $\mathscr{B}(\mathbb{I})$ of the) Lebesgue measure.

Perhaps naively, it was expected, in particular by the present author, that something similar would happen for a quasi–copula Q.

In the case of a quasi–copula Q repeating the same construction leads to the expression

$$\mu_Q(S) := \sum_{i=1}^n \mu_Q(R_i),$$

where $S = \cup_{i=1}^n R_i$, and the interiors of the rectangles R_i do not overlap. However, this only defines a *finitely* additive set function μ_Q on the ring of $\mathscr{R}$ of finite disjoint unions of rectangles. One may therefore wonder whether μ_Q can be extended to a real measure on the Borel sets of $\mathbb{I}^d$; we recall that a *real measure*, often called a *signed measure*, $\mu : \mathscr{F} \to \overline{\mathbb{R}}$ is a σ–additive set function defined on the measurable space $(\Omega, \mathscr{F})$ with the condition that $\mu(\emptyset) = 0$ and μ assumes at most one of the values ∞ and $-\infty$ (see, e.g., [18]). Equivalently, μ is the difference between two (positive) measures μ_1 and μ_2 (defined on the same measure space), such that at least one of them is finite. As a consequence of the definition of quasi–copulas, if Q induces a real measure μ_Q, this measure should be d–fold stochastic. In particular, any d-fold stochastic real measure μ on $\mathscr{B}(\mathbb{I}^d)$ should satisfy $\mu(\mathbb{I}^d) = 1$ and, in view of [18, page 119], $|\mu(E)| < \infty$ for every $E \in \mathscr{B}(\mathbb{I}^d)$.

It came as a surprise that the extension mentioned in the previous paragraph is not possible, as will be seen.

The following result provides information about the bounds of $V_Q(R)$.

Theorem 13.18. *For every 2-quasi-copula Q and for every rectangle R contained in $\mathbb{I}^2$, one has*

$$-\frac{1}{3} \leq V_Q(R) \leq 1\,.$$

Moreover, $V_Q(R) = 1$ if, and only if, $R = \mathbb{I}^2$ and $R = [1/3, 2/3]^2$ when $V_Q(R) = -1/3$.

Notice that, for $d = 2$, a unique rectangle exists on which the minimal mass (which turns out to be $-1/3$) can be spread, as well as a unique rectangle (the unit square itself) on which the maximal mass 1 can be accumulated. The situation is slightly different in higher dimensions. For the case $d = 3$, De Baets, De Meyer and Úbeda–Flores [6] showed that there still exists a unique 3–rectangle on which the minimal mass (which now turns out to be $-4/5$) can be spread, while there exist multiple 3–rectangles on which the maximal mass 1 can be spread. In principle, the methodology used in the proof may be applied to the case $d > 3$ as well (apart from an increasing complexity of the formulation).

In general, the area of rectangles contained in $\mathbb{I}^2$ with given Q_2–volume is subject to specific bounds.

Theorem 13.19. *Let $R = [u_1, u_2] \times [v_1, v_2]$ be a rectangle contained in $\mathbb{I}^2$ and let θ be in $[-1/3, 1]$. If there is a 2-quasi-copula Q for which $V_Q(R) = \theta$, then the area $A(R)$ of R satisfies*

$$\theta^2 \leq A(R) \leq \left(\frac{1+\theta}{2}\right)^2. \tag{13.17}$$

Moreover, $A(R)$ attains both bounds; in both cases R is necessarily a square.

It is important to note that a quasi–copula may have a negative mass as large as one wishes. The following result was first proved for the independence (quasi)–copula Π_2 and then extended to any quasi–copula.

Theorem 13.20. *For all given $\varepsilon > 0$ and $H > 0$, and for every 2–quasi–copula $\widetilde{Q}$, there exist a 2–quasi–copula Q and a subset S of $\mathbb{I}^2$ such that*

(a) $\mu_Q(S) < -H$;
(b) *for all u and v in $\mathbb{I}$, $|Q(u,v) - \widetilde{Q}(u,v)| < \varepsilon$.*

The fact that a quasi–copula need not induce a real measure on the Borel sets $\mathscr{B}(\mathbb{I}^2)$ of $\mathbb{I}^2$ is established below (see the papers [23], [28], [13] and [11]).

Theorem 13.21. *A 2-quasi-copula Q exists that does not induce a doubly stochastic real measure on $\mathbb{I}^2$.*

In general, i.e., for $d \geq 3$, one has

Theorem 13.22. *For every $d \geq 3$, W_d does not induce a d-fold stochastic real measure on $\mathbb{I}^d$.*

These results will be further analysed in the next section.

13.9 Category results

It has to be recalled that in the very first paper in which quasi–copulas were introduced it was conjectured that no absolutely continuous[3] proper quasi–copula existed, a conjecture proved wrong in [17] by using a result in Rodríguez Lallena's Ph.D. thesis [31]. Hence the importance of the recent extension [9] by Durante, Fernández–Sánchez and Trutschnig of their category results for copulas [7, 8] to the setting of bivariate quasi–copulas. Some of the results of [9] are summarised below.

Let $\mathscr{Q}_{2,R}$ denote the subset of $\mathscr{Q}_2$, for which an extension to a real measure, in the sense of the previous section, exists and let $\mathscr{Q}_{2,Rac} \subseteq \mathscr{Q}_{2,R}$ denote the subset of those quasi–copulas for which the associated real measure is absolutely continuous.

First of all, one has the following result.

Theorem 13.23. *The sets $\mathscr{Q}_{2,R}$, $\mathscr{Q}_{2,R}^c$ and $\mathscr{Q}_{2,Rac}$ are dense in $(\mathscr{Q}_2, d_\infty)$.*

We recall that a subset N of a metric space (Ω, d) is said to be *nowhere dense* if the interior of its closure is empty. A subset A of Ω is said to be of *first category* or *meager* if it is the countable union of a sequence of nowhere dense sets; otherwise A will be said to be of *second category*.

Theorem 13.24. *The set $\mathscr{C}_2$ of bivariate copulas is nowhere dense in $(\mathscr{Q}_2, d_\infty)$.*

This result implies that, in general, a quasi–copula is proper. A category argument allows to sharpen the result quoted above about the connexion between quasi–copulas and real measures. To this end, a quasi–copula Q will be said to be *locally extendable* when a point $\mathbf{u}$ in $\mathbb{I}^2$ and $\rho > 0$ exist such that V_Q of eq. (13.15) can be extended to a real measure on $\mathscr{B}(\mathbb{I}^2) \cap B(\mathbf{u}, \rho)$. The family of locally extendable quasi–copulas will be denoted by $\mathscr{Q}_{2,R}^{\text{loc}}$. Such quasi–copulas exist, see, for instance, [9, Example 2.1].

Theorem 13.25. *Both the sets $\mathscr{Q}_{2,R}$ and $\mathscr{Q}_{2,R}^{\text{loc}}$ are of first category in $(\mathscr{Q}_2, d_\infty)$.*

As a consequence, a typical quasi–copula cannot be associated with a doubly stochastic real measure on $\mathscr{B}(\mathbb{I}^2)$, not even locally.

13.10 Final comments

It is only appropriate to look at the list of open problems listed in Nelsen's survey [23] of 2005; however, there he dealt mainly with copulas and their relationships with triangular norms and the open problems he listed were questions about

[3] In view of the fact that in general the finitely additive measure induced by a quasi–copula cannot be extended to a real measure on the the Borel subsets $\mathscr{B}(\mathbb{I}^d)$, it is improper to speak of an absolutely continuous proper quasi–copula; here we continue to use this terminology, although more correctly one should speak of a quasi–copula Q in $\mathscr{Q}_2$, for instance, such that $D_2 D_1 Q = 0$ a.e..

copulas. But, of course, new questions may arise. The first, and most natural, of these concerns the search for best bounds, upper and lower, which are necessarily quasi–copulas, in all those instances in which one has to consider sets of copulas, either bivariate or multivariate, with special properties. This question will necessarily presents itself again whenever a new set of copulas is deemed important enough to be studied.

The category results of [9] will have to be extended to the case of d–dimensional quasi–copulas.

Finally, attention ought to be payed to the use of quasi–copulas in the construction of triangle functions (see [33]).

Acknowledgements I wish to thank Roger Nelsen, José Juan Quesada Molina, José Antonio Rodríguez Lallena, Manuel Úbeda Flores, Fabrizio Durante, Juan Fernández Sánchez, Christian Genest and Wolfgang Trutschnig with whom through the years I have had the chance to talk about quasi–copulas, and from whom I have learned much of what I wrote above.

This work has been also supported by the Ministerio de Economía y Competitividad (Spain) under research project MTM2014-60594-P and partially supported by FEDER.

References

1. Alsina, C., Nelsen, R.B., Schweizer, B.: On the characterization of a class of binary operations on distribution functions. Statist. Probab. Lett. **17**(2), 85–89 (1993)
2. Alsina, C., Schweizer, B.: Mixtures are not derivable. Found. Phys. Lett. **1**, 171–174 (1988)
3. Bertino, S.: Sulla dissomiglianza tra mutabili cicliche. Metron **35**, 53–88 (1977). In Italian
4. Cuculescu, I., Theodorescu, R.: Copulas: diagonals, tracks. Rev. Roumaine Math. Pures Appl. **46**, 731–742 (2001)
5. Davey, B.A., Priestley, H.A.: Introduction to lattices and order. Cambridge University Press (2002)
6. De Baets, B., De Meyer, H., Úbeda Flores, M.: Extremes of the mass distribution associated with a trivariate quasi–copula. C. R. Math. Acad. Sci. Paris **344**, 587–590 (2007)
7. Durante, F., Fernández-Sánchez, J., Trutschnig, W.: A typical copula is singular. J. Math. Anal. Appl. **430**, 517–527 (2015)
8. Durante, F., Fernández-Sánchez, J., Trutschnig, W.: Baire category results for exchangeable copulas. Fuzzy Sets and Systems **284**, 146–151 (2016)
9. Durante, F., Fernández-Sánchez, J., Trutschnig, W.: Baire category results for quasi–copulas. Depend. Model. **4**, 215–233 (2016)
10. Durante, F., Sempi, C.: Principles of copula theory. CRC Press, Boca Raton, FL (2016)
11. Fernández-Sánchez, J., Úbeda Flores, M.: A note on quasi-copulas and signed measures. Fuzzy Sets and Systems **234**, 109–112 (2014)
12. Fernández-Sánchez, J., Nelsen, R.B., Úbeda Flores, M.: Multivariate copulas, quasi–copulas and lattices. Stat. Probab. Lett. **81**, 1365–1369 (2011)
13. Fernández-Sánchez, J., Rodríguez-Lallena, J.A., Úbeda Flores, M.: Bivariate quasi–copulas and doubly stochastic signed measures. Fuzzy Sets and Systems **168**, 81–88 (2011)
14. Fernández-Sánchez, J., Trutschnig, W.: Some members of the class of (quasi-)copulas with given diagonal from the Markov kernel perspective. Comm. Statist. Theory Methods **45**, 1508–1526 (2015)
15. Fredricks, G.A., Nelsen, R.B.: The Bertino family of copulas. In: C.M. Cuadras, J. Fortiana, J.A. Rodríguez Lallena (eds.) Distributions with given marginals and statistical modelling, pp. 81–91. Kluwer, Dordrecht (2003)

16. Fredricks, G.A., Nelsen, R.B., Rodríguez-Lallena, J.A.: Copulas with fractal support. Insur. Math. Econom. **37**, 42–48 (2005)
17. Genest, C., Quesada Molina, J.J., Rodríguez Lallena, J.A., Sempi, C.: A characterization of quasi-copulas. J. Multivariate Anal. **69**, 193–205 (1999)
18. Halmos, P.R.: Measure theory. Springer (1974)
19. Kimberling, C.H.: A probabilistic interpretation of complete monotonicity. Aequationes Math. **10**, 152–164 (1974)
20. Klement, E.P., Kolesárová, A.: Intervals of 1-Lipschitz aggregation operators, quasi-copulas, and copulas with given affine section. Monatsh. Math. **152**(2), 151–167 (2007)
21. McNeil, A.J., Nešlehová, J.: Multivariate Archimedean copulas, d–monotone functions and l_1–norm symmetric distributions. Ann. Statist. **37**, 3059–3097 (2009)
22. Nelsen, R.B.: An introduction to copulas, *Lecture Notes in Statistics*, vol. 139. Springer, New York (1999). 2nd edn., Springer Series in Statistics, 2006
23. Nelsen, R.B.: Copulas and quasi–copulas: an introduction to their properties and applications. In: E.P. Klement, R. Mesiar (eds.) Logical, algebraic analytic, and probabilistic aspects of triangular norms, pp. 391–413. Elsevier, New York (2005)
24. Nelsen, R.B.: An introduction to copulas, *Springer Series in Statistics*, vol. 139. Springer, New York (2006). 2nd edn.
25. Nelsen, R.B., Úbeda Flores, M.: The lattice-theoretic structure of sets of bivariate copulas and quasi–copulas. C. R. Acad. Sci. Paris Sér. I **341**, 583–586 (2005)
26. Nelsen, R.B., Quesada Molina, J.J., Rodríguez Lallena, J.A., Úbeda Flores, M.: Multivariate Archimedean quasi–copulas. In: C.M. Cuadras, J. Fortiana, J.A. Rodríguez Lallena (eds.) Distributions with given marginals and statistical modelling, pp. 179–185. Kluwer, Dordrecht (2002)
27. Nelsen, R.B., Quesada Molina, J.J., Rodríguez Lallena, J.A., Úbeda Flores, M.: Best-possible bounds on sets of bivariate distribution functions. J. Multivariate Anal. **90**, 348–358 (2004)
28. Nelsen, R.B., Quesada Molina, J.J., Rodríguez Lallena, J.A., Úbeda Flores, M.: Quasi-copulas and signed measures. Fuzzy Sets and Systems **161**, 2328–2336 (2010)
29. Nelsen, R.B., Quesada Molina, J.J., Schweizer, B., Sempi, C.: Derivability of some operations on distribution functions. In: L. Rüschendorf, B. Schweizer, M.D. Taylor (eds.) Distributions with fixed marginals and related topics, *Lecture Notes–Monograph Series*, vol. 28, pp. 233–243. Institute of Mathematical Statistics, Hayward, CA (1996)
30. Quesada Molina, J., Sempi, C.: Discrete quasi–copulæ. Insurance: Math. and Econom. **37**, 27–41 (2005). Doi: 10.1016/j.insmatheco.2005.01.007
31. Rodríguez-Lallena, J.A.: Estudio de la compatibilidad y diseño de nuevas familias en la teoría de cópulas. Aplicaciones. Ph.D. thesis, Universidad de Granada, Spain (1993). Published in electronic form in 1996 by Servicio de Publicaciones de la Universidad de Almería.
32. Rodríguez-Lallena, J.A., Úbeda Flores, M.: Best-possible bounds on sets of multivariate distribution functions. Comm. Statist. Theory Methods **33**, 805–820 (2004)
33. Saminger-Platz, S., Sempi, C.: A primer on triangle functions. I. Aequationes Math. **76**, 201–240 (2008)
34. Schweizer, B., Sklar, A.: Operations on distribution functions not derivable from operations on random variables. Studia Math. **52**, 43–52 (1974)
35. Úbeda Flores, M.: On the best-possible upper bound on sets of copulas with given diagonal sections. Soft Comput. **12**, 1019–1025 (2008)

Chapter 14
Complete dependence everywhere?

Wolfgang Trutschnig

Abstract Describing the situation of full predictability of a random variable Y given the value of another random variable X, the notion of complete dependence might seem far too restrictive to be of any practical importance at first glance. Recent results have shown, however, that complete dependence naturally appears in various settings. This chapter will therefore sketch some problems related to complete dependence. Doing so, the focus will be on dependence measures strictly separating extreme dependence concepts, on a problem related to joint-default maximization, on a question from uniform distribution theory, and on the relationship between the two most well-known measures of concordance, Kendall's τ and Spearman's ρ. A short excursion to topology showing that complete dependence is not atypical at all complements the chapter.

14.1 Introduction

Given two random variables X, Y we call Y completely dependent on X if there exists a measurable function f such that $Y = f(X)$ holds with probability one (see [21] for the original definition). In other words: Knowing X means knowing Y but not necessarily vice versa. Although a dependence structure describing full predictability seems very pathological at first, research in the field of dependence modeling conducted in the last years clearly points in the direction that complete dependence is much more important than reflected by textbooks so far. Main objective of this chapter is to illustrate this observation by means of some fairly recent results.

The rest of this chapter is organized as follows: Section 2 gathers notation and preliminaries that will be used in the sequel and states various properties equivalent to complete dependence. Section 3 recalls one possible way to construct metrics that

Wolfgang Trutschnig
Department for Mathematics, University Salzburg, Hellbrunnerstrasse 34, 5020 Salzburg, Austria
e-mail: wolfgang@trutschnig.net

M. Úbeda Flores et al. (eds.), *Copulas and Dependence Models with Applications*,
DOI 10.1007/978-3-319-64221-5_14

clearly distinguish the two extremal dependence concepts, that of complete dependence and that of independence. Based on these metrics a new dependence measure ζ_2 is introduced and analyzed, and alternative approaches from the literature are mentioned. Section 4 first sketches why extreme points naturally comes into play in the context of optimization problems and then shows two concrete examples in both of which completely dependent copulas constitute solutions of the maximization problem. Section 5 contains a short excursion to topology showing that (in the language of Baire categories) with respect to weak convergence complete dependence is all but an atypical property of a copula. Finally, Section 6 focuses on recently established sharp inequalities between Kendall's τ and Spearman's ρ and points out that mutually completely dependent random variable cover all possible constellations of τ and ρ.

14.2 Notation and preliminaries

In the sequel $\mathscr{C}$ will denote the family of all two-dimensional *copulas*, $\mathscr{P}_{\mathscr{C}}$ the family of all *doubly stochastic measures*, i.e. the family of all probability measures on $[0,1]^2$ whose marginals are uniformly distributed on $[0,1]$. M will denote the lower Fréchet-Hoeffding bound, W the lower Fréchet-Hoeffding bound and Π the product copula; for background on copulas we refer to [7] and [25]. For every $C \in \mathscr{C}$ the corresponding doubly stochastic measure will be denoted by μ_C. Letting d_∞ denote the uniform metric on $\mathscr{C}$ it is well known that $(\mathscr{C}, d_\infty)$ is a compact metric space.

For every metric space (Ω, d) the Borel σ-field on Ω will be denoted by $\mathscr{B}(\Omega)$, δ_x will denote the Dirac measure (concentrated) at $x \in \Omega$. λ and λ_2 will denote the Lebesgue measure on $\mathscr{B}(\mathbb{R})$ and $\mathscr{B}(\mathbb{R}^2)$ respectively. For every probability measure ν on $\mathscr{B}(\Omega)$ the support of ν, i.e. the complement of the union of all open sets U fulfilling $\nu(U) = 0$, will be denoted by $Supp(\nu)$.

Suppose that (Ω_1, d_1) and (Ω_2, d_2) are metric spaces. A *Markov kernel* from Ω_1 to $\mathscr{B}(\Omega_2)$ is a mapping $K : \Omega_1 \times \mathscr{B}(\Omega_2) \to [0,1]$ such that $x \mapsto K(x,B)$ is measurable for every fixed $B \in \mathscr{B}(\Omega_2)$ and $B \mapsto K(x,B)$ is a probability measure for every fixed $x \in \Omega_1$. Given real-valued random variables X, Y on a probability space $(\Omega, \mathscr{A}, \mathscr{P})$, a Markov kernel $K : \mathbb{R} \times \mathscr{B}(\mathbb{R}) \to [0,1]$ is called a *regular conditional distribution of Y given X* if for every $B \in \mathscr{B}(\mathbb{R})$

$$K(X(\omega), B) = \mathbb{E}(\mathbf{1}_B \circ Y | X)(\omega) \tag{14.1}$$

holds $\mathscr{P}$-a.e. It is well known that for each pair (X,Y) of real-valued random variables a regular conditional distribution $K(\cdot,\cdot)$ of Y given X exists, that $K(\cdot,\cdot)$ is unique $\mathscr{P}^X$-a.s. (i.e. unique for $\mathscr{P}^X$-almost all $x \in \mathbb{R}$) and that $K(\cdot,\cdot)$ only depends on $\mathscr{P}^{X \otimes Y}$. Hence, given $A \in \mathscr{C}$ we will denote (a version of) the regular conditional distribution of Y given X by $K_A(\cdot,\cdot)$, directly view it as Markov kernel from $[0,1]$ to $\mathscr{B}([0,1])$, and refer to $K_A(\cdot,\cdot)$ simply as *regular conditional distribution of A* or as

Markov kernel of A. Note that for every $A \in \mathscr{C}$, its regular conditional distribution $K_A(\cdot,\cdot)$, and every Borel set $G \in \mathscr{B}([0,1]^2)$ we have the following *disintegration* (here $G_x := \{y \in [0,1] : (x,y) \in G\}$ denotes the x-section of G for every $x \in [0,1]$)

$$\int_{[0,1]} K_A(x, G_x)\, d\lambda(x) = \mu_A(G), \tag{14.2}$$

so in particular

$$\int_{[0,1]} K_A(x, F)\, d\lambda(x) = \lambda(F) \tag{14.3}$$

for every $F \in \mathscr{B}([0,1])$. On the other hand, every Markov kernel $K : [0,1] \times \mathscr{B}([0,1]) \to [0,1]$ fulfilling (14.3) induces a unique element $\mu \in \mathscr{P}_{\mathscr{C}}([0,1]^2)$ via (14.2). For every $A \in \mathscr{C}$ and $x \in [0,1]$ the function $y \mapsto F_x^A(y) := K_A(x,[0,y])$ will be called *conditional distribution function of A at x*. For more details and properties of conditional expectation, regular conditional distributions, and general disintegration see [15] and [18].

Viewing copulas as special Markov kernels (fulfilling eq. (14.3)) has proved surprisingly useful in the past. As an example, translating the so-called $*$-product of copulas introduced in [3] to the Markov kernel setting directly yields that the Markov kernel K_{A*B} of $A * B$ is nothing else but the standard composition of the Markov kernels K_A and K_B as well known in the context Markov processes, i.e.

$$K(x,F) := \int_{[0,1]} K_B(y,F) K_A(x,dy) \tag{14.4}$$

is a Markov kernel of $A * B$ (see [36]). For additional examples underlining the usefulness of Markov kernels we refer, for instance, to [7, 10, 36, 37] and the references therein.

A copula A is called *completely dependent* (see [21, 35]) if there exists a λ-preserving transformation $h : [0,1] \to [0,1]$ such that $K(x,E) = \mathbf{1}_E(h(x))$ is a Markov kernel of A. Slightly extending [35, Lemma 10] the subsequent characterization of complete dependence can be proved:

Lemma 14.1. *The following assertions are equivalent:*

(d1) A is completely dependent .
(d2) For $\mathscr{P}^{X \otimes Y} = \mu_A$ *the conditional variance* $\mathbb{V}(Y|X=x)$ *of Y given X fulfills* $\mathbb{V}(Y|X=x) = 0$ *for* λ*-a.e.* $x \in [0,1]$.
(d3) For λ*-a.e.* $x \in [0,1]$ *the conditional distribution function* F_x^A *is* $\{0,1\}$*-valued.*
(d4) There exists a λ*-preserving transformation* $h : [0,1] \to [0,1]$ *such that* $A(x,y) = \lambda\big([0,x] \cap h^{-1}([0,y])\big)$ *for all* $(x,y) \in [0,1]^2$.
(d5) There exists a λ*-preserving transformation* $h : [0,1] \to [0,1]$ *with* $\mu_A(\Gamma(h)) = 1$, *whereby* $\Gamma(h) = \{(x,h(x)) : x \in [0,1]\} \in \mathscr{B}([0,1]^2)$ *denotes the graph of h.*
(d6) A is left invertible w.r.t. the star product, i.e. there exists a copula $B \in \mathscr{C}$ *with* $B * A = M$.

In the sequel $\mathscr{T}$ will denote the family of all λ-preserving transformations h on $[0,1]$, $\mathscr{T}_b$ the subclass of all λ-preserving bijections, and $\mathscr{T}_s$ the subclass of all

λ-preserving, piecewise linear bijections. For every $h \in \mathscr{T}$ we will let A_h denote the corresponding completely dependent copula, $\mathscr{C}_d$ will denote the family of all completely dependent copulas. In case of $h \in \mathscr{T}_s$ we will refer to A_h as a (classical) *shuffle of M* and in case of $h \in \mathscr{T}_b$ the copula A_h will be called *mutually completely dependent*.

14.3 Quantifying dependence

According to [25] the class of all shuffles of M is dense in the compact metric space $(\mathscr{C}, d_\infty)$. Hence, as pointed out explicitly already in [22], the uniform distance d_∞ is not clearly 'distinguishing different types of statistical dependence' and the same holds for every dependence measure that is continuous w.r.t. d_∞, including Schweizer and Wolffs σ (see [25, 30]). Viewing copulas as Markov kernels allows for a simple way to construct stronger metrics on $\mathscr{C}$ that strictly separate extremal dependence concepts, i.e. that of independence and that of complete dependence. Following [35] and setting

$$D_p^p(A,B) := \int_{[0,1]^2} \big|K_A(x,[0,y]) - K_B(x,[0,y])\big|^p d\lambda_2(x,y) \tag{14.5}$$

defines a metric D_p on $\mathscr{C}$ for every $p \in [1,\infty)$. For a generalization to the multivariate setting we refer to [7, 9]. According to [35] the metrics D_2 and D_1 induce the same topology on $\mathscr{C}$ and the resulting metric spaces $(\mathscr{C}, D_1)$ and $(\mathscr{C}, D_2)$ are complete and separable. It is straightforward to verify that the same assertions hold for $(\mathscr{C}, D_p)$ and arbitrary $p \in [1,\infty)$. In fact, using Hölder inequality and considering $\big|K_A(x,[0,y]) - K_B(x,[0,y])\big| \leq 1$ directly yields

$$D_p^p(A,B) \leq D_1(A,B) \leq D_p(A,B), \tag{14.6}$$

from which separability and completeness of $(\mathscr{C}, D_p)$ directly follows from separability and completeness of $(\mathscr{C}, D_1)$. Although all metrics D_p induce the same topology on $\mathscr{C}$ they are not equivalent - the following result holds:

Lemma 14.2. *For any pair $p,q \in [1,\infty)$ with $p \neq q$ the metrics D_p and D_q are not equivalent.*

Proof: We start by showing that D_p^p and D_q^q coincide on $\mathscr{C}_d$ and consider $h_1, h_2 \in \mathscr{T}$:

$$\begin{aligned} D_p^p(A_{h_1}, A_{h_2}) &= \int_{[0,1]^2} \big|\mathbf{1}_{[0,y]}(h_1(x)) - \mathbf{1}_{[0,y]}(h_2(x))\big|^p d\lambda_2(x,y) = \|h_1 - h_2\|_1 \\ &= \int_{[0,1]^2} \big|\mathbf{1}_{[0,y]}(h_1(x)) - \mathbf{1}_{[0,y]}(h_2(x))\big|^q d\lambda_2(x,y) = D_q^q(A_{h_1}, A_{h_2}) \end{aligned}$$

From this, considering $q = 1$ it also follows that the first inequality in (14.6) can not be improved. For every $n \in \mathbb{N}$ define $h_n \in \mathscr{T}_s$ (see Figure 14.1) by

$$S_n(x) = \begin{cases} x + \left(1 - \frac{1}{2^n}\right) & \text{if } x \in \left(0, \frac{1}{2^n}\right] \\ x - \left(1 - \frac{1}{2^n}\right) & \text{if } x \in \left(1 - \frac{1}{2^n}, 1\right] \\ x & \text{otherwise,} \end{cases}$$

and set $h = id_{[0,1]} \in \mathscr{T}_s$. Then we get

$$D_p^p(A_{h_n}, M) = D_q^q(A_{h_n}, M) = \|h_n - h\|_1 = 2\int_{[0,\frac{1}{2^n}]} \left(1 - \frac{1}{2^n}\right) d\lambda = \frac{1}{2^{n-1}}\left(1 - \frac{1}{2^n}\right). \tag{14.7}$$

Suppose now that $p > q$. Then eq. (14.7) implies that the quotient $\frac{D_p(A_n,M)}{D_q(A_n,M)}$ is unbounded in n, so D_q and D_p can not be equivalent metrics. ■

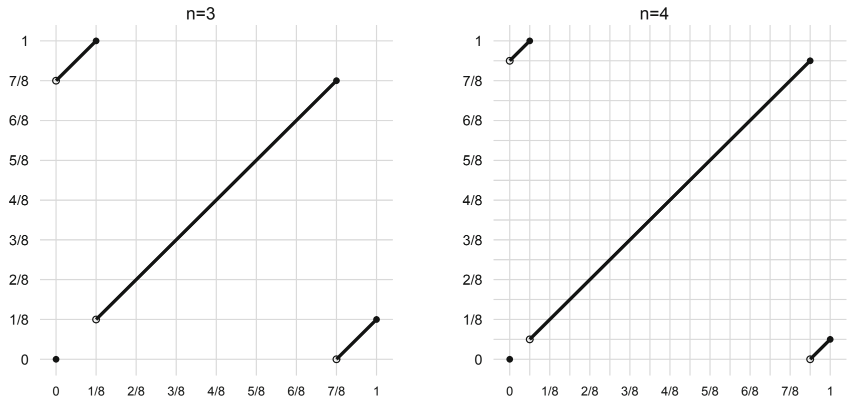


Fig. 14.1: The transformations h_n used in the proof of Lemma 14.2

In [35] the metric D_1 mainly served as a vehicle to construct the dependence measure $\zeta_1(A) = 3D_1(A, \Pi)$ for every $A \in \mathscr{C}$. The most important properties of ζ_1 are summarized in the following theorem.

Theorem 14.1 ([35]). *For every $A \in \mathscr{C}$ we have $\zeta_1(A) \in [0,1]$. Additionally, $\zeta_1(A) = 1$ holds if and only if $A \in \mathscr{C}_d$, and $\zeta_1(A) = 0$ implies $A = \Pi$. In other words: Exclusively all completely dependent copulas are assigned maximum dependence measure and the product copula is the only copula with zero dependence.*

Taking into account eq. (14.6) a similar result can be expected for the dependence measure ζ_p, defined by $\zeta_p(A) = c_p D_p(A, \Pi)$ for every $p \in [1, \infty)$. Thereby c_p is a normalizing constant assuring $\max_{A\in\mathscr{C}} \zeta_p(A) = 1$. In the sequel we will state and prove the result for the case $p = 2$ which allows for a more elegant proof than the original one for D_1.

Theorem 14.2. *For every $A \in \mathscr{C}$ we have $D_2^2(A,\Pi) \leq 1/6$ with equality if and only if $A \in \mathscr{C}_d$.*

Proof: Fix $A \in \mathscr{C}$ and $y \in [0,1]$, and define a random variable Z_y on the probability space $([0,1], \mathscr{B}([0,1], \lambda) \to [0,1]$ by

$$Z_y(x) := K_A(x,[0,y]).$$

Then

$$\mathbb{E}(Z_y) = \int_{[0,1]} K_A(x,[0,y])d\lambda(x) = \mu_A\big([0,1]\times[0,y]\big) = y$$

holds. Considering

$$\begin{aligned}\int_{[0,1]} \big(K_A(x,[0,y]) - y\big)^2 d\lambda(x) &= \mathbb{V}(Z_y) = \mathbb{E}(Z_y^2) - (\mathbb{E}(Z_y))^2 = \mathbb{E}(Z_y^2) - y^2 \\ &\leq \mathbb{E}(Z_y) - y^2 = y - y^2 \end{aligned} \tag{14.8}$$

we directly get $D_2^2(A,\Pi) \leq \int_{[0,1]}(y-y^2)d\lambda(y) = \frac{1}{6}$, which completes the proof of the first assertion.

Since ineq. (14.8) becomes an equality if and only if $\mathbb{E}(Z_y^2) = \mathbb{E}(Z_y)$ holds, it follows that $D_2^2(A,\Pi) = \frac{1}{6}$ is equivalent to the condition that $Z_y^2 = Z_y$ holds λ-a.e. The latter, however, is obviously equivalent to the existence of a set $\Lambda_y \in \mathscr{B}([0,1])$ with $\lambda(\Lambda_y) = 1$ such that $Z_y(x) = K_A(x,[0,y]) \in \{0,1\}$ for every $x \in \Lambda_y$.

Suppose now that $D_2^2(A,\Pi) = 1/6$. Repeating the last argument we can find a set $\Lambda \in \mathscr{B}([0,1])$ fulfilling $\lambda(\Lambda) = 1$ such that $F_x^A(y) = K_A(x,[0,y]) \in \{0,1\}$ holds for every $x \in \Lambda$ and every $y \in \mathbb{Q} \cap [0,1]$. Using right-continuity of distribution functions we immediately get that λ-a.e. conditional distribution functions F_x^A are $\{0,1\}$-valued, so Lemma 14.1 implies that A is completely dependent. Since, on the other hand, it is straightforward to verify $D_2^2(A_h,\Pi) = \frac{1}{6}$ for every $h \in \mathscr{T}$, the proof is complete. ■

As direct consequence of Theorem 14.1, setting $\zeta_2(A) = \sqrt{6}D_2(A,\Pi)$ we get the following result:

Proposition 14.1. *For every $A \in \mathscr{C}$ we have $\zeta_2(A) \in [0,1]$. Additionally, $\zeta_2(A) = 1$ holds if and only if $A \in \mathscr{C}_d$ and $\zeta_2(A) = 0$ implies $A = \Pi$. In other words: Exclusively all completely dependent copulas are assigned maximum ζ_2-value and the product copula is the only copula with $\zeta_2(A) = 0$.*

Independence of two random variables X,Y is a symmetric concept (knowing X does not change our knowledge about Y and vice versa) - nevertheless, from the author's point of view, notions quantifying dependence should not automatically be symmetric since in many situations one might also be interested in understanding causal effects between X and Y and the dependence structure might be strongly asymmetric. The latter is the case, for instance, for the copula $A_h \in \mathscr{C}_d$ with $h \in \mathscr{T}$ being the transformation $h(x) = 2^n (mod 1)$ for large $n \in \mathbb{N}$. Furthermore, having a unidirectional (i.e. non-mutual) dependence measure one can easily construct a mutual one: The mutual dependence measure ω studied by Siburg and Stoimenov (see [33]), for instance, can easily be expressed in terms of ζ_2 as

$$\omega^2(A) = 3\left(\zeta_2^2(A) + \zeta_2^2(A^t)\right), \tag{14.9}$$

whereby A^t denotes the transpose of A, defined by $A^t(x,y) = A(x,y)$. Proposition 14.1 directly yields that $\omega(A) = 1$ if and only if both A and A^t are completely dependent, i.e. if and only if A is mutually completely dependent.

Recently, dependence measures for (absolutely continuous) random vectors have been introduced using similar ideas as the afore-mentioned ones, see [2] and the references therein. For a dependence measure based on conditional variance we refer to [16].

14.4 Complete dependence in the context of optimization

Remember that a point x in a convex set Ω is called an extreme point of Ω if it is not an interior point of any line segment lying entirely in Ω, i.e. if $x = \alpha y + (1-\alpha)z$ for $y,z \in \Omega$ and $\alpha \in [0,1]$ implies $x = y$ or $x = z$.

It is straightforward to show that every completely dependent copula is an extreme point of $\mathscr{C}$. As a consequence, in the metric space $(\mathscr{C}, d_\infty)$ the set $Ex(\mathscr{C})$ of all extreme points of $\mathscr{C}$ is dense. Although a full and handy characterization of the set $Ex(\mathscr{C})$ seems out of reach it is known, that there are extreme points which are not completely dependent. In fact, in [31] (also see [26, 32]) so-called hairpin copulas, which concentrate their mass on the graphs of two functions were studied and shown to be elements of $Ex(\mathscr{C})$. For the generalization of hairpin copulas to the multivariate setting we refer to [5]. To the best of the author's knowledge the most striking example of an extreme point of $\mathscr{C}$ was given in [23], where the author proved the existence of a copula $A \in \mathscr{C}$ such that $Supp(\mu_A) = [0,1]^2$.

Extreme points of $\mathscr{C}$ naturally come into play in the context of optimization problems of the form

$$\overline{M}_H := \sup_{A \in \mathscr{C}} \int_{[0,1]^2} H(x,y)\, d\mu_A(x,y) \tag{14.10}$$

whereby H is a non-negative measurable function on $[0,1]^2$. In fact, if there is a unique $A \in \mathscr{C}$ attaining $\overline{M}_H$ then A has to be an extreme point of $\mathscr{C}$. Additionally, if H is continuous (hence bounded) then $\overline{M}_H$ is attained and, according to the Bauer Maximum Principle (see [1]), the maximum is also attained by an extreme point.

14.4.1 Distributions with fixed marginals maximizing the mass of the endograph of a function

Suppose that F and G are (continuous) distribution functions of two default times and let $\mathscr{F}_{F,G}$ denote the Fréchet class of F,G (i.e. the family of all two-dimensional distribution functions having F and G as marginals). It is well known from coupling theory (see [34]) that there exists a maximal coupling, i.e. a two-dimensional distribution function $H \in \mathscr{F}_{F,G}$ such that for the case of $(X,Y) \sim H$ the probability of a joint default $\mathbb{P}(X=Y)$ is maximal. Translating to the class of copulas maximizing the probability of a joint default means calculating $\sup_{A\in\mathscr{C}} \mu_A(\Gamma(T))$ for $T:[0,1]\to[0,1]$ being defined by $T = G\circ F^-$, F^- denoting the quasi-inverse of F and $\Gamma(T)$ the graph of T. Using coupling theory we can find a (not necessarily unique) copula A_0 with

$$\overline{M}_{\mathbf{1}_{\Gamma(T)}} = \sup_{A\in\mathscr{C}} \int_{[0,1]^2} \mathbf{1}_{\Gamma(T)} d\mu_A(x,y) = \sup_{A\in\mathscr{C}} \mu_A(\Gamma(T)) = \mu_{A_0}(\Gamma(T)) \tag{14.11}$$

that can even be computed in closed form. Additionally, applying the results from [28] or via manual calculations a very simple formula for $\overline{M}_{\mathbf{1}_{\Gamma(T)}}$ can be derived. Returning to the original problem of maximizing the probability of a joint default, considering $(U,V)\sim A_0$ and setting $(X,Y) = (F^-\circ U, G^-\circ V)$, it follows that the pair (X,Y) has marginal distribution functions F and G and maximizes the joint default probability. In general, A_0 is not completely dependent unless F and G coincide.

Slightly modifying the optimization problem and maximizing $\mathbb{P}(Y\leq X)$ instead of $\mathbb{P}(Y=X)$ brings us back to complete dependence. In fact, proceeding analogously as before and setting

$$\Gamma^{\leq}(T) = \big\{(x,y)\in[0,1]^2 : y \leq T(x)\big\} \in \mathscr{B}([0,1]^2) \tag{14.12}$$

the following results can be derived manually or using the results in [28]:

Theorem 14.3 ([24]). *For every non-decreasing $T:[0,1]\to[0,1]$ we have*

$$\overline{M}_{\mathbf{1}_{\Gamma^{\leq}(T)}} = 1 + \inf_{x\in[0,1]} (T(x)-x). \tag{14.13}$$

Additionally, there exists a shuffle $A_R \in \mathscr{C}_d$ fulfilling $\mu_{A_R}(\Gamma^{\leq}(T)) = \overline{M}_{\mathbf{1}_{\Gamma^{\leq}(T)}}$.

In other words, given continuous distribution functions F and G, considering $(U,V)\sim A_R$ and setting $(X,Y) = (F^-\circ U, G^-\circ V)$, it follows that for the pair (X,Y) the quantity $\mathbb{P}(Y\leq X)$ is maximal.

Example 14.1. We consider a very simple situation illustrating Theorem 14.3: Choosing F as the distribution function of $\mathscr{U}(0,1)$ and $G=\Phi$ as the distribution function of $\mathscr{N}(0,1)$ we immediately get $T=\Phi$ as well as $\overline{M}_{\mathbf{1}_{\Gamma^{\leq}(T)}} = \Phi(1)\approx 0.841$. Figure 14.2 denotes a sample of the random vector (X,Y) with marginal distribution functions F and G for which $\mathbb{P}(Y\leq X)$ is maximal.

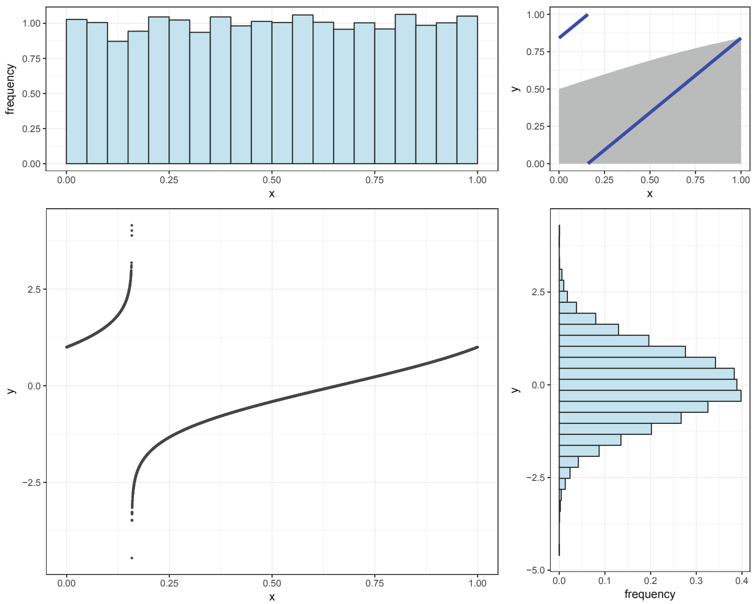


Fig. 14.2: Sample of size $n = 10.000$ of (X,Y) as in Example 14.2 and the corresponding marginal histrograms; the upper right panel depicts the endograph of T (gray) and the shuffle A_R according to Theorem 14.3 (blue).

14.4.2 A maximization problem from uniform distribution theory

Recall that a sequence $(x_n)_{n\in\mathbb{N}}$ in $[0,1]$ is called uniformly distributed if the induced empirical measure $\vartheta_n := \frac{1}{n}\sum_{i=1}^n \delta_{x_i}$ converges weakly to λ on $[0,1]$ for $n \to \infty$. For background on uniform distribution theory we refer to [4, 20].

Following [14] and the references therein one particularly interesting problem in the context of uniform distribution theory is the following one: Given uniformly distributed sequences $(x_n)_{n\in\mathbb{N}}$ and $(y_n)_{n\in\mathbb{N}}$ in $[0,1]$ and a real-valued continuous function H on $[0,1]^2$, determine

$$\limsup_{n\to\infty} \frac{1}{n}\sum_{i=1}^n H(x_i,y_i) \quad \text{and} \quad \liminf_{n\to\infty} \frac{1}{n}\sum_{i=1}^n H(x_i,y_i). \tag{14.14}$$

It is a straightforward exercise to show that for every accumulation point a of the sequence $(\frac{1}{n}\sum_{i=1}^n H(x_i,y_i))_{n\in\mathbb{N}}$ there exists a copula $A \in \mathscr{C}$ such that

$$a = \int_{[0,1]^2} H \, d\mu_A$$

holds. In other words, for calculating the quantities in eq (14.14) it suffices to calculate $\overline{M}_H$ and $-\overline{M}_{-H}$.

We will proceed as in [14] and consider the following special case for H: Fix $y_0 \in (0,1)$ and, suppose that H is continuous on $[0,1]^2$ and that $\frac{\partial^2 H(x,y)}{\partial y \partial x} > 0$ on $(0,1)\times(0,y_0)$ as well that $\frac{\partial^2 H(x,y)}{\partial y \partial x} < 0$ on $(0,1)\times(y_0,1)$ holds. Moreover we will let $\mathscr{S}_{y_0}$ denote the class of all y_0-sections of copulas, i.e. the family of all maps of the form $x \mapsto C(x,y_0)$, $x \in [0,1]$, with $C \in \mathscr{C}$. It is straightforward to verify that $s \in \mathscr{S}_{y_0}$ if and only if s fulfills the following three properties:

- $s(0) = 0, s(1) = y_0$
- s is non-decreasing and Lipschitz continuous with Lipschitz constant $L = 1$
- s fulfills $s(x) \in [W(x,y_0), M(x,y_0)]$ for all $x \in [0,1]$

For every $s \in \mathscr{S}_{y_0}$ in the rest of this section the copula $C^s \in \mathscr{C}$ be defined by

$$C^s(x,y) = \begin{cases} M(s(x),y) & \text{if } (x,y) \in [0,1]\times[0,y_0] \\ s(x) + (1-y_0)W\left(\frac{x-s(x)}{1-y_0}, \frac{y-y_0}{1-y_0}\right) & \text{if } (x,y) \in [0,1]\times(y_0,1]. \end{cases} \tag{14.15}$$

Obviously the y_0-section of C^s coincides with s and, setting $\bar{s}(x) = 1-(x-s(x))$ the copula C^s concentrates its mass on $\Gamma(s)\cup\Gamma(\bar{s})$ in the sense that

$$\mu_{C_s}(\Gamma(s)\cup\Gamma(\bar{s})) = 1.$$

The following reduction result can be shown:

Theorem 14.4 ([14]). *Under the afore-mentioned assumptions on H the following equality holds:*

$$\overline{M}_H = \max_{s\in\mathscr{S}_{y_0}} \int_{[0,1]^2} H\, d\mu_{C^s} \tag{14.16}$$

For general $s \in \mathscr{S}_{y_0}$ obviously the copula C^s need not be completely dependent. If, however, s is strictly increasing with $s' < 1$ then we get $(C^s)^t \in \mathscr{C}_d$, i.e. the transpose of C^s is completely dependent (also see Figure 14.3). We conclude this section with an example illustrating that a copula of the latter type may even be the unique maximizer.

Example 14.2. Consider $y_0 \in [\frac{1}{2},1)$ and suppose that H is given by

$$H = \begin{cases} xy \text{ if } (x,y) \in [0,1]\times[0,y_0] \\ xy_0 - x(y-y_0) \text{ if } (x,y) \in [0,1]\times(y_0,1]. \end{cases} \tag{14.17}$$

For arbitrary $s \in \mathscr{S}_{y_0}$ applying Theorem 14.4 and using integration by parts we finally get

$$\int_{[0,1]^2} H\, d\mu_{C^s} = y_0^2 - \frac{1}{2}\left\{ \int_{[0,1]} \left(s^2(x) + (2y_0 - 1 + x - s(x))^2\right) d\lambda(x) \right\}.$$

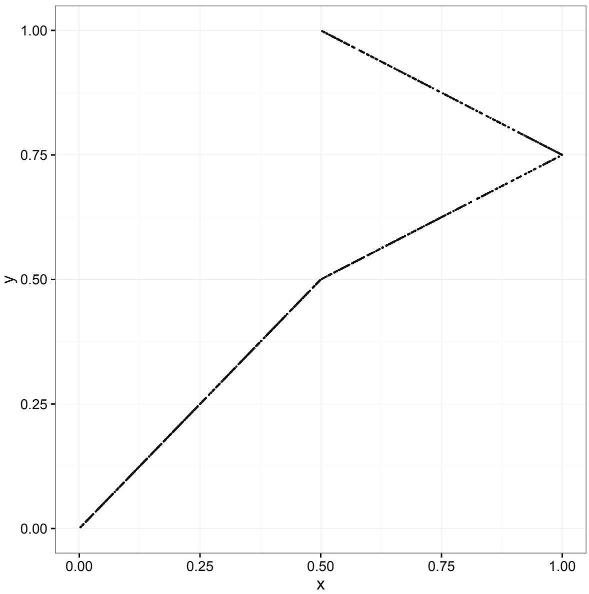


Fig. 14.3: Sample of size $n = 1.000$ of the unique maximizer C^s in Example 14.2.

For fixed x the integrand becomes minimal if $s(x) = y_0 - \frac{1}{2} + \frac{x}{2}$. The function $s_1 : x \mapsto y_0 - \frac{1}{2} + \frac{x}{2}$ is a global minimizer of the integral which, however, only lies in $\mathscr{S}_{y_0}$ for $y_0 = \frac{1}{2}$. It is straightforward to verify that for $y_0 \geq \frac{1}{2}$ the (piecewise linear) function h, defined by

$$s(x) = \begin{cases} x \text{ if } x \in [0, 2y_0 - 1] \\ y_0 - \frac{1}{2} + \frac{x}{2} \text{ if } x \in (2y_0 - 1, 1] \end{cases}$$

is the unique minimizer of the integral in eq. (14.17). As a consequence, the corresponding copula C^s, which fulfills $(C^s)^t \in \mathscr{C}_d$ is the unique copula attaining $\overline{M}_H$. Figure 14.3 depicts a sample of the corresponding copula C^s for the case $y_0 = \frac{3}{4}$.

14.5 A typical copula is (mutually) completely dependent

As already mentioned in Section 3 the set of all shuffles of M is dense in the metric space $(\mathscr{C}, d_\infty)$ (but nowhere dense in the metric space $(\mathscr{C}, D_1)$, see [35]). On the one hand, non-absolutely continuous copulas naturally appear in various problems, on the other hand, possibly due to their handy structure, absolutely continuous copulas are certainly not underrepresented in the literature.

Topology offers a way to quantify the size of sets in a binary manner through Baire categories (see [27]): A subset E of a general metric space (Ω,d) is considered small if it is of first category or meager, i.e. if it is the countable union of sets E_i whose topological closer has empty interior. If E is not of first category then, by definition, E is said to be of second category. Finally, if E is meager then E^c is considered big and referred to as co-meager. Following [6] we will call elements of a meager set E *atypical* and elements of a co-meager set *typical*. With this topological notions various interesting results can be shown, e.g. that the family $\mathscr{C}_{abs}$ of all absolutely continuous copulas is of first category both in $(\mathscr{C},d_\infty)$ and in $(\mathscr{C},D_1)$. Considering completeness of $(\mathscr{C},d_\infty)$ and $(\mathscr{C},D_1)$ it directly follows that the family $\mathscr{C}^c_{abs} = \mathscr{C}\setminus\mathscr{C}_{abs}$ of all copulas with non-degenerated singular component is co-meager and of second category in $(\mathscr{C},d_\infty)$ and in $(\mathscr{C},D_1)$, i.e. a typical copula has a non-degenerated singular component. As a matter of fact, the following much stronger result holds (and can be extended to the general multivariate setting):

Theorem 14.5 ([6]). *The family $\mathscr{C}_{sing}$ of all purely singular copulas is co-meager (hence of second category) in $(\mathscr{C},d_\infty)$.*

In other words: A typical copula in $(\mathscr{C},d_\infty)$ has no absolutely continuous component. It remains an open question if $\mathscr{C}_{sing}$ is also co-meager in $(\mathscr{C},D_1)$. As the authors of [6] discovered recently, Theorem 14.5 is not even close to the end of the story - the following striking result was proved already in 1968:

Theorem 14.6 ([17]). *$\mathscr{C}_d$ is co-meager (hence of second category) in $(\mathscr{C},d_\infty)$.*

Based on the elegant proof of Theorem 14.6 as given in [17] one gets the following even more striking corollary, saying that *in $(\mathscr{C},d_\infty)$ a typical copula is mutually completely dependent*, without any difficulty:

Corollary 14.1. *The family of all mutually completely dependent copulas is co-meager (hence of second category) in $(\mathscr{C},d_\infty)$.*

14.6 Sharp inequalities between Kendall's τ and Spearman's ρ

This section first recalls the main results from [29] and then sketches why 'complete dependence everywhere' particularly holds true in the situation of Kendall's τ and Spearman's ρ.

Suppose that X,Y are random variables with continuous distribution functions F and G respectively. Then Spearman's ρ is defined as the Pearson correlation coefficient of the $\mathscr{U}(0,1)$-distributed random variables $U := F\circ X$ and $V := G\circ Y$ and Kendall's τ is given by the probability of concordance minus the probability of discordance, i.e.

$$\rho(X,Y) = 12\big(\mathbb{E}(UV) - \tfrac{1}{4}\big)$$
$$\tau(X,Y) = \mathbb{P}\big((X_1-X_2)(Y_1-Y_2)>0\big) - \mathbb{P}\big((X_1-X_2)(Y_1-Y_2)<0\big),$$

where (X_1,Y_1) and (X_2,Y_2) are independent and have the same distribution as (X,Y). Clearly τ and ρ are the two most famous nonparametric measures of concordance. Both measures are scale invariant and only depend on the underlying (uniquely determined) copula A of (X,Y). It is well known and straightforward to verify (see [25]) that $\tau(X,Y)$ and $\rho(X,Y)$ can be expressed in terms of the underlying copula A as

$$\tau(X,Y) = 4\int_{[0,1]^2} A(x,y)\, d\mu_A(x,y) - 1 =: \tau(A) \tag{14.18}$$

$$\rho(X,Y) = 12\int_{[0,1]^2} xy\, d\mu_A(x,y) - 3 =: \rho(A) \tag{14.19}$$

Considering that τ and ρ quantify different aspects of the underlying dependence structure, it is natural to ask how much they can differ. Since the 1950s two universal inequalities between τ and ρ are known - the first one goes back to Daniels ([3]), the second one to Durbin and Stuart ([8]); for proofs alternative to the original ones see [19, 12, 25].

$$|3\tau - 2\rho| \leq 1 \tag{14.20}$$

$$\frac{(1+\tau)^2}{2} - 1 \leq \rho \leq 1 - \frac{(1-\tau)^2}{2} \tag{14.21}$$

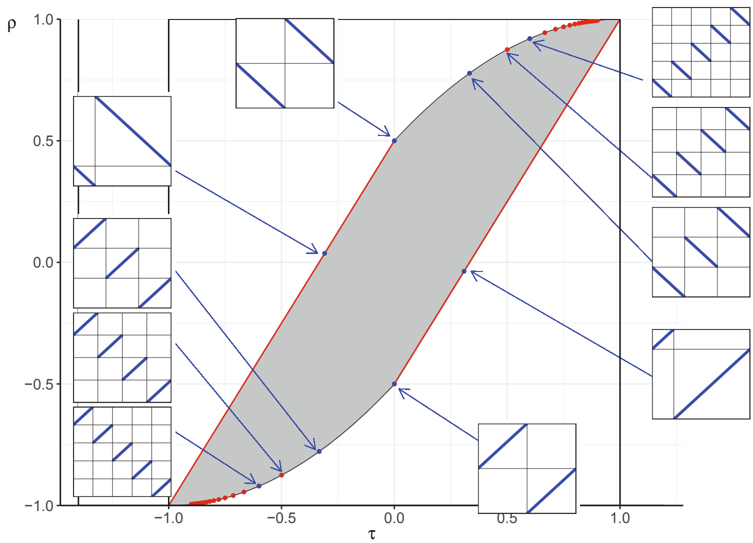


Fig. 14.4: The classical τ-ρ-region Ω_0 and some copulas (distributing mass uniformly on the blue segments) for which the inequality by Durbin and Stuart is sharp.

The inequalities together yield the set Ω_0 (see Figure 14.4) which, following [29] we will refer to as *classical* τ-ρ *region* in the sequel. Daniels' inequality was known to be sharp [25] whereas the first part of the inequality by Durbin and Stuart was only known to be sharp at the points $p_n = (-1+\frac{2}{n}, -1+\frac{2}{n^2})$ with $n \geq 2$ (which, using symmetry, is to say that the second part is sharp at the points $-p_n$). Although both inequalities were known since the 1950s and the interrelation between τ and ρ keeps receiving much attention in recent years (particularly in the context of the Hutchinson-Lai conjecture [11, 13]), only very recently a full characterization of the *exact* τ-ρ *region* Ω, defined by

$$\begin{aligned}\Omega &= \big\{(\tau(X,Y), \rho(X,Y)) : X,Y \text{ continuous random variables}\big\} \\ &= \big\{(\tau(A), \rho(A)) : A \in \mathscr{C}\big\},\end{aligned} \tag{14.22}$$

was given in [29]. One direct consequence of this characterization is the fact that inequality by Durbin and Stuart is not sharp outside the points $\pm p_n$.

Throughout the entire proof shuffles (hence complete dependence) played a crucial role: The authors first calculated τ and ρ for so-called prototypes, which, loosely speaking, are shuffles consisting of $n-1$ segments of equal length and a shorter one, arranged in decreasing order similar to the shuffles depicted in Figure 14.4. Based on these prototypes they defined $\Phi_n : [-1+\frac{2}{n}, 1] \to [-1,1]$ by

$$\Phi_n(x) = -1 - \frac{4}{n^2} + \frac{3}{n} + \frac{3x}{n} - \frac{n-2}{\sqrt{2}n^2\sqrt{n-1}}(n-2+nx)^{3/2} \tag{14.23}$$

and then set

$$\Phi(x) = \begin{cases} -1 & \text{if } x = -1, \\ \Phi_n(x) & \text{if } x \in \left[\frac{2-n}{n}, \frac{2-(n-1)}{n-1}\right] \text{ for some } n \geq 2. \end{cases} \tag{14.24}$$

Based on Φ the set Ω can be characterized as follows (this characterization was already conjectured by Manuel Úbeda-Flores in an unpublished working paper in 2011):

Theorem 14.7 ([29]). *The following equality holds:*

$$\Omega = \big\{(x,y) \in [-1,1]^2 : \Phi(x) \leq y \leq -\Phi(-x)\big\} \tag{14.25}$$

In particular, Ω is compact but not convex. For an animation showing for which copulas A we have $(\tau(A), \rho(A)) \in \partial\Omega$, where $\partial\Omega$ denotes the topological boundary of Ω, we refer to http://www.trutschnig.net/tau-rho-boundary.pdf

Returning to complete dependence, notice that continuity of τ and ρ with respect to d_∞ directly yields that $\{(\tau(A_h), \rho(A_h)) : h \in \mathscr{T}_s\}$ is dense in Ω. The proof of Theorem 14.7, however, produced the by-product that only for prototypes $A \in \mathscr{C}_d$ we can have $(\tau(A), \rho(A)) \in \partial\Omega$. In fact, using a homotopy argument it was possible to show the following corollary which underlines the importance of (mutual) complete dependence yet again.

Corollary 14.2. *For every point* $(x,y) \in \Omega$ *there exists a transformation* $h \in \mathscr{T}_s$ *such that we have* $(\tau(A_h), \rho(A_h)) = (x,y)$.

As pointed out in [29, Section 6] characterizing the exact τ-ρ-region for standard subclasses of copulas may in some cases be even more difficult than determining Ω was. The main reason for this fact is that not in all subclasses of $\mathscr{C}$ we may find dense subsets consisting of elements B for which $\tau(B)$ and $\rho(B)$ reduce to handy formulas (as it is the case for shuffles of M). The author conjectures, however, that the classical Hutchinson-Lai inequalities are not sharp for the class of all extreme-value copulas and that it might be possible to derive sharper inequalities in the near future.

Acknowledgements The authory gratefully acknowledge the support of the grant MTM2014-60594-P (partially supported by FEDER) from the Spanish Ministry of Economy and Competitiveness.

References

1. Aliprantis, C.D., Border, K.C.: Infinite dimensional analysis, third edn. Springer, Berlin (2006). A hitchhiker's guide
2. Boonmee, T., Tasena, S.: Measure of complete dependence of random vectors. J. Math. Anal. Appl. **443**(1), 585–595 (2016)
3. Daniels, H.E.: Rank correlation and population models. J. Roy. Statist. Soc. Ser. B. **12**, 171–181 (1950)
4. Drmota, M., Tichy, R.F.: Sequences, discrepancies and applications, *Lecture Notes in Mathematics*, vol. 1651. Springer-Verlag, Berlin (1997)
5. Durante, F., Fernández Sánchez, J., Trutschnig, W.: Multivariate copulas with hairpin support. J. Multivariate Anal. **130**, 323–334 (2014)
6. Durante, F., Fernández-Sánchez, J., Trutschnig, W.: A typical copula is singular. J. Math. Anal. Appl. **430**(1), 517–527 (2015)
7. Durante, F., Sempi, C.: Principles of copula theory. CRC Press, Boca Raton, FL (2016)
8. Durbin, J., Stuart, A.: Inversions and rank correlation coefficients. J. Roy. Statist. Soc. Ser. B. **13**, 303–309 (1951)
9. Fernández Sánchez, J., Trutschnig, W.: Conditioning-based metrics on the space of multivariate copulas and their interrelation with uniform and levelwise convergence and iterated function systems. J. Theoret. Probab. **28**(4), 1311–1336 (2015)
10. Fernández Sánchez, J., Trutschnig, W.: Singularity aspects of Archimedean copulas. J. Math. Anal. Appl. **432**(1), 103–113 (2015)
11. Fredricks, G.A., Nelsen, R.B.: On the relationship between Spearman's rho and Kendall's tau for pairs of continuous random variables. J. Statist. Plann. Inference **137**(7), 2143–2150 (2007)
12. Genest, C., Nešlehová, J.: Analytical proofs of classical inequalities between Spearman's ρ and Kendall's τ. J. Statist. Plann. Inference **139**(11), 3795–3798 (2009)
13. Hürlimann, W.: Hutchinson-Lai's conjecture for bivariate extreme value copulas. Statist. Probab. Lett. **61**(2), 191–198 (2003)
14. Iacò, M.R., Thonhauser, S., Tichy, R.F.: Distribution functions, extremal limits and optimal transport. Indag. Math. (N.S.) **26**(5), 823–841 (2015)
15. Kallenberg, O.: Foundations of modern probability, second edn. Probability and its Applications (New York). Springer-Verlag, New York (2002)

16. Kamnitui, N., Santiwipanont, T., Sumetkijakan, S.: Dependence measuring from conditional variances. Depend. Model. **3**(1), 98–112 (2015)
17. Kim, C.W.: Uniform approximation of doubly stochastic operators. Pacific J. Math. **26**, 515–527 (1968)
18. Klenke, A.: Probability theory, second edn. Universitext. Springer, London (2014)
19. Kruskal, W.H.: Ordinal measures of association. J. Amer. Statist. Assoc. **53**, 814–861 (1958)
20. Kuipers, L., Niederreiter, H.: Uniform distribution of sequences. Wiley-Interscience [John Wiley & Sons], New York-London-Sydney (1974). Pure and Applied Mathematics
21. Lancaster, H.O.: Correlation and complete dependence of random variables. Ann. Math. Statist. **34**, 1315–1321 (1963)
22. Li, X., Mikusiński, P., Taylor, M.D.: Strong approximation of copulas. J. Math. Anal. Appl. **225**(2), 608–623 (1998)
23. Losert, V.: Counterexamples to some conjectures about doubly stochastic measures. Pacific J. Math. **99**(2), 387–397 (1982)
24. Mroz, T., Trutschnig, W., Fernández Sánchez, J.: Distributions with fixed marginals maximizing the mass of the endograph of a function. preprint on aArXiv: 1602.05807
25. Nelsen, R.B.: An Introduction to Copulas, second edn. Springer Series in Statistics. Springer, New York (2006)
26. Nelsen, R.B., Fredricks, G.A.: Diagonal copulas. In: Distributions with given marginals and moment problems (Prague, 1996), pp. 121–128. Kluwer Acad. Publ., Dordrecht (1997)
27. Oxtoby, J.C.: Measure and category. A survey of the analogies between topological and measure spaces. Springer-Verlag, New York-Berlin (1971). Graduate Texts in Mathematics, Vol.2
28. Rüschendorf, L.: Random variables with maximum sums. Adv. in Appl. Probab. **14**(3), 623–632 (1982)
29. Schreyer, M., Paulin, R., Trutschnig, W.: On the exact region determined by Kendall's tau and Spearman's rho. J. Roy. Statist. Soc. Ser. B. **79**(2), 613–633 (2017)
30. Schweizer, B., Wolff, E.F.: On nonparametric measures of dependence for random variables. Ann. Statist. **9**(4), 879–885 (1981)
31. Seethoff, T.L., Shiflett, R.C.: Doubly stochastic measures with prescribed support. Z. Wahrscheinlichkeitstheorie und Verw. Gebiete **41**(4), 283–288 (1977/78)
32. Sherwood, H., Taylor, M.D.: Doubly stochastic measures with hairpin support. Probab. Theory Related Fields **78**(4), 617–626 (1988)
33. Siburg, K.F., Stoimenov, P.A.: A measure of mutual complete dependence. Metrika **71**(2), 239–251 (2010)
34. Thorisson, H.: Coupling, stationarity, and regeneration. Probability and its Applications (New York). Springer-Verlag, New York (2000)
35. Trutschnig, W.: On a strong metric on the space of copulas and its induced dependence measure. J. Math. Anal. Appl. **384**(2), 690–705 (2011)
36. Trutschnig, W., Fernández Sánchez, J.: Idempotent and multivariate copulas with fractal support. J. Statist. Plann. Inference **142**(12), 3086–3096 (2012)
37. Trutschnig, W., Fernández Sánchez, J.: Copulas with continuous, strictly increasing singular conditional distribution functions. J. Math. Anal. Appl. **410**(2), 1014–1027 (2014)

Chapter 15
Sklar's theorem: The cornerstone of the Theory of Copulas

Manuel Úbeda-Flores and Juan Fernández-Sánchez

Abstract In this contribution we present a complete study of the original proof of Sklar's theorem, provide an alternative proof by using Zorn's lemma, and review other proofs given in the literature.

15.1 Introduction

After the pioneering work of Maurice Fréchet in the study of the relationship between multivariate distribution functions and their one-dimensional marginals, Abe Sklar introduced the concept of copula to explain such a relation in a theorem that now bears his name. The proof of the theorem does not appear in Sklar's paper [21] and more than a decade had to pass before a proof was published in [22], even though it is not provided in detail. Thanks to Sklar's theorem, copulas are an essential tool for creating mathematical models since they capture the scale-invariant dependence properties of continuous random vectors. For a complete review on copulas, see [6, 16].

Although the original proof of Sklar's theorem can be found in some works (see [20, 22], for instance), we stress that these proofs are sketches, more or less extensive. Our aim in this work is to provide a complete proof of the theorem in the $n-$dimensional case—a preliminary study is done in [23]. Furthermore, we also review some alternative proofs which use different ideas.

Manuel Úbeda-Flores
Department of Mathematics, University of Almería, 04120 Almería, (Spain)
e-mail: mubeda@ual.es

Juan Fernández Sánchez
Research Group of Mathematical Analysis, University of Almería, 04120 Almería (Spain)
e-mail: juanfernandez@ual.es

M. Úbeda Flores et al. (eds.), *Copulas and Dependence Models with Applications*,
DOI 10.1007/978-3-319-64221-5_15

15.2 Preliminaries

We begin this section with some notations and definitions.

Let n be a positive integer. Let $\mathbf{a} = (a_1, a_2, \ldots, a_n)$ and $\mathbf{b} = (b_1, b_2, \ldots, b_n)$ be two points in $\overline{\mathbb{R}}^n$ $(= [-\infty, +\infty]^n)$. An *n-dimensional box* (briefly, *n-box*), $J = [\mathbf{a}, \mathbf{b}]$, is any cartesian product of n closed intervals of $\overline{\mathbb{R}}$, i.e., $J = \prod_{i=1}^n [a_i, b_i]$. If there is some i, $1 \le i \le n$, such that $a_i = b_i$, we say that J is a *degenerated* $n-$box; otherwise, we say that J is a *non-degenerated* $n-$box.

Given an $n-$box J, we define the *vertices* of J as the points $\mathbf{c} = (c_1, c_2, \ldots, c_n)$ in $\overline{\mathbb{R}}$ such that, for every $i = 1, 2, \ldots, n$, we have $c_i = a_i$ or $c_i = b_i$. We denote by $v(J)$ the set of all vertices of J.

If $J = [\mathbf{a}, \mathbf{b}]$ is a non-degenerated $n-$box, then, for any vertix $\mathbf{c}$ of J, we define the *sign* of $\mathbf{c}$—denoted by $\operatorname{sgn}(\mathbf{c})$—as follows: $\operatorname{sgn}(\mathbf{c}) = 1$ if $c_i = a_i$ for an even number of $i's$, and $\operatorname{sgn}(\mathbf{c}) = -1$ if $c_i = a_i$ for an odd number of $i's$. If J is a degenerated $n-$box, we define the sign of any vertix $\mathbf{c}$ as $\operatorname{sgn}(\mathbf{c}) = 0$.

Let D be a non-empty set of $\overline{\mathbb{R}}^n$. Let $G\colon D \longrightarrow \mathbb{R}$ be a function and let J be an $n-$box whose vertices belong to D. The *G−volume* of J is defined by

$$V_G(J) = \sum_{\mathbf{c} \in v(J)} \operatorname{sgn}(\mathbf{c}) G(\mathbf{c}).$$

Furthermore, G is said to be *n−increasing* if $V_G(J) \ge 0$ for any $n-$box J such that $v(J) \subset D$.

Definition 15.1. Let n be a natural number such that $n \ge 2$. An *n-dimensional subcopula* (briefly, $n-$subcopula) is a function $S\colon \prod_{i=1}^n A_i \longrightarrow \mathbb{I}\,(= [0,1])$, where $A_i \subseteq \mathbb{I}$, such that A_i contains both 0 and 1 for all $i = 1, 2, \ldots, n$, and fulfilling the following conditions:

1. $S(\mathbf{u}) = 0$ if $u_i = 0$ for some $i = 1, 2, \ldots, n$;
2. $S(1, \ldots, 1, u_i, 1, \ldots, 1) = u_i$ for all $i = 1, 2, \ldots, n$ and $u_i \in A_i$; and
3. S is $n-$increasing.

Note that, as a consequence of Definition 15.1, any subcopula is non-decreasing in each variable, is 1-Lipschitz, i.e., $|S(\mathbf{u}) - S(\mathbf{v})| \le \sum_{i=1}^n |u_i - v_i|$ for all $\mathbf{u}, \mathbf{v} \in \operatorname{Dom}(S)$, and, therefore, is uniformly continuous in its domain.

Let $\mathscr{S}$ denote the set of all $n-$subcopulas, and $\mathscr{S}'$ denotes the set of all $n-$subcopulas with closed domain.

An *n-dimensional copula* (briefly, *n-copula*) C is an $n-$subcopula S with domain $\mathbb{I}^n$. Therefore, an $n-$copula C is an extension of S if it agrees with S on its domain. We denote by $\mathscr{C}_n$ the set of all $n-$copulas.

We recall the concept of $n-$dimensional distribution function.

Definition 15.2. Let n be a positive integer. The function F from $\overline{\mathbb{R}}^n$ onto $\mathbb{R}$ is an $n-$dimensional distribution function if it satisfies the following conditions:

1. $F(+\infty, +\infty, \ldots, +\infty) = 1$;

2. $F(x_1,x_2,\dots,x_n) = 0$ if $x_i = -\infty$ for some i, $1 \le i \le n$;
3. F is left-continuous in each variable; and
4. F is $n-$increasing.

We note that, for historic reasons, the distribution functions defined in Definition 15.2 are left−continuous, but the result is similar if we consider right−continuity.

If $n \ge 2$ and $1 \le i \le n$, the $1-$dimensional (or univariate) $i-$marginal of F is the distribution function F_i defined by

$$F_i(x_i) = F(+\infty,\dots,+\infty,x_i,+\infty,\dots,+\infty), \quad x_i \in \overline{\mathbb{R}}.$$

Similarly, for $1 < m < n$, the $m-$dimensional marginals of F can be defined as the distribution functions obtained by taking $n-m$ variables equal to $+\infty$.

Let $F_1,F_2,\dots,F_n$ be n univariate distribution functions. The *Fréchet class* associated with these functions, and which we denote by $\Gamma(F_1,F_2,\dots,F_n)$, is the family of all $n-$dimensional distribution functions whose univariate margins are $F_1,F_2,\dots,F_n$. Fréchet studied this class (see [8]), and in 1959, with the aid of copulas, Sklar [21] provided the relationship between a multivariate distribution function and its univariate marginals.

Theorem 15.1 (Sklar). *Let $F_1,F_2,\dots,F_n$ be n univariate distribution functions. Then $F \in \Gamma(F_1,F_2,\dots,F_n)$ if and only if there exists an $n-$copula C, which is uniquely determined on $\prod_{i=1}^n \operatorname{Range} F_i$, such that*

$$F(\mathbf{x}) = C(F_1(x_1),F_2(x_2),\dots,F_n(x_n)),\ \forall \mathbf{x} = (x_1,x_2,\dots,x_n) \in \overline{\mathbb{R}}^n.$$

Furthermore, if $F_1,F_2,\dots,F_n$ are continuous, then C is unique.

We want to stress that in the original proof by Schweizer and Sklar [20], the authors considered distribution functions defined on $\overline{\mathbb{R}}^n$, but in most of the probabilistic approaches to Sklar's theorem—for example, in [15, 18]—the distribution functions are defined on $\mathbb{R}^n$. Here we consider $\overline{\mathbb{R}}^n$.

15.3 Extension theorem

We begin this section with a preliminary lemma.

Lemma 15.1. *Let $F_1,F_2,\dots,F_n$ be n univariate distribution functions and let $F \in \Gamma(F_1,F_2,\dots,F_n)$. Then we have*

$$|F(\mathbf{y}) - F(\mathbf{x})| \le \sum_{i=1}^n |F_i(y_i) - F_i(x_i)|$$

for all $\mathbf{x},\mathbf{y}$ in $\overline{\mathbb{R}}^n$.

Proof. Let $i \in \{1,2,\ldots,n\}$, let x_i, y_i be two points in $\overline{\mathbb{R}}$ such that $x_i \le y_i$. First we prove that

$$F(x_1,\ldots,x_{i-1},y_i,x_{i+1},\ldots,x_n) - F(x_1,\ldots,x_i,\ldots,x_n) \le F_i(y_i) - F_i(x_i)$$

for $x_1,\ldots,x_{i-1},x_{i+1},\ldots,x_n$ in $\overline{\mathbb{R}}$. For that, let $k = 1,\ldots,i-1,i+1,\ldots,n$. If $k < i$, let $(\mathbf{x}_i)_{k-}, (\mathbf{y}_i)_{k+}$ be the points

$$\begin{aligned}(\mathbf{x}_i)_{k-} &= (-\infty,\ldots,-\infty,x_k,-\infty,\ldots,-\infty,x_i,-\infty,\ldots,-\infty)\\ (\mathbf{y}_i)_{k+} &= (+\infty,\ldots,+\infty,x_k,\ldots,x_{i-1},y_i,x_{i+1},\ldots,x_n),\end{aligned}$$

and if $k > i$, let $(\mathbf{x}_i)^*_{k-}, (\mathbf{y}_i)^*_{k+}$ be the points

$$\begin{aligned}(\mathbf{x}_i)^*_{k-} &= (-\infty,\ldots,-\infty,x_i,-\infty,\ldots,-\infty,x_k,-\infty,\ldots,-\infty)\\ (\mathbf{y}_i)^*_{k+} &= (+\infty,\ldots,+\infty,y_i,+\infty,\ldots,+\infty,x_k,\ldots,x_n).\end{aligned}$$

We consider the following $n-$boxes:

$$J_k = \begin{cases} \left[(\mathbf{x}_i)_{k-}, (\mathbf{y}_i)_{(k+1)+}\right], & \text{if } k < i \\ \left[(\mathbf{x}_i)^*_{k-}, (\mathbf{y}_i)^*_{(k+1)+}\right], & \text{if } k > i. \end{cases}$$

After some elementary computations, we have

$$V_F(J_k) = \begin{cases} F\left((\mathbf{y}_i)_{(k+1)+}\right) - F\left((\mathbf{y}_i)_{k+}\right) - F\left((\mathbf{x}_i)_{(k+1)+}\right) - F\left((\mathbf{x}_i)_{k+}\right), & \text{if } k < i \\ F\left((\mathbf{y}_i)^*_{(k+1)+}\right) - F\left((\mathbf{y}_i)^*_{k+}\right) - F\left((\mathbf{x}_i)^*_{(k+1)+}\right) - F\left((\mathbf{x}_i)^*_{k+}\right), & \text{if } k > i, \end{cases}$$

and thus

$$\begin{aligned}\sum_{\substack{k=1\\k\neq i}}^{n} V_F(J_k) &= F_i(y_i) - F\left((\mathbf{y}_i)_{1+}\right) - F_i(x_i) + F\left((\mathbf{x}_i)_{1+}\right)\\ &= F_i(y_i) - F(x_1,\ldots,x_{i-1},y_i,x_{i+1},\ldots,x_n) - F_i(x_i) + F(\mathbf{x}).\end{aligned}$$

Since F is $n-$increasing, then F is non-decreasing in each variable and

$$F_i(y_i) - F(x_1,\ldots,x_{i-1},y_i,x_{i+1},\ldots,x_n) - F_i(x_i) + F(\mathbf{x}) \ge 0;$$

therefore,

$$|F(x_1,\ldots,x_{i-1},y_i,x_{i+1},\ldots,x_n) - F(\mathbf{x})| \le |F_i(y_i) - F_i(x_i)|$$

for every $i = 1,2,\ldots,n$.

Now, note

$$\begin{aligned}|F(\mathbf{y})-F(\mathbf{x})| &= |F(\mathbf{y})-F(x_1,y_2,\ldots,y_n)+F(x_1,y_2,\ldots,y_n)\\ &\quad -F(x_1,x_2,y_3,\ldots,y_n)+F(x_1,x_2,y_3,\ldots,y_n)\\ &\quad -\cdots+F(x_1,x_2,\ldots,x_{n-1},y_n)-F(\mathbf{x})|\\ &\le \sum_{i=1}^{n}|F_i(y_i)-F_i(x_i)|,\end{aligned}$$

which completes the proof.

Remark 15.1. We want to stress that Lemma 15.1 is also true for functions defined from $\overline{\mathbb{R}}^n$ onto $\mathbb{R}$ such that they only satisfy conditions 1,2 and 4 in Definition 15.2, but we have adapted it for our purposes.

Now, in order to prove Sklar's theorem, we first prove the following result.

Theorem 15.2. *Let $F_1,F_2,\ldots,F_n$ be n univariate distribution functions and let $F \in \Gamma(F_1,F_2,\ldots,F_n)$. Then there is a unique $n-$subcopula $S\colon \prod_{i=1}^n \text{Range } F_i \longrightarrow \mathbb{I}$ such that*

$$F(\mathbf{x}) = S(F_1(x_1),F_2(x_2),\ldots,F_n(x_n)) \tag{15.1}$$

for every $\mathbf{x}\in\overline{\mathbb{R}}^n$. The $n-$subcopula S is given by

$$S(\mathbf{u}) = F\left(F_1^{(-1)}(u_1),F_2^{(-1)}(u_2),\ldots,F_n^{(-1)}(u_n)\right)$$

for all $\mathbf{u}\in\prod_{i=1}^n \text{Range } F_i$, where, for each $i\in\{1,2,\ldots,n\}$, $F_i^{(-1)}$ is the quasi-inverse of F_i, i.e., $F_i^{(-1)}(t)=\inf\{x\in\overline{\mathbb{R}}\colon F_i(x)\ge t\}$.

Conversely, if $S\colon \prod_{i=1}^n A_i \longrightarrow \mathbb{I}$ is an $n-$subcopula such that $\text{Range}\, F_i \subseteq A_i$, for every $i=1,2,\ldots,n$, then the function F defined by (15.1) satisfies $F\in\Gamma(F_1,F_2,\ldots,F_n)$.

Proof. Suppose $F\in\Gamma(F_1,F_2,\ldots,F_n)$. From Lemma 15.1, if $F_i(x_i)=F_i(y_i)$ for each $i=1,2,\ldots,n$, then $F(\mathbf{x})=F(\mathbf{y})$ for $\mathbf{x},\mathbf{y}$ in $\overline{\mathbb{R}}^n$, i.e., there is a unique function S such that (15.1) holds.

We prove that S is an $n-$subcopula. Firstly, observe that both 0 and 1 are in $\text{Range}\, F_i$ for every $i=1,2,\ldots,n$. Moreover, we have

$$\begin{aligned}&S(F_1(x_1),\ldots,F_{i-1}(x_{i-1}),0,F_{i+1}(x_{i+1}),\ldots,F_n(x_n))\\ &\quad= S(F_1(x_1),\ldots,F_{i-1}(x_{i-1}),F_i(-\infty),F_{i+1}(x_{i+1}),\ldots,F_n(x_n))\\ &\quad= F(x_1,\ldots,x_{i-1},-\infty,x_{i+1},\ldots,x_n)=0.\end{aligned}$$

and

$$\begin{aligned}&S(1,\ldots,1,F_i(x_i),1,\ldots,1)\\ &\quad= S(F_1(+\infty),\ldots,F_{i-1}(+\infty),F_i(x_i),F_{i+1}(+\infty),\ldots,F_n(+\infty))\\ &\quad= F(+\infty,\ldots,+\infty,x_i,+\infty,\ldots,+\infty)=F_i(x_i)\end{aligned}$$

for every $x_i\in\overline{\mathbb{R}}$.

Let $J = [\mathbf{a}, \mathbf{b}]$ be a non-degenerated $n-$box whose vertices are in $\prod_{i=1}^{n} \operatorname{Range} F_i$. Then we have

$$V_S\left((F_1(a_1), F_2(a_2), \ldots, F_n(a_n)), (F_1(b_1), F_2(b_2), \ldots, F_n(b_n))\right) = V_F(J) \geq 0,$$

which proves that S is an $n-$subcopula.

Conversely, if $S\colon \prod_{i=1}^{n} A_i \longrightarrow \mathbb{I}$ is an $n-$subcopula with $\operatorname{Range} F_i \subseteq A_i$, we define

$$F(\mathbf{x}) = S(F_1(x_1), F_2(x_2), \ldots, F_n(x_n))$$

for every $\mathbf{x} \in \overline{\mathbb{R}}^n$. We now prove that $F \in \Gamma(F_1, F_2, \ldots, F_n)$. For that, note the following facts:

1. $F(+\infty, +\infty, \ldots, +\infty) = S(1, 1, \ldots, 1) = 1$.
2. $F(x_1, \ldots, x_{i-1}, -\infty, x_{i+1}, \ldots, x_n) =$
$$= S(F_1(x_1), \ldots, F_{i-1}(x_{i-1}), F_i(-\infty), F_{i+1}(x_{i+1}), \ldots, F_n(x_n)) = 0$$
since $F_i(-\infty) = 0$.
3. Since F_i is left-continuous for every i and S is continuous, the composition is left-continuous, and thus F is left-continuous in each variable.
4. F is $n-$increasing since S is $n-$increasing.

Moreover,

$$F(+\infty, \ldots, +\infty, x_i, +\infty, \ldots, +\infty) = S(1, \ldots, 1, F_i(x_i), 1, \ldots, 1) = F_i(x_i),$$

which completes the proof.

Observe that if the functions $F_1, F_2, \ldots, F_n$ in Theorem 15.2 are continuous, the function S is an $n-$copula. Therefore, to prove the Sklar's theorem is equivalent to prove that every $n-$subcopula can be extended to an $n-$copula, which is the "hard" part. We call this result the *Extension theorem*. It appears for first time—without proof—in [21]. Its proof can be found in [20] for the bivariate case, and in [19, 22] for the $n-$dimensional case, although they are not made with the same detail than ours. Other proofs concerning the next results—both the bivariate case and the general case—can be found in [5, 13, 15, 24], even though their proofs are provided in probabilistic terms and do not use either copulas or subcopulas.

Theorem 15.3 (Extension). *Every $n-$subcopula can be extended to an $n-$copula, i.e., given an $n-$subcopula $S\colon \prod_{i=1}^{n} A_i \longrightarrow \mathbb{I}$, there exists an $n-$copula C such that $C|_{\prod_{i=1}^{n} A_i} = S$.*

To prove Theorem 15.3 we need some preliminary results which we now provide.

Given an $n-$box $J = [\mathbf{a}, \mathbf{b}]$ and a function $f\colon v(J) \longrightarrow \mathbb{R}$, we define the *standard extension E from f to J* as

$$E(\mathbf{x}) = \sum_{\mathbf{c} \in v(J)} \left[\prod_{i=1}^{n} K(c_i, x_i)\right] f(\mathbf{c})$$

with $\mathbf{x} = (x_1, x_2, \ldots, x_n)$ and $\mathbf{c} = (c_1, c_2, \ldots, c_n)$, and where

$$K(c_i, x_i) = \begin{cases} 1 - \dfrac{|c_i - x_i|}{b_i - a_i}, & \text{if } a_i < b_i \\ 1, & \text{if } a_i = b_i. \end{cases}$$

Observe that, for $a_i < b_i$, we have

$$K(c_i, x_i) = \begin{cases} \dfrac{b_i - x_i}{b_i - a_i}, & \text{if } c_i = a_i \\ \\ \dfrac{x_i - a_i}{b_i - a_i}, & \text{if } c_i = b_i. \end{cases}$$

Moreover, if $a_i < b_i$, then $K(a_i, x_i) + K(b_i, x_i) = 1$. Therefore, we have

$$E(\mathbf{x}) = \sum_{\mathbf{c} \in v(J)} \left[\prod_{\substack{c_i = a_i \\ a_i < b_i}} \frac{b_i - x_i}{b_i - a_i} \prod_{\substack{c_i = b_i \\ a_i < b_i}} \frac{x_i - a_i}{b_i - a_i} \right] f(\mathbf{c}).$$

The next result shows that the standard extension uses one of the forms in which every point of an $n-$box can be extended as a convex linear combination of the vertices.

Lemma 15.2. *Let $J = [\mathbf{a}, \mathbf{b}]$ be an $n-$box, and let $\mathbf{x}$ be a point in J. Then we have*

$$\mathbf{x} = \sum_{\mathbf{c} \in v(J)} \mathbf{c} \cdot \prod_{i=1}^{n} K(c_i, x_i).$$

Proof. Let $j \in \{1, 2, \ldots, n\}$. If J is a non-degenerated $n-$box, then

$$\begin{aligned} &\sum_{\mathbf{c} \in v(J)} c_j \prod_{i=1}^{n} K(c_i, x_i) \\ &\quad = \sum_{\substack{\mathbf{c} \in v(J) \\ c_1 = a_1}} c_j K(a_1, x_1) \prod_{i=2}^{n} K(c_i, x_i) + \sum_{\substack{\mathbf{c} \in v(J) \\ c_1 = b_1}} c_j K(b_1, x_1) \prod_{i=2}^{n} K(c_i, x_i). \end{aligned}$$

Since

$$\sum_{\substack{\mathbf{c} \in v(J) \\ c_1 = a_1}} c_j \prod_{i=2}^{n} K(c_i, x_i) = \sum_{\substack{\mathbf{c} \in v(J) \\ c_1 = b_1}} c_j \prod_{i=2}^{n} K(c_i, x_i),$$

we have

$$\sum_{\mathbf{c}\in v(J)} c_j \prod_{i=1}^{n} K(c_i,x_i) = [K(a_1,x_1)+K(b_1,x_1)] \sum_{\substack{\mathbf{c}\in v(J)\\ c_1=a_1}} c_j \prod_{i=2}^{n} K(c_i,x_i)$$

$$= \sum_{\substack{\mathbf{c}\in v(J)\\ c_1=a_1}} c_j \prod_{i=2}^{n} K(c_i,x_i).$$

By using similar arguments—by repeating the same process or by induction—we have

$$\sum_{\mathbf{c}\in v(J)} c_j \prod_{i=1}^{n} K(c_i,x_i) = \sum_{\substack{\mathbf{c}\in v(J)\\ (c_1,c_2,\dots,c_k)=(a_1,a_2,\dots,a_k)}} c_j \prod_{i=k+1}^{n} K(c_i,x_i)$$

as long as $k < j$, and

$$\sum_{\mathbf{c}\in v(J)} c_j \prod_{i=1}^{n} K(c_i,x_i) = \sum_{\substack{\mathbf{c}\in v(J)\\ (c_1,c_2,\dots,c_{j-1})=(a_1,a_2,\dots,a_{j-1})}} c_j \prod_{i=j}^{n} K(c_i,x_i)$$

$$= \sum_{\substack{\mathbf{c}\in v(J)\\ (c_1,c_2,\dots,c_{j-1},c_j)=(a_1,a_2,\dots,a_{j-1},a_j)}} \left[K(a_j,x_j) \prod_{i=j+1}^{n} K(c_i,x_i)\right] a_j$$

$$+ \sum_{\substack{\mathbf{c}\in v(J)\\ (c_1,c_2,\dots,c_{j-1},c_j)=(a_1,a_2,\dots,a_{j-1},b_j)}} \left[K(b_j,x_j) \prod_{i=j+1}^{n} K(c_i,x_i)\right] b_j$$

$$= [a_jK(a_j,x_j)+b_jK(b_j,x_j)] \sum_{\substack{\mathbf{c}\in v(J)\\ (c_1,c_2,\dots,c_j)=(a_1,a_2,\dots,a_j)}} \left[\prod_{i=j+1}^{n} K(c_i,x_i)\right]$$

$$= \left[a_j\frac{b_j-x_j}{b_j-a_j} + b_j\frac{x_j-a_j}{b_j-a_j}\right] \sum_{\substack{\mathbf{c}\in v(J)\\ (c_1,c_2,\dots,c_j)=(a_1,a_2,\dots,a_j)}} \left[\prod_{i=j+1}^{n} K(c_i,x_i)\right]$$

$$= x_j \sum_{\substack{\mathbf{c}\in v(J)\\ (c_1,c_2,\dots,c_j)=(a_1,a_2,\dots,a_j)}} \left[\prod_{i=j+1}^{n} K(c_i,x_i)\right].$$

By continuing the process for $(c_1,c_2,\dots,c_k) = (a_1,a_2,\dots,a_k)$, with $k > j$, we easily obtain the result.

If J is degenerated, then there exists $i \in \{1,2,\dots,n\}$ such that $a_i = b_i$ and, in such a case, there is a unique $K(c_i,x_i)$ for that i, and $K(c_i,x_i) = 1$; therefore, we obtain the same result, which concludes the proof.

We now give an interesting property of the standard extension.

Lemma 15.3. *Let $J=[\mathbf{a},\mathbf{b}]$ be an $n-$box, let $f\colon v(J)\longrightarrow\mathbb{R}$ be a function, and let E be a standard extension from f to J. If $1\le j\le n$ is such that $a_j<b_j$, then*

$$E(\mathbf{x})=\frac{b_j-x_j}{b_j-a_j}E(\mathbf{x}_{a_j})+\frac{x_j-a_j}{b_j-a_j}E(\mathbf{x}_{b_j}).$$

where $\mathbf{x}_{c_j}=(x_1,\ldots,x_{j-1},c_j,x_{j+1},\ldots,x_n)$ for $c_j=a_j,b_j$.

Proof. Firstly, note

$$K(c_j,a_j)=\begin{cases}1 \text{ si } c_j=a_j\\ 0 \text{ si } c_j=b_j\end{cases}\quad\text{and}\quad K(c_j,b_j)=\begin{cases}0 \text{ si } c_j=a_j\\ 1 \text{ si } c_j=b_j,\end{cases}$$

so that

$$\begin{aligned}\frac{b_j-x_j}{b_j-a_j}E(\mathbf{x}_{a_j})+\frac{x_j-a_j}{b_j-a_j}E(\mathbf{x}_{b_j})&=K(a_j,x_j)\sum_{\substack{\mathbf{c}\in v(J)\\ c_j=a_j}}\left[\prod_{\substack{i=1\\ i\neq j}}^{n}K(c_i,x_i)\right]f(\mathbf{c})\\ &\quad+K(b_j,x_j)\sum_{\substack{\mathbf{c}\in v(J)\\ c_j=b_j}}\left[\prod_{\substack{i=1\\ i\neq j}}^{n}K(c_i,x_i)\right]f(\mathbf{c})\\ &=\sum_{\mathbf{c}\in v(J)}\left[\prod_{i=1}^{n}K(c_i,x_i)\right]f(\mathbf{c})=E(\mathbf{x}),\end{aligned}$$

which completes the proof.

Remark 15.2. We want to stress that Lemma 15.3 provides an important fact: When making the standard extension from f to any $n-$box J, that extension is compatible with all the extensions of 1-boxes that could be done in J; and, thus, it is also compatible with the extensions that can be done with m-boxes, $m<n$, in J.

We are now in position to prove Theorem 15.3.

Proof (of Theorem 15.3). Let $S:\prod_{i=1}^n A_i\longrightarrow\mathbb{I}$ be an $n-$subcopula. To extend S to an $n-$copula C, we follow five steps:

1. We extend S to $\prod_{i=1}^n\overline{A_i}$ by continuity, where $\overline{A_i}$ denotes the closure of A_i). This extension will be an $n-$subcopula as well.
   We know that $\prod_{i=1}^n A_i$ is non-empty set since both 0 and 1 are in A_i for every $i=1,2,\ldots,n$. Let $\mathbf{x}\in\prod_{i=1}^n\overline{A_i}$—note that $\prod_{i=1}^n\overline{A_i}=\overline{\prod_{i=1}^n A_i}$. Then there exists a sequence $\{\alpha_m\}=\{(a_1,a_2,\ldots,a_n)_m\}$ of elements of $\prod_{i=1}^n A_i$ such that $\{\alpha_m\}$ converges to $\mathbf{x}$—denoted by $\{\alpha_m\}\to\mathbf{x}$. It is clear that $\{\alpha_m\}$ is a Cauchy sequence and, since S is uniformly continuous, we have that $\{S(\alpha_m)\}$ is a Cauchy sequence in $\mathbb{I}$. Since $\mathbb{I}$ is a complete set, there exists $u\in\mathbb{I}$ such that $\{S(\alpha_m)\}\to u$. We note that u depends only on $\mathbf{x}$, since if $\{\beta_m\}=\{(b_1,b_2,\ldots,b_n)_m\}$ is another

sequence of elements of $\prod_{i=1}^{n} A_i$ such that $\{\beta_m\} \to \mathbf{x}$ and $\{S(\beta_m)\} \to z \in \mathbb{I}$, then $u = z$, because S is uniformly continuous and, for every $\varepsilon > 0$, there exists $\delta > 0$ such that if $||\mathbf{x}-\mathbf{y}|| < \delta$ then $|S(\mathbf{x}) - S(\mathbf{y})| < \varepsilon/3$. Now then, there also exists a natural number m_1 such that if $m > m_1$, then $||\alpha_m - \mathbf{x}|| < \delta/2$ and $||\beta_m - \mathbf{x}|| < \delta/2$. Therefore, $||\beta_m - \alpha_m|| < \delta$ and thus $|S(\alpha_m) - S(\beta_m)| < \varepsilon/3$ for $m > m_1$. On the other hand, there exists another natural number m_2 such that if $m > m_2$ then $|S(\alpha_m) - u| < \varepsilon/3$ and $|S(\beta_m) - z| < \varepsilon/3$. In short, if $m > m_0 = \max\{m_1, m_2\}$, then we have $|u - z| \le |u - S(\alpha_m)| + |S(\alpha_m) - S(\beta_m)| + |S(\beta_m) - z| < \varepsilon$ for all $\varepsilon > 0$, that is, $u = z$. We denote by $S^*(\mathbf{x})$ the unique element u obtained from $\mathbf{x}$; so that we can extend S to a function $S^*\colon \prod_{i=1}^{n} \overline{A_i} \longrightarrow \mathbb{I}$. Since S is an $n-$subcopula and S^* extends to S by continuity, we obtain that S^* is an $n-$subcopula.

2. For every non$-$degenerated $n-$box J of $\mathbb{I}^n$ such that all its vertices—and only its vertices—belong to $\prod_{i=1}^{n} \overline{A_i}$, we extend S^* to S_2 according the standard extension. The same process is done with S_2 and the degenerated $n-$boxes in which $a_i = b_i$ for a unique i; and so on up to complete the extension to $\mathbb{I}^n$.
3. All possible extensions from the previous step, and taking into account Lemma 15.3, lead to an extension C defined on $\mathbb{I}^n$ for which the values of the extension coincide where two $n-$boxes overlap.
4. We check that C satisfies the boundary conditions of a copula.
   Let $(a_1, a_2, \ldots, a_n)$ be a point in $\mathbb{I}^n$ such that $a_i = 0$ for some i. Suppose there exists a unique $h \in \{1, 2, \ldots, n\}$ such that a_h is not in $\overline{A_h}$. Then we have

$$a_h = \frac{w - a_h}{w - t} t + \frac{a_h - t}{w - t} w,$$

   where $t = \max\{x \in \overline{A_h} | x < a_h\}$ and $w = \min\{x \in \overline{A_h} | x > a_h\}$. Note that both t and w are in $\overline{A_h}$. Suppose $h < i$ (similarly if $h > i$). Then, by applying Lemma 15.3, we have

$$\begin{aligned} & C(a_1, \ldots, a_h, \ldots, a_{i-1}, 0, a_{i+1}, \ldots, a_n) \\ & \quad = \frac{w - a_h}{w - t} S^*(a_1, \ldots, t, \ldots, a_{i-1}, 0, a_{i+1}, \ldots, a_n) \\ & \qquad + \frac{a_h - t}{w - t} S^*(a_1, \ldots, w, \ldots, a_{i-1}, 0, a_{i+1}, \ldots, a_n) \\ & \quad = \frac{w - a_h}{w - t} 0 + \frac{a_h - t}{w - t} 0 = 0, \end{aligned}$$

   since C coincides with S^* in those points whose coordinates belong to $\overline{A_h}$. When the number of $h's$ such that a_h is not in $\overline{A_h}$ is greater than one, we lead to the same conclusion by induction.
   Consider now a point $(1, \ldots, 1, a_h, 1, \ldots, 1)$ in $\mathbb{I}^n$ such that a_h is not $\overline{A_h}$ (if a_h is in $\overline{A_h}$, there is nothing to prove). Then we have

$$\begin{aligned} & C(1,\ldots,1,a_h,1,\ldots,1) \\ & \quad = \frac{w-a_h}{w-t}S^*(1,\ldots,1,t,1,\ldots,1)+\frac{a_h-t}{w-t}S^*(1,\ldots,1,w,1,\ldots,1) \\ & \quad = \frac{w-a_h}{w-t}t+\frac{a_h-t}{w-t}w=\frac{a_h(w-t)}{w-t}=a_h. \end{aligned}$$

5. We check that C is $n-$increasing.
Let $\mathscr{I}_s$ be the class of the intervals $[a,b]$ such that a and b are in $\overline{A_s}$, and let $\mathscr{J}_s$ be the class of the intervals that are not in $\mathscr{I}_s$. Let $J=[\mathbf{a},\mathbf{b}]$ be a non$-$degenerated $n-$box in $\mathbb{I}^n$, $[a_s,b_s]$ is in $\mathscr{J}_s$, $t=\max\{x\in\overline{A_s}|x\le a_s\}$, $u=\min\{x\in\overline{A_s}|x\ge a_s\}$, $v=\max\{x\in\overline{A_s}|x\le b_s\}$, and $w=\min\{x\in\overline{A_s}|x\ge b_s\}$. It is clear $t<w$, $t\le v$, and $u\le w$. Then, to prove that C is $n-$increasing, we apply induction on the number of intervals $[a_s,b_s]$ in $\mathscr{J}_s$ of the $n-$box J. For $k=0$, we have that $[a_s,b_s]$ is in $\mathscr{I}_s$ for all s, so that $V_C(J)=V_{S^*}(J)\ge 0$, since S^* is an $n-$subcopula. Suppose now the result is true for $k=i<n$, and let us check that it also holds for $k=i+1$. In this case, J has $i+1$ intervals $[a_s,b_s]$ in $\mathscr{J}_s$, and suppose that $[a_j,b_j]$ is in $\mathscr{J}_j$. We study two cases, where we use the notation $\mathbf{x}_{j,y}=(x_1,\ldots,x_{j-1},y,x_{j+1},\ldots,x_n)$, for $\mathbf{x}=\mathbf{a},\mathbf{b},\mathbf{c}$ and $y=a_j,b_j,t,u,v,w$.

   a. $\overline{A_j}\cap(a_j,b_j)=\emptyset$.
   In this case, we have $t\le a_j<b_j\le w$. Then, by applying Lemma 15.3, we obtain

$$\begin{aligned} V_C(J) &= \sum_{\mathbf{c}\in v(J)}\operatorname{sgn}(\mathbf{c})C(\mathbf{c}) \\ &= \sum_{\mathbf{c}\in v(J)}\operatorname{sgn}(\mathbf{c})\left[\frac{w-c_j}{w-t}C(\mathbf{c}_{j,t})+\frac{c_j-t}{w-t}C(\mathbf{c}_{j,w})\right]. \end{aligned}$$

   Therefore,

$$\begin{aligned} V_C(J) = &\sum_{\substack{\mathbf{c}\in v(J)\\ c_j=a_j}}\operatorname{sgn}(\mathbf{c})\left[\frac{w-a_j}{w-t}C(\mathbf{c}_{j,t})+\frac{a_j-t}{w-t}C(\mathbf{c}_{j,w})\right] \\ &+\sum_{\substack{\mathbf{c}\in v(J)\\ c_j=b_j}}\operatorname{sgn}(\mathbf{c})\left[\frac{w-b_j}{w-t}C(\mathbf{c}_{j,t})+\frac{b_j-t}{w-t}C(\mathbf{c}_{j,w})\right]. \end{aligned}$$

   Consider the $n-$box $J_{t,w}=[\mathbf{a}_{j,t},\mathbf{b}_{j,w}]$, which has only i intervals satisfying $[a_s,b_s]\in\mathscr{J}_s$, and whose vertices are $\mathbf{c}_{j,t}$ and $\mathbf{c}_{j,w}$. For the first sum we have $\operatorname{sgn}(\mathbf{c})=\operatorname{sgn}(\mathbf{c}_{j,t})=-\operatorname{sgn}(\mathbf{c}_{j,w})$, and for the second sum we have $\operatorname{sgn}(\mathbf{c})=-\operatorname{sgn}(\mathbf{c}_{j,t})=\operatorname{sgn}(\mathbf{c}_{j,w})$—note that the sign of $\mathbf{c}$ is considered with respect to the $n-$box J, while the signs of $\mathbf{c}_{j,t}$ and $\mathbf{c}_{j,w}$ are considered with respect to the $n-$box $J_{t,w}$. Thus, we obtain

$$V_C(J) = \sum_{\mathbf{c}_{j,t}\in v(J_{t,w})} \operatorname{sgn}(\mathbf{c}_{j,t}) C(\mathbf{c}_{j,t}) \left[\frac{w-a_j}{w-t} - \frac{w-b_j}{w-t}\right] + \sum_{\mathbf{c}_{j,w}\in v(J_{t,w})} \operatorname{sgn}(\mathbf{c}_{j,w}) C(\mathbf{c}_{j,w}) \left[\frac{b_j-t}{w-t} - \frac{a_j-t}{w-t}\right] = \frac{b_j-a_j}{w-t} V_C(J_{t,w}) \geq 0,$$

by hypothesis of induction.

b. $\overline{A_j} \cap (a_j, b_j) \neq \emptyset$.

Firstly, we write the $C-$volume of the $n-$box J as

$$V_C(J) = \sum_{\substack{\mathbf{c}\in v(J)\\ c_j=a_j}} \operatorname{sgn}(\mathbf{c}) C(\mathbf{c}) + \sum_{\substack{\mathbf{c}\in v(J)\\ c_j=b_j}} \operatorname{sgn}(\mathbf{c}) C(\mathbf{c}).$$

Now, suppose $a_j, b_j \notin \overline{A_j}$, then $t < a_j < u \leq v < b_j < w$. By applying Lemma 15.3 we have

$$C\left(\mathbf{c}_{j,a_j}\right) = \frac{u-a_j}{u-t} C(\mathbf{c}_{j,t}) + \frac{a_j-t}{u-t} C(\mathbf{c}_{j,u}) \tag{15.2}$$

and

$$C\left(\mathbf{c}_{j,b_j}\right) = \frac{w-b_j}{w-v} C(\mathbf{c}_{j,v}) + \frac{b_j-v}{w-v} C(\mathbf{c}_{j,w}). \tag{15.3}$$

Thus, we obtain

$$V_C(J) = \sum_{\substack{\mathbf{c}\in v(J)\\ c_j=a_j}} \operatorname{sgn}(\mathbf{c}) \frac{u-a_j}{u-t} C(\mathbf{c}_{j,t}) + \sum_{\substack{\mathbf{c}\in v(J)\\ c_j=a_j}} \operatorname{sgn}(\mathbf{c}) \frac{a_j-t}{u-t} C(\mathbf{c}_{j,u}) + \sum_{\substack{\mathbf{c}\in v(J)\\ c_j=b_j}} \operatorname{sgn}(\mathbf{c}) \frac{w-b_j}{w-v} C(\mathbf{c}_{j,v}) + \sum_{\substack{\mathbf{c}\in v(J)\\ c_j=b_j}} \operatorname{sgn}(\mathbf{c}) \frac{b_j-v}{w-v} C(\mathbf{c}_{j,w}).$$

Consider the following $n-$boxes: $J_{t,u} = [\mathbf{a}_{j,t}, \mathbf{b}_{j,u}]$, $J_{u,v} = [\mathbf{a}_{j,u}, \mathbf{b}_{j,v}]$ and $J_{v,w} = [\mathbf{a}_{j,v}, \mathbf{b}_{j,w}]$. Note that the $n-$boxes $J_{t,u}$ and $J_{v,w}$ are non-degenerated, but $J_{u,v}$ may be degenerated if $u = v$. The three $n-$boxes have exactly i intervals satisfying that $[a_s, b_s]$ is in $\mathscr{J}_s$. We denote by $\mathbf{c}_{t,u}$, $\mathbf{c}_{u,v}$ and $\mathbf{c}_{v,w}$ their respective vertices, and by $(c_{p,q})_j$ the element of $\mathbf{c}_{p,q}$ in position j. Then, clearly

$$\sum_{\substack{\mathbf{c}\in v(J)\\ c_j=a_j}} \operatorname{sgn}(\mathbf{c})\frac{u-a_j}{u-t}C(\mathbf{c}_{j,t}) = \frac{u-a_j}{u-t}\sum_{\substack{\mathbf{c}_{t,u}\in v\left(J_{t,u}\right)\\ (c_{t,u})_j=t}} \operatorname{sgn}(\mathbf{c}_{t,u})C(\mathbf{c}_{j,t})$$

$$= \frac{u-a_j}{u-t}\left[V_C(J_{t,u}) - \sum_{\substack{\mathbf{c}_{t,u}\in v\left(J_{t,u}\right)\\ (c_{t,u})_j=u}} \operatorname{sgn}(\mathbf{c}_{t,u})C(\mathbf{c}_{j,u})\right]$$

and

$$\sum_{\substack{\mathbf{c}\in v(J)\\ c_j=b_j}} \operatorname{sgn}(\mathbf{c})\frac{b_j-v}{w-v}C(\mathbf{c}_{j,w}) = \frac{b_j-v}{w-v}\sum_{\substack{\mathbf{c}_{v,w}\in v\left(J_{v,w}\right)\\ (c_{v,w})_j=w}} \operatorname{sgn}(\mathbf{c}_{v,w})C(\mathbf{c}_{j,w})$$

$$= \frac{b_j-v}{w-v}\left[V_C(J_{v,w}) - \sum_{\substack{\mathbf{c}_{v,w}\in v\left(J_{v,w}\right)\\ (c_{v,w})_j=v}} \operatorname{sgn}(\mathbf{c}_{v,w})C(\mathbf{c}_{j,v})\right].$$

Moreover,

$$\sum_{\substack{\mathbf{c}\in v(J)\\ c_j=a_j}} \operatorname{sgn}(\mathbf{c})\frac{a_j-t}{u-t}C(\mathbf{c}_{j,u}) = \frac{a_j-t}{u-t}\sum_{\substack{\mathbf{c}_{t,u}\in v\left(J_{t,u}\right)\\ (c_{u,v})_j=u}} -\operatorname{sgn}(\mathbf{c}_{t,u})C(\mathbf{c}_{j,u})$$

and

$$\sum_{\substack{\mathbf{c}\in v(J)\\ c_j=b_j}} \operatorname{sgn}(\mathbf{c})\frac{w-b_j}{w-v}C(\mathbf{c}_{j,v}) = \frac{w-b_j}{w-v}\sum_{\substack{\mathbf{c}_{v,w}\in v\left(J_{v,w}\right)\\ (c_{v,w})_j=v}} -\operatorname{sgn}(\mathbf{c}_{v,w})C(\mathbf{c}_{j,v}).$$

Thus, by summing the four previous terms, we obtain

$$
\begin{aligned}
V_C(J) = {} & \frac{u-a_j}{u-t} V_C\left(J_{t,u}\right) + \frac{b_j - v}{w-v} V_C\left(J_{v,w}\right) \\
& - \left[\sum_{\substack{\mathbf{c}_{t,u} \in v\left(J_{t,u}\right) \\ \left(c_{t,u}\right)_j = u}} \operatorname{sgn}\left(\mathbf{c}_{t,u}\right) C\left(\mathbf{c}_{j,u}\right) \right] \left(\frac{u-a_j}{u-t} + \frac{a_j - t}{u-t} \right) \\
& - \left[\sum_{\substack{\mathbf{c}_{v,w} \in v\left(J_{v,w}\right) \\ \left(c_{v,w}\right)_j = v}} \operatorname{sgn}\left(\mathbf{c}_{v,w}\right) C\left(\mathbf{c}_{j,v}\right) \right] \left(\frac{b_j - v}{w-v} + \frac{w-b_j}{w-v} \right) \\
= {} & \frac{u-a_j}{u-t} V_C\left(J_{t,u}\right) + \frac{b_j - v}{w-v} V_C\left(J_{v,w}\right) \\
& + \sum_{\substack{\mathbf{c}_{u,v} \in v\left(J_{u,v}\right) \\ \left(c_{u,v}\right)_j = u}} \operatorname{sgn}\left(\mathbf{c}_{u,v}\right) C\left(\mathbf{c}_{j,u}\right) \\
& + \sum_{\substack{\mathbf{c}_{u,v} \in v\left(J_{u,v}\right) \\ \left(c_{u,v}\right)_j = v}} \operatorname{sgn}\left(\mathbf{c}_{u,v}\right) C\left(\mathbf{c}_{j,v}\right) \\
= {} & \frac{u-a_j}{u-t} V_C\left(J_{t,u}\right) + V_C\left(J_{u,v}\right) + \frac{b_j - v}{w-v} V_C\left(J_{v,w}\right) \geq 0,
\end{aligned}
$$

by hypothesis of induction.
Note that the previous reasoning has been done by assuming that the $n-$box J^2 is non-degenerated, i.e., with $u < v$. A light modification in the proof leads to the same result when $u = v$.
On the other hand, we observe that the proof would be simplified in the case that a_j or b_j belong to $\overline{A_j}$. If a_j is in $\overline{A_j}$, then the decomposition (15.2) would not be necessary, since $t = a_j = u$; if $b_j \in \overline{A_j}$, it would not be necessary to apply (15.3); and the rest of the proof would be similar.

The proof is complete.

Remark 15.3. We stress that not only does Theorem 15.3 show that an $n-$subcopula S can be extended to an $n-$copula, but also C; the result shows that if $a_j = \max\left\{x\colon x \in \overline{A_j} \text{ and } x \leq x_j\right\}$ and $b_j = \min\left\{x\colon x \in \overline{A_j} \text{ and } x_j \leq x\right\}$, then

$$
C(\mathbf{x}) = \sum_{\mathbf{c} \in v(J_{\mathbf{x}})} \prod_{\substack{c_i = a_i \\ a_i < b_i}} \frac{b_i - x_i}{b_i - a_i} \prod_{\substack{c_i = b_i \\ a_i < b_i}} \frac{x_i - a_i}{b_i - a_i} S^*(\mathbf{c})
$$

is an extension of S, where $J_{\mathbf{x}} := \prod_{i=1}^n [a_i, b_i]$ (the empty product is interpreted as 1).

15.3.1 An alternative proof of Extension theorem by using Zorn's lemma

Along the proof of Theorem 15.3, it has been essential the use of the linear interpolation in one variable showed in Lemma 15.3. Next, we sketch out—we refer [11] for a complete study—that if we use this interpolation in an alternative way, the proof of Theorem 15.3 can be shortened by using the Zorn's lemma, although we do not obtain an explicit expression for the $n-$copula C. This is what it happens with the rest of the known proofs of the Sklar's theorem that we discuss in the last section of this chapter.

We begin with some basic definitions. Let P be a *poset* (partially ordered set), i.e., a (non-empty) set taken together with a partial order "$\leq$" on it. A *chain* in P is a totally ordered subset, i.e., a subset in which any two elements are comparable. An element $m \in P$ is called *maximal* if there is no $x \in P$ with $x > m$.

We now recall Zorn's lemma [14].

Lemma 15.4 (Zorn). *Let P be a poset in which every chain has an upper bound. Then P has at least one maximal element.*

Given two subcopulas S_1 and S_2 with respective domains D_1 and D_2 in $\mathbb{I}^n$, we say that S_2 *extends* S_1 if $D_1 \subset D_2$ and $S_1(\mathbf{u}) = S_2(\mathbf{u})$ when $\mathbf{u} \in D_1$.

Let $A_1, A_2, \ldots, A_n$ be n subsets of $\mathbb{I}$ containing 0 and 1. Given an $n-$subcopula $T\colon \prod_{i=1}^n A_i \longrightarrow \mathbb{I}$, let $\mathscr{S}_T$ denote the set of all $n-$subcopulas which extend T. Then we have the following lemma, whose proof can be found in [11].

Lemma 15.5. *In $\mathscr{S}_T$, every chain has an upper bound.*

As a consequence of Lemmas 15.4 and 15.5, we have the following result.

Corollary 15.1. *$\mathscr{S}_T$ has at least one maximal element.*

The next result shows how to extend an $n-$subcopula to another $n-$subcopula.

Lemma 15.6. *Let $S\colon \prod_{i=1}^n A_i (= D) \longrightarrow \mathbb{I}$ be an $n-$subcopula such that $D \neq \mathbb{I}^n$. Then S can be extended to an $n-$subcopula S' with $D \subsetneq Dom(S')$.*

Proof. If D is not a closed set, then we extend the $n-$subcopula S to an $n-$subcopula S' on its closure.

If D is a closed set, there exists an index i such that $A_i \neq \mathbb{I}$—otherwise, Sklar's theorem would be proved. Suppose, without loss of generality, $i = 1$. Let p, q be two real numbers such that both are in A_1 and $]p, q[\cap A_1 = \emptyset$. We define $r = (p+q)/2$, $A_1' = A_1 \cup \{r\}$, and the function $S'\colon A_1' \times \prod_{i=2}^n A_i \longrightarrow \mathbb{I}$ given by

$$S'(\mathbf{u}) = \begin{cases} S(\mathbf{u}), & u_1 \neq r, \\ \dfrac{S(p, \mathbf{u}_1) + S(q, \mathbf{u}_1)}{2}, & u_1 = r, \end{cases}$$

where $\mathbf{u}_1 = (u_2, \ldots, u_n)$. It is easy to check that $V_S(B) = V_{S'}(B)$ when the $n-$box B has its vertices in D. In the case that $B = [(r, \mathbf{u}_1), \mathbf{u}^*]$, we consider the $n-$boxes $B_1' = [(p, \mathbf{u}_1), (q, \mathbf{u}_1^*)]$, $B_1 = [(r, \mathbf{u}_1), (q, \mathbf{u}_1^*)]$, and $B_2 = [(q, \mathbf{u}_1), \mathbf{u}^*)]$, it is clear that $B = B_1 \cup B_2$ and $\mathrm{int}(B_1) \cap \mathrm{int}(B_2) = \emptyset$, and we have

$$V_{S'}(B) = V_{S'}(B_1) + V_S(B_2) = \frac{V_S(B_1')}{2} + V_S(B_2) \geq 0.$$

If $B = [\mathbf{u}^*, (r, \mathbf{u}_1)]$, by using a similar reasoning, we obtain the same result, and we conclude that S' is an $n-$subcopula.

From Corollary 15.1, we know that there exists a maximal element C which extends an $n-$subcopula S. If C is not an $n-$copula, then, from Lemma 15.6, C can be extended, but this contradicts the maximality of C; thus, we obtain again Sklar's theorem.

15.4 Other proofs of Sklar's theorem

After Schweizer and Sklar's works in [20, 22], different proofs of Sklar's theorem have been published in the literature. These publications use a wide variety of viewpoints and mathematical tools that we now review.

The first one is due to Moore and Spruill [15], and is based on probabilistic arguments (see also [5]). Similar ideas can be found in the proof given by Rüschendorf [18] (see [17] for additional details of this proof). Specifically, let $\mathbf{X} = (X_1, X_2, \ldots, X_n)$ be a random vector on a probability space $(\Omega, \mathscr{A}, P)$ with distribution function F, and let $V_1, V_2, \ldots, V_n$ be n independent random variables uniformly distributed on $\mathbb{I}$ and independent of $\mathbf{X}$. Then the basis of the proof is that the distribution functions of the random variables $U_i = (1 - V_i)F_i(X_i^-) + V_i F_i(X_i)$ are uniformly distributed on $\mathbb{I}$.

A novel idea appears in the work due to Carley and Taylor [4] (see also [2]). It deals with the use of the compactness of $\mathscr{C}_n$ endowed with the topology induced by the uniform metric d. For that, a sequence $\{C_n\}$ of copulas (checkerboard) is constructed by using the subcopula S given in Theorem 15.2. The compactness of $(\mathscr{C}_n, d)$ implies that there exists, at least, a copula C which is the limit of a subsequence of the sequence of approximations of S. The copula C has the property of extending to S, and thus obtaining Sklar's theorem. Another proof based on the compactness of $(\mathscr{C}_n, d)$ can be found in [9].

Another idea of proof, which uses the compactness as well, is based on the regularization of distribution functions. This idea appears in the works due to Durante *et al.* [7] (see also [8]) and Faugeras [10]. In [7], the convolutions of F with the functions $\varphi_m(\mathbf{x}) := m^n \varphi(m\mathbf{x})$ is used, where m is a positive integer and $\varphi\colon \overline{\mathbb{R}}^n \longrightarrow \overline{\mathbb{R}}$ is the function defined by $\varphi(\mathbf{x}) := k \exp\left(\frac{1}{|\mathbf{x}|^2 - 1}\right) \mathbf{1}_{B_1(\mathbf{0})}(\mathbf{x})$, where $\mathbf{1}_{B_1(\mathbf{0})}(\mathbf{x})$ denotes the characteristic function of the unit open ball and the constant k is such that the L^1 norm $\|\varphi\|_1$ of φ is equal to 1. In [10], the regularization is given by a continuous random variable Z, independent of a random variable X and the definition of the random variables $X_m = X + h_m Z$, $h_m \in \mathbb{R}$, $h_m \downarrow 0$.

In the context of rearrangement inequalities, Burchard and Hajaiej [3] prove a result which contains as a particular case the Sklar's theorem. The proof is based on the reiterated application of an univariate result. To be exact, let g be a nondecreasing real-valued function defined on an interval I. Then, for every nondecreasing function f on I satisfying $|f(z) - f(y)| \leq K(g(z) - g(y))$ for all points $y < z \in I$ with some constant K, there exists a Lipschitz continuous nondecreasing function $h\colon \mathbb{R} \longrightarrow [\inf f, \sup f]$ such that $f = h \circ g$.

The last proof can be found in [1], altough this paper has not been published yet. As main tool, it uses the density of the subcopulas defined on a net of points in the metric space $(\mathscr{S}', \xi)$, where, for every S_1 and S_2 in $\mathscr{S}$, it is defined

$$\xi\left(S_1, S_2\right) := \mathrm{d}_H\left(\mathrm{Graph}\left(S_1\right), \mathrm{Graph}\left(S_2\right)\right),$$

where $\mathrm{Graph}(S)$ denotes the graph of S and d_H is the Hausdorff distance.

Acknowledgements We have done this chapter with our best intention and we would like to dedicate it to Professor Roger B. Nelsen in occasion of his 75th birthday with our gratitude and appreciation. We have had the honor and pleasure of working together. He is a very good teacher and researcher, but, above all, he is a great person. Thanks a lot, Roger!

This contribution has been supported by the Ministerio de Economía y Competitividad (Spain) under research project MTM2014-60594-P. The authors acknowledge the comments of an anonymous referee.

References

1. E. de Amo, M. Díaz Carrilo, F. Durante, J. Fernández Sánchez, Extensions of subcopulas, submitted
2. W. Anakkamatee, S. Dhompongsa, S. Tasena, A constructive proof of the Sklar's theorem on copulas, J. Nonlinear Convex Anal. **15**(6), 1137–1145 (2014)
3. A. Burchard, H. Hajaiej, Rearrangement inequalities for functionals with monotone integrands. J. Funct. Anal. **233**(2), 561–582 (2006)
4. H. Carley, M.D. Taylor, A new proof of Sklar's theorem, in: *Distributions with Given Marginals and Statistical Modelling*, ed. by C.M. Cuadras, J. Fortiana, J.A. Rodríguez-Lallena (Kluwer Academic Publishers, Dordrecht, 2002), pp. 29–34
5. P. Deheuvels, Caractérisation complète des lois extrêmes multivariées et de la convergence des types extrêmes, Publ. Inst. Stat. Univ. Paris **23**(3–4), 1–36 (1978)
6. F. Durante, C. Sempi, *Principles of Copula Theory* (Chapman & Hall/CRC, Boca Raton, FL, 2016)
7. F. Durante, J. Fernández-Sánchez, C. Sempi, Sklar's theorem obtained via regularization techniques, Nonlinear Anal. **75**(2), 769–774 (2012)
8. F. Durante, J. Fernández-Sánchez, C. Sempi, How to prove Sklar's theorem, in *Aggregation Functions in Theory and in Practise. Proceedings of the 7th International Summer School on Aggregation Operators at the Public University of Navarra, Pamplona, Spain, July 16–20, 2013*, ed. by H. Bustince, J. Fernández, R. Mesiar, T. Calvo (Springer, Berlin, 2013), pp. 85–90

9. F. Durante, J. Fernández-Sánchez, C. Sempi, A topological proof of Sklar's theorem, Appl. Math. Lett. **26**(9), 945–948 (2013)
10. O. Faugeras, Sklar's theorem derived using probabilistic continuation and two consistency results, J. Multivar. Anal. **122**, pp. 271–277 (2013)
11. J. Fernández-Sánchez, M. Úbeda-Flores, Proving Sklar's Theorem via Zorn's Lemma. To appear in Int. J. Uncertainty, Fuzziness and Knowledge-Based Systems
12. M. Fréchet, Sur les tableaux de corrélation dont les marges sont données, Ann. Univ. Lyon 9, Sect. A, 53–77 (1951)
13. J. Galambos, *The Asymptotic Theory of Extreme Order Statistics* (John Wiley and Sons, New York, 1978)
14. S. Lang, *Real and Functional Analysis*, 3rd edn. (Springer, Berlin, 1993)
15. D.S. Moore, M.C. Spruill, Unified large-sample theory of general chi-squared statistics for tests of fit, Ann. Statist. **3**, 599–616 (1975)
16. R.B. Nelsen, *An Introduction to Copulas*, 2nd edn. (Springer, New York, 2006)
17. F. Oertel, An analysis of the Rüschendorf transform–with a view towards Sklar's Theorem, Dependence Modeling **3**(1), 113–125 (2015)
18. L. Rüschendorf, On the distributional transform, Sklar's Theorem, and the empirical copula process. J. Statist. Plan. Infer. **139**(11), 3921–3927 (2009)
19. B. Schweizer, A. Sklar, *Probabilistic Metric Spaces* (Dover Publications, New York, 2005)
20. B. Schweizer, A. Sklar, Operations on distribution functions not derivable from operations on random variables, Studia Math. **52**, 43–52 (1974)
21. A. Sklar, Fonctions de répartition à *n* dimensions et leurs marges, Publ. Inst. Statist. Univ. Paris **8**, 229–231 (1959)
22. A. Sklar, Random variables, joint distribution functions, and copulas, Kybernetika **9**, 449–460 (1973)
23. M. Úbeda-Flores, *Introducción a la Teoría de Cópulas. Aplicaciones* (Research work, University of Almería, Spain, 1998)
24. W. Whitt, Bivariate distributions with given marginals, Ann. Statist. **4**, 1280–1289 (1976)

Erratum to: When Gumbel met Galambos

Christian Genest and Johanna G. Nešlehová

Erratum to:
Chapter 6 in: M. Úbeda Flores et al. (eds.), *Copulas and Dependence Models with Applications*,
https://doi.org/10.1007/978-3-319-64221-5_6

This chapter was originally published under © Springer International Publishing AG 2017, but has now been republished under © The Author(s).

The updated online version of this chapter can be found at
https://doi.org/10.1007/978-3-319-64221-5_6

M. Úbeda Flores et al. (eds.), *Copulas and Dependence Models with Applications*,
https://doi.org/10.1007/978-3-319-64221-5_16