Multivariate Extremal Dependence and Risk Measures

by

Lei Hua

B.Econ., University of Science and Technology of China, 2002

M.Sc. University of Calgary, 2008

A THESIS SUBMITTED IN PARTIAL FULFILLMENT
OF THE REQUIREMENTS FOR THE DEGREE OF

Doctor of Philosophy

in

THE FACULTY OF GRADUATE STUDIES

(Statistics)

The University of British Columbia

(Vancouver)

June 2012

Abstract

Overlooking non-Gaussian and tail dependence phenomena has emerged as an important reason of underestimating aggregate financial or insurance risks. For modeling the dependence structures between non-Gaussian random variables, the concept of copula plays an important role and provides practitioners with promising quantitative tools. In order to study copula families that have different tail patterns and tail asymmetry than multivariate Gaussian and t copulas, we introduce the concepts of tail order and tail order functions. These provide a unified way to study three types of dependence in the tails: tail dependence, intermediate tail dependence and tail orthant independence. Some fundamental properties of tail order and tail order functions are obtained. For multivariate Archimedean copulas, we relate the tail heaviness of a positive random variable to the tail behavior of the Archimedean copula constructed by the Laplace transform of the random variable.

Quantitative risk measurements pay more attention on large losses. A good statistical approach for the whole data does not guarantee a good way for risk assessments. We use tail comonotonicity as a conservative dependence structure for modeling multivariate dependent losses. By this way, we do not lose too much accuracy but gain reasonable conservative risk measures, especially when we consider high-risk scenarios. We have conducted a thorough investigation on the properties and constructions of tail comonotonicity, and found interesting properties such as asymptotic additivity properties of risk measures. Sufficient conditions have also been obtained to justify the conservativity of tail comonotonicity.

For large losses, tail behavior of loss distributions is more critical than the whole distributions. Asymptotic study assuming that each marginal risk goes to infinity is more mathematically tractable. However, the asymptotic study that leads to a first order approximation is only a crude way and may not be sufficient. To this end, we study the second order conditions for risk measures of sub-extremal multiple risks. Some relationships between Value at Risk and Conditional Tail Expectation have been obtained under the condition of Second Order Regular Variation. We also find that the second order parameter determines whether a higher order approximation is necessary.

Preface

This thesis has been written up under the supervision of my advisor Professor Harry Joe, and it is based on four papers coauthored with Dr. Harry Joe. Two of them have been published, one is to be published and the other is under review. This thesis also contains a small portion of research that has not been published or submitted anywhere by the date of completing the thesis.

Chapter 1 was prepared for the purpose of presenting the general shape or taste of the main research within the thesis. More than half of the contents are newly written and the others are scattered in the following Chapters 3 to 6.

Chapter 3 is based on a published paper: Hua and Joe (2011a), Tail order and intermediate tail dependence of multivariate copulas. *Journal of Multivariate Analysis*, 102: 1454–1471. This chapter covers the topics in the paper but not exclusively. Dr. Joe identified the main research questions and I contributed to the paper mainly through completing the majority of the proofs, writing the first version of the manuscripts and conducting follow-up revisions. The key idea for the proof of the important Lemma 3.4 was contributed by Dr. Joe, and the proof for Proposition 3.5 also benefited from the idea.

Chapter 4 is based on a submitted paper: Hua and Joe (2012b), Tail comonotonicity: properties, constructions, and asymptotic additivity of risk measures. The concept of tail comonotonicity was partially motivated by Dr. Joe, especially with his concrete examples. I proposed some main ideas such as the asymptotic additivity of risk measures, completing almost all of the proofs, writing the first version of the manuscripts and taking care of follow-up tasks.

Chapter 5 is based on a paper in press: Hua and Joe (2012a), Tail comonotonicity and conservative risk measures. *to appear in ASTIN Bulletin*, 2012. I conducted all of the research here, including theoretical development, simulation study and data analysis under the supervision of Dr. Joe. I wrote the first version of the manuscripts and conducted follow-up tasks.

Chapter 6 is based on a published paper: Hua and Joe (2011b), Second order regular variation and conditional tail expectation of multiple risks. *Insurance Math. Econom.*, 49: 537–546, 2011. I proposed the main ideas and conducted all of the research, including mathematical proofs and simulations under the supervision of Dr. Joe. I wrote up the first version of the manuscripts and dealt with other follow-up tasks.

For all of the above referred research, in addition to the contributions mentioned, Dr. Joe spent a huge amount of time motivating research questions, checking the results and proofs, improving the writing and making all the other contributions to make the research contained in this thesis better.

Table of Contents

List of Tables

List of Figures

Glossary

cdf cumulative distribution function

ACIG Archimedean Copula based on the Laplace Transform of Inverse Gamma

MTCJ Mardia-Takahasi-Cook-Johnson copula

MRV Multivariate Regular Variation

HRV Hidden Regular Variation

2RV Second Order Regular Variation

CTE Conditional Tail Expectation

VaR Value at Risk

MDA Maximum Domain of Attraction

ALAE Allocated Loss Adjustment Expenses

LOSS Indemnity Payment

DJI Dow Jones Industrial Average Index

FTSE FTSE 100 Index

N225 NIKKEI 225 Index

AMA Advanced Measurement Approaches

LT Laplace Transform

Acknowledgments

I am greatly indebted to my supervisor Professor Harry Joe. Professor Joe is an encyclopedia to me. Every time when I asked questions, I can get prompt and instructive suggestions from him. Professor Joe has written up many inspiring notes to motivate research questions, and to help me develop my own research skills; I feel that he cares about my research and even my future careers very much. He masters the art of supervision: he gave me sufficient pressure to move forward, and also allowed me to adjust my own pace to it. With the excellent supervision of Harry, I, just like a baby who starts learning to walk, have accumulated great encouragement to step forward, because I know that a powerful arm is just around and is always ready to pull me up.

I would like to express my sincere thanks to Professors Haijun Li, Natalia Nolde and Matías Salibiá-Barrera, who are members of the Comprehensive Committee and/or the Supervision Committee. Professor Li is very warm-hearted and he has pointed out several interesting topics to motivate my research interests. Special thanks are given to Natalia; she helped me a lot for improving my writing, not only for the thesis but also for other application documents. I also want to thank Professor Christian Genest for acting as the external examiner, and Professors Edwin Perkins and Ruben Zamar for serving as the university examiners.

I want to extend my sincere gratitude to the faculty and staff of the department, to those who have directly or indirectly helped me during my study at UBC. I have taken two courses in statistical consulting with Professor John Petkau, who helped me a lot in improving my statistical thinking and writing. I also want to say Thank You to Peggy Ng, Elaine Salameh, Andrea Sollberger, and Viena Tran for lots of prompt helps and routine work.

I also thank the department Ping Pang players Professors Ruben Zamar and Lang Wu, and Chen Xu, Xu Wang and Yongliang Zhai, for many fun times, and thank the other fellow students who have helped me during these enjoyable four years.

Special thanks to Professor Ka Chun Cheung, my Master's supervisor, and the Department of Statistics and Actuarial Science at the University of Hong Kong; I was preparing the thesis proposal with the kind hospitality of Ka Chun and the financial support from Canada Graduate Scholarships - Michael Smith Foreign Study Supplements Program. This thesis has been written up with the generous financial support from the Natural Sciences and Engineering Research Council of Canada, to which I want to express my sincere thankfulness.

Lastly and most importantly, I sincerely thank my wife, Chaoxiong Xia, for her love, encouragement and support in every moment of my life, and my two little kids Claire and Daniel, for bringing me so much happiness during the journey of completing the thesis.

Dedication

献给我的爸妈

Chapter 1

Introduction

1.1 Motivation

> "The interviewed firms described some general common lessons learned from the crisis, of which probably the most commonly mentioned was the necessity (but difficulty) of capturing tail risks and dependencies resulting from tail events."
>
> *Developments in Modelling Risk Aggregation*
>
> Basel Committee on Banking Supervision, 2010

The research activities on dependence modeling of large losses and their risk assessment have been sparked by the contagion of the global financial crisis arising in the year of 2007. There have been tons of articles debating the roots of the crisis from many perspectives. One common opinion from the viewpoint of researchers in quantitative fields is that the high-risk scenarios as well as their comovements in the global financial system have been ignored, which hides the high possibility of extremal losses that could happen simultaneously. This thesis is not going to join the discussion, but attempts to contribute to the understanding of quantitative and/or statistical methodologies that account for the phenomenon of heavy tails of each individual risk and their dependence as well.

Among many issues of modeling insurance losses or financial returns, estimating, predicting and pricing potential extremal risks are particularly important.

Moreover, many extremal losses have been observed to have occurred consecutively or simultaneously. The comovement, in the meanwhile, could considerably increase the entire risk. Thus, the dependence of extremal risks is expected to play an important role for integrated risk management. Although the statistical methodologies for studying overall properties of random samples such as means are relatively mature, sound methodologies for studying the dependence relationship between high quantities are far from being well-developed.

1.2 Notation

The following notation will be used throughout the thesis.

We use bold letters for vectors, such as $\mathbf{x} := (x_1, \dots, x_d)^{\mathrm{T}}$, where T is the transpose operation of a matrix. $\mathbf{1}_d = (1, \dots, 1)$ with d elements, and then $u\mathbf{1}_d = (u, \dots, u)$; $I_d = \{1, \dots, d\}$.

$\mathbb{N}_+ = \{1, 2, \dots\}$; $\mathbb{R} := (-\infty, \infty)$; $\mathbb{R}_+ := [0, \infty)$; $\overline{\mathbb{R}}_+ := [0, \infty]$; $\overline{\mathbb{R}} := [-\infty, \infty]$; $A \subset B$ means that A is a subset (not necessarily a proper subset) of B; Cartesian product is denoted as $[\mathbf{a}, \mathbf{b}] := [a_1, b_1] \times \cdots \times [a_d, b_d]$; let $\mathbf{x} := (x_1, \dots, x_d)$ and $\mathbf{y} := (y_1, \dots, y_d)$, then $\mathbf{x} \le \mathbf{y}$ means that $x_i \le y_i$ for all $i = 1, \dots, d$.

$[x] = \max\{t \text{ integer} : t \le x\}$; $\psi^{(i)}(s)$ is the ith order derivative of ψ evaluated at s; $1_A(\mathbf{x})$ is the indicator function such that $1_A(\mathbf{x}) = 1$ if $\mathbf{x} \in A$, and $1_A(\mathbf{x}) = 0$ if $\mathbf{x} \notin A$.

For a random variable X, the generalized inverse of its cumulative distribution function (cdf) is defined as $F_X^{-1}(p) = \inf\{x \in \mathbb{R} : F_X(x) \ge p\}$. Also, note that for all $x \in \mathbb{R}$ and $p \in [0, 1]$, we have (e.g., Dhaene et al., 2002a)

$$F_X^{-1}(p) \le x \iff p \le F_X(x). \tag{1.1}$$

A random variable is said to be continuous if its cdf is continuous. The survival function is $\overline{F}_j = 1 - F_j$ if F_j is a cdf. For a d-dimensional multivariate cdf F, the survival function is $\overline{F} = 1 + (-1)^{|I|} \sum_{\emptyset \ne I \subset \{1, \dots, d\}} F_I$, where F_I is the cdf for the I-margin. $X \overset{d}{=} Y$ means X and Y equal in distribution. For a bivariate differentiable cdf F, we write $(\partial_1 F)(x_1, x_2) := \partial F(x_1, x_2)/\partial x_1 = F_{2|1}(x_2|x_1) f_1(x_1)$, where $f_1 = F_1'$. Unless otherwise mentioned, a distribution is assumed to be absolutely continuous with respect to Lebesgue measure. Throughout the thesis, we use $\mathcal{R}(F_1, \ldots, F_d)$ to represent the Fréchet space of all d-dimensional random vectors with $F_1, \ldots, F_d$ as the marginal distributions.

$T_\nu(\cdot)$ is the cdf of the standard univariate Student t distribution with degree of freedom ν, and $t_\nu(\cdot)$ is the corresponding density function. $T_{\nu,\Sigma}(\cdot)$ is the cdf of the standard multivariate Student t distribution with degree of freedom ν and covariance matrix Σ, and $t_{\nu,\Sigma}(\cdot)$ is the corresponding density function. $\Phi(\cdot)$ is the cdf of the standard univariate Normal distribution and $\phi(\cdot)$ is the density function. For the standard multivariate Normal distribution, the cdf and density function are Φ_Σ and ϕ_Σ, respectively. Throughout this thesis, ϱ will be used to represent a correlation coefficient and ρ is reserved for the second order parameter of Second Order Regular Variation (2RV).

The notation C will always represent a copula function, a multivariate cdf with $U(0, 1)$ univariate margins. The conditional distributions of a bivariate copula are written as $C_{1|2}(u|v) = \partial C(u, v)/\partial v$ and $C_{2|1}(v|u) = \partial C(u, v)/\partial u$. A density function of copula is written as $c(u, v) := \partial^2 C(u, v)/\partial u \partial v$. Second order partial derivatives are $\mathsf{D}_{22}C(u, v) = \partial^2 C(u, v)/\partial^2 v$.

For a copula, say, the Gumbel copula, the corresponding survival copula is denoted as “s.Gumbel”; that is,“s.” symbolizes the survival copula. A lower tail dependence parameter is denoted by λ_L, and an upper tail dependence parameter is denoted by λ_U. We use κ for tail order, and κ_L and κ_U are lower and upper tail orders, respectively.

For positive functions f, g, asymptotic equivalence is denoted as $f \sim g$ if $\lim f/g = 1$, asymptotic inequality $f \gtrsim g$ if $\liminf f/g \geq 1$, and asymptotic inequality $f \lesssim g$ if $\limsup f/g \leq 1$. The notation $f(t) \succsim g(t), t \to \infty$ means that f is ultimately greater than g; that is, $\exists t_0$ such that $t \geq t_0$ implies that $f(t) \geq g(t)$. $h(t) = o(g(t))$ means $\lim_{t\to t_0} h(t)/g(t) = 0$, and $h(t) = O(g(t))$ means $0 \leq \lim_{t\to t_0} h(t)/g(t) < \infty$, where $-\infty \leq t_0 \leq \infty$ is a given limiting point.

The notation RV_α represents the class of functions that are regularly varying at ∞ with index $\alpha \in \mathbb{R}$, and $\mathrm{RV}_\alpha(0^+)$ represents the class of functions that are regularly varying at 0 with index $\alpha \in \mathbb{R}$. $\mathrm{MRV}_d(\alpha)$ represents the class of d-dimensional multivariate regularly varying random vectors with $\alpha \in \mathbb{R}$ the regular variation index. $\ell(x)$ is used as a slowly varying function.

The letter $\mathbb{E}$ represents mathematical expectation, and Var is for variance. Note that the Var here is different than the notation VaR (Value at Risk) defined in Definition 2.1; the latter is used in insurance and finance. $x \vee y := \max\{x, y\}$; $x \wedge y := \min\{x, y\}$.

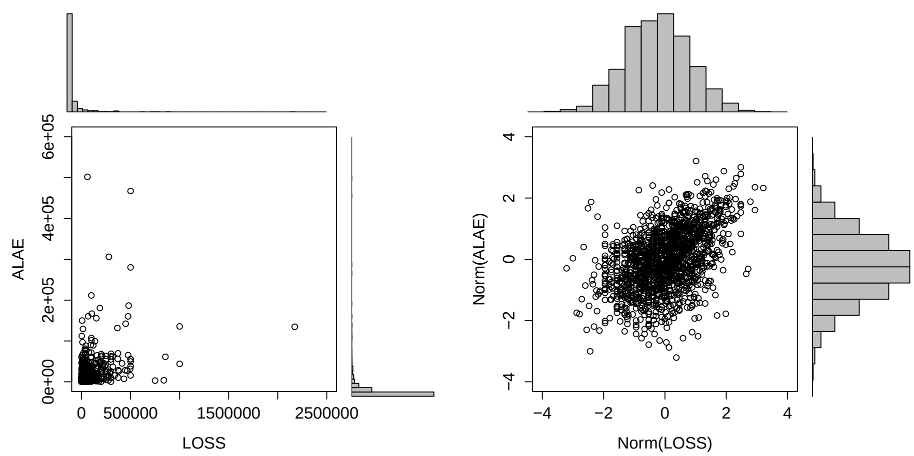
1.3 Data examples

In this subsection, we will present some data examples to motivate various patterns of tail dependence and tail asymmetry. The marginal distributions are usually transformed to a standard Normal distribution to visualize the patterns of upper and lower tails.

Let us look at a well known dataset in the actuarial literature (Frees and Valdez, 1998). This dataset comprises 1500 general liability claims with Indemnity Payment (LOSS) representing the amount of payment and Allocated Loss Adjustment Expenses (ALAE) that are specifically attributable to each claim and may include legal expenses, investigation expenses, etc. A typical pattern for this dataset is that the margins are heavy-tailed and the degree of positive dependence between large values is larger than that between small values. This kind of tail asymmetry pattern is not clear as one looks at the scatter plot for the original data (left panel of Figure 1.1). In order to visualize the asymmetric tail dependence, we usual transform each margin to be distributed as univariate standard normal, and then look at the scatter plot for the transformed pairs (right panel of Figure 1.1); such transformed values are referred to as normal scores. A stronger upper tail dependence for this dataset appears, and it makes sense since a claim with a larger loss amount tends to involve more legal expenses and demands more time and costs for investigation.

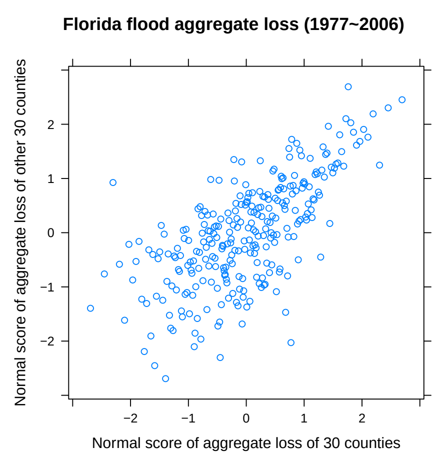
The asymmetric tail dependence pattern can also be observed in environmental data, such as a Florida flood loss dataset. There are 67 counties in Florida. The dataset comprises the monthly amount of losses for each county from 1977 to 2006, and it is a part of the Spatial Hazard Events and Losses Database for the United States (SHELDUS) that is maintained by the Hazards and Vulnerability Research Institution. We choose randomly 30 counties and add the loss amount among these counties together to get a monthly aggregate loss for these counties. Then we

Figure 1.1: Data examples - LOSS vs Expense (ALAE)



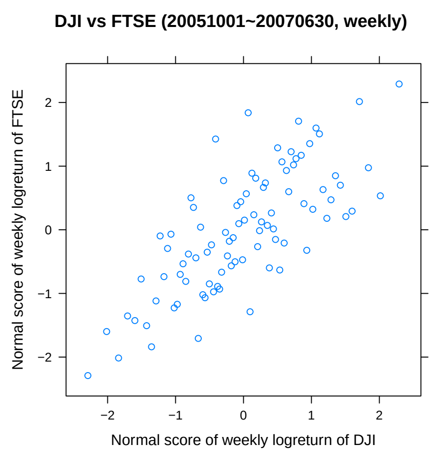
randomly choose another 30 counties from the rest counties to get the monthly aggregate loss similarly. The dependence between these two aggregate losses can be demonstrated by Figure 1.2. It is very clear that the upper tail dependence is stronger than the lower tail dependence. This phenomenon of stronger upper tail dependence for aggregate losses is related to a concept of micro-correlation that is studied in Cooke et al. (2011).

Figure 1.2: Data examples - Florida flood



In the natural world or in our human society, there seems to be certain mechanism that drives the happening of clusters of high-risk events. Not only in insurance practice, but also in the financial market, we may observe such tail asymmetry phenomena. Figure 1.3 is a scatter plot of normal scores for weekly log returns between Dow Jones Industrial Average Index (DJI) and FTSE 100 Index (FTSE) during the period of October 1st, 2005 to June 30th, 2007. The plot illustrates that the dependence between large losses of two stock markets is higher than the dependence between large returns.

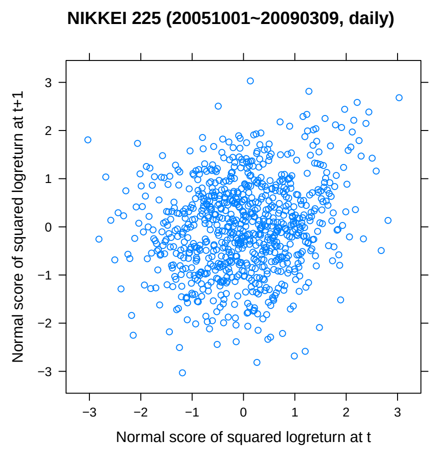
Figure 1.3: Data examples - DJI vs FTSE



Serial dependence within a univariate financial time series may also appear to be tail asymmetric. For example, if we take the squared daily log returns as approximations to observed volatilities, then the asymmetric pattern of serial dependence between such observed volatilities can be illustrated by Figure 1.4. Here we use NIKKEI 225 Index (N225) during the period of October 1st, 2005 to March 9th, 2009 as an example. The observed volatilities tend to clutter together, which has already been noticed by many researchers. We refer to Engle (2004), written by one of the two holders of Nobel Prize of 2003 in economics, for an excellent review on volatility clustering.

Note that the above scatter plots are simply used to visualize the asymmetric tail dependence patterns; those are not necessarily meaning that certain bivariate

Figure 1.4: Data examples - N225(t) vs N225$(t+1)$



distributions are ready to be used for fitting those data (or transformed data), as the data points in the plots may not even be a random sample. For specific questions, one needs to fully consider the dependence structures that saturate the data of interest to decide where a dependence model such as a copula can be appropriately applied.

1.4 Relevant practical needs

The term of tail modeling in this thesis is referred to as statistical modeling for marginal distributional tails and their dependence structure in the tails of the joint distribution.

Traditional areas that require tail modeling include rate makings for reinsurance, auto insurance, etc., and statistical modeling in finance and various subject scientific fields such as environmental science and Internet engineering. There are also some emerging needs for tail modeling in quantitative risk management for the bank and insurance sectors. For instance, under Basel III – a global regulatory standard on bank capital adequacy, stress testing and market liquidity risks – operational risk charges can be calculated based on the so-called Advanced Measurement Approaches (AMA). Operational risk is typically very heavy-tailed (Dutta and Perry, 2006), and the modeling and risk assessment is actually very challenging (Chavez-Demoulin et al., 2006; Cope et al., 2009). In parallel, under Solvency II, there are relevant regulatory requirements on operational risks for insurers.

First of all, AMA itself is far from being well developed. Sound methodologies/strategies for risk assessment and management are in high demand. Secondly, the developing market of operational risk mitigation products would entail huge potential for research on tail modeling; for example, how to conduct tail risk securitization and how to design and price relevant products?

1.5 Overview of literature

Before dealing with any particular models, a fundamental question is how to model the dependence between various univariate random variables, especially for those with non-Gaussian distributions. In the past practice of actuarial science, quantitative finance and so on, dependence structures between random variables were often studied with some summary quantities, such as, correlation coefficient, Kendall's τ and Spearman's ρ. Especially the usage of correlation coefficient has been popular. It is well known that the maximum correlation of 1 will be reached if and only if the two random variables have the same distribution up to a linear transformation (Rodgers and Nicewander, 1988). Shortcomings of using correlation coefficient to model the dependence for insurance and finance have been observed by many papers. We refer to Embrechts et al. (2002) for an inspiring reading. Apparently, every single summary statistic is not able to fully describe the dependence structure between random variables. Instead, the copula approach becomes more and more popular for modeling dependence structures of insurance losses or financial returns where marginal distributions are often heavy-tailed and skewed. We list here, among many of others, some advantages of copula approaches: (1) the copula itself is a joint cdf, which captures much more information than any single quantity; (2) when the univariate margins are all continuous, the copula function can be uniquely determined, thus fully characterizing the dependence structure; (3) as a multivariate cdf, copula can be parametrized and is useful for statistical modeling and inference; (4) the copula provides a great deal of flexibility for constructing multivariate distributions.

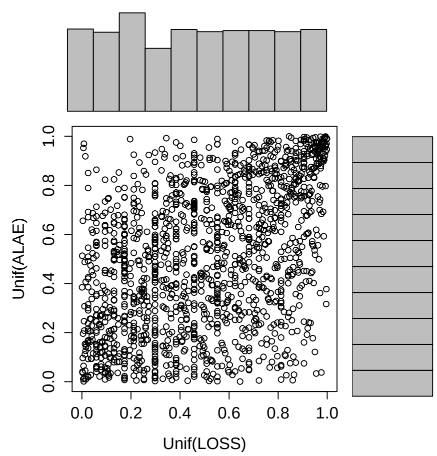
The copula itself is a joint cdf with Uniform$(0, 1)$ marginal distributions. Let

F be the joint cdf, $F_1, \ldots, F_d$ be the univariate cdfs, and C be a copula, then F can be written as

$$F(x_1, \ldots, x_d) = C(F_1(x_1), \ldots, F_d(x_d)).$$

For the insurance loss data in Figure 1.1, if we transform each marginal data via the corresponding empirical distribution to get data that is approximately uniformly distributed on $[0, 1]$, then the scatter plot for the pairs of transformed data is Figure 1.5. Roughly speaking, the copula C can be used to fit such transformed data. A more formal introduction to copulas is in Section 2.2.

Figure 1.5: Copula is a joint cdf of uniform margins



A pioneer paper on copula theory is Sklar (1959), which established the well known Sklar's theorem: there is a copula function for joint distributions and the copula is unique if the univariate marginal distributions are continuous. Standard references for dependence modeling and copulas are Joe (1997) and Nelsen (2006). The former is a very informative monograph on dependence concepts and relevant statistical modelings, where copula is one fundamental tool for statistical dependence modeling. The latter deals with the concept of copula in a more mathematical way, and focuses on relevant theories of copulas. For contributions to the classic statistical issues on copula modeling, such as statistical inference and goodness-of-fit tests, we refer the reader to the following contributions and references therein:

Genest and Rivest (1993), Genest and Favre (2007), Genest et al. (2009b) and Brechmann et al. (2012). Research activities on copula modeling in actuarial science and quantitative finance have sparked since Frees and Valdez (1998), and Li (2000). Two excellent reference books on dependence modeling in actuarial science and quantitative risk management are Denuit et al. (2005) and McNeil et al. (2005). Most recently, Kurowicka and Joe (2011) and Jaworski et al. (2010) have provided good summaries of copula theories and applications: the former promotes a relatively new concept of vine copulas, and the latter contains several survey papers on copulas.

The copula approach is particularly useful for studying tail behavior of multivariate non-Gaussian distributions (Joe, 2006; Mikosch, 2006). It is extremely flexible to use copula to construct multivariate distributions that are able to account for various distributional tails. In the practice of actuarial science and quantitative risk management, a more relevant task is to model and conduct risk assessment for the joint tails of dependent losses, especially for large or extremal risks. The classic univariate extreme value theory provides a sound theoretical ground for building up useful approaches for assessing univariate marginal tails. Compared with the univariate extreme value theory, the multivariate extreme value theory has not been well developed. In parallel with the copula approach, researchers find that transforming the marginal distributions to those having the same power law can be more convenient in some situations. This approach is referred to as Multivariate Regular Variation (MRV), and its root, the theory of univariate regular variation has been nicely developed.

There are some excellent reference books for extreme value theory, and for regular variation. de Haan and Ferreira (2006) is a nice monograph on theoretical aspects of extreme value theory and includes some discussion on 2RV, Embrechts et al. (1997) has become a classic for modeling univariate extremal events, and Coles (2001) provides a clear and more intuitive introduction to statistical modeling of univariate extreme values. The books Beirlant et al. (2004) and Reiss and Thomas (1997) comprise many interesting topics for statistical modeling of extreme values that include not only univariate but also multivariate extreme values, and Falk et al. (2010) is a comprehensive monograph on theories and applications of multivariate extreme values such as multivariate Peaks Over Threshold (POT) approaches. For researchers who study the theory of regular variation and who may use regular variation as a tool, Bingham et al. (1987) should be a must-have and it provides encyclopedic reference on the theory of univariate regular variation. For MRV, Resnick (1987) and Resnick (2007) are standard reference books, and the latter also reflects some recent development of MRV.

The above literature review is about the general idea of copula theory and extreme value theory. More specific literature reviews for concrete topics that belong to these research themes will be conducted in later chapters when it becomes necessary.

1.6 Summary of main concepts

In this section, we will highlight different concepts for copulas and their joint tails; this is to emphasize that different copula families have different tail properties which may have a large effect on inferences such as joint tail probabilities (chance of simultaneously large losses from different risks) and conditional tail expectations. The bivariate case will be employed here to explain the basic ideas, and those ideas can be extended to the multivariate case.

1.6.1 Tail order and strength of dependence in the tails

One theme of the thesis is to study the strength of dependence in the tails. In statistical modeling, a basic idea is to find a summary quantity that can capture useful information of interest. Such a quantity should be simple and be able to account for as much as possible relevant and critical information. If we use a copula C to model the dependence between two continuous random variables X and Y, then the graph of the map $u \mapsto C(u, u), u \in [0, 1]$ describes the evolution of joint probabilities of X and Y. In order to study the tail behavior of the map and thus the tail dependence between X and Y, we hope to have a certain simple function g that can be used to approximate $C(u, u)$ as $u \to 0^+$; that is, $C(u, u) \sim g(u), u \to 0^+$. Such a function $g(u)$ should be bounded by u and go to 0 as $u \to 0^+$. Therefore, a suitable class of functions for g can be the class of regularly varying functions with the minimum regular variation index being 1, e.g., $g(u) = u^\kappa \ell(u)$ where $1 \le \kappa$ and ℓ is a slowly varying function such as $\ell(u) = (-\log u)^{-1/3}$. A formal introduction to regular variation is in Section 2.3.

For the function g, the term u^κ dominates the behavior of $g(u)$ as $u \to 0^+$. So we choose κ as such a summary quantity that can be used to capture the information of tail association patterns, and κ is referred to as *tail order*. A smaller κ means higher joint probability in the lower corner as $u \to 0$. Although obviously a single quantity such as the tail order κ can not capture all the information on the tail, it provides relevant and critical information of interest, especially when we compare tail orders of copulas within the same copula family, or when the information of interest is for a sufficiently small u. Therefore, such a simple quantity may also be useful in statistical modeling.

When $\kappa = 1$, that is, $C(u, u) \sim u\ell(u), u \to 0^+$, and $\lim_{u \to 0^+} \ell(u) = \lambda$, if $0 < \lambda \leq 1$, it is said that the copula C has lower tail dependence and the limit λ (if exists) is referred to the *tail dependence parameter* (some papers use "tail dependence coefficient" for the same meaning). The concept of tail dependence has been studied intensively in the literature, e.g., Joe (1997) and Schmidt (2002). As a special case of the usual tail dependence, the so-called *tail comonotonicity* corresponds to that the tail order $\kappa = 1$ and the tail order parameter $\lambda = 1$ as well. Tail comonotonicity is the strongest tail dependence structure. In situations where no sufficient data in the upper tail of certain joint distributions is available and the statistical inference of interest is aimed at the region beyond a very high threshold, which is often the situation that one meets in quantitative risk management, tail comonotonicity may provide a conservative dependence structure.

When $1 < \kappa < 2$, we refer to this case as *intermediate tail dependence*, and the limit (if exists) of the slowly varying function ℓ is referred to as the *tail order parameter*.

When $\kappa = 2$ and the limit (if exists) of the slowly varying function ℓ is finite and nonzero, then we use the notion of *tail quadrant independence* (Reiss, 1989). For d-dimensional cases with $\kappa = d > 2$, we employ the term of *tail orthant independence*.

Note that, in the literature of the extreme value theory, terms of "asymptotic independence" or "tail independence" have often been used to represent the cases where $\kappa > 1$. However, in this thesis, in order to discriminate the two cases of $1 < \kappa < 2$ and $\kappa = 2$, we use the notation of "intermediate tail dependence" and "tail quadrant independence", respectively. Intermediate tail dependence for the multivariate case corresponds to $1 < \kappa < d$.

Upper tails can be dealt with similarly: the upper tail behavior of copula $C(u, v)$ is simply the lower tail behavior of its survival copula $\widehat{C}(u, v) := u + v - 1 + C(1-u, 1-v)$. If $\widehat{C}(u, u) \sim u^{\kappa}\ell(u), u \to 0^+$, then the κ here is referred to as the upper tail order of the copula C.

For any of the above cases, the limit of the slowly varying function ℓ can be referred to as the tail order parameter no matter it is the case of tail dependence, intermediate tail dependence or tail quadrant independence. For comparing the strength of dependence in the tails, we first compare tail orders that dominate the tail dependence pattern. If tail orders are the same, then tail order parameters can be used.

Examples of lower and upper tail orders for commonly used Archimedean copulas are listed in Table 1.1, and the contour plots of these copulas are given in Figures 1.6, 1.7, 1.8, 1.9, 1.10 and 1.11. Note that, the first three examples of Gaussian, Student t and Frank copulas are all *reflection symmetric*; that is, the pattern of the upper tail is the same as the pattern of the lower tail. But *tail asymmetry* appears in the contour plots of the other three examples. When tail order $\kappa > 1$, such as the tails of Gaussian copula and the lower tail of the Gumbel, the contour plots look more round than the contour plots for tails with $\kappa = 1$. We refer to Balkema and Nolde (2010) for sufficient conditions of $\kappa > 1$ from a geometric perspective.

Table 1.1: Tail order of various bivariate copulas

Copula	cdf	κ of $C(u,v)$		κ of $c(u,v)$	
		Lower	Upper	Lower	Upper
Gaussian	$\Phi_\Sigma\left(\Phi^{-1}(u),\Phi^{-1}(v)\right)$; Σ : correlation matrix, Φ : cdf of normal	$\frac{2}{1+\varrho}$	$\frac{2}{1+\varrho}$	$\frac{-2\varrho}{1+\varrho}$	$\frac{-2\varrho}{1+\varrho}$
Student t	$T_{\nu,\Sigma}\left(T_\nu^{-1}(u),T_\nu^{-1}(v)\right)$; $0<\nu,\Sigma$: correlation matrix, T_ν : cdf of Student t	1	1	-1	-1
Frank	$-\log\left([1-e^{-\delta}-(1-e^{-\delta u})(1-e^{-\delta v})]/(1-e^{-\delta})\right)/\delta; 0\le\delta$	2	2	0	0
Gumbel	$\exp\left\{-\left((-\log u)^\delta+(-\log v)^\delta\right)^{1/\delta}\right\}; 1\le\delta$	$2^{1/\delta}$	1	$2^{1/\delta}$-2	-1
MTCJ	$\left(u^{-\delta}+v^{-\delta}-1\right)^{-1/\delta}; 0\le\delta$	1	2	-1	0
BB2	$C=\left[1+\delta^{-1}\log\left(e^{\delta(u^{-\theta}-1)}+e^{\delta(v^{-\theta}-1)}-1\right)\right]^{-1/\theta}; 0<\theta, 0<\delta$	1	2	-1	0

When the tail order for the copula density function is negative such as Gaussian copula (Table 1.1), the density at the boundaries $(0,0)$ and $(1,1)$ of the copula is infinity.

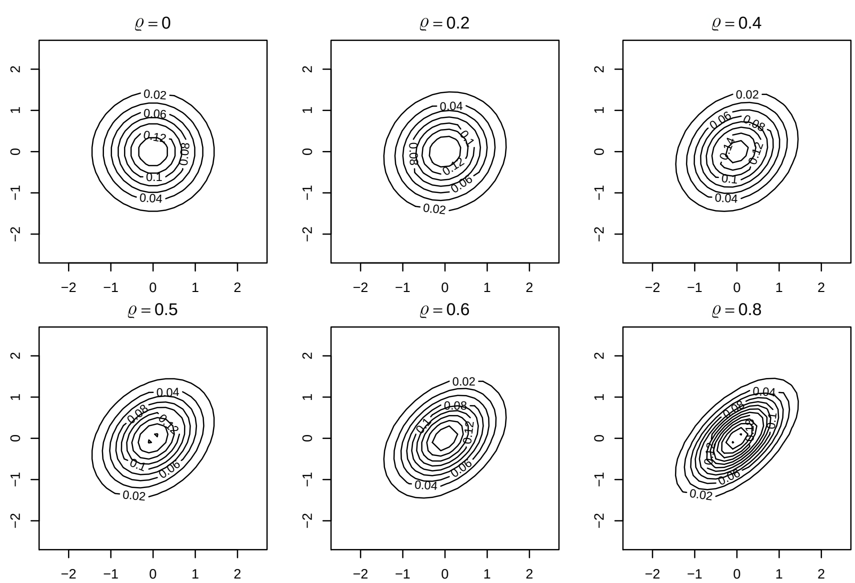
The quantities of tail order and tail order parameters only capture the tail behavior along the diagonal of the copula. We may use a function b to obtain more information about limiting behavior of copulas, namely,

$$\lim_{u\to 0^+} C(uw_1, uw_2)/u^\kappa \ell(u) =: b(w_1, w_2), \quad w_1, w_2 > 0.$$

The b function is referred to as the *tail order function* of the copula C.

For detailed study on tail order and tail order functions, we refer the reader to Chapter 3.

Figure 1.6: Contour plots: Gaussian copula + standard normal margins



1.6.2 Tail order and conditional copula

Now let us look at how tail order will affect conditional copulas of the form $C_{1|2}(u|v)$ as v goes to 0 or 1. We use $C_{1|2}(u|0)$ and $C_{1|2}(u|1)$ to represent these

Figure 1.7: Contour plots: Student t copula + standard normal margins

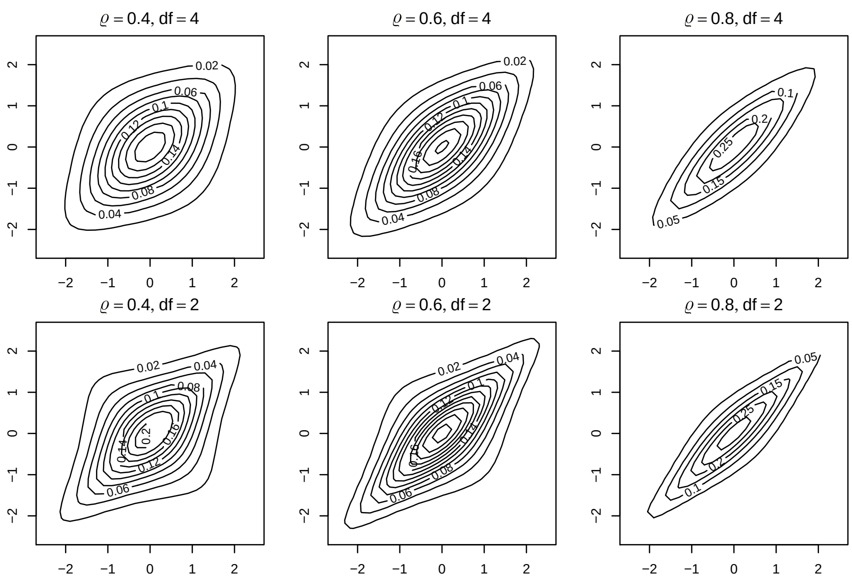


Figure 1.8: Contour plots: Frank copula + standard normal margins

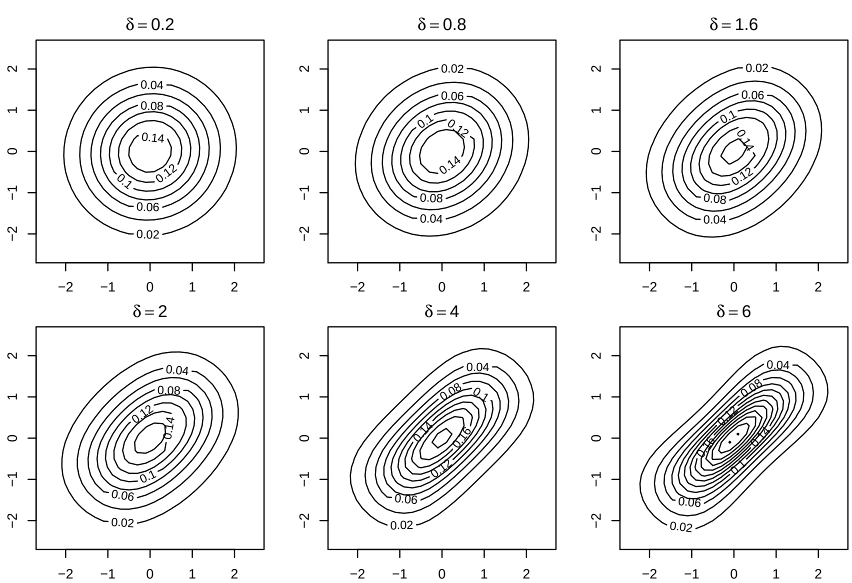


Figure 1.9: Contour plots: Gumbel copula + standard normal margins

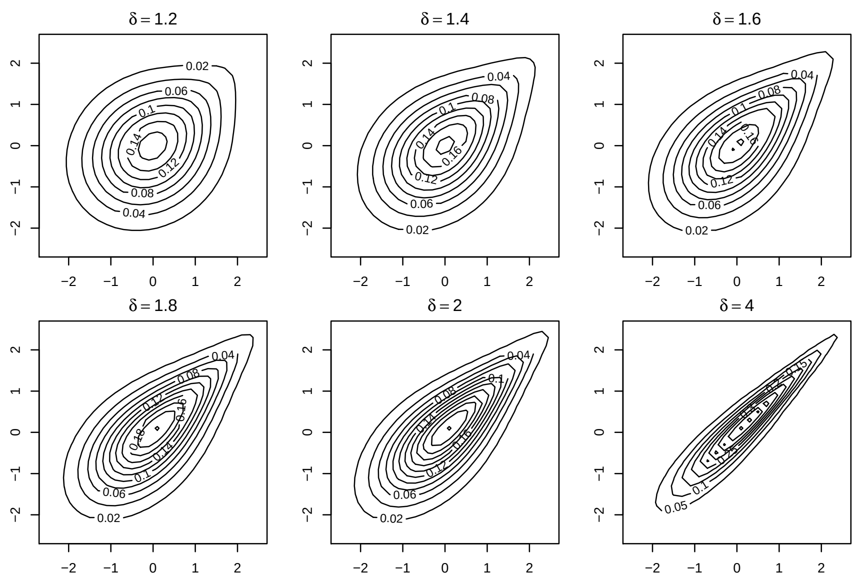


Figure 1.10: Contour plots: MTCJ copula + standard normal margins

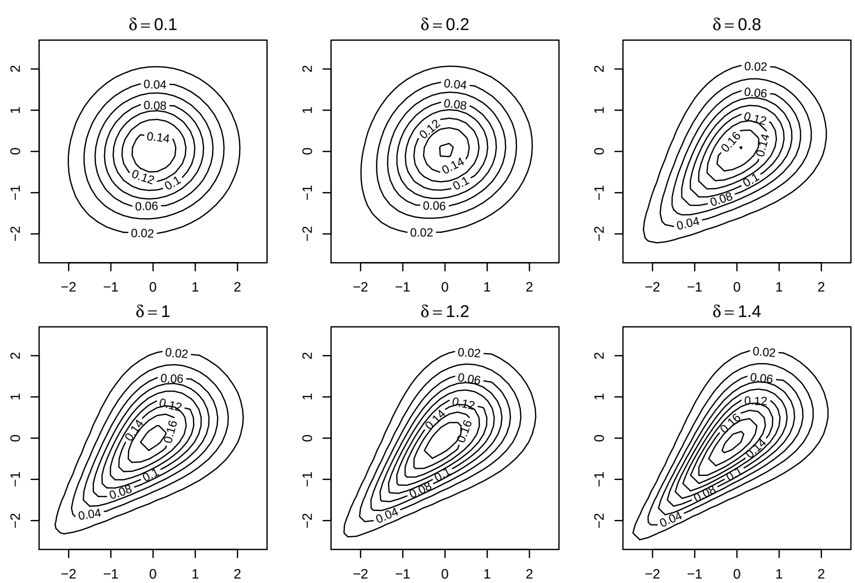
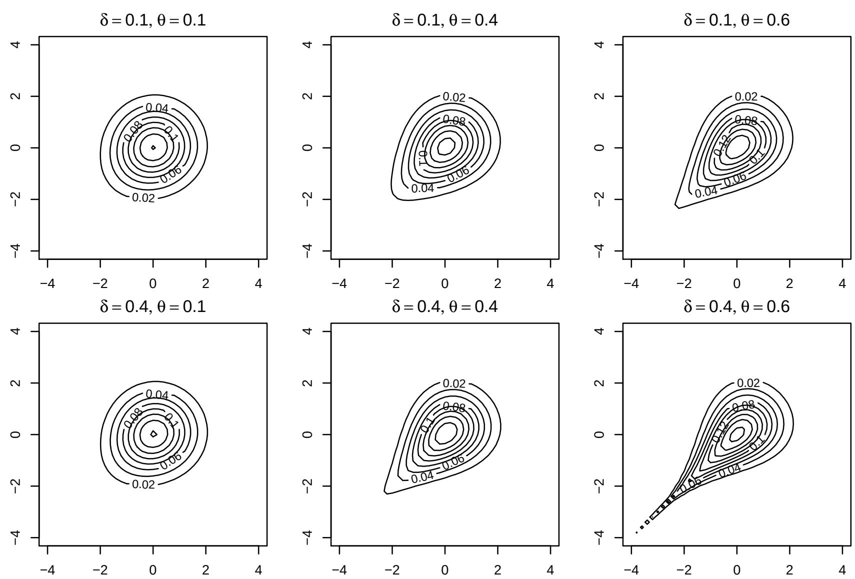


Figure 1.11: Contour plots: BB2 copula + standard normal margins



two limits, respectively. Explicit formulas for conditional Gaussian, Student t, Frank, Gumbel, MTCJ and BB2 copulas are presented in what follows. The limits for those conditional copulas are summarized in Table 1.2.

For those examples, if $1 \le \kappa_L < 2$, then $C_{1|2}(u|0) =: \varsigma_L$ with $0 < \varsigma_L \le 1$. The limit $\varsigma_L = 1$ for such a κ_L, except for the conditional Student t copula of which ς_L is strictly less than 1. The case $\varsigma_L = 1$ means that the limit of the conditional copula degenerates and becomes a single point at 0 as the conditioning random variable goes to 0. But for the Student t copula that has $0 < \varsigma_L < 1$, the limit of the conditional copula degenerates to be two points at $u = 0$ and $u = 1$, respectively, as the conditioning random variable $V = v \to 0^+$. Looking at Figure 1.7, one may find that the bivariate Student t copula possesses tail dependence in all of the four directions, which explains why the limit of the conditional copula degenerates as two points. In parallel, for upper tails, if $1 \le \kappa_U < 2$, then the limit $C_{1|2}(u|1)$ degenerates to a single point at $u = 1$ except that for the Student t copula, it degenerates into two points at $u = 0$ and $u = 1$, respectively.

If $\kappa_L = 2$, that is, if the copula is lower tail quadrant independent, then the limit of the conditional copula as $V = v \to 0^+$ is still a non-degenerate cdf. Similarly, if $\kappa_U = 2$, then the limit is a non-degenerate cdf as well.

In the sense of such a limit of conditional copulas, that $1 \le \kappa < 2$ indicates a relatively higher degree of positive association in the tail. This is also a reason why we want to use different notions to discriminate intermediate tail dependence ($1 < \kappa < 2$) and tail quadrant independence ($\kappa = 2$), although in the literature of the extreme value theory, these two cases are both referred to as asymptotic independence or tail independence.

Gaussian copula

Let (X_1, X_2) be standard bivariate Normal with correlation ϱ. Since $X_1|X_2 = x \sim Normal(\varrho x, 1-\varrho^2)$, the conditional Gaussian copula is

$$C_{1|2}(u|v) = \Phi\left(\frac{\Phi^{-1}(u) - \varrho\Phi^{-1}(v)}{\sqrt{1-\varrho^2}}\right).$$

Thus $C(u|0) = 1$ for $0 < u \le 1$, and $C(u|1) = 0$ for $0 \le u < 1$.

Student t copula

Let (X_1, X_2) be standard bivariate Student t distributed with ν the degree of freedom and ϱ the correlation coefficient in the correlation matrix. Since by (5.30) of McNeil et al. (2005),

$$\sqrt{\frac{\nu+1}{\nu+x^2}}\frac{X_2-\varrho x}{\sqrt{1-\varrho^2}}\,\bigg|\, X_1 = x \sim t_{\nu+1}(0,1),$$

the conditional Student t copula is

$$C_{1|2}(u|v) = T_{\nu+1}\left(\sqrt{\frac{\nu+1}{\nu+(T_\nu^{-1}(v))^2}} \times \frac{T_\nu^{-1}(u) - \varrho T_\nu^{-1}(v)}{\sqrt{1-\varrho^2}}\right).$$

Thus $C_{1|2}(u|0) = T_{\nu+1}\left(\varrho\sqrt{\nu+1}/\sqrt{1-\varrho^2}\right)$ for $0 < u < 1$, and $C(u|1) = T_{\nu+1}\left(-\varrho\sqrt{\nu+1}/\sqrt{1-\varrho^2}\right)$ for $0 < u < 1$.

Frank copula

The conditional Frank copula is straightforward to get, and it is

$$C_{1|2}(u|v) = \frac{(1-e^{-\delta u})e^{-\delta v}}{1-e^{-\delta}-(1-e^{-\delta v})(1-e^{-\delta u})}.$$

Therefore, $C_{1|2}(u|0) = (1-e^{-\delta u})/(1-e^{-\delta})$ for $0 \le u \le 1$, and $C_{1|2}(u|1) = ((1-e^{-\delta u})e^{-\delta})/[(1-e^{-\delta})e^{-\delta u}]$ for $0 \le u \le 1$.

Gumbel copula

The conditional Gumbel copula is

$$C_{1|2}(u|v) = \frac{1}{v}\exp\left(-\left(\tilde{u}^{\delta}+\tilde{v}^{\delta}\right)^{1/\delta}\right)\left[1+\left(\frac{\tilde{u}}{\tilde{v}}\right)^{\delta}\right]^{-1+1/\delta},$$

where $\tilde{v} := -\log v$ and $\tilde{u} := -\log u$. Thus, $C_{1|2}(u|0) = 1$ for $0 < u \le 1$, and $C_{1|2}(u|1) = 0$ for $0 \le u < 1$.

MTCJ copula

The conditional Mardia-Takahasi-Cook-Johnson copula (MTCJ) copula is

$$C_{1|2}(u|v) = (v^{-\delta}+u^{-\delta}-1)^{-1/\delta-1}v^{-\delta-1},$$

and $C_{1|2}(u|0) = 1$ for $0 < u \le 1$, and $C_{1|2}(u|1) = u^{1+\delta}$ for $0 \le u \le 1$.

BB2 copula

The conditional BB2 copula is

$$\begin{aligned} &C_{1|2}(u|v) \\ &= \left[1+\delta^{-1}\log\left(e^{\delta(v^{-\theta}-1)}+e^{\delta(u^{-\theta}-1)}-1\right)\right]^{-1/\theta-1}\frac{e^{\delta(v^{-\theta}-1)}v^{-\theta-1}}{e^{\delta(v^{-\theta}-1)}+e^{\delta(u^{-\theta}-1)}-1}. \end{aligned}$$

Thus, $C_{1|2}(u|0) = 1$ for $0 < u \le 1$, and $C_{1|2}(u|1) = u^{1+\theta}e^{\delta(1-u^{-\theta})}$ for $0 \le u \le 1$.

Table 1.2: Tail order and conditional copula

Copula	κ_L	$C_{1\|2}(u\|0)$
Gaussian	$2/(1+\varrho)$	1 for $0 < u \leq 1$
Student t	1	$T_{\nu+1}\left(\varrho\sqrt{\nu+1}/\sqrt{1-\varrho^2}\right)$ for $0 < u < 1$
Frank	2	$(1-e^{-\delta u})/(1-e^{-\delta})$ for $0 \leq u \leq 1$
Gumbel	$2^{1/\delta}$	1 for $0 < u \leq 1$
MTCJ	1	1 for $0 < u \leq 1$
BB2	1	1 for $0 < u \leq 1$

Copula	κ_U	$C_{1\|2}(u\|1)$
Gaussian	$2/(1+\varrho)$	0 for $0 \leq u < 1$
Student t	1	$T_{\nu+1}\left(-\varrho\sqrt{\nu+1}/\sqrt{1-\varrho^2}\right)$ for $0 < u < 1$
Frank	2	$((1-e^{-\delta u})e^{-\delta})/[(1-e^{-\delta})e^{-\delta u}]$ for $0 \leq u \leq 1$
Gumbel	1	0 for $0 \leq u < 1$
MTCJ	2	$u^{1+\delta}$ for $0 \leq u \leq 1$
BB2	2	$u^{1+\theta}e^{\delta(1-u^{-\theta})}$ for $0 \leq u \leq 1$

1.6.3 Tail order and Conditional Tail Expectation

Two fundamental conditional specifications that are often useful in more specific modelings are $\mathbb{E}[X_1|X_2 > t]$ and $\mathbb{E}[X_1|X_2 = t]$. In actuarial science, these two quantities are often referred to as certain forms of *Conditional Tail Expectation*, and they are useful as risk measures to assessing magnitude of losses. Moreover, the study of the tail behavior as $t \to \infty$ of these two conditional specifications may also be meaningful when one needs to develop diagnostic plots for discriminating types of tail dependence, or formulate certain regression models that account for tail dependence patterns.

The case: $\mathbb{E}[X_1|X_2 > t]$

Let X_1, X_2 be non-negative random variables with copula C and the same Pareto distribution $F(x) = 1-(1+x)^{-\theta}, \theta > 1$. Now we consider the effect of tail order on the following conditional specification. Letting $v := F(t)$,

$$\begin{aligned}&\mathbb{E}[X_1|X_2 > t]\\&= (1-v)^{-1}\left\{1/(\theta-1) + \int_0^1 \{[-v + C(u,v)] \cdot [f(F^{-1}(u))]^{-1}\}du\right\}. \quad (1.2)\end{aligned}$$

Similar to that we use regularly varying functions to describe the tail behavior of copula functions, we may assume that there exists a regularly varying function $g(t) := t^{\beta}\ell(t)$ such that

$$\mathbb{E}[X_1|X_2 > t] = g(t) = t^{\beta}\ell(t), \quad t \to \infty.$$

Then we can briefly discuss the following three cases:

1. For the usual tail dependence case, $\beta = 1$ with some regularity conditions. The derivation of this form can be conducted through the theory of MRV.

2. For the intermediate tail dependence case, with certain regularity conditions, $0 < \beta < 1$ can be a function of the shape parameter α for the Pareto margins.

3. For the tail quadrant independence case, with some regularity conditions, $\beta = 0$. A special case is when X_1 and X_2 are independent, then clearly $\mathbb{E}[X_1|X_2 > t] = \mathbb{E}[X_1]$ is a constant and does not rely on t.

In Figure 1.12, a comparison between Gumbel, s.Gumbel and MTCJ copulas is illustrated. Based on it, we find that when the upper tail order is 1 (Gumbel), the CTE plot seems to be linear in t on the whole support no matter what θ is. When the upper tail order is 2 (MTCJ), the CTE becomes very flat. The CTE plots for the intermediate upper tail dependence with the upper tail order $1 < \kappa < 2$ (s.Gumbel) are located between the above two cases.

The case: $\mathbb{E}[X_1|X_2 = t]$

Let X_1, X_2 be non-negative random variables with copula C and the same Pareto distribution $F(x) = 1-(1+x)^{-\theta}, \theta > 1$. Then we compare the effect of tail order on the following conditional specification. Letting $v := F(t)$,

$$\mathbb{E}[X_1|X_2 = t] = \int_0^1 \left\{1 - C_{1|2}(u|v)\right\} \cdot [f(F^{-1}(u))]^{-1}du.$$

Figure 1.12: Tail order and CTE plots of the form $\mathbb{E}[X_1|X_2 > t]$. The range of t in each plot was chosen to cover the support of t between the 1% and 95% quantiles. Gumbel($\delta = 2$), s.Gumbel($\delta = 2$) and MTCJ($\delta = 1$).

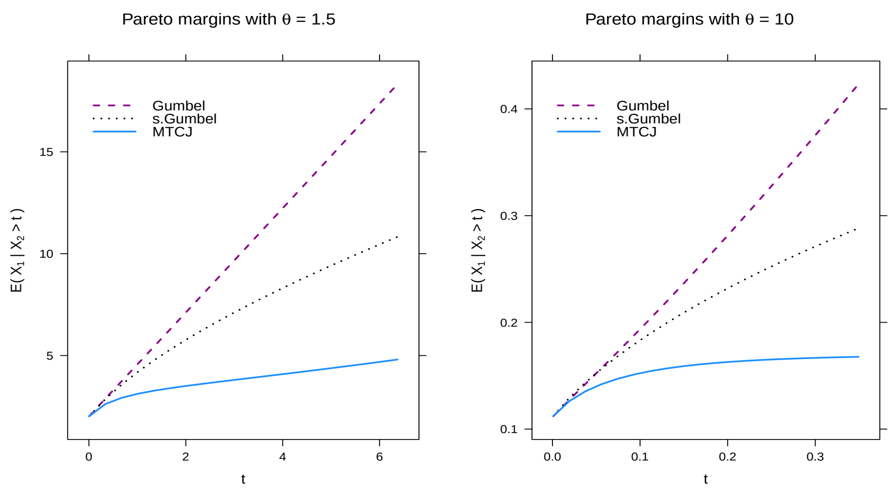
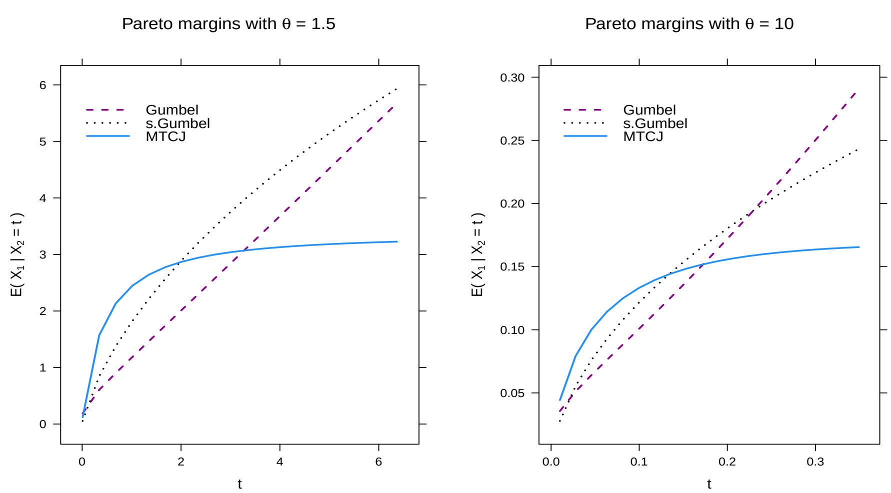


Figure 1.13: Tail order and CTE plots of the form $\mathbb{E}[X_1|X_2 = t]$. The range of t in each plot was chosen to cover the support of t between the 1% and 95% quantiles. Gumbel($\delta = 2$), s.Gumbel($\delta = 2$) and MTCJ($\delta = 1$).



From Figure 1.13, we can also observe the similar pattern of linearity for the Gumbel copula, and that the line for MTCJ copula becomes flat.

Those are just illustrations of the effect of tail order on Conditional Tail Expectation (CTE)s. In Chapter 5, more detailed study will be conducted for comparisons of CTEs under different dependence in the tails; in Chapter 7, more detailed study on such CTE plots are presented to illuminate future research on this topic.

1.6.4 Tail order and tail order functions

In this subsection, we briefly introduce the concepts of tail order and tail order functions for the following three general copula families: extreme value copula, Archimedean copula and elliptical copula. More detailed study is in Chapter 3.

Extreme value copula

If a copula C satisfies $C(u_1^t, \dots, u_d^t) = C^t(u_1, \dots, u_d)$ for any $(u_1, \dots, u_d) \in [0, 1]^d$ and $t > 0$, then we refer to C as an extreme value copula.

For any extreme value copula C, there exists a function $A : [0, \infty)^d \to [0, \infty)$ such that $C(u_1, \dots, u_d) = \exp\{-A(-\log u_1, \dots, -\log u_d)\}$, where A is convex, homogeneous of order 1 and satisfies $\max(x_1, \dots, x_d) \le A(x_1, \dots, x_d) \le x_1 + \dots + x_d$.

For the bivariate case, $C(u, u) = \exp\{A(1, 1) \log u\} = u^{A(1,1)}$, and the lower tail order is $\kappa_L = A(1, 1)$. Also, $\overline{C}(1 - u, 1 - u) = 2u - 1 + (1 - u)^{A(1,1)-1} \sim [2 - A(1, 1)]u, u \to 0^+$. Therefore, the upper order is $\kappa_U = 1$.

Tail order functions for extreme value copulas are relatively easy to get. For the lower tail, $b(w_1, w_2) = \lim_{u \to 0^+} C(uw_1, uw_2)/u^{A(1,1)} = w_1^{A_1(1,1)} w_2^{A_2(1,1)}$, where $A_i = \partial A(x_1, x_2)/\partial x_i, i = 1, 2$. For the upper tail, $b^*(w_1, w_2) = w_1 + w_2 - A(w_1, w_2)$.

Archimedean copula

The Archimedean copula studied in this thesis is constructed by a Laplace Transform (LT) ψ of a positive random variable; that is

$$C(u_1, \dots, u_d) = \psi\left(\psi^{-1}(u_1) + \dots + \psi^{-1}(u_d)\right).$$

This form covers most of the commonly used Archimedean copulas, and it has the following representation

$$C(u_1, \ldots, u_d) = \int_0^\infty \prod_{j=1}^{d} G_j^\eta(u_j)\, dF_H(\eta),$$

where F_H is the cdf of the resilience random variable H, $G_j(u) = \exp\{-\psi^{-1}(u)\}$ $(0 \le u \le 1)$ for all j, and $\psi(s) = \psi_H(s) = \int_0^\infty e^{-s\eta} dF_H(\eta)$ with $\lim_{s\to\infty} \psi(s) = 0$.

When η becomes larger, the density of the joint distribution $\prod_{j=1}^{d} G_j^\eta(u_j)$ will be pushed to the right end of the support; as η goes to a smaller value, the density of the joint distribution $\prod_{j=1}^{d} G_j^\eta(u_j)$ will be pushed towards the left end of the support. Therefore, the behavior of H at ∞ will affect the upper tail of the copula, and the behavior of H at 0 will affect the lower tail of the copula.

For a positive random variable Y with LT ψ, if the *maximal non-negative moment* $M_Y := \sup\{m \ge 0 : \mathbb{E}(Y^m) < \infty\}$ is non-integer such that $1 < M_Y < d$, then the upper tail order of C_ψ is $\kappa_U = M_Y$ under some mild regularity conditions; if $0 \le M_Y < 1$, then the upper tail order of C_ψ is $\kappa_U = 1$. In other words, the behavior of $\psi(u)$ as $u \to 0^+$ influences the upper tail behavior of the copula C_ψ. The tail behavior of $\psi(u)$ as $u \to \infty$, however, will affect the lower tail behavior of the copula C_ψ.

When tail order of an Archimedean copula has been obtained, we can get the corresponding tail order functions usually by applying the l'Hôpital's rule.

Elliptical copula

An elliptical random vector $\mathbf{X}$ has the stochastic representation $\mathbf{X} \stackrel{d}{=} RA\mathbf{U}$, where A is a deterministic matrix, and $R \ge 0$ is independent of $\mathbf{U}$ and referred to as the radial random variable, and $\mathbf{U}$ is uniformly distributed on the surface of a unit ball. The tail behavior of R will affect the tail behavior of corresponding elliptical copula. If R is in the Maximum Domain of Attraction (MDA) of Fréchet, then the copula has tail dependence, i.e., $\kappa = 1$; if R is in the MDA of Gumbel, then the tail order $\kappa > 1$ and the value of κ relies on both the tail behavior of R and the associated correlation coefficient ϱ.

Tail order functions of elliptical copulas for the first case ($R \in \text{MDA}(\text{Fréchet})$) can be calculated directly by considering the above stochastic representation, but a general form for the second case remains to be an open question. We refer to Chapter 3 for more detailed arguments about tail order and tail order function for elliptical copulas.

Inference for tail probability

Tail order functions can be useful for statistical inference on multivariate extremal events. Let b^* be the upper tail order function for a copula C, and $\widehat{C}$ be the corresponding survival copula, then we have $\widehat{C}(u\mathbf{w}) \sim u^\kappa \ell(u) b^*(\mathbf{w}), u \to 0^+$. Since the term u^κ is tractable and dominates the tail behavior, the tail order function $b^*(\mathbf{w})$ may approximate the tail behavior of the copula C. Tail order functions are limiting functions and it would be more robust than using the copula itself for statistical inference on tail events.

1.6.5 Concordance ordering in the tail

Let $\mathbf{X}$ and $\mathbf{Y}$ be d-dimensional random vectors with distribution function $F_\mathbf{X}$ and $F_\mathbf{Y}$ such that $X_i \stackrel{d}{=} Y_i, i = 1, \ldots, d$. Then $\mathbf{X}$ is less concordant than $\mathbf{Y}$ if $F_\mathbf{X}(\mathbf{z}) \le F_\mathbf{Y}(\mathbf{z})$ and $\overline{F}_\mathbf{X}(\mathbf{z}) \le \overline{F}_\mathbf{Y}(\mathbf{z})$ for any $\mathbf{z}$ in the support of $F_\mathbf{X}$ and $F_\mathbf{Y}$. The concordance order is an important concept in the theory of stochastic orders (Joe, 1990). That means small values of $\mathbf{Y}$ are more likely to occur together than those of $\mathbf{X}$, and large values of $\mathbf{Y}$ are also more likely to occur together than those of $\mathbf{X}$. If $\mathbf{X}$ is less concordant than $\mathbf{Y}$, then the upper and lower tail orders of the copula of $\mathbf{X}$ are larger than those of the copula of $\mathbf{Y}$, respectively.

For comparisons between tails, especially between tails in a single direction, the concordance order is a relatively strong condition. Here we actually only need the conditions such as $F_\mathbf{X}(\mathbf{z}) \le F_\mathbf{Y}(\mathbf{z})$ holds ultimately as $\mathbf{z}$ is sufficiently small for comparing the lower tail, and $\overline{F}_\mathbf{X}(\mathbf{z}) \le \overline{F}_\mathbf{Y}(\mathbf{z})$ holds ultimately as $\mathbf{z}$ is sufficiently large for comparing the upper tail.

1.6.6 Sub-extremes and second order conditions

In real applications of tail modeling, one often has interests not only in the limiting properties but also in the tail part beyond a high threshold. If the limit is referred to as an extreme, then the tail part beyond the high threshold will be referred to as sub-extremes. We use limiting properties – such as tail index for margins and tail order for dependence structures – as summary quantities for the tail, in hope that the limiting quantity may provide a reasonable approximation for the extreme and sub-extremal levels as well.

Apparently, the speed of decay for the convergence involved to get the limiting quantity may influence the goodness of such an approximation. We will use second order conditions as a generic term to represent those conditions that may affect rates of certain convergence. For instance, in Chapter 6, we study second order conditions in the form of the so-called second order regular variation. For example, if $g(x) := x^{-2} + x^{-3}$, then for any given $t > 0$, $g(xt)/g(x) \to t^{-2}$ as $x \to \infty$. Since $g(xt)/g(x) - t^{-2} = (x+1)^{-1}(t^{-3} - t^{-2})$, the term $(x+1)^{-1} \sim x^{-1}$ dominates the speed of convergence of $g(xt)/g(x) \to t^{-2}$. We refer to -1 here as a second order parameter. When the second order parameter is more negative, the limiting quantity t^{-2} becomes a better approximation to $g(xt)/g(x)$ for large x.

1.7 Outline

This thesis is organized as following: In Chapter 2, motivations, definitions and related known results for some fundamental concepts and technical tools will be presented. Chapter 3 includes the study of using tail order to quantify the degree of positive tail association. Some relevant studies on multivariate copulas such as the Archimedean copula have also been done in this chapter. In Chapter 4, tail comonotonicity, the strongest tail dependence case will be studied for its fundamental properties, constructions and influence on risk measures. Chapter 5 will be used to report the findings of how tail behavior of copulas affects commonly-used risk measures, and in particular, of the conservativity of tail comonotonicity. In Chapter 6, we will study how 2RV affects the performance of risk measures. At last, in Chapter 7, we will conclude the thesis and propose some relevant topics for future research.

1.8 Highlights of selected new results

- A new concept called tail order has been proposed in Definition 3.1, and the corresponding tail order function is defined by Definition 3.2. A formula for deriving tail order of a bivariate elliptical copula where the radial random variable belongs to MDA(Gumbel) is given in Proposition 3.1. Some fundamental properties for these two concepts are given in Proposition 3.2 and 3.3.

- In order to study the tail behavior of the Archimedean copula, we first prove a result in Proposition 3.5 that relates the tail heaviness of a positive random variable to the Taylor expansion of its LT.

- For the Archimedean copula, we relate the tail heaviness of a positive random variable to the tail behavior of the Archimedean copula constructed from the LT of the random variable, and extend the results of Charpentier and Segers (2009) for the upper tail in Proposition 3.6. And the corresponding upper tail order function has been proved in Proposition 3.7 as well. For the lower tail, we study a more concrete form by assuming some mild conditions, and the relevant result is Proposition 3.8.

- A new notion of tail comonotonicity is proposed in Definition 4.1, and relevant results about their construction include Proposition 4.3 and Proposition 4.8.

- Asymptotic additivity of risk measures Value at Risk (VaR) and CTE under tail comonotonicity is reported in Propositions 4.13, 4.14, 4.15 and 4.16.

- Effect of tail behavior of copulas on $\mathbb{E}[X_1|X_2 > t]$ and $\mathbb{E}[X_1|X_2 = t]$ are mainly summarized in Propositions 5.1, 5.4 and 5.9, and their corollaries.

- For the study of 2RV, results for asymptotic analysis of CTE and VaR has been reported in Proposition 6.4 for the univariate case, and in Proposition 6.6 for the multivariate case.

Chapter 2

Preliminaries

2.1 Risk measures

Some statistical quantities about a random variable X are often referred to as risk measures. Among many of them, Value at Risk (VaR) and conditional tail expectation (CTE) are probably the most popular risk measures. Both of them have been adopted by regulations of insurers. For example, CTE has been required for calculating the relevant risks of segregated fund in Canada (OSFI, 2011).

Definition 2.1 *Let be given a random variable X the amount of risk, and $p \in (0,1)$ a probability level, then the corresponding VaR, denoted by $\mathrm{VaR}_p(X)$, is defined as*

$$\mathrm{VaR}_p(X) = F_X^{-1}(p) := \inf\{x \in \mathbb{R} : F_X(x) \geq p\}.$$

If $\mathbb{E}[X] < \infty$, then the corresponding CTE is defined as

$$\mathrm{CTE}_p(X) = \mathbb{E}[X|X > \mathrm{VaR}_p(X)]. \tag{2.1}$$

Note that the VaR defined above is a left-continuous generalized inverse of F_X, and CTE is a coherent risk measure for continuous random variables (Artzner et al., 1999), and it can better account for tail heaviness of a loss distribution than VaR does. We refer to Denuit et al. (2005) for some other relevant risk measures and their relationships.

Risk measures are used to help people make decisions through certain decision principles. For instance, one use risk measures to determine premiums of insurance products or calculate solvency capital charges. We refer to Goovaerts et al. (2010) for a comprehensive overview of risk measures and relevant decision principles.

The risk measures mentioned in Definition 2.1 are both quantile-based risk measures. In this thesis, the emphasis will be risk measures for high-risk scenarios, for which p is close to 1. We will conduct asymptotic comparisons of risk measures in the sense that, when p is close to 1 what will be the ordering for those risk measures. For stochastic comparisons of risks, Müller and Stoyan (2002) is popular reference. Recently, Mainik and Rüschendorf (2012) conduct asymptotic comparisons for risks of portfolios, where a concept of asymptotic portfolio loss order is studied.

A classic question in quantitative risk management is to study the so-called diversification benefit of a portfolio of risks.

Definition 2.2 Let $\mathbf{X} := (X_1, \ldots, X_d)^{\mathrm{T}}$ be a d-dimensional random vector. Then risk concentration $\varpi(\mathbf{X}|Q_p)$ of $\mathbf{X}$ with respect to a risk measure Q_p at level $p \in (0, 1)$ is defined as

$$\varpi(\mathbf{X}|Q_p) := \frac{Q_p(X_1 + \cdots + X_d)}{\sum_{i=1}^{d} Q_p(X_i)},$$

where $Q_p : \mathbb{R} \to \mathbb{R}_+$ is a risk measure associated with a probability level p. The corresponding diversification benefit is $1 - \varpi(\mathbf{X}|Q_p)$.

Here two relevant concepts are subadditivity and superadditivity of risk measures. If $\varpi(\mathbf{X}|Q_p) \leq 1$, then Q_p is said to be subadditive with respect to $\mathbf{X}$, and $\varpi(\mathbf{X}|Q_p) \geq 1$ superadditive. It is well known that CTE defined in (2.1) is subadditive with respect to continuous margins, but VaR is not necessarily subadditive for such a general case (e.g., Denuit et al., 2005). In this thesis, we will consider asymptotic additivity of CTE and VaR in the sense that $p \to 1^-$ in the risk concentration $\varpi(\mathbf{X}|Q_p)$.

2.2 Copula

The study of tail behavior of random vectors has received increasing attention, especially in the framework of quantitative risk management. Let $\mathbf{X} = (X_1, \ldots, X_d)^{\mathrm{T}}$ be a random vector with distribution function F and continuous univariate marginal distribution functions F_i, $i = 1, \ldots, d$. Due to Sklar's theorem Sklar (1959),

$$F(x_1, \ldots, x_d) = C(F_1(x_1), \ldots, F_d(x_d)), \tag{2.2}$$

in which the copula function $C : [0, 1]^d \to [0, 1]$ is uniquely determined by

$$C(u_1, \ldots, u_d) = F(F_1^{-1}(u_1), \ldots, F_d^{-1}(u_d)), \tag{2.3}$$

where F_i^{-1} is the inverse function of $F_i, i = 1, \ldots, d$. The corresponding survival function $\overline{C}$ is defined as $\overline{C}(u_1, \ldots, u_d) = 1 + \sum_{\emptyset \neq I \subset \{1,\ldots,d\}} (-1)^{|I|} C_I(u_i, i \in I)$, where C_I is the I-margin of the copula C with $|I|$ the cardinality of the set I.

There exist several related methodologies. The lower tail dependence parameter is defined as

$$\lambda_L = \lim_{u \to 0^+} u^{-1} C(u, \ldots, u),$$

and the upper tail dependence parameter is defined similarly with the survival function $\overline{C}$. As extensions, Juri and Wüthrich (2002, 2003) studied tail dependence from a distributional point of view. Klüppelberg et al. (2008) defined the so-called tail dependence function of $\mathbf{X}$ as

$$\lambda^{\mathbf{X}}(x_1, \ldots, x_d) = \lim_{t \to 0} t^{-1} \mathbb{P}[1 - F_1(X_1) \leq tx_1, \ldots, 1 - F_d(X_d) \leq tx_d],$$

and Nikoloulopoulos et al. (2009); Joe et al. (2010); Li and Sun (2009) further studied the properties of the tail dependence function and their applications for multivariate t copulas, vine copulas and heavy-tailed scale mixtures of multivariate distributions, respectively. We refer to the above papers for details and properties of tail dependence functions.

Multivariate Archimedean copulas are widely used in insurance and financial risk analyses as they can be used for sensitivity analyses to changes in dependence, tail asymmetry, and lower/upper tail behavior. They are not flexible enough as models for high-dimensional data but can be extended to more flexible copula families (Joe and Hu, 1996) based on mixtures of max-infinitely divisible distributions. In the literature, a d-dimensional Archimedean copula C is often (e.g., Genest and MacKay, 1986; Nelsen, 2006) defined as $C(u_1, \ldots, u_d) = \phi^{-1}(\phi(u_1) + \cdots + \phi(u_d))$, $(u_1, \ldots, u_d) \in [0, 1]^d$. McNeil and Nešlehová (2009) showed that d-monotonicity is a sufficient and necessary condition on the Archimedean generator ϕ^{-1} so that the above form is a copula; a real-valued function g defined on (a, b) with $a, b \in \overline{\mathbb{R}}$ is said to be d-monotone ($d \geq 2$) if it is differentiable up to the order $d-2$, $(-1)^k g^{(k)}(x) \geq 0$ for any $x \in (a, b)$ and $k = 0, \ldots, d-2$, and $(-1)^{d-2} g^{(d-2)}$ is non-increasing and convex in (a, b).

To get a better understanding of tail dependence (intermediate or usual tail dependence), we use the mixture of power or LT representation in Marshall and Olkin (1988) and Joe (1997); for mixing distribution functions [resp. survival functions], the power is called *resilience* [resp. *frailty*] in Marshall and Olkin (2007). Let

$$C_\psi(u_1, \ldots, u_d) = \psi(\psi^{-1}(u_1) + \cdots + \psi^{-1}(u_d)), \quad (u_1, \ldots, u_d) \in [0, 1]^d, \tag{2.4}$$

where ψ is the LT of a positive random variable. Note that as $d \geq 3$ increases, Archimedean copulas extend less into the region of negative dependence (Joe, 1997, Sections 4.4 and 5.4) and hence the restriction to LTs does not lose much generality for defining an Archimedean copula. Since a LT is completely monotone, it can be used to construct copulas of any dimension.

In this thesis, we assume that the function ψ is a LT of a positive random variable; that is, $\psi(t) := \int_0^\infty e^{-xt} F_X(dx)$, where F_X is the distribution function of a positive random variable X. Such choices of ψ cover most of the commonly used Archimedean copulas.

The tail behavior of a LT function plays a critical role in analyzing the dependence in the tails of the Archimedean copula constructed by the LT function. Here we list basic properties of a LT $\psi : [0, \infty) \to [1, 0)$ as follows: $\psi(0) = 1, \lim_{s\to\infty} \psi(s) = 0$, ψ is continuous and strictly decreasing, and $\psi(s)$ is completely monotone; that is, the derivatives $\psi^{(n)}$ exist for $n = 0, 1, 2, \ldots$ and

$(-1)^n\psi^{(n)}(s) \geq 0$ for all $s \geq 0$.

If $C(u_1, \dots, u_d)$ is a copula, then the corresponding survival copula is defined as

$$\widehat{C}(u_1, \dots, u_d) = 1 + \sum_{I \subset \{1, \dots, d\}} (-1)^{|I|} C_I(1 - u_i, i \in I),$$

where C_I is the I-margin of the copula C with $|I|$ the cardinality of the set I; in particular, for bivariate case, $\widehat{C}(u, v) = u + v - 1 + C(1 - u, 1 - v)$. If C_ψ is an Archimedean copula constructed by LT ψ, then the corresponding survival copula is referred to as a survival Archimedean copula constructed by LT ψ. The survival copula itself is a copula, and it should be distinguished from the survival function of a copula (denoted as $\overline{C}$); for bivariate case, the survival function of a copula C is $\overline{C}(u, v) = 1 - u - v + C(u, v)$.

Definition 2.3 A copula C is said to be reflection symmetric if the copula is the same as its survival copula $\widehat{C}$; that is, $C(\mathbf{u}) \equiv \widehat{C}(\mathbf{u})$ for any $\mathbf{u} \in [0, 1]^d$.

Another copula family that will be studied is extreme value copula. If a copula C satisfies

$$C(u_1^t, \dots, u_d^t) = C^t(u_1, \dots, u_d)$$

for any $(u_1, \dots, u_d) \in [0, 1]^d$ and $t > 0$, then we refer to C as an extreme value copula.

When we compare certain properties for different copulas, we usually assume that those copulas share some common summary quantities such as the following Blomqvist's β (Blomqvist, 1950) for the bivariate case.

Definition 2.4 The Blomqvist's β for a bivariate copula C is defined as $\beta := 4C(1/2, 1/2) - 1$.

So $-1 \leq \beta \leq 1$; $\beta = 0$ holds for the independent copula and a larger positive β suggests a higher positive dependence. Moreover, $\beta(C) = \beta(\widehat{C})$ for the bivariate case.

Standard references for copula theory are Joe (1997) and Nelsen (2006). For applications of copulas in actuarial science and quantitative risk management, we refer to Denuit et al. (2005), McNeil et al. (2005) and Genest et al. (2009a).

2.3 Regular variation

For asymptotic analysis of tail behavior of random variables or equivalently their distribution functions, the theory of regular variation provides a powerful platform. Here we only give some most fundamental concepts. More relevant concepts and results about regular variation will be used throughout the thesis, and will be introduced when they are needed. Standard references on regular variation are Bingham et al. (1987), Resnick (1987), and Geluk and de Haan (1987); Embrechts et al. (1997) and Resnick (2007) are more relevant for applications in actuarial science, quantitative finance and risk management.

Definition 2.5 A measurable function $g : \mathbb{R}_+ \to \mathbb{R}_+$ is *regularly varying at* ∞ *with index* $\alpha \neq 0$ (written $g \in \mathrm{RV}_\alpha$) if for any $t > 0$,

$$\lim_{x\to\infty} \frac{g(xt)}{g(x)} = t^\alpha. \tag{2.5}$$

If equation (2.5) holds with $\alpha = 0$ for any $t > 0$, then g is said to be slowly varying at ∞ and written as $g \in \mathrm{RV}_0$.

For the lower limit at 0^+, if for any $t > 0$, $\lim_{x\to 0^+} g(xt)/g(x) = t^\alpha$, then g is regularly varying at 0^+ and denoted by $g \in \mathrm{RV}_\alpha(0^+)$. Note that $g(t) \in \mathrm{RV}_\alpha \iff g(1/t) \in \mathrm{RV}_{-\alpha}(0^+)$. Similarly, $\mathrm{RV}_0(0^+)$ is defined. We will use $\ell(x)$ to represent a slowly varying function, and a regularly varying function g can be written as $g(x) = x^\alpha \ell(x)$.

Note that, when we say that a random variable is regularly varying, it actually means that the survival function of the random variable is regularly varying.

Definition 2.6 A measurable function $g : \mathbb{R}_+ \to \mathbb{R}_+$ is *rapidly varying at* ∞ *with* *index* ∞ (written $g \in \mathrm{RV}_\infty$) if for any $t > 0$,

$$\lim_{x\to\infty} \frac{g(xt)}{g(x)} = t^\infty = \begin{cases} 0, & \text{if } t < 1, \\ 1, & \text{if } t = 1, \\ \infty, & \text{if } t > 1. \end{cases}$$

Similarly, $g \in \mathrm{RV}_{-\infty}$ if for any $t > 0$,

$$\lim_{x\to\infty} \frac{g(xt)}{g(x)} = t^{-\infty} = \begin{cases} \infty, & \text{if } t < 1, \\ 1, & \text{if } t = 1, \\ 0, & \text{if } t > 1. \end{cases}$$

Proposition 2.1 *If $U \in \mathrm{RV}_\alpha$ for $\alpha \in \mathbb{R}$, then*

$$\lim_{x\to\infty} U(tx)/U(x) = t^\alpha,$$

locally uniformly in t on $(0, \infty)$. If $\alpha < 0$, then uniform convergence holds on intervals of the form $(b, \infty), b > 0$. If $\alpha > 0$, uniform convergence holds on intervals $(0, b]$ provided U is bounded on $(0, b]$ for all $b > 0$.

The following result is the famous Karamata's theorem. It tells us that if we write some regularly varying function $U(x) = x^\alpha \ell(x)$ with $\ell(x)$ slowly varying, then $U(x)$ behaves like x^α in terms of integration.

Theorem 2.2 *Suppose $\alpha \geq -1$ and $U \in \mathrm{RV}_\alpha$. Then $\int_0^x U(t)dt \in \mathrm{RV}_{\alpha+1}$ and*

$$\lim_{x\to\infty} \frac{xU(x)}{\int_0^x U(t)dt} = \alpha + 1. \tag{2.6}$$

If $\alpha < -1$ (or if $\alpha = -1$ and $\int_x^\infty U(s)ds < \infty$), then $U \in \mathrm{RV}_\alpha$ implies that $\int_x^\infty U(t)dt$ is finite, $\int_x^\infty U(t)dt \in \mathrm{RV}_{\alpha+1}$, and

$$\lim_{x\to\infty} \frac{xU(x)}{\int_x^\infty U(t)dt} = -\alpha - 1. \tag{2.7}$$

Definition 2.7 A function $g : \mathbb{R} \to (0, \infty)$ is Γ-varying if it is non-decreasing and right-continuous, and there exists a measurable function $h : \mathbb{R} \to (0, \infty)$ such that for any $t \in \mathbb{R}$

$$\lim_{x\to\infty} \frac{g(x+th(x))}{g(x)} = e^t, \tag{2.8}$$

where $h(\cdot)$ is called an auxiliary function.

Definition 2.8 A measurable function $g : \mathbb{R}^+ \to \mathbb{R}$ is Π-varying if there exits a function $h : \mathbb{R}^+ \to \mathbb{R}^+$ such that for all $t > 0$,

$$\lim_{x\to\infty} \frac{g(tx) - g(x)}{h(x)} = \log t, \tag{2.9}$$

where $h(\cdot)$ is called an auxiliary function.

The following is a definition for MRV. Due to the limitation of space, we will not give more details about MRV here. For interested readers, Resnick (2007) is a nice reference.

Definition 2.9 A random vector $\mathbf{X} \geq \mathbf{0}$ is MRV if there exists a Radon measure μ (i.e., finite on compact sets) on $E := [\mathbf{0}, \boldsymbol{\infty}] \setminus \{\mathbf{0}\}$ such that,

$$\lim_{t\to\infty} \frac{\mathbb{P}[\mathbf{X} \in tB]}{\mathbb{P}[||\mathbf{X}|| > t]} = \mu(B)$$

for any relatively compact set $B \subset E$ with $\mu(\partial B) = 0$, where $||\cdot||$ denotes a norm on $\mathbb{R}^d$.

2.4 Maximum Domain of Attraction

Definition 2.10 A random variable X is said to belong to the MDA of an extreme value distribution H if there exist normalizing constants $\sigma_n > 0$ and $\mu_n \in \mathbb{R}$ such that

$$(M_n - \mu_n)/\sigma_n \xrightarrow{d} H, \quad n \to \infty,$$

where $M_n := X_{[1]}$ is the first order statistics (i.e., maximum) of a random sample of X with the sample size being n. This is written as $X \in \text{MDA}(H)$.

It is well known that there are only three non-degenerate univariate extreme value distributions: Fréchet, Gumbel and Weibull. Since MDA(Weibull) corresponds to bounded random variables that are irrelevant to quantitative risk management, we will only consider the first two cases. The following two theorems characterize the classes of MDA(Fréchet) and MDA(Gumbel), respectively.

Theorem 2.3 *A random variable X with cdf F is said to belong to the Fréchet maximum domain of attraction if and only if $\overline{F} \in \mathrm{RV}_{-\alpha}, \alpha > 0$, and the corresponding Fréchet cdf is $\exp(-x^{-\alpha})$.*

Theorem 2.4 *A random variable X with cdf F is said to belong to the Gumbel maximum domain of attraction if and only if there exists a positive auxiliary function $a(\cdot)$ such that*

$$\lim_{x\to\infty} \frac{\overline{F}(x+ta(x))}{\overline{F}(x)} = e^{-t}, \quad t \in \mathbb{R}, \tag{2.10}$$

where $a(\cdot)$ can be chosen as $a(x) = \int_x^\infty \overline{F}(t)/\overline{F}(x)dt$.

The concept of MDA(Gumbel) here is closely related to Γ-variation: if $g(x) := 1/\overline{F}(x)$ with $\overline{F}$ satisfying (2.10), then $g(x)$ is non-decreasing and right-continuous, and satisfies the Γ-variation condition (2.8) with $h(\cdot) = a(\cdot)$ the auxiliary function in (2.10).

Note that, Marshall and Olkin (1983) has a characterization of MDA of a multivariate extreme value distribution. This paper also shows that a random vector $\mathbf{X}$ that follows a multivariate extreme value distribution must be associated; that is, for any real-valued increasing functions g_1, g_2, $\mathbb{E}[g_1(\mathbf{X})g_2(\mathbf{X})] \geq \mathbb{E}[g_1(\mathbf{X})]\mathbb{E}[g_2(\mathbf{X})]$, provided that the expectations exist.

Chapter 3

Tail order and intermediate tail dependence

3.1 Introduction

For statistical modeling with copulas, properties such as strengths of upper/lower tail dependence and reflection symmetry or direction of reflection asymmetry are important in deciding on appropriate copulas. For example, for the tail asymmetry phenomena of financial markets (Patton, 2006; Okimoto, 2008), copula families with a variety of tail behavior are useful for statistical modeling. Although the multivariate Gaussian and t copula families have a wide range of dependence, they are not appropriate when there is reflection or tail asymmetry. But copulas can be constructed from other methods to get different joint tail behavior. Then for use of copulas for inference for joint tail probabilities, sensitivity analysis over different families can be performed.

For the study of tail dependence behavior of random vectors, we not only have interest in the cases where the random vector is asymptotically dependent, but also where asymptotic independence exhibits. Ledford and Tawn (1996) proposed the following model for a bivariate random vector $(X_1, X_2)^{\mathrm{T}}$, where X_1 and X_2 are unit Fréchet distributed with cdf $F_i(x) = e^{-1/x}, x \geq 0, i = 1, 2$, and are nonnegatively associated,

$$\mathbb{P}[X_1 > r, X_2 > r] \sim \ell(r) r^{-1/\eta}, \quad r \to \infty, \tag{3.1}$$

where $1/2 \le \eta \le 1$. If we let $U_i = F_i(X_i), i = 1, 2$, where F_i is the cdf of the unit Fréchet and $r = (-\log(u))^{-1}$, then clearly

$$\begin{aligned} \lim_{u \to 1^-} \frac{\mathbb{P}[U_1 > u, U_2 > u]}{(\mathbb{P}[U_1 > u])^\kappa} &= \lim_{r \to \infty} \frac{\mathbb{P}[X_1 > r, X_2 > r]}{(\mathbb{P}[X_1 > r])^\kappa} \\ &= \lim_{r \to \infty} \frac{\ell(r) r^{-1/\eta}}{[1 - \exp(-r^{-1})]^\kappa} = \lim_{r \to \infty} \frac{\ell(r) r^{-1/\eta}}{r^{-\kappa}}. \end{aligned} \tag{3.2}$$

Thus the "tail order" κ that we will introduce in Definition 3.1 corresponds to $1/\eta$ of Ledford and Tawn's representation. If $\eta = 1, i.e., \kappa = 1$ and $\ell(r) \nrightarrow 0$, X_1 and X_2 are upper tail dependent with upper tail dependence parameter $\lambda_U = \lim_{r \to \infty} \ell(r)$; if $1/2 < \eta < 1$, they are positively dependent; if $\eta = 1/2$ and $\ell(r) \ge 1$, they are "tail quadrant independence". A lot of research has been done following this direction. We refer to Ledford and Tawn (1996, 1997); Coles et al. (1999); Heffernan (2000); Ramos and Ledford (2009) for further development of this idea.

The relation (3.1) tells us that the power term $1/\eta$ dominates the speed of decay of the joint tail probability. We believe that the parameter $1/\eta$ plays an important role in the study of tail dependence behavior, and deserves a new name "tail order" that is explained in Section 3.2.1, based on copula functions. Moreover, analogously to the tail dependence function, we will propose the tail order function, which includes the information of the convergence along routes other than the diagonal.

In this chapter, the emphasis is on the case where the tail order is between 1 and d for a d-dimensional random vector. We refer to this case as "intermediate tail dependence" under some positive dependence assumptions; this is explained before Example 3.1.

Our main contributions span the following aspects: 1. We propose the concepts of tail order and tail order functions as an integrated way to study tail behavior of multivariate copulas. 2. We relate the tail heaviness of a positive random variable to the tail behavior of the Archimedean copula constructed by the LT of the random variable. In our opinion, it is an insightful way to better understand the tail behavior of Archimedean copulas. 3. Our theoretical study of tail behavior of Archimedean copulas leads to a new one-parameter Archimedean copula family, based on the LT of the inverse Gamma distribution, which shows patterns of upper and lower tails not seen in commonly used copula families.

The remainder of this chapter is organized as follows. Section 3.2 introduces the concepts of tail order and tail order functions, and some properties of them. In particular, some results on relations of tail orders of marginal copulas are given. Sections 3.3 and 3.4 contain studies of intermediate tail dependence for Archimedean copulas and copulas constructed by mixture of max-id distributions, respectively. For multivariate Archimedean copulas, we have a more concrete result than Charpentier and Segers (2009) for the lower tail, and new results for the upper tail. Asymptotic behavior of LTs of positive random variables is studied in Section 3.3.1, and the new Archimedean copula family is presented in Section 3.3.4. Finally, Section 3.5 concludes with some topics of further research. The main proofs are put in Section 3.6.

3.2 Tail orders: definitions and properties

In this section, we define the concepts of tail order and tail order functions, indicate their use for reflection asymmetry and derive some properties.

3.2.1 Multivariate tail order and tail order functions

To avoid technicalities for tail orders, we assume conditions involving regular variation of tails of copula and other functions.

Definition 3.1 Suppose C is a d-dimensional copula. If there exists some $\kappa_L(C) > 0$ such that, with some $\ell(u) \in \mathrm{RV}_0(0^+)$

$$C(u\mathbf{1}_d) \sim u^{\kappa_L(C)}\ell(u), \quad u \to 0^+,$$

then we refer to $\kappa_L(C)$ as the *lower tail order* of C and refer to $\lambda_L(C) = \lim_{u\to 0^+} \ell(u)$ as the *lower tail order parameter*, provided the limit exists. Similarly, the *upper* *tail order* is defined as $\kappa_U(C)$ such that

$$\overline{C}((1-u)\mathbf{1}_d) \sim u^{\kappa_U(C)}\ell(u), \quad u \to 0^+,$$

with the *upper tail order parameter* $\lambda_U(C) = \lim_{u\to 0^+} \ell(u)$, provided the limit exists.

When no confusion arises, we use the notation κ to represent lower or upper tail orders, and λ for tail order parameters. $\kappa_L(C) = 1$ [resp. $\kappa_U(C) = 1$] and $\ell(u) \nrightarrow 0$ corresponds to the usual definition of upper [resp. lower] tail dependence. We will assume that $\lim_{s\to 0^+} \ell(s) = h \in [0,\infty]$. But $h = 0$ or $h = \infty$ correspond to boundary cases, in which case more care is needed. In these boundary cases, the "speed" of decrease or increase of $\ell(u)$ affects the tail dependence behavior. For example, if $\ell(u) \to 0$, then a lower speed indicates a stronger tail dependence; if $\ell(u) \to +\infty$, then a higher speed indicates a stronger tail dependence. Note that with $\ell(u) \to h$, if $\kappa(C) = 1$ then $0 \le h \le 1$; if $\kappa(C) > 1$ then $0 \le h \le \infty$. Note also that $\kappa_L(C) = \kappa_U(C) = d$ for the d-dimensional independence copula. It is not possible for $\kappa < 1$ (refer to Proposition 3.3), but it is possible for $\kappa_L(C)$ and $\kappa_U(C)$ to be greater than d for copulas with negative dependence. For example, as a boundary case, for the bivariate counter-monotonic copula, $\kappa_L(C)$ and $\kappa_U(C)$ can be considered as $+\infty$ because $C(u,u)$ and $\overline{C}(1-u, 1-u)$ are zero for $0 < u < 1/2$.

The cases of $\kappa = 1$ or d have been well studied in the literature, while not much research exists for $1 < \kappa < d$. For the bivariate case, $1 < \kappa < 2$ represents some level of positive dependence in the tail, but not as strong as tail dependence. For multivariate cases, without any further conditions, the meaning of $1 < \kappa < d$ is complicated. We refer to the case $1 < \kappa < d$ as lower [resp. upper] intermediate tail dependence only when all marginal copulas (ultimately) possess positive lower [resp. upper] orthant dependence, of which a formal definition will be given in Definition 3.3. Unless otherwise specified, when a copula is said to possess intermediate tail dependence, the orthant dependence condition is assumed implicitly. The following is an example of intermediate tail dependence for Gaussian copulas.

Example 3.1 (Gaussian copula) Consider a multivariate Gaussian copula, constructed by

$$C_{\Phi_d}(u_1, \ldots, u_d) = \Phi_d(\Phi^{-1}(u_1), \ldots, \Phi^{-1}(u_d); \Sigma), \tag{3.3}$$

where $\Phi_d(\cdot; \Sigma)$ is the joint cdf of a standard d-variate Gaussian random vector with positive definite correlation matrix Σ. The multivariate Gaussian copula defined in (3.3) has intermediate tail dependence with the tail order $\kappa = \mathbf{1}_d \Sigma^{-1} \mathbf{1}_d^{\mathrm{T}}$, the sum of all elements of Σ^{-1}. It can be verified by noticing that $\lim_{u \to 0^+} C_{\Phi_d}(u\mathbf{1}_d)/u^\kappa = \lim_{t \to -\infty} \Phi_d(t\mathbf{1}_d)/[\Phi(t)]^\kappa$, and as $t \to -\infty$, $\Phi(t) \sim \phi(t)/|t|$ and $\Phi_d(t\mathbf{1}_d)$ is dominated by the exponent term $\exp(-t^2 \mathbf{1}_d \Sigma^{-1} \mathbf{1}_d^{\mathrm{T}}/2)$ (Hashorva and Hüsler, 2003, Corollary 4.1).

The bivariate Gaussian copula with $\varrho > 0$ has intermediate tail dependence with the tail order $\kappa = 2/(1 + \varrho)$ and the slowly varying function at 0^+ being $\ell(u) = (-\log u)^{-\varrho/(1+\varrho)}$. A related result without using copula functions has been given in Ledford and Tawn (1996). For dimension d with constant correlation ϱ, the tail order is $\kappa = d/[1+(d-1)\varrho]$. For the trivariate case with $\varrho_{12} = \varrho_{23} = \varrho$, we have $\kappa = [3 + \varrho_{13} - 4\varrho]/[1 + \varrho_{13} - 2\varrho^2]$.

Gaussian copulas are reflection symmetric and have intermediate tail dependence when correlations are positive. They are a subfamily of the elliptical copulas. Under some regularity conditions, tail orders of elliptical copulas will be determined by the tail behavior of corresponding radial random variable R.

Since copula is invariant to a strict increasing transformation on margins, for the study of elliptical copula, we may omit the location and scale parameters of joint elliptical distributions, and consider the following representation: let $\mathbf{X} := (X_1, X_2)$ be an elliptical random vector such that

$$\mathbf{X} \stackrel{d}{=} RA\mathbf{U}, \tag{3.4}$$

where the radial random variable $R \geq 0$ is independent of $\mathbf{U}$, $\mathbf{U}$ is an bivariate random vector uniformly distributed on the surface of the unit hypersphere $\{\mathbf{z} \in \mathbb{R}^k | \mathbf{z}^{\mathrm{T}}\mathbf{z} = 1\}$, A is a 2×2 matrix such that $AA^{\mathrm{T}} = \Sigma$ where $\Sigma_{11} = \Sigma_{22} = 1$

and $\Sigma_{12} = \Sigma_{21} = \varrho$ with $-1 < \varrho < 1$, eg., $A = \begin{pmatrix} 1 & 0 \\ \varrho & \sqrt{1-\varrho^2} \end{pmatrix}$. For such an elliptical distribution, the margins have the same distribution assumed to be F.

The tail behavior of R may influence the tail order of elliptical copulas. For the usual tail dependence case, Schmidt (2002) proved that when the radial random variable R has a regularly varying tail then X_1 and X_2 are tail dependent, and thus the tail order of the corresponding elliptical copula is $\kappa = 1$.

Example 3.2 (Student t copula) The radial random variable R for Student t distributions is a generalized inverse Gamma distribution such that R^2 follows an inverse Gamma distribution with the shape and scale parameters being $\nu/2$, where ν is the degree of freedom. It can be verified that $\overline{F}_R \in \mathrm{RV}_{-\nu}$ (see Example 6.4). So the tail order for Student t copula is $\kappa = 1$.

For the case where R has lighter tails than any regularly varying tails, some asymptotic study has been conducted for elliptical distributions where R belongs to the Gumbel Maximum Domain of Attraction. We refer to Hashorva (2007), Hashorva (2010) and Hashorva (2008) for relevant references.

Here we are ready to present a result that is useful to find the tail order of a bivariate elliptical copula where the radial random variable R belongs to Gumbel MDA.

Proposition 3.1 *Let C be the copula for an elliptical random vector $\mathbf{X} := (X_1, X_2)$ constructed as (3.4), and $b_\varrho = \sqrt{2/(1+\varrho)}$. If $R \in \mathrm{MDA(Gumbel)}$, then the upper and lower tail orders of C is*

$$\kappa = \lim_{r\to\infty} \frac{\log\left(1 - F_R(b_\varrho r)\right)}{\log\left(1 - F_R(r)\right)}, \tag{3.5}$$

provided the limit exists.

This result is very convenient for us to derive the tail order if we know the tail behavior of R.

Example 3.3 (Bivariate symmetric Kotz type (Fang et al., 1990) copula) The density generator

$$g(x) = Kx^{N-1}\exp\{-\beta x^{\xi}\}, \quad \beta, \xi, N > 0,$$

where K is a normalizing constant. By Theorem 2.9 of Fang et al. (1990), the density function of R is

$$f_R(x) = 2\pi x g(x^2) = 2K\pi x^{2N-1}\exp\{-\beta x^{2\xi}\}.$$

So, the survival function is

$$\begin{aligned}
1 - F_R(x) &= \int_x^{\infty} 2K\pi t^{2N-1}\exp\{-\beta t^{2\xi}\}dt \\
&= \int_{\beta x^{2\xi}}^{\infty} \frac{K\pi}{\xi}\beta^{-N/\xi}w^{N/\xi-1}\exp\{-w\}dw \\
&= \frac{K\pi}{\xi}\beta^{-N/\xi}\Gamma(N/\xi, \beta x^{2\xi}), \quad \Gamma(\cdot,\cdot) \text{ incomplete Gamma function} \\
&\sim \frac{K\pi}{\xi}\beta^{-1}x^{2N-2\xi}\exp\{-\beta x^{2\xi}\}, \quad x \to \infty.
\end{aligned}$$

Then by (3.5), we can easily get that

$$\kappa = b_{\varrho}^{2\xi} = [2/(1+\varrho)]^{\xi}.$$

Therefore, the tail order for the symmetric Kotz type copula is $\kappa = [2/(1+\varrho)]^{\xi}$. Gaussian copula belongs to this class with $\xi = 1$, so its tail order is $2/(1+\varrho)$ which is consistent to what we have obtained.

Definition 3.2 Suppose C is a d-dimensional copula and $C(u\mathbf{1}_d) \sim u^{\kappa}\ell(u)$, $u \to 0^+$ for some $\ell(u) \in \mathrm{RV}_0(0^+)$. The *lower tail order function* $b : \mathbb{R}_+^d \to \mathbb{R}_+$ is defined as

$$b(\mathbf{w}; C, \kappa) = \lim_{u \to 0^+} \frac{C(uw_j, 1 \le j \le d)}{u^{\kappa}\ell(u)},$$

provided the limit function exists. In parallel, if $\overline{C}((1-u)\mathbf{1}_d) \sim u^{\kappa}\ell(u)$, $u \to 0^+$ for some $\ell(u) \in \mathrm{RV}_0(0^+)$, the *upper tail order function* $b^* : \mathbb{R}_+^d \to \mathbb{R}_+$ is defined as

$$b^*(\mathbf{w}; C, \kappa) = \lim_{u \to 0^+} \frac{\overline{C}(1 - uw_j, 1 \le j \le d)}{u^\kappa \ell(u)},$$

provided the limit function exists. If $\ell(u) \to h \neq 0$, then $hb(\mathbf{w}; C, 1)$ and $hb^*(\mathbf{w}; C, 1)$ become the tail dependence functions in Joe et al. (2010).

Note that the copula C that satisfies the conditions of the above definition is said to be multivariate regularly varying with a limit function b or b^* (Resnick, 2007). Although the general theory and definitions must accommodate an arbitrary slowly varying function ℓ, in specific parametric families of copulas that have tractable forms, we find that either $\ell(u)$ is a constant or proportional to a power of $(-\log u)$.

Example 3.4 (Extreme value copula) For any multivariate extreme value copula C^{EV}, there exists a function $A : [0, \infty)^d \to [0, \infty)$ such that

$$C^{EV}(u_1, \ldots, u_d) = \exp\{-A(-\log u_1, \ldots, -\log u_d)\},$$

where A is convex, homogeneous of order 1 and satisfies

$$\max(x_1, \ldots, x_d) \le A(x_1, \ldots, x_d) \le x_1 + \cdots + x_d.$$

We refer to Chapter 6 of Joe (1997) for details of multivariate extreme value copulas. Thus,

$$C^{EV}(u\mathbf{1}_d) = \exp\{A(\mathbf{1}_d) \log u\} = u^{A(\mathbf{1}_d)}.$$

That is, for any extreme value copula C^{EV}, the lower tail order is $\kappa_L(C^{EV}) = A(\mathbf{1}_d)$ and there is intermediate lower tail dependence except for the boundary cases such as independence copula and comonotonicity copula, where $A(\mathbf{1}_d) = d$ and 1, respectively.

In order to get the lower tail order function of extreme value copulas, first consider the bivariate case, for which

$$
\begin{aligned}
& C^{EV}(uw_1, uw_2) \\
& = \exp\{-A(-\log uw_1, -\log uw_2)\} \\
& = \exp\left\{(\log u) A\left(1+\frac{\log w_1}{\log u}, 1+\frac{\log w_2}{\log u}\right)\right\} \\
& \sim \exp\left\{(\log u)\left[A(1,1)+A_1(1,1)\left(\frac{\log w_1}{\log u}\right)+A_2(1,1)\left(\frac{\log w_2}{\log u}\right)\right]\right\}, u \to 0^+ \\
& = u^{A(1,1)} w_1^{A_1(1,1)} w_2^{A_2(1,1)},
\end{aligned}
$$

where $A_i = \partial A/\partial x_i, i = 1, 2$. Therefore, the lower tail order function is $b(w_1, w_2) = w_1^{A_1(1,1)} w_2^{A_2(1,1)}$. Similarly, for a d-variate extreme value copula, $b(w_1, \ldots, w_d) = w_1^{A_1(\mathbf{1}_d)} \ldots w_d^{A_d(\mathbf{1}_d)}$. By Euler's formula for homogeneous functions, $A(\mathbf{1}_d) = \sum_{i=1}^d A_i(\mathbf{1}_d)$. Then it can be verified that b is homogeneous of order $A(\mathbf{1}_d)$.

In the bivariate case, $\kappa_U = 1$, $\lambda_U = 2 - A(1,1)$, and $\kappa_L = A(1,1)$. That is, a larger value of the upper tail dependence parameter implies stronger lower intermediate tail dependence.

Example 3.5 (Elliptical copula) Tail order functions of bivariate elliptical copulas can be derived as following: for any given $0 < w_1, w_2 < \infty$, as $0 < v < 1$ is sufficiently close to 0,

$$
\begin{aligned}
& \overline{C}(1-vw_1, 1-vw_2) \\
& = \mathbb{P}[F(X_1) \geq 1-vw_1, F(X_2) \geq 1-vw_2] \\
& = \mathbb{P}[X_1 \geq F^{-1}(1-vw_1), X_2 \geq F^{-1}(1-vw_2)] \\
& = \mathbb{P}[R\cos\varphi \geq F^{-1}(1-vw_1), R[\varrho\cos\varphi + \sqrt{1-\varrho^2}\sin\varphi] \geq F^{-1}(1-vw_2)] \\
& = \frac{1}{2\pi}\int_{\theta\in\Theta_1} \mathbb{P}\left[R \geq \frac{F^{-1}(1-vw_1)}{\cos\theta}, R \geq \frac{F^{-1}(1-vw_2)}{\varrho\cos\theta+\sqrt{1-\varrho^2}\sin\theta}\right] d\theta,
\end{aligned}
$$

where $\varphi \sim$ Uniform$(-\pi, \pi)$ and $\Theta_1 := (-\pi/2, \pi/2) \cap \{\theta : \tan\theta \geq -\varrho/\sqrt{1-\varrho^2}\}$.

Also, letting $\Theta_2 := (-\pi/2, \pi/2)$, as $0 < v < 1$ is sufficiently close to 0

$$v = \mathbb{P}\left[R\cos\varphi \geq F^{-1}(1-v)\right] = \frac{1}{2\pi}\int_{\theta\in\Theta_2} \mathbb{P}\left[R \geq \frac{F^{-1}(1-v)}{\cos\theta}\right] d\theta.$$

If $R \in$ MDA(Fréchet), i.e., $R \in \mathrm{RV}_{-\alpha}$ for some $\alpha > 0$, then the margins $X_1, X_2 \in \mathrm{RV}_{-\alpha}$. Therefore, the map $v \mapsto F^{-1}(1-v) \in \mathrm{RV}_{-1/\alpha}(0^+)$. Then, by the uniform convergence of regularly varying functions (e.g., Proposition 2.4 of Resnick (2007))

$$\begin{aligned}
&\lim_{v\to 0^+} \frac{\overline{C}(1-vw_1, 1-vw_2)}{v} \\
&= \lim_{v\to 0^+} \frac{\int_{\theta\in\Theta_1} \mathbb{P}\left[R \geq \frac{F^{-1}(1-vw_1)}{\cos\theta} \vee \frac{F^{-1}(1-vw_2)}{\varrho\cos\theta+\sqrt{1-\varrho^2}\sin\theta}\right] d\theta / \mathbb{P}\left[R > F^{-1}(1-v)\right]}{\int_{\theta\in\Theta_2} \mathbb{P}\left[R \geq \frac{F^{-1}(1-v)}{\cos\theta}\right] d\theta / \mathbb{P}\left[R > F^{-1}(1-v)\right]} \\
&= \lim_{v\to 0^+} \frac{\int_{\theta\in\Theta_1} \mathbb{P}\left[R \geq \frac{w_1^{-1/\alpha}F^{-1}(1-v)}{\cos\theta} \vee \frac{w_2^{-1/\alpha}F^{-1}(1-v)}{\varrho\cos\theta+\sqrt{1-\varrho^2}\sin\theta}\right] d\theta / \mathbb{P}\left[R > F^{-1}(1-v)\right]}{\int_{\theta\in\Theta_2} \mathbb{P}\left[R \geq \frac{F^{-1}(1-v)}{\cos\theta}\right] d\theta / \mathbb{P}\left[R > F^{-1}(1-v)\right]} \\
&= \frac{\int_{\theta\in\Theta_1} \max\left\{w_1(\cos\theta)^\alpha, w_2(\varrho\cos\theta + \sqrt{1-\varrho^2}\sin\theta)^\alpha\right\} d\theta}{\int_{\theta\in\Theta_2}(\cos\theta)^\alpha d\theta}.
\end{aligned}$$

The multivariate case is similar (Klüppelberg et al., 2008). However, if $R \in$ MDA(Gumbel), a general form of tail order functions of the whole class remains unsolved.

We next mention how the upper and lower tail orders are useful to establish the direction of reflection asymmetry. Let $\widehat{C}$ be the copula of $(1-U_1, \ldots, 1-U_d)$ when the copula of $(U_1, \ldots, U_d)$ is C, where U_i's are standard uniform variables. Reflection symmetry means that $\widehat{C} \equiv C$ (Definition 2.3) and otherwise we say that there is reflection asymmetry. If $C(u\mathbf{1}_d) \geq \widehat{C}(u\mathbf{1}_d)$ for all $0 < u < u_0$, for some $0 < u_0 \leq 1/2$, then the copula has more probability in the lower tail (reflection asymmetry with skewness to lower tail). If the inequality is reversed leading to $C(u\mathbf{1}_d) \leq \widehat{C}(u\mathbf{1}_d)$, then the copula has more probability in the upper tail (reflection asymmetry with skewness to upper tail). For most existing parametric families of copulas, it is difficult to analytically compare $C(u\mathbf{1}_d)$ and $\widehat{C}(u\mathbf{1}_d)$, so the direction of reflection asymmetry is analytically easier via the upper and lower tail orders. For example, if $\kappa_L(C) > \kappa_U(C)$, then C has reflection asymmetry skewed to the upper tail (smaller κ means slower convergence to 0), and if $C(u\mathbf{1}_d) \sim \lambda_L u^\kappa$ and $\widehat{C}(u\mathbf{1}_d) \sim \lambda_U u^\kappa$ as $u \to 0^+$ with $\lambda_L > \lambda_U > 0$, then C has reflection asymmetry skewed to the lower tail. For many parametric copula families where we have done numerical computations, u_0 can be taken as $1/2$.

The following are some elementary properties of the lower and upper tail order functions b and b^*. Obvious properties of tail order for $\widehat{C}$ are the following: $\kappa_L(C) = \kappa_U(\widehat{C})$, $\kappa_U(C) = \kappa_L(\widehat{C})$, $b(\mathbf{w}; C, \kappa) = b^*(\mathbf{w}; \widehat{C}, \kappa)$ and $b^*(\mathbf{w}; C, \kappa) = b(\mathbf{w}; \widehat{C}, \kappa)$.

Proposition 3.2 *A lower tail order function $b(\mathbf{w}) = b(\mathbf{w}; C, \kappa)$ has following properties:*

1. *$b(\mathbf{1}) \equiv 1$, and $b(\mathbf{w}) = 0$ if there exists an $i \in I_d$ with $w_i = 0$;*
2. *$b(\mathbf{w})$ is increasing in $w_i, i \in I_d$;*
3. *for any fixed $t > 0$,*

$$\begin{aligned} b(t\mathbf{w}) &= \lim_{u \to 0^+} \frac{C(tuw_j, 1 \le j \le d)}{u^\kappa \ell(u)} \\ &= t^\kappa \lim_{u \to 0^+} \frac{C(tuw_j, 1 \le j \le d)}{(tu)^\kappa \ell(tu)} = t^\kappa b(\mathbf{w}). \end{aligned} \tag{3.6}$$

Thus, $b(\mathbf{w})$ is homogeneous of order κ.

If $b(\mathbf{w})$ is partially differentiable with respect to each w_i on $(0, +\infty)$, then by the Euler's formula on homogeneous functions, we can write

$$b(\mathbf{w}) = \frac{1}{\kappa} \sum_{j=1}^{d} \frac{\partial b}{\partial w_j} w_j, \quad \forall \mathbf{w} \in \mathbb{R}_+^d.$$

Remark 3.1 Since $C(u\mathbf{w}) \sim u^\kappa \ell(u) b(\mathbf{w}) = b(u\mathbf{w})\ell(u)$, $u \to 0^+$, the tail order function b captures the tail behavior of the copula C in different directions.

3.2.2 Further properties of tail orders

In this subsection, we obtain some general properties of tail orders of multivariate copulas, especially on inequalities on tail orders of marginal copulas. There is an "obvious" property in terms of concordance. For two joint cdfs $H, G \in \mathcal{R}(F_1, \ldots, F_d)$, we say that H is less concordant than G, if $H(\mathbf{x}) \leq G(\mathbf{x})$ and $\overline{H}(\mathbf{x}) \leq \overline{G}(\mathbf{x})$ for any $\mathbf{x}$ in the support of H and G. If C_1 is less concordant than C_2, then $\kappa_L(C_1) \geq \kappa_L(C_2)$ and $\kappa_U(C_1) \geq \kappa_U(C_2)$.

Next we introduce some concepts of positive dependence, under which multivariate copulas may have some particular properties on tail orders. We refer to Joe (1997), Colangelo et al. (2005) for details.

Definition 3.3 Suppose that $F(\mathbf{x})$ is the cdf of a d-variate random vector $\mathbf{X} = (X_1, \ldots, X_d)^{\mathrm{T}}$, then $\mathbf{X}$ or F is said to be

1. positive lower orthant dependent (PLOD) if $\mathbb{P}[X_i \leq x_i, \forall i \in I_d] \geq \prod_{i=1}^d \mathbb{P}[X_i \leq x_i]$ for any $\mathbf{x} \in \mathbb{R}^d$;

2. left tail decreasing in sequence (LTDS) if $\mathbb{P}[X_i \leq x_i | X_1 \leq x_1, \ldots, X_{i-1} \leq x_{i-1}]$ is decreasing in $x_1, \ldots, x_{i-1}$ for all $x_i, i \in \{2, \ldots, d\}$;

3. multivariate left tail decreasing (MLTD) if $(X_{i_1}, \ldots, X_{i_d})$ is LTDS for all permutation $(i_1, \ldots, i_d)$ of $(1, \ldots, d)$.

Proposition 3.3 *Suppose a multivariate copula $C(u_1, \ldots, u_d)$ has a lower tail order $\kappa_L(C)$, then $\kappa_L(C) \geq 1$. Moreover,*

1. *if C is (ultimately) positive lower orthant dependent (PLOD), then $\kappa_L(C) \leq d$;*

2. *for any $S_1 \subset S_2 \subseteq I_d$ with $|S_1| \geq 2$, $\kappa_L(C_{S_2}) - \kappa_L(C_{S_1}) \geq 0$. In particular, if $\kappa_L(C) = 1$, then for any $S \subset I_d$ with $|S| \geq 2$, $\kappa_L(C_S) = 1$; if C is multivariate left tail decreasing (MLTD), then $\kappa_L(C_{S_2}) - \kappa_L(C_{S_1}) \leq |S_2| - |S_1|$.*

Analogous results hold with κ_L replaced by κ_U, and conditions of positive upper orthant dependence and multivariate right tail increasing.

Remark 3.2 The above result says that when some regularity condition holds, marginality will keep the order of tail orders in the sense that margins have smaller tail orders. However, marginality does not inherit the inequality between tail orders of lower and upper tails. For example, take the trivariate Archimedean copula with the ψ function in Example 3.6 in Section 3.3.1 below. Then $3^\alpha > 1 + \alpha$ for $0 < \alpha < 1$ so that $\kappa_L(C) > \kappa_U(C)$ (see Table 3.1). But $2^\alpha < 1+\alpha$ for $0 < \alpha < 1$ so that for the bivariate margins, $\kappa_L(C_I) < \kappa_U(C_I)$ with $|I| = 2$.

Sometimes partial derivatives and the density have a simpler form than the copula cdf. We hope to know what tail properties will be inherited if we take partial derivatives of the copula. For example, for the lower tail, if

$$C(uw_1, \dots, uw_d) \sim u^\kappa \ell(u) b(w_1, ..., w_d), \quad u \to 0^+,$$

then we want to differentiate both sides of the above with respect to the w_j's to get:

$$u \frac{\partial C(uw_1, \dots, uw_d)}{\partial w_j} \sim u^\kappa \ell(u) \frac{\partial b(w_1, ..., w_d)}{\partial w_j}, \quad u \to 0^+,$$

and higher order derivatives up to:

$$u^d \frac{\partial^d C(uw_1, \dots, uw_d)}{\partial w_1 \cdots \partial w_d} \sim u^\kappa \ell(u) \frac{\partial^d b(w_1, ..., w_d)}{\partial w_1 \cdots \partial w_d}, \quad u \to 0^+.$$

A sufficient condition is ultimate monotonicity of partial derivatives of the copula (eg: $\partial C / \partial u_j$ is ultimately monotone in u_j at 0^+, and similar conditions are sufficient for higher orders). A proof is similar to that in Theorem 1.7.2 (Monotone density theorem) in Bingham et al. (1987).

As an example of using the density to get the tail order, consider a multivariate Gaussian copula with positive definite correlation matrix Σ which satisfies $C_\Phi(u\mathbf{1}_d) \sim u^\kappa \ell(u) = u^\kappa(-\log u)^\zeta$, $u \to 0^+$. Then (as can be shown directly with the monotone density theorem), this would be equivalent to $c_\Phi(u\mathbf{1}_d) \sim hu^{\kappa-d}(-\log u)^\zeta$, $u \to 0^+$, where h is a constant. Thus, with ϕ_d for the multivariate Gaussian density,

$$
\begin{aligned}
1 &= \lim_{u\to 0^+} \frac{c_\Phi(u\mathbf{1}_d)}{hu^{\kappa-d}(-\log u)^\zeta} = \lim_{u\to 0^+} \frac{\phi_d\left(\Phi^{-1}(u)\mathbf{1}_d;\Sigma\right)}{\phi^d(\Phi^{-1}(u))u^{\kappa-d}(-\log u)^\zeta h} \\
&= \lim_{z\to-\infty} \frac{\phi_d\left(z\mathbf{1}_d;\Sigma\right)}{\phi^d(z)[\Phi(z)]^{\kappa-d}[-\log(\Phi(z))]^\zeta h} \\
&= \lim_{z\to-\infty} \frac{\phi_d\left(z\mathbf{1}_d;\Sigma\right)}{\phi^\kappa(z)|z|^{d-\kappa}[-\log(\phi(z)/|z|)]^\zeta h}
\end{aligned} \tag{3.7}
$$

Since the exponent terms dominate the numerator and denominator of (3.7), to cancel the exponent terms, a necessary condition is that $\kappa = \mathbf{1}_d\Sigma^{-1}\mathbf{1}_d^{\mathrm{T}}$, which turns out to be the tail order of the copula C_Φ. Also, to cancel the term of $|z|$ in (3.7), we need that $d-\kappa+2\zeta=0$, so $\zeta=(\kappa-d)/2$.

3.3 Intermediate tail dependence : Archimedean copulas

Archimedean copulas are reflection asymmetric except for the bivariate Frank copula, and have a variety of tail behavior. In this section, we will study the upper/lower tail orders and tail order functions for Archimedean copulas. A new family of one-parameter Archimedean copulas, that interpolates independence and comonotonicity, will be given. This copula possesses intermediate upper and lower tail dependence, and has patterns of tail orders different from existing parametric families.

Before getting to the main results, we provide some intuition on conditions on ψ for intermediate upper and lower tail dependence for C_ψ.

Let $G_1,\dots,G_d$ be univariate cdfs. For $\eta>0$, $G_1^\eta,\dots,G_d^\eta$ are cdfs, and η is called a *resilience parameter*. As the parameter $\eta\to 0$, then random variables with distributions $G_1^\eta,\dots,G_d^\eta$ tend towards the lower endpoint of support of $G_1,\dots,G_d$, and as $\eta\to\infty$, random variables with distributions $G_1^\eta,\dots,G_d^\eta$ tend towards the upper endpoint of support of $G_1,\dots,G_d$. There is also a parallel for survival functions and frailty, where the conclusions are reversed when the frailty parameter goes to 0 or ∞.

In this way, an Archimedean copula C_ψ has a mixture representation with LT ψ. That is, $C_\psi(u_1,\dots,u_d)=\int_0^\infty \prod_{j=1}^d G_j^\eta(u_j)\,dF_H(\eta)$, where F_H is the cdf of the resilience random variable H, $G(u)=\exp\{-\psi^{-1}(u)\}$ $(0\le u\le 1)$, and

$\psi(s) = \psi_H(s) = \int_0^\infty e^{-s\eta} dF_H(\eta)$. The mixture means that: there are random variables $X_1, \ldots, X_d$ such that given $H = \eta$, they are conditionally independent with respective cdfs $G_1^\eta, \ldots, G_d^\eta$. If the random variable H has heavy tail at ∞, then there is a "chance" that $H = \eta$ is large and hence conditionally, $X_1, \ldots, X_d$ are all close to their upper endpoints of support (i.e., dependence in the upper tail). Hence conditions on the heaviness of the upper tail of the distribution of H lead to intermediate upper tail dependence. For the opposite tail, if the random variable H has concentration of density near 0, then there is a "chance" that $H = \eta$ is near zero and hence conditionally, $X_1, \ldots, X_d$ are all close to their lower endpoints of support (i.e., dependence in the lower tail). Hence conditions on the density of the lower tail of the distribution of H lead to intermediate lower tail dependence.

3.3.1 Laplace transform and univariate tail heaviness

In this subsection, we relate the asymptotic behavior of a LT to the maximal moment of the positive random variable with the given LT.

Definition 3.4 For a positive random variable Y with LT ψ, the *maximal non-negative moment* is

$$M_Y = M_\psi = \sup\{m \geq 0 : \mathbb{E}(Y^m) < \infty\}. \tag{3.8}$$

M_Y is 0 if no moments exist and M_Y is ∞ if all moments exist. A smaller value of M_Y means that Y has a heavier tail at ∞.

The next lemma shows that M_ψ is related to the behavior of ψ at 0 when $0 < M_\psi < 1$. The result for a general non-integer M_ψ such that $k < M_\psi < k + 1$ will be derived subsequently.

Lemma 3.4 *Suppose $\psi(s)$ is the LT of a positive random variable Y, with $0 < M_Y < 1$. If $1 - \psi(s)$ is regularly varying at 0^+, then $1 - \psi(s) \in \mathrm{RV}_{M_Y}(0^+)$.*

Remark 3.3 Even if $\mathbb{E}(Y) = \infty$, we may also have $M_Y = 1$. However, Lemma 3.4 does not hold in general for this case.

Remark 3.4 If we write $1 - \psi(s) = s^M \ell(s)$ and $\ell(s) \to h_1$ with $0 < h_1 < \infty$ as $s \to 0^+$, then clearly $\psi(s) = 1 - h_1 s^M + o(s^M)$, $s \to 0^+$.

Proposition 3.5 *Suppose $\psi(s)$ is the LT of a positive random variable Y, with $k < M_Y < k+1$ where $k \in \{0\} \cup \mathbb{N}_+$. If $\left|\psi^{(k)}(0) - \psi^{(k)}(s)\right|$ is regularly varying at 0^+, then $\left|\psi^{(k)}(0) - \psi^{(k)}(s)\right| \in \mathrm{RV}_{M_Y-k}(0^+)$. In particular, if the slowly varying component is $\ell(s)$ and $\lim_{s\to 0^+} \ell(s) = h'_{k+1}$ with $0 < h'_{k+1} < \infty$, then $s \to 0^+$,*

$$\psi(s) = 1 - h_1 s + h_2 s^2 - \cdots + (-1)^k h_k s^k + (-1)^{k+1} h_{k+1} s^{M_Y} + o(s^{M_Y}), \tag{3.9}$$

where $0 < h_i < \infty$ for $i = 1, \ldots, k+1$.

The above results can be summarized as follows. If $M_\psi = \infty$, the LT $\psi(s)$ has an infinite Taylor expansion about $s = 0$. If M_ψ is finite and non-integer-valued, then with some regularity conditions, $\psi(s)$ has a Taylor expansion about $s = 0$ up to order $[M_\psi]$, and the next term after this has order M_ψ.

3.3.2 Upper tail

Based on the results in Section 3.3.1, we derive upper tail orders and corresponding tail order functions of multivariate Archimedean copulas; the results extend those of Charpentier and Segers (2009).

Proposition 3.6 *Let ψ be the LT of a positive random variable and assume that ψ satisfies the condition of Proposition 3.5. Assume that $k < M_\psi < k+1$ with some $k \in \{1, \ldots, d-1\}$, then the Archimedean copula C_ψ in (2.4) has upper intermediate tail dependence. The corresponding tail order is $\kappa_U = M_\psi$. If $\psi^{(i)}(0)$ is finite for all $i \in I_d$, then the upper tail order $\kappa_U = d$. If $\psi'(0)$ is infinite and $0 < M_\psi < 1$, then the upper tail order is $\kappa_U = 1$, and particularly for the bivariate case, $\lambda_U = 2 - 2^{M_\psi}$.*

Remark 3.5 If we know the value of a in (3.22) in the proof of Proposition 3.6, then the tail order parameter is

$$\begin{aligned}\lim_{u\to 1^-} \overline{C}_\psi(u,u)/(1-u)^{M_\psi} &= 2a[-\psi'(0)]^{-1-M}(2^M - 1)/(1+M) \\ &= a[-\psi'(0)]^{-M_\psi}(2^{M_\psi} - 2)/M_\psi.\end{aligned}$$

Example 3.6 Consider the LT of Example 4.2 in Joe and Ma (2000) with parameter $0 < \alpha < 1$ (see Joe-Ma in Table 1). We refer to this as the normalized integral of the positive stable LT. Note that $m = \psi'(0) = -1/\Gamma(1+\alpha^{-1})$ is finite, $\psi''(0) = \infty$ and $\psi(s) \sim 1 - s/\Gamma(1+\alpha^{-1})$ as $s \to 0^+$. We can write $\psi(s) = 1 + \psi'(0)s + o(s), s \to 0^+$. Let $g(s) = \psi'(s) - \psi'(0) = (1 - \exp\{-s^\alpha\})/\Gamma(1+\alpha^{-1}) \sim s^\alpha/\Gamma(1+\alpha^{-1}), s \to 0^+$. Then clearly, $g(s) \in \mathrm{RV}_\alpha(0^+)$ and can be written as $g(s) = s^\alpha \ell(s)$ with $\ell(s) \to 1/\Gamma(1+\alpha^{-1})$ as $s \to 0^+$. So $g(s) = as^\alpha + o(s^\alpha), s \to 0^+$, with $a = 1/\Gamma(1+\alpha^{-1}) > 0$. By Proposition 3.6, the copula C_ψ has intermediate upper tail dependence when $0 < \alpha < 1$. Also, $\kappa_U = 1 + \alpha$ and

$$\lim_{u \to 0^+} \frac{\overline{C}_\psi(1-u, 1-u)}{u^{1+\alpha}} = \frac{2[\Gamma(1+\alpha^{-1})]^\alpha (2^\alpha - 1)}{1+\alpha}.$$

It can be shown numerically that the d-variate Archimedean copula with this 1-parameter LT family is decreasing in concordance as α increases. As $\alpha \to 1^-$, numerically, the limit is close to the independence copula; as $\alpha \to 0^+$, the limit is close to the comonotonic copula.

In the next proposition, we state a result for the upper tail order function of Archimedean copulas.

Proposition 3.7 *Let C_ψ be a multivariate Archimedean copula with $\kappa_U = M_\psi$ being a non-integer in the interval $(1, d)$, and suppose ψ satisfies the condition of Proposition 3.5. With the notation $M = M_\psi - [M_\psi]$ and $k = [M_\psi]$, the upper tail order parameter is*

$$\lambda_U(C_\psi) = \frac{Mh}{[-\psi'(0)]^{M_\psi} \prod_{j=0}^{k}(M_\psi - j)} \sum_{\emptyset \neq I \subset I_d} (-1)^{|I|+k+1} |I|^{M_\psi},$$

and the upper tail order function is

$$b^*(\mathbf{w}) = \frac{\sum_{\emptyset \neq I \subset I_d} (-1)^{|I|} \left(\sum_{i \in I} w_i\right)^{M_\psi}}{\sum_{\emptyset \neq I \subset I_d} (-1)^{|I|} \left|I\right|^{M_\psi}},$$

where $h = \lim_{s \to 0^+} \ell(s)$ with $|\psi^{(k)}(s) - \psi^{(k)}(0)| = s^M \ell(s)$ as $s \to 0^+$.

Remark 3.6 For a d-variate Archimedean copula, the pattern of the upper tail order function also depends on the upper tail order κ. For example, in $d = 3$, the homogeneous function b^* is positively proportional to

$$-w_1^\kappa - w_2^\kappa - w_3^\kappa + (w_1+w_2)^\kappa + (w_1+w_3)^\kappa + (w_2+w_3)^\kappa - (w_1+w_2+w_3)^\kappa, \ 1 < \kappa < 2;$$

$$w_1^\kappa + w_2^\kappa + w_3^\kappa - (w_1+w_2)^\kappa - (w_1+w_3)^\kappa - (w_2+w_3)^\kappa + (w_1+w_2+w_3)^\kappa, \ 2 < \kappa < 3.$$

The signs of all terms depend on whether $1 < \kappa < 2$ or $2 < \kappa < 3$. The pattern of alternating signs extends to $d > 3$. This pattern, together with Lemma 3.12, also shows why we don't have a general form of the tail order function when M_ψ is a positive integer.

Recently, we have noticed that Larsson and Nešlehová (2011) uses survival copulas for a random vector $\mathbf{X} := (X_1, \ldots, X_d) \stackrel{d}{=} R\mathbf{S}_d$ as another representation for Archimedean copulas, where R is a radial random variable, the random vector $\mathbf{S}_d$ is uniformly distributed on the simplex $\{(x_1, \ldots, x_d) \in \mathbb{R}_+^d : \sum_{i=1}^d x_i = 1\}$, and R and $\mathbf{S}_d$ are independent. In parallel to elliptical copulas that are constructed based on random vectors that have a similar scale mixture representation, the authors find that the tail behavior of the radial random variable R will affect the tail dependence patterns of corresponding Archimedean copulas. This representation provides an alternative way to study the dependence in the tails of Archimedean copulas.

3.3.3 Lower tail

For intermediate lower tail dependence of Archimedean copulas, a general result has been obtained in Theorem 3.3 of Charpentier and Segers (2009). We will derive a more concrete and usable result that involves the slowly varying function ℓ, and give an interpretation in terms of the (resilience) random variable H which has LT ψ.

The condition below on the LT $\psi(s)$ as $s \to \infty$ covers almost all of the LT families in the Appendix of Joe (1997), as well as other LT families that can be obtained by integration or differentiation. Suppose

$$\psi(s) \sim T(s) = a_1 s^q \exp\{-a_2 s^r\}$$
$$\text{and } \psi'(s) \sim T'(s), \quad s \to \infty, \text{ with } a_1 > 0, a_2 \geq 0, \tag{3.10}$$

where $r = 0$ implies $a_2 = 0$ and $q < 0$, and $r > 0$ implies $r \leq 1$ and q can be 0, negative or positive. Note that $r > 1$ is not possible because of the complete monotonicity property of a LT.

The condition can be interpreted as follows. As $\psi(s)$ decreases to 0 more slowly as $s \to \infty$, then the random variable H with LT ψ has a heavier "tail" at 0. Let $z = \lim_{\eta \to 0} f_H(\eta) \in [0, \infty]$, where f_H is the density of H and is assumed well-behaved near 0. As z increases, then the "tail" at 0 is heavier. If $z = 0$, then the tail is lighter as the rate of decrease to 0 is faster. If $z = \infty$, then the tail is heavier as the rate of increase to ∞ is faster. In terms of the LT and the condition in (3.10), as r increases (with fixed q), the tail of H at 0 gets lighter, and as q increases (with fixed r), the tail of H at 0 gets heavier.

The next proposition shows that lower tail dependence behavior is influenced by r.

Proposition 3.8 *Suppose a LT ψ satisfies the condition in (3.10) with $0 \leq r \leq 1$. If $r = 0$, then C_ψ has lower tail dependence or lower tail order is 1. If $r = 1$, then $\kappa_L(C_\psi) = d$. If $0 < r < 1$, then C_ψ has intermediate lower tail dependence with $1 < \kappa_L(C_\psi) = d^r < d$, $\ell(u) = d^q a_1^{1-\kappa} a_2^{-\zeta} (-\log u)^\zeta$ with $\zeta = (q/r)(1 - d^r)$, and the tail order function is $b(\mathbf{w}) = \prod_{i=1}^d w_i^{d^{r-1}}$.*

Remark 3.7 Condition (3.10) does not cover all possibilities. It is possible that as $s \to \infty$, $\psi(s)$ goes to 0 slower than anything of form (3.10). Examples are given by LT families LTF and LTG in Joe (1997), leading to Archimedean families such that $\lim_{u \to 0^+} C_\psi(u\mathbf{1}_d)/u = 1$ (for the bivariate case, see families BB2 and BB3 in Joe and Hu (1996), and Joe (1997)). Note that, for LTF, $\psi(s) = [1 + \delta^{-1}\log(1+s)]^{-1/\theta}$ with $\delta > 0$ and $\theta > 0$ and as $s \to \infty$, $\psi(s) \sim \delta^{1/\theta}(\log s)^{-1/\theta}$; for LTG, $\psi(s) = \exp\{-[\delta^{-1}\log(1+s)]^{1/\theta}\}$ with $\delta > 0$, $\theta > 1$ and as $s \to \infty$, $\psi(s) \sim \exp\{-\delta^{-1/\theta}(\log s)^{1/\theta}\}$. We refer to Chapters 4 and 5 for relevant research on LTF and LTG.

Remark 3.8 Consider the pair (ϕ, ψ) of LTs where (a) $\phi'(0)$ is finite and $\psi(s) = \phi'(s)/\phi'(0)$ or (b) $\int_0^\infty \psi(v)dv$ is finite and $\phi(s) = \int_s^\infty \psi(v)dv / \int_0^\infty \psi(v)dv$. For the upper tail, we get $M_\psi = M_\phi - 1$ so that LT ψ has a heavier tail and $\kappa_U(C_\psi)$ is smaller (stronger intermediate tail dependence) if $\kappa_U(C_\phi) < d$. Proposition 3.8 implies that $\kappa_L(C_\psi) = \kappa_L(C_\phi)$. But the second level of tail dependence strength comes from the slowly varying function $\ell(u) = d^q a_1^{1-\kappa} a_2^{-\zeta}(-\log u)^\zeta$. Since $C_\psi(u\mathbf{1}_d) \sim u^\kappa \ell(u)$, $u \to 0^+$, a smaller κ means stronger intermediate lower tail dependence at the first level, and a faster $\ell(u) \to +\infty$ or a slower $\ell(u) \to 0^+$ means stronger intermediate lower tail dependence at the second level. For the LT tail, $a_1 s^q \exp(-a_2 s^r)$, a smaller r means slower decrease to 0 as $s \to +\infty$ and the resilience random variable has more probability near 0 and C_ψ has more dependence in the lower tail. This can be shown by a smaller tail order d^r. A larger q means slower decrease to 0 as $s \to +\infty$, which also implies more lower tail dependence. This is seen from a faster increase of $\ell(u) \to +\infty$ as $u \to 0^+$ when $q < 0$ and increases, or a slower decrease of $\ell(u) \to 0^+$ as $u \to 0^+$ when $q > 0$ and increases. Note that when u is small enough, $(-\log u)^\zeta$ dominates $\ell(u)$.

3.3.4 A new parametric Archimedean copula

By applying the LT of the inverse Gamma distribution, we present a new one-parameter Archimedean copula that exhibits intermediate upper and lower tail dependence, and have essentially a full range of positive dependence from independence to comonotonicity.

Example 3.7 (Archimedean Copula based on the Laplace Transform of Inverse Gamma (ACIG)) Let $Y = X^{-1}$ have the inverse Gamma (IΓ) distribution, where $X \sim \text{Gamma}(\alpha, 1)$ for $\alpha > 0$. Then it is straightforward to derive that $M_Y = \alpha$. The LT of the inverse Gamma distribution:

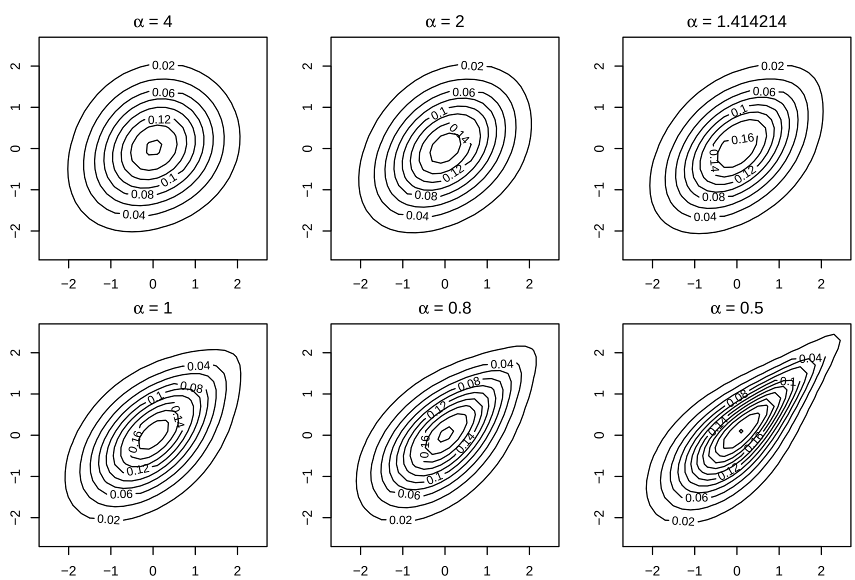
$$\psi(s;\alpha) = \frac{2}{\Gamma(\alpha)} s^{\alpha/2} K_\alpha(2\sqrt{s}), \quad s \geq 0,\ \alpha > 0, \tag{3.11}$$

where K_α is the modified Bessel function of the second kind. (Please see Section 3.6 for the derivation of (3.11).) It can be shown numerically that the d-variate Archimedean copula with this 1-parameter LT family is decreasing in concordance as α increases, with limits of the independence copula as $\alpha \to \infty$ and the comonotonic copula as $\alpha \to 0$.

Proposition 3.9 *Let C_ψ be an Archimedean copula constructed by (3.11). If $\alpha \in (0, +\infty)$ is not an integer, then the upper tail order is $\max\{1, \min\{\alpha, d\}\}$. The lower tail order is $\sqrt{d}$.*

Remark 3.9 For the bivariate case, $\kappa_U = \max\{1, \min\{\alpha, 2\}\}$ and $\kappa_L = \sqrt{2}$. Hence there is reflection asymmetry with skewness to the upper tail for $0 < \alpha < \sqrt{2}$ and skewness to the lower tail for $\alpha > \sqrt{2}$. Contour plots of this new Archimedean copula are given in Figure 3.1.

Figure 3.1: Contour plots: ACIG copula + standard normal margins



To conclude this subsection, we list the tail orders for some Archimedean copulas that interpolate independence and comonotonicity in Table 3.1. A variety of tail behavior obtains from known parametric Archimedean families and the new Archimedean family. Note that the bivariate Frank copula is reflection symmetric. But for d-dimensional Frank copula with $d \geq 3$, it can be shown numerically that $\overline{C}_\psi(\frac{1}{2}\mathbf{1}_d) > C_\psi(\frac{1}{2}\mathbf{1}_d)$ for parameters $\theta > 0$, although the lower tail order and upper tail order are the same. Some of the results in this table can be found in Charpentier and Segers (2009) and Heffernan (2000). For all of the examples in Table 3.1, the upper and lower tail orders decrease or remain constant as the dependence parameter(s) leads to increased dependence/concordance.

Table 3.1: Tail order of some Archimedean copulas that interpolate independence and comonotonicity

Copula/LT family	κ_L	κ_U
Frank; log-series LT $-\theta^{-1}\log[1-(1-e^{-\theta})e^{-s}]$ $(\theta > 0)$	d	d
[1]MTCJ; gamma LT $(1+s)^{-1/\theta}$ $(\theta > 0)$	1	d
Joe; Sibuya LT $1-(1-e^{-s})^{1/\theta}$ $(\theta > 1)$	d	1
Gumbel; positive stable LT $\exp\{-s^{1/\theta}\}$ $(\theta > 1)$	$d^{1/\theta}$	1
Joe-Hu; BB1 extension; $(1+s^{1/\delta})^{-1/\theta}$ $(\theta > 0, \delta > 1)$	1	1
Joe-Hu; BB7 extension; $1-[1-(1+s)^{-1/\delta}]^{1/\theta}$ $(\theta > 1, \delta > 0)$	1	1
Crowder; BB9 extension; $\exp\{-(\alpha^\theta + s)^{1/\theta} + \alpha\}$ $(\theta > 1)$	$d^{1/\theta}$	d
Joe-Ma; $\int_s^\infty e^{-v^\alpha} dv/\Gamma(1+\alpha^{-1})$ $(0 < \alpha < 1)$	d^α	$1+\alpha$
ACIG; LT of inverse gamma $2\Gamma^{-1}(\alpha)s^{\alpha/2}K_\alpha(2\sqrt{s})$ $(\alpha > 0)$	$d^{1/2}$	$(d \wedge \alpha) \vee 1$

1. Mardia-Takahasi-Cook-Johnson, (see Cook and Johnson, 1981).

3.4 Intermediate tail dependence: Mixture of max-id copulas

As an extension of Archimedean copulas, we study in this section the tail orders for copulas that are constructed with mixtures of max-id copulas. Some results studied in Joe (1997) are extended to intermediate tail dependence. Let F be a d-variate cdf. If F^t is also a cdf function for all $t > 0$, then F is max-id (Joe and Hu, 1996). The class of copulas based on mixture of max-id distributions has led to interesting classes of bivariate two-parameter copula families with both upper and lower tail dependence (e.g., labeled as BB1, BB4, BB7 in Joe (1997)). As well, other forms of intermediate tail dependence behavior are possible. These types of copulas will give us more flexibility in choices of bivariate linking copulas in vines (Aas et al., 2009; Joe et al., 2010).

Here we generalize Theorems 4.13 and 4.16 in Joe (1997) to multivariate versions and intermediate tail dependence. In the earlier research on copulas, the analyses determined when tail dependence (tail order $\kappa = 1$) can occur for different copula families; in that setting, the tail order occurred within the sufficient condition in Theorem 4.16 of Joe (1997). Let K be a multivariate max-id copula and ψ be a LT of a positive random variable, and consider the copulas that are of the following form

$$C(u_1, \dots, u_d) = \psi\Big(-\log K\big(e^{-\psi^{-1}(u_1)}, \dots, e^{-\psi^{-1}(u_d)}\big)\Big) \tag{3.12}$$

Proposition 3.10 *Suppose that a copula C be constructed by (3.12).*

1. *If ψ satisfies the condition of Proposition 3.5 with some $k \in \{1, \dots, d-1\}$ and $\kappa_U(K_I) > 1$ for any marginal copula K_I, then C has upper intermediate tail dependence and $\kappa_U(C) = \kappa_U(C_\psi)$.*

2. *If $1 - \psi(s) \in \mathrm{RV}_\beta(0^+)$, $\kappa_U(K) = 1$ with marginal copula $K_I(u\mathbf{1}_{|I|}) \sim u\ell_I(u), u \to 0^+$ such that $\lim_{u\to 0^+} \ell_I(u) = h_I \in (0,1]$, $0 < h_I^* = \sum_{\emptyset \neq J \subset I} (-1)^{|J|-1} h_J \leq 1$ and $0 < \sum_{\emptyset \neq I \subset I_d} (-1)^{|I|-1} (h_I^*)^\beta \leq 1$, then $\kappa_U(C) = 1$ with $\lambda_U(C) = \sum_{\emptyset \neq I \subset I_d} (-1)^{|I|-1} (h_I^*)^\beta$.*

Proposition 3.11 *Suppose that a copula C be constructed by (3.12) with $1 \leq \alpha = \kappa_L(K) \leq d$. If $-\psi(s)/\psi'(s) \in \mathrm{RV}_\beta$ with $0 < \beta \leq 1$, and $1 < \alpha^{1-\beta} < d$, then the copula C has lower intermediate tail dependence $\kappa_L(C) = \alpha^{1-\beta}$, and $\kappa_L(C) = \xi(\alpha, \beta) \cdot \kappa_L(C_\psi)$ with $\xi(\alpha, \beta) = (\alpha/d)^{1-\beta} \in (0, 1]$. Also, $\kappa_L(K) = 1$ implies that $\kappa_L(C) = 1$.*

Remark 3.10 Note that $\kappa_L(C)$ is less than or equal to both $\kappa_L(K)$ and $\kappa_L(C_\psi)$. K can be the independence copula or have intermediate lower tail dependence. The lower tail order of the copula C is increasing in $\kappa_L(K)$. One consequence of Propositions 3.10 and 3.11 is that if $\kappa_U(K) = d$ and $\kappa_L(K) = d$ then $\kappa_U(C) = \kappa_U(C_\psi)$ and $\kappa_L(C) = \kappa_L(C_\psi)$. Hence if K is chosen as the parametric Frank copula family with parameter $\theta \geq 0$, then $C(u_1, ..., u_d; \theta)$ as given in (3.12) will be increasing in concordance as θ increases. The parameter θ affects dependence only, while the LT ψ controls the upper and lower tail orders.

When we take K as the independence copula or the Frank copula with positive dependence, and the LT has tail of the form $\psi(s) \sim a_1 s^q \exp\{-a_2 s^r\}$, $s \to \infty$, $0 \le r < 1$, where a_1, a_2 are some positive constants, then we can construct a new family of Archimedean copulas that satisfies the condition of Proposition 3.11.

In dimensions $d \ge 3$, Archimedean and mixture of max-id copula families cannot achieve the range of dependence available from vine copulas (Bedford and Cooke, 2001; Aas et al., 2009; Joe et al., 2010). But for $d = 2$, the mixture of max-id approach can lead to more candidates, with a variety of upper and lower tail behavior, to be used as bivariate linking copulas in vines. For instance, from Table 3.1, the preceding subsections and propositions, the Joe-Ma ψ function, which is the normalized integral of the positive stable LT, combined with the bivariate Gaussian copula with $\varrho \ge 0$ can lead to a two-parameter family with more flexible upper and lower tail orders. Note that, the bivariate Gaussian density is Totally Positive of order 2 if $\varrho \ge 0$, and hence max-id (Joe, 1997, Theorem 2.6).

3.5 Discussion

We have shown how the concept of tail order is useful to quantify the strength of dependence in the upper and lower tails, as well as the direction of reflection asymmetry. One- and two-parameter families that are Archimedean copulas and mixture of max-id copulas together can cover a wide range of tail orders. The interpretation through the latent resilience variable shows why Archimedean copulas can obtain a full range of tail orders by varying the density of the resilience at 0 and ∞. In order to get our results for Archimedean copulas, we needed Proposition 3.5 which, on its own, contributes knowledge about LTs.

Archimedean copulas only have exchangeable dependence but their bivariate versions can be used within vines. Vine copulas (Bedford and Cooke, 2001; Aas et al., 2009; Joe et al., 2010) in dimension d, which include multivariate Gaussian and t copulas as special cases, are built from $d(d-1)/2$ bivariate linking copulas, of which $d-1$ are bivariate marginal copulas and the remainder are conditional bivariate copulas with the number of conditioning variables between 1 to $d-2$. By choosing bivariate linking copulas with flexible tail orders and reflecting symmetry/asymmetry, we can get vine copulas to cover a wide range of tail behavior, as well as dependence structures. The study of tail orders of vine copulas in terms of the tail orders of the bivariate linking copulas will be studied in future research. For vine copulas, we are also interested in conditions that retain consistent relation of upper and lower tail orders for all margins.

3.6 Proofs

Derivation of LT of the inverse Gamma distribution: With $Y = X^{-1}$ and $X \sim \mathrm{Gamma}(\alpha, 1)$, the LT is derived as

$$\psi(s) = \psi(s;\alpha) = \mathbb{E}(e^{-sY}) = \mathbb{E}(e^{-s/X}) = [\Gamma(\alpha)]^{-1} \int_0^\infty e^{-s/x} x^{\alpha-1} e^{-x} dx.$$

From the GIG(ν, χ, φ) density (McNeil et al., 2005),

$$\int_0^\infty w^{\nu-1} \exp\{-\tfrac{1}{2}(\chi w^{-1} + \varphi w)\}\, dw = 2(\chi/\varphi)^{\nu/2}\, K_\nu(\sqrt{\chi\varphi}\,).$$

Note that $K_\nu = K_{-\nu}$. Hence with $\chi = 2s$, $\varphi = 2$, $\nu = \alpha$

$$\psi(s;\alpha) = 2\Gamma^{-1}(\alpha)(2s/2)^{\alpha/2} K_\alpha(\sqrt{2s \cdot 2}\,) = 2\Gamma^{-1}(\alpha) s^{\alpha/2} K_\alpha(2\sqrt{s}\,).$$

$\square$

Proof of Proposition 3.1: Letting $r := F^{-1}(1-u)$ and $b_\varrho = \sqrt{2/(1+\varrho)}$, then by Example 6.2 (i) of Hashorva (2007), as $u \to 0^+$ and thus $r \to \infty$,

$$\begin{aligned}
&\overline{C}(1-u, 1-u) \\
&= \mathbb{P}\left[X_1 > F^{-1}(1-u), X_2 > F^{-1}(1-u)\right] = \mathbb{P}[X_1 > r, X_2 > r] \\
&= (1+o(1)) \frac{(1-\varrho^2)^{3/2}}{2\pi(1-\varrho)^2} [a(b_\varrho r)/r][1 - F_R(b_\varrho r)], \qquad (3.13)
\end{aligned}$$

where F_R is the cdf of R and $a(\cdot)$ is an auxiliary function of R with respect to the Gumbel Maximum Domain of Attraction in the sense of (2.10). As $u \to 0^+$, i.e., $r \to \infty$, both $a(b_\varrho r)/r \to 0$ (see Theorems 3.3.26 and 3.3.27 of Embrechts et al. (1997)) and $1 - F_R(b_\varrho r) \to 0$. Let $G(x) := 1/[1 - F_R(x)]$, then $G : \mathbb{R} \to \mathbb{R}_+$ is increasing and the condition of (2.10) is equivalent to that $G \in \Gamma$-varying with auxiliary function $a(\cdot)$ (de Haan, 1970, Definition 1.5.1). The inverse function of a Γ-varying function is a Π-varying function (de Haan, 1974, Corollary 1.10). Therefore, $G^{-1} \in \Pi$-varying. Assuming that an auxiliary function of G^{-1} is $a_0(\cdot)$, by Lemma 1.2.9 of de Haan and Ferreira (2006), the auxiliary function $a_0(\cdot)$ of the Π-varying function G^{-1} is slowly varying at ∞. Moreover, $a_0(t) = a(G^{-1}(t))$ (de Haan, 1974, Corollary 1.10). So, $a(x) = a_0(G(x))$. Then in (3.13),

$$a(b_\varrho r)/r = a_0(G(b_\varrho r))/r = a_0(1/[1 - F_R(b_\varrho r)])/r,$$

while $1 - F_R(b_\varrho r)$ is rapidly varying in r at ∞ due to the fact that G is Γ-varying and any Γ-varying function is rapidly varying (de Haan, 1970, Theorem 1.5.1). Therefore,

$$1 - F_R(b_\varrho r) = 1 - F_R\left(\sqrt{2/(1+\varrho)}F^{-1}(1-u)\right)$$

dominates the tail behavior of (3.13) as $u \to 0$, and thus determines the corresponding tail order of the elliptical copula. By the definition of tail order in Definition 3.1, we may also obtain the upper tail order by the following

$$\kappa = \lim_{u \to 0^+} \frac{\log \overline{C}(1-u, 1-u)}{\log u} = \lim_{r \to \infty} \frac{\log\left(1 - F_R(b_\varrho r)\right)}{\log\left(1 - F(r)\right)}.$$

By Example 6.2 (iii) of Hashorva (2007), as $r \to \infty$,

$$\mathbb{P}[X_1 > r] = (1 + o(1))(2\pi)^{-1/2}[a(r)/r]^{1/2}[1 - F_R(r)].$$

Due to the similar argument as before, $1 - F_R(r)$ dominates the tail behavior of $\mathbb{P}[X_1 > r]$, as $r \to \infty$. Therefore, we may write

$$\kappa = \lim_{r \to \infty} \frac{\log\left(1 - F_R(b_\varrho r)\right)}{\log\left(1 - F_R(r)\right)},$$

which completes the proof. □

Proof of Proposition 3.3: Assuming $C(u\mathbf{1}_d) \sim u^{\kappa_L(C)}\ell(u)$, $u \to 0^+$, with $\ell(u) \in \mathrm{RV}_0(0^+)$, for any copula C and $0 \le u \le 1$, $C(u\mathbf{1}_d) \le u$. Therefore, $\kappa_L(C) \ge 1$.

To prove (1), by the condition of PLOD, we have $C(u\mathbf{1}_d) \geq u^d$ for any $0 \leq u \leq 1$ and thus, $\kappa_L(C) \leq d$.

To prove (2), choosing $S_1 \subset S_2$ with $|S_2| - |S_1| = j \in \mathbb{N}_+$. Let us consider the case where $j = 1$ first, for some $l \in \{1, \ldots, |S_2|\}$ and any $0 \leq u \leq 1$,

$$\begin{aligned}
C_{S_2}(u\mathbf{1}_{|S_2|}) &= \mathbb{P}[U_l \leq u | U_1 \leq u, \ldots, U_{l-1} \leq u, U_{l+1} \leq u, \ldots, U_{|S_2|} \leq u] \\
&\quad \times \mathbb{P}[U_1 \leq u, \ldots, U_{l-1} \leq u, U_{l+1} \leq u, \ldots, U_{|S_2|} \leq u] \\
&\geq \mathbb{P}[U_l \leq u] \times \mathbb{P}[U_1 \leq u, \ldots, U_{l-1} \leq u, U_{l+1} \leq u, \ldots, U_{|S_2|} \leq u] \\
&= uC_{S_1}(u\mathbf{1}_{|S_1|}).
\end{aligned}$$

The inequality is due to the MLTD of C. Clearly, $\kappa_L(C_{S_2}) - \kappa_L(C_{S_1}) \leq 1$. Since $\mathbb{P}[U_l \leq u | U_1 \leq u, \ldots, U_{l-1} \leq u, U_{l+1} \leq u, \ldots, U_{|S_2|} \leq u] \leq 1$, $C_{S_2}(u\mathbf{1}_{|S_2|}) \leq C_{S_1}(u\mathbf{1}_{|S_1|})$ and thus, $\kappa_L(C_{S_2}) - \kappa_L(C_{S_1}) \geq 0$. An iterated argument will prove the case for a general j: $0 \leq \kappa_L(C_{S_2}) - \kappa_L(C_{S_1}) \leq |S_2| - |S_1|$. If $\kappa_L(C) = 1$, then for any $S \subset I_d$ with $|S| \geq 2$, we have $1 \leq \kappa_L(C_S) \leq \kappa_L(C) = 1$, which completes the proof. For $\kappa_L(C) = 1$, note that the MLTD condition is not needed. □

Proof of Lemma 3.4: Let Z be an exponential random variable, independent of Y, with mean 1. Choose any fixed m with $0 < m < 1$. Then $\mathbb{E}(Z^{-m}) = \Gamma(1-m)$, and if we define $W_m = (Y/Z)^m$, then for any $w > 0$,

$$\begin{aligned}
\mathbb{P}\left[W_m \geq w\right] &= \mathbb{P}[Z \leq Yw^{-1/m}] \\
&= \int_0^\infty \left(1 - \exp\{-yw^{-1/m}\}\right) F_Y(dy) = 1 - \psi(w^{-1/m}),
\end{aligned}$$

where F_Y is the cdf of Y. Therefore, $\mathbb{E}(Y^m) < \infty$ implies $\mathbb{E}(W_m) < \infty$ and $\lim_{w\to\infty} w[1 - \psi(w^{-1/m})] = 0$, i.e.,

$$\lim_{s\to 0^+} [1 - \psi(s)]/s^m = 0. \tag{3.14}$$

If $1 - \psi(s)$ is regularly varying at 0^+, then we can write $1 - \psi(s) = s^\alpha \ell(s)$ with $\alpha \neq 0$, where $\ell(s) \in \mathrm{RV}_0(0^+)$. Then, (3.14) implies that $\lim_{s\to 0^+} s^{\alpha - m}\ell(s) = 0$. Let $\epsilon > 0$ be arbitrarily small. If $m = M_Y - \epsilon$, then we have $\mathbb{E}(Y^{M_Y - \epsilon}) < \infty$ and thus $\lim_{s\to 0^+} s^{\alpha - M_Y + \epsilon}\ell(s) = 0$. Therefore, $\alpha \geq M_Y - \epsilon$.

Also by a result on page 49 of Chung (1974), (3.14) implies that for any $0 < \delta < 1$, $\mathbb{E}\big(Y^{m(1-\delta)}\big) < \infty$. If we assume that there exists an $\epsilon > 0$ with $m = M_Y + \epsilon$ such that, $\lim_{s\to 0^+} s^{\alpha - M_Y - \epsilon}\ell(s) = 0$, then for any small $\delta > 0$, $\mathbb{E}\big(Y^{(M_Y+\epsilon)(1-\delta)}\big) < \infty$. Then we may choose some $\delta_\epsilon < \epsilon/(\epsilon + M_Y)$, and get $\mathbb{E}\big(Y^{(M_Y+\epsilon)(1-\delta_\epsilon)}\big) < \infty$ with $(M_Y + \epsilon)(1 - \delta_\epsilon) > M_Y$, which gives rise to a contradiction. Thus, for any $\epsilon > 0$, we must have $\lim_{s\to 0^+} s^{\alpha - M_Y - \epsilon}\ell(s) \neq 0$, and hence, $\alpha - M_Y - \epsilon \leq 0$. So,

$$M_Y - \epsilon \leq \alpha \leq M_Y + \epsilon,$$

which completes the proof. $\square$

Proof of Proposition 3.5: This proof extends that in Lemma 3.4, which corresponds to the case where $k = 0$. For a positive integer j, let $Z_j \sim \mathrm{Gamma}(j+1, 1)$ so that $\mathbb{E}(Z_j^{-m}) = \Gamma(j+1-m)/\Gamma(j+1)$ if $0 < m < j+1$. Let $W_{m,j} = (Y/Z_j)^m$ where Y is independent of Z_j. Then for $0 < m < j+1$, $\mathbb{E}(W_{m,j}) < \infty$ if and only if $\mathbb{E}(Y^m) < \infty$. Next, similar to the proof of Lemma 3.4, if Y has LT ψ and moments up to order k, for $j \in \{0, 1, \ldots, k\}$ and $0 < m < j+1$,

$$\begin{aligned}
&\Pr\big[W_{m,j} = (Y/Z_j)^m \geq w\big] = \Pr(Z_j \leq Y w^{-1/m}) = \int_0^\infty F_{Z_j}(y w^{-1/m})\, dF_Y(y) \\
&= \int_0^\infty \Big[1 - \sum_{i=0}^{j} \frac{y^i w^{-i/m}}{i!} \exp\{-y w^{-1/m}\}\Big]\, dF_Y(y) \\
&= 1 - \sum_{i=0}^{j} \frac{w^{-i/m}}{i!} (-1)^i \psi^{(i)}(w^{-1/m}).
\end{aligned}$$

Suppose $0 < m < \min\{j+1, M_Y\}$. Then $\mathbb{E}(Y^m) < \infty$ implies that

$$w\Big[1 - \sum_{i=0}^{j} \frac{w^{-i/m}}{i!} (-1)^i \psi^{(i)}(w^{-1/m})\Big] \to 0, \quad w \to \infty,$$

i.e.,

$$s^{-m}\Big[1-\sum_{i=0}^{j}\frac{s^i}{i!}(-1)^i\psi^{(i)}(s)\Big]\to 0,\quad s\to 0^+. \tag{3.15}$$

Assuming ψ has derivatives at zero up to kth order, then for positive integer $j\le k$, the main term in (3.15) is

$$\begin{aligned}
&1-\sum_{i=0}^{j}\frac{s^i}{i!}(-1)^i\psi^{(i)}(s)\\
&=1-\sum_{i=0}^{j-1}\frac{s^i}{i!}(-1)^i\Big[\sum_{l=0}^{j-i}s^l\psi^{(i+l)}(0)/l!+o(s^{j-i})\Big]-\frac{s^j}{j!}(-1)^j\psi^{(j)}(s)\\
&=1-\sum_{i=0}^{j-1}\frac{s^i}{i!}(-1)^i\sum_{l=i}^{j}s^{l-i}\psi^{(l)}(0)/(l-i)!-\frac{s^j}{j!}(-1)^j\psi^{(j)}(s)+o(s^j)\\
&=1-\sum_{l=0}^{j}\psi^{(l)}(0)\frac{s^l}{l!}\Big\{\sum_{i=0}^{l\wedge(j-1)}(-1)^i\frac{l!}{i!(l-i)!}\Big\}-\frac{s^j}{j!}(-1)^j\psi^{(j)}(s)+o(s^j)\\
&=1-\psi(0)-\psi^{(j)}(0)\frac{s^j}{j!}\big[-(-1)^j\big]-\frac{s^j}{j!}(-1)^j\psi^{(j)}(s)+o(s^j)\\
&=(-1)^{j-1}\frac{s^j}{j!}\big[\psi^{(j)}(s)-\psi^{(j)}(0)\big]+o(s^j).
\end{aligned} \tag{3.16}$$

Hence (3.15) implies that

$$s^{j-m}\big[\psi^{(j)}(s)-\psi^{(j)}(0)\big]\to 0,\quad s\to 0^+,$$

if j is a non-negative integer less than M_Y and $m<M_Y$. In particular, if k is a non-negative integer such that $k<m<M_Y<k+1$, then

$$s^{k-m}\big[\psi^{(k)}(s)-\psi^{(k)}(0)\big]\to 0,\quad s\to 0^+.$$

If $\big|\psi^{(k)}(0)-\psi^{(k)}(s)\big|$ is regularly varying at 0^+, we write $\big|\psi^{(k)}(0)-\psi^{(k)}(s)\big| = s^{\alpha}\ell(s)$. For any $\epsilon>0$, a similar argument in the proof of Lemma 3.4 will prove that $\alpha\ge M_Y-k-\epsilon$. Now we prove the other direction. We assume that there exists an $\epsilon > 0$ with $m = M_Y + \epsilon$ such that,

$$\lim_{s \to 0^+} s^{\alpha+k-M_Y-\epsilon}\ell(s) = 0, \tag{3.17}$$

that is, $s^{k-M_Y-\epsilon}\big[\psi^{(k)}(s) - \psi^{(k)}(0)\big] \to 0$ as $s \to 0$. Since ψ is completely monotonic, $\psi^{(k)}(0) - \psi^{(k)}(s)$ is either negative or positive as $s \to 0^+$. That is, $(-1)^k[\psi^{(k)}(0) - \psi^{(k)}(s)] > 0$. The following argument is for an even k, and similar when k is odd. Then by the Karamata's theorem (Resnick, 2007), regular variation of $\big|\psi^{(k)}(0) - \psi^{(k)}(s)\big|$ implies that

$$\begin{aligned}
&-\psi^{(k-1)}(x) + \psi^{(k-1)}(0) + x\psi^{(k)}(0) \\
&= \int_0^x \big[\psi^{(k)}(0) - \psi^{(k)}(s)\big]ds \sim (\alpha+1)^{-1}x^{\alpha+1}\ell(x), \quad x \to 0^+.
\end{aligned} \tag{3.18}$$

Since $\int_0^x[\psi^{(k)}(0) - \psi^{(k)}(s)]ds$ is again regularly varying, we can take the integration on both sides repeatedly and obtain for $j = 0, 1, \ldots, k$,

$$-\psi^{(k-j)}(x) + \sum_{i=0}^{j} \frac{x^{j-i}}{(j-i)!}\psi^{(k-i)}(0) \sim \left(\alpha \prod_{i=0}^{j} \frac{1}{\alpha+i}\right) x^{\alpha+j}\ell(x), \quad x \to 0^+. \tag{3.19}$$

Multiplying both sides of (3.19) by $\frac{x^{k-j}}{(k-j)!}(-1)^{k-j}$ leads to

$$\begin{aligned}
LHS_j =: &-\frac{x^{k-j}}{(k-j)!}(-1)^{k-j}\psi^{(k-j)}(x) + \sum_{i=0}^{j} \frac{(-1)^{k-j}}{(k-j)!(j-i)!}x^{k-i}\psi^{(k-i)}(0) \\
&\sim \frac{(-1)^{k-j}}{(k-j)!}\left(\alpha \prod_{i=0}^{j} \frac{1}{\alpha+i}\right) x^{k+\alpha}\ell(x), \quad x \to 0^+.
\end{aligned} \tag{3.20}$$

Then we add the left-hand side of (3.20) for $j = 0, \ldots, k$, and after rearranging the summand, we have

$$\sum_{j=0}^{k} LHS_j = -\sum_{i=0}^{k} \frac{x^i}{i!}(-1)^i\psi^{(i)}(x) + 1 + \sum_{i=0}^{k-1}\sum_{j=i}^{k} \frac{(-1)^{k-j}}{(k-j)!(j-i)!}x^{k-i}\psi^{(k-i)}(0). \tag{3.21}$$

By the binomial theorem, for each given $i \in (0, \dots, k-1)$, $\sum_{j=i}^{k} \frac{(-1)^{k-j}}{(k-j)!(j-i)!} \equiv 0$. Then, from (3.20) and (3.21) we can conclude that

$$\sum_{j=0}^{k} LHS_j = 1 - \sum_{i=0}^{k} \frac{x^i}{i!}(-1)^i \psi^{(i)}(x) = O\big(x^{k+\alpha}\ell(x)\big).$$

Therefore, multiplying both sides of the above by $s^{-M_Y-\epsilon}$ and using (3.17),

$$s^{-M_Y-\epsilon}\Big[1 - \sum_{i=0}^{k} \frac{s^i}{i!}(-1)^i \psi^{(i)}(s)\Big] \to 0, \quad s \to 0^+.$$

Then for any small $\delta > 0$, $\mathbb{E}\big(Y^{(M_Y+\epsilon)(1-\delta)}\big) < \infty$. Then we may choose some $\delta_\epsilon < \epsilon/(\epsilon + M_Y)$, and get $\mathbb{E}\big(Y^{(M_Y+\epsilon)(1-\delta_\epsilon)}\big) < \infty$ with $(M_Y + \epsilon)(1-\delta_\epsilon) > M_Y$, which gives rise to a contradiction to the fact that M_Y is the maximal moment. Thus, for any $\epsilon > 0$, we must have $\lim_{s\to 0^+} s^{\alpha+k-M_Y-\epsilon}\ell(s) \neq 0$, and hence, $\alpha \le M_Y - k + \epsilon$. Thus, $\alpha = M_Y - k$.

To prove the last statement of the proposition, since

$$1-\phi(s) = 1-\psi^{(k)}(s)/\psi^{(k)}(0) = \big[\psi^{(k)}(0)-\psi^{(k)}(s)\big]/\psi^{(k)}(0) \in \mathrm{RV}_{M_Y-k}(0^+),$$

by Remark 3.4, we have

$$\phi(s) = \psi^{(k)}(s)/\psi^{(k)}(0) = 1 - h'_{k+1}s^{M_Y-k} + o(s^{M_Y-k}).$$

Then, by integration, we will have

$$\begin{aligned}\psi(s) &= 1 + \psi^{(1)}(0)s + \tfrac{1}{2}\psi^{(2)}(0)s^2 + \cdots + (-1)^{k+1}h_{k+1}s^{k+M_Y-k} + o(s^{k+M_Y-k}) \\ &= 1 - h_1 s + h_2 s^2 - \cdots + (-1)^{k+1}h_{k+1}s^{M_Y} + o(s^{M_Y}) \quad s \to 0^+,\end{aligned}$$

where $0 < h_i < \infty$. The integration is due to Lemma 31 of Breitung (1994). $\square$

Proof of Proposition 3.6: We provide the proof only for the bivariate case. For $d \ge 3$, the intermediate upper tail dependence can be studied analogously, and the (omitted) proof is similar but with more complicated notation.

Let $\psi'(0) = m$ with $-\infty < m < 0$, then by Proposition 3.5, as $s \to 0^+$, letting $M = M_\psi - 1$,

$$g(s) = \psi'(s) - m = as^M + o(s^M), \quad (0 < a < \infty; 0 < M < 1).$$

Since $g'(s) = \psi''(s)$ is increasing as $s \to 0^+$, if we write $g(s) \sim as^M \ell(s), s \to 0^+$, where $\ell(s) \in \mathrm{RV}_0(0^+)$ and $\ell(s) \to 1$ as $s \to 0^+$. Note that

$$g(s) = \psi'(s) - m = \int_0^s \psi''(x)dx = \int_0^s g'(x)dx.$$

By the Monotone Density Theorem (Bingham et al., 1987, Theorem 1.7.2),

$$\psi''(s) = g'(s) \sim aMs^{M-1}\ell(s), \quad s \to 0^+. \tag{3.22}$$

Thus, $\lim_{s\to 0^+} \psi''(2s)/\psi''(s) = 2^{M-1}$. Observe that for $0 < \zeta < 1$,

$$\begin{aligned}
&\lim_{u\to 1^-} \frac{\overline{C}_\psi(u,u)}{(1-u)^{1+\zeta}} \\
&= \lim_{u\to 1^-} \frac{1 - 2u + \psi(2\psi^{-1}(u))}{(1-u)^{1+\zeta}} = \lim_{u\to 1^-} \frac{-2 + 2\psi'(2\psi^{-1}(u))/\psi'(\psi^{-1}(u))}{-(1+\zeta)(1-u)^\zeta} \\
&= \lim_{u\to 1^-} \frac{4\psi''(2\psi^{-1}(u))/[\psi'(\psi^{-1}(u))]^2 - 2\psi''(\psi^{-1}(u))\psi'(2\psi^{-1}(u))/[\psi'(\psi^{-1}(u))]^3}{\zeta(1+\zeta)(1-u)^{\zeta-1}} \\
&= \lim_{s\to 0^+} \frac{4\psi''(2s)/[\psi'(s)]^2 - 2\psi''(s)\psi'(2s)/[\psi'(s)]^3}{\zeta(1+\zeta)(1-\psi(s))^{\zeta-1}} \quad \text{(letting } s = \psi^{-1}(u)\text{)} \\
&= \lim_{s\to 0^+} \frac{4m^{-2}\psi''(2s) - 2m^{-2}\psi''(s)}{\zeta(1+\zeta)(1-\psi(s))^{\zeta-1}} = \lim_{s\to 0^+} \frac{4m^{-2}\frac{\psi''(2s)}{\psi''(s)} - 2m^{-2}}{\zeta(1+\zeta)\frac{(1-\psi(s))^{\zeta-1}}{\psi''(s)}} \\
&= \lim_{s\to 0^+} \frac{2m^{-2}(2^M - 1)}{\zeta(1+\zeta)\frac{(1-\psi(s))^{\zeta-1}}{\psi''(s)}}.
\end{aligned}$$

By Proposition 3.5, there is a constant $h > 0$ such that

$$1 - \psi(s) = -ms - hs^{M+1} + o(s^{M+1}), \quad s \to 0^+.$$

Then, as $s \to 0^+$,

$$[1 - \psi(s)]^{\zeta-1} \sim (-m)^{\zeta-1} s^{\zeta-1}.$$

In addition, it has been shown that $\psi''(s) \sim aMs^{M-1}\ell(s)$ as $s \to 0^+$ and $2m^{-2}(2^M - 1)$ is finite. Hence, the intermediate tail dependence exists if and only if $\zeta = M$ and $\kappa_U = 1 + M = M_\psi$.

The proof for the case of $k = 0$ is similar, by applying Proposition 3.5. $\square$

Lemma 3.12 *Let $d \geq 2$ be a positive integer and let j be a positive integer that is less than d. Let*

$$S_{dj}(w_1, \ldots, w_d) = \sum_{\emptyset \neq I \subset I_d} (-1)^{|I|-1} \Big(\sum_{i \in I} w_i \Big)^j. \tag{3.23}$$

Then $S_{dj} \equiv 0$.

Proof of Lemma 3.12: When $j = 1$, by the binomial theorem, for any $n \in \mathbb{N}_+$, $\sum_{i=0}^n (-1)^i \binom{n}{i} = 0$, so that $S_{d1} \equiv 0$ for $d \geq 2$.

For $1 < j < d$, S_{dj} is a symmetric homogeneous function of order j, and its first order partial derivatives are homogeneous of order $j - 1$. By recursion with Euler's formula for homogeneous functions to the jth order partial derivatives

$$\begin{aligned} S_{dj}(\mathbf{w}) &= \frac{1}{j} \sum_{i=1}^d \frac{\partial S_{dj}(\mathbf{w})}{\partial w_i} w_i = \frac{1}{j(j-1)} \sum_{i_1=1}^d \sum_{i_2=1}^d \frac{\partial^2 S_{dj}(\mathbf{w})}{\partial w_{i_1} \partial w_{i_2}} w_{i_1} w_{i_2} \\ &= \frac{1}{j!} \sum_{i_1=1}^d \cdots \sum_{i_j=1}^d \frac{\partial^j S_{dj}(\mathbf{w})}{\partial w_{i_1} \cdots \partial w_{i_j}} w_{i_1} \cdots w_{i_j}. \end{aligned} \tag{3.24}$$

We will show that all the jth order partial derivatives are 0. Because of symmetry, we consider only terms for which $w_{i_1} \ldots w_{i_j} = w_1^{n_1} \cdots w_p^{n_p}$ where $1 \leq p \leq j < d$,

$n_1 > 0, \ldots, n_p > 0$ and $n_1 + \cdots + n_p = j$. Then

$$\begin{aligned}
\frac{\partial^j S_{dj}(\mathbf{w})}{\partial^{n_1} w_1 \cdots \partial^{n_p} w_p} &= j!(-1)^{p-1} + \sum_{\emptyset \neq J \subset \{p+1,\ldots,d\}} (-1)^{p+|J|-1} j! \\
&= j!(-1)^{p-1}\Big(1 + \sum_{\emptyset \neq J \subset \{p+1,\ldots,d\}} (-1)^{|J|}\Big) \\
&= j!(-1)^{p-1} \sum_{i=0}^{d-p} (-1)^i \binom{d-p}{i} = 0,
\end{aligned}$$

which completes the proof. $\square$

Note that (3.23) is not zero for $j = d$ because (3.24) would include a non-zero term such as $\partial^d S_{dj}/\partial w_1 \cdots \partial w_d = (-1)^{d-1} d!$. In fact, there are $d!$ non-zero terms in (3.24) when $(i_1, \ldots, i_d)$ is a permutation of $(1, \ldots, d)$, and $S_{dd}(w_1, \ldots, w_d) = (-1)^{d-1} d! \prod_{i=1}^{d} w_i$.

Proof of Proposition 3.7: Consider

$$\begin{aligned}
&\lim_{u \to 0^+} \frac{\mathbb{P}\left[\bigcap_{i \in I_d} \{U_i \geq 1 - uw_i\}\right]}{u^{k+M}} \\
&= \sum_{\emptyset \neq I \subset I_d} (-1)^{|I|-1} \lim_{u \to 0^+} \frac{\mathbb{P}\left[\bigcup_{i \in I} \{U_i \geq 1 - uw_i\}\right]}{u^{k+M}}.
\end{aligned}$$

By Proposition 3.5, since the function $w \mapsto 1 - \psi(w) \in \mathrm{RV}_1(0^+)$, then we have

$w \mapsto \psi^{-1}(1-w) \in \mathrm{RV}_1(0^+)$. Thus,

$$
\begin{aligned}
&\lim_{u\to 0^+} \frac{\mathbb{P}\left[\bigcup_{i\in I}\{U_i \geq 1-uw_i\}\right]}{u^{k+M}} \\
&= \lim_{u\to 0^+} \frac{1-\psi[\psi^{-1}(1-uw_1)+\cdots+\psi^{-1}(1-uw_d)]}{u^{k+M}} \\
&= \lim_{u\to 0^+} \frac{1-\psi[\psi^{-1}(1-uw_1)+\cdots+\psi^{-1}(1-uw_d)]}{\{1-\psi[\psi^{-1}(1-u)]\}^{k+M}} \\
&= \lim_{u\to 0^+} \frac{1-\psi\left[\psi^{-1}(1-u)\left(\frac{\psi^{-1}(1-uw_1)}{\psi^{-1}(1-u)}+\cdots+\frac{\psi^{-1}(1-uw_d)}{\psi^{-1}(1-u)}\right)\right]}{\{1-\psi[\psi^{-1}(1-u)]\}^{k+M}} \\
&= \lim_{u\to 0^+} \frac{1-\psi\left[\psi^{-1}(1-u)\left(\sum_{i\in I} w_i\right)\right]}{\{1-\psi[\psi^{-1}(1-u)]\}^{k+M}}.
\end{aligned}
$$

Let $s = \psi^{-1}(1-u)$ and

$$
Q(\mathbf{w}) = \sum_{\emptyset\neq I\subset I_d} (-1)^{|I|-1} \lim_{s\to 0^+} \frac{1-\psi\left[s\left(\sum_{i\in I} w_i\right)\right]}{\{1-\psi(s)\}^{k+M}}. \tag{3.25}
$$

To obtain the limit in (3.25), we may use the l'Hôpital's rule. For the first k derivatives of the numerator, for fixed $\mathbf{w}$ and $\psi^{(j)}(0)$ finite for $j=1,\ldots,k$,

$$
\lim_{s\to 0} \sum_{\emptyset\neq I\subset I_d} (-1)^{|I|-1} \psi^{(j)}\Big[s\Big(\sum_{i\in I} w_i\Big)\Big]\Big(\sum_{i\in I} w_i\Big)^j = 0, \quad j \in \{1,\ldots,k\},
$$

because by Lemma 3.12,

$$
\sum_{\emptyset\neq I\subset I_d} (-1)^{|I|-1}\Big(\sum_{i\in I} w_i\Big)^j = 0, \quad j\in\{1,\ldots,k\},\ 1\leq k<d.
$$

Then by the l'Hôpital's rule ($k+1$ applications)

$$
\begin{aligned}
Q(\mathbf{w}) = &\sum_{\emptyset\neq I\subset I_d} (-1)^{|I|-1} \times \\
&\times \lim_{s\to 0^+} \frac{-\psi^{(k+1)}[s\left(\sum_{i\in I} w_i\right)]\left(\sum_{i\in I} w_i\right)^{k+1}/\psi^{(k+1)}(s)}{[-\psi^{(1)}(s)]^{k+1}\{\prod_{j=0}^{k}(k+M-j)\}\,(1-\psi(s))^{M-1}/\psi^{(k+1)}(s)}.
\end{aligned}
$$

Since $g(s) = |\psi^{(k)}(s) - \psi^{(k)}(0)| = hs^M + o(s^M)$ with $h > 0$, we can write $g(s) \sim s^M \ell(s), s \to 0^+$ with a slowly varying function $\ell(s) \to h$ as $s \to 0^+$. Note that, $g(s) = \int_0^s |\psi^{(k+1)}(x)|dx = \int_0^s g'(x)dx$, and $g'(s)$ is monotonic as $s \to 0^+$, then by the Monotone density theorem, $|\psi^{(k+1)}(s)| = g'(s) \sim Ms^{M-1}\ell(s)$ and $\psi^{(k+1)}(s) \sim (-1)^{k+1} Ms^{M-1}\ell(s)$. Therefore,

$$Q(\mathbf{w}) = \frac{Mh}{[-\psi^{(1)}(0)]^{k+M} \prod_{j=0}^{k}(k+M-j)} \sum_{\emptyset \neq I \subset I_d} (-1)^{|I|+k+1} \Big(\sum_{i \in I} w_i \Big)^{k+M}.$$

Then, the upper tail order function is

$$b^*(\mathbf{w}) = \frac{Q(\mathbf{w})}{Q(\mathbf{1}_d)} = \frac{\sum_{\emptyset \neq I \subset I_d} (-1)^{|I|} \big(\sum_{i \in I} w_i \big)^{k+M}}{\sum_{\emptyset \neq I \subset I_d} (-1)^{|I|} \big| I \big|^{k+M}}.$$

Note that this is a homogeneous function in $\mathbf{w}$ of order $\kappa_U = M_\psi = k + M$. This completes the proof . $\square$

Proof of Proposition 3.8: If $r = 0$, then $\psi^{-1}(t) \sim (t/a_1)^{1/q}$ as $t \to 0^+$ (where $q < 0$). If $r > 0$, then for large s and small t, in

$$\log \psi(s) = \log t \sim \log a_1 + q \log s - a_2 s^r, \quad s \to \infty,$$

the third term dominates, so that

$$\psi^{-1}(t) \sim [(-\log t)/a_2]^{1/r}, \quad t \to 0^+.$$

Next, consider $C_\psi(u\mathbf{1}_d) = \psi(d\psi^{-1}(u))$.

For $r = 0$, one gets $\psi(d\psi^{-1}(u)) \sim \psi(da_1^{-1/q}u^{1/q}) \sim d^q u$, as $u \to 0^+$, with $d^q \in (0, 1)$, so that $\kappa_L = 1$.

For $0 < r < 1$, suppose

$$\Delta_{L,\kappa} = \lim_{u \to 0^+} \frac{\psi\big(d\psi^{-1}(u)\big)}{u^\kappa(-\log u)^\zeta} > 0.$$

Then by l'Hôpital's rule,

$$\Delta_{L,\kappa} = \lim_{u\to 0^+} \frac{d\psi'\big(d\psi^{-1}(u)\big)/\psi'\big(\psi^{-1}(u)\big)}{\kappa u^{\kappa-1}(-\log u)^{\zeta}} = \lim_{s\to\infty} \frac{d\psi'(ds)/\psi'(s)}{\kappa[\psi(s)]^{\kappa-1}[-\log\psi(s)]^{\zeta}}. \tag{3.26}$$

By condition (3.10), the dominating term of $\psi'(s)$ or $T'(s)$ is

$$\psi'(s) \sim -a_1a_2rs^{q+r-1}\exp\{-a_2s^r\}, \quad s\to\infty.$$

Consider the limit of the right-hand side of (3.26) without the factor d/κ:

$$\psi'(ds)/\psi'(s) \sim d^{q+r-1}\exp\{-a_2(d^r-1)s^r\}, \quad s\to\infty;$$

$$\begin{aligned}[\psi(s)]^{\kappa-1}[-\log\psi(s)]^{\zeta} &\sim [a_1s^q]^{\kappa-1}\exp\{-a_2(\kappa-1)s^r\}\cdot[a_2s^r-q\log s-\log a_1]^{\zeta}\\ &\sim a_1^{\kappa-1}a_2^{\zeta}s^{q(\kappa-1)+r\zeta}\exp\{-a_2(\kappa-1)s^r\}, \quad s\to\infty.\end{aligned}$$

Hence $\kappa = d^r$, $q(\kappa-1)+r\zeta = 0$ or $\zeta = (q/r)(1-\kappa) = (q/r)(1-d^r)$ and

$$\Delta_{L,\kappa} = \frac{dd^{q+r-1}}{\kappa a_1^{\kappa-1}a_2^{\zeta}} = \frac{d^q}{a_1^{\kappa-1}a_2^{\zeta}}.$$

So $\ell(u) = \Delta_{L,\kappa}\cdot(-\log u)^{\zeta} = d^q a_1^{1-\kappa}a_2^{-\zeta}(-\log u)^{\zeta}$.

Under the condition in (3.10), it can be verified that $-\psi(s)/\psi'(s)\in \mathrm{RV}_{1-r}$, which satisfies the condition in Theorem 3.3 of Charpentier and Segers (2009). So the tail order function is obtained. □

Proof of Proposition 3.9: When $0<\nu<1$,

$$K_\nu(s) \sim \tfrac{1}{2}\Big(\Gamma(\nu)\big(s/2\big)^{-\nu} + \Gamma(-\nu)\big(s/2\big)^{\nu}\Big). \tag{3.27}$$

We refer to the website of Wolfram Research Bes for asymptotic behavior of modified Bessel function of the second kind. For $0<\alpha<1$,

$$\psi(s;\alpha) \sim 1 + \frac{\Gamma(-\alpha)}{\Gamma(\alpha)}s^{\alpha}, \quad s\to 0^+,$$

and $1 - \psi(s;\alpha) \sim -s^{\alpha}\Gamma(-\alpha)/\Gamma(\alpha) \in \mathrm{RV}_{\alpha}(0^{+})$. This is consistent with Lemma 3.4.

Now, let us consider the case where α is non-integer with $\alpha > 1$. For an integer j with $0 < j < \alpha$ and $X' \sim \text{Gamma}(\alpha - j, 1)$,

$$\begin{aligned}\psi^{(j)}(s;\alpha) &= [\Gamma(\alpha)]^{-1}(-1)^{j}\int_{0}^{\infty} e^{-s/x}x^{\alpha-j-1}e^{-x}\,dx = (-1)^{j}\frac{\Gamma(\alpha-j)}{\Gamma(\alpha)}\mathbb{E}(e^{-s/X'})\\ &= (-1)^{j}2\Gamma^{-1}(\alpha)s^{(\alpha-j)/2}K_{\alpha-j}(2\sqrt{s}\,).\end{aligned}$$

So, $\phi(s;\alpha,j) = \psi^{(j)}(s;\alpha)/\psi^{(j)}(0;\alpha)$ is the LT of $Z = 1/X' \sim \mathrm{I}\Gamma(\alpha - j, 1)$ with $M_Z = \alpha - j$. When ν is non-integer with $|\nu| > 1$, the behavior near 0 of K_{ν} is

$$K_{\nu}(x) \sim x^{-|\nu|}2^{|\nu|-1}\Gamma(|\nu|)\Big[1 + \frac{x^{2}}{4(1-|\nu|)}\Big].$$

So, for integer $j < \alpha - 1$

$$\begin{aligned}&(-1)^{j}\psi^{(j)}(s;\alpha)\\ &= 2\Gamma^{-1}(\alpha)s^{(\alpha-j)/2}K_{\alpha-j}(2\sqrt{s}\,) \sim \frac{\Gamma(\alpha-j)}{\Gamma(\alpha)}\Big(1 + \frac{s}{1-\alpha+j}\Big), \quad s \to 0;\end{aligned}$$

for $k < \alpha < k + 1$, where $k \in \mathbb{N}_{+}$, then by (3.27)

$$\begin{aligned}&(-1)^{k}\psi^{(k)}(s;\alpha)\\ &= 2\Gamma^{-1}(\alpha)s^{(\alpha-k)/2}K_{\alpha-k}(2\sqrt{s}\,) \sim \frac{\Gamma(\alpha-k)}{\Gamma(\alpha)} + \frac{\Gamma(-\alpha+k)}{\Gamma(\alpha)}s^{\alpha-k}, \quad s \to 0.\end{aligned}$$

Therefore, $|\psi^{(k)}(s) - \psi^{(k)}(0)| \in \mathrm{RV}_{\alpha-k}(0^{+})$, which is consistent with Proposition 3.5. Then, by Proposition 3.5, there is a positive constant h_{k+1} such that

$$\begin{aligned}\psi(s) =& 1 + \psi^{(1)}(0)s + \tfrac{1}{2}\psi^{(2)}(0)s^{2} + \cdots + (-1)^{k}\psi^{(k)}(0)s^{k}/k! +\\ &+ (-1)^{k+1}h_{k+1}s^{\alpha} + o(s^{\alpha}).\end{aligned}$$

The upper tail order of the d-variate Archimedean copula C_{ψ} follows from Propositions 3.6 and 3.7. Therefore, if $\alpha \in (0, +\infty)$ is not an integer, the upper tail order is $\max\{1, \min\{\alpha, d\}\}$.

Next we investigate the lower tail. From Abramowitz and Stegun (1964), p. 378: for large z,

$$K_\nu(z) \sim \sqrt{\frac{\pi}{2z}}\, e^{-z}\Big\{1 + \frac{4\nu^2 - 1}{8z} + O(z^{-2})\Big\}.$$

Hence,

$$\begin{aligned}\psi(s;\alpha) &= 2\Gamma^{-1}(\alpha)s^{\alpha/2}K_\alpha(2\sqrt{s}\,) \\ &\sim 2\Gamma^{-1}(\alpha)s^{\alpha/2}\sqrt{\frac{\pi}{4s^{1/2}}}\, e^{-2s^{1/2}} = \pi^{1/2}\Gamma^{-1}(\alpha)s^{\alpha/2-1/4}e^{-2s^{1/2}}, \quad s \to \infty.\end{aligned} \tag{3.28}$$

Also, as $s \to \infty$,

$$\psi^{(1)}(s;\alpha) = -2\Gamma^{-1}(\alpha)s^{(\alpha-1)/2}K_{\alpha-1}(2\sqrt{s}\,) \sim -\pi^{1/2}\Gamma^{-1}(\alpha)s^{\alpha/2-3/4}e^{-2s^{1/2}}.$$

For the d-dimensional Archimedean copula, then by Proposition 3.8 with $a_1 = \pi^{1/2}\Gamma^{-1}(\alpha)$, $q = \alpha/2 - 1/4$, $a_2 = 2$ and $r = 1/2$ in (3.10), as $u \to 0$,

$$\begin{aligned}\psi\big(d\psi^{-1}(u)\big) &\sim \frac{d^{\alpha/2-1/4}}{a_1^{\kappa-1}2^{\zeta}}(-\log u)^{\zeta}u^{\sqrt{d}}, \\ &a_1 = \pi^{1/2}\Gamma^{-1}(\alpha),\ \zeta = (\alpha - 1/2)(1 - \sqrt{d}\,).\end{aligned}$$

Thus $\kappa_L(C_\psi) = \sqrt{d}$. $\square$

Proof of Proposition 3.10: Suppose that K is a multivariate max-id copula such that, for any index set $\emptyset \neq I \subset I_d$, $\overline{K}_I((1-s)\mathbf{1}_{|I|}) \sim s^{a_I}\ell_I(s)$, $s \to 0^+$, with $1 < a_I$ and $\ell_I(s) \in \mathrm{RV}_0(0^+)$. Note that

$$K((1-s)\mathbf{1}_d) = 1 + \sum_{\emptyset\neq I\subset I_d}(-1)^{|I|}\overline{K}_I\big((1-s)\mathbf{1}_{|I|}\big), \tag{3.29}$$

where $\overline{K}_I$ is the survival function of the I-margin copula K_I and let $\overline{K}_{\{i\}}(1-s) = s$ for any $i \in I_d$. Letting $s = 1 - \exp\{-\psi^{-1}(u)\}$, as $u \to 1^-$, i.e., $s \to 0^+$, since $a_I > 1$ for any $\emptyset \neq I \subset I_d$, $1 - ds$ dominates the right-hand side of (3.29), and thus,

$$-\log K\left(e^{-\psi^{-1}(u)}\mathbf{1}_d\right) = -\log K((1-s)\mathbf{1}_d) \sim -\log(1-ds)$$
$$\sim ds = d(1-\exp\{-\psi^{-1}(u)\}) \sim d\psi^{-1}(u).$$

Therefore,

$$C(u\mathbf{1}_d) \sim \psi\left(d\psi^{-1}(u)\right) = C_\psi(u\mathbf{1}_d), \quad u \to 1^-.$$

By Proposition 3.6, we know that C has intermediate upper tail dependence, and moreover, $\kappa_U(C) = \kappa_U(C_\psi)$. This proves (*a*).

To prove (*b*), note from Proposition 3.3 that $\kappa_U(K_I) = 1$ for any marginal copula K_I. Assuming $\ell_I(s) \to h_I \in (0,1]$ as $s \to 0^+$ and $h_{\{i\}} = 1$ for all i,

$$-\log K_I\left(e^{-\psi^{-1}(u)}\mathbf{1}_{|I|}\right) = -\log K_I((1-s)\mathbf{1}_{|I|}) \sim -\log(1-h_I^* s) \sim h_I^*\psi^{-1}(u),$$

where $h_I^* = \sum_{\emptyset \neq J \subset I}(-1)^{|J|-1}h_J$. By the construction of (3.12), for any I-margin copula of C,

$$C_I(u\mathbf{1}_{|I|}) = \psi\big(-\log K_I(e^{-\psi^{-1}(u)}\mathbf{1}_{|I|})\big) \sim \psi\big(h_I^*\psi^{-1}(u)\big).$$

Thus, as $s \to 0^+$, i.e., $u \to 1^-$,

$$\overline{C}(u\mathbf{1}_d) = 1 + \sum_{\emptyset \neq I \subset I_d}(-1)^{|I|}C_I(u\mathbf{1}_{|I|})$$
$$\sim 1 + \sum_{\emptyset \neq I \subset I_d}(-1)^{|I|}\psi\big(h_I^*\psi^{-1}(u)\big) = \sum_{\emptyset \neq I \subset I_d}(-1)^{|I|-1}\Big(1-\psi\big[h_I^*\psi^{-1}(u)\big]\Big).$$

If $1-\psi(x) \in \mathrm{RV}_\beta(0^+)$, then clearly,

$$\overline{C}((1-u)\mathbf{1}_d) \sim u\sum_{\emptyset \neq I \subset I_d}(-1)^{|I|-1}(h_I^*)^\beta, \quad u \to 0^+.$$

So, $\kappa_U(C) = 1$ and $\lambda_U(C) = \sum_{\emptyset \neq I \subset I_d}(-1)^{|I|-1}(h_I^*)^\beta$. $\square$

Proof of Proposition 3.11: Suppose that a d-variate max-id copula $K(s\mathbf{1}_d) \sim s^{\alpha}\ell(s)$ as $s \to 0^+$ and let $s = \exp\{-\psi^{-1}(u)\}$. As $u \to 0^+$, thus $s \to 0^+$,

$$-\log K\left(e^{-\psi^{-1}(u)}\mathbf{1}_d\right) = -\log K(s\mathbf{1}_d) \sim -\log(s^{\alpha}\ell(s)) \sim -\alpha\log s = \alpha\psi^{-1}(u).$$

Therefore,

$$C(u\mathbf{1}_d) = \psi\left(-\log K\big(e^{-\psi^{-1}(u)}\mathbf{1}_d\big)\right) \sim \psi\left(\alpha\psi^{-1}(u)\right), \quad u \to 0^+.$$

With some modification of the proof of Theorem 3.3 of Charpentier and Segers (2009), we can prove the rest. For purpose of notational convenience, we include the modification in the following. Letting $\Lambda = (\alpha + d\nu)/(1+\nu)$ and $\omega(s) = -\psi(s)/\psi'(s)$, then we know that

$$\lim_{s\to 0^+} \frac{\psi^{-1}(sx) - \psi^{-1}(s)}{\omega(\psi^{-1}(s))} = -\log(x),$$

and if $y(t) \to y \in \mathbb{R}$ as $t \to \infty$, then

$$\lim_{t\to\infty} \frac{\psi(t + y(t)\omega(t))}{\psi(t)} = \exp(-y).$$

For any $t > 0$, write $\psi(\alpha\psi^{-1}(ut)) = \psi\left[\alpha\psi^{-1}(u) + y(u,t)\omega(\alpha\psi^{-1}(u))\right]$, where

$$y(u,t) = \left(\frac{\alpha[\psi^{-1}(ut) - \psi^{-1}(u)]}{\omega(\psi^{-1}(u))}\right)\frac{\omega(\psi^{-1}(u))}{\omega(\alpha\psi^{-1}(u))}.$$

As $u \to 0^+$, $y(u,t) \to -\alpha\log(t)\alpha^{-\beta} = -\alpha^{1-\beta}\log(t)$. Therefore,

$$\lim_{u\to 0^+} \frac{\psi^{-1}(\alpha\psi^{-1}(ut))}{\psi^{-1}(\alpha\psi^{-1}(u))} = \exp(\alpha^{1-\beta}\log(t)) = t^{\alpha^{1-\beta}},$$

and thus $C(u\mathbf{1}_d) \in \mathrm{RV}_{\alpha^{1-\beta}}(0^+)$. We have also known from Theorem 3.3 of Charpentier and Segers (2009) that $\kappa_L(C_\psi) = d^{1-\beta}$. This completes the proof. $\square$

Chapter 4

Tail comonotonicity

4.1 Introduction

Suppose we have bivariate loss data and hope to estimate some high-risk scenarios, say by VaR or CTE. A widely used method is to fit some parametric models based on the bivariate Student t or other parametric copula families. Then risk measures or tail dependence can be derived from these fitted models. Due to model uncertainty, all of these methods are not conservative from the viewpoint of an actuary. The middle part might influence our estimation more than the important tail part does. These traditional methods are too sensitive to the middle part which contains most of the data (Nikoloulopoulos et al., 2012).

Given that we do not have enough information for the joint tail of losses, we conservatively assume that it is *upper tail comonotonic* (as defined in Section 4.2), and the conservativity can be justified in Chapter 5. In this way, we actually give up estimating the "first order tail parameter" (i.e., the usual tail dependence parameter λ) by assuming a conservative one (i.e., $\lambda = 1$), and let the likelihood of data contribute to the estimation of the "second order tail parameter". Since the first order parameter is only for an asymptotic property, the conservative assumption on it would not put too much constraint on the model. This approach will give us a more robust method of measuring risks.

In the literature on actuarial science and quantitative finance, many efforts have been done to seek finer upper bounds for dependence structures. The concepts of comonotonicity, conditional comonotonicity, and most recently, the upper comonotonicity have been studied to provide theoretically tractable bounds. However, these conditions are either too strong or not tailored for the tail, and could lead to an over-conservative risk measure. Moreover, those dependence structures lack flexible distribution families that can be used to model real data. We refer to Dhaene et al. (2002a,b); Cheung (2007, 2009) for the reference of these concepts. Tail comonotonicity, on the other hand, needs a weaker condition and requires that the degree of positive dependence approaches its maximum only when all the marginal losses go to infinity. The degree of dependence at the sub-extremal level can still be estimated from the data. This approach should better balance the requirements of safety and accuracy for risk management.

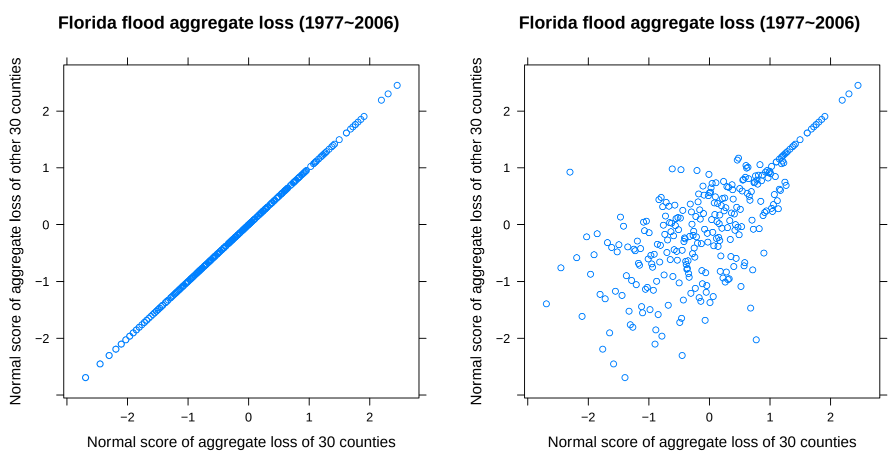
Tail comonotonicity is also referred to as asymptotic full dependence or a dependence structure where the tail dependence parameter satisfies $\lambda = 1$. Although such a dependence structure is not new, we find that it has interesting properties, such as asymptotic additivity of VaR and CTE, that are analogous to the usual comonotonicity. Moreover, some parametric families illustrate its suitability for modeling loss data that may appear to have tail dependence. Among many copula families, Archimedean copulas and copulas constructed from a scale mixture of a non-negative random vector can be used to provide a tail comonotonic dependence structure.

This chapter is organized as follows. In Section 4.2, the concept of tail comonotonicity, its properties and some parametric examples will be studied. Methods of constructing copulas with tail comonotonicity or the strongest tail dependence are given in Section 4.3. Asymptotic additivity of VaR and CTE under the assumption of tail comonotonicity is shown in Section 4.4. Finally, in Section 4.5, we conclude the chapter and propose some directions for future research.

4.2 Definitions of tail comonotonicity and properties

Let $\mathbf{X} = (X_1, \ldots, X_d)^{\mathrm{T}}$ be a non-negative random vector, representing amounts of d losses, and the univariate marginal cdf's are all continuous and denoted as $F_1, \ldots, F_d$. The Fréchet space containing all the random vectors with these univariate margins is denoted as $\mathcal{R}(F_1, \ldots, F_d)$. For risk management, the joint tail probability $\mathbb{P}[X_1 > x_1, \ldots, X_d > x_d]$ is relevant, especially when $x_1, \ldots, x_d$ are large values. Cheung (2009) studied an upper comonotonicity structure: roughly speaking, beyond a finite threshold $\mathbf{a}^*$, the dependence structure becomes comonotonicity, while keeping the dependence structure below the threshold flexible; this means that the distribution is not absolutely continuous. Using the Florida flooding dataset mentioned in Figure 1.2, the idea of comonotonicity and upper comonotonicity can be illustrated by Figure 4.1.

Figure 4.1: Florida flood data - Comonotonicity vs Upper comonotonicity



Under this dependence structure, if $\mathbf{x} \geq \mathbf{a}^*$, then clearly

$$\begin{aligned}\mathbb{P}[X_1 > x_1, \ldots, X_d > x_d] &= \min\{\overline{F}_1(x_1), \ldots, \overline{F}_d(x_d)\} \\ &= 1 - \max\{F_1(x_1), \ldots, F_d(x_d)\},\end{aligned}$$

which is the upper bound for such a tail probability of any $\mathbf{X} \in \mathcal{R}(F_1, \ldots, F_d)$. The upper bound coincides with the tail probability for the usual comonotonicity.

Letting $\widehat{C}$ be the survival copula of $\mathbf{X}$,

$$\mathbb{P}[X_1 > x_1, \ldots, X_d > x_d] = \widehat{C}(\overline{F}_1(x_1), \ldots, \overline{F}_d(x_d)).$$

A proper threshold $\mathbf{a}^*$ might not exist in real applications. However, note that our aim of proposing a conservative dependence structure will be met if the tail probability can be well approximated by the upper bound when $\mathbf{x}$ is *sufficiently large*; that is,

$$\widehat{C}(\overline{F}_1(x_1), \ldots, \overline{F}_d(x_d)) \approx \min\{\overline{F}_1(x_1), \ldots, \overline{F}_d(x_d)\}, \quad x_i\text{'s are sufficiently large.}$$

When $\mathbf{x}$ is sufficiently large, $\overline{F}_i(x_i)$'s become sufficiently close to 0. But how x_i's converge to the right end points of their supports will affect the joint tail probability. Considering that the usual comonotonicity is a copula property and does not depend on the margins, here we want to define a dependence concept which will not rely on margins as well. Therefore, letting $u_i := \overline{F}_i(x_i)$, it suffices to have that

$$\widehat{C}(u_1, \ldots, u_d) \approx \min\{u_1, \ldots, u_d\}, \quad u_i\text{'s are sufficiently small.}$$

To be convenient in real applications, this condition should also be satisfied with copulas $\widehat{C}$ that are absolutely continuous. Of course, how u_i's approach 0 will affect the above approximation. However, as long as the rates of convergence of u_i's to 0 are comparable (in the sense that $u_i = uw_i$ for any given $0 < w_i < +\infty$), the above approximation should be good. Then $\widehat{C}$ would be what we want if it satisfies

$$\widehat{C}(uw_1, \ldots, uw_d) \sim \min\{uw_1, \ldots, uw_d\}, \quad u \to 0^+, \quad w_i, \ldots, w_d \in [0, +\infty),$$

i.e.,

$$\lim_{u \to 0^+} \frac{\widehat{C}(uw_1, \ldots, uw_d)}{u} = \min\{w_1, \ldots, w_d\}, \quad w_i, \ldots, w_d \in [0, +\infty), \quad (4.1)$$

where the trivial case of $w_i = 0$ for some i is also included.

Definition 4.1 A random vector $\mathbf{X}$ is said to be *upper tail comonotonic* if $\mathbf{X}$ has a copula C and its survival copula $\widehat{C}$ satisfies (4.1); the copula C is said to be an upper tail comonotonic copula. $\mathbf{X}$ is said to be *lower tail comonotonic* if $\mathbf{X}$ has a copula C that satisfies $\lim_{u \to 0^+} C(uw_1, \ldots, uw_d)/u = \min\{w_1, \ldots, w_d\}, w_i \in [0, +\infty)$; the copula C is said to be a lower tail comonotonic copula.

Remark 4.1 Based on the above definition, tail comonotonicity is a concept about copula and does not rely on the marginal distributions as long as the conditions hold. In Definition 4.1, the random vector $\mathbf{X}$ is not necessarily continuous and the right end points of each univariate margin can be finite or infinite.

Remark 4.2 In the framework of multivariate regular variation (MRV), the concept of upper tail comonotonicity is represented by the limit Radon measure (Resnick 2007 Resnick (2007))

$$\nu([\mathbf{0}, \mathbf{x}]^c) = (\min\{x_1, \ldots, x_d\})^{-\alpha}.$$

So asymptotic full dependence for MRV is a case of upper tail comonotonicity with the univariate margins being power laws.

Recall from Joe et al. (2010) that, when C is a d-variate copula with the survival function $\overline{C}$, then lower tail dependence means that there exists a non-zero homogeneous function b of order 1 such that

$$\lim_{u \to 0^+} C(uw_1, \ldots, uw_d)/u = b(w_1, \ldots, w_d), \quad w_i \in [0, +\infty), \tag{4.2}$$

and upper tail dependence means that there exists a non-zero homogeneous function b^* of order 1 such that

$$\lim_{u \to 0^+} \overline{C}(1 - uw_1, \ldots, 1 - uw_d)/u = b^*(w_1, \ldots, w_d), \quad w_i \in [0, +\infty). \tag{4.3}$$

So, tail comonotonicity simply means that the lower and/or upper tail dependence functions are $\min\{w_1, \ldots, w_d\}$. Also, note that for any tail dependence function $b(\mathbf{w})$, we must have $b(\mathbf{w}) \leq \min(w_1, \ldots, w_d)$.

Some relevant properties for upper/lower tail comonotonicity are mentioned in Beirlant et al. (2004), we give some alternative results in the following two propositions.

Proposition 4.1 *Suppose C is a copula. Then $\lambda_U(C) = 1$ if and only if the upper tail dependence function exists and $b^*(w_1, \ldots, w_d) = \min(w_1, \ldots, w_d)$. In* *parallel, $\lambda_L(C) = 1$ if and only if the lower tail dependence function exists and $b(w_1, \ldots, w_d) = \min(w_1, \ldots, w_d)$.*

Remark 4.3 From Proposition 4.1, we can conclude that the concept of upper/lower tail comonotonicity is nothing but the tail dependence parameter $\lambda = 1$. The designation of this concept is just consistent with the names of upper/lower tail dependence. If the tail dependence function exists, then it must have the unique form.

It is well known that all pairs of a random vector are comonotonic is equivalent to that the random vector is comonotonic. A parallel result also holds for upper/lower tail comonotonicity.

Proposition 4.2 *Suppose C is a d-variate copula, then any bivariate marginal copula $C_{\{ij\}}, i \neq j$, is upper [resp. lower] tail comonotonic if and only if C is upper [resp. lower] tail comonotonic.*

Next, we give two examples of parametric tail comonotonic copulas that are two-parameter Archimedean copulas.

Example 4.1 (BB2 in Joe and Hu (1996), Joe (1997)) With LT

$$\psi(s) = [1 + \delta^{-1} \log(1 + s)]^{-1/\theta},$$

the bivariate Archimedean copula is

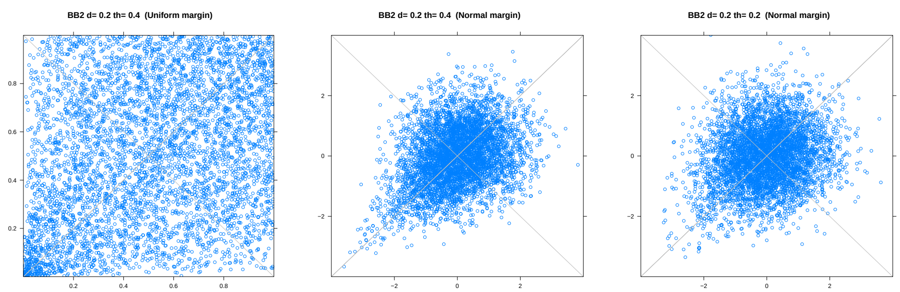
$$C_\psi(u, v) = \left[1 + \delta^{-1} \log\left(e^{\delta(u^{-\theta}-1)} + e^{\delta(v^{-\theta}-1)} - 1\right)\right]^{-1/\theta}, \quad \theta > 0, \delta > 0.$$

Then as $u \to 0^+$,

$$\begin{aligned}\frac{C_\psi(uw_1, uw_2)}{u} &\sim \left[1 + \delta^{-1} \max\{\delta((uw_1)^{-\theta} - 1), \delta((uw_2)^{-\theta} - 1)\}\right]^{-1/\theta} u^{-1} \\ &= \min(w_1, w_2).\end{aligned}$$

Thus C_ψ is lower tail comonotonic. Scatter plots of the BB2 copula with parameters $\delta = 0.2$ and $\theta = 0.4$ or 0.2 are in Figure 4.2 ($N = 2000$); there is no upper tail dependence for this copula.

Figure 4.2: Simulation of BB2 with/without the univariate margins being transformed to the standard Normal; in the left and middle plots $\delta = 0.2$, $\theta = 0.4$ and in the right plot $\delta = 0.2$, $\theta = 0.2$



Example 4.2 (BB3 in Joe and Hu (1996), Joe (1997)) With LT

$$\psi(s) = \exp\left\{-[\delta^{-1}\log(1+s)]^{1/\theta}\right\},$$

the bivariate Archimedean copula is

$$C_\psi(u,v) = \exp\left\{-[\delta^{-1}\log(e^{\delta\tilde{u}^\theta} + e^{\delta\tilde{v}^\theta} - 1)]^{1/\theta}\right\}, \quad \theta > 1, \delta > 0,$$

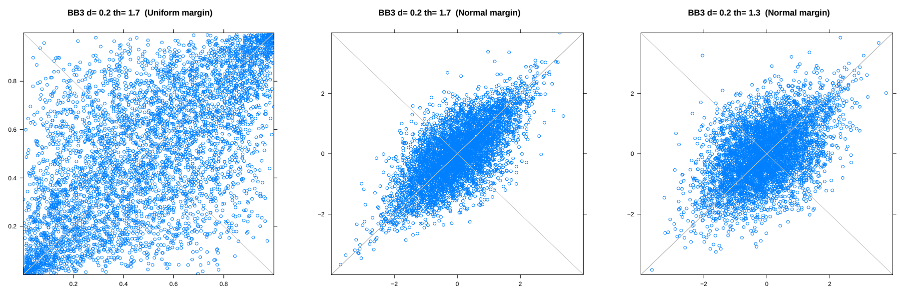
where $\tilde{u} = -\log u, \tilde{v} = -\log v$. Then as $u \to 0^+$,

$$\frac{C_\psi(uw_1, uw_2)}{u} \sim \exp\left\{-[\delta^{-1}\log(e^{\delta(-\log[u\min(w_1,w_2)])^\theta})]^{1/\theta}\right\} u^{-1} = \min(w_1, w_2).$$

Thus C_ψ is also lower tail comonotonic. Scatter plots of the BB3 copula with parameters $\delta = 0.2$ and $\theta = 1.7$ or 1.3 are in Figure 4.3 ($N = 2000$); there is also upper tail dependence for this copula and $\lambda_U = 2 - 2^{1/\theta}$.

Remark 4.4 Looking at the plots for the lower comonotonic copulas, they appear suitable as survival copulas to be used to get conservative dependence structure for joint large losses. Although they are both lower tail comonotonic, there is not much constraint on the sub-extremal level.

Figure 4.3: Simulation of BB3 with/without the univariate margins being transformed to the standard Normal; in the left and middle plots $\delta = 0.2$, $\theta = 1.7$ and in the right plot $\delta = 0.2$, $\theta = 1.3$



4.3 Construction of tail comonotonic copulas

In this section, we propose methods to construct tail comonotonic copulas based on mixing distributions with very heavy tails; "very heavy" means there are no moments of any positive/negative orders or the corresponding survival function is slowly varying. Precise conditions will be presented in the following subsections.

4.3.1 Archimedean copulas

In this subsection, we will relate the tail behavior of LTs to upper/lower tail comonotonicity of corresponding Archimedean copulas.

The next result says that if the LT ψ is slowly varying at the tails, then C_ψ must have the tail dependence function $\min(w_1, \ldots, w_d)$.

Proposition 4.3 *Let the Archimedean copula C_ψ be based on the LT ψ satisfying $\psi(\infty) = 0$. If $\psi(s) \in \mathrm{RV}_0$ then the lower tail dependence function exists and C_ψ is lower tail comonotonic; if $1 - \psi(s) \in \mathrm{RV}_0(0^+)$ then the upper tail dependence function exists and C_ψ is upper tail comonotonic.*

Remark 4.5 Write the Archimedean copula as the mixture

$$C_\psi(u_1, \ldots, u_d) = \int_0^\infty \prod_{i=1}^{d} G^\eta(u_i) dF_H(\eta),$$

where $G(u) = \exp\{-\psi^{-1}(u)\}$ is a cdf on $[0,1]$ and H is a resilience random variable with LT ψ. The condition $1-\psi \in \mathrm{RV}_0(0^+)$ means that the density $f_H(\eta)$ has a very heavy tail as $\eta \to \infty$ so that $\mathbb{E}[H^m] = \infty$ for all $m > 0$. The condition $\psi \in \mathrm{RV}_0$ means that the density $f_H(\eta)$ has a very heavy tail as $\eta \to 0$ so that $\mathbb{E}[H^{-m}] = \infty$ for all $m > 0$. The implication of the condition of $1-\psi$ at 0 follows by the proof of Lemma 3.4. Similarly, we can show the implication of the condition of ψ at ∞ as follows. Suppose to the contrary that $\mathbb{E}[H^{-m}] < \infty$ for some $m > 0$. Let $W_m = (Z/H)^m$, where $Z \sim$ Exponential(1), independent of H, and H has the LT ψ. Then $\mathbb{P}[W_m \geq w] = \mathbb{P}[Z \geq Hw^{1/m}] = \int_0^\infty \exp(-yw^{1/m})F_H(dy) = \psi(w^{1/m})$. Then $\mathbb{E}[H^{-m}] < \infty$ and Z having all positive moments implies that $\mathbb{E}[W_m] < \infty$ and thus $w\mathbb{P}[W_m \geq w] \to 0$, as $w \to \infty$; that is, $w\psi(w^{1/m}) \to 0$, or equivalently $s^m\psi(s) \to 0$, as $s \to \infty$. It is well known that if $\psi(s) \in \mathrm{RV}_0$ and $m > 0$, we must have $s^m\psi(s) \to \infty$. The contradiction implies that $\mathbb{E}[H^{-m}] = \infty$ for all $m > 0$.

It can be verified that the LTs for BB2 and BB3 are both slowly varying at $+\infty$, except in the lower boundary case (i.e., $\theta = 1$ for BB3); that is,

$$\psi_{BB2}(s) = [1 + \delta^{-1}\log(1+s)]^{-1/\theta} \in \mathrm{RV}_0, \quad \theta > 0, \delta > 0;$$
$$\psi_{BB3}(s) = \exp\left\{-[\delta^{-1}\log(1+s)]^{1/\theta}\right\} \in \mathrm{RV}_0, \quad \theta > 1, \delta > 0.$$

Now, consider the tail behavior of the LT studied in (3.10); that is, assume that a LT η satisfies

$$\eta(s) \sim T(s) = a_1 s^q \exp\{-a_2 s^r\}$$
$$\text{and } \eta'(s) \sim T'(s), \quad s \to \infty, \text{ with } a_1 > 0, a_2 \geq 0, \tag{4.4}$$

where $r = 0$ implies $a_2 = 0$ and $q < 0$, and $r > 0$ implies $r \leq 1$ and q can be 0, negative or positive. Note that $r > 1$ or $r < 0$ is not possible because of the complete monotonicity property of a LT. This condition covers almost all of the LT families in the Appendix of Joe (1997), as well as other LT families that can be obtained by integration or differentiation.

The next result is contained in the Appendix of Joe (1997), where $\mathcal{L}^*_{+\infty}$ is the class of infinitely differentiable increasing functions of $[0,\infty)$ onto $[0,\infty)$, with alternating signs for derivatives. This combines some results from pages 441 and 450 of Feller (1971); the condition on ϕ implies that it is the LT of an infinitely divisible random variable.

Lemma 4.4 *If ϕ is a LT such that $-\log\phi \in \mathcal{L}^*_{+\infty}$ and η is another LT, then $\psi(s) = \eta(-\log\phi(s))$ is a LT.*

Proposition 4.5 *Suppose the LT $\eta(t)$ satisfies condition (4.4), and take LT of Gamma $\phi(s) = (1+s)^{-1/\theta}$, thus $\psi(s) := \eta(-\log(\phi(s)))$ is also a LT. Then $0 \le r < 1$ implies that $\psi(s) \in \mathrm{RV}_0$.*

In the literature, risks are compared with respect to some stochastic orders, say, the usual stochastic order and the increasing convex order (Müller and Stoyan, 2002). In our opinion, the comparisons of risks are more meaningful for high-risk scenarios. However, the conditions these stochastic orders need to satisfy are too strong to be flexible for comparing tail risks. To this end, we may define some new concepts of stochastic order that are particularly used to compare the tails of univariate/multivariate cdf's. For example, we may define a stochastic order based on asymptotic properties of distribution functions as what follows.

Definition 4.2 (Ultimate usual stochastic orders) Let X, Y be random variables, X is said to be greater than Y in the *upper ultimate usual stochastic order* if there exists a finite q such that $\overline{F}_X(t) \ge \overline{F}_Y(t)$ for all $t \ge q$, written as $X \gtrsim^{st} Y$. Let $\mathbf{X}$ and $\mathbf{Y}$ be random vectors, then $\mathbf{X}$ is said to be greater than $\mathbf{Y}$ in the *upper ultimate usual stochastic order* if there exists a finite threshold $\mathbf{q}$ such that $\overline{F}_{\mathbf{X}}(\mathbf{t}) \ge \overline{F}_{\mathbf{Y}}(\mathbf{t})$ for all $\mathbf{t} \ge \mathbf{q}$, written as $\mathbf{X} \gtrsim^{st} \mathbf{Y}$. For lower ultimate usual stochastic order, the notation is ${}_{st}\gtrsim$. Then X is said to be greater than Y in the *lower ultimate usual stochastic order* (written as $X{}_{st}\gtrsim Y$) if there exists a finite q such that $F_X(t) \ge F_Y(t)$ for all $t \le q$, and $\mathbf{X}$ is said to be greater than $\mathbf{Y}$ in the *lower ultimate usual stochastic order* if there exists a finite threshold $\mathbf{q}$ such that $F_{\mathbf{X}}(\mathbf{t}) \ge F_{\mathbf{Y}}(\mathbf{t})$ for all $\mathbf{t} \le \mathbf{q}$, written as $\mathbf{X}{}_{st}\gtrsim\mathbf{Y}$. The relationships can also be symbolized by using corresponding cdfs, such as, $F_{\mathbf{X}}{}_{st}\gtrsim F_{\mathbf{Y}}$ for $\mathbf{X}{}_{st}\gtrsim\mathbf{Y}$.

If X, Y represent amounts of losses, then $X \gtrsim^{st} Y$ implies that X is riskier than Y. If the right tail of X is heavier than that of Y, then this is a sufficient condition for $X \gtrsim^{st} Y$.

Proposition 4.6 *Let ψ_1, ψ_2 be given LTs. If $\psi_1^{-1} \circ \psi_2(s)$ is superadditive for sufficiently large s, then there exists a finite threshold $\mathbf{q}$ such that for any $\mathbf{u} \leq \mathbf{q}$, $C_{\psi_1}(\mathbf{u}) \geq C_{\psi_2}(\mathbf{u})$, i.e., $C_{\psi_1\,st}\!\gtrsim C_{\psi_2}$.*

Note: using notation of Archimedean copula in Nelsen (2006), the generator should be strict in order to apply Proposition 4.6.

Remark 4.6 From Proposition 4.6, we expect that BB2 is more lower tail positive dependent than BB3 below a sufficiently low threshold, although they are both lower tail comonotonic. Let $\psi_1(s) := \psi_{\mathrm{BB2}}(s) = [1 + \delta_1^{-1} \log(1+s)]^{-1/\theta_1}$ and $\psi_2(s) := \psi_{\mathrm{BB3}}(s) = \exp\left\{-[\delta_2^{-1} \log(1+s)]^{1/\theta_2}\right\}$, then

$$g(s) := \psi_1^{-1} \circ \psi_2(s) = \exp\{\delta_1(e^{\theta_1 (\log(1+s)/\delta_2)^{1/\theta_2}} - 1)\} - 1.$$

Letting $h(s) := \log(1+s)/\delta_2$,

$$g'(s) = \delta_1 \theta_1 h(s)^{1/\theta_2} e^{\theta_1 h(s)^{1/\theta_2}} e^{\delta_1 \left(e^{\theta_1 h(s)^{1/\theta_2}} - 1\right)} / (\theta_2 (1+s) \log(1+s)).$$

By observing $g'(s)$, we know that $g(s)$ is strictly increasing and ultimately strictly convex. Assume that $g(s)$ is strictly convex as $s \geq s_0$, and let $y_0 := g(s_0)$ and $z(s) := g(s+s_0) - y_0$, then $z(s)$ is superadditive for $s \in [0, \infty)$. Therefore, for $x, y \geq 0$, $z(x+y) \geq z(x) + z(y)$; that is, $g(x+y+s_0) + g(s_0) \geq g(x+s_0) + g(y+s_0)$. Since $g(s)$ is strictly increasing and strictly convex as $s \geq s_0$, we must have, when x, y are sufficiently large, $g(x+y+2s_0) \geq g(x+y+s_0) + g(s_0)$. Therefore, $\psi_1^{-1} \circ \psi_2(s)$ is ultimately superadditive.

4.3.2 Heavy tail mixtures

Consider the heavy tail scale mixture

$$\mathbf{X} = (RT_1, \ldots, RT_d), \tag{4.5}$$

where R and T_i's are all non-negative random variables, R is independent of the T_i's, and the dependence structure between T_i's is not specified. This form covers truncated elliptical distributions. The CTE for such random vectors has been studied in Zhu and Li (2012), and more relevant study is in Chapter 6, where R is assumed to be regularly varying and second order regularly varying in the right tail, respectively.

The following lemma is often referred to as the Breiman's Theorem (Breiman (1965)). Although the proof in Breiman (1965) is not for a general α, it can be adapted for proving a more general case where $0 \leq \alpha < +\infty$. To the best of our knowledge, we have not found a complete and detailed proof for this case. So we include a proof in Section 4.6, and emphasize that the result for $\alpha = 0$ corresponding to slowly variation of Y also holds.

Lemma 4.7 *Suppose a random variable $R \geq 0$ with $\overline{F}_R(y) \in \mathrm{RV}_{-\alpha}, (0 \leq \alpha < +\infty)$. $T \geq 0$ is independent of R and $\mathbb{E}[T^{\alpha+\delta}] < \infty$ for some $\delta > 0$. Then*

$$\lim_{y\to\infty} \frac{\mathbb{P}[RT > y]}{\mathbb{P}[R > y]} = \mathbb{E}[T^\alpha]. \tag{4.6}$$

Proposition 4.8 *Let $\mathbf{X} = (RT_1, \ldots, RT_d)$ be defined as in (4.5). If $R \in \mathrm{RV}_0$, $\mathbb{E}[T_i^{\delta_i}] < \infty, i = 1, \ldots, d$, for some $\delta_i > 0$, then $\mathbf{X}$ is upper tail comonotonic.*

The above scale mixture can be used to construct upper tail comonotonic copulas. Note that, the margins T_i's are not necessarily identical and their dependence structure is not specified.

An elliptical random vector $\mathbf{X}$ can be written as

$$\mathbf{X} \overset{d}{=} \boldsymbol{\mu} + RA\mathbf{U} \sim E_d(\boldsymbol{\mu}, \Sigma, \phi), \tag{4.7}$$

where the radial random variable $R \geq 0$ is independent of $\mathbf{U}$, $\mathbf{U}$ is an k-dimensional random vector uniformly distributed on the surface of the unit hypersphere $\mathbb{S}_2^{k-1} = \{\mathbf{z} \in \mathbb{R}^k | \mathbf{z}^\mathrm{T}\mathbf{z} = 1\}$, A is a $d \times k$ matrix with $rank(A) = k$ and $AA^\mathrm{T} = \Sigma$ that is positive semidefinite, ϕ is the characteristic generator. When Σ is positive definite, write $\rho_{ij} := \Sigma_{ij}/\sqrt{\Sigma_{ii}\Sigma_{jj}}$, and $\rho_{ij} \neq \pm 1$ in the following to avoid trivial cases. We refer to Fang et al. (1990) for a comprehensive reference for elliptical distributions.

Now, we are studying whether there exist tail comonotonic elliptical distributions. Note that the T_i's in Proposition 4.8 are all non-negative, so it does not cover the usual elliptical distributions. For tail dependence of elliptical distributions, Schmidt (2002), and Hult and Lindskog (2002) study the case where the radial random variable R has certain tail patterns such as regularly varying right tails. Frahm et al. (2003) shows that the two different representations of tail dependence parameters in Schmidt (2002) and Hult and Lindskog (2002) for the regularly varying case are equivalent. From their results, we know that, with a regularly varying radial random variable R (tail index $-\alpha < 0$), the non-degenerate elliptical distributions have tail dependence parameters that are strictly less than 1. We now show that, even when $\mathbb{P}[R > s]$ is slowly varying in s, tail comonotonicity still does not hold. The intuitive reasoning for this result is that for bivariate elliptical distributions that have usual tail dependence, there is tail dependence in all corners/quadrants, so that the tail dependence in the upper positive quadrant is not 1.

Proposition 4.9 *Suppose* $\mathbf{X}$ *is an elliptical random vector as defined in (4.7) with* $\Sigma_{ii} > 0$ *for* $i = 1, \ldots, d$, *if* $\mathbb{P}[R > s] \in \mathrm{RV}_0$, *then any bivariate margins of* $\mathbf{X}$ *is upper (and lower) tail dependent, and the tail dependence parameter* $\lambda_{ij} = 1/2 + (1/\pi) \arcsin \rho_{ij}$.

Remark 4.7 In the proof of Theorem 4.3 in Hult and Lindskog (2002), it is claimed that $\mathbf{X} \stackrel{d}{=} RA\mathbf{U}$ implies that

$$\begin{pmatrix} X_i \\ X_j \end{pmatrix} \stackrel{d}{=} R \begin{pmatrix} \sqrt{\Sigma_{ii}} & 0 \\ \sqrt{\Sigma_{jj}}\rho_{jj} & \sqrt{\Sigma_{jj}}\sqrt{1-\rho_{ij}^2} \end{pmatrix} \begin{pmatrix} \cos\varphi \\ \sin\varphi \end{pmatrix},$$

where $\varphi \sim \mathrm{Uniform}(-\pi, \pi)$. Here we need to make a note that the above claim is true only when $\mathbf{X}$ is bivariate. For a general d-dimensional $\mathbf{X}$ ($d > 2$), radial random variables for margins will not be the same as the original R for the whole distribution. We make a more detailed argument about this in what follows.

Due to the Corollary in page 43 of Fang et al. (1990), margins of elliptical distributions are still elliptical with the same characteristic generator ϕ. Moreover, due to Lemma 5.3 of Schmidt (2002), the radial random variable R depends on the dimension d of the elliptical distribution, and there exists a constant $k > 0$ such that

$$R_s \stackrel{d}{=} kR_d\sqrt{B}, \quad B \sim \text{Beta}(s/2, (d-s)/2), \quad s < d,$$

and B is independent of R_d. Since a Beta distribution has moments of all orders, by Lemma 4.7, R_s inherits the tail behavior of R_d when R_d is regularly varying or slowly varying, which is why the original proof in Hult and Lindskog (2002) can still yield a correct conclusion.

Let (U_1, U_2) follows a bivariate copula, where $U_i \sim \mathrm{Uniform}(0,1), i = 1, 2$. From Embrechts et al. (2009a), the tail dependence parameters in the "North East (NE)" and "South East (SE)" for the bivariate Student t copula are the following:

$$\lambda_{\mathrm{NE}} := \lim_{u \to 1^-} \mathbb{P}[U_2 > u, |U_1 > u] = \frac{\int_{(\pi/2-\arcsin\rho)/2}^{\pi/2} (\cos t)^\nu dt}{\int_0^{\pi/2} (\cos t)^\nu dt} =: \Lambda(\nu, \rho); \tag{4.8}$$

$$\lambda_{\mathrm{SE}} := \lim_{u \to 1^-} \mathbb{P}[U_2 \le 1-u, |U_1 > u] = \frac{\int_{(\pi/2+\arcsin\rho)/2}^{\pi/2} (\cos t)^\nu dt}{\int_0^{\pi/2} (\cos t)^\nu dt} =: \Lambda(\nu, -\rho). \tag{4.9}$$

Note that $\lambda_{\mathrm{NE}} + \lambda_{\mathrm{SE}}$ is still a conditional probability and thus $0 \le \lambda_{\mathrm{NE}} + \lambda_{\mathrm{SE}} \le 1$; in fact, for the above $0 < \Lambda(\nu, \rho) + \Lambda(\nu, -\rho) < 1$ for $\nu > 0$. Moreover, both λ_{NE} and λ_{SE} are decreasing in ν: for a bivariate t_ν copula (Embrechts et al. (2002), Joe (2011)),

$$\lambda_{\mathrm{NE}} = 2T_{\nu+1}\left(-\sqrt{(\nu+1)(1-\rho)/(1+\rho)}\right); \tag{4.10}$$

$$\lambda_{\mathrm{SE}} = 2T_{\nu+1}\left(-\sqrt{(\nu+1)(1+\rho)/(1-\rho)}\right), \tag{4.11}$$

and it can be verified that (4.8) and (4.10) are equivalent, and (4.9) and (4.11) are equivalent (see Section 4.6 for a direct derivation). Since $T_{\nu+1}(x)$ is decreasing in ν for a fixed $x \le 0$, and $T_{\nu+1}(x)$ is increasing in x for a fixed ν, both λ_{NE} and λ_{SE} are decreasing in ν.

From Theorem 5.2 of Schmidt (2002), we know that if there exists a bivariate tail dependent margin, then we must have

$$0 < \liminf_{x\to\infty} \frac{\overline{F}(tx)}{\overline{F}(x)} \leq \limsup_{x\to\infty} \frac{\overline{F}(tx)}{\overline{F}(x)} \leq 1, \quad t \geq 1,$$

where $\overline{F}(\cdot)$ is the survival function of the radial random variable R. If R is slowly varying, then the upper bound in the above necessary condition is reached.

Although we have not proved it for now, we conjecture that there are no tail comonotonic non-degenerate elliptical distributions and the maximum tail dependence parameter for a bivariate elliptical copula is $\lambda = 1/2 + (1/\pi)\arcsin\rho$. As $\nu \to 0^+$ in (4.8) and (4.9), the sum converges to $[1/2 + (1/\pi)\arcsin\rho] + [1/2 + (1/\pi)\arcsin(-\rho)] = 1$ and this is the maximum possible value of $\lim_{x\to\infty}\mathbb{P}(|X_2| > x|X_1 > x)$ when X_1, X_2 have a common distribution. Hence when all corners are considered, elliptical distributions with a slowly variable radial random variable have the strongest possible tail dependence. In contrast, note that Archimedean copulas based on LTs do not have tail dependence on quadrants that have different signs.

4.3.3 Extreme value copulas

We now show that for extreme value copulas, tail comonotonicity is equivalent to comonotonicity.

Proposition 4.10 *Suppose C is an extreme value copula, then*

$$\lambda_U = 1 \iff C \text{ is a comonotonic copula} \iff \lambda_L = 1.$$

4.4 Asymptotic additivity of risk measures

As a dependence structure, tail comonotonicity may affect the risk measures of aggregated losses. In this section, we mainly study impacts of tail comonotonicity on additivity of commonly used risk measure such as VaR and CTE.

It is well known that VaR and CTE are additive when the loss random variables are comonotonic (see Dhaene et al. (2006)); that is, if $(X_1, \ldots, X_d)$ is comonotonic, then for all $p \in (0,1)$,

$$\mathrm{VaR}_p\left(\sum_{i=1}^{d} X_i\right) = \sum_{i=1}^{d} \mathrm{VaR}_p(X_i); \quad \mathrm{CTE}_p\left(\sum_{i=1}^{d} X_i\right) = \sum_{i=1}^{d} \mathrm{CTE}_p(X_i).$$

The additivity property also holds for the upper comonotonicity in the sense of Cheung (2009) when the probability level p associated with the risk measures is larger than a threshold specified by the upper comonotonicity structure. A natural question is whether such an additivity property can be kept asymptotically as $p \to 1^-$.

Asymptotic super and/or sub additivity of risk measures has been studied explicitly or implicitly in several papers. For example, Embrechts et al. (2009b) studied asymptotic additivity properties of VaR for multivariate ($d > 2$) dependent loss random variables that have regularly varying survival functions (with index $-\beta$) and Archimedean dependence structures. It is shown that for a probability level $p < 1$ and sufficiently close to 1, whether *strict* super or sub additivity depends on whether $\beta < 1$ or $\beta > 1$. The Archimedean copula has upper tail dependence but the upper tail dependence parameter $\lambda < 1$ (since the generator of the Archimedean copula is assumed to be regularly varying). In Section 3.3 of Alink et al. (2007), an example of lower tail comonotonic copula has been shown together with Corollary 2.4 of Embrechts et al. (2009b) that VaR is asymptotically additive for exchangeable regularly varying dependent random variables that have an upper tail comonotonic copula.

Analogous to the additivity property of VaR and CTE for a comonotonic random vector, we find that, asymptotic additivity of VaR and CTE still holds for a large class of random vectors that are relevant for quantitative risk management. We now prove some results for asymptotic additivity of VaR and CTE for random variables that are in the maximum domain of attraction (MDA) of Fréchet and Gumbel, respectively. The case for MDA of Weibull corresponds to loss random variables that are bounded above, and is not very relevant for actuarial applications, so we do not consider it here.

Only upper tails of losses are to be studied as they are more relevant for risk measures; analogous results for lower tails also hold but are omitted here. It suffices to consider non-negative random variables to study upper tails, so in the following the random vector $\mathbf{X}$ is assumed to be non-negative. However, for lower tails, we have simpler notation for tail comonotonicity. Therefore, instead of studying $\mathbf{X}$ directly, in what follows, we always define $\mathbf{Y} := -\mathbf{X}$ and prove the results based on $\mathbf{Y}$. The following corresponding assumptions are also assigned on $\mathbf{Y}$, which have corresponding natural meanings for the non-negative random vector $\mathbf{X}$.

For study of tail behavior of random variables, we always assume that the distribution function is continuous (or at least continuous in the tail regions). So $\mathbb{P}[X < x]$ coincides with $\mathbb{P}[X \leq x]$ in what follows.

Assumption 4.1 Let $\mathbf{Y}$ be a non-positive continuous random vector with marginal distributions $F_1, \ldots, F_d$ defined on $(-\infty, 0]$ such that $F_i(-t) \in \mathrm{RV}_{-\alpha}$ with $\alpha > 0$, and

$$\lim_{t\to\infty} F_i(-t)/F_1(-t) = k_i, \quad 0 < k_i < \infty,$$

for any $i = 1, \ldots, d$.

Remark 4.8 If $\mathbf{Y} := -\mathbf{X}$ satisfies Assumption 4.1, thus $\mathbb{P}[X_i > t] \in \mathrm{RV}_{-\alpha}$, then each univariate margin of $\mathbf{X}$ is in the MDA of Fréchet (Theorem 3.3.7 of Embrechts et al. (1997)).

The following lemmas are analogous to Lemma 6.1 of Resnick (2007). They are useful in proving Propositions 4.13 and 4.14. For $E \subset \overline{\mathbb{R}}^d$, denote

$$\mathcal{M}_+(E) := \{\mu : \text{ Radon, non-negative measures on the Borel } \sigma\text{-algebra } \mathcal{E} \text{ of } E\},$$

and let the lower boundary $LB := \{(y_1, \ldots, y_d) : \text{ some } y_i = 0\}$ and the upper boundary $UB := \{(y_1, \ldots, y_d) : \text{ some } y_i = \infty\}$, then define

$$E_1 := [\mathbf{0}, \boldsymbol{\infty}] \setminus LB;$$
$$E_2 := [-\boldsymbol{\infty}, \boldsymbol{\infty}] \setminus UB.$$

Lemma 4.11 *Suppose $\mu_n, \mu \in \mathcal{M}_+(E_1)$. Then, as $n \to \infty$*

$$\mu_n \stackrel{v}{\to} \mu \text{ in } \mathcal{M}_+(E_1) \iff \mu_n([\mathbf{y}, \boldsymbol{\infty}]) \to \mu([\mathbf{y}, \boldsymbol{\infty}])$$

for any $\mathbf{y} \in [\mathbf{0}, \boldsymbol{\infty}) \setminus LB$ such that $\mu(\partial[\mathbf{y}, \boldsymbol{\infty}]) = 0$.

Remark 4.9 The set E_1 excludes the lower boundary, since otherwise rectangles of the form $[\mathbf{y}, \boldsymbol{\infty}]$ can not determine the vague convergence. If μ puts mass on axes such as the usual normalization for asymptotic independence in Section 6.5.1 of Resnick (2007), then $\mu \notin \mathcal{M}_+(E_1)$. The condition $\mu(\partial[\mathbf{y}, \boldsymbol{\infty}]) = 0$ also requires that μ does not put mass on the upper boundary.

Lemma 4.12 *Suppose $\mu_n, \mu \in \mathcal{M}_+(E_2)$. Then, as $n \to \infty$*

$$\mu_n \stackrel{v}{\to} \mu \text{ in } \mathcal{M}_+(E_2) \iff \mu_n([-\boldsymbol{\infty}, \mathbf{y}]) \to \mu([-\boldsymbol{\infty}, \mathbf{y}])$$

for any $\mathbf{y} \in (-\boldsymbol{\infty}, \boldsymbol{\infty}] \setminus UB$ such that $\mu(\partial[-\boldsymbol{\infty}, \mathbf{y}]) = 0$.

Proposition 4.13 *(Asymptotic additivity of VaR: Fréchet case) Suppose $\mathbf{X}$ is non-negative and upper tail comonotonic, and $\mathbf{Y} := -\mathbf{X}$ satisfies Assumption 4.1. If $S = X_1 + \cdots + X_d$, then*

$$\text{VaR}_p(S) \sim \sum_{i=1}^{d} \text{VaR}_p(X_i), \quad p \to 1^-.$$

Assumption 4.2 Let $\mathbf{Y}$ be a non-positive continuous random vector with marginal distributions $F_1, \ldots, F_d$ defined on $(-\infty, 0]$ and there exists a positive measurable function $a(\cdot)$ such that

$$\lim_{t \to \infty} \frac{F_i(-t + a(t)s)}{F_i(-t)} = e^s, \quad \forall s \in \mathbb{R}, \tag{4.12}$$

and $\lim_{t\to\infty} F_i(-t)/F_1(-t) = k_i$ with $0 < k_i < \infty$ for any $i = 1, \ldots, d$.

Remark 4.10 Note that the condition $F_i(-t)/F_1(-t) \to k_i$ implies that we can take $a_i(t) = a(t)$ for all i without loss of generality, and $a(t) = o(t), t \to \infty$. If $\mathbf{Y} := -\mathbf{X}$ satisfies Assumption 4.2, then due to Theorem 3.3.27 of Embrechts et al. (1997), each univariate margin of $\mathbf{X}$ is in the MDA of Gumbel; that is $\overline{F}_{X_i}(t+a(t)s)/\overline{F}_{X_i}(t) \to e^{-s}, s \in \mathbb{R}$.

Remark 4.11 A mixture distribution can satisfy the tail equivalence condition $\overline{F}_{X_i}(t)/\overline{F}_{X_1}(t) \to k_i$. Without loss of generality, we can let $0 < k_i \leq 1$. For example, let random variables $X_1 \sim$ Exponential(1) and

$$X_2 = \begin{cases} \text{Exponential(1)}, & \text{with probability } k_2 \\ 0, & \text{with probability } 1 - k_2. \end{cases}$$

Then $\overline{F}_{X_2}(x)/\overline{F}_{X_1}(x) \to k_2$.

Proposition 4.14 *(Asymptotic additivity of VaR: Gumbel case) Suppose* $\mathbf{X}$ *is non-negative and upper tail comonotonic, and* $\mathbf{Y} := -\mathbf{X}$ *satisfies Assumption 4.2. If* $S = X_1 + \cdots + X_d$*, then*

$$\mathrm{VaR}_p(S) \sim \sum_{i=1}^{d} \mathrm{VaR}_p(X_i), \quad p \to 1^-.$$

The conditions that are studied for asymptotic additivity of VaR satisfy assumptions considered in Asimit et al. (2011), which investigates asymptotic proportionality between CTE and VaR. With the help of their results, we may conclude the asymptotic additivity of CTE in what follows.

Proposition 4.15 *(Asymptotic additivity of CTE: Fréchet case) Suppose* $\mathbf{X}$ *is non-negative and upper tail comonotonic, and* $\mathbf{Y} := -\mathbf{X}$ *satisfies Assumption 4.1 with* $\alpha > 1$*. If* $S = X_1 + \cdots + X_d$*, then*

$$\mathrm{CTE}_p(S) \sim \sum_{i=1}^{d} \mathrm{CTE}_p(X_i), \quad p \to 1^-.$$

Proposition 4.16 *(Asymptotic additivity of CTE: Gumbel case) Suppose* $\mathbf{X}$ *is non-negative and upper tail comonotonic, and* $\mathbf{Y} := -\mathbf{X}$ *satisfies Assumption 4.2. If*

$S = X_1 + \cdots + X_d$, *then*

$$\mathrm{CTE}_p(S) \sim \sum_{i=1}^{d} \mathrm{CTE}_p(X_i), \quad p \to 1^-.$$

The above asymptotic additivity results suggest that when we use tail comonotonicity as the dependence structure for such marginal distributions, we should expect that the diversification benefit (see Definition 2.2) will decrease to 0 as $p \to 1^-$. However, the speed of decay is unknown by observing the asymptotic relationship only; that is, we do not know when p is sufficiently close 1 so that the additivity relationship of the risk measures is sufficiently good. The relevant study will involve a second order approximation, and it is out of the scope of the thesis.

4.5 Concluding remarks and future research

Tail comonotonicity, like comonotonicity and upper comonotonicity, provides a bound-like dependence structure, and it is more reasonable to be used to capture information from data. Tail comonotonicity has some parallel properties of the usual comonotonicity, such as asymptotic additivity of VaR and CTE. Among many different copula families, Archimedean copulas with a mixing distribution that has no moments of any positive orders (for upper tail comonotonicity) or no moments of any negative orders (for lower tail comonotonicity), and copulas based on scale mixtures with a slowly varying non-negative random variable can be used to construct tail comonotonic copulas. However, elliptical and extreme value copulas cannot provide useful tail comonotonic copulas.

Since tail comonotonicity is only an asymptotic property, it may or may not provide a conservative dependence structure for sub-extremal levels of risks. In Chapter 5, we will study how conservative are risk measures, under the tail comonotonic dependence structure.

Although asymptotic additivity properties of VaR and CTE we have proved has already covered a wide range of random vectors, an elegant proof for a most general case (if exists; note that additivity of VaR and CTE for the usual comonotonicity does not depend on margins) must be welcome. Many other topics, such as optimal portfolio design, could be studied under the assumption of tail comonotonicity. We think that tail comonotonicity is particularly useful for dealing with high dimensional data for high-risk scenarios. Due to the curse of dimensionality, it is very hard (or impossible) to accurately model the tail behavior of aggregate high dimensional risks. For dependence modeling of multivariate random variables, vine copula is promising (Bedford and Cooke, 2002; Aas et al., 2009; Joe et al., 2010). By Proposition 4.2, it is feasible to use bivariate tail comonotonic copulas to build up a vine copula that is still tail comonotonic. The vine copula with tail comonotonicity may provide a reasonable cushion for the overall dependence structure.

4.6 Proofs

Proof of Proposition 4.1: If we take $(w_1, \dots, w_d) = (1, \dots, 1)$, then it is obvious that $\lambda = 1$. For the other direction, letting $w^* = \min(w_1, \dots, w_d) \neq 0$, then

$$
\begin{aligned}
&\frac{C(uw^*, \dots, uw^*)}{u} \leq \frac{C(uw_1, \dots, uw_d)}{u} \leq \frac{\min(uw_1, \dots, uw_d)}{u} = w^* \\
&\Rightarrow w^* = \lim_{uw^* \to 0^+} \frac{C(uw^*, \dots, uw^*)}{uw^*} w^* \leq \liminf_{u \to 0^+} \frac{C(uw_1, \dots, uw_d)}{u} \\
&\leq \limsup_{u \to 0^+} \frac{C(uw_1, \dots, uw_d)}{u} \leq w^*.
\end{aligned}
$$

Thus, $b(w_1, \dots, w_d) = \min(w_1, \dots, w_d)$. Similar for the upper tail. $\square$

Proof of Proposition 4.2: By Proposition 4.1, it suffices to prove that $\lambda = 1$. We only state the proof for lower tail comonotonicity here. Suppose $U_i \sim \text{Uniform}(0, 1)$ with joint distribution C. Since

$$
\frac{\mathbb{P}[U_1 \leq u] - \sum_{i=2}^{d} \mathbb{P}[U_1 \leq u < U_i]}{u} \leq \frac{C(u\mathbf{1}_d)}{u} \leq 1,
$$

and pairwise lower tail comonotonicity implies that $\lim_{u \to 0^+} \sum_{i=2}^{d} \mathbb{P}[U_1 \leq u < U_i]/u = 0$, which implies that C is lower tail comonotonic. The other direction is due to the fact that if $\lambda_L(C) = 1$, then $\lambda_L(C_I) = 1$ for every marginal copula $C_I, 1 < |I| < d$. $\square$

Proof of Proposition 4.3: By Proposition 4.1, we only need to prove the tail dependence parameter $\lambda = 1$. For the lower tail, letting $s := \psi^{-1}(u)$, then because $\psi(s) \in \mathrm{RV}_0$,

$$\lambda_L = \lim_{u \to 0^+} \frac{C_\psi(u\mathbf{1}_d)}{u} = \lim_{u \to 0^+} \frac{\psi(d\psi^{-1}(u))}{u} = \lim_{s \to +\infty} \frac{\psi(ds)}{\psi(s)} = 1.$$

For the upper tail, by Proposition 4.2, it suffices to prove the bivariate case. Let $s := \psi^{-1}(1-u)$, then

$$\begin{aligned} \lambda_U = \lim_{u \to 0^+} \frac{\widehat{C}_\psi(u,u)}{u} &= \lim_{u \to 0^+} \left[2 + \frac{\psi(2\psi^{-1}(1-u)) - 1}{u}\right] \\ &= \lim_{s \to 0^+} \left[2 - \frac{1-\psi(2s)}{1-\psi(s)}\right] = 1, \end{aligned}$$

which finishes the proof for the upper tail. □

Proof of Proposition 4.5: Clearly, $-\log(\phi(s)) = (1/\theta)\log(1+s) \in \mathcal{L}^*_{+\infty}$, by Lemma 4.4, $\psi(s)$ is a LT. Then the conclusion is straightforward by plugging $-\log(\phi(s))$ into $\eta(\cdot)$ of (4.4) since $\exp(-[\log(1+s)]^r) \in \mathrm{RV}_0$ as $0 < r < 1$. □

Proof of Proposition 4.6: Superadditivity of $\psi_1^{-1} \circ \psi_2(s)$ for sufficiently large s implies that there exists sufficiently large s_0 such that $s_1 \geq s_0$ and $s_2 \geq s_0$ lead to

$$\psi_1^{-1} \circ \psi_2(s_1 + s_2) \geq \psi_1^{-1} \circ \psi_2(s_1) + \psi_1^{-1} \circ \psi_2(s_2).$$

Let $s_1 := \psi_2^{-1}(u)$ and $s_2 := \psi_2^{-1}(v)$, then $u, v \leq \psi_2(s_0)$ implies that

$$\psi_1^{-1} \circ \psi_2(\psi_2^{-1}(u) + \psi_2^{-1}(v)) \geq \psi_1^{-1} \circ \psi_2(\psi_2^{-1}(u)) + \psi_1^{-1} \circ \psi_2(\psi_2^{-1}(v));$$

that is, $\psi_2(\psi_2^{-1}(u) + \psi_2^{-1}(v)) \leq \psi_1(\psi_1^{-1}(u) + \psi_1^{-1}(v))$, which completes the proof. □

Proof of Lemma 4.7: Choose some $b(y)$, such as $y\ell_0(y)$ for a suitably chosen slowly varying function $\ell_0(y)$, satisfying that as $y \to \infty$, $b(y) \to \infty$, and

$$b^{\alpha+\delta}(y)\overline{F}_R(y) \to \infty, \quad \forall \delta > 0 \tag{4.13}$$

$$b(y)/y \to 0. \tag{4.14}$$

Then,

$$\begin{aligned}
&\lim_{y\to\infty} \frac{\mathbb{P}[RT > y]}{\mathbb{P}[R > y]} \\
&= \lim_{y\to\infty} \frac{\int_0^\infty \overline{F}_R(y/t)F_T(dt)}{\overline{F}_R(y)} \\
&= \lim_{y\to\infty} \frac{\int_0^{b(y)} \overline{F}_R(y/t)F_T(dt)}{\overline{F}_R(y)} + \lim_{y\to\infty} \frac{\int_{b(y)}^\infty \overline{F}_R(y/t)F_T(dt)}{\overline{F}_R(y)} \\
&= \lim_{y\to\infty} \int_0^\infty 1_{(0,b(y)]}(t)\frac{\overline{F}_R(y/t)}{\overline{F}_R(y)}F_T(dt) + \lim_{y\to\infty} \int_{b(y)}^\infty \frac{\overline{F}_R(y/t)}{\overline{F}_R(y)}F_T(dt) \\
&=: \lim_{y\to\infty} I_1 + \lim_{y\to\infty} I_2.
\end{aligned}$$

We know $1_{(0,b(y)]}(t)\frac{\overline{F}_R(y/t)}{\overline{F}_R(y)} \le 1_{(0,1]}(t) + 1_{(1,b(y)]}(t)\frac{\overline{F}_R(y/t)}{\overline{F}_R(y)}$, and also, when $t \in (1, b(y)]$, (4.14) implies that as $y \to \infty$, $y/t \to \infty$. Then as y is sufficiently large, by the Karamata's representation of regularly varying function, when $t \in (1, b(y)]$,

$$\frac{\overline{F}_R(y/t)}{\overline{F}_R(y)} \le t^{\alpha+\delta}.$$

Since $\mathbb{E}[T^{\alpha+\delta}] < \infty$, by the dominated convergence theorem,

$$\lim_{y\to\infty} I_1 = \int_0^\infty \lim_{y\to\infty} 1_{(0,b(y)]}(t)\frac{\overline{F}_R(y/t)}{\overline{F}_R(y)}F_T(dt) = \int_0^\infty t^\alpha F_T(dt) = \mathbb{E}[T^\alpha].$$

For I_2, we have

$$I_2 \le \int_{b(y)}^\infty \frac{1}{\overline{F}_R(y)}F_T(dt) \le \int_{b(y)}^\infty \frac{t^{\alpha+\delta}}{\overline{F}_R(y)b^{\alpha+\delta}(y)}F_T(dt) \to 0, \quad y \to \infty,$$

and the convergence to 0 is due to $\mathbb{E}[T^{\alpha+\delta}] < \infty$ and (4.13). □

Proof of Proposition 4.8: Due to Lemma 4.7, $\overline{F}_{X_1}(s) \sim \overline{F}_{X_i}(s), s \to +\infty$ for $i = 2, \ldots, d$, thus, $s = \overline{F}_{X_i}^{-1}(\overline{F}_{X_i}(s)) \sim \overline{F}_{X_i}^{-1}(\overline{F}_{X_1}(s))$. Since $\overline{F}_{X_1}^{-1}(t) = F_{X_1}^{-1}(1-t)$, clearly,

$$F_{X_i}^{-1}(F_{X_1}(s)) \sim s, \quad s \to +\infty; \quad i = 2, \ldots, d. \tag{4.15}$$

Assuming F is the joint cdf of $(T_1, \ldots, T_d)$, and $h_i(s) := F_{X_i}^{-1}(F_{X_1}(s))/s, i = 1, \ldots, d$, the upper tail dependence parameter

$$\begin{aligned}
\lambda_U &= \lim_{t \to 1^-} \frac{\mathbb{P}\left[X_1 > F_{X_1}^{-1}(t), \ldots, X_d > F_{X_d}^{-1}(t)\right]}{\mathbb{P}\left[X_1 > F_{X_1}^{-1}(t)\right]} \\
&= \lim_{s \to +\infty} \frac{\mathbb{P}\left[RT_1 > s, \ldots, RT_d > F_{X_d}^{-1}(F_{X_1}(s))\right] / \mathbb{P}[R > s]}{\mathbb{P}[RT_1 > s]/\mathbb{P}[R > s]} \\
&= \lim_{s \to +\infty} \frac{\int_{\overline{\mathbb{R}}_+^d} \mathbb{P}\left[R > s\left(\max\{h_1(s)/t_1, \ldots, h_d(s)/t_d\}\right)\right] F(dt_1, \ldots, dt_d)/\mathbb{P}[R > s]}{\mathbb{P}[RT_1 > s]/\mathbb{P}[R > s]}.
\end{aligned} \tag{4.16}$$

Due to Lemma 4.7, $\lim_{s \to +\infty} \mathbb{P}[RT_1 > s]/\mathbb{P}[R > s] = 1$. For the numerator of (4.16), choose some $b(s)$ such as $s\ell_0(s)$ with a proper slowly varying function ℓ_0 such that, $s \to \infty$ implies $b(s) \to \infty$, and

$$b^{\epsilon}(s)\overline{F}_R(s) \to \infty, \quad \forall \epsilon > 0 \tag{4.17}$$

$$b(s)/s \to 0. \tag{4.18}$$

Denote $(b(s)\mathbf{1}_d, +\boldsymbol{\infty}]^c := \overline{\mathbb{R}}^d_+ \setminus (b(s)\mathbf{1}_d, +\boldsymbol{\infty}]$, then

$$
\begin{aligned}
&\lim_{s\to+\infty} \int_{\overline{\mathbb{R}}^d_+} \frac{\mathbb{P}\left[R > s\left(\max\{h_1(s)/t_1, \ldots, h_d(s)/t_d\}\right)\right]}{\mathbb{P}[R > s]} F(dt_1, \ldots, dt_d) \\
&= \lim_{s\to+\infty} \int_{\overline{\mathbb{R}}^d_+} 1_{(b(s)\mathbf{1}_d, +\boldsymbol{\infty}]^c}(t_1, \ldots, t_d) \times \\
&\qquad\qquad \times \frac{\mathbb{P}\left[R > s\left(\max\{h_1(s)/t_1, \ldots, h_d(s)/t_d\}\right)\right]}{\mathbb{P}[R > s]} F(dt_1, \ldots, dt_d) \\
&\quad + \lim_{s\to+\infty} \int_{\overline{\mathbb{R}}^d_+} 1_{(b(s)\mathbf{1}_d, +\boldsymbol{\infty}]}(t_1, \ldots, t_d) \times \\
&\qquad\qquad \times \frac{\mathbb{P}\left[R > s\left(\max\{h_1(s)/t_1, \ldots, h_d(s)/t_d\}\right)\right]}{\mathbb{P}[R > s]} F(dt_1, \ldots, dt_d) \\
&=: \lim_{s\to+\infty} I_1 + \lim_{s\to+\infty} I_2.
\end{aligned}
$$

Due to (4.15), for any $0 < \gamma < 1$, there exists an $s_0 < \infty$ such that, $s > s_0$ implies that $h_i(s) > \gamma$ for all i. Then $s > s_0$ implies that

$$
\begin{aligned}
&1_{(b(s)\mathbf{1}_d, +\boldsymbol{\infty}]^c}(t_1, \ldots, t_d) \frac{\mathbb{P}\left[R > s\left(\max\{h_1(s)/t_1, \ldots, h_d(s)/t_d\}\right)\right]}{\mathbb{P}[R > s]} \\
&\leq 1_{[\mathbf{0}, \gamma\mathbf{1}_d)}(t_1, \ldots, t_d) + 1_{(b(s)\mathbf{1}_d, +\boldsymbol{\infty}]^c \setminus [\mathbf{0}, \gamma\mathbf{1}_d)}(t_1, \ldots, t_d) \times \\
&\qquad\qquad\qquad \times \frac{\mathbb{P}\left[R > s\gamma\left(\max\{1/t_1, \ldots, 1/t_d\}\right)\right]}{\mathbb{P}[R > s]}.
\end{aligned}
$$

It is well known that $\mathbb{P}[R > s]$ is slowly varying implies $\mathbb{P}[R > st]/\mathbb{P}[R > s]$ converges uniformly in $t \in [a, b]$ as $s \to \infty$, where $0 < a, b < \infty$ (see Bingham et al., 1987). Then

$$
\mathbb{P}[R > s\gamma(\max\{1/t_1, \ldots, 1/t_d\})]/\mathbb{P}[R > s]
$$

converges uniformly in $\mathbf{t} \in (b(s)\mathbf{1}_d, +\boldsymbol{\infty}]^c \setminus [\mathbf{0}, \gamma\mathbf{1}_d)$ to 1. Moreover,

$$
\mathbf{t} \in (b(s)\mathbf{1}_d, +\boldsymbol{\infty}]^c \setminus [\mathbf{0}, \gamma\mathbf{1}_d)
$$

and (4.18) implies that, as $s \to +\infty$, $s\gamma \max\{1/t_1, \ldots, 1/t_d\} \to +\infty$. Then slow variation of $\mathbb{P}[R > s]$ and dominated convergence theorem implies that

$$\begin{aligned}
&\lim_{s\to+\infty} I_1 \\
&= \int_{\overline{\mathbb{R}}_+^d} \lim_{s\to+\infty} 1_{(b(s)\mathbf{1}_d,+\boldsymbol{\infty}]^c}(t_1,\ldots,t_d)\times \\
&\quad \times \frac{\mathbb{P}\left[R > s\left(\max\{h_1(s)/t_1,\ldots,h_d(s)/t_d\}\right)\right]}{\mathbb{P}[R > s]} F(dt_1,\ldots,dt_d) \\
&=: \int_{\overline{\mathbb{R}}_+^d} \lim_{s\to+\infty} 1_{(b(s)\mathbf{1}_d,+\boldsymbol{\infty}]^c}(t_1,\ldots,t_d)\Omega(s,t_1,\ldots,t_d)F(dt_1,\ldots,dt_d).
\end{aligned}$$

Since (4.15), for any $0 < \epsilon_1, \epsilon_2 < 1$, there exists an s_1 such that $s > s_1$ implies that $1-\epsilon_1 \le h_i(s) \le 1+\epsilon_2$ for all i. Thus, for any $t_1,\ldots,t_d$ such that $\min\{t_1,\ldots,t_d\} > 0$,

$$\begin{aligned}
1 = \lim_{s\to\infty} \frac{\mathbb{P}[R > s(1+\epsilon_2)/\min\{t_1,\ldots,t_d\}]}{\mathbb{P}[R > s]} &\le \limsup_{s_1 < s\to\infty} \Omega(s,t_1,\ldots,t_d) \\
\le \lim_{s\to\infty} \frac{\mathbb{P}[R > s(1-\epsilon_1)/\min\{t_1,\ldots,t_d\}]}{\mathbb{P}[R > s]} &= 1.
\end{aligned}$$

Therefore,

$$\lim_{s\to\infty} I_1 = 1.$$

For I_2,

$$\begin{aligned}
&1_{(b(s)\mathbf{1}_d,+\boldsymbol{\infty}]}(t_1,\ldots,t_d)\frac{\mathbb{P}\left[R > s\left(\max\{h_1(s)/t_1,\ldots,h_d(s)/t_d\}\right)\right]}{\mathbb{P}[R > s]} \\
&\le 1_{(b(s)\mathbf{1}_d,+\boldsymbol{\infty}]}(t_1,\ldots,t_d)\frac{1}{\mathbb{P}[R > s]} \le 1_{(b(s)\mathbf{1}_d,+\boldsymbol{\infty}]}(t_1,\ldots,t_d)\frac{(\min\{t_1,\ldots,t_d\})^\epsilon}{\mathbb{P}[R > s]b^\epsilon(s)}.
\end{aligned}$$

Note that, $\mathbb{E}[T_i^{\delta_i}] < \infty$ implies that $\mathbb{E}[(\min\{T_1,\ldots,T_d\})^\epsilon] < \infty$ for some $\epsilon > 0$, because $\mathbb{E}[(\min\{T_1,\ldots,T_d\})^{\delta_i}] \le \int_{\overline{\mathbb{R}}_+^d} t_i^{\delta_i} F(dt_1,\ldots,dt_d) = \mathbb{E}[T_i^{\delta_i}] < \infty$. Since $\mathbb{E}[(\min\{T_1,\ldots,T_d\})^\epsilon] < \infty$, and $\mathbb{P}[R > s]b^\epsilon(s) \to +\infty$ (due to (4.17)), by dominated convergence theorem,

$$\lim_{s\to+\infty} I_2 = 0.$$

Thus, the claim is proved. $\square$

Proof of Proposition 4.9: The proof in Theorem 4.3 of Hult and Lindskog (2002) remains valid for this case where R is slowly varying, although in their proof, R is required to be regularly varying with a tail index $-\alpha < 0$. $\square$

Proof of equivalence between (4.8) and (4.10): The cdf of $t_{\nu+1}$ can be written as a regularized Beta function as follows: (Press, 2007, page 323)

$$T_{\nu+1}(t) = \frac{1}{2} - sign(t)\left[\frac{1}{2}\frac{B\left(x; \frac{\nu+1}{2}, \frac{1}{2}\right)}{B\left(\frac{\nu+1}{2}, \frac{1}{2}\right)} - \frac{1}{2}\right], \nu > 0; 0 < x := \frac{\nu+1}{t^2+\nu+1} < 1.$$

where $B(x; ., .)$ and $B(., .)$ are incomplete Beta and Beta functions. Using trigonometric function representation of Beta functions, we have with $t = \sin^2\theta$,

$$\begin{aligned} & B\left(\frac{\nu+1}{2}, \frac{1}{2}\right) \\ &= \int_0^1 t^{(\nu+1)/2-1}(1-t)^{1/2-1}dt = 2\int_0^{\pi/2}(\sin\theta)^\nu d\theta = 2\int_0^{\pi/2}(\cos\theta)^\nu d\theta. \end{aligned}$$

Next, with $B\left(x; \frac{\nu+1}{2}, \frac{1}{2}\right)$, where $t = -\sqrt{\frac{(\nu+1)(1-\varrho)}{1+\varrho}}$ and $x = \frac{\nu+1}{t^2+\nu+1} = (1+\varrho)/2$, then

$$\begin{aligned} B\left(\frac{1+\varrho}{2}; \frac{\nu+1}{2}, \frac{1}{2}\right) &= \int_0^{(1+\varrho)/2} s^{(v-1)/2}(1-s)^{-1/2}ds \\ = 2\int_0^{\arcsin\sqrt{(1+\varrho)/2}}(\sin\theta)^\nu d\theta &= 2\int_{\pi/2-\arcsin\sqrt{(1+\varrho)/2}}^{\pi/2}(\cos\theta)^\nu d\theta. \end{aligned}$$

Now using the identity $\cos(2\theta) = \cos^2\theta - \sin^2\theta$, it can be verified that $2(\pi/2 - \arcsin\sqrt{(1+\varrho)/2}) = \pi/2 - \arcsin\varrho$ by taking "cos" on both sides and getting ϱ. Therefore, the claim is proved. $\square$

Proof of Proposition 4.10: By Proposition 4.2, it suffices to prove the bivariate case. For any multivariate extreme value copula C, there exists a function $A : [0,\infty)^d \to [0,\infty)$ such that $C(u_1, \ldots, u_d) = \exp\{-A(-\log u_1, \ldots, -\log u_d)\}$, where A is convex, homogeneous of order 1 and satisfies that $\max\{x_1, \ldots, x_d\} \leq A(x_1, \ldots, x_d) \leq x_1 + \cdots + x_d$. For a bivariate extreme value copula, the function $A(x, y)$ can be written as $A(x, y) = (x + y)B(\frac{x}{x+y})$, where $B(\cdot)$ is convex and

$\max\{w, 1-w\} \le B(w) \le 1$ for $0 \le w \le 1$ (see Joe, 1997, Theorem 6.4).

The upper tail dependence parameter is $\lambda_U = 2 - A(1,1)$. If we let $\lambda_U = 1$, then $B(1/2) = 1/2$, also $B(w)$ must be convex, so we have $B(w) = \max\{w, 1-w\}$; that is, $A(x,y) = \max\{x,y\}$ and thus

$$C(u_1,u_2) = \exp\{-A(-\log u_1, -\log u_2)\} = \min\{u_1,u_2\}.$$

The other direction is straightforward.

The *lower tail order* for a bivariate extreme value copula is $\kappa = A(1,1)$ (see Example 3.4). If $\lambda_L = 1$, we must have the lower tail order $A(1,1) = 1$. Then the subsequent argument is the same as for the upper tail. $\square$

Proof of Lemma 4.11: The proof is the same as Lemma 6.1 of Resnick (2007). We rewrite the proof here for completeness. $(\Rightarrow)$ is due to Theorem 3.2 of Resnick (2007). For $(\Leftarrow)$, let $g \in \mathcal{C}_K^+(E_1)$, where

$$\mathcal{C}_K^+(E_1) := \{g : E_1 \to \mathbb{R}_+, \text{continuous with compact support}\},$$

then the support of g must be contained in some $[\mathbf{y}, \boldsymbol{\infty}]$ such that $\mu(\partial[\mathbf{y}, \boldsymbol{\infty}]) = 0$. Since convergence on this set holds, $\sup_n \mu_n([\mathbf{y}, \boldsymbol{\infty}]) < \infty$ and thus,

$$\sup_n \mu_n(g) \le \sup_{\mathbf{x}\in E_1} g(\mathbf{x}) \cdot \sup_n \mu_n([\mathbf{y}, \boldsymbol{\infty}]) < \infty.$$

This is true for any $g \in \mathcal{C}_K^+(E_1)$, so $\{\mu_n\}$ is relatively compact due to (3.16) of Resnick (2007). If μ and μ' are two subsequential limits, then μ and μ' agree on the continuity sets $[\mathbf{y}, \boldsymbol{\infty}]$. Additionally, the rectangles of those continuity sets $[\mathbf{y}, \boldsymbol{\infty}]$ constitute the π-system which generates $\mathcal{B}(K^o)$ the Borel σ-algebra of K^o, where $K^o := \{B \subset E_1 : B \text{ is relatively compact }, \mu(\partial B) = 0\}$. Then $\mu' = \mu$ on E_1 by Theorem 3.2 of Resnick (2007) again. $\square$

Proof of Lemma 4.12: The proof here is similar to the proof of Lemma 4.11 by replacing $[\mathbf{y}, \boldsymbol{\infty}]$ by $[-\boldsymbol{\infty}, \mathbf{y}]$. $\square$

Proof of Proposition 4.13: Let $W = -S$ and $Y_i = -X_i, i = 1, \dots, d$, then $(Y_1, \dots, Y_d)$ is lower tail comonotonic, and the map $y \mapsto \mathbb{P}[Y_i \le -y] \in \mathrm{RV}_{-\alpha}$.

Moreover, since $\lim_{t\to\infty} F_i(-t)/F_1(-t) = k_i$ for any $i = 1, \ldots, d$, clearly, $F_i^{-1}(p) \sim k_i^{1/\alpha} F_1^{-1}(p)$ as $p \to 0^+$ (Lemma 2.1 of Asimit et al. (2011), or Proposition 0.8(vi) of Resnick (1987)). It suffices to prove that,

$$F_W^{-1}(p) \sim F_1^{-1}(p) \sum_{i=1}^{d} k_i^{1/\alpha}, \quad p \to 0^+.$$

Also, $\mathbb{P}[W \le -t] \in \mathrm{RV}_{-\alpha}$ (Proposition 7.3 of Resnick (2007)) and apply Proposition 0.8(vi) again. Then it suffices to show as $t \to \infty$,

$$\mathbb{P}[W \le -t] \sim \mathbb{P}[Y_1 \le -t] \left(\sum_{i=1}^{d} k_i^{1/\alpha} \right)^{\alpha}. \tag{4.19}$$

Under Assumption 4.1, if $\mathbf{Y}$ is also lower tail comonotonic, and C is the copula, then we have for any $\mathbf{y} \in [\mathbf{0}, \boldsymbol{\infty}) \setminus LB$,

$$\begin{aligned} &\lim_{t\to\infty} \frac{\mathbb{P}[Y_1 \le -ty_1, \ldots, Y_d \le -ty_d]}{\mathbb{P}[Y_1 \le -t]} \\ &= \lim_{t\to\infty} \frac{C\left(\frac{F_1(-ty_1)}{F_1(-t)} F_1(-t), \ldots, \frac{F_d(-ty_d)}{F_1(-t)} F_1(-t)\right)}{F_1(-t)} = \min\{k_1 y_1^{-\alpha}, \ldots, k_d y_d^{-\alpha}\}, \end{aligned} \tag{4.20}$$

where the last equality is due to lower tail comonotonicity of C and uniform continuity of copula functions (Nelsen, 2006, see). Define a measure μ on E_1 as

$$\mu([y_1, \infty] \times \cdots \times [y_d, \infty]) := \min\{k_1 y_1^{-\alpha}, \ldots, k_d y_d^{-\alpha}\}.$$

This measure μ only puts mass on the line $\{(y_1, \ldots, y_d) : k_1 y_1^{-\alpha} = \cdots = k_d y_d^{-\alpha}\}$. If $k_1 y_1^{-\alpha} = \cdots = k_d y_d^{-\alpha} =: h$, then $\mu([y_1, \infty] \times \cdots \times [y_d, \infty]) = h$. Clearly, μ is Radon and every set $[\mathbf{y}, \boldsymbol{\infty}]$ is a continuity set. Define

$$\mu_t(\cdot) := \frac{\mathbb{P}\left[(-Y_1/t, \ldots, -Y_d/t) \in \cdot\right]}{\mathbb{P}[Y_1 \le -t]},$$

then by Lemma 4.12 and (4.20),

$$\mu_t(\cdot) \xrightarrow{v} \mu(\cdot), \quad t \to \infty.$$

where $\xrightarrow{v}$ is vague convergence. Define a set

$$H := \left(\mathbf{y} \in [\mathbf{0}, \boldsymbol{\infty}) \setminus LB : \sum_{i=1}^{d} y_i \geq 1 \right).$$

Then clearly, $\mu(\partial H) = 0$ and H is relatively compact. Thus by Theorem 3.2 of Resnick (2007),

$$\mu_t(H) \to \mu(H), \quad t \to \infty.$$

That is, letting $z_i := k_i y_i^{-\alpha}$, as $t \to \infty$,

$$\begin{aligned} \frac{\mathbb{P}[W \leq -t]}{\mathbb{P}[Y_1 \leq -t]} = \mu_t(H) \to \min\Big\{ k_1 y_1^{-\alpha}, \ldots, k_d y_d^{-\alpha} : \sum_{i=1}^{d} y_i \geq 1 \Big\} \\ = \mu_{Lebesgue}\left(z : \sum_{i=1}^{d} (z/k_i)^{-1/\alpha} \geq 1 \right) = \left(\sum_{i=1}^{d} k_i^{1/\alpha} \right)^{\alpha}, \end{aligned}$$

which justifies (4.19), thus finishing the proof. □

Proof of Proposition 4.14: Let $W = -S$ and $Y_i = -X_i, i = 1, \ldots, d$, then $(Y_1, \ldots, Y_d)$ is lower tail comonotonic. Moreover, since

$$\lim_{t \to \infty} F_i(-t + a(t)s)/F_1(-t) = k_i e^s$$

for any $s \in \mathbb{R}$ and $i = 1, \ldots, d$, taking $s = -\log k_i$ leads to

$$\mathrm{VaR}_p(Y_i) \sim \mathrm{VaR}_p(Y_1) - a(-\mathrm{VaR}_p(Y_1)) \log k_i, \quad p \to 0^+ \tag{4.21}$$

Then it suffices to show as $t \to \infty$,

$$\mathbb{P}\left[W \leq -dt - a(t)\sum_{i=1}^{d}\log k_i\right] \sim \mathbb{P}[Y_1 \leq -t]. \tag{4.22}$$

Under Assumption 4.2, if $\mathbf{Y}$ is also lower tail comonotonic, and C is the copula, then we have for any $\mathbf{y} \in (-\boldsymbol{\infty}, \boldsymbol{\infty}] \setminus UB$,

$$\begin{aligned}
&\lim_{t\to\infty} \frac{\mathbb{P}[Y_i \leq -t - a(t)\log k_i + a(t)y_i, i = 1, \ldots, d]}{\mathbb{P}[Y_1 \leq -t]} \\
&= \lim_{t\to\infty} \frac{C\left(\frac{F_i(-t-a(t)\log k_i + a(t)y_i)}{F_1(-t)} \cdot F_1(-t), i = 1, \ldots, d\right)}{F_1(-t)} \\
&= \lim_{t\to\infty} \frac{C\left(k_i e^{y_i - \log k_i} F_1(-t), i = 1, \ldots, d\right)}{F_1(-t)} = \min\{e^{y_1}, \ldots, e^{y_d}\},
\end{aligned} \tag{4.23}$$

where the last equality is due to lower tail comonotonicity of C and uniform continuity of copula functions (Nelsen, 2006). Define a measure μ on E_2 as

$$\mu([-\infty, y_1] \times \cdots \times [-\infty, y_d]) := \min\{e^{y_1}, \ldots, e^{y_d}\}.$$

This measure μ only puts mass on the line $\{(y_1, \ldots, y_d) : e^{y_1} = \cdots = e^{y_d}\}$. If $e^{y_1} = \cdots = e^{y_d} =: h$, then $\mu([-\infty, y_1] \times \cdots \times [-\infty, y_d]) = h$. Clearly, μ is Radon and every set $[-\boldsymbol{\infty}, \mathbf{y}]$ is a continuity set. Define

$$\mu_t(\cdot) := \frac{\mathbb{P}\left[((Y_1 + t)/a(t) + \log k_1, \ldots, (Y_d + t)/a(t) + \log k_d) \in \cdot\right]}{\mathbb{P}[Y_1 \leq -t]},$$

then by Lemma 4.12 and (4.23),

$$\mu_t(\cdot) \xrightarrow{v} \mu(\cdot), \quad t \to \infty.$$

where $\xrightarrow{v}$ is vague convergence. Define a set

$$H := \left(\mathbf{y} \in (-\boldsymbol{\infty}, \boldsymbol{\infty}] \setminus UB : \sum_{i=1}^{d} y_i \leq 0\right).$$

Then clearly, $\mu(\partial H) = 0$ and H is relatively compact. Thus by Theorem 3.2 of Resnick (2007),

$$\mu_t(H) \to \mu(H), \quad t \to \infty.$$

That is, letting $z_i := e^{y_i}$, as $t \to \infty$,

$$\frac{\mathbb{P}[W \leq -dt - a(t)\sum_{i=1}^{d} \log k_i]}{\mathbb{P}[Y_1 \leq -t]} = \mu_t(H)$$
$$\to \min\left\{e^{y_1}, \ldots, e^{y_d} : \sum_{i=1}^{d} y_i \leq 0\right\} = \mu_{Lebesgue}\left(z > 0 : \sum_{i=1}^{d} \log z \leq 0\right) = 1.$$

which justifies (4.22) and asymptotic additivity holds. □

Proof of Proposition 4.15: It is well known that under Assumption 4.1 with $\alpha > 1$, $\text{CTE}_p(X_i) \sim \frac{\alpha}{\alpha-1}\text{VaR}_p(X_i), p \to 1^-$, for $i = 1, \ldots, d$ (e.g., Zhu and Li, 2012). By Theorem 2.1 in Asimit et al. (2011), and Proposition 4.13,

$$\sum_{i=1}^{d} \text{CTE}_p(X_i) \sim \frac{\alpha}{\alpha - 1} \sum_{i=1}^{d} \text{VaR}_p(X_i) \sim \frac{\alpha}{\alpha - 1}\text{VaR}_p(S) \sim \text{CTE}_p(S), \quad p \to 1^-,$$

which completes the proof. □

Proof of Proposition 4.16: It is well known that under Assumption 4.2, $\text{CTE}_p(X_i) \sim \text{VaR}_p(X_i), p \to 1^-$ (e.g., Section 3.3.3 of Embrechts et al. (1997)). Also, $\text{CTE}_p(S) \sim \text{VaR}_p(S), p \to 1^-$ (see (2.15) of Asimit et al. (2011)). Therefore, by Proposition 4.14,

$$\text{CTE}_p(S) \sim \text{VaR}_p(S) \sim \sum_{i=1}^{d} \text{VaR}_p(X_i) \sim \sum_{i=1}^{d} \text{CTE}_p(X_i), \quad p \to 1^-,$$

which completes the proof. □

Chapter 5

Tail dependence and conservativity

5.1 Introduction

Dependence modeling with copulas for multivariate random losses plays an important role in accounting for a nonlinear dependence structure in insurance and financial data. For example, the dependence between losses and associated expenses often appears to be asymmetric in upper and lower tails for auto insurance claim data (Frees and Valdez, 1998).

There are several important statistical issues for dependence modeling with copulas, such as statistical inference, goodness-of-fit testing and model selection. Among many others, we refer to Joe (1997), Genest and Favre (2007), Genest et al. (2009b) and Brechmann et al. (2012) for statistical issues in copula modeling.

When introducing various statistical methodologies into the actuarial community, an important question that should not be avoided is the sensitivity of commonly used risk measures to statistical modeling. As we know, nice properties of statistical methodologies are often derived based on certain nice assumptions. Moreover, there are not sufficient approaches that are tailored for tail risks. It is very challenging for a statistical method to accurately account for tail risks, especially when data in the tail are sparse and rare events leading to large losses happen unexpectedly. Due to the nature of the insurance industry, financial security is particularly important. Therefore, in practice, actuaries often need to consider more conservative modeling when there is less certainty, and conduct scenario testing accordingly.

In Chapter 4, we have studied a concept of tail comonotonicity, which can be a property for absolutely continuous copulas so it is not as restrictive as the comonotonic copula. Upper tail comonotonic copulas have an upper tail dependence parameter of 1 in the sense of the usual tail dependence (Joe, 1997, section 2.1.10), and include the comonotonic copula as well as copulas satisfying the properties of upper comonotonicity (Cheung, 2009) as special cases; we refer to Dhaene et al. (2002a,b) for the former concept, and Dong et al. (2010) and Nam et al. (2011) for further study of the latter concept. For upper comonotonicity, there is a constraint of comonotonicity beyond a certain threshold. But with tail comonotonicity, the constraint applies only for the limit and the data can contribute to the likelihood for the whole support of the copula. For joint non-functional-related losses, tail comonotonicity is a reasonably conservative assumption. The benefit of tail comonotonicity is that it introduces a balance between accuracy and security for assessing dependent tail risks; the less information we have and the larger losses we care about, the more conservative assumptions are adopted. We refer to Chapter 4 for more relevant research on tail comonotonicity.

If X_1 and X_2 are dependent losses, two fundamental conditional specifications $\mathbb{E}[X_1|X_2 > t]$ and $\mathbb{E}[X_1|X_2 = t]$ that are relevant to quantitative risk management are the emphasis of this chapter. As $t \to \infty$, for fixed margins, dependence structures will play a role in ordering the above conditional expectations. Sufficient conditions that lead to such a comparison have been derived for bivariate copulas; in particular, for Archimedean copula families, tail behavior of their generators may affect the asymptotic orders, and tail comonotonicity is proved to be conservative in term of the two conditional specifications. For tail comonotonic copulas, second order conditions determine the degree of conservativity, and we propose a way to compare the second order conditions. Other than the two conditional specifications, tail comonotonicity is also conservative in terms of diversification effects, thus we refer to tail comonotonicity as an asymptotically worst dependence structure.

Simulations are conducted to further understand the conditions and their influence on risk measures; tail comonotonicity is shown to be conservative as high-risk scenarios are considered. Finally, an absolutely continuous copula model with tail comonotonicity is applied to an auto insurance claim data; it was used to assess the magnitude of associated expenses based on the information of loss.

This chapter is organized as following: Section 5.2 reports results on the two conditional specifications and second order properties of tail comonotonicity. The asymptotically worst dependence structure is briefly introduced in Section 5.3. Some findings of the simulation study are presented in Section 5.4, and a data analysis with tail comonotonicity is conducted in Section 5.5. Section 5.6 concludes with discussions, and the proofs are collected in Section 5.7.

5.2 Conditional specifications

For risk management in actuarial science or quantitative finance, one often concerns the influence of risks from an individual asset over the whole portfolio. If $(X_1, \ldots, X_d)$ are unbounded dependent losses, then conditional tail expectation of the forms $\mathbb{E}[\sum_{i=1}^d X_i | X_k > t]$ or $\mathbb{E}[\sum_{i=1}^d X_i | X_k = t]$ can be used as relevant risk measures, where t is usually a high quantile of X_k. Due to linearity of probability expectations, it suffices to study the conditional tail expectation of the forms $\mathbb{E}[X_i | X_k > t]$ or $\mathbb{E}[X_i | X_k = t]$ for a given pair of $i, k \in \{1, \ldots, d\}$. For notational convenience, they are written as the following two forms:

$$\mathbb{E}[X_1 | X_2 > t]; \tag{5.1}$$

$$\mathbb{E}[X_1 | X_2 = t]. \tag{5.2}$$

Moreover, these forms of conditional specifications may also be used to model a univariate time series, such as, $\mathbb{E}[X_{i+1} | X_i > t]$ where copulas can be used to model the serial dependence between X_i and X_{i+1} (Patton, 2009).

The study of these two forms in the literature has been mainly on the influence of dependence and asymptotic analysis for the threshold t in (5.1) or (5.2) goes to infinity. For example, Landsman and Valdez (2003) has a formula for (5.1) when (X_1, X_2) follows a bivariate elliptical distribution; Zhu and Li (2012), and Chapter 6 conduct asymptotic analysis under the condition of (second order) regular variation.

In this section, we will investigate the effects of dependence structures between X_1 and X_2 on these two conditional specifications: $\mathbb{E}[X_1|X_2 > t]$ and $\mathbb{E}[X_1|X_2 = t]$ as $t \to +\infty$. In particular, we will study how conservative tail comonotonicity is for modeling dependent risks. From Chapters 3 and 4, we know conditions for Archimedean copulas to be tail comonotonic or satisfy other upper/lower tail properties. Therefore, we illustrate general results in the special case of Archimedean copulas. In order to illustrate the conservativity of tail comonotonicity, we will focus on the bivariate case although tail comonotonicity is a multivariate concept. For multivariate extensions, the conditional specification can be of the following forms: $\mathbb{E}[X_1|X_2 > t_1, \ldots, X_d > t_d]$ or $\mathbb{E}[X_1|X_2 = t_1, \ldots, X_d = t_d]$. These two forms may be useful in quantitative risk management for portfolios or constructing models for multiple regression analysis. However, these multivariate extensions are harder to analyze and are left for future study.

In this section, X_1, X_2, Y_1 and Y_2 are assumed to be absolutely continuous random variables defined on $[0, \infty)$, $X_i, Y_i \sim F_i, i = 1, 2$, written as $\mathbf{X}, \mathbf{Y} \in \mathcal{R}(F_1, F_2)$, the Fréchet space of all bivariate random vectors whose marginal distributions are F_1 and F_2, respectively, and F_1, F_2 are absolutely continuous and defined on $[0, \infty)$. Furthermore, $\mathbb{E}[X_i] = \mathbb{E}[Y_i] < \infty, i = 1, 2$. Then sufficient conditions that lead to the following two asymptotic inequalities will be studied.

$$\text{Relation (I):} \quad \mathbb{E}[X_1|X_2 > t] \ \gtrsim\ \mathbb{E}[Y_1|Y_2 > t], \quad t \to \infty; \tag{5.3}$$

$$\text{Relation (II):} \quad \mathbb{E}[X_1|X_2 = t] \ \gtrsim\ \mathbb{E}[Y_1|Y_2 = t], \quad t \to \infty. \tag{5.4}$$

In our analysis of the above two conditions, the limits of the conditional distributions of bivariate copulas will appear, specifically $C_{1|2}(u|1)$ and $C_{1|2}(u|0)$ for $0 < u < 1$. If $(U, V) \sim C$, these are the limiting distributions of $[U|V = v]$ as $v \to 1^-$ or $v \to 0^+$. These limiting conditional distributions appear in the tail dependence results in Cooke et al. (2011).

With this form of dependence in the tails, $C_{1|2}(u|1) = 0$ for $0 \leq u < 1$ or $[U|V = 1]$ degenerates at 1 is the strongest upper tail dependence and $C_{1|2}(u|0) = 1$ for $0 < u \leq 1$ or $[U|V = 0]$ degenerates at 0 is the strongest lower tail dependence. Note that, this tail dependence form is different from the widely-used form of tail dependence defined in Joe (1997, section 2.1.10); we refer to the latter as "usual tail dependence" in this chapter. Next consider the special case of a bivariate Archimedean copula constructed by a LT ψ. The condition $C_{1|2}(u|1) = 0$ for $0 \le u < 1$ occurs when $\psi'(0) = -\infty$; this is also the condition to achieve the usual upper tail dependence. The condition $C_{1|2}(u|0) = 1$ for $0 < u \le 1$ occurs as $0 < r < 1$ under the mild assumption $\psi(s) \sim T(s) = a_1 s^q \exp\{-a_2 s^r\}$ with $a_1 > 0$, $a_2 \ge 0$, and $\psi'(s) = T'(s)$, as $s \to \infty$; this matches conditions to achieve either lower tail dependence or intermediate lower tail dependence, i.e., a lower tail order of $1 \le \kappa < 2$.

5.2.1 The case: $\mathbb{E}[X_1|X_2 > t]$

Proposition 5.1 *Suppose* $\mathbf{X} := (X_1, X_2)$, $\mathbf{Y} := (Y_1, Y_2)$ *and* $\mathbf{X}, \mathbf{Y} \in \mathcal{R}(F_1, F_2)$. *(a) If* $\mathbf{X} \sim C^*(F_1, F_2)$, $\mathbf{Y} \sim C(F_1, F_2)$ *and* $\lim_{v \to 1^-}[C^*_{1|2}(u|v)/C_{1|2}(u|v)] < 1$ *for all* $0 < u < 1$, *then Relation (I) holds. (b) If* $\widehat{C}^*$ *and* $\widehat{C}$ *are survival copulas of* C^* *and* C, $\mathbf{X} \sim \widehat{C}^*(F_1, F_2)$, $\mathbf{Y} \sim \widehat{C}(F_1, F_2)$, *and* $\lim_{v \to 0^+}[1 - C^*_{1|2}(u|v)]/[1 - C_{1|2}(u|v)] < 1$ *for all* $0 < u < 1$, *then Relation (I) holds.*

For constructing Relations (I) and (II) for Archimedean copulas, the upper and lower tail behavior of the following function is critical:

$$\Upsilon(u) := \frac{\varphi'(\varphi^{-1}(u))}{\psi'(\psi^{-1}(u))}, \quad u \in [0, 1], \tag{5.5}$$

where ψ and φ are LTs for the copulas or survival copulas of $\mathbf{X}$ and $\mathbf{Y}$. Note that $\Upsilon(0)$ and $\Upsilon(1)$ are defined as the corresponding limits, provided that the limits exist.

Corollary 5.2 *Suppose* $\mathbf{X} := (X_1, X_2)$, $\mathbf{Y} := (Y_1, Y_2)$, $\mathbf{X}, \mathbf{Y} \in \mathcal{R}(F_1, F_2)$, *and their dependence structures are* C^* *and* C *that are bivariate Archimedean copulas constructed as in (2.4) with LTs* ψ *and* φ, *respectively. If* $\Upsilon(1)$ *exists and* $\Upsilon(u) > \Upsilon(1)$ *for any* $0 < u < 1$, *then Relation (I) holds.*

Remark 5.1 The condition (5.12) in the proof of Proposition 5.1 holds if C^* is larger than C in the PQD or concordance ordering (Joe, 1997; Shaked and Shanthikumar, 2007). The concordance ordering holds for copulas from a single commonly used 1-parameter bivariate copula family, but Proposition 5.1 and Corollary 5.2 can also be applied to two copulas from two different parametric families, as shown in Example 5.1.

Remark 5.2 The condition $\Upsilon(u) > \Upsilon(1)$ is the same as $\xi'(s) > \xi'(0)$, where $\xi = \psi^{-1} \circ \varphi$. For concordance ordering of the two Archimedean copulas, a sufficient condition is ξ convex or $\xi'(s)$ increasing in s (Joe, 1997, Corollary 4.2). So the tail comparison here is satisfied with the weaker condition.

Example 5.1 (a) (bivariate MTCJ vs Gumbel). The MTCJ copula is based on the gamma LT $(1 + s)^{-1/\delta_M}$ for $\delta_M > 0$ and the Gumbel copula is based on the positive stable LT $\exp\{-s^{-1/\delta_G}\}$ for $\delta_G > 1$. Suppose the parameters for MTCJ and Gumbel copulas are δ_M and δ_G, respectively, and the LTs are φ and ψ, respectively. Then $\Upsilon(u) = (\delta_G/\delta_M)u^{\delta_M}(-\log u)^{\delta_G - 1}$, and $\Upsilon(u) > \Upsilon(1) = 0$ for any $0 < u < 1$. So if (X_1, X_2) has an arbitrary Gumbel copula and (Y_1, Y_2) has an arbitrary MTCJ copula, and $X_i, Y_i \sim F_i$, $i = 1, 2$, Relation (I) must hold. This case is not surprising as the Gumbel copula has upper tail dependence for any $\delta_G > 1$ and the MTCJ has upper tail order of 2 (upper tail independence) for any $\delta_M > 0$. That is, the Gumbel copula has more probability in the upper tail and $\overline{C}^*(u, v; \delta_G) \geq \overline{C}(u, v; \delta_M)$ or equivalently $C^*(u, v; \delta_G) \geq C(u, v; \delta_M)$ for (u, v) in an upper set S_{δ_G, δ_M} that can include most of $[0, 1] \times [v_0, 1]$ for some v_0 near 1.

(b) (Joe/B5 vs Gumbel). The Joe/B5 copula is based on the Sibuya LT $1 - (1 - e^{-s})^{1/\delta_J}$ for $\delta_J > 1$. This also has upper tail dependence so the comparison with Gumbel requires a careful check of the conditions in Proposition 5.1. Let C^{Joe} and C^{Gum} denote these two copula families. Using the conditional distributions on page 147 of Joe (1997), as $v \to 1^-$,

$$\frac{C_{1|2}^{Joe}(u|v; \delta_J)}{C_{1|2}^{Gum}(u|v; \delta_G)} \sim \frac{(1 - u)^{-\delta_J + 1}\{1 - (1 - u)^{\delta_J}\}}{u(-\log u)^{-\delta_G + 1}} \cdot (1 - v)^{\delta_G - \delta_J}.$$

The limit is 0 if $\delta_G > \delta_J$ and ∞ if $\delta_G < \delta_J$ and finite if $\delta_G = \delta_J$. In the latter case, the ratio is shown numerically to be > 1 for $0 < u < 1$. Therefore if $\delta_G = \delta_J = \delta > 1$ and $C^* = C^{Gum}(; \delta)$, $C = C^{Joe}(; \delta)$, $(X_1, X_2) \sim C^*(F_1, F_2)$, $(Y_1, Y_2) \sim C(F_1, F_2)$, then Relation (I) must hold.

For applications in insurance and finance, we often need to consider a copula that has upper tail dependence. In Table 5.1, we summarize the comparison in the sense of Corollary 5.2 for some commonly used upper tail dependent copulas.

Table 5.1: Comparisons between Gumbel, Joe, BB1 and BB7 copula families

Copula	Parameter(s)	$\psi_\delta'(\psi_\delta^{-1}(u))$	Domination term as $u \to 1^-$
Gumbel	$1 \le \delta$	$-\frac{1}{\delta}u(-\log u)^{1-\delta}$	$(-\log u)^{1-\delta} \sim (1-u)^{1-\delta}$
Joe	$1 \le \delta$	$-\frac{1}{\delta}(1-u)^{1-\delta}[1-(1-u)^\delta]$	$(1-u)^{1-\delta}$
BB1	$1 \le \delta; 0 < \theta$	$-\frac{1}{\theta\delta}u^{\theta+1}(u^{-\theta}-1)^{1-\delta}$	$(u^{-\theta}-1)^{1-\delta} \sim \theta^{1-\delta}(1-u)^{1-\delta}$
[1] BB7	$1 \le \delta; 0 < \theta$	$-\frac{1}{\theta\delta}(1-u)^{1-\delta}[1-(1-u)^\delta]^{1+\theta}$	$(1-u)^{1-\delta}$

1. The notation of δ, θ is exchanged from their original form in Joe (1997).

Note that the upper tail dependence parameters for all of the copulas in Table 5.1 are $2 - 2^{1/\delta}$. The following result tells us how to use Table 5.1. It includes part (b) of Example 5.1 as a special case.

Corollary 5.3 *Suppose* $\mathbf{X} := (X_1, X_2)$*,* $\mathbf{Y} := (Y_1, Y_2)$*,* $\mathbf{X}, \mathbf{Y} \in \mathcal{R}(F_1, F_2)$*, and their dependence structures are copulas* $C^*, C \in \{Gumbel, Joe, BB1, BB7\}$*. If* $\delta^* > \delta$*, that is, the upper tail dependence parameters* $\lambda^* > \lambda$*, then Relation (I) holds.*

For modeling dependent risks, we often need to reflect a copula to get its survival copula. Now we consider the condition in the sense of Corollary 5.2 when survival Archimedean copulas are used to model the risks.

Proposition 5.4 *Suppose* $\mathbf{X} := (X_1, X_2)$*,* $\mathbf{Y} := (Y_1, Y_2)$*,* $\mathbf{X}, \mathbf{Y} \in \mathcal{R}(F_1, F_2), \mathbf{X} \sim \widehat{C}^*(F_1, F_2)$ *and* $\mathbf{Y} \sim \widehat{C}(F_1, F_2)$*, where* C^* *and* C *are Archimedean copulas constructed as in (2.4) with LTs* ψ *and* φ*, respectively. If* $\psi^{-1} \circ \varphi(s)$ *is strictly convex for large* s*, and in particular, if* $\Upsilon(u)$ *is strictly decreasing in* u *near* 0*, then Relation (I) holds.*

Remark 5.3 The above result is on the ordering in the lower tail of two copulas. For the concordance ordering of Archimedean copulas C_ψ and C_φ, a sufficient condition is that $\psi^{-1} \circ \varphi$ is superadditive (Joe, 1997, Theorem 4.1), which holds if $\psi^{-1} \circ \varphi$ is convex due to $\psi^{-1}(\varphi(0)) = 0$ (Marshall and Olkin, 2007, Proposition

A.11). Obviously, when $\psi^{-1} \circ \varphi$ is superadditive, Relation (I) must hold. But the condition in Proposition 5.4 only requires the superadditive inequality where one argument is large and the other is arbitrary, and this can be checked via the behavior of $\Upsilon(u)$ as $u \to 0^+$.

From Corollary 5.2 and Proposition 5.4, the tail behavior of $\psi'(\psi^{-1}(u))$ for $u \to 1^-$ and $u \to 0^+$ determine the upper and lower tail association in the sense of Relation (I). So we list in Table 5.2 the conditions for commonly used 1- or 2-parameter Archimedean copulas that appear in Joe (1997). By looking at the conditions listed in Table 5.2, we can easily compare the conditional tail expectations in the sense of Proposition 5.4 (or Corollary 5.2). Those conditions will also be used in the next subsection for establishing Relation (II).

Example 5.2 (BB3 vs BB2) As $u \to 0^+$, observed from Table 5.2, the term $e^{-\delta(u^{-\theta}-1)}$ for BB2 and the term $e^{-\delta(-\log u)^\theta}$ for BB3 dominate the condition in Proposition 5.4 for comparing these two copulas. It is clear that the former term goes to 0 much faster than the latter as $u \to 0^+$. Therefore, if C^* is BB2 and C is BB3, and $\widehat{C}^*$ and $\widehat{C}$ are copulas for $\mathbf{X}$ and $\mathbf{Y}$ respectively, then Relation (I) holds. It is also consistent to the fact that BB2 is more lower tail positive dependent than BB3 below a sufficiently low threshold, although they are both lower tail comonotonic (see Chapter 4).

Table 5.2 shows that $\psi'(\psi^{-1}(0)) \equiv 0$ for all of the Archimedean copula families listed, which reminds us of the following fact: suppose $\psi(s)$ is the LT of a positive random variable, then $\psi^{(n)}(s) \to 0$ as $s \to \infty$ for $n = 1, 2, \dots$.

The following lemmas are the Monotone Density Theorem (Bingham et al., 1987; Embrechts et al., 1997) applied to slowly varying functions.

Lemma 5.5 *Let $\ell \in \mathrm{RV}_0$ and $\ell(s) = \int_0^s \ell'(u)du$ (or $\ell(s) = \int_s^\infty \ell'(u)du$). If $\ell'(s)$ is ultimately monotone as $s \to \infty$, then $\ell'(s) = o(s^{-1}\ell(s)), s \to \infty$.*

Lemma 5.6 *Let $\ell \in \mathrm{RV}_0(0^+)$ and $\ell(s) = \int_0^s \ell'(u)du$. If $\ell'(s)$ is ultimately monotone as $s \to 0^+$, then $\ell'(s) = o(s^{-1}\ell(s)), s \to 0^+$.*

Table 5.2: Condition $\psi'(\psi^{-1}(u))$ for parametric Archimedean copula families

Copula	Parameter(s)	$\psi'(\psi^{-1}(u))$
Frank(B3)	$0 \le \delta$	$(1/\delta)(1-e^{\delta u})$
MTCJ(B4)	$0 \le \delta$	$-(1/\delta)u^{1+\delta}$
Joe(B5)	$1 \le \delta$	$-(1/\delta)(1-u)^{1-\delta}[1-(1-u)^{\delta}]$
Gumbel(B6)	$1 \le \delta$	$-(1/\delta)u(-\log u)^{1-\delta}$
BB1	$1 \le \delta; 0 < \theta$	$-(1/(\theta\delta))u^{\theta+1}(u^{-\theta}-1)^{1-\delta}$
BB2	$0 < \delta, \theta$	$-(1/(\theta\delta))u^{1+\theta}e^{-\delta(u^{-\theta}-1)}$
BB3	$0 < \delta; 1 \le \theta$	$-(1/(\theta\delta))u(-\log u)^{1-\theta}e^{-\delta(-\log u)^{\theta}}$
BB6	$1 \le \delta, \theta$	$-(1/(\theta\delta))(1-u)^{1-\theta}[1-(1-u)^{\theta}][-\log[1-(1-u)^{\theta}]]^{1-\delta}$
[1] BB7	$1 \le \delta; 0 < \theta$	$-(1/(\theta\delta))(1-u)^{1-\delta}[1-(1-u)^{\delta}]^{1+\theta}$
BB8	$0 < \delta \le 1; 1 \le \theta$	$-(1/(\theta\delta))(1-\delta u)^{1-\theta}+(1/(\theta\delta))(1-\delta u)$
BB9	$0 \le \alpha; 1 \le \theta$	$-(1/\theta)u(\alpha-\log u)^{1-\theta}$
BB10	$0 < \alpha; 0 \le \theta \le 1$	$-\alpha u-\alpha\theta(1-\theta)^{-1}u^{1+1/\alpha}$

1. The notation of δ, θ is exchanged from their original form in Joe (1997).

The following two results establish that tail comonotonicity is conservative in the sense of Relation (I). By Proposition 4.3, the condition $1-\psi \in \mathrm{RV}_0(0^+)$ in Corollary 5.7 implies that the Archimedean copula C_ψ is upper tail comonotonic, and the condition $\psi \in \mathrm{RV}_0$ in Corollary 5.8 implies that the Archimedean copula C_ψ is lower tail comonotonic.

Corollary 5.7 *Suppose $\mathbf{X} := (X_1, X_2)$, $\mathbf{Y} := (Y_1, Y_2)$, $\mathbf{X}, \mathbf{Y} \in \mathcal{R}(F_1, F_2)$, and their dependence structures are C^* and C that are Archimedean copulas constructed as in (2.4) with LTs ψ and φ, respectively. If $1-\psi \in \mathrm{RV}_0(0^+)$ and $\liminf_{s\to 0^+}(-s\varphi'(s))/(1-\varphi(s)) > 0$, then Relation (I) holds.*

Corollary 5.8 *Suppose $\mathbf{X} := (X_1, X_2)$, $\mathbf{Y} := (Y_1, Y_2)$, $\mathbf{X}, \mathbf{Y} \in \mathcal{R}(F_1, F_2), \mathbf{X} \sim \widehat{C}^*(F_1, F_2)$ and $\mathbf{Y} \sim \widehat{C}(F_1, F_2)$, where C^* and C are Archimedean copulas constructed as in (2.4) with LTs ψ and φ, respectively. If $\Upsilon(u)$ is ultimately monotone as $u \to 0^+$, $\psi \in \mathrm{RV}_0$ and $\liminf_{s\to\infty}(-s\varphi'(s))/(\varphi(s)) > 0$, then Relation (I) holds.*

Remark 5.4 The conditions

$$\liminf_{s\to 0^+}(-s\varphi'(s))/(1-\varphi(s)) > 0 \quad \text{(for copula upper tail)},$$
$$\liminf_{s\to\infty}(-s\varphi'(s))/\varphi(s) > 0 \quad \text{(for copula lower tail)}$$

cover tail dependence, intermediate tail dependence and tail orthant independence cases. For instance, for the case of lower tail, let $\pi := -\lim_{s\to\infty}(s\varphi'(s))/\varphi(s)$, then $\pi = 0$ is tail comonotonicity, $0 < \pi < \infty$ corresponds to tail dependence and $\pi = \infty$ leads to intermediate tail dependence or near independence (tail order $\kappa = d$). We refer to Charpentier and Segers (2009) and Chapter 3 for a comprehensive study of tail behavior of Archimedean copulas.

5.2.2 The case: $\mathbb{E}[X_1|X_2 = t]$

In this subsection, we have some parallel results to the previous subsection by conditioning on $X_2 = t$ instead of $X_2 > t$.

Proposition 5.9 *Suppose* $\mathbf{X} := (X_1, X_2)$*,* $\mathbf{Y} := (Y_1, Y_2)$ *and* $\mathbf{X}, \mathbf{Y} \in \mathcal{R}(F_1, F_2)$*. (a) If* $\mathbf{X} \sim C^*(F_1, F_2)$*,* $\mathbf{Y} \sim C(F_1, F_2)$ *and* $\lim_{v\to 1^-}[C^*_{1|2}(u|v)/C_{1|2}(u|v)] < 1$ *for all* $0 < u < 1$*, then Relation (II) holds. (b) If* $\widehat{C}^*$ *and* $\widehat{C}$ *are survival copulas of* C^* *and* C*,* $\mathbf{X} \sim \widehat{C}^*(F_1, F_2)$*,* $\mathbf{Y} \sim \widehat{C}(F_1, F_2)$ *and* $\lim_{v\to 0^+}[1 - C^*_{1|2}(u|v)]/[1 - C_{1|2}(u|v)] < 1$ *for all* $0 < u < 1$*, then Relation (II) holds.*

Corollary 5.10 *Suppose* $\mathbf{X} := (X_1, X_2)$*,* $\mathbf{Y} := (Y_1, Y_2)$*,* $\mathbf{X}, \mathbf{Y} \in \mathcal{R}(F_1, F_2)$*, and their dependence structures are* C^* *and* C *that are Archimedean copulas constructed as in (2.4) with LTs* ψ *and* φ*, respectively. If* $\Upsilon(1)$ *exists and* $\Upsilon(u) > \Upsilon(1)$ *for any* $0 < u < 1$*, then Relation (II) holds.*

Corollary 5.11 *Suppose* $\mathbf{X} := (X_1, X_2)$*,* $\mathbf{Y} := (Y_1, Y_2)$*,* $\mathbf{X}, \mathbf{Y} \in \mathcal{R}(F_1, F_2)$*,* $\mathbf{X} \sim \widehat{C}^*(F_1, F_2)$ *and* $\mathbf{Y} \sim \widehat{C}(F_1, F_2)$*, where* C^* *and* C *are Archimedean copulas constructed as in (2.4) with LTs* ψ *and* φ*, respectively. If* $\psi^{-1} \circ \varphi(s)$ *is strictly convex for large* s*, and in particular if for* u *sufficiently small,* $\Upsilon(u)$ *is strictly decreasing in* u*, then Relation (II) holds.*

Corollary 5.12 *Suppose* $\mathbf{X} := (X_1, X_2)$, $\mathbf{Y} := (Y_1, Y_2)$, $\mathbf{X}, \mathbf{Y} \in \mathcal{R}(F_1, F_2)$, *and their dependence structures are* C^* *and* C *that are Archimedean copulas constructed as in (2.4) with LTs* ψ *and* φ*, respectively. If* $1 - \psi \in \mathrm{RV}_0(0^+)$ *and* $\liminf_{s\to 0^+}(-s\varphi'(s))/(1-\varphi(s)) > 0$*, then Relation (II) holds.*

Corollary 5.13 *Suppose* $\mathbf{X} := (X_1, X_2)$, $\mathbf{Y} := (Y_1, Y_2)$, $\mathbf{X}, \mathbf{Y} \in \mathcal{R}(F_1, F_2)$, $\mathbf{X} \sim \widehat{C}^*(F_1, F_2)$ *and* $\mathbf{Y} \sim \widehat{C}(F_1, F_2)$*, where* C^* *and* C *are Archimedean copulas constructed as in (2.4) with LTs* ψ *and* φ*, respectively. If* $\Upsilon(u)$ *is ultimately monotone,* $\psi \in \mathrm{RV}_0$ *and* $\liminf_{s\to\infty}(-s\varphi'(s))/(\varphi(s)) > 0$*, then Relation (II) holds.*

5.2.3 Second order conditions and conservativity

From Example 5.2, we notice that BB2 in general has a stronger positive association in the lower tail than BB3, although BB2 and BB3 are both lower tail comonotonic. In this subsection, we will study second order conditions under which tail comonotonicity may appear various strength of lower tail dependence.

Suppose that copula C is lower tail comonotonic, if there exist auxiliary functions $A(t), B(t)$ such that

$$\lim_{u\to 0^+} \frac{B\left(\frac{C(uw_1,\ldots,uw_d)}{u} - \min(w_1,\ldots,w_d)\right)}{A(u)} = H(w_1,\ldots,w_d) \not\equiv 0, \quad \mathbf{w} > 0, \tag{5.6}$$

then the speed of convergence can be characterized by the auxiliary functions $A(t)$ and $B(t)$.

Second order conditions may play an important role in modeling high-risk scenarios. We refer to Degen et al. (2010) and Chapter 6 for relevant studies under the condition of second order regular variation (de Haan and Stadtmüller, 1996). The idea here is similar to second order regular variation, but the auxiliary functions $A(t)$ and $B(t)$ may have various forms other than a power function as in the second order regular variation.

Lemma 5.14 *Let* C *be a* d*-dimensional Archimedean copula constructed with LT* $\psi \in \mathrm{RV}_0$ *as in (2.4). Letting* $w^* := \min\{w_1, \ldots, w_d\}$*, and* w^* *is unique in* $\mathbf{w}$*,* *then as* $u \to 0^+$*,*

$$C(uw_1, \ldots, uw_d) \sim uw^* + \psi'(\psi^{-1}(uw^*)) \\ \times \sum_{i \in \{j : w_j \neq w^*\}} \big(\psi^{-1}(uw_i)\big) =: uw^* + \epsilon(u, \mathbf{w}),$$

where the second order term $\epsilon(u, \mathbf{w}) = o(u), u \to 0^+$.

Based on Lemma 5.14, one can obtain concrete forms of second order terms. The following are examples for the bivariate case of BB2 and BB3, for which the auxiliary functions have simple forms. The derivation is straightforward by using Lemma 5.14, thus omitted.

Example 5.3 (BB2) The auxiliary functions for BB2 copula are

$$A(t) = t^{-\theta}, \theta > 0, \\ B(t) = \log|t|;$$

that is,

$$\lim_{u \to 0^+} \frac{\log\left|\frac{C_\psi(uw_1, uw_2)}{u} - \min(w_1, w_2)\right|}{u^{-\theta}} =: H(w_1, w_2) = -\delta|w_1^{-\theta} - w_2^{-\theta}|.$$

Therefore, θ is the second order parameter that dominates the speed of convergence.

Example 5.4 (BB3) The auxiliary functions for BB3 copula are

$$A(t) = (-\log t)^{\theta - 1}, \theta > 1, \\ B(t) = \log|t|;$$

that is,

$$\lim_{u \to 0^+} \frac{\log\left|\frac{C_\psi(uw_1, uw_2)}{u} - \min(w_1, w_2)\right|}{(-\log u)^{\theta - 1}} =: H(w_1, w_2) = -\delta\theta|\log w_1 - \log w_2|.$$

So θ is the second order parameter.

For statistical inference with BB2 or BB3 in the lower tail, if we want to reduce the number of the parameters we can also consider to simply assume δ to be a fixed number, say $\delta = 0.01$ or $\delta = 0.1$. From computational experiments, we find that there are some redundancies when we involve the parameter δ. For both BB2 and BB3, when θ is larger, both $u^{-\theta}$ and $(-\log u)^{\theta-1}$ converges to $+\infty$ faster as $u \to 0^+$ and the copula has more positive dependence at the lower sub-extreme.

By looking at the auxiliary functions of BB2 and BB3, we may also conclude that BB2 is more lower tail positive dependent than BB3 below a sufficiently low threshold, which is consistent to Example 5.2. The reason is due to $u^{-\theta_{BB2}}$ goes to $+\infty$ faster than $(-\log u)^{\theta_{BB3}-1}$ as $u \to 0^+$, and this just implies that the convergence of $C_{BB2}(uw_1, uw_2)/u \to \min\{w_1, w_2\}$ is faster than the convergence of $C_{BB3}(uw_1, uw_2)/u \to \min\{w_1, w_2\}$, as $u \to 0^+$.

5.3 Asymptotically worst dependence structures

As we have shown in Section 5.2, tail comonotonicity affects the conditional specifications. Moreover, tail comonotonicity may also influence risk measures for aggregate losses. For instance, if $X_i \stackrel{d}{=} Y_i, i = 1, \ldots, d$, and the dependence structures of $(X_1, \ldots, X_d)$ and $(Y_1, \ldots, Y_d)$ are different, then risk measures on $\sum X_i$ and $\sum Y_i$ may be different, although they have the same marginal distributions. In this section, we will briefly study an asymptotically worst dependence concept brought about by tail comonotonicity.

Let $\mathbf{X} = (X_1, \ldots, X_d)$ be a d-dimensional real-valued random vector, with univariate cdf's $F_1, \ldots, F_d$. The Fréchet space containing all the random vectors that possess these margins is denoted as $\mathcal{R}(F_1, \ldots, F_d)$. Suppose $g : \mathbb{R}^d \to \mathbb{R}$ is an aggregate function, and $Q_p : \mathbb{R} \to \mathbb{R}_+$ is a quantile based risk measure, where p is the probability level and is usually close to 1. Then $\mathbf{X}$ is said to have an *asymptotically worst dependence structure* in terms of the risk measure Q_p and the aggregate function g, if for any $\mathbf{X}' \in \mathcal{R}(F_1, \ldots, F_d)$,

$$Q_p[g(\mathbf{X})] \gtrsim Q_p[g(\mathbf{X}')], \quad p \to 1^-. \tag{5.7}$$

Note that, in Mainik and Rüschendorf (2012), an asymptotic ordering called asymptotic portfolio loss order has been used to compare the tail probabilities of two portfolios.

In this section, we will study situations where tail comonotonicity becomes an asymptotically worst dependence structure. For the choice of risk measures, we restrict ourself to VaR and CTE.

For aggregate functions, we only consider the case where $g(x_1, \ldots, x_d) = \sum_{i=1}^{d} x_i$. Then this concept is also relevant to conservative assessments for diversification effects of risk aggregation.

5.3.1 Conditional tail expectation

Assumption 5.1 Let $\mathbf{X}$ be an upper tail comonotonic non-negative random vector with continuous marginal distributions $F_1, \ldots, F_d$ defined on $[0, \infty)$ such that $1 - F_i(t) \in \mathrm{RV}_{-\alpha}$ with $\alpha > 1$, and

$$\lim_{t\to\infty} [1 - F_i(t)]/[1 - F_1(t)] = k_i, \quad 0 < k_i < \infty$$

for any $i = 1, \ldots, d$.

The next result shows that tail comonotonicity is an asymptotically worst dependence structure for tail equivalent random variables that have regularly varying upper tails.

Proposition 5.15 *If $\mathbf{X}$ satisfies Assumption 5.1, then for any $\mathbf{X}' \in \mathcal{R}(F_1, \ldots, F_d)$,*

$$\mathrm{CTE}_p(X_1 + \cdots + X_d) \gtrsim \mathrm{CTE}_p(X_1' + \cdots + X_d'), \quad p \to 1^-.$$

In this sense, upper tail comonotonicity provides a conservative dependence structure.

Assumption 5.2 Let $\mathbf{X}$ be an upper tail comonotonic non-negative random vector with continuous marginal distributions $F_1, \ldots, F_d$ defined on $[0, \infty)$ and there exists a positive measurable function $a(\cdot)$ such that

$$\lim_{t\to\infty} \frac{1 - F_i(t + a(t)s)}{1 - F_i(t)} = e^s, \quad \forall s \in \mathbb{R}, \tag{5.8}$$

and $\lim_{t\to\infty}[1 - F_i(t)]/[1 - F_1(t)] = k_i$ with $0 < k_i < \infty$ for any $i = 1, \ldots, d$.

Due to Proposition 4.16, we also have the following result.

Proposition 5.16 *If $\mathbf{X}$ satisfies Assumption 5.2, then for any $\mathbf{X}' \in \mathcal{R}(F_1, \ldots, F_d)$,*

$$\text{CTE}_p(X_1 + \cdots + X_d) \gtrsim \text{CTE}_p(X_1' + \cdots + X_d'), \quad p \to 1^-.$$

5.3.2 Value at risk

Based on Propositions 4.13 and 4.14, it suffices to find conditions on the dependence structure for $\mathbf{X}$, where the margins are specified by either Assumptions A or B, so that VaR_p becomes subadditive as p is in a left-neighbor of 1. It is well known that, in general VaR is not subadditive. However, with certain regularity conditions, VaR may also be subadditive; the study of the regularity conditions has been an important topic in the literature but is out of the scope of this chapter. In what follows, some results for asymptotically worst dependence structure in terms of VaR will be given. Those are based on known facts of situations where VaR becomes (asymptotically) subadditive.

Proposition 5.17 *Let $\mathbf{X}' = (X_1', \ldots, X_d') \in \text{MRV}_d(-\beta), \beta > 1$, with identically distributed margins on $[0, \infty)$, and $\mathbf{X}$ is upper tail comonotonic with $X_i \overset{d}{=} X_i', i = 1, \ldots, d$. Then*

$$\text{VaR}_p(X_1 + \cdots + X_d) \gtrsim \text{VaR}_p(X_1' + \cdots + X_d'), \quad p \to 1^-.$$

For marginal distributions that are in the Maximum Domain of Attraction of Gumbel, subadditivity of VaR may hold for multivariate elliptical distributions (McNeil et al., 2005, Theorem 6.8). So the following result holds.

Proposition 5.18 *Let $\mathbf{X}' = (X_1', \ldots, X_d')$ be a d-dimensional elliptical distribution where the margins satisfying (5.8) with support $(-\infty, \infty)$, and $\mathbf{X}$ is upper tail comonotonic with $X_i \stackrel{d}{=} X_i', i = 1, \ldots, d$. Then*

$$\text{VaR}_p(X_1 + \cdots + X_d) \gtrsim \text{VaR}_p(X_1' + \cdots + X_d'), \quad p \to 1^-.$$

5.4 Simulation study

In this section, we report on simulations to illustrate the conservativity of tail comonotonicity in terms of $\mathbb{E}[X_1|X_2 > \text{VaR}_p(X_2)]$, $\mathbb{E}[X_1|X_2 = \text{VaR}_p(X_2)]$, $\text{VaR}_p(X_1 + X_2)$ and $\text{CTE}_p(X_1 + X_2)$ as $p \to 1^-$, respectively.

5.4.1 Conditional specifications

Now we compare the effects on $\mathbb{E}[X_1|X_2 > \text{VaR}_p(X_2)]$ and $\mathbb{E}[X_1|X_2 = \text{VaR}_p(X_2)]$ for different Archimedean copulas. Comonotonicity and the survival copulas of BB2, MTCJ, Gumbel and Frank (denoted as s.BB2, s.MTCJ, s.Gumbel and s.Frank) are used to model the dependence structures of two margins X_1 and X_2; that is, the strength of upper tail positive dependence are different. From Chapter 3, considering both tail order and tail dependence parameters, the strength of dependence in the upper tail can be ordered as:

$$\text{Comonotonicity} \succ \text{s.BB2} \succ \text{s.MTCJ} \succ \text{s.Gumbel} \succ \text{s.Frank}. \tag{5.9}$$

Choosing parameters for these Archimedean copulas to have the same Blomqvist's β (since MLEs for different copula families often lead to a similar Blomqvist's β), we can compare the values of $\mathbb{E}[X_1|X_2 > \text{VaR}_p(X_2)]$ or $\mathbb{E}[X_1|X_2 = \text{VaR}_p(X_2)]$ for different Archimedean copulas. Figures 5.1 and 5.2 contain the plots for the two conditional specifications, respectively, with p ranging from 80% to 99%; this range is most relevant for applications. The calculations were based on numerical integration through Monte Carlo simulations. For both figures, the copula parameters used in the simulation were the following: BB2: $\delta = 2, \theta = 0.4$; MTCJ: $\delta = 1.424$; Gumbel: $\delta = 1.729$ and Frank: $\delta = 3.844$. The common Blomqvist's β was 0.421.

From Figure 5.1, it is interesting that, when p is sufficiently large (say, $p \geq 80\%$ for the plots), the order of the values of $\mathbb{E}[X_1|X_2 > \text{VaR}_p(X_2)]$ is the same as the order in (5.9). Moreover, the same pattern can be observed regardless of the margins. The pattern for Figure 5.2 is different than that for $\mathbb{E}[X_1|X_2 > \text{VaR}_p(X_2)]$. But when p is sufficiently large (say, $p \geq 95\%$ in this example), the rank of $\mathbb{E}[X_1|X_2 = \text{VaR}_p(X_2)]$ is still kept as the same as (5.9).

Based on these comparisons, the BB2 tail comonotonic copula is reasonable and better than the usual comonotonicity when considering a relatively lower probability level p where more data are available. However, as less data are available when considering a higher probability level, the tail comonotonic copula becomes more conservative and approaches the upper bound provided by the usual comonotonicity. This kind of conservatism makes more sense for quantitative risk management.

5.4.2 Asymptotically worst dependence structures

To illustrate the idea of applying tail comonotonicity for a reasonably conservative modeling of high-risk scenarios, we generated a bivariate data with s.BB1 as the copula and standard Pareto as the margins. We transformed the data to rank scores and fitted these bivariate rank scores by copulas that possess upper tail dependence, such as t, Gumbel, s.BB2 and s.BB3 copulas. The estimation was based on maximum quasi-likelihood estimation and the margins are assumed to be given. The Monte Carlo method was then used to approximate the corresponding VaRs and CTEs, based on the fitted copulas and the known margins.

Table 5.3 and 5.4 present comparisons between different copulas, for the cases of more dependence and less dependence, respectively. The margins were standard Pareto distributions with cdf $F(x) = 1 - x^{-3}$. For each run, the sample size was 10^5 and "se", the simulation error for the Monte Carlo method was based on $1,000$ runs. In the tables, "ind" is for the independent case, "s.BB1.o" means s.BB1 is the true model, and λ is the corresponding upper tail dependence parameter.

Table 5.3 shows that: (1) Overlooking the dependence between margins will seriously underestimate the aggregate risks. (2) The Gumbel copula has the smallest AIC, but it underestimates the risks. (3) The BB2 and BB3 copulas overesti-

Figure 5.1: The value of y-axis is $\mathbb{E}[X_1|X_2 > \text{VaR}_p(X_2)]$. In the first plot, X_1, X_2 have Exponential distributions with cdf $F(x) = 1 - Exp(-x/\sigma)$. In the second plot, X_1, X_2 have Pareto distributions with cdf $F(x) = 1 - (1 + x/\sigma)^{-\theta}$.

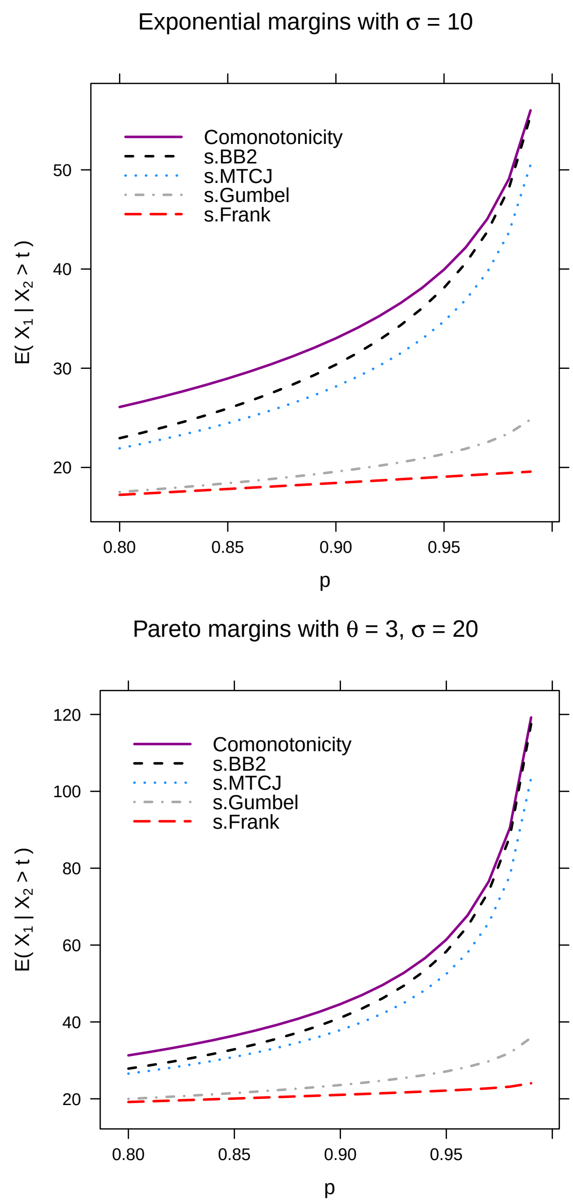
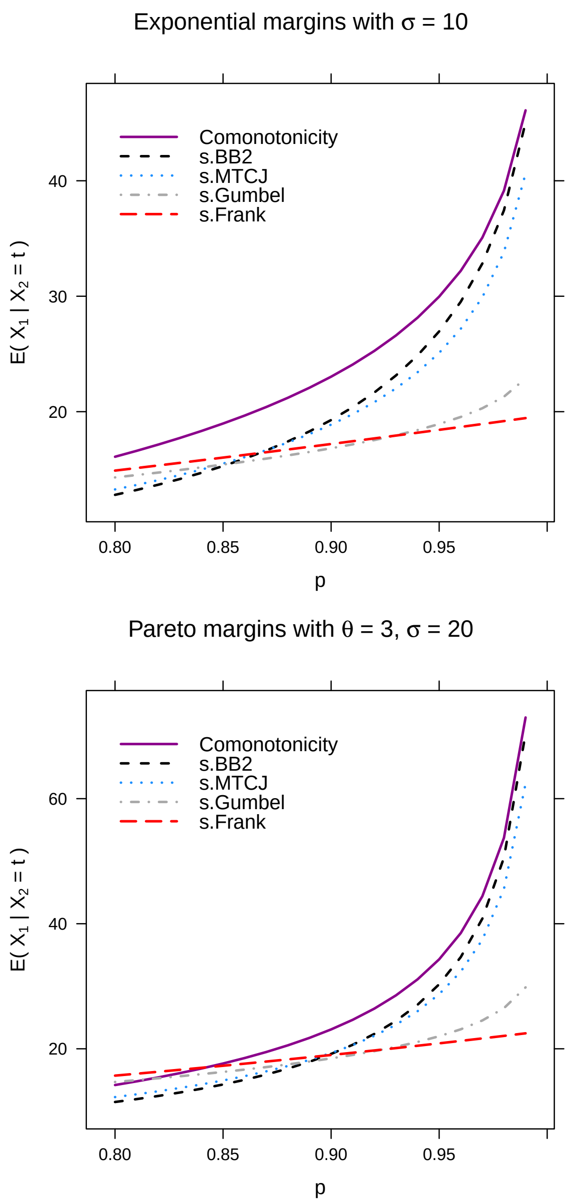


Figure 5.2: The value of y-axis is $\mathbb{E}[X_1|X_2 = \text{VaR}_p(X_2)]$. In the first plot, X_1, X_2 have Exponential distributions with cdf $F(x) = 1 - Exp(-x/\sigma)$. In the second plot, X_1, X_2 have Pareto distributions with cdf $F(x) = 1 - (1 + x/\sigma)^{-\theta}$.



mate the risk a little bit and could provide a reasonable conservative estimate for this case, especially when the probability level p is very high. (4) The tail dependence parameters seem relevant to the magnitude of the aggregate risks: a higher tail dependence parameter tends to lead to a larger aggregate risk, especially when the probability level is quite high. This also suggests that tail comonotonicity is conservative under some conditions. Comparing Tables 5.3 and 5.4, we can also find that overestimation arising from tail comonotonicity is milder for the case of stronger tail dependence than for the case of less tail dependence.

Table 5.3: (More dependence) VaR and CTE for $X_1 + X_2$. The MLEs were based on the whole sample generated from s.BB1 ($\delta = 1.57, \theta = 1.68, \lambda = 0.77$, more dependence) with a sample size of 2000. The bold AIC value is the smallest.

	VaR90	VaR995	CTE90	CTE995	AIC	λ
ind	4.084	8.949	5.532	12.631	-	-
se	<0.001	0.003	0.001	0.010		
s.BB1.o	4.270	11.573	6.400	17.369	-	0.77
se	<0.001	0.005	0.001	0.017		
t	4.269	11.275	6.315	16.842	-2678.4	0.53
se	<0.001	0.005	0.001	0.016		
Gumbel	4.272	11.537	6.389	17.289	**-2887.7**	0.73
se	<0.001	0.005	0.001	0.016		
s.BB2	4.279	11.616	6.419	17.411	-2737.6	1
se	<0.001	0.005	0.001	0.017		
s.BB3	4.263	11.643	6.416	17.440	-2782.7	1
se	<0.001	0.005	0.001	0.016		

With the same simulation settings, we repeated the experiment for 40 times. Table 5.5 is a summary of these simulations. The approach is conservative only when the probability level is very high (say, $> 99\%$). With a moderately high probability level, say, 90%, this approach would be not conservative. Moreover, tail comonotonicity worked better for the case where the original dependence is actually stronger.

Table 5.4: (Less dependence) VaR and CTE for $X_1 + X_2$. The MLEs were based on the whole sample generated from s.BB1 ($\delta = 2, \theta = 0.4, \lambda = 0.42$, less dependence) with a sample size of 2000. The bold AIC value is the smallest.

	VaR90	VaR995	CTE90	CTE995	AIC	λ
ind	4.083	8.951	5.531	12.641	-	-
se	<0.001	0.004	0.001	0.010		
s.BB1.o	4.252	11.094	6.249	16.515	-	0.42
se	<0.001	0.005	0.001	0.015		
t	4.262	11.100	6.261	16.465	**-2099.0**	0.39
se	<0.001	0.005	0.001	0.015		
Gumbel	4.245	11.417	6.331	17.105	-1946.5	0.65
se	<0.001	0.006	0.001	0.015		
s.BB2	4.234	11.480	6.345	17.225	-1574.3	1
se	<0.001	0.005	0.001	0.016		
s.BB3	4.217	11.192	6.220	16.994	-2054.5	1
se	<0.001	0.005	0.001	0.017		

5.5 Application on a claim dataset

We now illustrate how tail comonotonicity can be used in real applications with the dataset of Loss (X_2) versus Allocated Loss Adjustment Expense (ALAE) (X_1) that has been studied in the literature such as Frees and Valdez (1998), and Klugman and Parsa (1999). The margins will be fitted by a Pareto distribution with the following distribution function (same as in Frees and Valdez (1998)) $F(x) = 1 - (1 + x/\sigma)^{-\theta}$. Its density and inverse functions are $f(x) = (\theta/\sigma)\,(1 + x/\sigma)^{-\theta-1}$ and $F^{-1}(x) = \sigma(1-x)^{-1/\theta} - \sigma$. In addition to the Gumbel copula fitted in Frees and Valdez (1998), we use s.BB2 as a conservative model to fit the dependence structure as well.

ALAE for auto insurance claims often involves legal expenses that could be very large. Just like the current dataset, for a relatively small amount of loss, the ALAE can be very large and there are usually no limits that can be set for the expense. If one wants to assess the risk of future total loss and expense, one way is to think that the historical dataset reflects the future claims. However, this method may be too optimistic and does not account for the uncertainty beyond the historical dataset.

Table 5.5: VaR and CTE for $X_1 + X_2$. The first 4 columns of values are the means of corresponding quantities calculated from 40 random samples generated from the same settings as before. The rest are frequency of those quantities being greater than those for s.BB1.o. (A: more dependent, large sample size; B: less dependent, large sample size; C: more dependent, small sample size; D: less dependent, small sample size. More dependence: $\delta = 1.57, \theta = 1.68, \lambda = 0.77$; less dependence: $\delta = 2, \theta = 0.4, \lambda = 0.42$. Large sample: $N = 2000$; small sample: $N = 50$.)

		VaR90	VaR995	CTE90	CTE995	VaR90	VaR995	CTE90	CTE995
A	s.BB1.o	4.270	11.574	6.400	17.365				
	t	4.266	11.281	6.313	16.845	2	0	0	0
	Gumbel	4.270	11.536	6.387	17.303	15	0	0	0
	s.BB2	4.277	11.612	6.417	17.423	40	40	40	40
	s.BB3	4.249	11.643	6.408	17.490	0	40	33	40
B	s.BB1.o	4.251	11.090	6.247	16.510				
	t	4.257	11.093	6.255	16.488	38	21	35	13
	Gumbel	4.243	11.397	6.324	17.085	0	40	40	40
	s.BB2	4.230	11.462	6.337	17.198	0	40	40	40
	s.BB3	4.219	11.168	6.218	16.960	0	31	1	40
C	s.BB1.o	4.270	11.573	6.399	17.355				
	t	4.263	11.279	6.310	16.818	18	0	0	0
	Gumbel	4.270	11.530	6.385	17.293	18	8	10	9
	s.BB2	4.274	11.624	6.421	17.439	23	34	34	34
	s.BB3	4.253	11.620	6.398	17.459	15	35	24	36
D	s.BB1.o	4.251	11.087	6.247	16.499				
	t	4.255	11.129	6.262	16.527	22	26	26	22
	Gumbel	4.248	11.420	6.335	17.120	15	40	40	40
	s.BB2	4.236	11.529	6.359	17.298	9	40	40	40
	s.BB3	4.228	11.325	6.279	17.062	7	32	25	36

We now remove the data where ALAE is among the 1% largest, pretending that we have not observed those large values yet. Then we use a Pareto distribution to fit the margins for the rest of the dataset, and use the Gumbel and s.BB2 copulas to fit the dependence structure, respectively. Then simulations based on the two fitted models will be used to assess high senario risks in terms of $\mathbb{E}[\text{ALAE}|\text{LOSS} \geq \text{VaR}_p(\text{LOSS})]$; the results for the Gumbel, s.BB2 copulas, and empirical assessement based on the data will be compared. See Figure 5.3 for a comparison, where calculations were based on the mean of 40 simulations, and for each simulation a random sample of size 10^6 was used. When the number of simulations was about 40, the mean became stable.

First, we fit the two margins separately and get the estimates (Table 5.6). Compared to the estimates for the original data (replicated from Frees and Valdez (1998)), the upper tails of both X_1 and X_2 become a bit lighter (because of larger $\hat{\theta}$).

Table 5.6: Estimates for margins with/without 1% largest ALAE removed

		LOSS(X_2)		ALAE(X_1)	
Data	Parameters	Estimate	Standard error	Estimate	Standard error
1% removed	σ	17189	1714	21686	2964
	θ	1.310	0.082	3.045	0.323
original	σ	14453	1397	15133	1633
	θ	1.135	0.066	2.223	0.175

Then we assume that the margins are known as what we have estimated in Table 5.6 and plug them into likelihood functions. Considering there are 31 (originally 34) right-censored loss data (X_2), the likelihood contributed by a right-censored data value (x_1, x_2) is

$$f_1(x_1) - (\partial_1 F)(x_1, x_2) = f_1(x_1)\left(1 - C_{2|1}(F_2(x_2)|F_1(x_1))\right).$$

The estimation is reported in Table 5.7, and a comparison for different models is illustrated in Figure 5.3. From this example, we can observe that, as the probability level p is very large (eg, 99.5%), "empirical" is not able to extrapolate the increasing trend from the original data and even gives a counter-intuitive decreasing trend, but s.BB2 leads to a reasonable extrapolation beyond the original data although s.BB2 was fitted based on the data with the 1% largest removed. The Gumbel copula seems to be reasonably good as well but has a limited ability in extrapolation compared with s.BB2. In this particular example, s.BB2 outperforms the other models and provides reasonable extrapolation into high-risk scenarios. Furthermore, the choice of δ controls the degree of conservativity of s.BB2.

Table 5.7: Comparison between the s.BB2 and Gumbel copulas: from the second order condition of BB2, we know that θ is the second order parameter that dominates the tail behavior of the copula. If we fit δ, θ of s.BB2 simultaneously for the data, then the MLE of δ tends to be very close to 0; that is, s.MTCJ could better fit the data than s.BB2 (since as $\delta \to 0$, BB2 becomes MTCJ). But s.MTCJ will lose the asymptotic full dependence structure. The aim of this comparison is not to find the best fitting copula but to study how conservative s.BB2 is. So we fix two values for the parameter δ for s.BB2 ($\delta = 0.01, 0.1$), and then obtain the MLEs of θ.

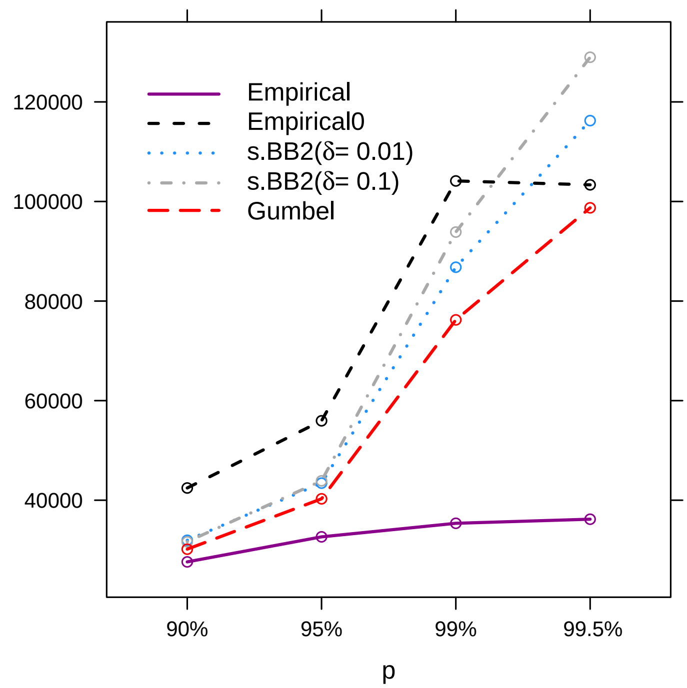
Model	Parameters	Estimate	Standard error	Blomqvist's β
s.BB2($\delta = 0.01$)	θ	0.679	0.041	0.254
s.BB2($\delta = 0.10$)	θ	0.520	0.026	0.226
Gumbel	δ	1.399	0.028	0.283

5.6 Discussion

"Essentially, all models are wrong, but some are useful." (Box and Draper, 1987). By showing the conservativity of tail comonotonicity and its usage in accounting for unexpected rare events (especially with the promising result illustrated in Figure 5.3), we are not conveying the information that tail comonotonicity is better than other candidate models. Through this study, we hope to be able to raise an issue that has not been attracted enough attention, but may be useful in dependence modeling when one has to assess dependent large risks relying on insufficient information. Since there are very few insurance claim datasets available to the academic researchers and actuaries, more empirical evidence for the usefulness of tail comonotonicity must be welcome.

An absolutely continuous copula that is tail comonotonic can provide more realistic conservative bounds on risk measures than the comonotonic copula. Of the simple 2-parameter bivariate copula families, the BB2 family is the most tail asymmetric with lower tail comonotonicity and upper tail independence (upper tail order 2). The BB3 copula is also lower tail comonotonic but has upper tail dependence. As shown in Example 5.2 and Section 5.2.3, BB3 has a slower rate of convergence to the limiting comonotonic tail dependence function. It is shown that the survival

Figure 5.3: The value of y-axis is $\mathbb{E}[\text{ALAE}|\text{LOSS} \geq \text{VaR}_p(\text{LOSS})]$. "Empirical" is based on the data with the 1% largest removed; "Empirical0" is based on the original data.



BB2 copula family provides conservative bounds as a model for bivariate loss data if in fact there is upper tail dependence but not tail comonotonicity. The survival BB3 copula family may also provide bounds if one does not want to use a copula as conservative as BB2.

When a tail comonotonic copula is used to provide a data-driven conservative dependence structure, it is suggested to first fit the margins separately then fit the copula based on the uniform scores derived from the fitted margins. If one wants to fit the copula and margins simultaneously, misspecification of the copula model may affect the estimation for margins severely. Moreover, under such a situation, computation may become more complicated.

For multivariate loss data, one can use multivariate extensions of the BB2 and

BB3 copulas. The BB2 and BB3 families are bivariate Archimedean copulas based on LT families given in Examples 4.1 and 4.2, so there is an obvious extension to exchangeable multivariate Archimedean copulas by the definition. Because the two LT families belong to RV_0, they can be used to get multivariate tail comonotonic copulas with more flexible non-extremal dependence via the mixture of max-id copula families (Joe and Hu, 1996).

Although the influence is not large, the choice of the BB2 dependence parameter δ in the data analysis example was arbitrary. The δ can be thought as a parameter to control the degree of conservativity. For future research, we expect that there are certain criteria due to statistical and/or economic reasons to get the parameter δ appropriately tuned. Moreover, we will also look more closely at the tail expansions of $\mathbb{E}[X_1|X_2 > t]$ and $\mathbb{E}[X_1|X_2 = t]$ for different marginal distributions of X_1 and X_2. The strength of the dependence in the tail can affect whether the conditional tail expectation behaves like $O(t)$, $O(t^\gamma)$ for some $0 \le \gamma < 1$ as $t \to \infty$.

5.7 Proofs

Proof of Proposition 5.1: (a) Write

$$\begin{aligned}
\mathbb{E}[X_1|X_2 > t; C^*] &= \int_0^{+\infty} \mathbb{P}[X_1 > x|X_2 > t]dx = \int_0^{+\infty} \frac{\mathbb{P}[X_1 > x, X_2 > t]}{\mathbb{P}[X_2 > t]}dx \\
&= \int_0^{+\infty} \frac{1 - F_1(x) - F_2(t) + C^*(F_1(x), F_2(t))}{\overline{F}_2(t)}dx \\
&= (\overline{F}_2(t))^{-1}\left\{\mathbb{E}[X_1] + \int_0^{\infty}\{-F_2(t) + C^*(F_1(x), F_2(t))\}dx\right\}.
\end{aligned} \tag{5.10}$$

Since $\mathbb{E}[X_1] < \infty$, and $\mathbb{P}[X_1 > x, X_2 > t] \le \mathbb{P}[X_1 > x]$, the integral $\int_0^\infty\{-F_2(t)+C^*(F_1(x), F_2(t))\}dx$ is finite. The conclusion follows if for all v close to 1 from below,

$$\int_0^\infty \{-v + C^*(F_1(x), v)\}dx = \int_0^1 \{[-v + C^*(u, v)] \cdot [f_1(F_1^{-1}(u))]^{-1}\}du$$
$$\geq \int_0^1 \{[-v + C(u, v)] \cdot [f_1(F_1^{-1}(u))]^{-1}\}du = \int_0^\infty \{-v + C(F_1(x), v)\}dx. \tag{5.11}$$

The above follows if for any given $0 < u < 1$, there exists a $v_0 = v_0(u)$ such that $1 \geq v > v_0$ implies that

$$C^*(u, v) \geq C(u, v). \tag{5.12}$$

Taking a Taylor expansion for $g(v) := C^*(u, v) - C(u, v)$ about v' with $v < v' < 1$ leads to

$$C^*(u, v) - C(u, v) \sim (v - v')\big[C^*_{1|2}(u|v')] - C_{1|2}(u|v')\big]$$
$$= (v - v')\, C_{1|2}(u|v') \left\{ \frac{C^*_{1|2}(u|v')}{C_{1|2}(u|v')} - 1 \right\}.$$

Then (5.12) holds if $\lim_{v' \to 1} C^*_{1|2}(u|v')/C_{1|2}(u|v') < 1$. (If the limit of the ratio is 1, then possibly a non-standard second order expansion is needed.)

(b) The copulas for F_1, F_2 are $\widehat{C}^*(u, v) = u + v - 1 + C^*(1 - u, 1 - v)$ and $\widehat{C}(u, v) = u + v - 1 + C(1 - u, 1 - v)$, so that $\widehat{C}^*_{1|2}(u|v) = 1 - C^*_{1|2}(1 - u|1 - v)$ and $\widehat{C}_{1|2}(u|v) = 1 - C_{1|2}(1 - u|1 - v)$. Now part (a) applies to $\widehat{C}^*_{1|2}$ and $\widehat{C}_{1|2}$. □

Proof of Corollary 5.2: For any given $0 < u < 1$, write

$$\frac{C^*_{1|2}(u|v)}{C_{1|2}(u|v)} = \frac{\psi'(\psi^{-1}(u) + \psi^{-1}(v))/\psi'(\psi^{-1}(v))}{\varphi'(\varphi^{-1}(u) + \varphi^{-1}(v))/\varphi'(\varphi^{-1}(v))},$$

so $\lim_{v \to 1^-} C^*_{1|2}(u|v)/C_{1|2}(u|v) < 1$ if

$$\lim_{v \to 1^-} \frac{\varphi'(\varphi^{-1}(v))}{\psi'(\psi^{-1}(v)} < \frac{\varphi'(\varphi^{-1}(u))}{\psi'(\psi^{-1}(u)},$$

or $\Upsilon(u) > \Upsilon(1)$. The conclusion now follows from Proposition 5.1. □

Proof of Corollary 5.3: Let $\psi = \psi_{\delta^*}$ and $\varphi = \varphi_\delta$ be the LTs for C^* and C respectively. By considering the domination term as $u \to 1^-$ in Table 5.1, $\Upsilon(1) = k \lim_{u \to 1^-} (1-u)^{\delta^* - \delta} = 0$ where k is a positive constant. Hence $\Upsilon(u) > \Upsilon(1)$ for $0 < u < 1$ and Corollary 5.2 applies. $\square$

Proof of Proposition 5.4: Rather than applying part (b) of Proposition 5.1, we provide a more direct proof. Because we are using survival copulas, rather than (5.12), we want to show that for any given $0 < u < 1$, there is a v_0 such that $\widehat{C}^*(u,v) \geq \widehat{C}(u,v)$ or $C^*(1-u, 1-v) = C_\psi(1-u, 1-v) \geq C_\varphi(1-u, 1-v) = C(1-u, 1-v)$ for $1 \geq v > v_0$. The inequality is the same as

$$\psi(\psi^{-1}(u_1) + \psi^{-1}(v_1)) \geq \varphi(\varphi^{-1}(u_1) + \varphi^{-1}(v_1)), \tag{5.13}$$

with $u_1 = 1 - u$ and $v_1 = 1 - v$ for v_1 near 0. Let $u_1 = \varphi(s_u)$ and $v_1 = \varphi(s_v)$. Since ψ is decreasing, inequality (5.13) is the same as

$$\psi^{-1}(u_1) + \psi^{-1}(v_1) \leq (\psi^{-1} \circ \varphi)(\varphi^{-1}(u_1) + \varphi^{-1}(v_1))$$

or

$$(\psi^{-1} \circ \varphi)(s_u) + (\psi^{-1} \circ \varphi)(s_v) \leq (\psi^{-1} \circ \varphi)(s_u + s_v).$$

For s sufficiently large, if $\xi(s) := \psi^{-1} \circ \varphi(s)$ is strictly convex, then for any $w > 0$, $\xi(s+w) - \xi(s)$ is strictly increasing in s with limit of ∞, since $\xi(s)$ is also a strictly increasing function. Therefore, for s_v large enough, $\xi(s_u + s_v) - \xi(s_v) > \xi(s_u)$.

If $\Upsilon(u)$ is strictly decreasing in u near 0, that is, $\xi'(s) = \varphi'(s)/\psi'(\xi(s))$ is strictly increasing for large s by letting $s = \varphi^{-1}(u)$, then $\xi(s)$ is strictly convex for large s, which completes the proof. $\square$

Proof of Corollary 5.7: Since $1 - \psi(s) = \int_0^s -\psi'(t)dt$ and $-\psi'(t)$ is strictly increasing in t as $t \to 0^+$, by Lemma 5.6, $-\psi'(s)/(1-\psi(s)) = o(s^{-1})$, $s \to 0^+$. Therefore, by L'Hôpital's rule,

$$\lim_{s \to 0^+} \frac{\log(1-\psi(s))}{\log(1-\varphi(s))} = \lim_{s \to 0^+} \frac{[\log(1-\psi(s))]'}{[\log(1-\varphi(s))]'} = \lim_{s \to 0^+} \frac{\psi'(s)/[1-\psi(s)]}{\varphi'(s)/[1-\varphi(s)]} = 0,$$

which implies $\varphi(s)$ approaches 1 faster than $\psi(s)$ as $s \to 0^+$, and $\lim_{s \to 0^+}[1 -$

$\psi(s)]/[1-\varphi(s)] = \infty$. Thus, $\lim_{u\to 1^-} \psi^{-1}(u)/\varphi^{-1}(u) = 0$, and

$$\lim_{u\to 1^-} \Upsilon(u) = \lim_{u\to 1^-} \frac{\varphi'(\varphi^{-1}(u))}{\psi'(\psi^{-1}(u))} = \lim_{u\to 1^-} \frac{[\psi^{-1}(u)]'}{[\varphi^{-1}(u)]'} = \lim_{u\to 1^-} \frac{\psi^{-1}(u)}{\varphi^{-1}(u)} = 0.$$

Then Corollary 5.2 implies the result. □

Proof of Corollary 5.8: Since $\psi(s) = \int_s^\infty -\psi'(t)dt$ and $-\psi'(t)$ is strictly decreasing in t as $t \to \infty$, by Lemma 5.5, $-\psi'(s)/\psi(s) = o(s^{-1})$, $s \to \infty$. Therefore, by L'Hôpital's rule,

$$\lim_{s\to\infty} \frac{\log\psi(s)}{\log\varphi(s)} = \lim_{s\to\infty} \frac{[\log\psi(s)]'}{[\log\varphi(s)]'} = \lim_{s\to\infty} \frac{\psi'(s)/\psi(s)}{\varphi'(s)/\varphi(s)} = 0,$$

which implies that $\varphi(s)$ goes to 0 faster than $\psi(s)$ as $s \to \infty$ and $\lim_{s\to\infty} \varphi(s)/\psi(s) = 0$. Thus, $\lim_{u\to 0^+} \varphi^{-1}(u)/\psi^{-1}(u) = 0$. Then write

$$\frac{1}{\Upsilon(u)} := \frac{\psi'(\psi^{-1}(u))}{\varphi'(\varphi^{-1}(u))} = \frac{-\psi'(\psi^{-1}(u))\psi^{-1}(u)}{\psi(\psi^{-1}(u))} \times \frac{\varphi(\varphi^{-1}(u))}{-\varphi'(\varphi^{-1}(u))\varphi^{-1}(u)} \times \frac{\varphi^{-1}(u)}{\psi^{-1}(u)}.$$

So there exists a $0 < K < \infty$, such that

$$0 \le \lim_{u\to 0^+} \frac{1}{\Upsilon(u)} \le \lim_{s\to\infty} K \frac{-s\psi'(s)}{\psi(s)} \times \lim_{u\to 0^+} \frac{\varphi^{-1}(u)}{\psi^{-1}(u)} = 0.$$

Then Proposition 5.4 implies the result. □

Proof of Proposition 5.9: We prove only part (a), since part (b) follows from (a) as in Proposition 5.1. Since $\mathbb{E}[X_1]$ is finite, $\mathbb{E}[X_1|X_2]$ is finite a.e. F_2. Write

$$\begin{aligned}
&\mathbb{E}[X_1|X_2 = t; C^*] \\
&= \int_0^{+\infty} \mathbb{P}[X_1 > x | X_2 = t] dx = \int_0^{+\infty} \frac{f_2(t) - (\partial_2 F)(x,t)}{f_2(t)} dx \\
&= \int_0^{+\infty} \frac{f_2(t) - C^*_{1|2}(F_1(x)|F_2(t)) f_2(t)}{f_2(t)} dx = \int_0^{+\infty} \{1 - C^*_{1|2}(F_1(x)|F_2(t))\} dx,
\end{aligned}$$

The conclusion follows if with $v = F_2(t) < 1$,

$$\begin{aligned} &\int_0^\infty \{C_{1|2}(F_1(x)|v) - C^*_{1|2}(F_1(x)|v)\}dx \\ &= \int_0^\infty C_{1|2}(F_1(x)|v)\left\{1 - \frac{C^*_{1|2}(F_1(x)|v)}{C_{1|2}(F_1(x)|v)}\right\}dx \geq 0, \end{aligned} \tag{5.14}$$

as $t \to \infty$. A sufficient condition is

$$\lim_{v\to 1^-} \frac{C^*_{1|2}(u|v)}{C_{1|2}(u|v)} < 1, \quad 0 < u < 1.$$

□

Proof of Corollary 5.10: From the assumption,

$$\lim_{v\to 1^-} \frac{C^*_{1|2}(u|v)}{C_{1|2}(u|v)} = \lim_{v\to 1^-} \frac{\psi'\big(\psi^{-1}(u)+\psi^{-1}(v)\big)/\psi'\big(\psi^{-1}(v)\big)}{\varphi'\big(\varphi^{-1}(u)+\varphi^{-1}(v)\big)/\varphi'\big(\varphi^{-1}(v)\big)} = \frac{\lim_{v\to 1^-}\Upsilon(v)}{\Upsilon(u)} < 1$$

so that Proposition 5.9 applies. □

Proof of Corollary 5.11: From part (b) of Proposition 5.9, we want to show that for fixed $0 < u < 1$, $\lim_{v\to 0^+}[1 - C^*_{1|2}(u|v)]/[1 - C_{1|2}(u|v)] < 1$, or

$$\lim_{v\to 0^+} C^*_{1|2}(u|v)/C_{1|2}(u|v) > 1,$$

i.e.,

$$\lim_{v\to 0^+} \frac{\psi'\left(\psi^{-1}(u)+\psi^{-1}(v)\right)}{\psi'(\psi^{-1}(v))} =: g(0) > h(0) := \lim_{v\to 0^+} \frac{\varphi'\left(\varphi^{-1}(u)+\varphi^{-1}(v)\right)}{\varphi'(\varphi^{-1}(v))}. \tag{5.15}$$

The strict convexity condition on $\psi^{-1}\circ\varphi$ and the proof of Proposition 5.4 imply that with u fixed,

$$\psi(\psi^{-1}(u)+\psi^{-1}(v)) =: G(v) > H(v) := \varphi(\varphi^{-1}(u)+\varphi^{-1}(v)), \tag{5.16}$$

for $v > 0$ sufficiently small. Also, $G(0) = H(0) = 0$, so we must have $G'(0) >$

$H'(0)$ for any $0 < u < 1$, thus (5.15) holds. since $G' = g$ and $H' = h$. $\square$

Proof of Corollary 5.12: The proof is straightforward due to Corollary 5.10 and the proof of Corollary 5.7. $\square$

Proof of Corollary 5.13: The proof is straightforward due to Corollary 5.11 and the proof of Corollary 5.8. $\square$

Proof of Lemma 5.14: It can be written that

$$C(uw_1, \ldots, uw_d) = \psi\left(\psi^{-1}(uw^*)\left(1 + \sum_{i\in\{j:w_j\neq w^*\}} \frac{\psi^{-1}(uw_i)}{\psi^{-1}(uw^*)}\right)\right).$$

Letting $t_i(u) := \psi^{-1}(uw_i)/\psi^{-1}(uw^*)$, since $\psi \in \mathrm{RV}_0$, $\psi^{-1} \in \mathrm{RV}_\infty(0^+)$, and thus $w_i > w^*$ implies that $\lim_{u\to 0^+} t_i(u) = 0$ for any $i \in \{j : w_j \neq w^*\}$. Let $\mathbf{t}^* := (t_i)_{i\in\{j:w_j\neq w^*\}}$, and define $h(\mathbf{t}^*) := \psi(s(1 + \sum_{i\in\{j:w_j\neq w^*\}} t_i))$. A Taylor expansion for $h(\mathbf{t}^*)$ about $\mathbf{0}$ leads that

$$h(\mathbf{t}^*) \sim \psi(s) + s\psi'(s) \times \sum_{i\in\{j:w_j\neq w^*\}} t_i. \tag{5.17}$$

Then choosing $s = \psi^{-1}(uw^*)$ in (5.17) proves the claim, and $\epsilon(u, \mathbf{w}) = o(u), u \to 0^+$ is due to Lemma 5.5. $\square$

Proof of Proposition 5.15: It is well known that CTE is subadditive for continuous risks (Denuit et al., 2005); that is, for any $\mathbf{X}' \in \mathcal{R}(F_1, \ldots, F_d)$

$$\mathrm{CTE}_p(X_1' + \cdots + X_d') \leq \sum_{i=1}^{d} \mathrm{CTE}_p(X_i'), \quad 0 < p < 1.$$

From Proposition 4.15, for such an $\mathbf{X}$,

$$\mathrm{CTE}_p(X_1 + \cdots + X_d) \sim \sum_{i=1}^{d} \mathrm{CTE}_p(X_i) = \sum_{i=1}^{d} \mathrm{CTE}_p(X_i'), \quad p \to 1^-.$$

Thus,

$$\text{CTE}_p(X_1 + \cdots + X_d) \gtrsim \text{CTE}_p(X'_1 + \cdots + X'_d), \quad p \to 1^-.$$

$\square$

Proof of Proposition 5.17: By Theorem 4.3 of Embrechts et al. (2009a), VaR_p is asymptotically subadditive for $\mathbf{X}'$ as $p \to 1^-$. Also, $\mathbf{X}$ satisfies Assumption 5.1, and thus VaR_p is asymptotically additive for $\mathbf{X}$. So the above asymptotic inequality holds, and tail comonotonicity becomes an asymptotically worst dependence structure in this sense. $\square$

Proof of Proposition 5.18: By Theorem 6.8 of McNeil et al. (2005), VaR_p is subadditive for $\mathbf{X}'$ when $0.5 \leq p < 1$. Also, notice that if the assumption that $\mathbf{X}$ is non-negative in Proposition 4.14 is relaxed, the asymptotic additivity of VaR still holds. Note that the $\mathbf{X}$ here is not elliptical; within elliptical distribution families, there are not non-trivial tail comonotonic distributions (see Chapter 4). $\square$

Chapter 6

Second order regular variation and risk measures

6.1 Introduction

In actuarial science, some statistical quantities about a random variable X are often referred to as risk measures. Among many of them, value at risk (VaR) and conditional tail expectation (CTE) are the most popular risk measures. For $0 < p < 1$ and usually with $p > 0.9$, VaR_p refers to the $100p$ percentile of loss and CTE_p is a conditional expectation given the loss exceeds VaR_p. Both of them have been adopted by regulations of insurers. For example, CTE has been required for calculating the relevant risks of segregated fund in Canada (OSFI, 2011, Chapter 8). VaR has been commonly-used for financial and insurance risks while recently CTE has been suggested to be a more conservative risk measure compared to VaR, especially when loss distributions are heavy-tailed. We refer to Denuit et al. (2005) and McNeil et al. (2005) for relevant risk measures and their relationships. More precisely, the class of heavy-tailed distributions is a very large family and can be defined as $\mathcal{K} = \{F \text{ df on } (0, \infty) : \int_0^{+\infty} e^{\epsilon x} dF(x) = \infty \text{ for all } \epsilon > 0\}$ (See Figure 1.4.1 of Embrechts et al. (1997) for the classes of heavy-tailed distributions). In this chapter, we only study an important subclass of the heavy-tailed distributions, consisting of distributions that are supported on the positive real line and have regularly varying survival functions. For narrative convenience, we use the term “heavy-tailed distribution” for this subclass of heavy-tailed distributions.

From the viewpoint of risk management, tail behavior of loss random variables is important. In the literature, much work has been done to better understand the extremal patterns of aggregate losses by doing asymptotic analysis, such as, Alink et al. (2004, 2005, 2007); Barbe et al. (2006); Embrechts et al. (2009b). In these papers, asymptotic behavior of aggregate dependent losses has been studied. We believe that the asymptotic study on the tail behavior is relevant to understand the magnitude of large losses and the performance of corresponding risk measures. The benefit of doing asymptotic analysis on risk measures, on the other hand, may establish asymptotic relationships between risk measures. For example, as a coherent risk measure, CTE has many advantages over VaR; we refer to Artzner et al. (1999) for limitations of VaR compared with a coherent risk measure. However, computing CTE is often more costly than estimating VaR since the former often involves Monte Carlo simulations for non-closed forms of integrations. To this end, studying asymptotic relationships between CTE and VaR becomes promising. Moreover, by studying the asymptotic relationship between CTE and VaR, we will be able to have better sense of how VaR may underestimate heavy-tailed risks, and we refer to Figure 6.1 for an example.

Some asymptotic analysis of CTE in terms of VaR has been done most recently by Zhu and Li (2012). In their paper, the non-negative random vector under study has the form of $\mathbf{X} = (T_1R, \dots, T_dR)$, where R is non-negative, independent of $\mathbf{T} = (T_1, \dots, T_d) \geq \mathbf{0}$ and has a regularly varying survival function. Zhu and Li (2012) derived the closed-form first order approximation for CTE of the following form: $\mathbb{E}[X_1 | \|\mathbf{X}\| > \mathrm{VaR}_p(\|\mathbf{X}\|)]$, as $p \to 1$. Their result is interesting since the closed-form approximation of CTE only relies on some mild moment conditions on T_i's. The results obtained in Zhu and Li (2012) is only a first order approximation of the relationship between CTE and VaR, as $p \to 1$. The first order property is useful since we have known more about VaR while the asymptotic relationship will provide some insights about CTE. However, the limitation is also obvious that the first order asymptotic approximation may not be good for a sub-extremal threshold. So, second order properties would be able to provide more insightful information. On one hand, the second order properties determine the rates of convergence for the first order approximation, and when the speed of convergence is higher, the asymptotic approximation by only the first order term has a better chance to provide a good approximation for the sub-extremal level; on the other hand, the second order term itself improves the approximation.

In the asymptotic analysis of CTE, second order regular variation (2RV) provides a tractable tool for studying second order properties. 2RV was originally studied in the extreme value theory, to study the speed of convergence of the extreme value condition (de Haan and Ferreira, 2006, Section 2). For a general theory of 2RV, we refer to de Haan and Ferreira (2006); de Haan and Stadtmüller (1996) for references. In this chapter, we use 2RV to study the tail behavior of the distribution function $F(x)$, which is more naturally related to the study of risk measures. In Geluk et al. (1997), 2RV on survival functions was studied in the context of convolution and so on. We refer to Degen et al. (2010) for a recent study on risk concentration and diversification under the assumption of 2RV, which studied risk concentration and diversification benefit of multiple losses.

In general, we are interested in the behavior of

$$\mathbb{E}[X_i | g(X_1, \ldots, X_d) > t], \quad t \to \infty, \tag{6.1}$$

where g is a loss aggregate function, and $(X_1, \ldots, X_d)$ has a general dependence structure and the margins have some particular tail patterns. However, the scale mixture approach studied in Zhu and Li (2012) provides a wide subset of such random vector $\mathbf{X}$. We will follow this approach in this chapter and study the second order property of (6.1) for the case where g is a homogeneous function of order 1 and the non-negative $\mathbf{X}$ is constructed by the scale mixture approach. More specifically, we will study the rates of convergence of asymptotic relationships between CTE and VaR under the framework of 2RV. Closed-form second order approximations have been obtained for both the univariate case and the multivariate case where the random vector is constructed from a scale mixture with a heavy-tailed non-negative random variable. For both cases, we find that the first order approximation is affected by only the regular variation index $-\alpha$ of marginal survival functions, while the second order approximation is influenced by both the parameters for first and second order regular variation, and the convergence speed is dominated by the second order parameter only. Many well-known continuous distributions for modeling univariate/multivariate losses or financial returns — such as the Student t_ν, Burr, multivariate t_ν, multivariate distributions constructed by regularly varying margins and some upper tail dependent copulas, and some elliptical distributions with regularly varying margins — satisfy the second order conditions that we will study in this chapter. Some side results are second order properties for the well-known Karamata's Theorem and Breiman's Theorem.

This chapter is organized as following: In Section 6.2, notation, preliminary concepts and results are presented. In Section 6.3, the univariate case with 2RV conditions is studied to understand the asymptotic relationship between CTE and VaR. The study on the multivariate case is reported in Section 6.4. Some discussion of future research is in Section 6.5. Finally, Section 6.6 presents all the proofs.

6.2 Preliminaries

Fundamental knowledge about regular variation has been introduced in Section 2.3. For the multivariate case to be studied in Section 6.4, we will study positive random vector $\mathbf{X}$ constructed as this form $(X_1, \ldots, X_d) = (RT_1, \ldots, RT_d)$, where the tail behavior of product of two random variables plays an important role. The following result just tells us how the product will inherit the tail behavior of the random variable that has a heavier tail.

Proposition 6.1 *(Breiman's Theorem) Suppose X is a non-negative random variable such that its survival function $\overline{F}(x) \in \mathrm{RV}_{-\alpha}$ with $\alpha > 0$, and $Y \geq 0$ is a random variable, independent of X, with $\mathbb{E}[Y^{\alpha+\epsilon}] < \infty$ for some $\epsilon > 0$. Then*

$$\mathbb{P}[XY > x] \sim \mathbb{E}[Y^\alpha]\mathbb{P}[X > x], \quad x \to \infty. \tag{6.2}$$

That is, if X is regularly varying, then its right tail behavior will be inherited by the product XY, where Y has a lighter right tail and a finite moment of order higher than $\mathbb{E}[Y^\alpha]$. We refer to Breiman (1965) for a proof of the case where $0 < \alpha < 1$, and the proof is also adaptable for proving the case where $0 < \alpha$. For some multivariate versions of Breiman's Theorem, we refer to Resnick (2007).

The following definition of the second order regular variation comes from de Haan and Stadtmüller (1996), de Haan and Ferreira (2006) and Geluk et al.

(1997). See also Neves (2009) for a slightly different version of extended second order regular variation and Wang and Cheng (2005) for third and higher order regular variation.

Definition 6.1 If the survival function of a non-negative random variable X is $\overline{F} := 1 - F$ and $\overline{F} : [0, \infty) \to (0, 1]$ satisfies that $\overline{F} \in \mathrm{RV}_{-\alpha}$ with $\alpha > 0$. Then $\overline{F}$ is said to be of second-order regular variation with parameter $\rho \leq 0$, if there exists a function $A(t)$ that ultimately has a constant sign with $lim_{t\to\infty} A(t) = 0$ and a constant $c \neq 0$ such that

$$\lim_{t\to\infty} \frac{\overline{F}(tx)/\overline{F}(t) - x^{-\alpha}}{A(t)} = H_{\alpha,\rho}(x) = cx^{-\alpha} \int_1^x u^{\rho-1} du, \quad x > 0. \tag{6.3}$$

Then it is written as $\overline{F} \in 2\mathrm{RV}_{-\alpha,\rho}$ and $A(t)$ is referred to as the auxiliary function of $\overline{F}$.

It is known from de Haan and Stadtmüller (1996) or a more relevant form in Geluk et al. (1997) that if $H_{\alpha,\rho}(x)$ is not a multiple of $x^{-\alpha}$ then $\rho < 0$ implies that there exists a $c \neq 0$ such that

$$H_{\alpha,\rho}(x) = cx^{-\alpha} \frac{x^\rho - 1}{\rho},$$

$|A| \in \mathrm{RV}_\rho$ and no other choices of ρ are consistent with $A(t) \to 0$. Moreover, convergence in (6.3) is uniform in x on compact intervals of $(0, \infty)$. By adjusting $A(t)$, we can always let $c = 1$, and unless otherwise specified, c is assumed to be 1. We only study the case where $\rho < 0$ and $\alpha > 1$ in this chapter. Also, note that for second order regular variation, $\rho \neq -\infty$ (Resnick, 2007, page 68). Therefore, the Pareto distribution with the cdf $F(x) = 1 - (k/x)^\alpha, x \geq k$, where $H_{\alpha,\rho}(x)$ in (6.3) would be 0 for any function $A(t)$, does not satisfy the conditions of second order regular variation and thus will not be studied in this chapter.

To illustrate the main result for the multivariate case, we will use the concept of Archimedean copula as a mixture.

6.3 Univariate cases

For the univariate case, a direct application of Karamata's Theorem (Theorem 2.2) shows that, for a non-negative random variable $X \in \mathrm{RV}_{-\alpha}$ with $\alpha > 1$, $\mathrm{CTE}_p(X)$ is finite and

$$\mathrm{CTE}_p(X) \sim \frac{\alpha}{\alpha - 1} \mathrm{VaR}_p(X), \quad p \to 1. \tag{6.4}$$

From (6.4), we know that when p is *sufficiently* large, CTE and VaR of a regularly varying random variable have a deterministic relationship that is determined by the regular variation index $-\alpha$. However, we do not know how close is the right-hand side of (6.4) to its left-hand side. In practice, we always choose a level p that is strictly less than 1 to evaluate risks. To this end, a higher order approximation is useful. A closed-form second order approximation is ideal. If one can not obtain the closed-form second order approximation, knowledge about how fast the right-hand side converges to the left-hand size of (6.4) is also informative. In this section, we will derive the closed-form second order approximation for the univariate case assuming that the loss random variable is not only regularly varying with index $-\alpha$, but also second regularly varying. This assumption is mild, since some useful parametric distributions, such as Student t_ν, satisfy it (see Example 6.3). Moreover, as a technical tool for our proof, Proposition 6.5 at the end of this section shows some interesting second order properties with respect to Karamata's Theorem.

Now, we are ready to give a technical lemma showing that the slowly varying function involved in the second order regular variation has a certain nice form.

Lemma 6.2 *Suppose* $\overline{F} \in 2\mathrm{RV}_{-\alpha,\rho}, \alpha > 0, \rho < 0$*, then we can write* $\overline{F}(t) = kt^{-a}\ell(t), t > 0$ *with some* $k > 0$ *and* $\lim_{t\to\infty} \ell(t) = 1$*, and moreover we have* $|1 - \ell(t)| \in \mathrm{RV}_\rho$.

In what follows, we give a result of a uniform inequality for the second order regular varying survival distribution function. This result is essentially Theorem 2.3.9 of de Haan and Ferreira (2006) and will be useful to prove our subsequent propositions. We refer to de Haan and Ferreira (2006) and Cheng and Jiang (2001) for relevant discussion.

Proposition 6.3 *Suppose $\overline{F} \in 2\mathrm{RV}_{-\alpha,\rho}$ with $\alpha > 0$ and $\rho < 0$, then for any $\epsilon, \delta > 0$, there exists $t_0 = t_0(\epsilon, \delta)$ such that for all $t, tx \geq t_0$ and $x > 0$,*

$$\left| \frac{\overline{F}(tx)/\overline{F}(t) - x^{-\alpha}}{a(t)} - x^{-\alpha} \frac{x^{\rho} - 1}{\rho} \right| \leq \epsilon \max \left(x^{-\alpha+\rho+\delta}, x^{-\alpha+\rho-\delta} \right), \qquad (6.5)$$

where $a(t) = -\rho[1 - \ell(t)]/\ell(t)$ with $\overline{F}(t) = kt^{-\alpha}\ell(t)$, $0 < k < +\infty$ and $\lim_{t\to\infty} \ell(t) = 1$.

Remark 6.1 From Proposition 6.3, we may choose the $A(t)$ in Definition 6.1 as $A(t) := a(t) = -\rho[1 - \ell(t)]/\ell(t)$, and clearly $|A(t)| \in \mathrm{RV}_{\rho}$.

The following result tells us that the rates of approximation of $\mathrm{CTE}_p(X)$ to $\frac{\alpha}{\alpha-1}\mathrm{VaR}_p(X)$ are determined by the second order parameter ρ and are of the order of $[\mathrm{VaR}_p(X)]^{\rho+1}$. When ρ is smaller (more negative), the speed of convergence is faster. In other words, when ρ is more negative, the first order term can provide a better approximation.

Proposition 6.4 *Suppose that a non-negative random variable $X \in 2\mathrm{RV}_{-\alpha,\rho}$ with $\alpha > 1$ and $\rho < 0$, and write its survival function $\overline{F}(t) := kt^{-\alpha}\ell(t)$. Then*

$$\mathrm{CTE}_p(X) \sim \frac{\alpha}{\alpha - 1}\mathrm{VaR}_p(X) + \eta(\mathrm{VaR}_p(X)), \quad p \to 1, \qquad (6.6)$$

where

$$\eta(t) = \frac{ta(t)}{(\alpha - 1 - \rho)(\alpha - 1)}, \qquad (6.7)$$

and $a(t)$ can be chosen as $a(t) := -\rho[1 - \ell(t)]/\ell(t)$.

Remark 6.2 Since $|a(t)| \in \mathrm{RV}_{\rho}$ and thus $|ta(t)| \in \mathrm{RV}_{\rho+1}$, Equation (6.7) just shows that the second order term becomes more important as $\rho > -1$ (because $|ta(t)|$ is increasing as $t \to \infty$). So we comment that the second order approximation is especially useful for the case where $-1 < \rho < 0$, in which the second order term may not be ignored. In Example 6.5, we will illustrate by an example for a multivariate case that the second order term significantly improves the approximation.

Remark 6.3 Proposition 6.4 just indicates that the first order term depends on α only, and the second order term depends on both α and ρ. The proof in Proposition 6.4 also shows that for $\alpha > 1$ and $\rho < 0$, as $t \to \infty$, $\frac{\overline{F}(tx)/\overline{F}(t)-x^{-\alpha}}{A(t)}$ converges to $x^{-\alpha}\frac{x^{\rho}-1}{\rho}$ uniformly in x on any set of the form $[x_0, \infty)$, where $x_0 > 0$. Then we can exchange the orders of limit and integral operations in order to get $\lim_{t\to\infty}\int_a^{\infty}\frac{\mathbb{P}[X>ts]/\mathbb{P}[X>t]-s^{-\alpha}}{A(t)}ds$ for some $a > 0$.

Remark 6.4 A necessary condition for the asymptotic relationship between CTE and VaR is that, the random variable can not be too heavy-tailed; that is, the regular variation index α should be greater than 1. When $\alpha \le 1$, the tail behavior becomes much more complicated. In particular, as $\alpha < 1$, $\text{CTE}_p(X) = \infty, \forall p < 1$, and under the current framework, one cannot study the asymptotic relationship between CTE and VaR.

Example 6.1 (Hall/Weiss class) Suppose X is a random variable with support $[0, \infty)$ and distribution function F such that with $x \ge 1, \alpha > 1, \rho < 0$, $\overline{F}(x) = \frac{1}{2}x^{-\alpha}(1 + x^{\rho})$. Then clearly, $\overline{F} \in 2\text{RV}_{-\alpha,\rho}$, and when $t := \text{VaR}_p(X)$ is big enough,

$$\text{CTE}_p(X) = t + \frac{\int_t^{\infty}\frac{1}{2}x^{-\alpha}(1+x^{\rho})dx}{\frac{1}{2}t^{-\alpha}(1+t^{\rho})} = t + \frac{t/(\alpha-1) + t^{\rho+1}/(\alpha-1-\rho)}{1+t^{\rho}},$$

which implies, with $(1+t^{\rho})^{-1} \sim 1 - t^{\rho}, t \to +\infty$,

$$\text{CTE}_p(X) - \frac{\alpha}{\alpha-1}\text{VaR}_p(X) \sim \frac{\rho}{(\alpha-1-\rho)\times(\alpha-1)}[\text{VaR}_p(X)]^{\rho+1}, \quad p \to 1.$$

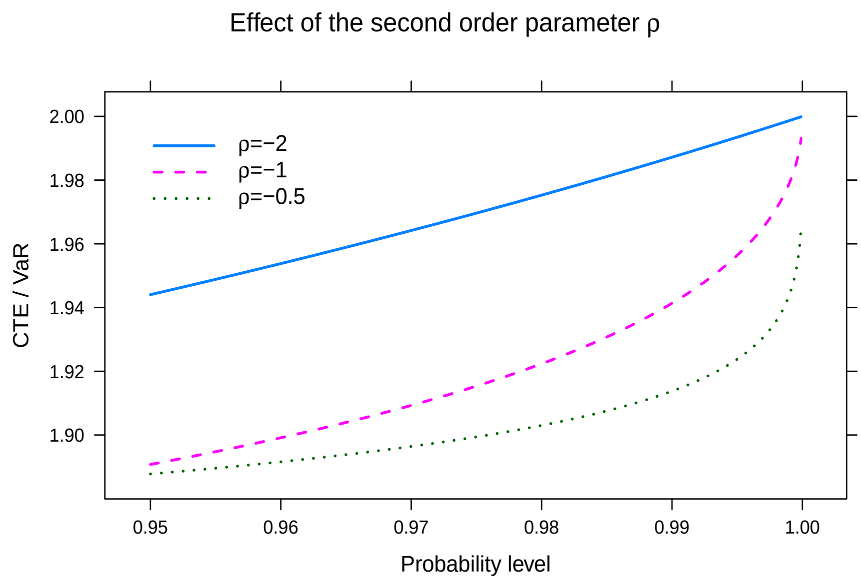
It means that the rate of convergence is determined by the second order parameter ρ and is of the order of $[\text{VaR}_p(X)]^{\rho+1}$. Figure 6.1 illustrates the effects of ρ.

Example 6.2 (Burr distribution) Consider the Burr distribution with the survival function $\overline{F}(x) = (1+x^b)^{-a}, x, a, b > 0$, and the second order expansion leads to

$$\overline{F}(x) = x^{-ab}[1 - ax^{-b} + o(x^{-b})], \quad x \to \infty,$$

so that $\overline{F} \in 2\text{RV}_{-ab,-b}$.

Figure 6.1: Sub-extremal relationships between CTE and VaR ($\alpha = 2$) for Hall/Weiss class



Now we evaluate how large p must be for the second order approximation of Proposition 6.4 to be decent for the Burr distribution. Assume $X \sim Burr(a, b, \sigma)$, where a, b are the two shape parameters and σ is the scale parameter; that is, $X \sim F(x) := 1 - \left(1 + (x/\sigma)^b\right)^{-a}$. Write $\overline{F}(x) = (1 + (x/\sigma)^b)^{-a} = \sigma^{ab} x^{-ab} \ell(x)$, where $\ell(t) = ((t/\sigma)^{-b} + 1)^{-a}$ is a slowly varying function. According to Proposition 6.4, an auxiliary function for the Burr distribution can be chosen as

$$A(t) := b[1 - \ell(t)]/\ell(t) = b\left[\left(1 + (t/\sigma)^{-b}\right)^a - 1\right].$$

Clearly, $A \in \mathrm{RV}_{-b}$. Then, letting $t := \mathrm{VaR}_p(X)$,

$$
\begin{aligned}
\mathrm{CTE}_p(X) &= t + (1-p)^{-1} \int_t^\infty \left(1 + (x/\sigma)^b\right)^{-a} dx \\
&= t + \frac{\sigma}{b(1-p)} \int_0^{(1+(t/\sigma)^b)^{-1}} (1-y)^{1/b-1} y^{a-1/b-1} dy \\
&= t + \frac{\sigma}{b(1-p)} Beta\left((1+(t/\sigma)^b)^{-1}; a-1/b, 1/b\right), \qquad (6.8)
\end{aligned}
$$

where $Beta(\cdot; a-1/b, 1/b)$ is an incomplete Beta function with parameters $a - 1/b, 1/b$. We can use (6.8) to get the exact calculation, and then compare it to the second order approximation.

Table 6.1 presents the values of p for which a decent second order approximation can be obtained. For a given leading parameter (ab) for the tail heaviness, a larger b leads to a smaller p; that is, the second order parameter $(\rho = -b)$ is more negative and thus the rates of convergence becomes faster. More interestingly, for a given b, the second order approximation for a lighter tail (larger α) is not as good as that for a heavier tail. When the tail becomes very light (say, $\alpha \geq 4$), it was hard to get a good second order approximation within the 5% error band. The smallest p's required that are reported in Table 6.1 do not rely on the scale parameter. However, theoretically, there are infinitely many choices for the auxiliary function $A(t)$. The smallest p required to get a decent second order approximation heavily rely on the choice of $A(t)$.

Table 6.1: How large p must be to get a good second order approximation: the values are the corresponding p for which the absolute difference between the second order approximation and the true value is 5% of the true value.

	$\alpha(=ab)$			
b	1.1	1.5	2	3
0.5	0.86977	0.98828	0.99859	0.99994
1	0.32947	0.73711	0.88890	0.97825
1.5	0.11240	0.44426	0.65444	0.86267

The next result extends Karamata's theorem to a second order regular condition for the case where regular variation index $-\alpha < -1$. Since some continuous random variables, such as the Student t_ν random variable, have closed-form density functions but do not have closed-form cdf's, the following result is very useful to derive second order properties, such as the second order parameter and the corresponding auxiliary function, for their survival functions from their density functions.

Proposition 6.5 *Let $g \in 2\mathrm{RV}_{-\alpha,\rho}, \alpha > 1, \rho < 0$, with an auxiliary function $A(x)$, and define $g^*(t) := \int_t^\infty g(x)dx$. Then $g^* \in 2\mathrm{RV}_{-\alpha+1,\rho}$, the corresponding auxiliary function is $A^*(t) = \frac{\alpha-1}{\alpha-1-\rho}A(t)$, and*

$$g^*(t) \sim \frac{t}{\alpha-1}g(t) + \frac{tA(t)}{\rho}\left(\frac{1}{\alpha-\rho-1} - \frac{1}{\alpha-1}\right)g(t), \quad t \to \infty. \tag{6.9}$$

Example 6.3 (Student t_ν distribution) Consider the standard Student t_ν distribution with density function

$$f(x) = \frac{\Gamma((\nu+1)/2)}{\sqrt{\nu\pi}\Gamma(\nu/2)}\left(1+\frac{x^2}{\nu}\right)^{-(\nu+1)/2}, \quad \nu > 1.$$

Then it can be verified that

$$f(x) \sim \frac{\Gamma((\nu+1)/2)}{\sqrt{\nu\pi}\Gamma(\nu/2)}\nu^{(\nu+1)/2}x^{-\nu-1}\left(1 - 2^{-1}\nu(\nu+1)x^{-2}\right), \quad x \to \infty.$$

That is, $f(x) \in 2\mathrm{RV}_{-\nu-1,-2}$. By Proposition 6.5, its survival function $\overline{F} \in 2\mathrm{RV}_{-\nu,-2}$. For the Student t_ν distribution, the second order parameter is always that $\rho = -2$. The second order property can also be obtained from scale mixture representation of the t_ν distribution. In Example 6.4 in the next section, we will derive the same thing from a mixture with a generalized inverse Gamma distribution.

6.4 Multivariate cases

From the actuarial viewpoint, we are interested in CTEs of the form $\mathbb{E}[X_1|g(X_1,...,X_d) > t]$, and in particular $\mathbb{E}[X_1|X_2 > t]$ for the bivariate case, where t is large, and $X_1,...,X_d$ possess some general dependence structure and are heavy-tailed loss random variables (with the same RV index). In Chapter 3, the concept of *tail order* has been proposed to study the degree of positive tail association that covers asymptotic dependence and asymptotic independence. In Chapter 3, the upper tail order of a copula C is denoted as $\kappa_U(C)$. And $\kappa_U(C) = 1$ corresponds to the usual upper tail dependence case and $\kappa_U(C) > 1$ covers asymptotic independence and even negative dependence cases. For $E(X_1|X_2 > t)$, some examples and rough analysis suggests that under some regularity conditions on the copula C of (X_1, X_2),

$$\begin{aligned} \kappa_U(C) = 1 &\Rightarrow \mathbb{E}[X_1|X_2 > t] = O(t), \quad t \to \infty; \\ \kappa_U(C) > 1 &\Rightarrow \mathbb{E}[X_1|X_2 > t] = o(t), \quad t \to \infty. \end{aligned}$$

However, it is difficult to find sufficient conditions to deduce general results for the order of t and the leading coefficient. In Chapter 7, we have more discussions on the tail behavior of the above conditional specifications. The study of second order terms would depend on developing a theory for the second order tail functions of copulas (Joe and Li, 2011).

A special example with multivariate regular variation and upper tail dependence is the case of heavy-tailed scale mixtures studied by Zhu and Li (2012). They obtained $\mathbb{E}[X_1|X_2 > t] \sim at, t \to \infty$, with a being a constant relying on some finite moments. In this section, one of our main results is to assume 2RV and get an extension to Zhu and Li's result to a second order approximation. We also show that heavy-tailed scale mixtures with 2RV conditions covers some parametric distribution families that are commonly used in actuarial science; that is, it is not too restrictive for applications.

Now, consider bivariate cases first. Let

$$(X_1, X_2) = (RT_1, RT_2),$$

where $(T_1, T_2) \geq \mathbf{0}$, independent of $R \geq 0$. Note here we do not specify the dependence structure between T_1 and T_2. This class of distributions covers multivariate Pareto distributions and truncated multivariate elliptical distributions, etc. With some regularity conditions on R and moment conditions on T_1 and T_2, the following asymptotic relationship has been derived in Zhu and Li (2012),

$$\mathbb{E}[X_1|X_2 > \text{VaR}_p(X_2)] \sim \frac{\alpha}{\alpha-1}\frac{\mathbb{E}[T_1T_2^{\alpha-1}]}{\mathbb{E}[T_2^{\alpha}]}\text{VaR}_p(X_2), \quad p \to 1. \tag{6.10}$$

In this section, we obtain the rate of the convergence of (6.10) and the second order approximation.

Proposition 6.6 *Let $(X_1, X_2) = (RT_1, RT_2)$, where $(T_1, T_2) \geq \mathbf{0}$, independent of $R \geq 0$ such that the survival function of R is $\overline{F}_R(t) \in 2\text{RV}_{-\alpha,\rho}$ with $\alpha > 1$ and $\rho < 0$. If $\mathbb{E}[T_1] < \infty$, $\mathbb{E}[T_1T_2^{\alpha-1-\rho+\xi}] < \infty$ for some $\xi > 0$ and $\mathbb{E}[T_2^{\alpha-\rho+\delta}] < \infty$ for some $\delta > 0$, then an extension of (6.10) can be represented by*

$$\mathbb{E}[X_1|X_2 > \text{VaR}_p(X_2)] \sim \eta_1(\text{VaR}_p(X_2)) + \eta_2(\text{VaR}_p(X_2)), \quad p \to 1,$$

where

$$\begin{aligned}
\eta_1(t) &= \frac{\alpha}{\alpha-1}\frac{\mathbb{E}[T_1T_2^{\alpha-1}]}{\mathbb{E}[T_2^{\alpha}]}t;\\
\eta_2(t) &= \frac{A(t)t}{\rho}\left\{\frac{\alpha-\rho}{\alpha-1-\rho}\frac{\mathbb{E}[T_1T_2^{\alpha-1-\rho}]}{\mathbb{E}[T_2^{\alpha}]} - \frac{\alpha}{\alpha-1}\frac{\mathbb{E}[T_1T_2^{\alpha-1}]\mathbb{E}[T_2^{\alpha-\rho}]}{(\mathbb{E}[T_2^{\alpha}])^2}\right\}.
\end{aligned}$$

Here, $A(t)$ is the auxiliary function for $\overline{F}_R(t)$.

Remark 6.5 Similarly to the univariate case, the first order term only depends on α, and the second order term depends on α and ρ. In order to get a higher order approximation, existence of higher moments of the form $\mathbb{E}[T_1T_2^{\alpha-1-\rho+\xi}]$ for some $\xi > 0$ and $\mathbb{E}[T_2^{\alpha-\rho+\delta}] < \infty$ for some $\delta > 0$ are necessary. The speed of convergence is dominated by the second order parameter ρ only. Since $|A(t)| \in \text{RV}_\rho$, the convergence speed of $\mathbb{E}[X_1|X_2 > \text{VaR}_p(X_2)] \to \frac{\alpha}{\alpha-1}\frac{\mathbb{E}[T_1T_2^{\alpha-1}]}{\mathbb{E}[T_2^{\alpha}]}\text{VaR}_p(X_2)$ is of the order of $[\text{VaR}_p(X_2)]^{\rho+1}$; that is, when ρ is more negative, the convergence is faster at the sub-extremal level. It is also consistent with the univariate case.

In Proposition 6.6, we do not specify the dependence structure between T_1 and T_2. If we have a random vector $\mathbf{T} \geq 0$ and the conditioning event can be of the form $h(\mathbf{X}) > t$ where $h : \mathbb{R}_+^d \to \mathbb{R}_+$ is homogeneous of order 1, then $h(\mathbf{X}) = h(\mathbf{T})R$ and we can take $T_2 := h(\mathbf{T})$.

Proposition 6.7 *Suppose there is a function $h : \mathbb{R}_+^d \to \mathbb{R}_+$, homogeneous of order 1, and a random vector $\mathbf{X} := (RT_1, \ldots, RT_d)$ where $\mathbf{T} = (T_1, \ldots, T_d) \geq \mathbf{0}$ independent of $R \geq 0$ and $\overline{F}_R \in 2\text{RV}_{-\alpha,\rho}$ with $\alpha > 1$ and $\rho < 0$. If $\mathbb{E}[T_1] < \infty$, $\mathbb{E}[T_1 h^{\alpha-1-\rho+\xi}(\mathbf{T})] < \infty$ for some $\xi > 0$ and $\mathbb{E}[h^{\alpha-\rho+\delta}(\mathbf{T})] < \infty$ for some $\delta > 0$, then*

$$\mathbb{E}[X_1 | h(\mathbf{X}) > \text{VaR}_p(h(\mathbf{X}))] \sim \eta_1(\text{VaR}_p(h(\mathbf{X}))) + \eta_2(\text{VaR}_p(h(\mathbf{X}))), \quad p \to 1,$$

where

$$\eta_1(t) = \frac{\alpha}{\alpha - 1} \frac{\mathbb{E}[T_1 h^{\alpha-1}(\mathbf{T})]}{\mathbb{E}[h^\alpha(\mathbf{T})]} t;$$
$$\eta_2(t) = \frac{A(t)t}{\rho} \left\{ \frac{\alpha - \rho}{\alpha - 1 - \rho} \frac{\mathbb{E}[T_1 h^{\alpha-1-\rho}(\mathbf{T})]}{\mathbb{E}[h^\alpha(\mathbf{T})]} - \frac{\alpha}{\alpha - 1} \frac{\mathbb{E}[T_1 h^{\alpha-1}(\mathbf{T})]\mathbb{E}[h^{\alpha-\rho}(\mathbf{T})]}{(\mathbb{E}[h^\alpha(\mathbf{T})])^2} \right\}.$$

Remark 6.6 Examples of the homogeneous h function are norms of vectors, such as the l_1-norm and l_∞-norm. These two norms correspond to the CTEs of a marginal risk X_1 respectively conditioning on the events that the sum of all marginal risks and the maximum of all marginal risks are higher than a threshold.

Example 6.4 A d-dimensional elliptical distribution can be constructed by $\mathbf{X} = (X_1, \ldots, X_d) = RD(U_1, \ldots, U_m)$, where $(U_1, \ldots, U_m)$ is uniformly distributed on the surface of the unit hypersphere in $\mathbb{R}^m$, D is a $d \times m$ matrix with DD^{T} positive semi-definite, and R is a positive random variable with $\overline{F}_R \in 2\text{RV}_{-\alpha,\rho}$ and the auxiliary function is $A(x)$. We may take $\mathbf{T} = (T_1, \ldots, T_d) = D(U_1, \ldots, U_m)$, and $T_i^+ = 0 \vee T_i$ for $i = 1, \ldots, d$, then the form $\mathbf{X}^+ = R\mathbf{T}^+$ fits into the study of this chapter. $\mathbf{X}^+$ follows a left-truncated [1] elliptical distribution. Taking the homogeneous function h in Proposition 6.7 to be the l_1 norm, and letting $t := \text{VaR}_p\left(\sum_{i=1}^d X_i^+\right)$, then we have for any $k \in \{1, \ldots, d\}$,

$$\mathbb{E}\left[X_k^+ \middle| \sum_{i=1}^d X_i^+ > t\right] \sim \eta_1(t) + \eta_2(t), \quad p \to 1,$$

[1] Unless otherwise stated, left-truncated in this chapter refers to left-truncating below zeros for each element of a random vector.

where

$$\eta_1(t) = \frac{\alpha}{\alpha-1}\frac{\mathbb{E}[T_k(\sum_{i=1}^d T_i)^{\alpha-1}]}{\mathbb{E}[(\sum_{i=1}^d T_i)^{\alpha}]}t;$$

$$\eta_2(t) = \frac{A(t)t}{\rho}\left\{\frac{\alpha-\rho}{\alpha-\rho-1}\frac{\mathbb{E}[T_k(\sum_{i=1}^d T_i)^{\alpha-1-\rho}]}{\mathbb{E}[(\sum_{i=1}^d T_i)^{\alpha}]}\right.$$
$$\left. -\frac{\alpha}{\alpha-1}\frac{\mathbb{E}[T_k(\sum_{i=1}^d T_i)^{\alpha-1}]\mathbb{E}[(\sum_{i=1}^d T_i)^{\alpha-\rho}]}{(\mathbb{E}[(\sum_{i=1}^d T_i)^{\alpha}])^2}\right\}.$$

In particular, we may choose a bivariate standard normal random vector (T_1, T_2) with identical margins, correlation $0 < \varrho < 1$, and let $T_i^+ = 0 \vee T_i$ for $i = 1, 2$, and $R = Y^{-1/2}$ with Y following Gamma$(\nu/2, \nu/2)$ in which $\nu/2$ are both shape and rate parameters. Then the scale mixture $(X_1, X_2) = (RT_1, RT_2)$ is a bivariate t_ν distribution, and $(X_1^+, X_2^+) = (RT_1^+, RT_2^+)$ is a left-truncated bivariate t_ν distribution. Note that Y^{-1} follows an Inverse Gamma distribution with both the shape and scale parameters being $\nu/2$. Thus the density function for Y^{-1} is

$$f_{Y^{-1}}(x) = \frac{(\nu/2)^{\nu/2}}{\Gamma(\nu/2)}x^{-\nu/2-1}e^{-x^{-1}\nu/2}, \quad x > 0; \nu > 0.$$

Therefore, the density function of R can be derived as

$$f_R(x) = 2f_{Y^{-1}}(x^2)x = 2\frac{(\nu/2)^{\nu/2}}{\Gamma(\nu/2)}x^{-\nu-1}e^{-x^{-2}\nu/2}$$
$$\sim 2\frac{(\nu/2)^{\nu/2}}{\Gamma(\nu/2)}x^{-\nu-1}(1 - x^{-2}\nu/2), \quad x \to \infty.$$

Then it can be verified that $f_R(x) \in 2\mathrm{RV}_{-\nu-1,-2}$ and the corresponding auxiliary function $A_{f_R}(x) = \nu x^{-2}$. Hence $\overline{F}_R(x) \in 2\mathrm{RV}_{-\nu,-2}$ by Proposition 6.5 and the corresponding auxiliary function $A_{\overline{F}_R}(x) = \frac{(\nu+1)-1}{(\nu+1)-1-(-2)}A_{f_R}(x) = \frac{\nu^2}{\nu+2}x^{-2}$. Moreover, from (6.25), we know that the product $X_i = RT_i$ will inherit the right tail behavior of R; that is, the univariate Student t_ν random variable $X_i \in 2\mathrm{RV}_{-\nu,-2}$, which is also consistent to Example 6.3, where the same conclusion is drawn directly from the density function of the t_ν distribution. Then more specifically, letting $\nu = 2$ and $t := \mathrm{VaR}_p(X_1^+ + X_2^+)$, then for the bivariate left-truncated

t_ν distribution,

$$\mathbb{E}\left[X_1^+ \,\middle|\, X_1^+ + X_2^+ > t\right] \sim \eta_1(t) + \eta_2(t), \quad p \to 1,$$

where

$$\eta_1(t) = 2\frac{\mathbb{E}[T_1^{+2}] + \mathbb{E}[T_1^+ T_2^+]}{\mathbb{E}[(T_1^+ + T_2^+)^2]}t;$$
$$\eta_2(t) = \left\{-\frac{2}{3}\frac{\mathbb{E}[T_1^+(T_1^+ + T_2^+)^3]}{\mathbb{E}[(T_1^+ + T_2^+)^2]} + \frac{\mathbb{E}[T_1^{+2} + T_1^+ T_2^+]\mathbb{E}[(T_1^+ + T_2^+)^4]}{(\mathbb{E}[(T_1^+ + T_2^+)^2])^2}\right\} A_{\overline{F}_R}(t)t.$$

The higher moments of truncated standard bivariate t_ν random variables can be calculated in what follows by some recurrence relations in Shah and Parikh (1964) or page 313–314 of Kotz et al. (2000). That is,

$$\eta_1(t) = t;$$
$$\eta_2(t) = g(\varrho)A_{\overline{F}_R}(t)t = g(\varrho)t^{-1}. \tag{6.11}$$

With $v_{r,s} := v_{r,s}(\varrho) = \mathbb{E}[(T_1^+)^r (T_2^+)^s]$ for non-negative integers r, s,

$$g(\varrho) = \frac{v_{4,0} + 4v_{3,1} + 3v_{2,2}}{6(v_{2,0} + v_{1,1})} = \frac{(1+\varrho)\{3[\frac{1}{2}\pi + \arcsin(\varrho)] + (4-\varrho)\sqrt{1-\varrho^2}\,\}}{3\{\frac{1}{2}\pi + \arcsin(\varrho) + \sqrt{1-\varrho^2}\,\}}.$$

$g(\varrho)$ is strictly increasing in $\varrho \in [-1, 1]$. From the equations in (6.11), we know that for this truncated bivariate standard t_ν distribution, the convergence speed to the first order approximation is dominated by $[\mathrm{VaR}_p(X_1^+ + X_2^+)]^{-1}$. When the correlation coefficient ϱ is larger — that is, when the dependence between margins is stronger — the second order approximation is more important. Some relevant observations for second order approximation of vector-valued CTE have also been reported in Example 3.2 of Joe and Li (2011). The effect from the tail heaviness of margins may also have subtle influence on the second order approximation, however we have to calculate higher moments in this case. A general conclusion about the effect of marginal heaviness on the second order approximation seems to be interesting and non-trivial to obtain.

However, following the discussion in Remark 6.2, for the above bivariate t

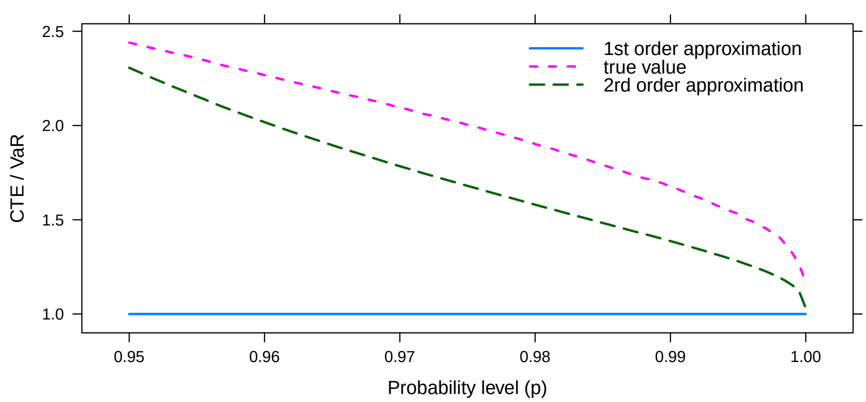
distribution, the second order parameter $\rho = -2$. So the second order term will be negligible when $[\mathrm{VaR}_p(X_1^+ + X_2^+)]$ is large enough. In the following Example 6.5, we will illustrate by an example for a case of $-1 < \rho < 0$ that the second order term significantly improves the approximation and can not be ignored.

Example 6.5 (Burr scale mixture of normal) Let $(X_1, X_2) := (Z_1^+ R, Z_2^+ R)$, where (Z_1^+, Z_2^+) is left-truncated standard bivariate normal with correlation $\varrho = 0.9$ and $R \sim Burr(a = 5, b = 0.4)$ (i.e., $\overline{F}_R \in 2\mathrm{RV}_{-2,-0.4}$, refer to Example 6.2) with a scale parameter 1000. In Figure 6.2, the vertical axis is the ratio of CTE to VaR

$$\frac{\mathrm{CTE}}{\mathrm{VaR}} := \frac{\mathbb{E}[X_1 | X_1 + X_2 > \mathrm{VaR}_p(X_1 + X_2)]}{\mathrm{VaR}_p(X_1 + X_2)},$$

defined as a function of p. The true value is approximated based on a Monte Carlo simulation with a sample size of 10^7, and the first and second order approximations are calculated based on Proposition 6.7.

Figure 6.2: First/Second order approximations for Burr scale mixture of normal



Example 6.6 Suppose T_1, T_2 are independent and exponentially distributed with survival function $F_{T_1}(x) = F_{T_2}(x) = 1 - e^{-x}$, $Q = 1/R$ and independent of

$T_1, T_2, \overline{F}_R(x) \in 2\mathrm{RV}_{-\alpha,\rho}$ and $(X_1, X_2) := (RT_1, RT_2)$. Then

$$\mathbb{P}[X_1 > x_1, X_2 > x_2] = \int_0^\infty e^{-(x_1+x_2)q} F_Q(dq) = \psi_Q(x_1 + x_2), \tag{6.12}$$

where F_Q and ψ_Q are the distribution function and LT of Q, respectively. Take $x_1 = \psi_Q^{-1}(u_1)$ and $x_2 = \psi_Q^{-1}(u_2)$, and then

$$\mathbb{P}[X_1 > \psi_Q^{-1}(u_1), X_2 > \psi_Q^{-1}(u_2)] = \psi_Q(\psi_Q^{-1}(u_1) + \psi_Q^{-1}(u_2)) =: C_{\psi_Q}(u_1, u_2) \tag{6.13}$$

Also, we have $\mathbb{P}[X_i > x] = \psi_Q(x)$ for $i = 1, 2$. Then

$$\mathbb{E}[X_1 | X_2 > t] = \int_0^\infty \frac{\psi_Q(x+t)}{\psi_Q(t)} dx. \tag{6.14}$$

For a random variable R following the inverse Gamma distribution with a shape parameter α and a scale parameter β, the density function $f_R(x) \in 2\mathrm{RV}_{-\alpha-1,-1}$ with an auxiliary function $A_{f_R}(x) = \beta x^{-1}$. We can derive from the density function and Proposition 6.5 that $\overline{F}_R(x) \in 2\mathrm{RV}_{-\alpha,-1}$ and the auxiliary function $A_{\overline{F}_R}(x) = \frac{\alpha\beta}{\alpha+1} x^{-1}$. If we let $Q \sim \mathrm{Gamma}(\alpha, 1)$; that is, α is the shape parameter and $\beta = 1$ is the rate parameter, since for the standard exponential random variable the a-th moment is $\Gamma(a+1)$, then we can get by Proposition 6.6 that

$$\mathbb{E}[X_1 | X_2 > t] \sim \frac{1}{\alpha - 1} t + \frac{1}{\alpha - 1}, \quad t \to \infty.$$

In another way, take $\psi_Q(x) = (1+x)^{-\alpha}$, the LT of $\mathrm{Gamma}(\alpha, 1)$ and from (6.14) it can be directly verified that $\mathbb{E}[X_1 | X_2 > t] = 1/(\alpha - 1) + t/(\alpha - 1)$, which is consistent with Proposition 6.6.

This example just shows that Proposition 6.6 can also be used for a random vector where margins possess the tail behavior of $\psi_Q(x)$ (i.e., $\mathbb{P}[X_i > x] = \psi_Q(x)$) and the dependence structure is the Archimedean copula constructed by $\psi_Q(x)$. This example is interesting since the tail behavior of the LT $\psi_Q(x)$ affects both the tail dependence structure and the tails of margins. $\overline{F}_R(x) \in \mathrm{RV}_{-\alpha}$ implies that $F_Q(x) \in \mathrm{RV}_\alpha(0)$. Then by the Karamata's Tauberian Theorem (e.g., Theorem

1.7.1' in Bingham et al. (1987)), $\psi_Q(x) \in \text{RV}_{-\alpha}$. That is, the margins are all regularly varying with index $-\alpha$. Moreover, there is upper tail dependence for (X_1, X_2), since $\psi_Q(t) \in \text{RV}_{-\alpha}$, the corresponding copula C_{ψ_Q} has lower tail dependence (refer to subsection 3.1 of Charpentier and Segers (2009)) and the copula for (X_1, X_2) is the survival copula of C_{ψ_Q}.

Based on (6.14), it seems that the second order property of $\psi_Q(x)$ could also be related to the second order parameter of $\overline{F}_R$. We conjecture that under some conditions, if $\overline{F}_R(t) \in 2\text{RV}_{-\alpha,\rho}$ and $Q := 1/R$, then $\psi_Q(t) \in 2\text{RV}_{-\alpha,\rho}$.

6.5 Concluding remarks

Second order regular variation provides a nice theoretical platform for studying second order approximations of limiting properties, like the asymptotic relationship between CTE and VaR that we have studied in this chapter. More importantly, many parametric distributions satisfy those theoretical assumptions so that the implementation of the main results is feasible. From the viewpoint of risk management, the study on risks at the sub-extremal level is more realistic and important, and we believe that the study involving the second order condition on asymptotic analysis for risk measures should be promising for this purpose.

For future research, it will be interesting to study how marginal tail heaviness and tail dependence structures affect the second order approximations of risk measures. Also, the second order regular variation we have studied on risk measures — no matter for the single risk or for the multiple risks — is univariate second order regular variation. To get a more general result for multiple risks, we need to study the performance of risk measures under multivariate second order regular variation. Moreover, whether the second order parameter ρ is greater than -1 is critical to decide on the necessity of a second order approximation. So it will be interesting to develop a methodology to test $H_0 : \rho \leq -1$ verses $H_1 : \rho > -1$.

6.6 Proofs

Proof of Lemma 6.2: Since $\overline{F} \in \text{RV}_{-\alpha}$, we can write $\overline{F}(t) = t^{-\alpha}\ell(t)$ where $\ell(t)$ is slowly varying and non-negative. By Definition 6.1, there is an auxiliary function

A and a constant $c \neq 0$ such that

$$\lim_{t\to\infty} \frac{x^{-\alpha}\ell(tx)/\ell(t) - x^{-\alpha}}{A(t)} = cx^{-\alpha}\frac{x^{\rho}-1}{\rho}.$$

Therefore,

$$\lim_{t\to\infty} \frac{\ell(tx)/\ell(t) - 1}{A(t)} = \lim_{t\to\infty} \frac{\ell(tx) - \ell(t)}{A(t)\ell(t)} = c\frac{x^{\rho}-1}{\rho}.$$

Since the sign of $A(t)$ can be adjusted by changing the sign of c due to Definition 6.1, if c and $A(t)$ have the same sign, then we assume that $c > 0$ and $A(t) > 0$, and thus $A(t)\ell(t) > 0$. By Theorem B.2.2 of de Haan and Ferreira (2006), then $\lim_{t\to\infty} \ell(t) = k$ with $0 < k < +\infty$ and $k - \ell(t) \in \mathrm{RV}_{\rho}$. So we can rewrite $\overline{F}(t) = kt^{-\alpha}\ell(t)$ with $\lim_{t\to\infty} \ell(t) = 1$ and also $1 - \ell(t) \in \mathrm{RV}_{\rho}$. If c and $A(t)$ have different signs, then let $A(t) > 0$, $-c > 0$ and then consider the function $-\ell(t)$ instead of $\ell(t)$. Also by Theorem B.2.2 of de Haan and Ferreira (2006), $\lim_{t\to\infty} -\ell(t) = -k$ with $0 < k < \infty$. So $\overline{F}(t) = kt^{-\alpha}\ell(t)$ with $\lim_{t\to\infty} \ell(t) = 1$ and also $\ell(t) - 1 \in \mathrm{RV}_{\rho}$. $\square$

Proof of Proposition 6.3: From Lemma 6.2, we can write $\overline{F}(t) = kt^{-\alpha}\ell(t)$ with $0 < k < +\infty$ and $\lim_{t\to\infty} \ell(t) = 1$. Then by Definition 6.1, with an auxiliary function $A(t)$ and a constant $c \neq 0$,

$$\lim_{t\to\infty} \frac{\overline{F}(tx)/\overline{F}(t) - x^{-\alpha}}{cA(t)} = \lim_{t\to\infty} \frac{x^{-\alpha}\ell(tx) - x^{-\alpha}\ell(t)}{cA(t)\ell(t)} = x^{-\alpha}\frac{x^{\rho}-1}{\rho}.$$

That is,

$$\lim_{t\to\infty} \frac{\ell(tx) - \ell(t)}{cA(t)\ell(t)} = \frac{x^{\rho}-1}{\rho}.$$

If c and $A(t)$ have the same sign, then $cA(t)\ell(t) > 0$ and by Theorem B.2.18 of de Haan and Ferreira (2006), we have for any $\epsilon, \delta > 0$, there exists a $t_0 = t_0(\epsilon, \delta)$ such that for all $t \wedge tx \geq t_0$,

$$\left| \frac{\ell(tx) - \ell(t)}{-\rho[1 - \ell(t)]} - \frac{x^{\rho}-1}{\rho} \right| \leq \epsilon x^{\rho} \max(x^{\delta}, x^{-\delta}). \tag{6.15}$$

A similar argument can be used for the case where c and $A(t)$ have different signs, and inequality (6.15) still holds. Multiplying both sides of (6.15) by $x^{-\alpha}$ finishes the proof. $\square$

Proof of Proposition 6.4: Denote that $t := \mathrm{VaR}_p(X)$, and the asymptotic relationship of (6.6) without the term $\eta(t)$ can be easily derived by applying the Karamata's theorem. Write $\overline{F}(t) := kt^{-\alpha}\ell(t)$ and choose $a(t) := -\rho[1-\ell(t)]/\ell(t)$. Then

$$\begin{aligned}
&\mathrm{CTE}_p(X)\\
&= t + \frac{\int_t^\infty \mathbb{P}[X > x]dx}{\mathbb{P}[X > t]} = t\left(1 + \frac{\int_t^\infty \mathbb{P}[X > x]dx}{t\mathbb{P}[X > t]}\right) = t\left(1 + \int_t^\infty \frac{\mathbb{P}[X > x]}{t\mathbb{P}[X > t]}dx\right)\\
&= t\left(1 + \int_1^\infty \frac{\mathbb{P}[X > ts]}{\mathbb{P}[X > t]}ds\right) = t\left(1 + a(t)\int_1^\infty \frac{\frac{\mathbb{P}[X>ts]}{\mathbb{P}[X>t]} - s^{-\alpha}}{a(t)}ds + \frac{1}{\alpha-1}\right)\\
&= \frac{\alpha}{\alpha-1}t + \frac{ta(t)}{(\alpha-1-\rho)(\alpha-1)} + ta(t)\int_1^\infty \frac{\frac{\mathbb{P}[X>ts]}{\mathbb{P}[X>t]} - s^{-\alpha}}{a(t)} - s^{-\alpha}\frac{s^\rho - 1}{\rho}ds,
\end{aligned} \tag{6.16}$$

where the last equality is due to $\int_1^{+\infty} s^{-\alpha}\frac{s^\rho-1}{\rho}ds = \frac{1}{(\alpha-1-\rho)(\alpha-1)}$. By Proposition 6.3, when $s \geq 1$, for any small ϵ, δ such that $\epsilon > 0$ and $\alpha - \rho - 1 > \delta > 0$

$$\left|\frac{\overline{F}(ts)/\overline{F}(t) - s^{-\alpha}}{a(t)} - s^{-\alpha}\frac{s^\rho - 1}{\rho}\right| \leq \epsilon \max\left(s^{-\alpha+\rho+\delta}, s^{-\alpha+\rho-\delta}\right) = \epsilon s^{-\alpha+\rho+\delta}.$$

Since ϵ is arbitrarily small and $s^{-\alpha+\rho+\delta}$ is integrable on $[1,\infty)$, by the dominated convergence theorem and Definition 6.1, the integration term in (6.16) is 0 as $t \to \infty$, which proves the claim. $\square$

Proof of Proposition 6.5: Since $g \in 2\mathrm{RV}_{-\alpha,\rho}$, letting $x = st$,

$$\begin{aligned}
&\lim_{t\to\infty} \frac{\int_t^\infty \frac{g(x)}{tg(t)}dx - \frac{1}{\alpha-1}}{A(t)} = \lim_{t\to\infty}\int_t^\infty \frac{\frac{g(x)}{tg(t)} - \frac{x^{-\alpha}}{t^{-\alpha+1}}}{A(t)}dx = \lim_{t\to\infty}\int_1^\infty \frac{\frac{g(st)}{g(t)} - s^{-\alpha}}{A(t)}ds\\
&= \int_1^\infty \lim_{t\to\infty}\frac{\frac{g(st)}{g(t)} - s^{-\alpha}}{A(t)}ds = \frac{1}{\rho}\left(\frac{1}{\alpha-\rho-1} - \frac{1}{\alpha-1}\right).
\end{aligned} \tag{6.17}$$

The third equality in (6.17) is due to Remark 6.3. Therefore,

$$\int_t^\infty \frac{g(x)}{tg(t)}dx \sim \frac{1}{\alpha-1} + \frac{1}{\rho}\left(\frac{1}{\alpha-\rho-1} - \frac{1}{\alpha-1}\right)A(t), \quad t \to \infty,$$

which proves (6.9). Then by (6.9) it is easy to verify that $g^* \in 2\mathrm{RV}_{-\alpha+1,\rho}$. Let $r := (\alpha-\rho-1)^{-1} - (\alpha-1)^{-1}$, then as $t \to \infty$,

$$\begin{aligned}
&\frac{g^*(ty)}{g^*(t)} - y^{-\alpha+1} \\
&\sim \frac{(\alpha-1)^{-1}yg(ty) + \rho^{-1}rA(ty)yg(ty)}{(\alpha-1)^{-1}g(t)+\rho^{-1}rA(t)g(t)} - y^{-\alpha+1} \\
&\sim \left\{\frac{1}{\alpha-1}\frac{y^{-\alpha+1}(y^\rho-1)}{\rho}A(t)\right. \\
&\quad \left. + \frac{r}{\rho}A(ty)\left(\frac{y^{-\alpha+1}(y^\rho-1)}{\rho}A(t) + y^{-\alpha+1}\right) - \frac{r}{\rho}A(t)y^{-\alpha+1}\right\}(\alpha-1) \\
&\sim \left\{\frac{1}{\alpha-1}\frac{y^{-\alpha+1}(y^\rho-1)}{\rho} + \frac{r}{\rho}\frac{A(ty)}{A(t)}y^{-a+1} - \frac{r}{\rho}y^{-\alpha+1}\right\}A(t)(\alpha-1) \\
&\sim \left(\frac{1}{\alpha-1}+r\right)\frac{y^{-\alpha+1}(y^\rho-1)}{\rho}A(t)(\alpha-1) = \frac{\alpha-1}{\alpha-\rho-1}A(t)\frac{y^{-\alpha+1}(y^\rho-1)}{\rho}.
\end{aligned}$$

Thus, $A^*(t) = \frac{\alpha-1}{\alpha-\rho-1}A(t)$. ☐

Proof of Proposition 6.6: Denoting $t = \mathrm{VaR}_p(X_2)$ and following the proof of Proposition 2.3 of Zhu and Li (2012), we have

$$\begin{aligned}
&\mathbb{E}[X_1|X_2>t] \\
&= \frac{t}{\mathbb{P}[X_2>t]}\int_{\mathbb{R}_+^2}\left\{\frac{t_1}{t_2}\mathbb{P}\left[R>\frac{t}{t_2}\right] + \int_{\frac{t_1}{t_2}}^\infty \mathbb{P}\left[R>\frac{tw}{t_1}\right]dw\right\}F_{T_1,T_2}(dt_1,dt_2).
\end{aligned} \tag{6.18}$$

Then it suffices to study the second order approximation of the following two asymptotic relationships.

$$D_1(t) := \int_{\mathbb{R}_+^2}\frac{t_1}{t_2}\frac{\mathbb{P}\left[R>\frac{t}{t_2}\right]}{\mathbb{P}[X_2>t]}F_{T_1,T_2}(dt_1,dt_2) \to \frac{\mathbb{E}[T_1T_2^{\alpha-1}]}{\mathbb{E}[T_2^\alpha]}, \quad t\to\infty, \tag{6.19}$$

and

$$D_2(t) := \int_{\mathbb{R}_+^2} \frac{\int_{\frac{t_1}{t_2}}^{\infty} \mathbb{P}\left[R > \frac{tw}{t_1}\right] dw}{\mathbb{P}[X_2 > t]} F_{T_1,T_2}(dt_1, td_2) \to \frac{1}{\alpha - 1} \frac{\mathbb{E}[T_1 T_2^{\alpha-1}]}{\mathbb{E}[T_2^{\alpha}]}, t \to \infty, \tag{6.20}$$

where both t_1, t_2 are positive.

The proof will include three parts: 1. the second order property of the Breiman's convergence: $\mathbb{P}[X_2 > t] \sim \mathbb{E}[T_2^{\alpha}]\mathbb{P}[R > t], t \to \infty$; 2. the second order property for (6.19); 3. the second order property for (6.20).

Part 1: The original proof of Breiman's theorem is referred to Breiman (1965). Here we need to derive its speed of convergence. Since T_2 and R are independent,

$$\begin{aligned}
\frac{\mathbb{P}[X_2 > t]}{\mathbb{P}[R > t]} = \frac{\mathbb{P}[T_2 R > t]}{\mathbb{P}[R > t]} &= \int_0^{\infty} \frac{\overline{F}_R\,(t/x)}{\overline{F}_R(t)} F_{T_2}(dx) \\
&= A(t) \int_0^{\infty} \frac{\frac{\overline{F}_R(t/x)}{\overline{F}_R(t)} - x^{\alpha}}{A(t)} F_{T_2}(dx) + \mathbb{E}[T_2^{\alpha}].
\end{aligned}$$

Now we claim that if $\mathbb{E}[T_2^{\alpha-\rho+\delta}] < \infty$ for some $\delta > 0$, then

$$\lim_{t\to\infty} \int_0^{\infty} \frac{\frac{\overline{F}_R(t/x)}{\overline{F}_R(t)} - x^{\alpha}}{A(t)} F_{T_2}(dx) = \int_0^{\infty} x^{\alpha} \frac{x^{-\rho} - 1}{\rho} F_{T_2}(dx). \tag{6.21}$$

Let

$$h(t,x) := \frac{\frac{\overline{F}_R(t/x)}{\overline{F}_R(t)} - x^{\alpha}}{A(t)}; \quad g(x) := x^{\alpha} \frac{x^{-\rho} - 1}{\rho}.$$

By Proposition 6.3, we can write $\overline{F}_R(t) := kt^{-\alpha}\ell(t)$ with $0 < k < \infty$ and $\lim_{t\to\infty} \ell(t) = 1$. Also, let $A(t) = -\rho[1 - \ell(t)]/\ell(t)$ as the one in Proposition 6.3 and write $|A(t)| := t^{\rho}\ell_A(t)$, where $\ell_A(t)$ is a slowly varying function, since $|A(t)| \in \mathrm{RV}_{\rho}$. Furthermore, for $\delta > 0$ such that $\mathbb{E}[T^{\alpha-\rho+\delta}] < \infty$ and any $\epsilon > 0$, there exists t_0 such that for all $t, t/x \geq t_0$ and $x > 0$,

$$|h(t,x) - g(x)| \leq \epsilon \max(x^{\alpha-\rho+\delta}, x^{\alpha-\rho-\delta}).$$

Hence, assuming $\delta < \alpha - \rho$, without loss of generality,

$$\int_0^{t/t_0} |h(t,x) - g(x)|\, F_{T_2}(dx) \leq \epsilon \int_0^{t/t_0} \max(x^{\alpha-\rho+\delta}, x^{\alpha-\rho-\delta}) F_{T_2}(dx) \leq \epsilon K_1, \tag{6.22}$$

where $0 < K_1 < \infty$ and does not depend on t. When $x \in (t/t_0, \infty)$, we have $xt_0 > t$ and $|A(t)|$ is ultimately decreasing implies that when t is sufficiently large $|A(t)| > |A(xt_0)| = x^\rho t_0^\rho |\ell_A(xt_0)|$; $\overline{F}_R(t) = kt^{-\alpha}\ell(t)$ and $\overline{F}_R(t/x) \leq 1$ imply that $\ell(t/x) \leq k^{-1}(t/x)^\alpha < k^{-1}t_0^\alpha$. Together with $\lim_{t\to\infty} \ell(t) = 1$, we have when t is sufficiently large and $x \in (t/t_0, \infty)$,

$$|h(t,x)| < \frac{x^\alpha |2k^{-1}t_0^\alpha - 1|}{x^\rho t_0^\rho |\ell_A(xt_0)|} < x^{\alpha-\rho+\delta} K_2, \quad (0 < K_2 < \infty)$$

and thus by the dominated convergence theorem, when t is sufficiently large

$$\int_{t/t_0}^\infty |h(t,x)| F_{T_2}(dx) < \epsilon. \tag{6.23}$$

Also, apparently, for t sufficiently large,

$$\int_{t/t_0}^\infty |g(x)| F_{T_2}(dx) < \epsilon. \tag{6.24}$$

Combining (6.22), (6.23) and (6.24) implies that, for any $\epsilon > 0$, when t is sufficiently large,

$$\begin{aligned}
&\left| \int_0^\infty h(t,x) F_{T_2}(dx) - \int_0^\infty g(x) F_{T_2}(dx) \right| \leq \int_0^\infty |h(t,x) - g(x)|\, F_{T_2}(dx) \\
&\leq \int_0^{t/t_0} |h(t,x) - g(x)|\, F_{T_2}(dx) + \int_{t/t_0}^\infty |h(t,x)| F_{T_2}(dx) + \int_{t/t_0}^\infty |g(x)| F_{T_2}(dx) \\
&< (K_1 + 2)\epsilon,
\end{aligned}$$

which proves the claim.

Hence, as $t \to +\infty$,

$$\mathbb{P}[X_2 > t] \sim \mathbb{E}[T_2^{\alpha}]\mathbb{P}[R > t] + \frac{A(t)}{\rho}\left(\mathbb{E}[T_2^{\alpha-\rho}] - \mathbb{E}[T_2^{\alpha}]\right)\mathbb{P}[R > t]. \tag{6.25}$$

Part 2: For (6.19), take $\overline{F}_{R*}(t) := t\overline{F}_R(t)$, then $\overline{F}_{R*}(t) \in 2\mathrm{RV}_{-\alpha+1,\rho}$ and the corresponding auxiliary function is $A_*(t) = A(t)$. Let

$$M_2(B) := \int_B \int_{\mathbb{R}_+} t_1 F_{T_1,T_2}(dt_1, dt_2), B \subset \mathbb{R}_+$$

denote the marginal mean measure induced by T_1. Since $M_2(\mathbb{R}_+) = \mathbb{E}[T_1] < \infty$, $M_2(\cdot)$ is a finite measure. Since $\mathbb{E}[T_1 T_2^{\alpha-1-\rho+\xi}] < \infty$, then by (6.21) in the proof of Part 1, applied to $\overline{F}_{R*}$, we have

$$\begin{aligned}
&\lim_{t\to\infty}\int_{\mathbb{R}_+^2}\frac{t_1}{t_2}\frac{\overline{F}_R(t/t_2)}{\overline{F}_R(t)}F_{T_1,T_2}(dt_1,dt_2) = \lim_{t\to\infty}\int_{\mathbb{R}_+}\frac{\overline{F}_{R*}(t/t_2)}{\overline{F}_{R*}(t)}M_2(dt_2)\\
&= \mathbb{E}[T_1T_2^{\alpha-1}] + \frac{A(t)}{\rho}\left(\mathbb{E}[T_1T_2^{\alpha-1-\rho}] - \mathbb{E}[T_1T_2^{\alpha-1}]\right). \qquad (6.26)
\end{aligned}$$

Also,

$$\begin{aligned}
\lim_{t\to\infty} D_1(t) &= \lim_{t\to\infty}\int_{\mathbb{R}_+^2}\frac{t_1}{t_2}\frac{\mathbb{P}\left[R > \frac{t}{t_2}\right]}{\mathbb{P}[X_2 > t]}F_{T_1,T_2}(dt_1,dt_2)\\
&= \lim_{t\to\infty}\frac{\mathbb{P}\left[R > t\right]}{\mathbb{P}[X_2 > t]} \times \lim_{t\to\infty}\int_{\mathbb{R}_+^2}\frac{t_1}{t_2}\frac{\overline{F}_R(t/t_2)}{\overline{F}_R(t)}F_{T_1,T_2}(dt_1,dt_2) \qquad (6.27)
\end{aligned}$$

Part 3: For (6.20), let $\overline{F}_R^*(t) := \int_t^\infty \overline{F}_R(x)dx$, since $\overline{F}_R(x) \in 2\mathrm{RV}_{-\alpha,\rho}$, then by Proposition 6.5, $\overline{F}_R^*(t) \in 2\mathrm{RV}_{-\alpha+1,\rho}$ and the corresponding auxiliary function is

$A^*(t) = \frac{\alpha-1}{\alpha-1-\rho}A(t)$. Also, letting $x = tw/t_1$,

$$
\begin{aligned}
D_2(t) &= \frac{\mathbb{P}[R > t]}{\mathbb{P}[X_2 > t]} \int_{\mathbb{R}_+^2} \frac{\int_{\frac{t_1}{t_2}}^{\infty} \overline{F}_R(tw/t_1)dw}{\overline{F}_R(t)} F_{T_1,T_2}(dt_1, dt_2) \\
&= \frac{\mathbb{P}[R > t]}{\mathbb{P}[X_2 > t]} \int_{\mathbb{R}_+^2} t_1 \left\{ \int_{\frac{t}{t_2}}^{\infty} \frac{\overline{F}_R(x)}{t\overline{F}_R(t)} dx \right\} F_{T_1,T_2}(dt_1, dt_2) \\
&= \frac{\mathbb{P}[R > t]}{\mathbb{P}[X_2 > t]} \int_{\mathbb{R}_+^2} t_1 \frac{\overline{F}_R^*(t/t_2)}{\overline{F}_R^*(t)} \frac{\int_t^{\infty} \overline{F}_R(x)dx}{t\overline{F}_R(t)} F_{T_1,T_2}(dt_1, dt_2).
\end{aligned}
$$

Therefore,

$$
\begin{aligned}
\lim_{t\to\infty} D_2(t) = \lim_{t\to\infty} \frac{\mathbb{P}[R > t]}{\mathbb{P}[X_2 > t]} &\times \lim_{t\to\infty} \frac{\int_t^{\infty} \overline{F}_R(x)dx}{t\overline{F}_R(t)} \\
&\times \lim_{t\to\infty} \int_{\mathbb{R}_+^2} t_1 \frac{\overline{F}_R^*(t/t_2)}{\overline{F}_R^*(t)} F_{T_1,T_2}(dt_1, dt_2). \quad (6.28)
\end{aligned}
$$

From (6.21), applied to $\overline{F}_R^*$, because $\mathbb{E}[T_1 T_2^{\alpha-1-\rho+\xi}] < \infty$, as $t \to \infty$,

$$
\begin{aligned}
&\int_{\mathbb{R}_+^2} t_1 \frac{\overline{F}_R^*(t/t_2)}{\overline{F}_R^*(t)} F_{T_1,T_2}(dt_1, dt_2) \\
&= \int_0^{\infty} \frac{\overline{F}_R^*(t/t_2)}{\overline{F}_R^*(t)} M_2(dt_2) \\
&\sim \mathbb{E}[T_1 T_2^{\alpha-1}] + \frac{A^*(t)}{\rho} \left(\mathbb{E}[T_1 T_2^{\alpha-1-\rho}] - \mathbb{E}[T_1 T_2^{\alpha-1}] \right) \\
&= \mathbb{E}[T_1 T_2^{\alpha-1}] + \frac{A(t)}{\rho} \frac{\alpha-1}{\alpha-1-\rho} \left(\mathbb{E}[T_1 T_2^{\alpha-1-\rho}] - \mathbb{E}[T_1 T_2^{\alpha-1}] \right) \qquad (6.29)
\end{aligned}
$$

Hence, combining (6.25), (6.26), (6.27), (6.9), (6.28) and (6.29) together, we will get

$$
\begin{aligned}
\mathbb{E}[X_1 | X_2 > t] \sim \frac{\alpha}{\alpha-1} &\frac{\mathbb{E}[T_1 T_2^{\alpha-1}]}{\mathbb{E}[T_2^{\alpha}]} t \\
&+ \frac{A(t)t}{\rho} \left\{ \frac{\alpha-\rho}{\alpha-1-\rho} \frac{\mathbb{E}[T_1 T_2^{\alpha-1-\rho}]}{\mathbb{E}[T_2^{\alpha}]} - \frac{\alpha}{\alpha-1} \frac{\mathbb{E}[T_1 T_2^{\alpha-1}]\mathbb{E}[T_2^{\alpha-\rho}]}{(\mathbb{E}[T_2^{\alpha}])^2} \right\}.
\end{aligned}
$$

This completes the proof . $\square$

Proof of Proposition 6.7: In Proposition 6.6, replace T_2 by $h(\mathbf{T})$ since $h(\mathbf{X}) = h(\mathbf{T})R$. $\square$

Chapter 7

Conclusions and future research

7.1 Summary

The topics for the thesis have shown their importance in quantitative modeling in insurance and finance where the assumption of normality would fail to explain the rare events and their nonlinear dependence structures.

Two themes of the thesis are strength of dependence in the tails, and influence of tail patterns on risk measures.

7.1.1 Strength of dependence in the tails

Tail order and intermediate tail dependence

For multivariate insurance and financial data, the degree of positive association often appears to be higher on the tail than in the middle. To this end, we have intensively investigated various tail dependence patterns and made contributions to the development of relevant theories, including characterizing an intermediate tail dependence structure.

Intermediate tail dependence can be explained as a continuum of strength of dependence in the tail between the tail quadrant independence (the tail behaves like the product of the univariate margins) and the usual tail dependence. We introduce the concepts of tail order and tail order functions for studying copula families that have various tail patterns, and find a new pattern of intermediate tail dependence for upper tails of Archimedean copulas and conditions that lead to such a pattern; this fills in gaps in the literature. In addition, we have developed theories that relate tail heaviness of a positive random variable to the tail behavior of the Archimedean copula constructed with the LT of the random variable, which is an insightful way to understand the tail behavior of Archimedean copulas.

7.1.2 Tail patterns and risk measures

Study of tail patterns such as tail heaviness and tail dependence plays a critical role in risk assessment. To this end, we have conducted research in the following two aspects.

Conservative tail dependence structure

The influence of dependence in the tails on risk measures for multivariate insurance or financial data has attracted more and more attention of researchers. However, it is very difficult to give a good estimation of the tail dependence as the data in the tail is sparse. To deal with this situation, we propose a data-driven conservative dependence structure to account for the uncertainty in tail dependence.

The conservative dependence structure is referred to as *tail comonotonicity*. It is assumed that the strongest dependence appears when losses go to infinity, while the dependence on the sub-extremal level is characterized by the second-order parameters that can be estimated using data. The data-driven merit provides a practical way to obtain a better guess on the worst dependence structure, while the strongest positive dependence on the tail guarantees that the less information we have and the larger the losses are, the more conservative assumptions are adopted. We have conducted a thorough investigation on the properties and constructions of this dependence structure, and found interesting asymptotic additivity properties of relevant risk measures. Sufficient conditions have been obtained to justify the conservativity of tail comonotonicity. In addition to theoretical studies, real data analysis and intensive simulations suggest that, by assuming tail comonotonicity, one does not lose much accuracy but gain reasonably conservative risk assessment, especially for high-risk scenarios.

Sub-extremal risks

Sub-extremal risks can be referred to as the risks that are higher than a certain high threshold (e.g., 90% quantile). For modeling sub-extremal losses, an important question is to understand the probabilistic structures in the sub-extremal level. To answer this question, we have studied a general second-order tail heaviness structure and established second-order approximations to conditional tail expectations (CTE) for both univariate and multivariate dependent risks.

We use 2RV conditions to study the properties of risk measures on the sub-extremal level. Under some mild conditions that can be easily satisfied, closed-form second-order approximations for both the univariate and multivariate cases of CTE have been derived. The results provide a clear and general probabilistic structure for characterizing the distributional tails, and indicate conditions under which a second order parameter becomes more important.

7.2 Future research

A couple of research questions that are relevant to the theory contained in this thesis are presented in what follows.

7.2.1 Tail behavior of CTEs

In Chapter 5, $\mathbb{E}[X_1|X_2 > t]$ and $\mathbb{E}[X_1|X_2 = t]$ have been studied in terms of asymptotic comparisons under the condition of different tail dependence patterns. That is, different tail properties of copulas affect the CTEs, and the asymptotic behavior of $E(X_1|X_2 > t)$ and $E(X_1|X_2 = t)$ would be another way to show how copula families can be differentiated, and to show that the choice of copula family has a big influence on tail inference.

We first use plots to illustrate comparisons between these three cases of usual tail dependence ($\kappa = 1$), intermediate tail dependence ($1 < \kappa < 2$) and tail quadrant independence ($\kappa = 2$) by Figure 7.1. The Blomqvist's β for the three columns of the plots in Figure 7.1 were chosen to be $\beta = 0.3, 0.6, 0.9$, respectively. In each plot, the parameters for Gumbel, s.Gumbel and MTCJ copulas were calculated based on the same Blomqvist's β. For the Gumbel copula, $\delta = 1.434, 2.484, 9.709$, respectively for each column, and so does the δ for the s.Gumbel copula. For the MTCJ copula, $\delta = 0.863, 2.764, 13.513$, respectively for each column. The parameter for the identical Pareto margins were chosen to be the same in each row; that is, $\theta = 2, 100, 1.1$, respectively.

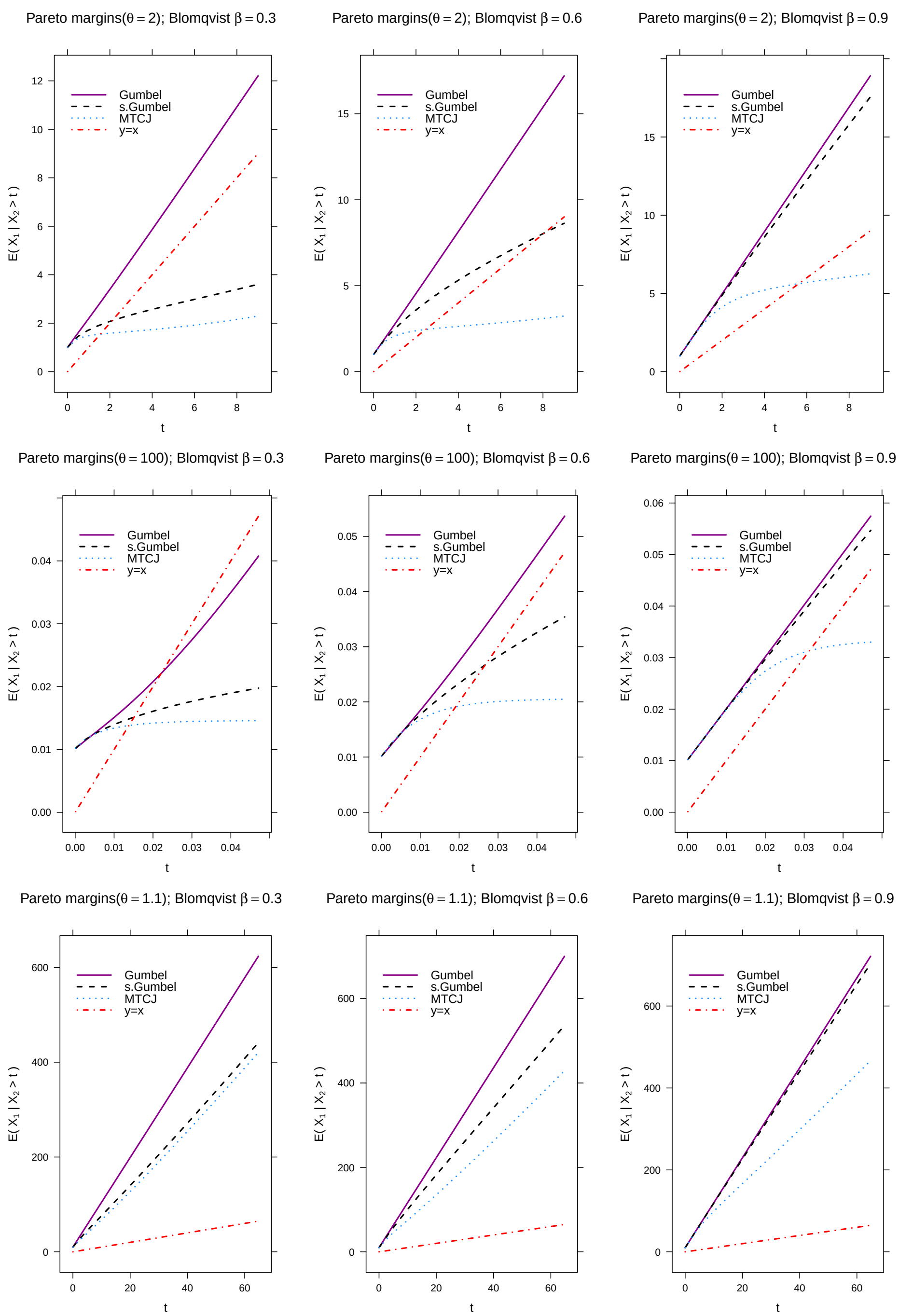
Based on Figure 7.1, we find that when the upper tail order is 1 (Gumbel), the CTE plots seem to be linear in t on the whole support no matter what the θ is. When the upper tail order is 2 (MTCJ), the CTE lines become very flat. The CTE plots for the intermediate upper tail dependence with the upper tail order $1 < \kappa < 2$ (s.Gumbel) are located between the above two cases. However, if the Pareto margins are too heavy-tailed with a very small parameter θ (e.g., $\theta = 1.1$ in the last row of Figure 7.1), the CTE plots are not suitable for discriminating the three cases of tail orders. So, it may suggest to choose a relatively large θ for the purpose of discriminating the types of tail orders. Moreover, when Blomqvist's β is very large (e.g., $\beta = 0.9$), it becomes very hard to distinguish tail dependence and intermediate tail dependence; similarly, when Blomqvist's β is very small, the plot may becomes difficult to differentiate intermediate tail dependence and tail quadrant independence.

In what follows, we will try to keep the arguments suitable to general cases when discussing the three types of tail behavior, otherwise, concrete examples will be given to show the possible behavior if a general discussion is not possible at the moment. The research goal will be to obtain checkable sufficient conditions to confirm whether CTE behavior is $t^1\ell(t)$, $t^\alpha\ell(t)$ with $0 < \alpha < 1$, or $t^0\ell(t)$, as $t \to \infty$.

Usual tail dependence

The linearity in t for the usual tail dependence case can be explained as what follows. If X_1 and X_2 have the same regularly varying upper tail and their copula possess an upper tail dependence function, then (X_1, X_2) is multivariate regularly varying (see Weng and Zhang, 2012); let the corresponding intensity measure be ν. Choosing $B := [0, \infty] \times [1, \infty]$ in Theorem 2.2 (1) of Joe and Li (2011), we

Figure 7.1: CTE plots for $\mathbb{E}[X_1|X_2 > t]$. The range of t in each plot was chosen to cover the support of t between the 1% and 99% quantiles.



have

$$\mathbb{E}[X_1|X_2 > t] \sim t \int_0^\infty \frac{\nu([w,\infty] \times [1,\infty])}{\nu B} dw, \quad t \to \infty. \tag{7.1}$$

So the CTE plot should be approximately linear in t when t is sufficiently large, as long as the integration in (7.1) is finite. Here, an open question is that, for the Gumbel copula, why the linearity always appears even when t is very small.

Intermediate tail dependence

The relevant theory for the other two cases is more tricky. Now, let us look at the example of Pareto margins with a survival Gumbel copula more closely, for the intermediate tail dependence case.

Example 7.1 (Pareto margins and survival Gumbel copula) Suppose X_1 and X_2 have a copula C that is the survival copula of a Gumbel copula $\widehat{C}$, so that the upper tail order of C is $1 < \kappa = 2^{1/\delta} < 2$. Then we can write

$$\begin{aligned}
&\mathbb{E}[X_1|X_2 > t] \\
&= [\overline{F}(t)]^{-1} \int_0^\infty \widehat{C}(\overline{F}(x), \overline{F}(t)) dx \\
&= [\overline{F}(t)]^{-1} \int_0^\infty \exp\left\{ -\left([-\log \overline{F}(x)]^\delta + [-\log \overline{F}(t)]^\delta \right)^{1/\delta} \right\} dx \\
&= (1+t)^\theta \int_0^\infty \exp\left\{ -\theta \left([\log(1+x)]^\delta + [\log(1+t)]^\delta \right)^{1/\delta} \right\} dx \\
&= \int_0^\infty \exp\left\{ s\theta - \theta \left(y^\delta + s^\delta \right)^{1/\delta} + y \right\} dy, \quad y := \log(1+x); s := \log(1+t) \\
&= s \int_0^\infty \exp\left\{ s \left[\theta - \theta \left(z^\delta + 1 \right)^{1/\delta} + z \right] \right\} dz, \quad z := y/s \\
&=: s \int_0^\infty e^{sg(z)} dz,
\end{aligned}$$

where $g(z) := \theta - \theta \left(z^\delta + 1\right)^{1/\delta} + z$. As $t \to \infty$, $s \to \infty$, so we may use a Laplace approximation to evaluate the behavior of $\int_0^\infty e^{sg(z)} dz$ as $s \to \infty$. In order to apply the Laplace approximation, the function $g(z)$ has to satisfy some regularity conditions (see Small, 2010) as follows.

(1) Clearly, $g(z)$ is differentiable on $(0, \infty)$, and the global maximum is attained when $z = z_0 = [\theta^{\delta/(\delta-1)} - 1]^{-1/\delta}$. As $\theta, \delta > 1$, $0 < z_0 < \infty$. While $g''(z) = (1-\delta)\theta z^{\delta-2}(1+z^\delta)^{1/\delta-2} < 0$ for $z \in (0, \infty)$ when $\theta, \delta > 1$.

(2) Since $g(0) = 0$, $g(\infty) = -\infty$, $g(z)$ is strictly increasing for $z \in (0, z_0]$ and $g(z)$ is strictly decreasing for $z \in [z_0, \infty)$, we may choose a $0 < \xi < z_0$ and $\epsilon = g(z_0) - \max\{g(z_0 - \xi), g(z_0 + \xi)\} > 0$ such that $g(z) < g(z_0) - \epsilon$ for all $z \in (0, \infty) \cap \{z : |z - z_0| \geq \xi\}$.

(3) Also, $\int_0^\infty \exp\{\theta - \theta\left(z^\delta + 1\right)^{1/\delta} + z\}dz < e^\theta \int_0^\infty e^{z(1-\theta)}dz < \infty$.

Therefore, as $t \to \infty$,

$$\mathbb{E}[X_1|X_2 > t] \sim s e^{sg(z_0)} \sqrt{\frac{2\pi}{-sg''(z_0)}} = (1+t)^{g(z_0)} \sqrt{\frac{2\pi \log(1+t)}{-g''(z_0)}},$$

and $g(z_0) = \theta - \left(\theta^{\delta/(\delta-1)} - 1\right)^{(\delta-1)/\delta} > 0$ affects the behavior of $\mathbb{E}[X_1|X_2 > t]$. Note that, the Laplace approximation involves a Taylor expansion to the second order of $sg(z)$ and reweighing of the Gaussian integral for a Gaussian random variable with the variance being $-(sg''(z_0))^{-1}$.

Let $\xi(\theta, \delta) := g(z_0)$, then $\partial\xi(\theta, \delta)/\partial\theta = 1 - \left(1 - \theta^{-\delta/(\delta-1)}\right)^{-1/\delta} < 0$ and thus $\xi(\theta, \delta)$ is deceasing in θ. So increasing θ, the parameter for the Pareto margins, will decrease the speed of $\mathbb{E}[X_1|X_2 > t]$ as $t \to \infty$. Similarly, it can be easily verified that, for $\theta, \delta > 1$, $\partial\xi(\theta, \delta)/\partial\delta > 0$. Therefore, increasing δ will increase the speed of $\mathbb{E}[X_1|X_2 > t]$ as $t \to \infty$. Those are consistent to the plots in Figure 7.1. Note that, the speed of $\mathbb{E}[X_1|X_2 > t]$ can be roughly compared by comparing the relative positions of the CTE lines to the "$y = x$" line.

Tail quadrant independence

From Figure 7.1, we can find that the CTE line for the tail quadrant independence case (MTCJ) tends to converge to a finite number, if the parameter θ for the Pareto margins is sufficiently large.

Now, suppose continuous random variables $X_1 \overset{d}{=} X_2 \sim F$ supported on $[0, \infty)$ with the copula C. Let $g(v) := C(u, v) - v$ for a given $0 \leq u \leq 1$.

Then a Taylor expansion at $v = 1$ leads to, as $v \to 1^-$,

$$g(v) = (u-1) + (1-v)[1 - C_{1|2}(u|1)] + 2^{-1}\mathsf{D}_{22}C(u,1)(v-1)^2 + R(u,v),$$

where $\mathsf{D}_{22}C(u,1) = \partial^2 C(u,v)/\partial^2 v|_{v=1}$ and the remainder $R(u,v) = o((v-1)^2)$. If there exists a $v_0 < 1$ such that $v > v_0$ implies that

$$2^{-1}\mathsf{D}_{22}C(u,1)(v-1)^2 + R(u,v) \le 0, \quad \text{for } u \text{ a.e.}; \tag{7.2}$$

that is, $g(v) \le h(v) := (u-1) + (1-v)[1 - C_{1|2}(u|1)]$, and let $u = F(x)$, then $C(F(x), v) - v \le (F(x) - 1) + (1-v)[1 - C_{1|2}(F(x)|1)]$. Therefore as $v > v_0$,

$$\begin{aligned}
&\mathbb{E}[X_1|X_2 > t] \\
&= (1-v)^{-1}\left\{\mathbb{E}[X_1] + \int_0^\infty [C(F(x),v) - v]dx\right\} \\
&\le (1-v)^{-1}\left\{\mathbb{E}[X_1] + \int_0^\infty (F(x)-1) + (1-v)[1 - C_{1|2}(F(x)|1)]dx\right\} \\
&= \int_0^\infty [1 - C_{1|2}(F(x)|1)]dx.
\end{aligned}$$

If

$$\begin{aligned}
&G(x) := \lim_{v\to 1^-} C_{1|2}(F(x)|v) \text{ exists} \\
&G(x) \text{ is the cdf of } Y, \text{ and } \mathbb{E}[Y] < \infty.
\end{aligned} \tag{7.3}$$

Then by the dominated convergence theorem,

$$\begin{aligned}
&\lim_{t\to\infty} \mathbb{E}[X_1|X_2 > t] = \lim_{v\to 1^-} \frac{\int_0^\infty \{C(F(x),v) - v + 1 - F(x)\}dx}{1-v} \\
&= \int_0^\infty \lim_{v\to 1^-} \frac{C(F(x),v) - v + 1 - F(x)}{1-v} dx \\
&= \int_0^\infty \left\{1 - C_{1|2}(F(x)|1)\right\} dx.
\end{aligned}$$

That is,

$$\mathbb{E}[X_1|X_2 > t] \to \mathbb{E}[Y] < \infty, \quad t \to \infty.$$

Condition (7.3) is easy to verify. From Table 1.2, tail quadrant independent copulas such as Frank, MTCJ and BB2 all satisfy the condition that $G(x) := C_{1|2}(F(x)|1)$ is the cdf of a random variable Y. By appropriately choosing $F(x)$, we can have $\mathbb{E}[Y] < \infty$. However, the condition (7.2) or related conditions require further investigation, and a similar final conclusion could be obtained with sufficient conditions other than (7.2).

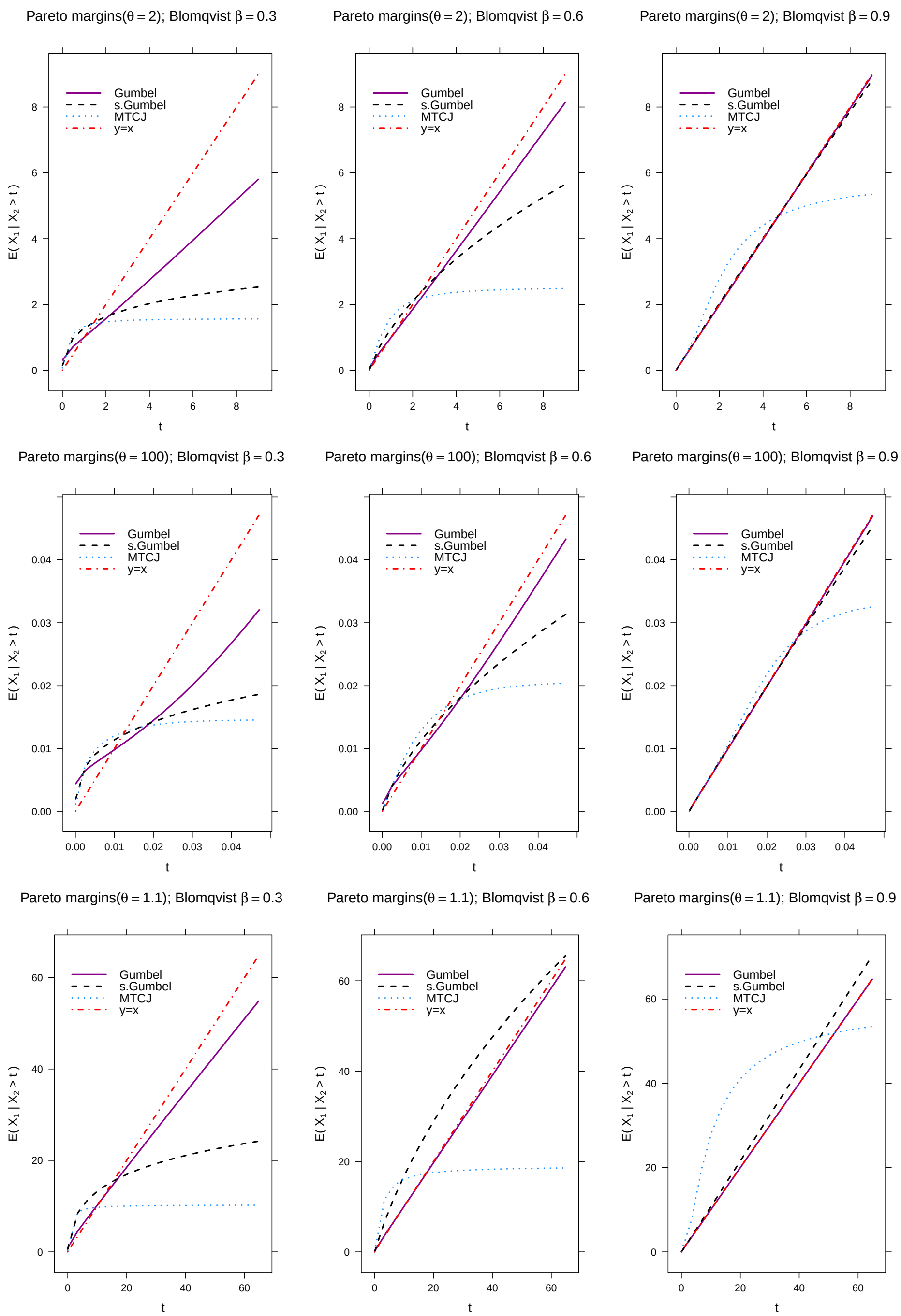
Similar patterns can be observed in Figure 7.2 for the corresponding CTE plots of $\mathbb{E}[X_1|X_2 = t]$.

Unlike the former case, $\mathbb{E}[X_1|X_2 = t]$ is relatively harder to get the empirical quantities. So CTE plots for $\mathbb{E}[X_1|X_2 > t]$ have a better potential to be further developed as a diagnostic tool for the patterns of dependence in the tails.

We now close this subsection by giving an example for the intermediate tail dependence case.

Example 7.2 (Pareto margins and survival Gumbel copula) Suppose X_1 and X_2 have a copula C that is the survival copula of a Gumbel copula $\widehat{C}$, so that the upper tail order of C is $1 < \kappa = 2^{1/\delta} < 2$. Letting $y := \log(1 + x)$, $s := \log(1 + t)$ and

Figure 7.2: CTE plots for $\mathbb{E}[X_1|X_2 = t]$. The range of t in each plot was chosen to cover the support of t between the 1% and 99% quantiles.



$z := y/s$, we can write

$$\begin{aligned}
&\mathbb{E}[X_1|X_2 = t] \\
&= \int_0^{+\infty} \{\widehat{C}_{1|2}(\overline{F}(x)|\overline{F}(t))\}dx \\
&= \int_0^{+\infty} \frac{1}{\overline{F}(t)} \exp\left\{-\left([-\log \overline{F}(x)]^\delta + [-\log \overline{F}(t)]^\delta\right)^{1/\delta}\right\} \\
&\qquad \times \left[1 + \left([-\log \overline{F}(x)]/[-\log \overline{F}(t)]\right)^\delta\right]^{-1+1/\delta} dx \\
&= \int_0^\infty \exp\left\{s\theta - \theta\left(y^\delta + s^\delta\right)^{1/\delta} + y\right\} \times \left[1 + (y/s)^\delta\right]^{-1+1/\delta} dy, \\
&= s\int_0^\infty \exp\left\{s\left[\theta - \theta\left(z^\delta + 1\right)^{1/\delta} + z\right]\right\} \times \left(1 + z^\delta\right)^{-1+1/\delta} dz, \\
&=: s\int_0^\infty h(z)e^{sg(z)}dz,
\end{aligned}$$

where $g(z) := \theta - \theta\left(z^\delta + 1\right)^{1/\delta} + z$ and $h(z) := \left(1 + z^\delta\right)^{-1+1/\delta}$. We may also use a Laplace approximation. From Example 7.1, the regularity conditions for $g(z)$ are hold. In addition, $h(z)$ is continuous on $(0, \infty)$, and $h(z_0) \neq 0$. Therefore (see Small, 2010), as $t \to \infty$,

$$\begin{aligned}
&\mathbb{E}[X_1|X_2 > t] \\
&\sim se^{sg(z_0)}h(z_0)\sqrt{\frac{2\pi}{-sg''(z_0)}} = (1+t)^{g(z_0)}h(z_0)\sqrt{\frac{2\pi\log(1+t)}{-g''(z_0)}}.
\end{aligned}$$

7.2.2 Intermediate tail dependence from hidden regular variation

Another relevant topic is the relationship[1] between intermediate tail dependence and Hidden Regular Variation (HRV) (Resnick, 2002).

In Resnick (2007), HRV is defined on the cone

$$E^0 := \{s \in E : \exists\, 1 \leq i < j \leq d, s_i \wedge s_j > 0\}; \tag{7.4}$$

[1]Thanks to Professor Haijun Li for pointing out the connection to Hidden Regular Variation.

that is, at most $d-2$ elements are 0, i.e., removing the axes of $E := [0, \infty]^d \setminus \{0\}$.

More precisely, if there exist scaling functions $b(t), b^0(t) \to \infty$, $b(t)/b^0(t) \to \infty$ as $t \to \infty$, such that as $t \to \infty$,

$$t\mathbb{P}\left(\frac{\mathbf{X}}{b(t)} \in B\right) \to \nu(B) \tag{7.5}$$

for all relatively compact sets $B \subset \overline{\mathbb{R}}_+^d \backslash \{0\}$ with $\nu(\partial B) = 0$, and

$$t\mathbb{P}\left(\frac{\mathbf{X}}{b^0(t)} \in B\right) \to \nu^0(B) \tag{7.6}$$

for all relatively compact sets $B \subseteq \overline{\mathbb{R}}_+^d \backslash \cup_{i=1}^d \{te^{(i)}, t \geq 0\}$ satisfying $\nu^0(\partial B) = 0$, then $\mathbf{X}$ is said to possess HRV on E^0.

The tail order functions (see Definition 3.2) we have studied is closely related to the concept of HRV. With certain regularity conditions, we can establish the relationship between tail order functions for the intermediate tail dependence case to the Radon measure ν^0 for the HRV. This is a technical result to be included in a future publication.

For further research on this topic, it would be interesting to study how the theory of HRV and tail order functions can be used to study the tail behavior of the conditional specification $\mathbb{E}[X_1 | X_2 > t]$ where (X_1, X_2) is intermediate tail dependent or tail quadrant independent in the upper tails, and how a useful statistical inference approach can be developed based on these theories to estimating the tail probability for these cases.

Bibliography

http://functions.wolfram.com/Bessel-TypeFunctions/BesselK/introductions/Bessels/ShowAll.html (accessed on June 04, 2012). → pages 74

K. Aas, C. Czado, A. Frigessi, and H. Bakken. Pair-copula constructions of multiple dependence. *Insurance: Mathematics and Economics*, 44(2):182–198, 2009. → pages 59, 61, 99

M. Abramowitz and I. A. Stegun. *Handbook of Mathematical Functions with Formulas, Graphs, and Mathematical Tables*, volume 55. Dover publications, 1964. → pages 76

S. Alink, M. Löwe, and M. V. Wüthrich. Diversification of aggregate dependent risks. *Insurance: Mathematics and Economics*, 35(1):77–95, 2004. → pages 144

S. Alink, M. Löwe, and M. V. Wüthrich. Analysis of the expected shortfall of aggregate dependent risks. *ASTIN Bulletin*, 35(1):25–43, 2005. → pages 144

S. Alink, M. Löwe, and M. V. Wüthrich. Diversification for general copula dependence. *Statistica Neerlandica*, 61(4):446–465, 2007. → pages 94, 144

P. Artzner, F. Delbaen, J.-M. Eber, and D. Heath. Coherent measures of risk. *Mathematical Finance*, 9(3):203–228, 1999. → pages 29, 144

A. Asimit, E. Furman, Q. Tang, and R. Vernic. Asymptotics for Risk Capital Allocations based on Conditional Tail Expectation. *Insurance: Mathematics and Economics*, 49(3):310–324, 2011. → pages 97, 107, 110

G. Balkema and N. Nolde. Asymptotic independence for unimodal densities. *Advances in Applied Probability*, 42(2):411–432, 2010. → pages 13

P. Barbe, A.-L. Fougères, and C. Genest. On the tail behavior of sums of dependent risks. *ASTIN Bulletin*, 36(2):361–373, 2006. → pages 144

T. Bedford and R. M. Cooke. Probability density decomposition for conditionally dependent random variables modeled by vines. *Annals of Mathematics and Artificial Intelligence.*, 32(1-4):245–268, 2001. → pages 61

T. Bedford and R. M. Cooke. Vines—a new graphical model for dependent random variables. *The Annals of Statistics*, 30(4):1031–1068, 2002. → pages 99

J. Beirlant, Y. Goegebeur, J. Teugels, and J. Segers. *Statistics of Extremes*. Wiley Series in Probability and Statistics. John Wiley & Sons Ltd., Chichester, 2004. → pages 10, 83

N. H. Bingham, C. M. Goldie, and J. L. Teugels. *Regular Variation*, volume 27 of *Encyclopedia of Mathematics and its Applications*. Cambridge University Press, Cambridge, 1987. → pages 11, 34, 50, 69, 103, 118, 161

N. Blomqvist. On a measure of dependence between two random variables. *Annals of Mathematical Statistics*, 21:593–600, 1950. → pages 33

G. Box and N. Draper. *Empirical Model-Building and Response Surfaces.* John Wiley & Sons, 1987. → pages 134

E. C. Brechmann, C. Czado, and K. Aas. Truncated regular vines in high dimensions with application to financial data. *Canadian Journal of Statistics*, 40(1):68–85, 2012. → pages 10, 111

L. Breiman. On some limit theorems similar to the arc-sin law. *Theory of Probability and its Applications*, 10(2):323–331, 1965. → pages 90, 146, 165

K. Breitung. *Asymptotic Approximations for Probability Integrals*. Springer, 1994. → pages 68

A. Charpentier and J. Segers. Tails of multivariate Archimedean copulas. *Journal of Multivariate Analysis*, 100(7):1521–1537, 2009. → pages 28, 40, 53, 55, 59, 74, 78, 120, 161

V. Chavez-Demoulin, P. Embrechts, and J. Nešlehová. Quantitative models for operational risk: extremes, dependence and aggregation. *Journal of Banking & Finance*, 30(10):2635–2658, 2006. → pages 8

S. Cheng and C. Jiang. The edgeworth expansion for distributions of extreme values. *Science in China. Series A. Mathematics*, 44(4):427–437, 2001. → pages 148

K. C. Cheung. Characterizations of conditional comonotonicity. *Journal of Applied Probability*, 44(3):607–617, 2007. → pages 80

K. C. Cheung. Upper comonotonicity. *Insurance: Mathematics and Economics*, 45(1):35–40, 2009. → pages 80, 94, 112

K. L. Chung. *A Course in Probability Theory, second edition*. Academic, New York, 1974. → pages 65

A. Colangelo, M. Scarsini, and M. Shaked. Some notions of multivariate positive dependence. *Insurance: Mathematics and Economics*, 37(1):13–26, 2005. → pages 49

S. Coles. *An Introduction to Statistical Modeling of Extreme Values*. Springer Series in Statistics. Springer-Verlag London Ltd., London, 2001. → pages 10

S. Coles, J. Heffernan, and J. Tawn. Dependence measures for extreme value analyses. *Extremes*, 2:339–365, 1999. → pages 39

R. D. Cook and M. E. Johnson. A family of distributions for modelling non-elliptically symmetric multivariate data. *Journal of the Royal Statistical Society B*, 43:210–218, 1981. → pages 59

R. Cooke, C. Kousky, and H. Joe. Micro correlations and tail dependence. In D. Kurowicka and H. Joe, editors, *Dependence Modeling: Vine Copula Handbook*, chapter 5, pages 89–112. World Scientific Publishing Company, Singapore, 2011. → pages 5, 114

E. Cope, G. Mignola, I. SpA, G. Antonini, and R. Ugoccioni. Challenges in Measuring Operational Risk from Loss Data. *IBM Zurich Research Lab, Intesa Sanpaolo SpA*, 2009. → pages 8

L. de Haan. *On Regular Variation and Its Application to The Weak Convergence of Sample Extremes*. Mathematisch Centrum, 1970. → pages 63

L. de Haan. Equivalence classes of regularly varying functions. *Stochastic Processes and their Applications*, 2(3):243–259, 1974. → pages 63

L. de Haan and A. Ferreira. *Extreme Value Theory: An introduction*. Springer Series in Operations Research and Financial Engineering. Springer, New York, 2006. → pages 10, 63, 145, 146, 148, 162

L. de Haan and U. Stadtmüller. Generalized regular variation of second order. *Journal of the Australian Mathematical Society*, 61(03):381–395, 1996. → pages 121, 145, 146, 147

M. Degen, D. D. Lambrigger, and J. Segers. Risk concentration and diversification: second-order properties. *Insurance: Mathematics and Economics*, 46(3):541–546, 2010. → pages 121, 145

M. Denuit, J. Dhaene, M. Goovaerts, and R. Kaas. *Actuarial Theory for Dependent Risks*. Wiley, 2005. → pages 10, 29, 30, 34, 141, 143

J. Dhaene, M. Denuit, M. J. Goovaerts, R. Kaas, and D. Vyncke. The concept of comonotonicity in actuarial science and finance: theory. *Insurance: Mathematics and Economics*, 31(1):3–33, 2002a. → pages 2, 80, 112

J. Dhaene, M. Denuit, M. J. Goovaerts, R. Kaas, and D. Vyncke. The concept of comonotonicity in actuarial science and finance: applications. *Insurance: Mathematics and Economics*, 31(2):133–161, 2002b. → pages 80, 112

J. Dhaene, S. Vanduffel, M. J. Goovaerts, R. Kaas, Q. Tang, and D. Vyncke. Risk measures and comonotonicity: a review. *Stochastic Models*, 22(4):573–606, 2006. → pages 93

J. Dong, K. Cheung, and H. Yang. Upper comonotonicity and convex upper bounds for sums of random variables. *Insurance: Mathematics and Economics*, 47(2):159–166, 2010. → pages 112

K. Dutta and J. Perry. A tale of tails: An empirical analysis of loss distribution models for estimating operational risk capital. *Federal Reserve Bank of Boston, Working Paper*, No 06-13, 2006. → pages 7

P. Embrechts, C. Klüppelberg, and T. Mikosch. *Modelling Extremal Events*, volume 33. Springer-Verlag, Berlin, 1997. → pages 10, 34, 62, 95, 96, 110, 118, 143

P. Embrechts, A. J. McNeil, and D. Straumann. Correlation and dependence in risk management: properties and pitfalls. In *Risk management: Value at Risk and Beyond*, pages 176–223. Cambridge Univ. Press, Cambridge, 2002. → pages 8, 92

P. Embrechts, D. Lambrigger, and M. Wüthrich. Multivariate extremes and the aggregation of dependent risks: examples and counter-examples. *Extremes*, 12 (2):107–127, 2009a. → pages 92, 142

P. Embrechts, J. Nešlehová, and M. V. Wüthrich. Additivity properties for value-at-risk under Archimedean dependence and heavy-tailedness. *Insurance Math. Econom.*, 44(2):164–169, 2009b. → pages 94, 144

R. Engle. Risk and volatility: Econometric models and financial practice. *The American Economic Review*, 94(3):405–420, 2004. → pages 6

M. Falk, J. Hüsler, and R. Reiss. *Laws of Small Numbers: Extremes and Rare Events*. Springer Basel, 2010. → pages 10

K. Fang, S. Kotz, and K. Ng. *Symmetric Multivariate and Related Distributions, Monographs on Statistics and Applied Probability, 36*. Chapman and Hall, 1990. → pages 43, 44, 90, 91

W. Feller. *An Introduction to Probability Theory and Its Applications. Vol. II.* Second edition. John Wiley & Sons Inc., New York, 1971. → pages 88

G. Frahm, M. Junker, and A. Szimayer. Elliptical copulas: applicability and limitations. *Statistics & Probability Letters*, 63(3):275–286, 2003. → pages 91

E. W. Frees and E. A. Valdez. Understanding relationships using copulas. *North American Actuarial Journal*, 2(1):1–25, 1998. → pages 4, 10, 111, 131, 133

J. Geluk and L. de Haan. *Regular Variation, Extensions and Tauberian Theorems*. CWI Tract 40, Center for Mathematics and Computer Science, Amsterdam, The Netherlands., 1987. → pages 34

J. Geluk, L. de Haan, S. Resnick, and C. Stărică. Second-order regular variation, convolution and the central limit theorem. *Stochastic Processes and their Applications*, 69(2):139–159, 1997. → pages 145, 146, 147

C. Genest and A.-C. Favre. Everything you always wanted to know about copula modeling but were afraid to ask. *Journal of Hydrologic Engineering*, 12: 347–368, 2007. → pages 10, 111

C. Genest and R. J. MacKay. Copules archimédiennes et familles de lois bidimensionnelles dont les marges sont données. *The Canadian Journal of Statistics*, 14(2):145–159, 1986. → pages 32

C. Genest and L. Rivest. Statistical inference procedures for bivariate archimedean copulas. *Journal of the American Statistical Association*, pages 1034–1043, 1993. → pages 10

C. Genest, M. Gendron, and M. Bourdeau-Brien. The advent of copulas in finance. *The European Journal of Finance*, 15(7-8):609–618, 2009a. → pages 34

C. Genest, B. Rémillard, and D. Beaudoin. Goodness-of-fit tests for copulas: a review and a power study. *Insurance: Mathematics and Economics*, 44(2): 199–213, 2009b. → pages 10, 111

M. Goovaerts, R. Kaas, and R. Laeven. Decision principles derived from risk measures. *Insurance: Mathematics and Economics*, 47(3):294–302, 2010. → pages 30

E. Hashorva. Asymptotic properties of type I elliptical random vectors. *Extremes*, 10(4):175–206, 2007. → pages 43, 62, 63

E. Hashorva. Tail asymptotic results for elliptical distributions. *Insurance: Mathematics and Economics*, 43(1):158–164, 2008. → pages 43

E. Hashorva. On the residual dependence index of elliptical distributions. *Statistics & Probability Letters*, 80(13-14):1070–1078, 2010. → pages 43

E. Hashorva and J. Hüsler. On multivariate Gaussian tails. *Annals of the Institute of Statistical Mathematics*, 55(3):507–522, 2003. → pages 42

J. E. Heffernan. A directory of coefficients of tail dependence. *Extremes*, 3(3): 279–290 (2001), 2000. → pages 39, 59

L. Hua and H. Joe. Tail order and intermediate tail dependence of multivariate copulas. *Journal of Multivariate Analysis*, 102:1454–1471, 2011a. → pages iv

L. Hua and H. Joe. Second order regular variation and conditional tail expectation of multiple risks. *Insurance: Mathematics and Economics*, 49:537–546, 2011b. → pages v

L. Hua and H. Joe. Tail comonotonicity and conservative risk measures. *to appear in ASTIN Bulletin*, 2012a. → pages iv

L. Hua and H. Joe. Tail comonotonicity: properties, constructions, and asymptotic additivity of risk measures. *in revision*, 2012b. → pages iv

H. Hult and F. Lindskog. Multivariate extremes, aggregation and dependence in elliptical distributions. *Advances in Applied Probability*, 34(3):587–608, 2002. → pages 91, 92, 105

P. Jaworski, F. Durante, W. K. Härdle, and T. Rychlik, editors. *Copula Theory and Its Applications: Proceedings of the Workshop Held in Warsaw, 25-26 September 2009*, volume 198. Springer Verlag, 2010. → pages 10

H. Joe. Multivariate concordance. *Journal of Multivariate Analysis*, 35(1):12–30, 1990. → pages 26

H. Joe. *Multivariate Models and Dependence Concepts*, volume 73 of *Monographs on Statistics and Applied Probability*. Chapman & Hall, London, 1997. → pages 9, 12, 32, 34, 45, 49, 55, 56, 59, 60, 61, 84, 85, 87, 106, 111, 112, 115, 116, 117, 118, 119

H. Joe. Discussion of: "Copulas: tales and facts" by T. Mikosch [Extremes **9** (2006), no. 1, 3–20]. *Extremes*, 9(1):37–41, 2006. → pages 10

H. Joe. Tail dependence in vine copulae. In D. Kurowicka and H. Joe, editors, *Dependence Modeling: Vine Copula Handbook*, chapter 8. World Scientific Publishing Company, Singapore, 2011. → pages 92

H. Joe and T. Hu. Multivariate distributions from mixtures of max-infinitely divisible distributions. *Journal of Multivariate Analysis*, 57(2):240–265, 1996. → pages 32, 56, 59, 84, 85, 136

H. Joe and H. Li. Tail risk of multivariate regular variation. *Methodology and Computing in Applied Probability*, 13:671–693, 2011. → pages 154, 158, 173

H. Joe and C. Ma. Multivariate survival functions with a min-stable property. *Journal of Multivariate Analysis*, 75(1):13–35, 2000. → pages 54

H. Joe, H. Li, and A. K. Nikoloulopoulos. Tail dependence functions and vine copulas. *Journal of Multivariate Analysis*, 101:252–270, 2010. → pages 31, 45, 59, 61, 83, 99

A. Juri and M. V. Wüthrich. Copula convergence theorems for tail events. *Insurance: Mathematics and Economics*, 30(3):405–420, 2002. → pages 31

A. Juri and M. V. Wüthrich. Tail dependence from a distributional point of view. *Extremes*, 6(3):213–246 (2004), 2003. → pages 31

S. A. Klugman and R. Parsa. Fitting bivariate loss distributions with copulas. *Insurance: Mathematics and Economics*, 24(1-2):139–148, 1999. 1st IME Conference (Amsterdam, 1997). → pages 131

C. Klüppelberg, G. Kuhn, and L. Peng. Semi-parametric models for the multivariate tail dependence function—the asymptotically dependent case. *Scandinavian Journal of Statistics*, 35(4):701–718, 2008. → pages 31, 47

S. Kotz, N. Balakrishnan, and N. L. Johnson. *Continuous Multivariate Distributions. Vol. 1*. Wiley Series in Probability and Statistics: Applied Probability and Statistics. Wiley-Interscience, New York, second edition, 2000. → pages 158

D. Kurowicka and H. Joe. *Dependence Modeling: Vine Copula Handbook*. World Scientific Publishing Company, Singapore, 2011. → pages 10

Z. Landsman and E. Valdez. Tail conditional expectations for elliptical distributions. *North American Actuarial Journal*, 7(4):55–71, 2003. → pages 113

M. Larsson and J. Nešlehová. Extremal behavior of archimedean copulas. *Advances in Applied Probability*, 43(1):195–216, 2011. → pages 55

A. W. Ledford and J. A. Tawn. Statistics for near independence in multivariate extreme values. *Biometrika*, 83(1):169–187, 1996. → pages 38, 39, 42

A. W. Ledford and J. A. Tawn. Modelling dependence within joint tail regions. *Journal of the Royal Statistical Society. Series B. Methodological*, 59(2): 475–499, 1997. → pages 39

D. Li. On default correlation: a copula function approach. *Journal of Fixed Income*, 9(4):43–54, 2000. → pages 10

H. Li and Y. Sun. Tail dependence for heavy-tailed scale mixtures of multivariate distributions. *Journal of Applied Probability*, 46(4):925–937, 2009. → pages 31

G. Mainik and L. Rüschendorf. Ordering of multivariate risk models with respect to extreme portfolio losses. *Statistics & Risk Modeling*, 29(1):73–106, 2012. → pages 30, 124

A. Marshall and I. Olkin. Domains of attraction of multivariate extreme value distributions. *The Annals of Probability*, 11:168–177, 1983. → pages 37

A. W. Marshall and I. Olkin. Families of multivariate distributions. *Journal of the American Statistical Association*, 83(403):834–841, 1988. → pages 32

A. W. Marshall and I. Olkin. *Life Distributions*. Springer Series in Statistics. Springer, New York, 2007. Structure of nonparametric, semiparametric, and parametric families. → pages 32, 117

A. J. McNeil and J. Nešlehová. Multivariate Archimedean copulas, d-monotone functions and l_1-norm symmetric distributions. *Annals of Statistics*, 37(5B): 3059–3097, 2009. → pages 32

A. J. McNeil, R. Frey, and P. Embrechts. *Quantitative Risk Management*. Princeton Series in Finance. Princeton University Press, Princeton, NJ, 2005. Concepts, techniques and tools. → pages 10, 19, 34, 62, 125, 142, 143

T. Mikosch. "Copulas: tales and facts" [Extremes **9** (2006), no. 1, 3–20; mr2327842]—rejoinder. *Extremes*, 9(1):55–62, 2006. → pages 10

A. Müller and D. Stoyan. *Comparison Methods for Stochastic Models and Risks*. Wiley Series in Probability and Statistics. John Wiley & Sons Ltd., Chichester, 2002. → pages 30, 88

H. Nam, Q. Tang, and F. Yang. Characterization of upper comonotonicity via tail convex order. *Insurance: Mathematics and Economics*, 48(3):368, 2011. → pages 112

R. B. Nelsen. *An Introduction to Copulas*. Springer Series in Statistics. Springer, New York, second edition, 2006. → pages 9, 32, 34, 89, 107, 109

C. Neves. From extended regular variation to regular variation with application in extreme value statistics. *Journal of Mathematical Analysis and Applications*, 355(1):216–230, 2009. → pages 147

A. Nikoloulopoulos, H. Joe, and H. Li. Vine copulas with asymmetric tail dependence and applications to financial return data. *Computational Statistics & Data Analysis*, in press, 2012. → pages 79

A. K. Nikoloulopoulos, H. Joe, and H. Li. Extreme value properties of multivariate t copulas. *Extremes*, 12:129–148, 2009. → pages 31

T. Okimoto. New evidence of asymmetric dependence structures in international equity markets. *Journal of Financial and Quantitative Analysis*, 43(03): 787–815, 2008. → pages 38

OSFI. Guideline A: Minimum Continuing Capital and Surplus Requirements (MCCSR) for Life Insurance Companies. 2011. → pages 29, 143

A. Patton. Copula–based models for financial time series. *Handbook of Financial Time Series*, pages 767–785, 2009. → pages 113

A. J. Patton. Modelling asymmetric exchange rate dependence. *International Economic Review*, 47(2):527–556, 2006. → pages 38

W. Press. *Numerical Recipes: the Art of Scientific Computing*. Cambridge Univ Pr, 2007. → pages 105

A. Ramos and A. Ledford. A new class of models for bivariate joint tails. *Journal of the Royal Statistical Society: Series B (Statistical Methodology)*, 71(1): 219–241, 2009. → pages 39

R. Reiss. *Approximate Distributions of Order Statistics: with Applications to Nonparametric Statistics*. Springer Series in Statistics. Springer-Verlag. New York, 1989. → pages 12

R. Reiss and M. Thomas. *Statistical Analysis of Extreme Values*. Birkhäuser Basel, 1997. → pages 10

S. Resnick. Hidden regular variation, second order regular variation and asymptotic independence. *Extremes*, 5(4):303–336 (2003), 2002. → pages 180

S. I. Resnick. *Extreme Values, Regular Variation, and Point Processes*, volume 4 of *Applied Probability. A Series of the Applied Probability Trust*. Springer-Verlag, New York, 1987. → pages 11, 34, 107

S. I. Resnick. *Heavy-tail Phenomena*. Springer Series in Operations Research and Financial Engineering. Springer, New York, 2007. → pages 11, 34, 36, 45, 47, 67, 83, 95, 96, 106, 107, 108, 110, 146, 147, 180

J. Rodgers and W. Nicewander. Thirteen ways to look at the correlation coefficient. *American Statistician*, 42(1):59–66, 1988. → pages 8

R. Schmidt. Tail dependence for elliptically contoured distributions. *Mathematical Methods of Operations Research*, 55(2):301–327, 2002. Special issue on mathematical models in market and credit risk. → pages 12, 43, 91, 92

S. M. Shah and N. T. Parikh. Moments of singly and doubly truncated standard bivariate normal distribution. *Vidya*, 7(1):81–91, 1964. → pages 158

M. Shaked and J. G. Shanthikumar. *Stochastic Orders*. Springer Series in Statistics. Springer, New York, 2007. → pages 115

A. Sklar. Fonctions de répartition à n dimensions et leurs marges. *Publ. Inst. Statist. Univ. Paris*, 8:229–231, 1959. → pages 9, 31

C. Small. *Expansions and Asymptotics for Statistics*, volume 115. CRC Press, 2010. → pages 175, 180

X. Wang and S. Cheng. General regular variation of n–th order and the 2nd order edgeworth expansion of the extreme value distribution (i). *Acta Mathematica Sinica*, 21(5):1121–1130, 2005. → pages 147

C. Weng and Y. Zhang. Characterization of multivariate heavy-tailed distribution families via copula. *Journal of Multivariate Analysis*, 106:178–186, 2012. → pages 173

L. Zhu and H. Li. Asymptotic analysis of multivariate tail conditional expectations. *to appear in North American Actuarial Journal*, 2012. → pages 90, 110, 113, 144, 145, 154, 155, 164