Using Copulae to bound the Value-at-Risk for functions of dependent risks

Paul Embrechts, Andrea Höing, Alessandro Juri

Department of Mathematics ETHZ, CH-8092 Zurich, Switzerland.
(e-mail: embrechts, ahoeing, juri@math.ethz.ch.)

JEL Classification: G10

Mathematics Subject Classification (1991): 90A09, 60E05, 60E15

Abstract. The theory of copulae is known to provide a useful tool for modelling dependence in integrated risk management. In the present paper we review and extend some of the more recent results for finding distributional bounds for functions of dependent risks. As an example, the main emphasis is put on Value-at-Risk as a risk measure.

Key words: comonotonicity, copula, dependent risks, Fréchet bounds, orthant dependence, risk management, Value-at-Risk

1 Introduction

Due to its simplicity, but also because of regulatory reasons, Value-at Risk (VaR) remains one of the most popular risk measures. This despite some fundamental criticism (see for instance Artzner et al. (1999)). The aim of this paper is to give more insight into the problem of managing the VaR of a joint position resulting from the combination of different dependent risks.

Suppose for example that we have a VaR-based risk management system both for market and credit risk denoted respectively by X_1 and X_2. The quantity $\mathrm{VaR}_\alpha(X_i)$, $i = 1, 2$ denotes the Value-at-Risk at $100\,\alpha\%$ for a 1-day holding period X_i, i.e. the α-quantile of the daily profit-and-loss distribution function F_i of the position X_i, the latter calculated (estimated) through bank-internal models. At the integrated level the bank has to measure the risk (compute the VaR) of the joint position $X_1 + X_2$. Since $\mathrm{VaR}_\alpha(X_1 + X_2)$ is just the α-quantile of $F_{X_1+X_2}$, it is clear that the knowledge of the joint distribution of X_1 and X_2 reduces the problem of determining $\mathrm{VaR}_\alpha(X_1 + X_2)$ to a computational issue. Unfortunately, in many situations only partial or no information at all about the dependence between the two risks is available, i.e. the joint distribution of X_1 and X_2 is unknown. In practice, most often the sum $\mathrm{VaR}_\alpha(X_1) + \mathrm{VaR}_\alpha(X_2)$ is taken

as a measure of risk in the worst case, i.e. as an upper bound for $\mathrm{VaR}_\alpha(X_1+X_2)$. However, Fallacy 3 in Embrechts et al. (2002) states that "The worst case VaR for a portfolio $X_1 + X_2$ occurs when the linear correlation function $\rho(X_1, X_2)$ is maximal". This statement, though perhaps intuitively clear is in fact wrong. It can be shown that $\rho(X_1, X_2)$ maximal occurs for so called comonotonic risks; see Section 2.3 below. Moreover, we will see in Proposition 3.1 that for comonotonic random variables $\mathrm{VaR}_\alpha(X_1 + X_2) = \mathrm{VaR}_\alpha(X_1) + \mathrm{VaR}_\alpha(X_2)$. However, it may be that $\mathrm{VaR}_\alpha(X_1 + X_2) > \mathrm{VaR}_\alpha(X_1) + \mathrm{VaR}_\alpha(X_2)$.

The question at hand, and the one lying at the heart of the work on coherent risk measurement in Artzner et al. (1999), concerns the fact that typically for non-elliptical portfolios, VaR will not be subadditive. Of course dependence will play a fundamental role in determining (or bounding) the value of the so called *measure of diversification* $\Delta(\mathrm{VaR}_\alpha) := \mathrm{VaR}_\alpha(X_1 + X_2) - (\mathrm{VaR}_\alpha(X_1) + \mathrm{VaR}_\alpha(X_2))$ as discussed in Embrechts et al. (2002). The latter paper has clearly shown that linear correlation is insufficient as a measure of dependence for studying $\Delta(\mathrm{VaR}_\alpha)$ across a wide range of potential portfolio structures for (X_1, X_2).

More generally, the problem becomes the following: how can we bound VaR, or indeed any reasonable measure of risk, on the global position $\psi(X_1, \ldots, X_n)$, where $\psi : \mathbb{R}^n \to \mathbb{R}$ is some function of interest, if we only know the marginal profit-and-loss distributions $F_1, \ldots, F_n$ of the n one-period risks $X_1, \ldots, X_n$. One may think of the X_i's as different types of financial or insurance risk. In the above case we took $n = 2$ and $\psi(x_1, x_2) = x_1 + x_2$. The key point being that we may not have any dependence information on the X_i's. Further, how do these bounds change when specific dependence information is assumed.

2 Definitions and preliminaries

Generalized inverses of increasing functions and copulae constitute the technical instruments underlying the solution to the Value-at-Risk problem in the introduction. Most of the results presented exist in some form or other in the literature, though many may be hard to find for the non-specialist. Specific references will be given where relevant.

In the rest of the paper, we will often consider the *componentwise order* on $\mathbb{R}^n$ defined for $x, y \in \mathbb{R}^n$ as $x \leq y$ if $x_i \leq y_i$ for all $i = 1, \ldots, n$. Further, a real-valued function defined on $\mathbb{R}^n$ is said to be *increasing* if it is so with respect to the componentwise order.

2.1 Generalized inverses and Value-at-Risk

Generalized inverses of distribution functions are essential for a precise formulation of the above described Value-at-Risk problem. They will also play an important role in the proofs of the results of the next section.

Let $\overline{\mathbb{R}} := \mathbb{R} \cup \{\pm\infty\}$ denote the extended real line. In the following we will make use of the convention $\inf \emptyset = \sup \emptyset = -\infty$.

Definition 2.1. Let $\varphi : \mathbb{R} \to \mathbb{R}$ be an increasing function. Its *generalized left* and *right continuous inverses* are the functions $\varphi^{-1} : \mathbb{R} \to \overline{\mathbb{R}}$ and $\varphi^{\wedge} : \mathbb{R} \to \overline{\mathbb{R}}$ defined by

$$\varphi^{-1}(y) := \inf\{x \in \mathbb{R} \,|\, \varphi(x) \geq y\} \quad \text{and} \quad \varphi^{\wedge}(y) := \sup\{x \in \mathbb{R} \,|\, \varphi(x) \leq y\}.$$

These generalized inverses possess the following properties; see Embrechts et al. (1997), p.130 and Appendix A1.6 for further details and references.

Lemma 2.1.

(i) φ^{-1} *and* $\varphi^{\wedge}$ *are increasing functions,*

(ii) φ^{-1} *and* $\varphi^{\wedge}$ *are left, respectively right continuous on* $\mathbb{R}$*,*

(iii) If φ *is right continuous and* $\varphi^{-1}(y) > -\infty$*, then* $\varphi(x) \geq y \Leftrightarrow x \geq \varphi^{-1}(y)$*,*

(iv) If φ *is left continuous and* $\varphi^{\wedge}(y) > -\infty$*, then* $\varphi(x) \leq y \Leftrightarrow x \leq \varphi^{\wedge}(y)$*.*

Definition 2.2. For $0 \leq \alpha \leq 1$ the *Value-at-Risk at probability level* α of a random variable X is its α-quantile, i.e. $\mathrm{VaR}_\alpha(X) := F_X^{-1}(\alpha)$.

For risk management applications, F_X could stand for a 1- or 10-day market profit-and-loss distribution for which, for a given level α, $F_X^{-1}(\alpha)$ would correspond to the market VaR at the level α. We will assume that losses are represented in the right tail of F_X and hence typical values for α in this (market risk) case are 0.95 or 0.99. Similar interpretations can be given for credit risk (leading to Credit-VaR) and operational risk (OpRisk-VaR). In the latter case one typically looks at yearly aggregate loss data with $\alpha = 0.9995$, say.

2.2 Dependence structures and copulae

Very often independence assumptions in stochastic models are more due to their tractability rather than to the nature of the phenomenon being modelled.

For the static one-period situation given by real-valued random variables $X_1, \ldots, X_n$, the dependence between $X_1, \ldots, X_n$ is completely determined by their joint distribution function $F(x_1, \cdots, x_n) = P[X_1 \leq x_1, \cdots, X_n \leq x_n]$. The idea of separating F into two parts, one describing the dependence structure and another part describing the marginal behaviour only, leads to the by now well known concept of a copula. Excellent introductions to copulae and related concepts are given in Nelsen (1999) and Joe (1997), where most of the material of this section can be found. Risk management applications of copulae are considered in Embrechts et al. and (2001,2002); the latter papers contain numerous references to the existing work on copulae.

Definition 2.3. An *n-dimensional copula* is an n-dimensional distribution function restricted to $[0,1]^n$ with uniform-$(0,1)$ marginals.

For a given copula C and marginals $F_1, \ldots, F_n$ one has that

$$F(x_1, \ldots, x_n) = C(F_1(x_1), \ldots, F_n(x_n)) \tag{2.1}$$

is a distribution function with these marginals. Conversely, for a given joint distribution function F with marginals $F_1, \ldots, F_n$ there is always a copula C satisfying (2.1). This copula is not necessarily unique, but it is if $F_1, \ldots, F_n$ are continuous and in this case we have that $C(u_1, \ldots, u_n) = F(F_1^{-1}(u_1), \ldots, F_n^{-1}(u_n))$. These results are known as Sklar's Theorem and are a motivation for calling a copula a *dependence structure*. In fact, equation (2.1) means that the copula C separates the marginal behaviour given by $F_1, \ldots, F_n$ from the dependence contained in the joint distribution function F.

The so called *survival copula* and the *dual* of a copula defined below will play an important role for the results of Sections 3 and 4.

Definition 2.4. Let $(U_1, \ldots, U_n)$ be an n-dimensional random vector with standard uniform marginals and C be its distribution function. The *dual of* C is defined by

$$C^d(u_1, \ldots, u_n) := P[\cup_{i=1}^n \{U_i \leq u_i\}]$$

and the *survival copula* $\hat{C}$ of C is the distribution function of $(1-U_1, \ldots, 1-U_n)$.

Observe that for random variables $X_1, \ldots, X_n$ with joint distribution function F, marginals $F_1, \ldots, F_n$ and copula C

$$C^d(F_1(x_1), \ldots, F_n(x_n)) = P\left[\cup_{i=1}^n \{X_i \leq x_i\}\right], \tag{2.2}$$

$$\hat{C}(\overline{F}_1(x_1), \ldots, \overline{F}_n(x_n)) = \overline{F}(x_1, \ldots, x_n), \tag{2.3}$$

where $\overline{F}_i(x_i) := 1 - F_i(x_i)$ and $\overline{F}(x_1, \ldots, x_n) := P[X_1 > x_1, \ldots, X_n > x_n]$. Note that $C^d(F_1(x_1), \ldots, F_n(x_n))$ is increasing in each of the arguments $x_1, \ldots, x_n$ and that (2.3) is the analogue of Sklar's Theorem for survival distribution functions. Finally, observe that contrary to $\hat{C}$ the dual C^d is not a copula and that

$$C^d(u_1, \ldots, u_n) = 1 - \hat{C}(1 - u_1, \ldots, 1 - u_n). \tag{2.4}$$

The natural question relative to dependence structures concerns their comparison, i.e. we have to specify which copula leads to a strong or to a weak kind of dependence. The most natural approach is to compare copulae pointwise as functions and to define the riskiness of a dependence structure through this comparison. At this extent, we want to recall that any copula C lies between the so called *lower* and *upper Fréchet bounds* $C_{\mathrm{L}}(u_1, \ldots, u_n) := \left(\sum_{i=1}^n u_i - n + 1\right)^+$ and $C_{\mathrm{U}}(u_1, \ldots, u_n) := \min_{1 \leq i \leq n} u_i$, namely

$$C_{\mathrm{L}} \leq C \leq C_{\mathrm{U}}. \tag{2.5}$$

Observe that, contrary to C_{U}, the lower Fréchet bound C_{L} is not a distribution function for $n \geq 3$.

2.3 Comonotonicity and orthant dependence

The issue of comparing copulae described above is part of the more general framework of comparing probability measures on $\mathbb{R}^n$ making use of partial orders

(see Müller (1996,1997a,1997b) and Bäuerle and Müller (1998) among others). For a considered partial order and two ordered distributions, the dependence between the marginals of the larger one are said to be more (or less) risky than the one between the marginals of the smaller one.

The upper Fréchet bound C_{U} corresponds to the riskiest possible dependence with respect to many partial orders for distributions as e.g. the *lower orthant* and the *supermodular order*; see Bäuerle and Müller (1998) for definitions of the latter. Random variables $X_1, \ldots, X_n$ with a C_{U}-dependence structure are called *comonotonic*. An equivalent formulation of comonotonicity allows for a representation

$$(X_1, \ldots, X_n) = (f_1(Z), \ldots, f_n(Z)) \quad \text{in law}, \tag{2.6}$$

where $f_1, \ldots, f_n : \mathbb{R} \to \mathbb{R}$ are increasing and Z is some random variable. We see from (2.6) that comonotonic random variables are increasing functions of a common random variable and therefore strongly dependent. It is precisely this representation which motivates the use of the concept of comonotonicity in financial applications. The random variable Z can be seen as a common (or even causal) underlying factor. Usually, one makes use of the representation

$$(X_1, \ldots, X_n) = (F_1^{-1}(U), \ldots, F_n^{-1}(U)) \quad \text{in law}, \tag{2.7}$$

where the F_i's are the distribution of the X_i's and U is any standard uniformly distributed variable.

Besides this extreme situation, there are other interesting ones. For example, random variables $X_1, \ldots, X_n$ having a copula C satisfying $C \geq C_{\mathrm{I}}$, where

$$C_{\mathrm{I}}(u_1, \ldots, u_n) := \prod_{i=1}^{n} u_i \tag{2.8}$$

is the copula corresponding to independent copies of the X_i's, are called *positive lower orthant dependent*, PLOD for short. A brief discussion on the orthant order, its interpretation and its connections with other related concepts is given below.

Definition 2.5. Let $X = (X_1, \ldots, X_n)$ and $Y = (Y_1, \ldots, Y_n)$ have pairwise equal marginal distributions and $\overline{F}_X(x_1, \ldots, x_n) := P[X_1 > x_1, \ldots, X_n > x_n]$, $\overline{F}_Y(x_1, \ldots, x_n) := P[Y_1 > x_1, \ldots, Y_n > x_n]$,. The random vector X is said to be smaller than Y in the *upper orthant order*, written $X \leq_{\mathrm{uo}} Y$, if $\overline{F}_X \leq \overline{F}_Y$. Similarly, X is said to be smaller than Y in the *lower orthant order*, written $X \leq_{\mathrm{lo}} Y$, if $F_X \leq F_Y$.

Denoting by $\tilde{X}_1, \ldots, \tilde{X}_n$ independent copies of $X_1, \ldots, X_n$ we see that $\tilde{X} \leq_{\mathrm{lo}} X$ is equivalent to saying that $X_1, \ldots, X_n$ are PLOD. Similarly, in the case $\tilde{X} \leq_{\mathrm{uo}} X$ the variables $X_1, \ldots, X_n$ are said to be *positive upper orthant dependent*, PUOD for short. In the case where we have both PLOD and PUOD, then one simply speaks about *positive orthant dependent* (POD) risks. Intuitively, PLOD (PUOD) means that the risks $X_1, \ldots, X_n$ are more likely to take simultaneously

small (large) values compared to independent copies of them. The important fact about PLOD, PUOD and $\leq_{\text{lo}}$, $\leq_{\text{uo}}$ is that they are implied by other stronger dependence concepts and relations such as *association*, *conditional increasing in sequence* (CIS) and *supermodular order* (see Joe (1997), Theorem 2.4 and Müller and Scarsini (2000)). These orders are also consistent with the intuition behind usual linear correlation. In fact (see Dhaene and Goovaerts (1996)) it can be shown that $X \leq_{\text{uo}} Y$ or $X \leq_{\text{lo}} Y$ implies for any f, g increasing and all $i \neq j$ that

$$\text{Cor}(f(X_i), g(X_j)) \leq \text{Cor}(f(Y_i), g(Y_j)). \tag{2.9}$$

In particular, for $X_1, \ldots, X_n$ PLOD or PUOD the correlation between increasing transformed X_i, X_j is non-negative, whence the name "positive dependence". Finally, when $n = 2$, the orders $\leq_{\text{lo}}$ and $\leq_{\text{uo}}$ coincide and one usually speaks about *positive quadrant dependent* (PQD) risks instead of PLOD, PUOD or POD risks.

Many copulae are governed by parameters controlling the amount of dependence between the components. Examples are the *Clayton* and the *Gumbel* copulae

$$C^{\text{Cl},\alpha}(u_1, \ldots, u_n) := \left(\sum_{i=1}^{n} u_i^{-\alpha} - n + 1 \right)^{-1/\alpha}, \tag{2.10}$$

$$C^{\text{Gu},\beta}(u_1, \ldots, u_n) := \exp\left(- \left(\sum_{i=1}^{n} (-\log u_i)^{1/\beta} \right)^{\beta} \right), \tag{2.11}$$

where $0 < \alpha < \infty$ and $0 < \beta \leq 1$. It can be shown (see Nelsen (1999), Corollary 4.4.5) that $C^{\text{Cl},\alpha}$ and $C^{\text{Gu},\beta}$ increase with respect to $\leq_{\text{lo}}$ as the parameters α and β increase. Moreover, for α and β at the boundary of the respective parameter intervals one gets (as a limit) independence and comonotonicity respectively. For $n = 2$, the copula $C^{\text{Gu},\beta}$ is consistent with bivariate extreme value theory and could be used to model the limiting dependence structure of componentwise maxima of bivariate random samples (Galambos (1987), Genest and Rivest (1989) and Joe (1997)). A motivation for the use of $C^{\text{Cl},\alpha}$ when dealing with conditional bivariate extremes is given in Juri and Wüthrich (2002).

3 Distributional Bounds

For given risks $X_1, \ldots, X_n$ and a function $\psi : \mathbb{R}^n \to \mathbb{R}$ one often is interested in computing certain quantities of $\psi(X_1, \ldots, X_n)$ like some moments or a quantile. For our discussion below, we shall concentrate on estimating quantiles of $\psi(X_1, \ldots, X_n)$ and always refer to this as estimating a VaR. Typical examples of ψ include:

- $\psi(x_1, \ldots, x_n) = x_1 + \cdots + x_n$, hence in the quantile case one would be interested in the Value-at-Risk of the joint position $X_1 + \cdots + X_n$.

- $\psi(x_1, \ldots, x_n) = \sum_{i=1}^n (x_i - k)^+$, $k \geq 0$. This corresponds to the functional underlying an excess-of-loss treaty in reinsurance. The X_i's could be individual claims or insurance losses due to different lines of business.

- $\psi(x_1, \ldots, x_n) = (\sum_{i=1}^n x_i - k)^+$, $k \geq 0$, with an obvious interpretation to financial derivatives (e.g. Asian options) or stop-loss reinsurance.

- The worlds of exotic options, basket derivatives, credit derivatives and operational risk insurance covers yield numerous further interesting examples.

As already mentioned in the introduction, determining $\mathrm{VaR}_\alpha(\psi(X_1, \ldots, X_n))$ reduces to a computational issue once the copula relative to $X_1, \ldots, X_n$ is specified. Since this is usually not the case, it would be useful to find bounds between which VaR must lie or equivalently bounds for the distribution function of $\psi(X_1, \ldots, X_n)$ as $\mathrm{VaR}_\alpha(\psi(X_1, \ldots, X_n))$ is just its α-quantile. Such bounds and the proof of their sharpness have been found first by Makarov (1981) in terms of generalized inverses in the case $n = 2$ and $\psi(x_1, x_2) = x_1 + x_2$. Frank et al. (1987) proved the same results using copulae, i.e introducing (without mentioning them explicitly) dependence structures. They also extended Makarov's results (except for the optimality of the bounds) to include arbitrary increasing continuous functions ψ. Williamson and Downs (1990) proved the pointwise best-possible nature of the bounds in the two-dimensional case and also developed an algorithm for computing these bounds numerically. Finally, Denuit et al. (1999) and Cossette et al. (2000) extended some of the existing theoretical results to $n \geq 3$. In the next section we further generalize these results by relaxing some of the continuity assumptions relative to ψ. Simpler and modified proofs of Theorems 3.1 and 3.2 of the next section are given in the Appendix.

3.1 Main results

For $\psi : \mathbb{R}^n \to \mathbb{R}$ and $1 \leq i_1 < \cdots < i_k \leq n$ denote by $\psi_{x_{i_1}, \ldots, x_{i_k}}$ the function ψ with the i_1-th, $\ldots, i_k-$th variables held fixed and taking the values $x_{i_1}, \ldots, x_{i_k}$. Further, for a copula C, marginals $F_1, \ldots, F_n$ and an increasing function $\psi : \mathbb{R}^n \to \mathbb{R}$, define

$$\tau_{C,\psi}(F_1, \ldots, F_n)(s) := \sup_{x_1, \ldots, x_{n-1} \in \mathbb{R}} C(F_1(x_1), \ldots, F_{n-1}(x_{n-1}), F_n(\psi^\wedge_{x_1, \ldots, x_{n-1}}(s))),$$

$$\sigma_{C,\psi}(F_1, \ldots, F_n)(s) := \int_{\{\psi \leq s\}} dC(F_1(x_1), \ldots, F_n(x_n)),$$

$$\rho_{C,\psi}(F_1, \ldots, F_n)(s) := \inf_{x_1, \ldots, x_{n-1} \in \mathbb{R}} C^d(F_1(x_1), \ldots, F_{n-1}(x_{n-1}), F_n(\psi^\wedge_{x_1, \ldots, x_{n-1}}(s))).$$

Observe that, if $(X_1, \ldots, X_n)$ has copula C and marginals $F_1, \ldots, F_n$, then $\sigma_{C,\psi}(F_1, \ldots, F_n) = F_{\psi(X_1, \ldots, X_n)}$. Further, it is easily seen that $\tau_{C,\psi}(F_1, \ldots, F_n)$ and $\rho_{C,\psi}(F_1, \ldots, F_n)$ are themselves distribution functions. Thus $\tau_{C,\psi}, \sigma_{C,\psi}$ and $\rho_{C,\psi}$ can be viewed as operations mapping Δ^n to Δ, where Δ is the set of one-dimensional distribution functions.

Theorem 3.1. *Let $(X_1, \dots, X_n)$ have marginal distribution functions $F_1, \dots, F_n$ and let $\psi : \mathbb{R}^n \to \mathbb{R}$ be increasing and left continuous in the last argument. If a copula C for $(X_1, \dots, X_n)$ satisfies $C \geq C_0$ and $C^d \leq C_1^d$ for some given copulae C_0 and C_1, then*

$$\tau_{C_0,\psi}(F_1, \dots, F_n) \leq \sigma_{C,\psi}(F_1, \dots, F_n) \leq \rho_{C_1,\psi}(F_1, \dots, F_n).$$

Translated in the language of Value-at-Risk, the statement of Theorem 3.1 becomes

$$\rho_{C_1,\psi}(F_1, \dots, F_n)^{-1}(\alpha) \leq \mathrm{VaR}_\alpha(\psi(X_1, \dots, X_n)) \leq \tau_{C_0,\psi}(F_1, \dots, F_n)^{-1}(\alpha).$$

The next result states that the bounds in Theorem 3.1 are pointwise best-possible. More precisely, there are dependence structures for $(X_1, \dots, X_n)$ such that the distribution function of $\psi(X_1, \dots, X_n)$ attains the bounds at least at one point.

Theorem 3.2. *Let the hypotheses of Theorem 3.1 be satisfied and for fixed $s \in \mathbb{R}$ consider $\alpha := \tau_{C_0,\psi}(F_1, \dots, F_n)(s) \leq \rho_{C_1,\psi}(F_1, \dots, F_n)(s) =: \beta$. Then there are copulae C^α and C^β such that*

$$\sigma_{C^\alpha,\psi}(F_1, \dots, F_n)(s) = \alpha \quad \textit{and} \quad \sigma_{C^\beta,\psi}(F_1, \dots, F_n)(s) = \beta.$$

Some remarks relative to the nature of the bounds of Theorems 3.1 and 3.2 and a detailed discussion of the assumptions $C \geq C_0$ and $C^d \leq C_1^d$ are given below.

3.2 Nature of the bounds

As stated above, the bounds $\tau_{C_0,\psi}(F_1, \dots, F_n)$ and $\rho_{C_1,\psi}(F_1, \dots, F_n)$ are themselves distribution functions. Nevertheless, it is not true in general that they are distribution functions of some random variable $\psi(Y_1, \dots, Y_n)$ such that Y_i is distributed as F_i, i.e. there is not necessarily a portfolio $\psi(Y_1, \dots, Y_n)$ such that $\tau_{C_0,\psi}(F_1, \dots, F_n) = F_{\psi(Y_1,\dots,Y_n)}$ or $\rho_{C_1,\psi}(F_1, \dots, F_n) = F_{\psi(Y_1,\dots,Y_n)}$. In fact, already in the two-dimensional case, we have from Theorem 1 in Schweizer and Sklar (1974) (see also Frank et al. (1987) p. 208 and Denuit et al. (1999) p. 87) that for $\psi(x_1, x_2) = x_1 + x_2$ and $C_0, C_1 \neq C_\mathrm{U}$ there is no Borel-measurable function f such that $\tau_{C_0,\psi}(F_1, F_2) = F_{f(X_1,X_2)}$ or $\rho_{C_1,\psi}(F_1, F_2) = F_{f(X_1,X_2)}$. In other words, the operations $\tau_{C_0,\psi}$ and $\rho_{C_0,\psi}$ cannot be derived form any binary operation on the random variables X_1, X_2.

3.3 The conditions $C \geq C_0$ and $C^d \leq C_1^d$

We have already emphasized that in general the copula C of $(X_1, \dots, X_n)$ is unknown. It could even be that C is non-unique. For an application of Theorems 3.1 and 3.2 the problem therefore reduces to verifying the assumptions $C \geq C_0$ and $C^d \leq C_1^d$. In practice these inequalities need not to be checked since C_0 and C_1 are chosen (engineered) depending on the situation considered and $C \geq C_0, C^d \leq C_1^d$ are assumed to hold for these choices. Indeed,

$C \geq C_0$ and $C^d \leq C_1^d$ represent the partial information available about the dependence between $X_1, \ldots, X_n$ or equivalently they define different dependence scenarios. For simplicity, consider first the two-dimensional case, where $C \geq C_0$ is equivalent with $C^d \leq C_0^d$, i.e. where we can take $C_1 = C_0$. Since C_{L} is the lower Fréchet bound, we have that any copula C satisfies $C \geq C_0 = C_{\mathrm{L}}$ and $C^d(u_1, u_2) \leq C_1^d(u_1, u_2) = C_{\mathrm{L}}^d(u_1, u_2) = (u_1 + u_2) \wedge 1$. The interpretation behind the bounds obtained with $C_0 = C_1 = C_{\mathrm{L}}$ is that no information about the dependence between X_1, X_2 is available or alternatively that no dependence restrictions are made.

Although for $n \geq 3$ the lower Fréchet bound C_{L} is not a copula, we have that Theorem 3.1 holds true even for C_0 and C_1^d replaced by C_{L} and $1 \wedge \sum_{i=1}^n u_i$ respectively. In fact, for the proof of Theorem 3.1 (see Appendix) only the property that C_0 and C_1^d are increasing in each argument is used. Moreover, any copula C satisfies $C^d(u_1, \ldots, u_n) \leq 1 \wedge \sum_{i=1}^n u_i$. The interpretation of the corresponding bounds is the same as above.

Alternatively, the choice $C_0 = C_1 = C_{\mathrm{I}}$ means that the risks are POD. In view of Section 2.3, this kind of partial information is a reasonable assumption for many scenarios where positive dependence has to be modelled.

In general, $C \geq C_0$ can be rewritten as $C \geq_{\mathrm{lo}} C_0$. Because of (2.4), we get that $C^d \leq C_1^d$ is equivalent to $\hat{C} \geq \hat{C}_1$, i.e. to $\hat{C} \geq_{\mathrm{lo}} \hat{C}_1$. Because of (2.9), these conditions have natural interpretation in terms of correlations. However, a choice like for instance $C_0 = C^{\mathrm{Cl},\alpha}$ (or $\hat{C}_1 = C^{\mathrm{Gu},\beta}$) is typically justified if we are in the presence of conditional bivariate extremes (componentwise maxima) as mentioned in Section 2.3.

From the definitions of $\tau_{C_0,\psi}$ and $\rho_{C_1,\psi}$ it is clear that the bounds obtained when sharpening the conditions $C \geq C_0$ and $C^d \leq C_1^d$ (or $\hat{C} \geq \hat{C}_1$) become tighter. This corresponds to the intuitive fact that the more information relative to the dependence structure of $X_1, \ldots, X_n$ is available the smaller becomes the range in which VaR can lie.

3.4 Extensions

In contrast to the existing literature on results like Theorems 3.1 and 3.2, we only require that ψ is increasing in each argument and left continuous in the last one. This allows for a more flexible application of the results in practice since especially in the financial derivatives world, discontinuous functions abound.

Theorems 3.1 and 3.2 can be modified for functions ψ which are increasing in some of their arguments and decreasing in the others. Further, although in the whole paper the main emphasis is put on Value-at-Risk, the above theorems can easily be adapted for risk measures different from VaR.

3.5 Some special cases: comonotonicity and independence

Comonotonicity and independence are particularly interesting cases. One reason for this is that in both cases VaR can be computed exactly and fairly easily. Further, comonotonicity represents an extreme dependence situation, whereas

independence often leads to a boundary case against which several dependence scenarios can be calibrated. In other words, formulae (3.1) and (3.3) below will be used in the example of Section 5 as a term of comparison for the bounds obtained.

Though a special case of the next result is well known and indeed easy to prove, we thought it important to state it in full detail. For risk management applications, it tells us when VaR calculations can be transported painlessly through functions of different risk types.

Proposition 3.1. *Let $\psi : \mathbb{R}^n \to \mathbb{R}$ be increasing and left continuous in each argument, $0 \leq \alpha \leq 1$ and $X_1, \ldots, X_n$ be comonotonic. Then*

$$\mathrm{VaR}_\alpha(\psi(X_1, \ldots, X_n)) = \psi(\mathrm{VaR}_\alpha(X_1), \ldots, \mathrm{VaR}_\alpha(X_n)), \tag{3.1}$$

provided that the above expressions are defined.

Proof. Let Z be a real valued random variable with range $\mathrm{Im}(Z)$ and suppose that $\varphi : \mathrm{Im}(Z) \subset \mathbb{R} \to \mathbb{R}$ is increasing and left continuous. Suppose that $\mathrm{VaR}_\alpha(Z)$ is finite for a given $\alpha \in [0,1]$. Using Lemma 2.1 we have that the distribution function of $\varphi(Z)$ is given by $F_{\varphi(Z)}(t) = P[\varphi(Z) \leq t] = P[Z \leq \varphi^\wedge(t)] = F_Z(\varphi^\wedge(t))$. Hence, again because of Lemma 2.1, we obtain

$$\begin{aligned} \mathrm{VaR}_\alpha(\varphi(Z)) &= \inf\{t \in \mathbb{R} \,|\, F_{\varphi(Z)}(t) \geq \alpha\} = \inf\{t \in \mathbb{R} \,|\, F_Z(\varphi^\wedge(t)) \geq \alpha\} \\ &= \inf\{t \in \mathbb{R} \,|\, \varphi^\wedge(t) \geq F_Z^{-1}(\alpha)\} \\ &= \inf\{t \in \mathbb{R} \,|\, t \geq \varphi(F_Z^{-1}(\alpha))\} = \varphi(F_Z^{-1}(\alpha)) = \varphi(\mathrm{VaR}_\alpha(Z)). \end{aligned} \tag{3.2}$$

Denote by $F_1, \ldots, F_n$ the distribution functions of the comonotonic random variables $X_1, \ldots, X_n$ and consider the increasing left continuous function $\varphi(\alpha) := \psi(F_1^{-1}(\alpha), \ldots, F_n^{-1}(\alpha))$. For an arbitrary, on $[0,1]$ uniformly distributed random variable U, we get from (3.2) that

$$\begin{aligned} \mathrm{VaR}_\alpha(\psi(X_1, \ldots, X_n)) &= \mathrm{VaR}_\alpha(\psi(F_1^{-1}(U), \ldots, F_n^{-1}(U))) = \mathrm{VaR}_\alpha(\varphi(U)) \\ &= \varphi(\mathrm{VaR}_\alpha(U)) = \varphi(\alpha) = \psi(F_1^{-1}(\alpha), \ldots, F_n^{-1}(\alpha)) \\ &= \psi(\mathrm{VaR}_\alpha(X_1), \ldots, \mathrm{VaR}_\alpha(X_n)). \end{aligned}$$

□

For independent random variables $X_1, \ldots, X_n$ the distribution function of $\psi(X_1, \ldots, X_n)$, and hence its VaR, can be computed by iterated conditioning. Since $X_1, \ldots, X_n$ are independent, we have that

$$\begin{aligned} F_{\psi(X_1,\ldots,X_n)}(s) &= P[\psi(X_1, \ldots, X_n) \leq s] \\ &= \int dF_n(t_n) P[\psi_{t_n}(X_1, \ldots, X_{n-1}) \leq s] \\ &= \int dF_n(t_n) \cdots \int dF_2(t_2) F_1(\psi^\wedge_{t_2,\ldots,t_n}(s)). \end{aligned} \tag{3.3}$$

When $n = 2$ this formula simplifies to $F_{\psi(X_1,X_2)}(s) = \int F_1(\psi^\wedge_{t_2}(s))\, dF_2(t_2)$.

4 Computational aspects

For notational convenience, we will write in the sequel $F_{\min}$ and $F_{\max}$ instead of $\tau_{C_0,\psi}(F_1,\ldots,F_n)$ and $\rho_{C_1,\psi}(F_1,\ldots,F_n)$ respectively. In most cases, the bounds $F_{\min}$ and $F_{\max}$ do not allow for a closed form expression. Exceptions are the cases where $\psi(x_1,\ldots,x_n) = \sum_{i=1}^n x_i$ and the F_i's all are of the same type, for example all are shifted exponential, shifted Pareto, Weibull or uniform (see Denuit et al. (1999)). In general, one has to resort to numerical approximations as proposed by Williamson and Downs (1990).

In the sequel, we denote again by C_0 a copula or the lower Fréchet bound C_{L}. Similarly, C_1^d can represent the dual of a copula C_1 as well as $1 \wedge \sum_{i=1}^n u_i$. Further, for $(U_1,\ldots,U_n)$ with distribution function C and indices $1 \le i_1 < \cdots < i_k \le n$ let $C_{i_1,\ldots,i_k}$ be the distribution function of $(U_{i_1},\ldots,U_{i_k})$ and $C^d_{i_1,\ldots,i_k}$ its dual, where for $C = C_{\mathrm{L}}$ and $C_1^d(u_1,\ldots,u_n) = 1 \wedge \sum_{i=1}^n u_i$ we obviously mean $(C_{\mathrm{L}})_{i_1,\ldots,i_k}(v_1,\ldots,v_k) = (\sum_{j=1}^k v_j - k + 1)^+$ and $(C_{\mathrm{L}})^d)_{i_1,\ldots,i_k}(v_1,\ldots,v_k) = 1 \wedge \sum_{j=1}^k v_j$.

4.1 Duality

The main problem relative to a practical computation of the expressions $F_{\min}(s)$ and $F_{\max}(s)$, thus of their inverses, is that they are defined as a supremum, respectively as an infimum, over the unbounded set $\mathbb{R}^{n-1}$. Hence, even the computation over a discretization of $\mathbb{R}^{n-1}$, amounts to an infinite number of comparisons for finding the maximum (minimum) of the relative expressions. This problem can be avoided applying the duality principle of Frank and Schweizer (1979). The special case useful for our purposes takes the following form.

Theorem 4.1 (Duality). *Let $-\infty \le a < b \le +\infty$ and $\psi : [a,b]^n \to [a,b]$ be an increasing continuous function with range $[a,b]$. For a copula C_0, marginals $F_1,\ldots,F_n$ and any $0 \le \alpha < 1$ one obtains*

$$F^{-1}_{\min}(\alpha) = \inf_{C_0(u_1,\ldots,u_n)=\alpha} \psi(F_1^{-1}(u_1),\ldots,F_n^{-1}(u_n)), \tag{4.1}$$

$$F^{-1}_{\max}(\alpha) = \sup_{C_1^d(u_1,\ldots,u_n)=\alpha} \psi(F_1^{-1}(u_1),\ldots,F_n^{-1}(u_n)). \tag{4.2}$$

Observe that, contrary to Theorems 3.1 and 3.2, the functional ψ in Theorem 4.1 has to satisfy the additional condition that its domain is $[a,b]^n$ and that its range is $[a,b]$. These constraints are satisfied for many functionals of interest and in particular in the example of Section 5.

The origin of the name "Duality Theorem" for the above result is due to the fact that (4.1) and (4.2) can be obtained from

$$\sup_{\psi(t_1,\ldots,t_n)=s} C_0(F_1(t_1),\ldots,F_n(t_n)), \tag{4.3}$$

$$\inf_{\psi(t_1,\ldots,t_n)=s} C_1^d(F_1(t_1),\ldots,F_n(t_n)), \tag{4.4}$$

replacing infima by suprema, the F_i's by the F_i^{-1}'s and exchanging the role of ψ and C_0 (or C_1). Moreover, formulae (4.3) and (4.4) look very similar to $F_{\min}(s)$ and $F_{\max}(s)$. In fact, for fixed s such that $\{(t_1,\dots,t_n)\in\mathbb{R}^n \,|\, \psi(t_1,\dots,t_n) = s\} \neq \emptyset$, the condition $\psi_{t_n}(t_1,\dots,t_{n-1}) = \psi(t_1,\dots,t_n) = s$ is equivalent with $\psi^{\wedge}_{t_1,\dots,t_{n-1}}(s-) \leq t_n \leq \psi^{\wedge}_{t_1,\dots,t_{n-1}}(s)$. Since $C_0(F_1(t_1),\dots,F_n(t_n))$ is increasing in $t_1,\dots,t_n$, it follows that its supremum over $\{(t_1,\dots,t_n)\in\mathbb{R}^n \,|\, \psi(t_1,\dots,t_n) = s\}$ is taken for t_n at the right end point of $[\psi^{\wedge}_{t_1,\dots,t_{n-1}}(s-), \psi^{\wedge}_{t_1,\dots,t_{n-1}}(s)]$. Hence,

$$\begin{aligned} F_{\min}(s) &= \sup_{t_1,\dots,t_{n-1}\in\mathbb{R}} C_0(F_1(t_1),\dots,F_{n-1}(t_{n-1}), F_n(\psi^{\wedge}_{t_1,\dots,t_{n-1}}(s))) \\ &= \sup_{\psi(t_1,\dots,t_n)=s} C_0(F_1(t_1),\dots,F_n(t_n)), \end{aligned}$$

which is an alternative representation of $F_{\min}(s)$ for all s such that $\{\psi = s\} \neq \emptyset$.

In general, the same cannot be repeated for $F_{\max}$, meaning that there are situations where (4.4) is at some points strict smaller than $F_{\max}$ and where (4.4) does not even yield an upper bound for $F_{\psi(X_1,\dots,X_n)}$ as the counterexample in the Appendix shows. However, both (4.3) and (4.4) can be used in order to obtain the quantile functions $F^{-1}_{\min}$ and $F^{-1}_{\max}$, i.e. bounds for $\mathrm{VaR}_\alpha(\psi(X_1,\dots,X_n))$. In effect, (4.4) is the left continuous version of $F_{\max}$ and hence it leads to the same quantiles. More precisely, considering left continuously defined distribution functions instead of right continuous ones, a representation like (4.4) holds even for $F_{\max}$.

Remark. In some of the classical literature on the subject left continuous versions of distribution functions are taken. Hence, when using results from those papers, care has to be taken.

4.2 Practical computation of $F^{-1}_{\min}(\alpha)$ and $F^{-1}_{\max}(\alpha)$

The advantage of (4.1) and (4.2) is that the infimum and the supremum are taken over the compact sets $\{C_0 = \alpha\}$ and $\{C_1^d = \alpha\}$ respectively. The practical idea, summarized by Figure 4.1 below, is to consider a suitable discretization of these sets and to minimize (maximize) over such a finite set of points.

For notational convenience, let $\alpha = r/N$, where N is some fixed integer and $r \in \{1,\dots,N-1\}$. Consider further, $l_1,\dots,l_{n-1} \in \{0,\dots,N\}$ and look for a solution $\nu_{r,l_1,\dots,l_{n-1}}$ to

$$C_0\left(l_1/N,\dots,l_{n-1}/N,\nu_{r,l_1\dots,l_{n-1}}\right) = r/N. \tag{4.5}$$

Since $C_0(l_1/N,\dots,l_{n-1}/N,\cdot)$ is a continuous function mapping the unit interval $[0,1]$ to $[0,(C_0)_{1,\dots,n-1}(l_1/N,\dots,l_{n-1}/N)]$, we have that a solution $\nu_{r,l_1,\dots,l_{n-1}}$ always exists if

$$(C_0)_{1,\dots,n-1}(l_1/N,\dots,l_{n-1}/N) \geq r/N.$$

Thus, because of (4.1) we have that

$$q_{\min}(r/N) := \min_{A_{r,l_1,\dots,l_{n-1}}} \psi(F_1^{-1}(l_1/N),\dots,F_{n-1}^{-1}(l_{n-1}/N), F_n^{-1}(\nu_{r,l_1\dots,l_{n-1}}))$$

is an approximation of $F_{\min}^{-1}(r/N)$, where the minimum is taken over the set $A_{r,l_1,\dots,l_{n-1}} := \{l_1,\dots,l_{n-1} \,|\, (C_0)_{1,\dots,n-1}(l_1/N,\dots,l_{n-1}/N) \geq r/N\}$. Similarly,

$$q_{\max}(r/N) := \max_{B_{r,l_1,\dots,l_{n-1}}} \psi(F_1^{-1}(l_1/N),\dots,F_{n-1}^{-1}(l_{n-1}/N),F_n^{-1}(\nu^*_{r,l_1\dots,l_{n-1}}))$$

is nearly $F_{\max}^{-1}(r/N)$, where $\nu^*_{r,l_1,\dots,l_{n-1}}$ is a solution to the equation

$$C_1^d\left(l_1/N,\dots,l_{n-1}/N,\nu^*_{r,l_1\dots,l_{n-1}}\right) = r/N \tag{4.6}$$

and $B_{r,l_1,\dots,l_{n-1}} := \{l_1,\dots,l_{n-1} \,|\, (C_1)^d_{1,\dots,n-1}(l_1/N,\dots,l_{n-1}/N) \leq r/N\}$.

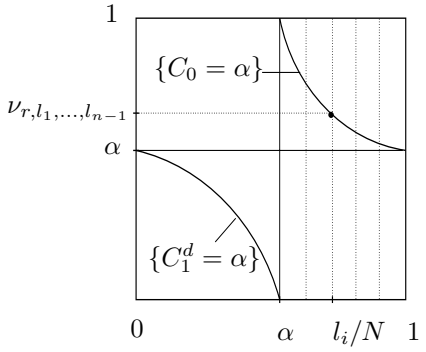

Figure 4.1: Discretization of $\{C_0 = \alpha\}$ and $\{C_1^d = \alpha\}$.

Remarks.

(a) In the two-dimensional case the above formulae for $q_{\min}$, $q_{\max}$ simplify to

$$\begin{aligned} q_{\min}(r/N) &= \min_{r\leq l\leq N} \psi(F_1^{-1}(l/N),F_2^{-1}(\nu_{r,l})), \\ q_{\max}(r/N) &= \max_{0\leq l\leq r} \psi(F_1^{-1}(l/N),F_2^{-1}(\nu^*_{r,l})), \end{aligned} \tag{4.7}$$

which can be easily implemented.

(b) The quantities $q_{\min}(r/N)$ and $q_{\max}(r/N)$ given by (4.7) are approximate bounds for $\mathrm{VaR}_{r/N}(\psi(X_1,\dots,X_n))$. More generally, one could be interested in approximating the whole distribution functions $F_{\min}$ and $F_{\max}$ as for instance in the case of risk measures different from VaR. This can be done approximating $F_{\min}$ and $F_{\max}$ from below and from above using the step functions

$$F_{\min,N} := \frac{1}{N}\sum_{r=1}^{N} 1_{[q_{\min}(r/N),\infty)}, \quad F_{\max,N} := \frac{1}{N}\sum_{r=0}^{N-1} 1_{[q_{\max}(r/N),\infty)}, \tag{4.8}$$

where $q_{\min}(1) := \sup \mathrm{supp}(F)$, $q_{\max}(0) := \inf \mathrm{supp}(F)$.

4.3 Archimedean scenarios

In many cases, $\nu_{r,l_1,\dots,l_{n-1}}$ can be found explicitly. For instance, if C_0 belongs to the wide class of strict archimedean copulae, i.e. if it is of the form

$$C_0(u_1,\dots,u_n) = \phi_0^{-1}\left(\sum_{i=1}^{n}\phi_0(u_i)\right), \tag{4.9}$$

where $\phi_0 : [0,1] \to [0,\infty]$ is a continuous strictly decreasing function with $\phi_0(0) = \infty$, $\phi_0(1) = 0$ and such that the inverse ϕ_0^{-1} is completely monotonic, then

$$\nu_{r,l_1,\dots,l_{n-1}} = \phi_0^{-1}\left(\phi_0(r/N) - \sum_{i=1}^{n-1}\phi_0(l_i/N)\right) \tag{4.10}$$

provided that $l_1,\dots,l_{n-1}$ satisfy $\phi_0(r/N) \geq \sum_{i=1}^{n-1}\phi_0(l_i/N)$. This is the case for the Clayton copula defined in (2.10) generated by $\phi_0(t) = t^{-\alpha} - 1$ and for the independent copula C_{I} generated by $\phi_0(t) = -\log t$.

On the contrary, finding $\nu^*_{r,l_1,\dots,l_{n-1}}$ may be not so easy. For instance, for a two-dimensional archimedean copula C_1 generated by some ϕ_1, we have that $\phi_1(\nu^*_{r,l}) + \phi_1(l/N) - \phi_1(\nu^*_{r,l} - (r-l)/N) = 0$ which cannot be solved explicitly in general. This problem can be avoided making assumptions on the survival copula $\hat{C}_1$. Indeed, assuming for instance that $\hat{C}_1$ is strict archimedean with generator ϕ_1, it follows by (2.4) that $\nu^*_{r,l_1,\dots,l_{n-1}}$ is a solution to $r/N = 1 - \hat{C}_1(1-l_1/N,\dots,1-l_{n-1}/N, 1-\nu^*_{r,l_1,\dots,l_{n-1}})$, i.e. that

$$\nu^*_{r,l_1,\dots,l_{n-1}} = 1 - \phi_1^{-1}\left(\phi_1(1-r/N) - \sum_{i=1}^{n-1}\phi_1(1-l_i/N)\right) \tag{4.11}$$

provided that $l_1,\dots,l_{n-1}$ satisfy $\phi_1(1-r/N) \geq \sum_{i=1}^{n-1}\phi_1(1-l_i/N)$.

Remark. The lower Fréchet bound C_{L} has also a representation like (4.9) for $\phi_0(t) = 1 - t$ and (4.10) corresponds to the solution $\nu_{r,l_1,\dots,l_{n-1}}$ of (4.5) and the relative boundary conditions for $l_1,\dots,l_{n-1}$ coincide. Similarly, the solution $\nu^*_{r,l_1,\dots,l_{n-1}}$ of (4.6) and the relative boundary conditions can be obtained from (4.11) for $\phi_1(t) = 1 - t$.

4.4 Alternative computational methods

The gridding procedure of Section 4.2 has the disadvantage that the computational complexity increases exponentially in the number of dimensions. Thus, an immediate generalization of the above procedure for large n translates into a massive increase of computational time. Further, the question relative to the convergence of $q_{\min}(\alpha), q_{\max}(\alpha)$ to $F^{-1}_{\min}(\alpha), F^{-1}_{\max}(\alpha)$ as the number N of discretization steps increases was not considered in the previous sections. Nevertheless, for many examples of interest, as the one discussed in Section 5 below, the above numerical procedure works well. An alternative numerical approach

that can be used for several applications of our results is based on the theory of *Semidefinite Programming* (SDP), where by means of *interior point methods* extremely efficient results are obtained for minimax problems having similarities with the computation of our bounds. The convergence properties of these methods have been studied in detail and one can in general expect that the increase of the numerical complexity is polynomial. We checked this approach for the example discussed in Section 5; the results obtained coincide. Further work along these lines is definitely necessary, especially as one may want to obtain numerical procedures for functions of high dimensional portfolios ($n \geq 100$, say). We plan to return to this in future work. For the relevant (linear, convex, SDP) techniques, see for instance Boyd and Vandenberghe (1999) and the references therein.

5 Example

The methodology of Section 4 can be applied to different concrete situations. Figure 5.1 below, gives the range for $\mathrm{VaR}_\alpha(X_1+X_2)$ in the case of standard normal marginals under various dependence scenarios. We have chosen to illustrate our results in the case of normal marginals in order to stress that the difficulties obtained, i.e. the non-subadditive of the risk measure is not stemming from the marginal distributions (P&L's) but from the dependence structure imposed. We also took the liberty to use the VaR language throughout the example, even if some of the obtained values become negative. This should not deter from the main message. We also have tested the results obtained on several other portfolios and alternative functions ψ, and also cross-checked the numerics with those obtained via convex programming techniques; these results are not reproduced here.

In the language of Theorem 3.1, the scenarios are:

(Sc1) $C_0 = C_1 = C_L$: no dependence restriction,

(Sc2) $C_0 = C_1 = C_I$: POD dependence,

(Sc3) $C_0 = C^{\mathrm{Cl},8}$, $\hat{C}_1 = C^{\mathrm{Gu},0.2}$,

where the choices of the parameter values $\alpha = 8$ and $\beta = 0.2$ correspond both to a Kendall's tau of 0.8. Also included are the independent and comonotonic cases. Figure 5.1 below was obtained computing $q_{\min}(r/N)$ and $q_{\max}(r/N)$ by means of (4.7) for $N = 1000$ points using `S-Plus` within seconds.

When dealing with two-dimensional portfolios as in this examples, a typical dependence scenario consists of PQD risks as in (Sc2). In this case the approximated lower bound $F_{\min}$ in Theorem 3.1 calculated by means of (4.8) with $C_0 = C_1 = C_{\mathrm{I}}$ (partial information) and the one computed with $C_0 = C_1 = C_{\mathrm{L}}$ (no information) do not look very much different in the tails as Figure 5.1 shows. A careful look at the numerical procedure used to calculate the quantiles gives more insight into this phenomenon which can also be observed for many different choices of ψ, F_1, F_2. The quantiles $q_{\min}(r/N)$ and $q_{\min}^{\mathrm{PQD}}(r/N)$ when no

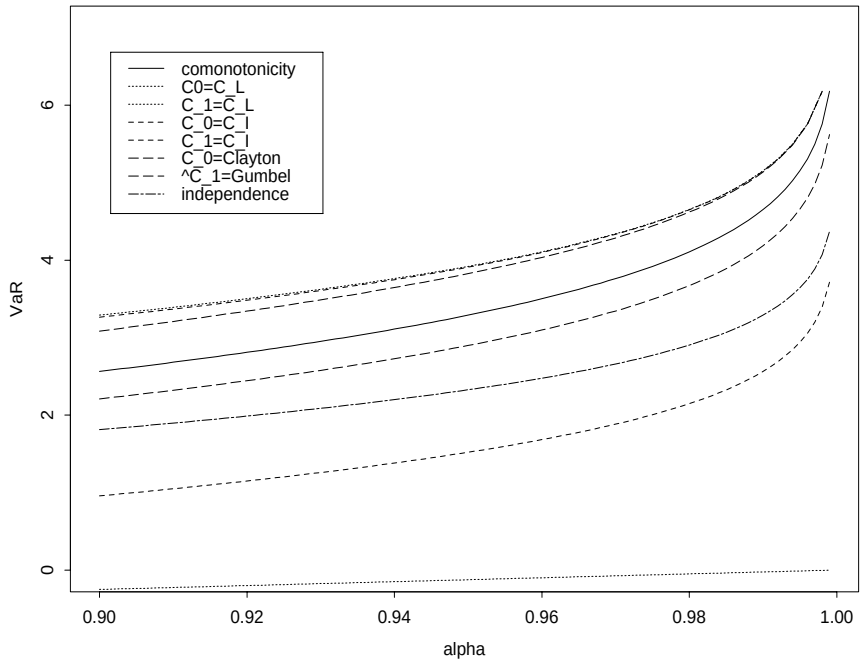

Figure 5.1: Range for $\mathrm{VaR}_\alpha(X_1 + X_2)$ for a standard normal portfolio under scenarios (Sc1), respectively (Sc2) and (Sc3).

information is available or a PQD assumption is made have the form

$$q_{\min}(r/N) = \min_{r \le l \le N} \psi\left(F_1^{-1}(l/N), F_2^{-1}\left(\frac{N-(l-r)}{N}\right)\right),$$
$$q_{\min}^{\mathrm{PQD}}(r/N) = \min_{r \le l \le N} \psi(F_1^{-1}(l/N), F_2^{-1}(r/l)).$$

For fixed r we have to take the minimum of

$$\begin{array}{ll} l = r & \psi(F_1^{-1}(r/N), F_2^{-1}(1)) \\ l = r+1 & \psi(F_1^{-1}((r+1)/N), F_2^{-1}((N-1)/N)) \\ l = r+2 & \psi(F_1^{-1}((r+2)/N), F_2^{-1}((N-2)/N)) \\ \dots & \dots \\ l = N & \psi(F_1^{-1}(1), F_2^{-1}(r/N)) \end{array}$$

in the first case and of

$$\begin{array}{ll} l = r & \psi(F_1^{-1}(r/N), F_2^{-1}(1)) \\ l = r+1 & \psi(F_1^{-1}((r+1)/N), F_2^{-1}(r/(r+1))) \\ l = r+2 & \psi(F_1^{-1}((r+2)/N), F_2^{-1}(r/(r+2))) \\ \dots & \dots \\ l = N & \psi(F_1^{-1}(1), F_2^{-1}(r/N)) \end{array}$$

in the second one. We observe that the quantiles of F_1 are the same in both cases whereas the quantiles of F_2 are evaluated at the points

$$1 \geq \frac{N-(l-r)}{N} \geq \frac{r}{N} = \alpha \quad \text{and} \quad 1 \geq \frac{r}{l} \geq \frac{r}{N} = \alpha,$$

respectively. If α is near to 1, then $(N-(l-r))/N \approx l/r$ for $N \geq l \geq r = N\alpha$. This means that the quantiles in the second argument of ψ are almost the same. Indeed, take for example $N = 100$ and $\alpha = 0.9$, then

	l=r+1	l=r+2	l=r+3	...	l=N-3	l=N-2	l=N-1	l=N
(N-(l-r))/N	0.99	0.98	0.97	...	0.93	0.92	0.91	0.90
r/l	0.989	0.978	0.968	...	0.928	0.918	0.909	0.90

The only way to see a difference would be to take F_2 heavy tailed and ψ strongly increasing in the second component at the same time.

To have an idea of the range of values obtained for a particular α, we have summarized this information in Table 5.1 for $\alpha = 0.95$ and $\alpha = 0.99$. The quantiles for the marginals are $\text{VaR}_{0.95}(X_i) = 1.96$, respectively $\text{VaR}_{0.99}(X_i) = 2.33$, $i = 1, 2$. Results as summarized in Figure 5.1 and Table 5.1 yield a measure of uncertainty for VaR calculations at the aggregate level $X_1 + X_2$ given VaR information at the sub-portfolio level X_1 and X_2. For instance, Table 5.1 tells

	$\alpha = 0.95$			$\alpha = 0.99$		
scenarios	exact	min	max	exact	min	max
1		-0.13	3.92		-0.03	5.15
2		1.52	3.91		2.56	5.15
3		2.90	3.83		4.19	5.14
$C = C_I$	2.33			3.29		
$C = C_U$	3.29			4.65		

Table 5.1: Range for $\text{VaR}_{0.95}(X_1 + X_2)$ and $\text{VaR}_{0.99}(X_1 + X_2)$ for a standard normal portfolio

us in the $\alpha = 0.95$ case that, whereas the marginal VaRs are 1.96 Mio. \$, say, for each position separately, the VaR for the joint portfolio $X_1 + X_2$ can reach a value between -0.13 and 3.92 Mio. \$. If one is prepared to accept extra dependence information as for instance PQD, then the bounds narrow to 1.52 and 3.91 Mio. \$. Only when much stronger assumptions can be made, in our situation $\hat{C} \geq C^{\text{Gu},0.2}$ and $C \geq C^{\text{Cl},8}$, then we have a substantial reduction for the interval size to $[2.90, 3.83]$ Mio. \$.

6 Conclusions

In this paper we focus on the construction of optimal bounds for risk measures of functions of dependent risks. The techniques introduced are exemplified in the case of Value-at-Risk. The methodology used is based on the theory of copulae. A numerical procedure for calculating these bounds in specific cases is provided.

Acknowledgments

The authors would like to thank two anonymous referees for helpful comments on an earlier version of the paper. They also thank Pablo Parrilo (ETHZ) and Stephen Boyd (Stanford) for several discussions on the use of convex programming in similar problems, and Francesco Audrino (ETHZ) for discussions on the numerics.

Appendix

Proof of Theorem 3.1

Proof. Let s be fixed and $t_1, \ldots, t_{n-1} \in \mathbb{R}$ such that $\psi^{\wedge}_{t_1,\ldots,t_{n-1}}(s)$ is finite. By Lemma 2.1, we have that $X_i > t_i$ for $i = 1, \ldots, n-1$ and $X_n > \psi^{\wedge}_{t_1,\ldots,t_{n-1}}(s)$ together imply that $\psi(X_1, \ldots, X_n) \geq \psi_{t_1,\ldots,t_{n-1}}(X_n) > s$. Hence,

$$\begin{aligned} P[\psi(X_1, \ldots, X_n) \leq s] &\leq P[\cup_{i=1}^{n-1}\{X_i \leq t_i\} \cup \{X_n \leq \psi^{\wedge}_{t_1,\ldots,t_{n-1}}(s)\}] \\ &= C^d(F_1(t_1), \ldots, F_{n-1}(t_{n-1}), F_n(\psi^{\wedge}_{t_1,\ldots,t_{n-1}}(s))) \\ &\leq C_1^d(F_1(t_1), \ldots, F_{n-1}(t_{n-1}), F_n(\psi^{\wedge}_{t_1,\ldots,t_{n-1}}(s))). \end{aligned}$$

If $\psi^{\wedge}_{t_1,\ldots,t_{n-1}}(s) = +\infty$, then we have that

$$C_1^d(F_1(t_1), \ldots, F_{n-1}(t_{n-1}), F_n(\psi^{\wedge}_{t_1,\ldots,t_{n-1}}(s))) = 1$$

which is greater or equal than $P[\psi(X_1, \ldots, X_n) \leq s]$. On the other hand, if $\psi^{\wedge}_{t_1,\ldots,t_{n-1}}(s) = -\infty$, then $\psi_{t_1,\ldots,t_{n-1}}(t_n) > s$ for all $t_n \in \mathbb{R}$ and therefore

$$\begin{aligned} P[\psi(X_1, \ldots, X_n) \leq s] &= P\left[\{\psi(X_1, \ldots, X_n) \leq s\} \cap \cup_{i=1}^{n-1}\{X_i \leq t_i\}\right] \\ &\quad + P[\{\psi(X_1, \ldots, X_n) \leq s\} \cap \cap_{i=1}^{n-1}\{X_i > t_i\}] \\ &\leq P[\cup_{i=1}^{n-1}\{X_i \leq t_i\}] + 0 \\ &= C^d(F_1(t_1), \ldots, F_{n-1}(t_{n-1}), 0) \\ &= C_1^d(F_1(t_1), \ldots, F_{n-1}(t_{n-1}), F_n(\psi^{\wedge}_{t_1,\ldots,t_{n-1}}(s))). \end{aligned}$$

Taking the infimum over all $t_1, \ldots, t_{n-1} \in \mathbb{R}$ leads to $F_{\psi(X_1,\ldots,X_n)}(s) \leq F_{\max}(s)$. Similarly, if $\psi^{\wedge}_{t_1,\ldots,t_{n-1}}(s)$ is finite, then $X_i \leq t_i$ for $i = 1, \ldots, n-1$ and

$X_n \leq \psi^{\wedge}_{t_1,\ldots,t_{n-1}}(s)$ together imply that $\psi(X_1,\ldots,X_n) \leq \psi_{t_1,\ldots,t_{n-1}}(X_n) \leq s$. Therefore,

$$\begin{aligned}
&P[\psi(X_1,\ldots,X_n) \leq s] = 1 - P[\psi(X_1,\ldots,X_n) > s] \\
&= 1 - P[\{\psi(X_1,\ldots,X_n) > s\} \cap \cup_{i=1}^{n-1}\{X_i > t_i\} \cup \{X_n > \psi^{\wedge}_{t_1,\ldots,t_n}(s)\}] \\
&\geq 1 - P[X_i > t_i, i = 1,\ldots,n, X_n > \psi^{\wedge}_{t_1,\ldots,t_n}(s)] \\
&= C(F_1(t_1),\ldots,F_{n-1}(t_{n-1}), F_n(\psi^{\wedge}_{t_1,\ldots,t_{n-1}}(s))) \\
&\geq C_0(F_1(t_1),\ldots,F_{n-1}(t_{n-1}), F_n(\psi^{\wedge}_{t_1,\ldots,t_{n-1}}(s))).
\end{aligned}$$

If $\psi^{\wedge}_{t_1,\ldots,t_{n-1}}(s) = +\infty$, then $\psi_{t_1,\ldots,t_{n-1}}(t_n) \leq s$ for all $t_n \in \mathbb{R}$ and therefore

$$\begin{aligned}
P[\psi(X_1,\ldots,X_n) \leq s] &\geq C_0(F_1(t_1),\ldots,F_{n-1}(t_{n-1}), 1) \\
&= C_0(F_1(t_1),\ldots,F_{n-1}(t_{n-1}), F_n(\psi^{\wedge}_{t_1,\ldots,t_{n-1}}(s))).
\end{aligned}$$

If $\psi^{\wedge}_{t_1,\ldots,t_{n-1}}(s) = -\infty$, then $C_0(F_1(x),\ldots,F_{n-1}(t_{n-1}), \psi^{\wedge}_{t_1,\ldots,t_{n-1}}(s)) = 0$ which is smaller or equal than $P[\psi(X_1,\ldots,X_n) \leq s]$. Taking the supremum over all $t_1,\ldots,t_{n-1} \in \mathbb{R}$ leads to $F_{\psi(X_1,\ldots,X_n)}(s) \geq F_{\min}(s)$. □

Proof of Theorem 3.2

Proof. For the optimality of $F_{\min}$ it suffices to define a copula C^{α} satisfying

$$C^{\alpha} \geq C_0, \tag{6.1}$$

$$\mu^{\alpha}(\{C_0 \leq \alpha\}) \leq \alpha, \tag{6.2}$$

where μ^{α} is the measure corresponding to C^{α}. Indeed, from (6.1) and Theorem 3.1 we have that $\alpha = \tau_{C_0,\psi}(F_1,\ldots,F_n)(s) \leq \sigma_{C^{\alpha},\psi}(F_1,\ldots,F_n)(s)$. Consequently, we only have to show that

$$\sigma_{C^{\alpha},\psi}(F_1,\ldots,F_n)(s) \leq \alpha. \tag{6.3}$$

The proof of (6.3) is essentially based on the idea of transporting the whole problem on the unit cube $[0,1]^n$. Let in fact $(U_1^{\alpha},\ldots,U_n^{\alpha})$ have distribution function C^{α}. Further, let $(X_1^{\alpha},\ldots,X_n^{\alpha}) := (F_1^{-1}(U_1^{\alpha}),\ldots,F_n^{-1}(U_n^{\alpha}))$ be a random vector with copula C^{α} and marginals $F_1,\ldots,F_n$. For the distribution ν^{α} of $(X_1^{\alpha},\ldots,X_n^{\alpha})$, the transform $h := (F_1^{-1},\ldots,F_n^{-1})$ and any measurable set $G \subset \mathbb{R}^n$ we have that

$$\nu^{\alpha}(G) = P[(X_1^{\alpha},\ldots,X_n^{\alpha}) \in G] = P[h(U_1^{\alpha},\ldots,U_n^{\alpha}) \in G] = \mu^{\alpha}(h^{-1}(G)).$$

Further, for $G = \{\psi \leq s\}$ we get that

$$h^{-1}(\{\psi \leq s\}) = \{(u_1,\ldots,u_n) \in [0,1]^n \,|\, \psi(F_1^{-1}(u_1),\ldots,F_n^{-1}(u_n)) \leq s\} =: A.$$

With this formalism, inequality (6.3) is equivalent to $\mu^{\alpha}(A) \leq \alpha$. We assume without loss of generality that $\{\psi \leq s\} \neq \emptyset$ i.e. that $A \neq \emptyset$ and we consider

$(t_1, \ldots, t_n) \in \{\psi \leq s\}$. For any such point we have by Lemma 2.1 that $t_n \leq \psi^{\wedge}_{t_1,\ldots,t_{n-1}}(s)$ and therefore that

$$\begin{aligned} C_0(F_1(t_1), \ldots, F_n(t_n)) &\leq C_0(F_1(t_1), \ldots, F_{n-1}(t_{n-1}), F_n(\psi^{\wedge}_{t_1,\ldots,t_{n-1}}(s))) \\ &\leq \tau_{C_0,\psi}(F_1, \ldots, F_n)(s) = \alpha. \end{aligned} \tag{6.4}$$

Equation (6.4) together with Lemma 2.1 implies for $(u_1, \ldots, u_n) \in A$ that

$$C_0(u_1, \ldots, u_n) \leq C_0(F_1(F_1^{-1}(u_1)), \ldots, F_n(F_n^{-1}(u_n))) \leq \alpha$$

i.e. that $A \subset \{C_0 \leq \alpha\}$. This means that (6.2) implies (6.3). The copula

$$C^{\alpha}(u_1, \ldots, u_n) := \begin{cases} C_0(u_1, \ldots, u_n) \vee \alpha & \text{when } (u_1, \ldots, u_n) \in [\alpha, 1]^n, \\ C_{\mathrm{U}}(u_1, \ldots, u_n) & \text{otherwise.} \end{cases}$$

satisfies condition (6.1) since C_{U} is the upper Fréchet bound. The values assigned by μ^{α} to subsets of $[0,1]^n$ are described in Figure 6.1 and in particular, μ^{α} assigns mass α to any set $[0, u_1] \times \cdots \times [0, u_n]$ such that $C_0(u_1, \ldots, u_n) = \alpha$. It follows

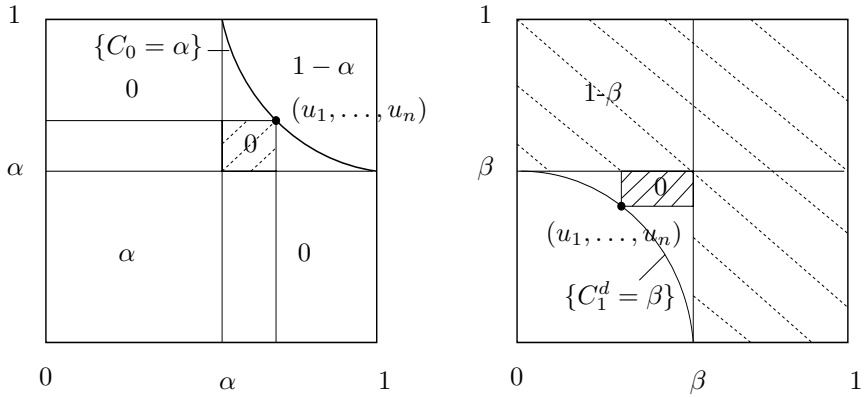

Figure 6.1: Mass values assigned by μ^{α} and μ^{β}.

that $\mu^{\alpha}(\{C_0 \leq \alpha\}) = \alpha$ and hence (6.3) has been proved.
The proof of the optimality of $F_{\max}$ is based on similar arguments. In fact it suffices to define a copula C^{β} such that

$$(C^{\beta})^d \leq C_1^d, \tag{6.5}$$

$$\mu^{\beta}(\{C_1^d \geq \beta\}) \leq 1 - \beta, \tag{6.6}$$

where μ^{β} is the measure corresponding to C^{β}. Indeed, from (6.5) and Theorem 3.1 we have that $\beta = \rho_{C_1,\psi}(F_1, \ldots, F_n)(s) \geq \sigma_{C^{\beta},\psi}(F_1, \ldots, F_n)(s)$. Consequently, we only have to show that

$$\sigma_{C^{\beta},\psi}(F_1, \ldots, F_n)(s) \geq \beta \tag{6.7}$$

which is the same as showing that $1-\sigma_{C^\beta,\psi}(F_1,\dots,F_n)(s)\le 1-\beta$. Using the above notation, we see that

$$\begin{aligned}1-\sigma_{C^\beta,\psi}(F_1,\dots,F_n)(s) &= P[\psi(X_1^\beta,\dots,X_n^\beta)>s]\\ &\le P[\psi(F_1^\wedge(U_1^\beta),\dots,F_n^\wedge(U_n^\beta))>s]=\mu^\beta(B),\end{aligned}$$

where $B:=g^{-1}(\{\psi>s\})$ and $g:=(F_1^\wedge,\dots,F_n^\wedge)$. We assume without loss of generality that $\{\psi>s\}\neq\emptyset$ i.e. that B is non-empty. Hence, for $(x_1,\dots,x_n)\in\{\psi>s\}$ it follows from Lemma 2.1 that $x_n>\psi^\wedge_{x_1,\dots,x_{n-1}}(s)$ and therefore that

$$\begin{aligned}\beta=\rho_{C_1}(F_1,\dots,F_n)(s)&\le C_1^d(F_1(x_1),\dots,F_{n-1}(x_{n-1}),F_n(\psi^\wedge_{x_1,\dots,x_{n-1}}(s)))\\ &\le C_1^d(F_1(x_1),\dots,F_{n-1}(x_{n-1}),F_n(x_n)).\end{aligned}$$

It follows for any $(u_1,\dots,u_n)\in B$ that

$$\beta\le C_1^d(F_1(F_1^\wedge(u_1)),\dots,F_1(F_n^\wedge(u_1)))\le C_1^d(u_1,\dots,u_n)$$

i.e. that $B\subset\{C_1^d\ge\beta\}$. This means that (6.6) implies (6.7). The idea behind the choice of C^β is better understood in the two-dimensional case with the help of Figure 6.1. In fact, for $n=2$, condition (6.5) is equivalent to $C^\beta\ge C_1$. Hence, we set first C^β equal to C_{U} on $[0,1]^2\setminus[0,\beta)^2$. Because of this, it follows that $\mu^\beta([0,\beta]^2)=\beta$. It suffices to define C^β such that μ^β assigns mass 0 to any rectangle $[u_1,\beta]\times[u_2,\beta]$ with $C_1^d(u_1,u_2)=\beta$. More precisely:

$$\begin{aligned}0&=C^\beta(\beta,\beta)-C^\beta(u_1,\beta)-C^\beta(\beta,u_2)+C^\beta(u_1,u_2)\\ &=\beta-u_1-u_2+C^\beta(u_1,u_2)=C_1^d(u_1,u_2)-u_1-u_2+C^\beta(u_1,u_2)\\ &=-C_1(u_1,u_2)+C^\beta(u_1,u_2).\end{aligned}$$

In other words, C^β must be equal to C_1 on $\{C_1^d=\beta\}$. In summary, define

$$C^\beta(u_1,u_2):=\begin{cases}u_1\wedge u_2 & \text{if } (u_1,u_2)\in[0,1]^2\setminus[0,\beta)^2\\ u_1+u_2-\beta & \text{if } (u_1,u_2)\in[0,\beta)^2\cap\{C_1^d\ge\beta\}\\ C_1(u_1,u_2) & \text{if } (u_1,u_2)\in[0,\beta)^2\cap\{C_1^d<\beta\}\end{cases}.$$

Notice that C^β is a copula satisfying both (6.5) and (6.6) since on $\{C_1^d\ge\beta\}$ holds $C_1(u_1,u_2)\le u_1+u_2-\beta$ and hence $(C^\beta)^d(u_1,u_2)\le C_1^d(u_1,u_2)$ for all u_1,u_2. These arguments can be generalized to $n\ge 2$, the idea being to define C^β such that $\mu^\beta([0,1]^n\setminus[0,\beta)^n)=1-\beta$ and $\mu^\beta(\prod_{i=1}^n[u_i,\beta])=0$ for any $(u_1,\dots,u_n)$ with $C_1^d(u_1,\dots,u_n)=\beta$. Moreover, the condition $(C^\beta)^d\le C_1^d$ is already taken into account when constructing C^β. □

Counterexample

In this paragraph we give a two-dimensional example where (4.4) is strictly smaller than $F_{\max}$ and not even an upper bound for $F_{\psi(X_1,X_2)}$. Denote by $G_{\max}$ the expression given by (4.4) i.e. $G_{\max}(s):=\inf_{\psi(t_1,t_2)=s}C_1^d(F_1(t_1),F_2(t_2))$.

Let $F_1 = F_2$ be uniform(0,1) distribution functions and $\psi : \mathbb{R}^2 \to \mathbb{R}$ be the increasing and continuous function defined by

$$\psi(x, y) := 2y1_{[0,1/2)}(y) + 1_{[1/2,2)}(y) + (y-1)1_{[2,\infty)}(y).$$

Note that ψ is actually a function in a single variable, namely y. For $s = 1$ we have that $\varphi_x^\wedge(s) = \varphi_x^\wedge(1) = 2$ for all $x \in \mathbb{R}$ and $\{(t_1, t_2) \,|\, \psi(t_1, t_2) = 1\} = \mathbb{R} \times [1/2, 2]$. Hence,

$$\begin{aligned} F_{\max}(1) &= \inf_{x\in\mathbb{R}} C_1^d(F_1(x), F_2(\varphi_x^\wedge(1))) = \inf_{x\in\mathbb{R}} C_1^d(F_1(x), 1) = C_1^d(0,1) = 1, \\ G_{\max}(1) &= \inf_{\psi(t_1,t_2)=1} C_1^d(F_1(t_1), F_2(t_2)) = \inf_{t_1\in\mathbb{R}, 1/2\le t_2\le 2} C_1^d(F_1(t_1), F_2(t_2)) \\ &= C_1^d(0, 1/2) = 1/2 \end{aligned}$$

and consequently $G_{\max}(1) < F_{\max}(1)$. Moreover, $G_{\max}(1)$ is not an upper bound for $F_{\psi(X_1,X_2)}(1)$. Indeed, $F_{\psi(X_1,X_2)}(1) = P[X_2 \le 2] = 1$ which is actually the bound given by $F_{\max}(1)$.

References

Artzner, P., Delbaen, F., Eber, J., Heath, D.: Coherent Measures of Risk. Math. Finance **9**, 203-228 (1999)

Bäuerle, N. and Müller, A.: Modelling and comparing dependencies in multivariate risk portfolios. Astin Bull. **28**, 59-76 (1998)

Boyd, S. and Vandenberghe, L.: Applications of semidefinite programming. Appl. Numer. Math. **29**, 283–299 (1999).

Cossette, H. and Denuit, M. and Marceau, E.: Distributional bounds for functions of dependent risks. Université Catholique de Louvain. Submitted (2000)

Denuit, M., Genest, C. and Marceau, E.: Stochastic bounds on sums of dependent risks. Insurance Math. Econom. **25**, 85-104 (1999)

Dhaene, J. and Goovaerts, M.J.: Dependency of risks and stop-loss order. Astin Bull. **26**, 201-212 (1996)

Embrechts, P., Klüppelberg, C. and Mikosch, T.: Modelling extremal events for insurance and finance. Berlin Heidelberg New York: Springer (1997)

Embrechts, P., McNeil, A. and Straumann, D.: Correlation: pitfalls and alternatives. Risk, 69-71 (1999a)

Embrechts, P., Lindskog, F. and McNeil, A.: Modelling dependence with copulas and applications to risk management. ETHZ Preprint (2001)

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

Frank, M.J., Nelsen, R.B. and Schweizer, B.: Best-possible bounds for the distribution for a sum — a problem of Kolmogorov. Probab. Th. Rel. Fields **74**, 199-211 (1987)

Frank, M.J. and Schweizer, B.: On the duality of generalized infimal and supremal convolutions. Rend. Mat. Ser. VI **12**, 1-23 (1979)

Galambos, J.: The Asymptotic Theory of Extreme Order Statistics. Kreiger Publishing Co., Melbourne, FL (1987)

Genest, C. and Rivest, L.P.: A characterization of Gumbel's family of extreme value distributions. Statist. Probab. Lett. **8**, 207-211 (1989)

Joe, H.: Multivariate Models and Dependence Concepts. London: Chapman and Hall 1997

Juri, A. and Wüthrich, M.V: Copula convergence theorems for tail events. ETHZ Preprint (2002)

Makarov, G.D.: Estimates for the distribution function of a sum of two random variables when the marginal distributions are fixed. Th. Probab. Appl. **26**, 803-806 (1981)

Müller, A.: Orderings of risks: A comparative study via stop-loss transforms. Insurance Math. Econom. **17**, 215-222 (1996)

Müller, A.: Stochastic orders generated by integrals: a unified study. Adv. Appl. Probab. **29**, 414-428 (1997a)

Müller, A.: Stop-loss ordering for portfolios of dependent risks. Insurance Math. Econom. **21**, 219-223 (1997b)

Müller, A. and Scarsini, M.: Some remarks on the supermodular order. J. Multivariate Anal. **73** no. 1, 107-119 (2000)

Nelsen, R.B.: An introduction to copulas. New York: Springer (1999)

Schweizer, B. and Sklar, A.: Operations on distribution functions not derivable from operations on random variables. Studia Math. **52**, 43-52 (1974).

Sklar, A.: Fonctions de répartition à n dimensions et leurs marges. Publ. Inst. Statist. Univ. Paris **8**, 229-231 (1959)

Williamson, R.C. and Downs, T.: Probabilistic arithmetic: numerical methods for calculating convolutions and dependency bounds. J. Approximate Reasoning **4**, 89-158 (1990)