On the Worst and Least Possible Asymptotic Dependence

Alexandru V. Asimit[1]

Cass Business School, City University London EC1Y 8TZ,
United Kingdom. E-mail: `asimit@city.ac.uk`

Russell Gerrard

Cass Business School, City University London EC1Y 8TZ,
United Kingdom. E-mail: `R.J.Gerrard@city.ac.uk`

October 30, 2015

Abstract. Multivariate extremes behave very differently under asymptotic dependence as compared to asymptotic independence. In the bivariate setting, we are able to characterise the extreme behaviour of the asymptotic dependent case by using the concept of the copula. As a result, we are able to identify the properties of the boundary cases, that are asymptotic independent but still have some asymptotic dependent features. These situations are the most problematic in statistical extreme, and, for this reason, distinguishing between asymptotic dependence and asymptotic independence represents a difficult problem. We propose a simple test to resolve this issue which is an alternative to the procedure based on the classical coefficient of tail dependence. In addition, we are able to identify the worst/least asymptotic dependence (in the presence of asymptotic dependence) that maximises/minimises the probability of a given extreme region if tail dependence parameter is fixed. It is found that the perfect extreme association is not the worst asymptotic dependence, which is consistent with the existing literature. We are able to find lower and upper bounds for some risk measures of functions of random variables. A particular example is the sum of random variables, for which a vivid academic effort has been noticed in the last decade, where bounds for a sum of random variables are sought. It is numerically shown that our approach provides a great improvement of the existing methods, which reiterates the sensible conclusion that any additional piece of information on dependence would help to reduce the spread of these bounds.

Keywords and phrases: Asymptotic dependence/independence; Copula; Extreme Value Theory; Gumbel Tail; Regular Variation; Risk measure.

1. Introduction

Estimation of multivariate extreme events is a challenging problem in *Extreme Value Theory* (EVT) and the starting point of non-parametric estimation is to decide if data exhibit the *asymptotic dependence* (AD) or *asymptotic independence* (AI) property. In simple words, under AD, concomitant extreme events are observed and both are at the same scale. Under AI, concomitant extreme events may occur

[1]Corresponding author. Phone: +44(0)2070405282. Fax: +44(0)2070408572.

but at different scales or may not even occur at the same time. Therefore, it is expected that extreme regions estimates to be very different in magnitude in the presence of AD than AI. It is well-known that statistical inferences in the presence of AI is very difficult, and many estimation methods are available if AD holds (see for example, de Haan and Ferreira, 2006). Since distinguishing between AD and AI plays an important role in predicting extreme events, Ledford and Tawn (1996, 1997) introduced the *coefficient of tail dependence* which has been extensively investigated in the literature. For example, nonparametric inference can be found in Peng (1999) and Draisma *et al.* (2004), while Goegebeur and Guillou (2012) considered an asymptotically unbiased estimator in the case of AI. The main disadvantage of the coefficient of tail dependence is that inconclusive results are possible, especially in situations which fall on the boundary between AD and AI. In order to help detect AI/AD, the recent paper of Asimit *et al.* (2015) proposes a conditional version of the classical measure of association Kendall's tau for absolutely continuous distributions.

The initial motivation of the paper was to examine in great details the joint tail behaviour of a bivariate random vector under AD and understand the differences between AD and almost AD (boundary between AD and AI) cases. Since we are interested in characterising the association of extreme events, the concept of the *copula* will be considered throughout this paper. Our properties will clarify the existing examples in the literature that pointed out naive conjectures of a link between some measure of tail dependence and the presence of AI/AD. Having in mind our AD characterisation, one may construct counterexample for such speculative conclusions and serve to provide a better understanding of extreme behaviour in the almost AD extreme behaviour. In fact, we exihibit one example, but many examples can be constructed in the same fashion, that can be useful as a model for any statistical extreme where the overlapping between AD and AI is of interest. We are able to identify the worst/least extreme dependence under AD with a fixed *tail dependence parameter*, which is a measure of tail dependence (for a summary of tail dependence concepts, we refer the reader to Hua and Joe, 2011). In our interpretation, worst (least) extreme dependence represents the least (most) favourable dependence that may occur and it really depends on the context. For example, when one deals with a sum of positive insurance losses, the worst(least) dependence is achieved when some tail risk measures of the aggregate risk is maximised (minimised). Note that focusing only on the tail dependence parameter, the overall dependence may be underestimated as argued in Furman *et al.* (2014). We can further find the upper and lower bounds for the tail distribution of a function of random variables (rv's). A special case is the sum of rv's that has been extensively studied in the literature as it can be seen below. Note that extreme quantile for a sum of rv's are of great interest in risk management among other areas (for example, see Asimit *et al.*, 2015). *Value-at-Risk* (which is in fact a quantile) is one of the most common risk measure used in practice in the banking and insurance industries, and therefore its evaluation has received particular attention in the last decade. The uncertainty with the dependence among rv's is huge, especially due to the data scarcity, and the choice of a parametric model is quite challenging even though such compromises are made in practice and are sometimes based on prior beliefs of the modeler. As a result, evaluating the range of values for the VaR of a sum of rv's is usually made when the marginal distributions are known and, possibly, an additional piece of information about dependence is known. This approach allows the decision-maker to understand the worst and least possible VaR-based risk. The best possible bounds for the distribution of a sum of rv's are described in Embrechts and Puccetti (2006 a and b) and the references therein. VaR bounds have been discussed in Embrechts *et al.* (2013), Wang *et al.* (2013) and Bernard *et al.* (2014), if no additional information about dependence is available. The recent paper of Bernard *et al.* (2015) investigates the VaR constrained set-up under an additional assumption that the aggregate variance is known. The same problem is investigated in Bernard *et al.* (2014) when the decision-maker has only a summary statistics of the individual risks (mean, variance, skewness etc, i.e. some high order expectations) instead of their distributions. Usually, these bounds are attained under extreme atomic dependence models which suggests that studying the constrained problem under a reduced set of feasible dependence structures represents the way forward in this field. As a result, Bignozzi *et al.* (2015) find VaR bounds under the assumption of lower orthant stochastic ordering with respect to a particular dependence model.

This paper first provides the necessary background in Section 2. The AD is fully characterised in Section 3, which enables us to identify the worst and least asymptotic dependence in Section 4. We propose a new procedure to identify the presence of AD/AI in Section 5. Section 6 numerically illustrates the advantages of our findings over the existing bounds available in the literature. Finally, all proofs are relegated in the Appendix.

2. Background

Let $X_1, \cdots, X_n$ be independent and identically distributed (iid) rv's with *cumulative distribution function* (cdf) F and infinite right-end point. EVT assumes that there are two sequences of constants $a_n > 0, b_n \in \Re$ such that

$$\lim_{n\to\infty} \mathbb{P}\left(a_n\left(\max_{1\le i\le n} X_i - b_n\right) \le x\right) = G(x), \qquad x \in \Re.$$

In this case, G is called an *Extreme Value Distribution* and F is said to belong to the *domain of attraction of G*. The Fisher-Tippett Theorem (see Fisher and Tippett, 1928) states that if the limit distribution is non-degenerate then $G(x) = \exp\{-x^{-\alpha}\}$ for all $x > 0$ with $\alpha > 0$ or $G(x) = \exp\{-e^{-x}\}$ for all $x \in \Re$, since the domain of F is assumed to be unbounded in the right tail. In the first case, X has the *regularly varying (RV)* property at ∞ with tail index α, i.e. the survival function $\bar{F} = 1 - F$ satisfies $\lim_{t\to\infty} \bar{F}(tx)/\bar{F}(t) = x^{-\alpha}$ for all $x > 0$, and we write $\bar{F} \in \mathcal{RV}_{-\alpha}$. In the second case, X has a *Gumbel* tail and it is well-known (see, for example, Resnick, 1987 or Embrechts *et al.*, 1997) that there exists a positive, measurable function a such that $\lim_{t\to\infty} \bar{F}(t + xa(t))/\bar{F}(t) = e^{-x}$ for all real x, and we write $\bar{F} \in \Lambda(a)$.

We now review the concept of *vague convergence*. Consider an n-dimensional cone $\mathcal{E}$ equipped with a Borel sigma-field $\mathcal{B}$. Two particular cones $\mathcal{E}_{\mathcal{F}} = [\mathbf{0}, \boldsymbol{\infty}]\backslash\{\mathbf{0}\}$ and $\mathcal{E}_{\mathcal{G}} = [-\boldsymbol{\infty}, \boldsymbol{\infty}]\backslash\{-\boldsymbol{\infty}\}$ will be of interest in this paper. In particular, $\mathcal{E}_{\mathcal{F}}$ is involved when the tails are RV, while $\mathcal{E}_{\mathcal{G}}$ becomes the main interest whenever we deal with Gumbel tails. A measure on the cone is called Radon if its value is finite for every compact set in $\mathcal{B}$. For a sequence of Radon measures $\{\nu, \nu_k, k = 1, 2, \ldots\}$ on $\mathcal{E}$, we say that ν_k vaguely converges to ν, written as $\nu_k \xrightarrow{v} \nu$, if

$$\lim_{k\to\infty} \int_{\mathcal{E}} h(\mathbf{z})\nu_k(\mathrm{d}\mathbf{z}) = \int_{\mathcal{E}} h(\mathbf{z})\nu(\mathrm{d}\mathbf{z})$$

holds for every nonnegative continuous function h with compact support. It is known that $\nu_k \xrightarrow{v} \nu$ on $\mathcal{E}_{\mathcal{F}}$ if and only if the convergence

$$\lim_{k\to\infty} \nu_k \left[\mathbf{0}, \mathbf{x}\right]^c = \nu \left[\mathbf{0}, \mathbf{x}\right]^c$$

holds for every continuity point $\mathbf{x} \in \mathcal{E}_{\mathcal{F}}$ of the limit. Obviously, $\mathbf{0}$ is replaced by $-\boldsymbol{\infty}$ if $\mathcal{E}_{\mathcal{G}}$ appears instead. For more details and related discussions, we refer the reader to Section 3.3.5 and Lemma 6.1 of Resnick (2007).

Dependence among rv's plays an important role in our paper, and we therefore introduce the concept of a copula. Let X and Y be two rv's with cdf's F and G, respectively. The dependence structure associated with the distribution of a random vector can be characterised in terms of its copula. A bivariate copula is a two-dimensional cdf defined on $[0,1]^2$ with uniformly distributed marginals. Due to Sklar's Theorem (see Sklar, 1959), if F and G are continuous, then there exists a unique copula, C, such that $\mathbb{P}(X \le x, Y \le y) = C\left(F(x), G(y)\right)$. Similarly, the *survival copula*, $\widehat{C}$, is defined as the copula corresponding to the joint tail function, i.e. the distribution of $\left(\bar{F}(X), \bar{G}(Y)\right)$ (see Nelsen, 2006).

Our main assumption on dependence is given as Assumption 2.1.

Assumption 2.1. *Assume that there exists a non-degenerate function* $H : [0,1]^2 \to [0,1]$ *such that*

$$H(x,y) = \lim_{u\downarrow 0} \frac{\widehat{C}(ux, uy)}{u}. \tag{2.1}$$

Consequently, $H(1,1) := c \in (0,1]$, *which is also called the tail dependence parameter.*

It is not difficult to find that H is a homogeneous function of order one, i.e. $H(t\cdot) = tH(\cdot)$. In addition, $H(\cdot) > 0$ on $(0,1]^2$, since otherwise the homogeneity property of function $H \equiv 0$ would make H degenerate. It is also true (see Nelsen, 2006) that $\widehat{C}(x,y) \le \min(x,y)$ and therefore $H(x,y) \le \min(x,y)$. Further, define $H_X(x) = H(x,1)/c$ and $H_Y(y) = H(1,y)/c$ the marginal cdf's of the joint cdf $H(\cdot)/c$. By setting $y = a$ and $x = az$, we see that $H(x,y) = cyH_X\left(\frac{x}{y}\right)$ if $x \le y$. In general,

$$H(x,y) = c\, yH_X\left(x/y\right)I\left(x \le y\right) + c\, xH_Y\left(y/x\right)I\left(y < x\right), \tag{2.2}$$

where I represents the *indicator function.* Moreover, $x \le H_X(x), H_Y(x) \le \min\left(x/c, 1\right)$ for all $0 \le x \le 1$ (for details, see Asimit *et al.*, 2015).

Note that $c > 0$ is assumed, which means that X and Y are AD (see de Haan and Ferreira, 2006 or Klüppelberg and Resnick, 2008). Alternatively, if $\lim_{u\downarrow 0} \frac{\hat{C}(u,u)}{u} = 0$, then we have AI. In order to distinguish between AD and AI, Ledford and Tawn (1996, 1997) introduced the concept of the coefficient of tail dependence $\eta \le 1$ by assuming that

$$\widehat{C}(u,u) = u^{1/\eta} s(u)\left(1 + o(1)\right) \text{ as } u \downarrow 0, \tag{2.3}$$

where s is a slowly varying function, i.e. $\lim_{u\downarrow 0} s(ux)/s(u) = 1$ for all $x > 0$. Thus, under condition (2.3), when $\eta = 1$ and $\lim_{u\downarrow 0} s(u) = c \in (0,1]$, AD property holds, while either $\eta < 1$ or $\eta = 1$ and $\lim_{u\downarrow 0} s(u) = 0$ implies AI. Therefore, η and the limit behaviour of function s can be used to distinguish between AD and AI. Note that standard estimators for c are available in Asimit *et al.* (2015) and Haug *et al.* (2011).

In order to explain the joint tail behaviour, we also need to assume that X and Y have similar tails.

Assumption 2.2. *The random variables X and Y are tail equivalent such that* $\lim_{t\to\infty} \bar{G}(t)/\bar{F}(t) = 1$.

Let $\overline{H} : \overline{\Re}_+^2 \setminus \{\mathbf{0}\} \to \Re_+$ such that

$$\overline{H}(x,y) := \max(x,y) H\left(\frac{x}{\max(x,y)}, \frac{y}{\max(x,y)}\right).$$

It can be shown that, for $\bar{F} \in \mathcal{RV}_{-\alpha}$,

$$\lim_{t\to\infty} \frac{\mathbb{P}\big(X > xt, Y > yt\big)}{\mathbb{P}(X > t)} = \overline{H}\big(x^{-\alpha}, y^{-\alpha}\big) \tag{2.4}$$

and, for $\bar{F} \in \Lambda(a)$,

$$\lim_{t\to\infty} \frac{\mathbb{P}\big(X > t + xa(t), Y > t + ya(t)\big)}{\mathbb{P}(X > t)} = \overline{H}\big(e^{-x}, e^{-y}\big) \tag{2.5}$$

(for details, see Alink *et al.*, 2007 and Kortschak and Albrecher, 2009). Consequently, if $\bar{F} \in \mathcal{RV}_{-\alpha}$, then

$$\frac{\mathbb{P}\Big(\big(X/t, Y/t\big) \in \cdot\Big)}{\bar{F}(t)} \overset{v}{\to} \mu_F(\cdot) \tag{2.6}$$

holds on $\mathcal{E}_{\mathcal{F}}$, where $\mu_F\big((x,\infty] \times (y,\infty]\big) := \overline{H}\big(x^{-\alpha}, y^{-\alpha}\big)$. Similarly, if $\bar{F} \in \Lambda(a)$, then

$$\frac{\mathbb{P}\bigg(\Big((X-t)/a(t), (Y-t)/a(t)\Big) \in \cdot\bigg)}{\bar{F}(t)} \overset{v}{\to} \mu_G(\cdot) \tag{2.7}$$

holds on $\mathcal{E}_{\mathcal{G}}$, where $\mu_G\big((x,\infty] \times (y,\infty]\big) := \overline{H}\big(e^{-x}, e^{-y}\big)$.

3. Characterisation of AD

This section provides a characterisation of the AD as defined in Assumption 2.1 and we show in Propositions 3.1 and 3.2 that the limiting dependence is fully described by its marginal cdf's, namely H_X and H_Y. The one-to-one relationship incentivise the authors even more to understand the properties of marginal cdf's. These technical results will help us later in Section 4 to find the worst and least possible extreme dependence, which is the main aim of our paper.

Proposition 3.1. *If Assumption 2.1 holds, then H_X and H_Y are continuous and possess right derivatives h_X and h_Y, which are themselves continuous and satisfy $h_X(1-) + h_Y(1-) = 1 + d$, for some $d \in [0,1]$. In addition, $x^{-1}H_X(x)$, $x^{-1}H_Y(x)$, $h_X(x)$ and $h_Y(x)$ are non-increasing functions of x. Moreover, $h_X(0+) = h_Y(0+) = 1/c$.*

A straightforward implication of Proposition 3.1 is given by Corollary 3.1, and its proof is left to the reader.

Corollary 3.1. *If Assumption 2.1 holds, then $H_X(x) \le 1 - h_X(1-)(1-x)$ and $H_Y(x) \le 1 - h_Y(1-)(1-x)$ for all $0 \le x \le 1$.*

It is interesting to find out whether, for any given pair of cdf's H_X and H_Y on $[0,1]$, possessing density functions h_X and h_Y, there exists a copula that satisfies (2.1). It is natural to believe that the bivariate cdf derived via (2.2) has a copula that holds the property from (2.1), which is established in the next proposition.

Proposition 3.2. *Let H_X and H_Y be two cdf's on $[0,1]$ with non-increasing density functions h_X and h_Y such that $h_X(0+) = h_Y(0+) = 1/c$ and $h_X(1-) + h_Y(1-) = 1 + d$, where $c \in (0,1]$ and $d \in [0,1]$. Define*

$$J(x,y) = \begin{cases} yH_X\left(\frac{x}{y}\right) & \text{if } 0 \le x < y \le 1 \\ xH_Y\left(\frac{y}{x}\right) & \text{if } 0 \le y \le x < 1 \end{cases}$$

Then J is a bivariate cdf with marginals H_X and H_Y, and its copula, $J\big(H_X^{-1}, H_Y^{-1}\big)$, satisfies (2.1) with $H \equiv cJ$, where H_X^{-1} and H_Y^{-1} are the left-continuous inverses of H_X and H_Y, respectively. Moreover, $x^{-1}H_X(x)$ and $x^{-1}H_Y(x)$ are non-increasing functions in $x \in (0,1]$.

Finally, we examine in Proposition 3.3 the almost AD cases, i.e. $\eta = 1$ and $c = 0$. The proof is left to the reader since it can be shown in the same manner as Propositions 3.1 and 3.2.

Proposition 3.3. *i) Assume that there exists a non-degenerate homogeneous of order one function $H : [0,1]^2 \to [0,1]$ such that*

$$H(x,y) = \lim_{u \downarrow 0} \frac{\widehat{C}(ux, uy)}{\widehat{C}(u,u)} \quad \text{and} \quad \lim_{u \downarrow 0} \frac{\widehat{C}(u,u)}{u} = 0. \tag{3.1}$$

Then $H_X(\cdot) = H(\cdot, 1)$ and $H_Y(\cdot) = H(1, \cdot)$ are continuous and possess right derivatives h_X and h_Y, which are themselves continuous and satisfy $h_X(1-) + h_Y(1-) \in [1,2]$. In addition, $x^{-1}H_X(x)$, $x^{-1}H_Y(x)$, $h_X(x)$ and $h_Y(x)$ are non-increasing functions of x. Moreover, $h_X(0+) = h_Y(0+) = \infty$.
ii) Let H_X and H_Y be two cdf's on $[0,1]$ with non-increasing density functions h_X and h_Y such that $h_X(0+) = h_Y(0+) = \infty$ and $h_X(1-) + h_Y(1-) \in [1,2]$. Define

$$J(x,y) = \begin{cases} yH_X\left(\frac{x}{y}\right) & \text{if } 0 \le x < y \le 1 \\ xH_Y\left(\frac{y}{x}\right) & \text{if } 0 \le y \le x < 1 \end{cases}.$$

Then J is a bivariate cdf with marginals H_X and H_Y, and its copula, $J\big(H_X^{-1}, H_Y^{-1}\big)$, satisfies (3.1) with $H \equiv J$, where H_X^{-1} and H_Y^{-1} are the left-continuous inverses of H_X and H_Y, respectively. Moreover, $x^{-1}H_X(x)$ and $x^{-1}H_Y(x)$ are non-increasing functions in $x \in (0,1]$.

Having in mind Proposition 3.3, one may easily construct examples that exhibit the almost AD property. Two examples are as follow:

$$H_X(x) = H_Y(x) := \frac{(x+1)\log(x+1) - x\log x}{2\log 2} \tag{3.2}$$

and

$$H_X(x) = H_Y(x) := x\left(1 - \frac{1}{2}\log x\right). \tag{3.3}$$

Note that the (3.2) appeared as Example 5.2 in Juri and Wüthrich (2003). Both examples are counterexamples to the naive conjecture that AI implies a joint extreme behaviour similar to independence:

$$\mathbb{P}\big((U,V) \in \cdot | U, V \le u\big) \simeq \mathbb{P}\big(U \in \cdot | U, V \le u\big)\, \mathbb{P}\big(V \in \cdot | U, V \le u\big),$$

for u sufficiently close to 0, where the random vector (U,V) has cdf H.

4. Worst and Least Dependence

The AD profile of a bivariate random vector is discussed in great detail in Section 3. These properties are useful to explain how to find the largest and lowest possible value (and their corresponding dependence structures) of an extreme event with a fixed positive value for $c > 0$. Examples include the tail probability of a function of rv's such as sum, product, absolute difference etc. We first provide the mathematical formulation of the chosen problems which are given in Theorem 4.1. These results are the key ingredient in establishing our bounds for the tail probability as obtained in Proposition 4.1 and Lemma 4.1.

Theorem 4.1. *Suppose that Assumptions 2.1 and 2.2 hold.*

i) *If $\bar{F} \in \mathcal{RV}_{-\alpha}$, then for any $b > 0$ we have that*

$$\lim_{t\to\infty} \frac{\mathbb{P}(X + bY > t)}{\mathbb{P}(X > t)} = 1 + b^\alpha - c(1+b)^\alpha \tag{4.1}$$
$$+bc\left(\int_0^1 z^{1/\alpha-1}\left(\left(1 + bz^{1/\alpha}\right)^{\alpha-1} h_X(z) + \left(b + z^{1/\alpha}\right)^{\alpha-1} h_Y(z)\right)\right) dz.$$

ii) *If $\bar{F} \in \mathcal{RV}_{-\alpha}$, then*

$$\lim_{t\to\infty} \frac{\mathbb{P}(XY > t^2)}{\mathbb{P}(X > t)} = -c + \frac{c}{2}\left(\int_0^1 z^{-1/2}\big(h_X(z) + h_Y(z)\big)\right) dz. \tag{4.2}$$

ii) *If $\bar{F} \in \Lambda(a)$, then*

$$\lim_{t\to\infty} \frac{\mathbb{P}(X + Y > 2t)}{\mathbb{P}(X > t)} = -c + \frac{c}{2}\left(\int_0^1 z^{-1/2}\big(h_X(z) + h_Y(z)\big)\right) dz. \tag{4.3}$$

Remark 4.1. *It is well-known (see Resnick, 2007) that the limit from (4.1) under AI becomes*

$$\lim_{t\to\infty} \frac{\mathbb{P}(X + bY > t)}{\mathbb{P}(X > t)} = \lim_{t\to\infty} \frac{\mathbb{P}(X > t) + \mathbb{P}(bY > t)}{\mathbb{P}(X > t)} = 1 + b^\alpha.$$

Now, the same limit is equal to 2 if $\alpha = b = 1$ for any $c \in [0, 1]$, which justifies the particular example from Section 3.2 of Klüppelberg and Resnick (2008). In other words, AD and AI provide the same limit whenever X and Y are tail equivalent and RV with tail index of 1. This is a another counterexample that a stronger positive dependence in the tail (usually simplified to the value of c) would increase the tail probability of $X + Y$. Recall that Embrechts et al. (2009) concluded that the marginal cdf's affect the tail behaviour and may have a greater impact than the dependence.

Let $\{a_X(x), a_Y(x) : 0 \le x \le 1\}$ be two continuous, monotone functions of x. As observed in Theorem 4.1, the aim is to find a pair (h_X, h_Y) of densities, satisfying the sufficient conditions stated in Proposition 3.2 in order to minimise (respectively maximise) an infinite dimensional optimisation problem with objective function given by:

$$J(a_X, a_Y) = \int_0^1 a_X(x)\, h_X(x)\, dx + \int_0^1 a_Y(y) h_Y(x)\, dx, \tag{4.4}$$

or at least to identify the infimum (supremum) of this quantity in the event that it is not attained. Thus, $h_X \in \mathcal{H}_{\xi,c,d}$ and $h_Y \in \mathcal{H}_{1+d-\xi,c,d}$, where

$$\mathcal{H}_{\xi,c,d} := \{h : h \text{ is a non-increasing density on } [0,1],\ h(0+) = 1/c, h(1-) = \xi\},\ \xi \in [d,1],\ c \in (0,1].$$

The infinite dimensionality issue is solved in Theorem 4.2 by reducing the set of feasible solutions.

Theorem 4.2. *Suppose $a:[0,1]\to\mathbb{R}$ is a continuous and monotone increasing function. Then*

$$\inf_{h\in\mathcal{H}_{\xi,c,d}}\int_0^1 a(x)h(x)\,dx=\int_0^1 a(x)h^*(x;\xi,c)\,dx \quad and \quad \sup_{h\in\mathcal{H}_{\xi,c,d}}\int_0^1 a(x)h(x)\,dx=\int_0^1 a(x)\,dx,$$

where

$$h^*(x;\xi,c,d)=\begin{cases}\frac{1}{c} & \text{if } 0<x<x(\xi,c)\\ \xi & \text{if } x(\xi,c)<x<1\end{cases}, \tag{4.5}$$

and $x(\xi,c)$ is chosen in order for $h^*(\cdot;\xi,c,d)$ is a valid density. That is, $x(\xi,c)=\frac{(1-\xi)c}{1-\xi c}$. If a is continuous and monotone decreasing, the extrema are interchanged.

With the help of Theorem 4.2, we can now solve the infinite dimensional optimisation problem defined in (4.4). Namely, if $a_X(\cdot)$ and $a_Y(\cdot)$ are continuous and monotone increasing functions on $[0,1]$, we have

$$\inf_{(h_X,h_Y)\in(\mathcal{H}_{\xi,c,d},\mathcal{H}_{1+d-\xi,c,d})}J(a_X,a_Y)=\inf_{0\le d\le\xi\le 1}\int_0^1\Big(a_X(x)h^*(x;\xi,c,d)+a_Y(x)h^*(x;1+d-\xi,c,d)\Big)dx,$$

$$\sup_{(h_X,h_Y)\in(\mathcal{H}_{\xi,c,d},\mathcal{H}_{1+d-\xi,c,d})}J(a_X,a_Y)=\int_0^1\big(a_X(x)+a_Y(x)\big)dx.$$

Note the the supremum is not attained unless $c=1$ while the inf and min are interchangeable (optimisation is made on a compact set, specifically $[0,1]$), and therefore we may obtain a sharp lower bound. Moreover, if a_X and a_Y are continuous and monotone decreasing functions, then the infimum and supremum may swap with each other. Sometimes, closed-form solutions can be found and are explicitly given in Proposition 4.1.

Proposition 4.1. *Denote $A_X(x)=\int_0^x a_X(y)\,dy$ and $A_Y(x)=\int_0^x a_Y(y)\,dy$.*

i) *Assume that $a_X\equiv a_Y\equiv a$ such that a is a continuous, monotone increasing function, the extremal values of $J(a,a)$ defined in (4.4) are*

$$\inf_{(h_X,h_Y)\in(\mathcal{H}_{\xi,c,d},\mathcal{H}_{1+d-\xi,c,d})}J(a,a)=A_X(1)+\frac{1}{r^*(c)}A_X\big(r^*(c)\big),$$

$$\sup_{(h_X,h_Y)\in(\mathcal{H}_{\xi,c,d},\mathcal{H}_{1+d-\xi,c,d})}J(a,a)=2A_X(1),$$

where $r^*(c)=c/(2-c)<1$. If a is a continuous and monotone decreasing function, then the infimum and supremum are reversed.

ii) *Assume that a_X and a_Y are monotone increasing and decreasing, respectively, continuous functions, then the extremal values of $J(a_X,a_Y)$ defined in (4.4) are*

$$\inf_{(h_X,h_Y)\in(\mathcal{H}_{\xi,c,d},\mathcal{H}_{1+d-\xi,c,d})}J(a_X,a_Y)=\frac{A_X(c)}{c}+A_Y(1),$$

$$\sup_{(h_X,h_Y)\in(\mathcal{H}_{\xi,c,d},\mathcal{H}_{1+d-\xi,c,d})}J(a_X,a_Y)=A_X(1)+\frac{A_Y(c)}{c}.$$

Remark 4.2. *Proposition 4.1i) tells us that for all symmetric problems (in a_X and a_Y) from Theorem 4.1, namely (4.1) with $b=1$ and $\alpha>1$, (4.2) and (4.3) have a lower bound when $H(x,y)=c\min(x,y)$, which can be achieved for many copulae (for example, take $\hat{C}(u,v)=c\min\{u,v\}+(1-c)uv$). That is, the least extreme dependence for a sum with a given c is the Fréchet-Hoeffding upper bound (when the upper copula, as explained in Juri and Wüthrich, 2003, is the Fréchet-Hoeffding upper bound). This confirms* *the fact that quantiles of a sum are maximised under negative association instead of a maximum positive associtíon (see for example, Embrechts et al., 2005). On the other side, the worst extreme dependence for a sum with a given c is given by*

$$H_*^{-1}\big(\max(u,v)\big)H_*\left(\frac{H_*^{-1}\big(\min(u,v)\big)}{H_*^{-1}\big(\max(u,v)\big)}\right),$$

where $H_*(x)=\frac{x}{c}I\big(0\le x\le r^*(c)\big)+\frac{1+x}{2}I\big(r^*(c)<x\le 1\big)$.

In the very end of this section we outline a variant of Lemma 7.1 when function $b(\cdot)$ is not always positive on $[0,1]$. The infinite dimensional optimisation problem is first solved over a reduced feasibility set given by

$$\overline{\mathcal{H}}_{\epsilon,\xi,y_0} := \big\{H : H \text{ is a cdf on } [0,1] \text{ with a non-increasing density } h \text{ such that (4.6) is satisfied}\big\},$$

$$H(x_0)=y_0, h(0+)=\frac{1}{c}, h(x_0+)\le\epsilon\le h(x_0-), h(1-)=\xi, \tag{4.6}$$

where $c\in(0,1]$ and $x_0\in(0,1)$ are some constants. In addition, the remaining parameters should satisfy

$$\xi\le\frac{1-y_0}{1-x_0}\le\epsilon\le\frac{y_0}{x_0}\le\frac{1}{c},\ d\le\xi\le 1,\ x_0\le y_0. \tag{4.7}$$

The final result is given below as Lemma 4.1 and its proof is left to the reader since one can follow similar arguments to the one used in the proof of Lemma 7.1.

Lemma 4.1. *Suppose $b:[0,1]\to\Re$ such that $\int_0^1|b(x)|\,dx<\infty$. In addition, there exists $0<x_0<1$ such that $b(x)\le 0$ and $b(x)\ge 0$ if $0\le x\le x_0$ and $x_0\le x\le 1$, respectively. Then*

$$\inf_{H\in\overline{\mathcal{H}}_{\epsilon,\xi,d,y_0}}\int_0^1 b(x)H(x)\,dx=\int_0^1 b(x)\underline{H}^*(x;\epsilon,\xi,d,y_0)\,dx,$$
$$\sup_{H\in\overline{\mathcal{H}}_{\epsilon,\xi,d,y_0}}\int_0^1 b(x)H(x)\,dx=\int_0^1 b(x)\overline{H}^*(x;\epsilon,\xi,d,y_0)\,dx$$

where

$$\underline{H}^*(x;\epsilon,\xi,d,y_0)=\begin{cases}\frac{x}{c} & \textit{if } 0\le x\le\frac{c(y_0-\epsilon x_0)}{1-c\epsilon}\\ y_0+\epsilon(x-x_0) & \textit{if } \frac{c(y_0-\epsilon x_0)}{1-c\epsilon}\le x\le x_0\\ 1-\frac{1-y_0}{1-x_0}(1-x) & \textit{if } x_0\le x\le 1\end{cases}$$

and

$$\overline{H}^*(x;\epsilon,\xi,d,y_0)=\begin{cases}\frac{y_0}{x_0}x & \textit{if } 1\le x\le x_0\\ y_0+\epsilon(x-x_0) & \textit{if } x_0\le x\le\frac{1-\xi-y_0+\epsilon x_0}{\epsilon-\xi}\\ 1-\xi(1-x) & \textit{if } \frac{1-\xi-y_0+\epsilon x_0}{\epsilon-\xi}\le x\le 1\end{cases}$$

As before, the desired bounds can be found via a finite dimensional constrained optimisation problem by varying the parameters ϵ, ξ, d and y_0 over the set defined in equation (4.7).

5. Detecting AD

It has been previously explained the importance of knowing whether or not AD represents a reasonable assumption. We already know that $\eta = 1$ may imply AD or AI. This section provides a new way of detecting AD and elaborates a simple test statistic to differentiate between AD and AI.

Let X and Y be two identically distributed truncated Pareto rv's with survival function $\bar{F}(t) = x^{-\alpha}$ for all $x \geq 1$. If the survival copula $\widehat{C}$ of (X, Y) satisfies Assumption 2.1, then from Theorem 4.1 i) we get that

$$\lim_{t\to\infty} \frac{\mathbb{P}(X+Y>t)}{\mathbb{P}(X>t)} = 2 - c\, 2^{\alpha} + c \int_0^1 z^{1/\alpha-1}\big(1+z^{1/\alpha}\big)^{\alpha-1}\big(h_X(z)+h_X(z)\big)\, dz.$$

The lower and upper bounds for the above limit can be found via Proposition 4.1 i) with

$$A(x;\alpha) = \left(1 + x^{1/\alpha}\right)^{\alpha} - 1.$$

If $\alpha > 1$ then

$$\underline{K}(\alpha, c) \leq \lim_{t\to\infty} \frac{\mathbb{P}(X+Y>t)}{\mathbb{P}(X>t)} \leq \overline{K}(\alpha, c), \tag{5.1}$$

where

$$\underline{K}(\alpha, c) := 2 - c\, 2^{\alpha} + 2cA(1;\alpha) = 2 + c\,(2^{\alpha} - 2)$$

and

$$\overline{K}(\alpha, c) := 2 - c\, 2^{\alpha} + c\left(A(1;\alpha) + \frac{2-c}{c} A\left(\frac{c}{2-c};\alpha\right)\right) = \left((2-c)^{1/\alpha} + c^{1/\alpha}\right)^{\alpha}.$$

Thus, the lower bound is strictly greater than 2 under AD, while Remark 4.1 tells us that the limit is always 2 under AI. These suggest a way of testing AD against AI as follows

$$H_0:\ K(\theta) > 2 \text{ versus } H_1:\ K(\theta) = 2,$$

for any fixed $\theta > 1$, where we define

$$K(\theta) = \lim_{u\downarrow 0} K(u;\theta) \text{ and } K(u;\theta) := \frac{\mathbb{P}\left(\left(U^{-1/\theta} + V^{-1/\theta}\right)^{-\theta} \leq u\right)}{u}.$$

Note that the asymptotic upper tail dependence of (X, Y) and the asymptotic lower tail dependence of $(U, V) = \big(\bar{F}(X), \bar{F}(Y)\big)$ are equal. Therefore, we check the AD/AI property for the pair of standard uniform (U, V) in the lower tail instead of the upper tail, and as a result, the assumption from (2.3) is replaced by

$$C(u, u) = u^{1/\eta} s(u)\big(1 + o(1)\big) \text{ as } u \downarrow 0.$$

Similarly, another way of testing AD against AI is as follows:

$$H_0:\ K(\theta) < 2 \text{ versus } H_1:\ K(\theta) = 2,$$

for any fixed $0 < \theta < 1$.

We now provide a brief simulation study for our proposed test for distinguishing between AD and AI. Obviously, a more detailed investigation is needed in order to grasp multiple potential problems that usually arise with such estimators (for example the optimal fraction problem), but these aspects are beyond the scope of this paper. Four dependence models are assumed as follow:

(A) Farlie-Gumbel-Morgenstern copula

$$C(u,v;\xi) := uv\big(1+\xi(1-u)(1-v)\big),\ -1 \le \xi \le 1.$$

The lower AI holds with $\eta = 1/2$.

(B) The first almost AD example with copula

$$H_X^{-1}\big(\max(u,v)\big)H_X\left(\frac{H_X^{-1}\big(\min(u,v)\big)}{H_X^{-1}\big(\max(u,v)\big)}\right),$$

where H_X is defined in (3.2). Recall that the lower AI holds with $\eta = 1$.

(C) The second almost AD example with copula

$$H_X^{-1}\big(\max(u,v)\big)H_X\left(\frac{H_X^{-1}\big(\min(u,v)\big)}{H_X^{-1}\big(\max(u,v)\big)}\right),$$

where H_X is defined in (3.3). Recall that the lower AI holds with $\eta = 1$.

(D) Clayton copula

$$C(u,v;\xi) := \left(u^{-1/\xi} + v^{-1/\xi} - 1\right)^{-\xi},\ \xi > 0.$$

The lower AD holds with $c = 2^{-\xi}$ and $H_X(x;\xi) = H_Y(x;\xi) = \left(\frac{1+x^{-1/\xi}}{2}\right)^{-\xi}$.

A sample (U_i, V_i) of size $n = 5,000$ is drawn from each copula and we plot in Figures 5.1, 5.2, 5.3 and 5.4 the tail dependence estimators, $\widehat{\eta}$ and $\widehat{K}(\theta)$ with $\theta = 1.2, 1.3$ for the four dependence models and different values of k. The value of k represents the fraction of the sample which is considered to be extreme behaviour of the sample. Recall that we investigate AD/AI at the lower end. The tail dependence estimator (at the upper end) and its properties have been investigated investigated in Draisma *et al.* (2004). In our setting, we have

$$\widehat{\eta}(k) = \frac{1}{k}\sum_{i=1}^{k} \log \frac{T_{(i)}}{T_{(k+1)}},$$

where $T_{(i)}$ is the ith largest order statistics of

$$T_i = \min\left(\frac{n+1}{R^{U_i}}, \frac{n+1}{R^{V_i}}\right)$$

with R^{U_i} being the rank of U_i among $U_1, U_2 \ldots, U_n$ and R^{V_i} being the rank of V_i among $V_1, V_2 \ldots, V_n$.

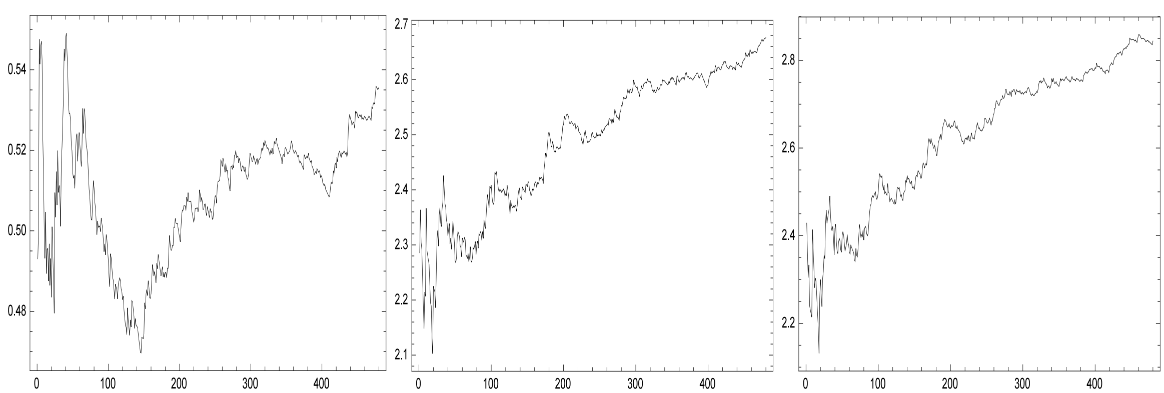

FIGURE 5.1. Estimators $\widehat{\eta}(k)$, $\widehat{K}(1.2;k)$ and $\widehat{K}(1.3;k)$ for copula (A) with $\xi = 0.5$ are plotted against $k = \{21, \ldots, 500\}$.

An estimator for $K(\theta)$ is

$$\widehat{K}(\theta;k) = \frac{1}{k}\sum_{i=1}^{n} I\left(\left(U_i^{-1/\theta} + V_i^{-1/\theta}\right)^{-\theta} \leq k/n\right).$$

The AI with $\eta < 1$ from Figure 5.1 seems quite clear and the η plot is more informative. The almost AD copulae from Figures 5.2 and 5.3 show that our proposed estimator could be carefully used in conjunction with the classical coefficient of tail dependence. The AD copula plots displayed in Figure 5.4 suggest that a significant change of $K(\theta)$ when θ marginally changes would be an indication that AD is present, but an extensive simulation study would provide a better understanding of how to interpret such plots. The horizontal lines in Figure 5.4 represent the theoretical values for $K(\theta)$ calculated via (4.1) as follows:

$$K(\theta) = 2 - 2^{\theta-\xi} + 2^{1-\xi}\int_0^1 z^{1/\theta-1}\left(1+z^{1/\theta}\right)^{\theta-1} H_X(dz;\xi).$$

Numerical evaluations show that $K(1.2) = 2.22974$ and $K(1.3) = 2.36934$.

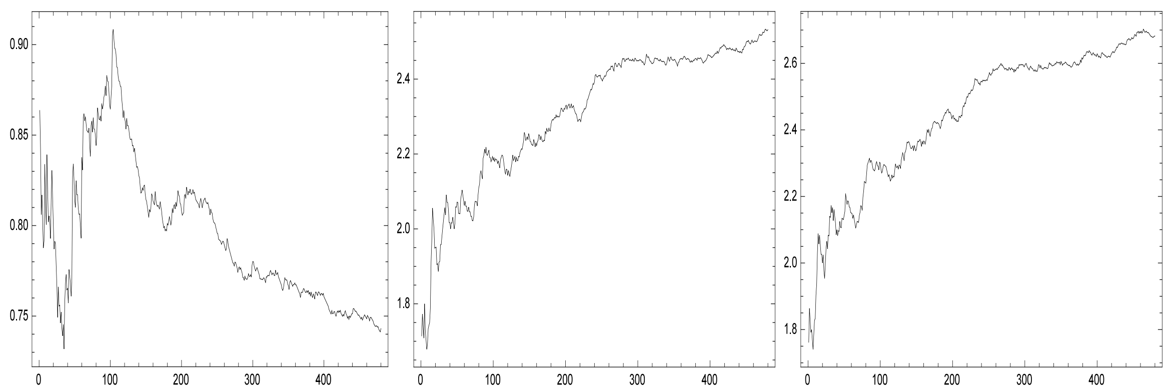

FIGURE 5.2. Estimators $\widehat{\eta}(k)$, $\widehat{K}(1.2;k)$ and $\widehat{K}(1.3;k)$ for copula (B) are plotted against $k = \{21,\ldots,500\}$.

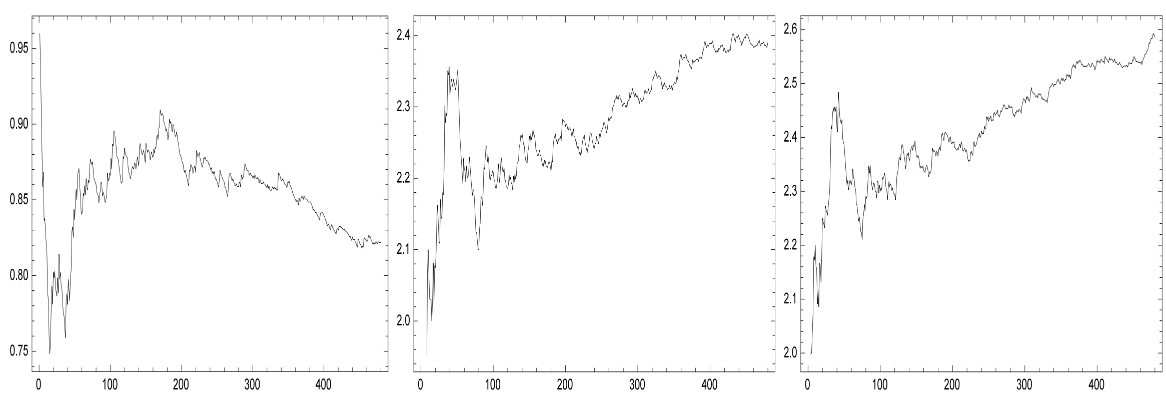

FIGURE 5.3. Estimators $\widehat{\eta}(k)$, $\widehat{K}(1.2;k)$ and $\widehat{K}(1.3;k)$ for copula (C) are plotted against $k = \{21,\ldots,500\}$.

Note that the behaviour of $\hat{K}$ (as shown in Figures 5.1-5.4) follows a similar pattern when k changes. For small values of k, the estimator behaves erratically due to small sized samples, while for large values of k,

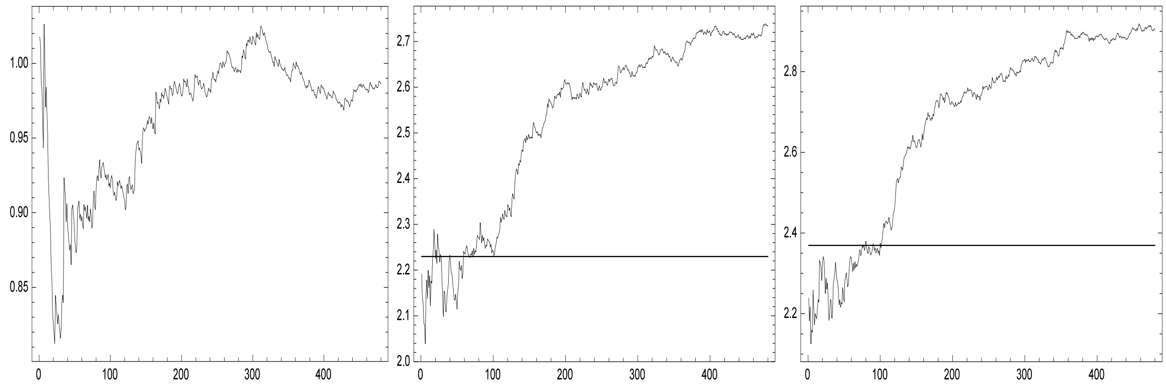

FIGURE 5.4. Estimators $\widehat{\eta}(k)$, $\widehat{K}(1.2;k)$ and $\widehat{K}(1.3;k)$ for copula (D) with $\xi = 1$ are plotted against $k = \{21, \ldots, 500\}$.

the rapidly growing bias is observed and poor estimates are obtained. In between those scenarios, there is a region of values for k, where the estimator is more stable and the actual estimate is chosen accordingly. Special attention should be paid to the theoretical optimal choice of k that is usually found by minimising the asymptotic mean squared error, but further research is needed to confirm this plausible choice. As a final comment, we would like to point out that further work is needed to show the consistency of our proposed estimator and other properties that will help us to produce confidence intervals. As explained in Asimit *et al.* (2015), we believe that a combination of some existing estimators (for example, the ones from Draisma *et al.*, 2004 and Asimit *et al.*, 2015) with our estimators would provide better statistical tools to distinguish between AD and AI.

6. Numerical Results

Some numerical examples are now given in order to justify the advantage of using our asymptotic approximations. As explained in Section 1, special attention has been given to evaluating the tail risk for a portfolio of risks for which the dependence is unknown or very little is known. According to our previous findings, we can answer the same questions by estimating the tail risk of a bivariate portfolio of risks where some partial information about dependence is known, namely, the tail dependence parameter c. Obviously, there is some uncertainty with the estimation of c, but confidence intervals can be found and in turn, the bounds are changed accordingly. Interestingly, we are able to find sharp upper bounds, which provide the most conservative scenario that a decision-maker might expected to encounter. The tail risk is based on one of the most popular risk measures, VaR. Its definition for a generic risk rv Z at a confidence level q is

$$VaR_q(Z) := \inf_t \left\{\mathbb{P}(Z \leq t) \geq q\right\}.$$

It is first assumed that X and Y are identically distributed Pareto rv's such that $\mathbb{P}(X > x) = (1 + x)^{-\alpha}$ for all $x \geq 0$. Lemma 2.1 of Asimit *et al.* (2011) and equation (5.1) show that

$$\left(\underline{K}(\alpha, c)\right)^{1/\alpha} \leq \lim_{q \downarrow 0} \frac{VaR_{1-q}(X + Y)}{VaR_{1-q}(X)} \leq \left(\overline{K}(\alpha, c)\right)^{1/\alpha}.$$

These bounds are plotted in Figure 6.1.

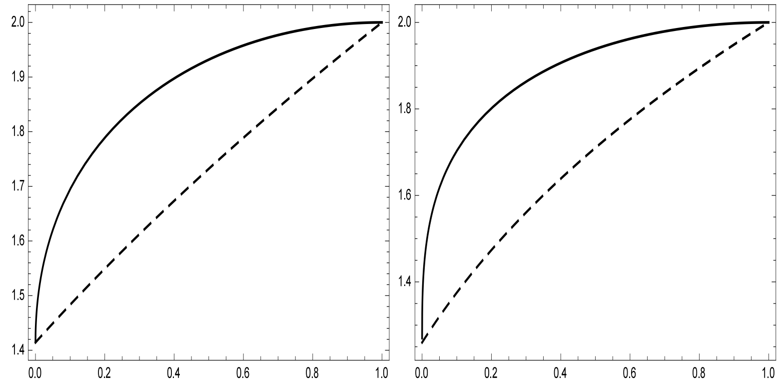

FIGURE 6.1. The upper bound (solid line) and lower bound (dashed line) for the ratio of $VaR_{1-q}(X+Y)/VaR_{1-q}(X)$ as a function of c with $\alpha = 2$ (left) and $\alpha = 3$ (right).

While both bounds are informative, the decision-maker is more keen to find the worst possible case, i.e. the upper bound. It is well-known that under AD, the joint tail behaviour exhibits the lower orthant property and therefore, we can compare our results with the one obtained by Bignozzi *et al.* (2015). Their Theorem 3.1 tells us that

$$VaR_{1-q}(X+Y) \leq \inf_{1-q \leq s \leq 1} \left\{ VaR_s(X) + VaR_{(1-q)/s}(X) \right\} = 2VaR_{\sqrt{1-q}}(X),$$

since the objective function from above is convex and symmetric due to the Pareto assumption. The VaR ratio upper bound found in Bignozzi *et al.* (2015) is depicted in Figure 6.2. Comparing Figures 6.1 and 6.2, it becomes clear that it is more advantageous to use our bounds if one has knowledge about the tail dependence parameter. The same conclusion can be drawn for the VaR ratio lower bounds, since the lower bound from Bignozzi *et al.* (2015) is 1 (see their Example 3.1). Recall that our asymptotic approximations displayed in Figure 6.1 depend only on c, but remain unchanged for different values of q. In turn, the alternative upper bounds from Figure 6.2 are only sensitive to changes of q.

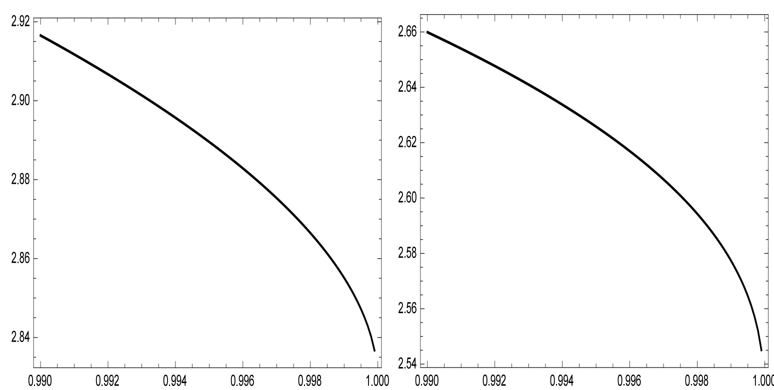

FIGURE 6.2. The upper bound for the ratio of $VaR_{1-q}(X+Y)/VaR_{1-q}(X)$ as a function of $1-q$ with $\alpha = 2$ (left) and $\alpha = 3$ (right).

Knowing the most and least conservative scenarios, it is interesting to understand how wide our confidence interval is. Thus, we plot in Figure 6.3 the relative spread (difference between the VaR ratio upper and lower bounds) based on our results (left) and Bignozzi *et al.* (2015) (right). Once again, our bounds are tighter since we include an additional piece of information about dependence, but it is fair to mention that our approach works only in the bivariate case. It is also worth mentioning that the bounds are less spread for relatively small values and large values of c, since the uncertainty with the tail dependence is reduced in these case when X and Y have RV tails.

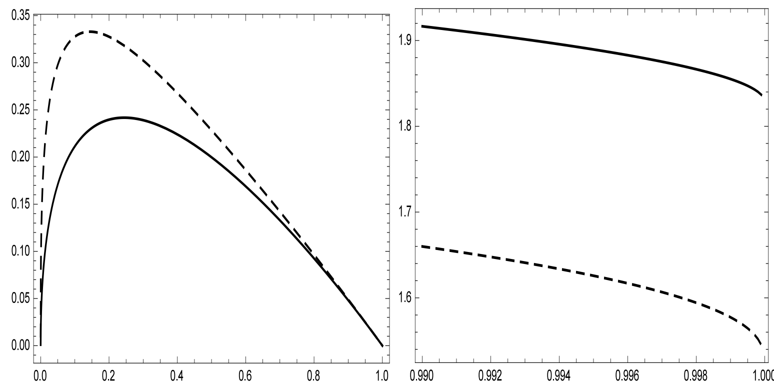

FIGURE 6.3. The relative spread for various values of c (left) and q (right) with $\alpha = 2$ (solid line) and $\alpha = 3$ (dashed line).

We now assume that the risks are exponentially distributed with mean 1 and perform the same analysis as before. Proposition 4.1i) and (4.3) yield that

$$c \leq \frac{\mathbb{P}(X+Y>2t)}{\mathbb{P}(X>t)} \leq \sqrt{c(2-c)}$$

and in turn, Lemma 2.4 of Asimit *et al.* (2011) implies that

$$2VaR_{1-q/c}(X) \leq VaR_{1-q}(X+Y) \leq 2VaR_{1-q/\sqrt{c(2-c)}}(X), \text{ for } q \text{ sufficiently close to 0.}$$

Our VaR ratios are depicted in Figure 6.4 and are calculated as above.

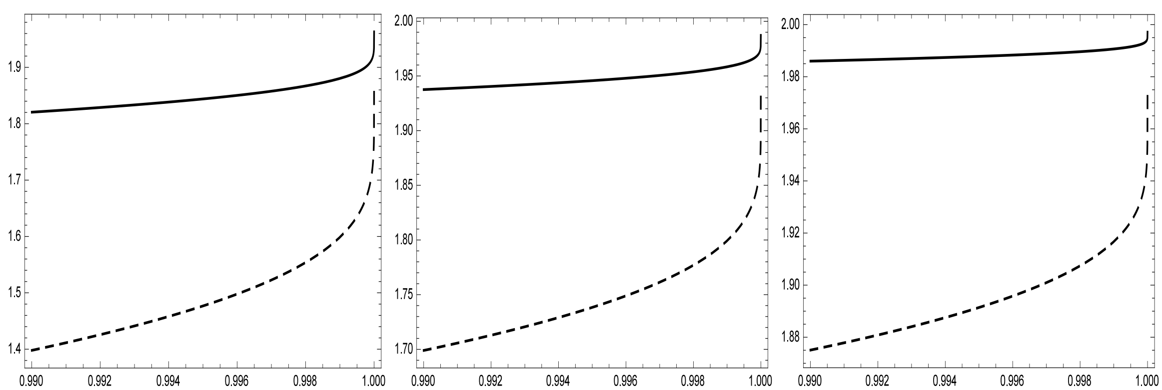

FIGURE 6.4. The upper bound (solid line) and lower bound (dashed line) for the ratio of $VaR_{1-q}(X+Y)/VaR_{1-q}(X)$ as a function of $1-q$ with $c = 0.25$ (left), $c = 0.5$ (middle) and $c = 0.75$ (right).

As before, the VaR ratio lower and upper bounds found in Bignozzi *et al.* (2015) are 1 and $2VaR_{\sqrt{1-q}}(X)$, respectively. Comparing these bounds with the values displayed in Figure 6.4, one can find that our bounds are tighter. Moreover, we compare in Figure 6.5 the relative spread of the VaR ratio based on our results (left) and Bignozzi *et al.* (2015) (right). Note that for larger values of c, our lower/upper bounds (see Figure 6.4) and relative spread (see Figure 6.5) increase.

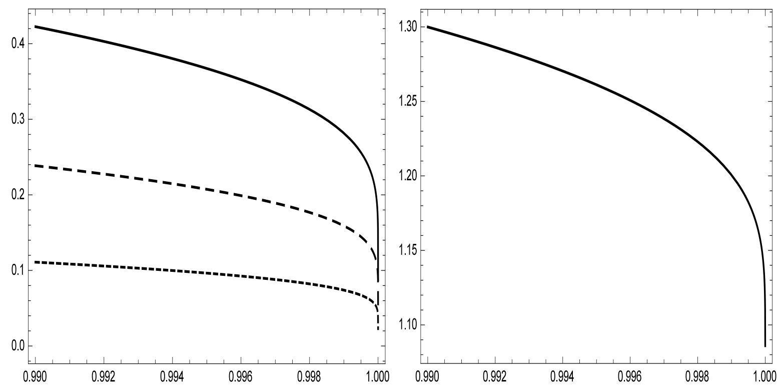

Figure 6.5. The relative spread as a function of q with (left) and without (right) asymptotic piece of information given by $c = 0.25$ (solid line), $c = 0.5$ (long dashed line) and $c = 0.75$ (short solid line).

Acknowledgments: We would like to thank to the Associate Editor and the three anonymous reviewers for their careful reading and helpful comments which improved the initial version of our manuscript.

7. Appendix

Proof of Proposition 3.1 It is sufficient to prove the properties for H_X, since the other case can be shown in the same fashion. Recall that H_X is right-continuous, as it is a cdf. Choose $\varepsilon > 0$ and for $0 < u < (1-\varepsilon)v$ we have

$$(1-\varepsilon)vH_X\left(\frac{u}{(1-\varepsilon)v}\right) = \frac{H\big(u,(1-\varepsilon)v\big)}{c} \leq \frac{H(u,v)}{c} = vH_X\left(\frac{u}{v}\right).$$

Write $x = \frac{u}{(1-\varepsilon)v}$ and in turn one may get that

$$\frac{H_X(x) - H_X\big((1-\varepsilon)x\big)}{H_X(x)} \leq \varepsilon.$$

Thus, H_X is a left-continuous function, and hence is a continuous function. Since it is non-decreasing, this means that it has a right derivative h_X, which must satisfy

$$\sup_{0<x<1} x\frac{h_X(x)}{H_X(x)} \leq 1.$$

In other words, we may write $H_X(x) = xJ_X(x)$, where J_X is a continuous, non-increasing function satisfying $J_X(1) = 1$. Taking this one step further, we observe for $0 < u < v$ that

$$ch_X\left(\frac{u}{v}\right) = \frac{\partial H(u,v)}{\partial u} \leq \frac{\partial H\big(u,v(1+\varepsilon)\big)}{\partial u} = ch_X\left(\frac{u}{v(1+\varepsilon)}\right).$$

Defining $x = \frac{u}{v(1+\epsilon)}$, one may get that $h_X((1+\varepsilon)x) \le h_X(x)$, and thus, h_X is right-continuous and non-increasing function on $(0,1)$. The left continuity of h_X is obtained in the same way was as the left continuity of H_X above.

Let (U,V) be two rv's on $[0,1]$ with joint cdf $G = H/c$, where H is defined in (2.2). For any $0 < x < 1$, equation (2.2) yields

$$\mathbb{P}(U,V \le x, U > V) = \int_0^x \left(H_Y(1) - h_Y(1-)\right)\, dz = \big(1 - h_Y(1-)\big)x.$$

Similarly, $\mathbb{P}(U,V \le x, U < V) = \big(1 - h_X(1-)\big)x$. These and the fact that $G(x,x) = x$ imply

$$\mathbb{P}(U = V \le x) = \big(h_X(1-) + h_Y(1-) - 1\big)x.$$

Thus, one may choose $d = \mathbb{P}(U = V) = h_X(1-) + h_Y(1-) - 1$, which clearly satisfies $d \in [0,1]$.

It only remains to justify $h_X(0+) = h_Y(0+) = 1/c$. Assume that the random vector (Z,T) has survival copula $\widehat{C}$ and $Z, T \in \mathcal{RV}_{-\alpha}$ are identically distributed and positive rv's. Clearly, relation (2.4) implies that

$$\lim_{t\to\infty} \mathbb{P}(T > ty | Z > t) = y^{-\alpha} H(y^\alpha, 1), \ \text{ for all } y < 1,$$

and since the limit is continuous, the limit holds uniformly in y as a result of Theorem 1.11 of Petrov (1995). Thus,

$$1 = \mu_F\big((1,\infty)\times(0,\infty]\big) = \lim_{y\downarrow 0} y^{-\alpha} H\big(y^\alpha, 1\big) = c \lim_{y\downarrow 0} \frac{H_X\big(y^\alpha\big)}{y^\alpha} = c\, h_X\big(0+\big),$$

which completes the proof.

Proof of Proposition 3.2 It is first proved that $x^{-1}H_X(x)$ is a non-increasing function in $x \in (0,1]$. Clearly, for any x

$$\frac{d}{dx}\frac{H_X(x)}{x} = -\frac{H_X(x)}{x^2} + \frac{h_X(x)}{x} = -\frac{1}{x^2}\int_0^x h_X(y)\, dy + \frac{h_X(x)}{x} = \frac{1}{x^2}\int_0^x \big(h_X(x) - h_X(y)\big)\, dy \le 0,$$

since h_X is non-increasing. The mirror result for H_Y can be shown in a similar manner.

Next, we show that $J(\cdot)$ is a valid cdf on $[0,1]^2$. Note that $J(x,0) = J(0,x) = 0$ for any $0 \le x \le 1$ and that $J(1,1) = 1$. It only remains to establish that

$$J\big(x_2,y_2\big) - J\big(x_2,y_1\big) - J\big(x_1,y_2\big) + J\big(x_1,y_1\big) \ge 0, \ \text{ for all } 0 \le x_1 \le x_2 \le 1,\ 0 \le y_1 \le y_2 \le 1.$$

We demonstrate that $\frac{\partial J}{\partial x}(x,y_2) - \frac{\partial J}{\partial x}(x,y_1) \ge 0$ for each x, from which the required result follows by integration. There are three cases to consider; in each case the result relies on the fact that h_X and h_Y are non-increasing, whilst in case ii) we also use the fact that $h_X(1-) + h_Y(1-) = 1 + d$.

i) Suppose first that $x \le y_1 \le y_2$. If $x < y_1$, the following holds

$$\frac{\partial J}{\partial x}(x,y_2) - \frac{\partial J}{\partial x}(x,y_1) = h_X\left(\frac{x}{y_2}\right) - h_X\left(\frac{x}{y_1}\right) \ge 0.$$

The remaining case in which $x = y_1$ is further shown. The right derivatives of J with respect to x at (x, y_1) is given by

$$\begin{aligned}\lim_{\epsilon\downarrow 0}\frac{J(x+\epsilon)-J(x,x)}{\epsilon} &= \lim_{\epsilon\downarrow 0}\frac{(x+\epsilon)H_Y\big(x/(x+\epsilon)\big)-x}{\epsilon}\\ &= \lim_{\epsilon\downarrow 0}\left(H_Y\big(x/(x+\epsilon)\big)+\frac{H_Y\big(1-\epsilon/(x+\epsilon)\big)-1}{\epsilon/(x+\epsilon)}\frac{x}{x+\epsilon}\right)\\ &= 1-h_Y(1-).\end{aligned}$$

Similarly, the left derivative becomes

$$\lim_{\epsilon\downarrow 0}\frac{J(x,x)-J(x-\epsilon)}{\epsilon} = h_X(1-),$$

which is always larger than the right derivative. Thus, our claim is true since $h_X\left(\frac{x}{y_2}\right)-h_X(1-)\geq 0$.
ii) Now, suppose that $y_1 < x \leq y_2$. If $y_1 < x < y_2$, we have that

$$\begin{aligned}&\frac{\partial J}{\partial x}(x,y_2)-\frac{\partial J}{\partial x}(x,y_1)\\ &= h_X\left(\frac{x}{y_2}\right)-\left[H_Y\left(\frac{y_1}{x}\right)-\frac{y_1}{x}h_Y\left(\frac{y_1}{x}\right)\right]\\ &= h_X\left(\frac{x}{y_2}\right)-h_X(1-)+1-h_Y(1-)-\left[H_Y\left(\frac{y_1}{x}\right)-\frac{y_1}{x}h_Y\left(\frac{y_1}{x}\right)\right]+d\\ &= h_X\left(\frac{x}{y_2}\right)-h_X(1-)+\int_{y/x}^{1}\big(h_Y(u)-h_Y(1-)\big)\,du+\frac{y_1}{x}\left(h_Y\left(\frac{y_1}{x}\right)-h_Y(1-)\right)+d\\ &\geq 0.\end{aligned}$$

As in setting i), the case $x = y_2$ is justified as follows:

$$\begin{aligned}\frac{\partial J}{\partial x}(x,y_2)-\frac{\partial J}{\partial x}(x,y_1) &\geq 1-h_Y(1-)-\left[H_Y\left(\frac{y_1}{x}\right)-\frac{y_1}{x}h_Y\left(\frac{y_1}{x}\right)\right]\\ &= \int_{y/x}^{1}\big(h_Y(u)-h_Y(1-)\big)\,du+\frac{y_1}{x}\left(h_Y\left(\frac{y_1}{x}\right)-h_Y(1-)\right)\\ &\geq 0.\end{aligned}$$

iii) Next, suppose $y_1 < y_2 < x$ and it yields

$$\begin{aligned}\frac{\partial J}{\partial x}(x,y_2)-\frac{\partial J}{\partial x}(x,y_1) &= \left[H_Y\left(\frac{y_2}{x}\right)-\frac{y_2}{x}h_Y\left(\frac{y_2}{x}\right)\right]-\left[H_Y\left(\frac{y_1}{x}\right)-\frac{y_1}{x}h_Y\left(\frac{y_1}{x}\right)\right]\\ &= \int_{y_1/x}^{y_2/x}\left(h_Y(u)-h_Y\left(\frac{y_2}{x}\right)\right)du+\frac{y_1}{x}\left(h_Y\left(\frac{y_1}{x}\right)-h_Y\left(\frac{y_2}{x}\right)\right)\\ &\geq 0.\end{aligned}$$

We also need to verify that C satisfies (2.1), where $C(u,v) := J\big(H_X^{-1}(u), H_Y^{-1}(v)\big)$ is the copula of J, which exists due to Sklar's Theorem. Elementary arguments may help to justify that

$$\lim_{u\downarrow 0}\frac{H_X^{-1}(u)}{u} = \lim_{u\downarrow 0}\frac{H_Y^{-1}(u)}{u} = c$$

as a result of $h_X(0+) = h_Y(0+) = 1/c$ and the fact that H_X^{-1} and H_Y^{-1} are non-decreasing. Thus, if $x < y$ and u is sufficiently close to 0, then $H_X^{-1}(ux) \leq H_Y^{-1}(uy)$ and we have

$$\lim_{u\downarrow 0}\frac{J\left(H_X^{-1}(ux), H_Y^{-1}(uy)\right)}{u} = \lim_{u\downarrow 0}\frac{H_Y^{-1}(uy)}{u}H_X\left(\frac{H_X^{-1}(ux)}{H_Y^{-1}(uy)}\right) = cyH_X\left(\frac{x}{y}\right) = cJ(x,y).$$

The case $y \leq x$ can be justified similarly, which completes the proof.

Proof of Theorem 4.1 i) Note first that equation (2.6) yields

$$\lim_{t\to\infty} \frac{\mathbb{P}(X+bY>t)}{\mathbb{P}(X>t)} = \mu_F\big(A_s\big), \quad \text{where } A_s := \big\{(x,y) : x+by>1\big\},$$

since $\mu_F\big(\partial A_s\big) = 0$ (for details, see Kortschak and Albrecher, 2009). An alternative proof is given in Resnick (2007) and Klüppelberg and Resnick (2008). Now, using equations (2.2), (2.4) repeatedly and obvious changes of variables, we get

$$\begin{aligned}
&\mu_F\big(A_s\big)\\
&= \mu_F\big((x,y) : x+by>1, x\geq y\big) + \mu_F\big((x,y) : x+by>1, x<y\big)\\
&= \int_{\frac{1}{1+b}}^{1} \mu_F\left(dx \times \left(\frac{1-x}{b}, x\right]\right) + \int_1^\infty \mu_F\big(dx\times(0,x]\big) + c\,d\,(1+b)^\alpha\\
&\quad + \int_{\frac{1}{1+b}}^{1/b} \mu_F\big((1-by,y]\times dy\big) + \int_{\frac{1}{b}}^{\infty} \mu_F\big((0,y]\times dy\big)\\
&= c\,\alpha \int_{\frac{1}{1+b}}^{1} x^{-\alpha-1} h_X\left(\frac{1-x}{ax}\right)\,dx + \mu_F\big((1,\infty]\times(0,\infty]\big) - \int_{\frac{1}{1+b}}^{\infty} \mu_F\big(dx\times(x,\infty]\big) + c\,d\,(1+b)^\alpha\\
&\quad + c\,\alpha \int_{\frac{1}{1+b}}^{\frac{1}{b}} y^{-\alpha-1} h_Y\left(\left(\frac{1-by}{y}\right)^\alpha\right)\,dy + \mu_F\big((0,\infty]\times(1/b,\infty]\big) - \int_{\frac{1}{1+b}}^{\infty} \mu_F\big((y,\infty]\times dy\big)\\
&= b\,c \int_0^1 z^{1/\alpha-1}\left(1+bz^{1/\alpha}\right)^{\alpha-1} h_X(z)\,dz + 1 - c\,h_X(1-)(1+b)^\alpha + c\,d\,(1+b)^\alpha\\
&\quad + b\,c \int_0^1 z^{1/\alpha-1}\left(b+z^{1/\alpha}\right)^{\alpha-1} h_Y(z)\,dx + b^\alpha - c\,h_Y(1-)(1+b)^\alpha,
\end{aligned}$$

which concludes (4.1) since $h_X(1-)+h_Y(1-) = 1+d$.

ii) The vague convergence from relation (2.6) yields

$$\lim_{t\to\infty} \frac{\mathbb{P}(XY>t^2)}{\mathbb{P}(X>t)} = \mu_F\big(A_p\big), \quad \text{where } A_p := \big\{(x,y) : xy>1\big\}, \tag{7.1}$$

as long as $\mu_F\big(\partial A_p\big) = 0$. Note that no mass is put in neighborhoods of ∞, and therefore, the only possible way to put same mass on the boundary of A_p is only on the curve $\{xy=1\}$. Assume that $\mu_F\Big(\big\{(x,1/x), x>0\big\}\Big) = m > 0$. Thus,

$$\begin{aligned}
\mu_F\Big(\big\{(x,y), 1<xy\leq 2\big\}\Big) &\geq \mu_F\left(\bigcup_{q\in Q\bigcap(1,2]} \big\{(x,y), xy=q\big\}\right)\\
&= \sum_{q\in Q\bigcap(1,2]} \mu_F\Big(\big\{(x,y), xy=q\big\}\Big)\\
&= \mu_F\Big(\big\{(x,y), xy=1\big\}\Big) \sum_{q\in Q\bigcap(1,2]} q^{-\alpha/2}\\
&= \infty,
\end{aligned}$$

where the second last step is due to the fact that $\mu_F(xA) = x^{-\alpha}\mu_F(A)$ holds for any relatively compact set, which contradicts our assumption that $m>0$, since μ_F is a Radon measure.

Some algebra that involves multiple use of equations (2.2), (2.4) and some obvious changes of variables lead to

$$
\begin{aligned}
&\mu_F\big(A_p\big)\\
&= \mu_F\big((x,y): xy>1, x\geq y\big) + \mu_F\big((x,y): xy>1, x<y\big)\\
&= \int_1^\infty \mu_F\Big(dx\times\big(1/x,\infty\big]\Big) - \int_1^\infty \mu_F\big(dx\times(x,\infty]\big) + c\,d\\
&\quad + \int_1^\infty \mu_F\big(\big(1/y,\infty\big]\times dy\big) - \int_1^\infty \mu_F\big((y,\infty]\times dy\big)\\
&= c\,\alpha\int_1^\infty x^{-\alpha-1}h_X\left(x^{-2\alpha}\right)\,dx - c h_X(1-) + c\,d + c\,\alpha\int_1^\infty y^{-\alpha-1}h_Y\left(y^{-2\alpha}\right)\,dy - c h_Y(1-)\\
&= \frac{c}{2}\int_0^1 z^{-1/2}h_X(z)\,dz + \frac{c}{2}\int_0^1 z^{-1/2}h_Y(z)\,dz - c,
\end{aligned}
$$

since $h_X(1-)+h_Y(1-)=1+d$. The latter and relation (7.1) conclude part ii).

iii) A consequence of equation (2.7) is that

$$\lim_{t\to\infty}\frac{\mathbb{P}(X+Y>2t)}{\mathbb{P}(X>t)} = \mu_G\big(B\big), \quad \text{where } B := \big\{(x,y): x+y>0\big\},$$

since $\mu_G\big(\partial B\big)=0$ (for details, see Kortschak and Albrecher, 2009). Note that there is an alternative approach, which is given in Klüppelberg and Resnick (2008). As before, by multiple use of equations (2.2), (2.5) and obvious changes of variables, we get

$$
\begin{aligned}
\mu_G(B) &= \mu_G\big((x,y): x+y>0, x>0\big) + \mu_G\big((x,y): x+y>0, y>0\big) - \mu_G\big((0,\infty]\times(0,\infty]\big)\\
&= \int_0^\infty \mu_G\big(dx\times(-x,\infty]\big) - \int_0^\infty \mu_G\big((-y,\infty]\times dy\big) - c\\
&= c\int_0^\infty e^{-x}h_X\left(e^{-2x}\right)\,dx + c\int_0^\infty e^{-y}h_X\left(e^{-2y}\right)\,dy - c\\
&= \frac{c}{2}\int_0^1 z^{-1/2}h_X(z)\,dz + \frac{c}{2}\int_0^1 z^{-1/2}h_Y(z)\,dz - c,
\end{aligned}
$$

The proof is now complete.

The first step in the proof of Theorem 4.2 is the next lemma. Let $\overline{\mathcal{H}}_{\xi,c,d}$ be the collection of cdf's whose densities h are elements of $\mathcal{H}_{\xi,c,d}$.

Lemma 7.1. *Suppose $b:[0,1]\to[0,\infty)$ satisfies $\int_0^1 b(x)\,dx<\infty$. Then*

$$\inf_{H\in\overline{\mathcal{H}}_{\xi,c,d}}\int_0^1 b(x)H(x)\,dx = \int_0^1 x b(x)\,dx, \tag{7.2}$$

$$\sup_{H\in\overline{\mathcal{H}}_{\xi,c,d}}\int_0^1 b(x)H(x)\,dx = \int_0^1 b(x)H^*(x;\xi,c)\,dx \tag{7.3}$$

where

$$H^*(x;\xi,c,d) = \begin{cases} \frac{x}{c} & \text{if } 0\leq x\leq \frac{(1-\xi)c}{1-\xi c}\\ 1-\xi+\xi x & \text{if } \frac{(1-\xi)c}{1-\xi c}\leq x\leq 1\end{cases}$$

Proof of Lemma 7.1 The proof of (7.3) is straightforward, since $H^*(\cdot;\xi,c,d)\in\overline{\mathcal{H}}_{\xi,c,d}$ and it follows from Corollary 3.1 and the fact that $H(x)\leq x/c$ that $H(\cdot)\leq H^*(\cdot;\xi,c,d)$ for any $H\in\overline{\mathcal{H}}_{\xi,c,d}$.

In addition, since $H(x) \geq x$, it is clear that the infimum in (7.2) is at least as large as $\int_0^1 xb(x)\,dx$. What remains is the proof of equality, which we accomplish by demonstrating a sequence of functions $H_n \in \overline{\mathcal{H}}_{\xi,c,d}$ such that $\int_0^1 b(x)H_n(x)\,dx \to \int_0^1 xb(x)\,dx$.

Define

$$H_n(x) = \begin{cases} \frac{x}{c} - \frac{n}{2}\left(\frac{1}{c}-1\right)^2 x^2 & \text{if } 0 \leq x \leq x_L(n) \\ x + \frac{1}{2n} & \text{if } x_L(n) \leq x \leq x_U(n) \\ 1-\xi(1-x) - \frac{n}{2}(1-\xi)^2(1-x)^2 & \text{if } x_U(n) \leq x \leq 1 \end{cases},$$

where $x_L(n) = \left[n\left(\frac{1}{c}-1\right)\right]^{-1}$ and $x_U(n) = 1-[n(1-\xi)]^{-1}$. It is not difficult to verify that $H_n \in \overline{\mathcal{H}}_{\xi,c,d}$. Further,

$$\begin{aligned} \int_0^1 b(x)\,|x - H_n(x)|\,dx &= \int_0^{x_L(n)} b(x)\left|x - \frac{x}{c} + \frac{n}{2}\left(\frac{1}{c}-1\right)^2 x^2\right| dx + \int_{x_L(n)}^{x_U(n)} \frac{1}{2n} b(x)\,dx \\ &\quad + \int_{x_U(n)}^1 b(x)\left|x - 1 + \xi(1-x) + \frac{n}{2}(1-\xi)^2(1-x)^2\right| dx \\ &= \left(1-\frac{1}{c}\right)\int_0^{x_L(n)} xb(x)\left[1+\frac{n}{2}\left(\frac{1}{c}-1\right)x\right] dx + \frac{1}{2n}\int_{x_L(n)}^{x_U(n)} b(x)\,dx \\ &\quad + (1-\xi)\int_{x_U(n)}^1 (1-x)b(x)\left|-1+\frac{n}{2}(1-\xi)(1-x)\right| dx \\ &\leq \frac{3}{2n}\int_0^{x_L(n)} b(x)\,dx + \frac{1}{2n}\int_{x_L(n)}^{x_U(n)} b(x)\,dx + \frac{3}{2n}\int_{x_U(n)}^1 b(x)\,dx \\ &\leq \frac{3}{2n}\int_0^1 b(x)\,dx, \end{aligned}$$

which justifies our results in full.

Proof of Theorem 4.2 Suppose $a : [0,1] \to \mathbb{R}$ is continuous and monotone increasing; then it has a non-negative derivative $b = a'$ almost everywhere. Thus, the proof becomes straightforward and it follows from integration by parts and Lemma 7.1.

Proof of Proposition 4.1 i) The supremum follows immediately from Theorem 4.2. For the infimum, we need to minimise

$$\begin{aligned} &\min_{0\leq d\leq\xi\leq 1} \int_0^1 a(x)\big(h^*(x;\xi,c,d) + h^*(x;1+d-\xi,c,d)\big)\,dx \\ &= \min_{0\leq d\leq\xi\leq 1}\left\{\frac{1}{c}A\left(\frac{(1-\xi)c}{1-\xi c}\right) + \xi\left(A(1) - A\left(\frac{(1-\xi)c}{1-\xi c}\right)\right) + \frac{1}{c}A\left(\frac{(\xi-d)c}{1+d-(1-\xi)c}\right)\right. \\ &\qquad \left. + (1+d-\xi)\left(A(1) - A\left(\frac{(\xi-d)c}{1+d-(1-\xi)c}\right)\right)\right\} \end{aligned} \tag{7.4}$$

It is first shown that the objective function is convex in ξ for any fixed d and attains its minimum at $\xi^*(d) = (1+d)/2$. We need to minimise $g(\xi) + g(1+d-\xi)$ on $[d,1]$, where $g : [d,1] \to \Re$ with

$$g(\xi) := \xi A(1) + \left(\frac{1}{c} - \xi\right) A\left(\frac{(1-\xi)c}{1-\xi c}\right).$$

Clearly,

$$\frac{\partial g}{\partial \xi} = A(1) - A\big(r(\xi)\big) - \big(1 - r(\xi)\big)a\big(r(\xi)\big), \quad \text{with} \quad r(\xi) := \frac{(1-\xi)c}{1-\xi c}.$$

Recall that $A(s) + (1-s)a(s)$ is a strictly increasing function, since its derivative is $(1-s)a'(s) > 0$. The latter and the fact that r is a strictly decreasing function suggest that $\frac{\partial g}{\partial \xi}$ is strictly increasing in ξ.

Consequently, the objective function from (7.4) is decreasing and increasing in ξ if and only if $\xi \le (1+d)/2$ and $\xi \ge (1+d)/2$, respectively, and therefore the minimum is reached at $\xi^*(d) = (1+d)/2$.

We now vary $d \in [0,1]$ in order to globally minimise (7.4). That is, we ought to minimise $2g\big((1+d)/2\big)$ or equivalently

$$min_{1\le t\le \frac{2-c}{2-2c}} h(t) := t\left(A\left(1-\frac{1}{t}\right) - A(1)\right) \text{ by denoting } t = \frac{\frac{1}{c} - \frac{1+d}{2}}{\frac{1}{c}-1}.$$

Now, $h''(t) = t^{-3}a'(1-1/t) > 0$ and it yields

$$h'(t) \le h'(1) = a(0) - A(1) = \int_0^1 \big(a(0) - a(y)\big)\, dy \le 0$$

and thus, the global minimum is attained when $t^* = \frac{2-c}{2-2c}$. Therefore, the infimum in (7.4) is $A(1) + \frac{1}{r^*}A(r^*)$ and is obtained at $\big(d^*, \xi^*\big) = (0, 1/2)$.

Finally, if a is monotone decreasing, we can apply the Proposition to $\tilde{a}(x) = -a(x)$ to obtain the required result.

ii) We only prove the infimum since supremum can be shown in a similar manner. The first stage optimisation problems are solved via Theorem 4.2 and we have

$$\begin{aligned} \inf_{(h_X,h_Y)\in(\mathcal{H}_{\xi,c,d},\mathcal{H}_{1+d-\xi,c,d})} J(a_X,a_Y) &= \inf_{0\le d\le\xi\le 1}\int_0^1 \Big(a_X(x)h^*(x;\xi,c,d) + a_Y(x)\Big)dx \\ &= \inf_{0\le d\le\xi\le 1} g(\xi) + A_Y(1), \end{aligned} \tag{7.5}$$

where g is defined in the proof of part i). For any fixed d, the above function is increasing in ξ since

$$g'(\xi) \ge g'(d) = A(1) - A\big(r(d)\big) - \big(1 - r(d)\big)a\big(r(d)\big) = \int_{r(d)}^1 a(y)\, dy - \big(1-r(d)\big)a\big(r(d)\big) \ge 0$$

by keeping in mind that g' and a are increasing functions. Thus, the minima is attained at $\xi^*(d) = d$ and by varying $d \in [0,1]$, it is not difficult to find that the infimum in (7.5) is obtained when $d^* = \xi^* = 0$, which completes the proof.

References

[1] Asimit, V., Furman, E., Tang, Q. and Vernic, R. 2011. "Asymptotics for Risk Capital Allocations based on Conditional Tail Expectation", *Insurance: Mathematics and Economics*, 49(3), 310–324.

[2] Asimit, A.V., Gerrard, R., Hou, Y. and Peng, L. 2015. "Tail Dependence Measure for Modeling Financial Extreme Co-movements", *Journal of Econometrics*, Forthcoming.

[3] Alink, S., Löwe, M. and Wüthrich, M.V. 2007. "Diversification for General Copula Dependence", *Statistica Neerlandica*, 61(4), 446–465.

[4] Bernard, C., Denuit, M. and Vanduffel, S. 2014. "Measuring Portfolio Risk Under Partial Dependence Information", Working paper. Available at `http://ssrn.com/abstract=2406377`.

[5] Bernard, C., Jiang, X. and Wang, R. 2014. "Risk Aggregation with Dependence Uncertainty", *Insurance: Mathematics and Economics*, 54, 93–108.

[6] Bernard, C., Rüschendorf, L. and Vanduffel, S. 2015. "Value-at-Risk Bounds with Variance Constraints", *Journal of Banking and Finance*, Forthcoming.

[7] Bignozzi, V., Puccetti, G. and Rüschendorf, L. 2015. "Reducing Model Risk via Positive and Negative Dependence Assumptions", *Insurance: Mathematics and Economics*, 61, 17–26.

[8] de Haan, L. and Ferreira, A. 2006. Extreme Value Theory: An Introduction. *Springer-Verlag, New York.*

[9] Draisma, G., Drees, H., Ferreira, A. and de Haan, L. 2004. "Bivariate Tail Estimation: Dependence in Asymptotic Independence", *Bernoulli* 10(2), 251–280.

[10] Embrechts, P., Höing, A. and Puccetti, G. 2005. "Worst VaR Scenarios", *Insurance: Mathematics and Economics*, 37(1), 115–134.

[11] Embrechts, P., Klüppelberg, C. and Mikosch, T. 1997. Modeling Extremal Events for Insurance and Finance. *Springer-Verlag, Berlin.*

[12] Embrechts, P., Nešlehová, J. and Wüthrich, M.V. 2009. "Additivity Properties for Value-at-Risk under Archimedean Dependence and Heavy-tailedness", *Insurance: Mathematics and Economics*, 44(2), 164–169.

[13] Embrechts, P. and Puccetti, G. 2006a. "Bounds for Functions of Dependent Risks", *Finance and Stochastics*, 10(3), 341–352.

[14] Embrechts, P. and Puccetti, G. 2006b. "Bounds for Functions of Multivariate Risks", *Journal of Multivariate Analysis*, 97(2), 526–547.

[15] Embrechts, P., Puccetti, G. and Rüschendorf, L. 2013. "Model Uncertainty and VaR Aggregation", *Journal of Banking and Finance*, 37(8), 2750–2764.

[16] Fisher, R.A. and Tippett, L.H.C. 1928. "Limiting Forms of the Frequency Distribution of the Largest or Smallest Member of a Sample", *Mathematical Proceedings of the Cambridge Philosophical Society*, 24(2), 180–190.

[17] Furman, E., Su, J. and Zitikis, R. 2014. "Paths and Indices of Maximal Tail Dependence", Working paper. Available at `http://ssrn.com/abstract=2433347`.

[18] Goegebeur, Y. and Guillou, A. 2012. "Asymptotically Unbiased Estimation of the Coefficient of Tail Dependence", *Scandinavian Journal of Statistics*, 40(1), 174–189.

[19] Haug, S., Klüppelberg, C. and Peng, L. 2011. "Statistical Models and Methods for Dependence in Insurance Data", *Journal of the Korean Statistical Society*, 40(2), 125–139.

[20] Hua, L. and Joe, H. 2011. "Tail Order and Intermediate Tail Dependence of Multivariate Copulas", *Journal of Multivariate Analysis 102(10), 1454–1471.*

[21] Juri, A. and Wüthrich, M. V. 2003. "Tail Dependence from a Distributional Point of View", *Extremes*, 6(3), 213–246.

[22] Klüppelberg, C. and Resnick, S.I. 2008. "The Pareto Copula, Aggregation of Risks, and the Emperor's Socks", *Journal of Applied Probability*, 45(1), 67–84.

[23] Kortschak, D. and Albrecher, H. 2009. "Asymptotic Results for the Sum of Dependent Non-identically Distributed Random Variables", *Methodology and Computing in Applied Probability*, 11(3), 279–306.

[24] Ledford, A.W. and Tawn, J. 1996. "Statistics for Near Independence in Multivariate Extreme Values", *Biometrika*, 83(1), 169–187.

[25] Ledford, A.W. and Tawn, J. 1997. "Modelling Dependence within Joint Tail Regions", *Journal of the Royal Statistical Society. Series B (Methodological)*, 59(2), 475–499.

[26] Nelsen, R. B. 2006. An Introduction to Copulas, 2nd ed. *Springer-Verlag, New York.*

[27] Peng, L. 1999. "Estimation of the Coefficient of Tail Dependence in Bivariate Extremes", *Statistics and Probability Letters*, 43(4), 399-409.

[28] Petrov, V. V. 1995. Limit Theorems of Probability Theory. Sequences of Independent Random Variables. *Oxford University Press, New York.*

[29] Resnick, S.I. 2007. Heavy-Tail Phenomena: Probabilistic and Statistical Modeling. *Springer-Verlag, New York.*

[30] Resnick, S.I. 1987. Extreme Values, Regular Variation and Point Processes. *Springer-Verlag, New York.*

[31] Sklar, A. 1959. "Fonctions de répartition à n dimensions et leurs marges", *Publications de l'Institut de Statistique de l'Université de Paris, 8, 229–231.*

[32] Wang, R., Peng, L. and Yang, J. 2013. "Bounds for the Sum of Dependent Risks and Worst Value-at-Risk with Monotone Marginal Densities", *Finance and Stochastics*, 17(2), 395–417.