Finance and Stochastics manuscript No.
(will be inserted by the editor)

Bounds for the Sum of Dependent Risks and Worst Value-at-Risk with Monotone Marginal Densities

Ruodu Wang · Liang Peng · Jingping Yang

the date of receipt and acceptance should be inserted later

Abstract In quantitative risk management, it is important and challenging to find sharp bounds for the distribution of the sum of dependent risks with given marginal distributions, but an unspecified dependence structure. These bounds are directly related to the problem of obtaining the worst Value-at-Risk of the total risk. Using the idea of the complete mixability, we provide a new lower bound for any given marginal distributions and give a necessary and sufficient condition for the sharpness of this new bound. For the sum of dependent risks with an identical distribution, which has either a monotone density or a tail-monotone density, the explicit values of the worst Value-at-Risk and the bounds on the distribution of the total risk are obtained. Some examples are given to illustrate the new results.

Keywords Complete mixability · Monotone density · Sum of dependent risks · Value-at-Risk

Mathematics Subject Classification (2010) 60E05, 60E15

JEL Classifications G10

1 Introduction

Let $\mathbf{X} = (X_1, \cdots, X_n)$ be a risk vector with known marginal distributions $F_1, \cdots, F_n$, denoted as $X_i \sim F_i, i = 1, \cdots, n$ and let $S = X_1 + \cdots + X_n$ be the total risk. For the purpose of risk management, it is of importance to find the best-possible bounds for the distribution of the total risk S when the dependence structure is unspecified, namely

$$m_+(s) = \inf\{\mathbb{P}(S < s) : X_i \sim F_i,\ i = 1, \cdots, n\}, \tag{1.1}$$

R. Wang · L. Peng
School of Mathematics, Georgia Institute of Technology, Atlanta, GA 30332-0160, USA.
E-mail: ruodu.wang@math.gatech.edu; peng@math.gatech.edu

J. Yang
Corresponding author. LMEQF, Department of Financial Mathematics, Center for Statistical Science, Peking University, Beijing, 100871, China. E-mail: yangjp@math.pku.edu.cn

and

$$M_+(s) = \sup\{\mathbb{P}(S < s) : X_i \sim F_i,\ i = 1, \cdots, n\}. \tag{1.2}$$

See Embrechts and Puccetti [6] for discussions on such problems in risk management. Since techniques for handling $M_+(s)$ are very similar to those for $m_+(s)$, we shall focus on $m_+(s)$ in this paper.

First let us review some known results on $m_+(s)$. Rüschendorf [11] found $m_+(s)$ when all marginal distributions have the same uniform or binomial distribution; Denuit et al. [1] and Embrechts et al. [2] used copulas to yield the so-called *standard bounds*, which are no longer sharp for $n \geq 3$, and discussed some applications; Embrechts and Puccetti [4] provided a better lower bound when all marginal distributions are the same and continuous, and some results when partial information on the dependence structure is available; Embrechts and Höing [3] provided a geometric interpretation to highlight the shape of the dependence structures with the worst VaR scenarios; Embrechts and Puccetti [5] extended this problem to multivariate marginal distributions and provided results similar to the univariate case. In summary, for $n \geq 3$, exact bounds were only found for the homogenous case ($F_1 = \cdots = F_n = F$) in Rüschendorf [11] where F is uniform or binomial and in Wang and Wang [14] where F has a monotone density on its support and satisfies a mean condition. Besides the above results on $m_+(s)$, Rüschendorf [11] associated an equivalent dual optimization problem with the bounds for a general function of $X_1, \cdots, X_n$ instead of the total risk S.

The bounds $m_+(s)$ and $M_+(s)$ directly lead to the sharp bounds on quantile-based risk measures of S. A widely used measure is the so-called Value-at-Risk (VaR) at level α, defined as

$$\mathrm{VaR}_\alpha(S) = \inf\{s \in \mathbb{R} : \mathbb{P}(S \leq s) \geq \alpha\}.$$

The bound on the above VaR is called the worst Value-at-Risk scenario. Deriving sharp bounds for the worst VaR is of great interest in the recent research of quantitative risk management; see Embrechts and Puccetti [6] and Kaas et al. [8] for more details.

In this paper, we first provide a new lower bound on $m_+(s)$, which is easy to calculate. Using the idea of the jointly mixable distributions, we give a necessary and sufficient condition for this bound to be the true value of $m_+(s)$. See Section 2 for details. In Section 3 we employ a special class of copulas to find $m_+(s)$ and the worst Value-at-Risk when all the marginal distributions are identical and have a monotone or tail-monotone density. The methods are illustrated by some examples. Some conclusions are drawn in Section 4. Some proofs are put in the appendix.

2 Bounds for the sum with general marginal distributions

Throughout, we identify probability measures with the corresponding distribution functions. Let $\mathbf{X} = (X_1, \cdots, X_n)$ and $S = X_1 + \cdots + X_n$. For any distribution F, we use $F^{-1}(t) = \inf\{s \in \mathbb{R} : F(s) \geq t\}$ to denote the (generalized) inverse function and denote by $\tilde{F}_a$ the conditional distribution of F on $[F^{-1}(a), \infty)$ for $a \in [0, 1)$, i.e., $\tilde{F}_a(x) = \max\left\{\frac{F(x)-a}{1-a}, 0\right\}$ for $x \in \mathbb{R}$. It is straightforward to check that for $u \in [0, 1]$,

$\tilde{F}_a^{-1}(u) = F^{-1}((1-a)u+a)$. In addition, we define $\tilde{F}_1(x) = \lim_{a\to 1-} \tilde{F}_a(x)$. In this paper, no specific probability space is assumed and discussions are focused on distributions, since $m_+(s)$ only depends on s and the distributions $F_1, \cdots, F_n$.

2.1 General bounds

In this section, we will give a general lower bound on $m_+(s)$. Before showing this bound, we need some definitions and lemmas.

Definition 2.1 The random vector $\mathbf{X} = (X_1, \cdots, X_n)$ with marginal distributions $F_1, \cdots, F_n$ is called an **optimal coupling** for $m_+(s)$ if

$$\mathbb{P}(X_1 + \cdots + X_n < s) = m_+(s).$$

It is known that the optimal coupling for $m_+(s)$ always exists (see the introduction in Rüschendorf [12] for instance). The following lemma is **Proposition 3(c)** of Rüschendorf [11], which will be used later.

Lemma 2.2 *Suppose $F_1, \cdots, F_n$ are continuous. Then there exists an optimal coupling $\mathbf{X} = (X_1, \cdots, X_n)$ for $m_+(s)$ such that $\{S \geq s\} = \{X_i \geq F_i^{-1}(m_+(s))\}$ for each $i = 1, \cdots, n$.*

Next we introduce the concept of completely mixable and jointly mixable distributions.

Definition 2.3 (completely mixable and jointly mixable distributions)

1. A univariate distribution function F is **n-completely mixable** (n-CM) if there exist n identically distributed random variables $X_1, \cdots, X_n$ with the same distribution F such that
$$\mathbb{P}(X_1 + \cdots + X_n = C) = 1 \tag{2.1}$$
for some $C \in \mathbb{R}$.
2. The univariate distribution functions $F_1, \cdots, F_n$ are **jointly mixable** (JM) if there exist n random variables $X_1, \cdots, X_n$ with distribution functions $F_1, \cdots, F_n$ respectively, such that (2.1) holds for some $C \in \mathbb{R}$.

The definition of the CM distributions is formally given in Wang and Wang [14] although the concept has been used in variance reduction problems earlier (see Gaffke and Rüschendorf [7], Knott and Smith [9], Rüschendorf and Uckelmann [13]). Some examples of n-CM distributions include the distribution of a constant (for $n \geq 1$), uniform distributions (for $n \geq 2$), Normal distributions (for $n \geq 2$), Cauchy distributions (for $n \geq 2$), binomial distributions $B(n, p/q)$ with $p, q \in \mathbb{N}$ (for $n = q$), bounded monotone distributions on $[0, 1]$ with $1/m \leq \mathbb{E}(X) \leq 1 - 1/m$ (for $n \geq m$). See Wang and Wang [14] for more details of the CM distributions.

The concept of JM distributions is first introduced in this paper as a generalization of the CM distributions. Obviously, $F_1, \cdots, F_n$ are JM distributions when $F_1 = \cdots = F_n = F$ and F is n-CM. The following proposition gives a necessary condition for JM distributions and the condition is sufficient for n normal distributions. The proof is given in the appendix.

Proposition 2.4

1. *Suppose $F_1, \cdots, F_n$ are JM with finite variance $\sigma_1^2, \cdots, \sigma_n^2$. Then*

$$\max_{1 \le i \le n} \sigma_i \le \frac{1}{2} \sum_{i=1}^{n} \sigma_i. \tag{2.2}$$

2. *Suppose F_i is $N(\mu_i, \sigma_i^2)$ for $i = 1, \cdots, n$. Then $F_1, \cdots, F_n$ are JM if and only if* (2.2) *holds.*

Remark 2.5 Due to the complexity of multivariate distributional problems, it remains unknown and extremely difficult to find general sufficient conditions for the JM distributions.

Before presenting the main results on the relationship between the bounds on $m_+(s)$ and the jointly mixable distributions, we define the conditional moment function $\Phi(t)$ which turns out to play an important role in the problem of finding $m_+(s)$. Suppose $X_i \sim F_i$ for $i = 1, \cdots, n$. Define

$$\Phi(t) = \sum_{i=1}^{n} \mathbb{E}(X_i | X_i \ge F_i^{-1}(t))$$

for $t \in (0, 1)$, and

$$\Phi(1) = \lim_{t \to 1-} \Phi(t), \quad \Phi(0) = \lim_{t \to 0+} \Phi(t).$$

Obviously $\Phi(t)$ is increasing and continuous when $F_i, i = 1, \cdots, n$ are continuous. Define

$$\Phi^{-1}(x) = \inf\{t \in [0, 1] \ : \ \Phi(t) \ge x\}$$

for $x \le \Phi(1)$ and $\Phi^{-1}(x) = 1$ for $x > \Phi(1)$.

Theorem 2.6 *Suppose the distributions $F_1, \cdots, F_n$ are continuous.*

(1) We have

$$m_+(s) \ge \Phi^{-1}(s); \tag{2.3}$$

(2) For each fixed $s \ge \Phi(0)$, the equality

$$m_+(s) = \Phi^{-1}(s) \tag{2.4}$$

holds if and only if the conditional distributions $\tilde{F}_{1,a}, \cdots, \tilde{F}_{n,a}$ are jointly mixable, where $a = \Phi^{-1}(s)$.

Proof

(1) It is trivial to prove the result when $\Phi(0) = \infty$. So we assume $\Phi(0) < \infty$. Note that from **Lemma 2.2** we know that there exists an optimal coupling $\mathbf{X} = (X_1, \ldots, X_n)$ for $m_+(s)$ such that $\{S \ge s\} = \{X_i \ge F_i^{-1}(m_+(s))\}$ for each $i = 1, \cdots, n$. Hence

$$s \le \mathbb{E}[S | S \ge s] = \sum_{i=1}^{n} \mathbb{E}[X_i | X_i \ge F_i^{-1}(m_+(s))] = \Phi(m_+(s)),$$

which implies (2.3).

(2) Suppose $\mathbf{X} = (X_1, \cdots, X_n)$ is an optimal coupling for $m_+(s)$ such that $\{S \geq s\} = \{X_i \geq F_i^{-1}(m_+(s))\}$ for each $i = 1, \cdots, n$. When $m_+(s) = \Phi^{-1}(s)$, it follows from the proof of part (1) that $\mathbb{E}(S|S \geq s) = s$, which implies that the conditional distributions of $X_1, \cdots, X_n$ on the set $\{S \geq s\}$ are JM, i.e., the conditional distributions $\tilde{F}_{1,a}, \cdots, \tilde{F}_{n,a}$ are JM.
Conversely, assume that $\tilde{F}_{1,a}, \cdots, \tilde{F}_{n,a}$ are JM. Then there exist $Y_1 \sim \tilde{F}_{1,a}, \cdots, Y_n \sim \tilde{F}_{n,a}$ such that

$$Y_1 + \cdots + Y_n = \mathbb{E}(Y_1 + \cdots + Y_n) = \Phi(a) \geq s.$$

Let

$$X_i = F_i^{-1}(U)\mathbf{1}_{\{U \leq a\}} + Y_i \mathbf{1}_{\{U > a\}}, \tag{2.5}$$

where $U \sim \mathrm{U}[0, 1]$ and is independent of $(Y_1, \cdots, Y_n)$. Then it is easy to verify that X_i has the distribution function F_i for $i = 1, \cdots, n$ and

$$m_+(s) \leq \mathbb{P}(S < s) \leq a = \Phi^{-1}(s).$$

The other inequality $m_+(s) \geq \Phi^{-1}(s)$ is shown in part (1).

Remark 2.7

1. It is seen from the proof that the continuity of F_i can be removed. In a recent preprint, Puccetti and Rüschendorff [10] established **Theorem 2.6** independently, where the equivalent form $\sup\{\mathbb{P}(S > s), X_1 \sim F_1, \cdots, X_n \sim F_n\} \leq 1 - \Phi^{-1}(s)$ is proved without assuming the continuity of F_i.
2. The optimal coupling is given in (2.5). Although the existence of such $Y_1, \cdots, Y_n$ is guaranteed by the mixable condition, finding $Y_1, \cdots, Y_n$ remains quite challenging. For example, when the marginal distributions F_i are identical and completely mixable, the dependence structure of random variables $Y_1, \cdots, Y_n$ may not be unique and is hard to be specified in general as discussed in Wang and Wang [14].

2.2 Bounds for the sum with identical marginal distributions

Here we consider $m_+(s)$ in the homogeneous case, i.e. $F_1 = \cdots = F_n \equiv F$. For $X \sim F$, define

$$\psi(t) = \mathbb{E}(X|X \geq F^{-1}(t))$$

for $t \in (0, 1)$,

$$\psi(1) = \lim_{t \to 1-} \psi(t), \quad \psi(0) = \lim_{t \to 0+} \psi(t),$$

$$\psi^{-1}(x) = \inf\{t \in [0, 1] \ : \ \psi(t) \geq x\}$$

for $x \leq \psi(1)$ and $\psi^{-1}(x) = 1$ for $x > \psi(1)$. The following result follows from **Theorem 2.6** immediately.

Corollary 2.8 *Suppose $F_1 = \cdots = F_n \equiv F$ and F is continuous.*

(1) We have

$$m_+(s) \geq \psi^{-1}(s/n). \tag{2.6}$$

(2) For each fixed $s \geq n\psi(0)$, the equality

$$m_+(s) = \psi^{-1}(s/n) \tag{2.7}$$

holds if and only if the conditional distribution function $\tilde{F}_a$ is n-completely mixable, where $a = \psi^{-1}(s/n)$.

Next we compare the bound in (2.6) with the bound obtained in Embrechts and Puccetti [4], which is

$$m_+(s) \geq 1 - n \inf_{r\in[0,s/n)} \frac{\int_r^{s-(n-1)r}(1 - F(t))\mathrm{d}t}{s - nr} \quad \text{for} \quad s > 0. \tag{2.8}$$

Proposition 2.9 *The bound* (2.8) *is greater than or equal to the bound* (2.6). *Moreover, both are equal if and only if $F^{-1}(1) < \infty$ and a solution to the infimum*

$$\inf_{r\in[0,s/n)} \frac{\int_r^{s-(n-1)r}(1 - F(t))\mathrm{d}t}{s - nr} \tag{2.9}$$

lies in $[0, \frac{s-F^{-1}(1)}{n-1}]$.

The proof of the above Proposition is given in the appendix.

Different from the bound in Embrechts and Puccetti [4], **Theorem 2.6** deals with a more general case, where the random variables $X_1, \cdots, X_n$ do not need to be identically distributed and positive. Moreover, the bound in **Theorem 2.6** is easier to calculate. Obviously, the bounds in **Corollary 2.8** and in Embrechts and Puccetti [4] are the same and both are sharp when the conditional distribution $\tilde{F}_a$ is completely mixable. A comparison of the two bounds is given in Figure 3.2 in Section 3 when the marginal distributions have infinite support (see also *Remark* 3.5). Note that infinite support generally implies that the mixable condition in **Theorem 2.6** and **Corollary 2.8** does not hold.

3 Bounds for identically distributed risks with monotone densities

In this section, we investigate the homogeneous case when $F_1 = \cdots = F_n = F$ and F has either a monotone density or a tail-monotone density on its support. Since the case of $n = 1$ is trivial, we assume $n \geq 2$. When the distribution F with support on $[0, 1]$ has a decreasing density and satisfies the regular condition $\psi(t) \geq t + \frac{1-t}{n}$, Wang and Wang [14] showed that $m_+(s) = \psi^{-1}(s/n)$, which now becomes a corollary of **Theorem 2.6**.

When the support of the distribution F is unbounded, the mixable condition in **Theorem 2.6** and **Corollary 2.8** is not satisfied (see **Proposition 2.1**(7) in Wang and Wang [14]), i.e., the bound $\psi^{-1}(s/n)$ is not sharp. In this section, we find a formula for calculating the bound $m_+(s)$ for any distribution with a monotone density or a tail-monotone density, and obtain the corresponding correlation structure. This partially answers the question of optimal coupling for $m_+(s)$, which has remained open for decades. As a direct application, the bounds on $\mathrm{VaR}_\alpha(S)$ are obtained as well.

3.1 Preliminaries

To calculate $m_+(s)$ for F having a monotone marginal density, we first review the construction of copula Q_n^F $(n \geq 2)$ in Wang and Wang [14], where F is a distribution function with an increasing (non-decreasing) density. More specifically, for some $0 \leq c \leq 1/n$ and random vector $(U_1, \cdots, U_n)$ with uniform marginal distributions on [0,1], we say $(U_1, \cdots, U_n) \sim Q_n^F(c)$ if

(a) for each $i = 1, \cdots, n$, given $U_i \in [0, c]$, we have $U_j = 1 - (n-1)U_i,\ \forall j \neq i$;
(b) $F^{-1}(U_1) + \cdots + F^{-1}(U_n)$ is a constant when any one of $U_i's$ lies in $(c, 1-(n-1)c)$.

Denote $Q_n^F = Q_n^F(c_n)$ where c_n is the smallest possible c such that $Q_n^F(c)$ exists. Note that $c_n = 0$ if and only if F is n-CM. Define

$$H(x) = F^{-1}(x) + (n-1)F^{-1}(1-(n-1)x) \quad \text{for } F \text{ with a non-decreasing density.} \quad (3.1)$$

Then the smallest possible c for F with an increasing density is

$$c_n = \min\{c \in [0, \frac{1}{n}] : \int_c^{\frac{1}{n}} H(t)\mathrm{d}t \leq (\frac{1}{n} - c)H(c)\} \quad (3.2)$$

and for any convex function f,

$$\min_{X_1,\cdots,X_n \sim F} \mathbb{E}f(X_1 + \cdots + X_n) = \mathbb{E}^{Q_n^F} f\left(F^{-1}(U_1) + \cdots + F^{-1}(U_n)\right). \quad (3.3)$$

Note that Q_n^F may not be unique. The existence of Q_n^F and details of the above results can be found in Section 3 of Wang and Wang [14].

For F with a decreasing density $(n \geq 2)$, we define $Q_n^F(c)$ similarly as follows. For some $0 \leq c \leq 1/n$, we say $(U_1, \cdots, U_n) \sim Q_n^F(c)$ if

(a') for each $i = 1, \cdots, n$, given $U_i \in [1-c, 1]$, we have $U_j = (n-1)(1-U_i),\ \forall j \neq i$;
(b') $F^{-1}(U_1) + \cdots + F^{-1}(U_n)$ is a constant when any one of U_i lies in $((n-1)c, 1-c)$.

Define

$$H(x) = (n-1)F^{-1}((n-1)x) + F^{-1}(1-x) \quad \text{for } F \text{ with a decreasing density.} \quad (3.4)$$

As for the distribution of Z with a decreasing density, the distribution of $-Z$ has an increasing density, thus the above properties hold for F with a decreasing density. That is, the smallest possible c for F with a decreasing density is

$$c_n = \min\{c \in [0, \frac{1}{n}] : \int_c^{\frac{1}{n}} H(t)\mathrm{d}t \geq (\frac{1}{n} - c)H(c)\}. \quad (3.5)$$

And for a distribution F with a decreasing density and any convex function f the equation (3.3) holds.

Although

$$m_+(s) = \min_{X_1,\cdots,X_n \sim F} \mathbb{E}(\mathbf{1}_{\{S<s\}}),$$

the above results can not be applied directly to solve $m_+(s)$ since the indicator function $\mathbf{1}_{(-\infty,s)}(\cdot)$ is not a concave function. Here we propose to find $m_+(s)$ for F with a monotone marginal density based on the following properties of Q_n^F.

Proposition 3.1 *Suppose F admits a monotone density on its support.*

1. *If $(U_1, \cdots, U_n) \sim Q_n^F(c)$ and F has an increasing density, then $\mathbf{1}_{\{U_i \in (c, 1-(n-1)c)\}} = \mathbf{1}_{\{U_1 \in (c, 1-(n-1)c)\}}$ a.s. for $i = 1, \cdots, n$.*
2. *If $X_1, \cdots, X_n \sim F$ with copula Q_n^F, then*

$$S = X_1 + \cdots + X_n = \begin{cases} H(U/n)\mathbf{1}_{\{U \le nc_n\}} + H(c_n)\mathbf{1}_{\{U > nc_n\}}, & c_n > 0; \\ n\mathbb{E}(X_1), & c_n = 0 \end{cases} \tag{3.6}$$

for some $U \sim \mathrm{U}[0, 1]$.

The proof of **Proposition 3.1** is given in the appendix. For more details of the copula Q_n^F, see Wang and Wang [14].

3.2 Monotone marginal densities

Now we are ready to give a computable formula for $m_+(s)$. In the following we define a function $\phi(x)$ which works similarly as $\Phi(x)$ in the mixable case.

For F with a decreasing density and $a \in [0, 1]$, define

$$H_a(x) = (n-1)F^{-1}(a + (n-1)x) + F^{-1}(1 - x) \tag{3.7}$$

for $x \in [0, \frac{1-a}{n}]$ and

$$c_n(a) = \min\{c \in [0, \frac{1}{n}(1-a)] : \int_c^{\frac{1}{n}(1-a)} H_a(t)\mathrm{d}t \ge (\frac{1}{n}(1-a) - c)H_a(c)\}. \tag{3.8}$$

Write

$$\phi(a) = \begin{cases} H_a(c_n(a)) & \text{if } c_n(a) > 0, \\ n\psi(a) & \text{if } c_n(a) = 0. \end{cases} \tag{3.9}$$

On the other hand, for F with an increasing density and $a \in [0, 1]$, define

$$H_a(x) = F^{-1}(a + x) + (n-1)F^{-1}(1 - (n-1)x), \tag{3.10}$$

$$c_n(a) = \min\{c \in [0, \frac{1}{n}(1-a)] : \int_c^{\frac{1}{n}(1-a)} H_a(t)\mathrm{d}t \le (\frac{1}{n}(1-a) - c)H_a(c)\} \tag{3.11}$$

and

$$\phi(a) = \begin{cases} H_a(0) \text{ if } c_n(a) > 0, \\ n\psi(a) \text{ if } c_n(a) = 0. \end{cases} \tag{3.12}$$

Some probabilistic interpretation of the functions $H_a(x)$ and $\phi(a)$ is given in the following remark. Technical details are put in **Lemma 3.3** later.

Remark 3.2 Suppose $Y_1, \cdots, Y_n \sim \tilde{F}_a$ with copula $Q_n^{\tilde{F}_a}$. By (3.6) we have

$$Y_1 + \cdots + Y_n = \begin{cases} \tilde{H}(U/n)\mathbf{1}_{\{U \le n\tilde{c}_n\}} + \tilde{H}(\tilde{c}_n)\mathbf{1}_{\{U > n\tilde{c}_n\}}, & \tilde{c}_n > 0, \\ n\mathbb{E}(Y_1), & \tilde{c}_n = 0 \end{cases}$$

for some $U \sim \mathrm{U}[0,1]$, where $\tilde{H}(x)$ and $\tilde{c}_n$ are $H(x)$ and c_n defined in (3.1), (3.2), (3.4) and (3.5) by replacing F with $\tilde{F}_a$. It is easy to check that $\tilde{H}(x) = H_a((1-a)x)$, $\tilde{c}_n = c_n(a)/(1-a)$ and $\tilde{H}(\tilde{c}_n) = H_a(c_n(a))$. For $c_n(a) > 0$, later we will show that $H_a(x)$, $x \in [0, c_n(a)]$ attains its minimum value at $H_a(c_n(a))$ for $\tilde{F}_a$ with a decreasing density and at $H_a(0)$ for $\tilde{F}_a$ with an increasing density. Therefore, the minimum possible value of $Y_1 + \cdots + Y_n$ is

$$\min_{x \in [0, c_n(a)]} H_a(x)\mathbf{1}_{\{c_n(a) > 0\}} + n\mathbb{E}(Y_1)\mathbf{1}_{\{c_n(a) = 0\}} = \phi(a).$$

Thus, $\mathbb{P}(Y_1 + \cdots + Y_n \ge \phi(a)) = 1$, which leads to $\mathbb{P}(S < \phi(a)) \le a$ by setting $X_i = F^{-1}(V)\mathbf{1}_{\{V \le a\}} + Y_i\mathbf{1}_{\{V > a\}}$ where $V \sim \mathrm{U}[0,1]$ is independent of $Y_1, \cdots, Y_n$. This suggests $m_+(s) \le \phi^{-1}(a)$, i.e., $\phi^{-1}(a)$ is potentially an optimal bound. In order to prove the optimality of $\phi^{-1}(a)$, more details of the functions $H_a(x)$ and $\phi(a)$ are given in the following lemma, whose proof is put in the appendix.

Lemma 3.3 *Suppose F admits a monotone density.*

(i) If F has a decreasing density, then given $a \in [0,1)$, $H_a(x)$ is decreasing and differentiable for $x \in [0, c_n(a)]$.
(ii) If F has an increasing density, then given $a \in [0,1)$, $H_a(x)$ is increasing and differentiable for $x \in [0, c_n(a)]$.
(iii) If F has a decreasing density, then $\phi(a) = n\mathbb{E}[F^{-1}(V_a)]$ where $V_a \sim \mathrm{U}[a + (n-1)c_n(a), 1 - c_n(a)]$.
(iv) For any random variables $U_1, \cdots, U_n \sim \mathrm{U}[a,1]$ and $0 \le a < b \le 1$, we have $\mathbb{E}(F^{-1}(U_i)|A) < \mathbb{E}[F^{-1}(V_b)]$ for $i = 1, \cdots, n$, where V_b is defined in (iii) and $A = \bigcap_{i=1}^n \{U_i \in [a, 1 - c_n(b)]\}$.
(v) Suppose $Y_1, \cdots, Y_n \sim \tilde{F}_a$ with copula $Q_n^{\tilde{F}_a}$, then $\mathbb{P}(Y_1 + \cdots + Y_n \ge \phi(a)) = 1$.
(vi) $\phi(a)$ is continuous and strictly increasing for $a \in [0,1)$.

Since $\phi(a)$ is continuous and strictly increasing, its inverse function $\phi^{-1}(a)$ exists. Put $\phi^{-1}(t) = 0$ if $t < \phi(0)$ and $\phi^{-1}(t) = 1$ if $t > \phi(1)$.

Theorem 3.4 *Suppose the distribution $F(x)$ has a decreasing density on its support and $\phi(a)$ is defined in* (3.9)*, or the distribution $F(x)$ has an increasing density on its support and $\phi(a)$ is defined in* (3.12)*. Then we have $m_+(s) = \phi^{-1}(s)$.*

Proof.

(a) We first prove $m_+(s) \le \phi^{-1}(s)$. Write $a = \phi^{-1}(s)$. For $i = 1, \cdots, n$, let $Y_1, \cdots, Y_n \sim \tilde{F}_a$ with copula $Q_n^{\tilde{F}_a}$ and $X_i = F^{-1}(V)\mathbf{1}_{\{V \le a\}} + Y_i\mathbf{1}_{\{V > a\}}$ where $V \sim \mathrm{U}[0,1]$ is independent of $Y_1, \cdots, Y_n$. It is easy to check that $X_i \sim F$ and by **Lemma 3.3**(v),

$$m_+(s) \le \mathbb{P}(S < \phi(a)) = 1 - \mathbb{P}(S \ge \phi(a)) \le 1 - \mathbb{P}(Y_1 + \cdots + Y_n \ge \phi(a))\mathbb{P}(V > a) = a.$$

Thus $m_+(s) \le \phi^{-1}(s)$.

(b) Next we prove $m_+(s) \geq \phi^{-1}(s)$ when $F(x)$ has a decreasing density.
Suppose $a = m_+(s) < \phi^{-1}(s) = b$ and $\mathbf{X} = (X_1, \cdots, X_n)$ is an optimal coupling for $m_+(s)$ such that $\{S \geq s\} = \{X_i \geq F^{-1}(a)\}$ for each $i = 1, \cdots, n$. Hence there exist $U_{a,1}, \cdots, U_{a,n} \sim \mathrm{U}[a, 1]$ such that $F^{-1}(U_{a,1}) + \cdots + F^{-1}(U_{a,n}) \geq s$ with probability 1. By **Lemma 3.3**(iii) and (iv), we have

$$s \leq \mathbb{E}[\sum_{i=1}^{n} F^{-1}(U_{a,i})|A] < n\mathbb{E}(F^{-1}(V_b)) = \phi(b) = s.$$

This leads to a contradiction. Thus $m_+(s) = \phi^{-1}(s)$.

(c) Finally we prove $m_+(s) \geq \phi^{-1}(s)$ when $F(x)$ has an increasing density. In this case $F^{-1}(1) < \infty$.
Write $a = m_+(s)$ and let $\mathbf{X} = (X_1, \cdots, X_n)$ be an optimal coupling for $m_+(s)$ such that $\{S \geq s\} = \{X_i \geq F^{-1}(a)\}$ for each $i = 1, \cdots, n$. It is clear that

$$\begin{aligned}&\mathbb{P}(S < F^{-1}(a) + (n-1)F^{-1}(1) + \epsilon | S \geq s)\\ &\quad \geq \mathbb{P}(X_i < F^{-1}(a) + \epsilon | X_i \geq F^{-1}(a)) > 0\end{aligned}$$

for any $\epsilon > 0$. Note that $\mathbb{P}(S < s | S \geq s) = 0$ and thus

$$s \leq F^{-1}(a) + (n-1)F^{-1}(1) = H_a(0).$$

This shows $s \leq H_a(0)$. The inequality $s \leq n\psi(a)$ is given by **Theorem 2.6**. Hence $s \leq \phi(a)$ and $a \geq \phi^{-1}(s)$. □

The proof of the above theorem suggests constructing the optimal correlation structure as follows. In both cases, for $a = \phi^{-1}(s)$ let $U_{a,1}, \cdots, U_{a,n} \sim \mathrm{U}[a, 1]$ with copula $Q_n^{\tilde{F}_a}$ and $U \sim$U[0,1] is independent of $(U_{a,1}, \cdots, U_{a,n})$. Define

$$U_i = U_{a,i}\mathbf{1}_{\{U \geq a\}} + U\mathbf{1}_{\{U < a\}} \tag{3.13}$$

for $i = 1, \cdots, n$. Then

$$\mathbb{P}(F^{-1}(U_1) + \cdots + F^{-1}(U_n) < s) = \phi^{-1}(s).$$

Remark 3.5

1. The copula Q_n^F plays an important role in deriving bounds for the convex minimization problem (3.3) and the $m_+(s)$ problem with monotone marginal densities. Note that Q_n^F may not be unique, hence the structure (3.13) may not be unique. Also, on the set $\{S < s\}$, the dependence structure of $X_1, \cdots, X_n$ can be arbitrary.
2. The value $\phi^{-1}(s)$ is accurate even when $\mathbb{E}(\max\{X_1, 0\}) = \infty$. When the distribution $\tilde{F}_a$ is n-CM, **Theorem 3.4** gives the sharp bound $\Phi^{-1}(s)$ in **Theorem 2.6**.
3. When a random variable X has a monotone density, $-X$ has a monotone density too. Hence the above theorem also solves the similar problem

$$M_+(s) = \sup_{X_i \sim P} \mathbb{P}(S < s) = 1 - \inf_{X_i \sim P} \mathbb{P}(-S \leq s) = 1 - \inf_{X_i \sim P} \mathbb{P}(-S < s),$$

where P has a monotone density.

4. Figure 3.1 shows the sketch of an optimal coupling for F with a decreasing density, some $a > 0$ and $c_n(a) > 0$. Here $U_1, \cdots, U_n \sim$ U[0, 1] and $\mathbb{P}(F^{-1}(U_1) + \cdots + F^{-1}(U_n) < s) = \phi^{-1}(s)$.
   (i) When $U_i \in [0, a]$, U_i is arbitrarily coupled to all other U_j in Part A.
   (ii) When $U_i \in [a, a + (n-1)c_n(a)]$, U_i is coupled to other U_j, $j \neq i$ in Part B and Part D. For $j \neq i$, either $U_i - a = (n-1)(1 - U_j)$ or $U_j = U_i$.
   (iii) When $U_i \in [a + (n-1)c_n(a), 1 - c_n(a)]$, U_i is coupled to all other U_j, $j \neq i$ in Part C, and $F^{-1}(U_1) + \cdots + F^{-1}(U_n) = \phi(a)$. It is the completely mixable part.
   (iv) When $U_i \in [1 - c_n(a), 1]$, U_i is coupled to other U_j, $j \neq i$ in Part B. For $j \neq i$, $U_j - a = (n-1)(1 - U_i)$.
5. Figure 3.2 shows the real values of $m_+(s)$ in **Theorem 3.4** and the lower bound $\psi^{-1}(s/n)$ in **Theorem 2.6** for the Pareto(2,1) distribution. We also calculate the bound (2.8) in Embrechts and Puccetti [4] (see Section 2.2). It turns out that in this case the real values are equal to the bound (2.8), which suggests that the bound (2.8) in [4] may be sharp for Pareto distributions.

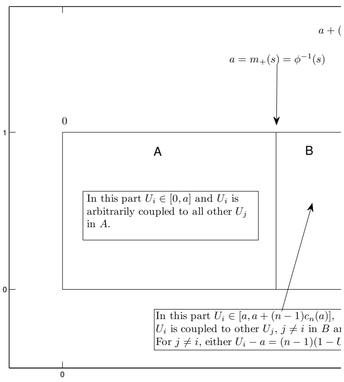

Fig. 3.1 Sketch of the optimal coupling

3.3 Tail-monotone marginal densities

For the distribution F with density $p(x)$, we say $p(x)$ is **tail-monotone**, if for some $b \in \mathbb{R}$, $p(x)$ is decreasing for $x > b$ or $p(x)$ is increasing for $x < b$. We are particularly

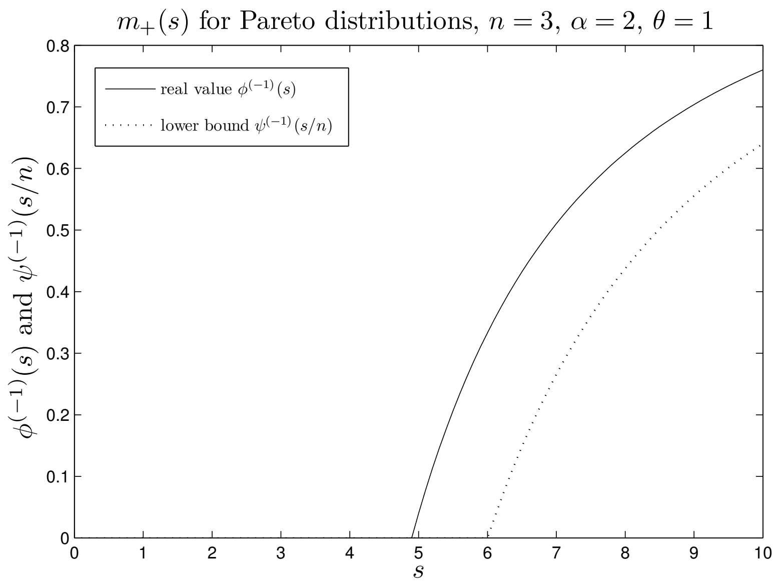

Fig. 3.2 $m_+(s)$ and $\psi^{-1}(s/n)$ for a Pareto distribution

interested in the case when $p(x)$ is **tail-decreasing** ($p(x)$ is decreasing for $x > b$) since the risks are usually positive random variables. For most risk distributions the tail-decreasing property is satisfied. For example, the Gamma distribution with shape parameter α for $\alpha > 1$ and the F-distribution with d_1, d_2 degrees of freedom with $d_1 > 2$ have a tail-decreasing density, but do not have a monotone density.

In the VaR problems, one is concerned with the tail behavior of the distribution. From the proof of **Theorem 3.4**, information on the left tail of F does not play any role in the calculation of $m_+(s)$. Based on this observation, we have the following theorem, which solves $m_+(s)$ for F with tail-decreasing density and some large s.

Theorem 3.6 *Suppose the density function of F is decreasing on $[b, \infty)$, and $\phi(a)$ is defined in* (3.9)*. Then for $s \geq \phi(F(b))$, $m_+(s) = \phi^{-1}(s)$.*

Proof Since the density function of F is decreasing on $[b, \infty)$, the conditional distribution $\tilde{F}_{F(b)}$ has a decreasing density. Note that $H_a(x)$, $c_n(a)$ and $\phi(a)$ only depend on the conditional distribution $\tilde{F}_a$, hence they are well defined for $F(b) \leq a \leq 1$.

Since $s \geq \phi(F(b))$, $\phi^{-1}(s) \geq F(b)$ and the conditional distribution $\tilde{F}_{\phi^{-1}(s)}$ has a decreasing density. **Theorem 3.6** follows from the same arguments as in the proof of **Theorem 3.4**, where no condition on the distribution of X_i on $\{X_i < F^{-1}(\phi^{-1}(s))\}$ is used.

3.4 The worst Value-at-Risk scenarios

The Value-at-Risk (VaR) is an important risk measure in risk management; see Embrechts and Puccetti [6] and references therein. Recall that VaR is the α-quantile of the distribution, i.e.,

$$\mathrm{VaR}_\alpha(S) = F_S^{-1}(\alpha) = \inf\{s \in \mathbb{R} : F_S(s) \geq \alpha\}, \tag{3.14}$$

where F_S is the distribution of S. Typical values of the level α are 0.95, 0.99 or even 0.999. As mentioned in Embrechts and Puccetti [6], banks are concerned with an upper bound on $\mathrm{VaR}(\sum_{i=1}^d X_i)$ when the correlation structure between $\mathbf{X} = (X_1, \cdots, X_d)$ is unspecified.

Finding the bounds on the VaR is equivalent to finding the inverse function of $m_+(s)$ (note that $m_+(s)$ is non-decreasing). Using **Theorem 3.4** and **Theorem 3.6**, we are able to obtain the explicit value of the upper bound on the VaR, namely, the worst Value-at-Risk. The proof follows directly from the fact that $\sup_{X_i \sim F, 1\leq i\leq n} \mathrm{VaR}_\alpha(S) = m_+^{-1}(\alpha)$ when $m_+(s)$ is continuous and strictly increasing.

Corollary 3.7 *Suppose that the density function of the marginal distribution F is decreasing on $[b, \infty)$ and $\phi(a)$ is defined in* (3.9). *Then for $\alpha \geq F(b)$, the worst VaR of $S = X_1 + \cdots + X_n$ is*

$$\sup_{X_i \sim F, 1\leq i\leq n} \mathrm{VaR}_\alpha(S) = m_+^{-1}(\alpha) = \phi(\alpha). \tag{3.15}$$

In particular, (3.15) *holds for all α if the marginal distribution F has decreasing density on its support and an optimal correlation structure is given by* (3.13).

For arbitrary marginal distributions $F_1, \cdots, F_n$, **Theorem 2.6** gives an upper bound for the worst-VaR problem as follows.

Corollary 3.8 *For arbitrary marginal distributions,*

$$\sup_{X_i \sim F_i, i=1,\cdots,n} \mathrm{VaR}_\alpha(S) \leq m_+^{-1}(\alpha) \leq \Phi(\alpha), \tag{3.16}$$

where $\Phi(\alpha)$ is defined in Section 2.

Figure 3.3 shows the explicit worst-VaR in (3.15) and the upper bound in (3.16) for the distribution Pareto(4,1) and $0.9 \leq \alpha \leq 0.995$.

3.5 Examples

Here we give some examples to show how to compute $m_+(s)$.

Example 3.9 Assume that $X \sim \mathrm{U}[0, 1]$, the uniform distribution on [0,1]. Then

$$p(x) = 1,\ F(x) = x,\ x \in [0, 1],\ F^{-1}(t) = t,\ t \in [0, 1].$$

Further we have $c_n(a) = 0$ for all $0 \leq a \leq 1$ and $\phi(t) = n\psi(t) = n\mathbb{E}(X|X > t) = \frac{n(1+t)}{2}$ for $t \in [0, 1]$. Thus

$$m_+(s) = \phi^{-1}(s) = 1 \wedge \left(\frac{2s}{n} - 1\right)_+ .$$

This result indeed is the same as that in Rüschendorf [11]. One optimal correlation structure is also given in Rüschendorf and Uckelmann [13].

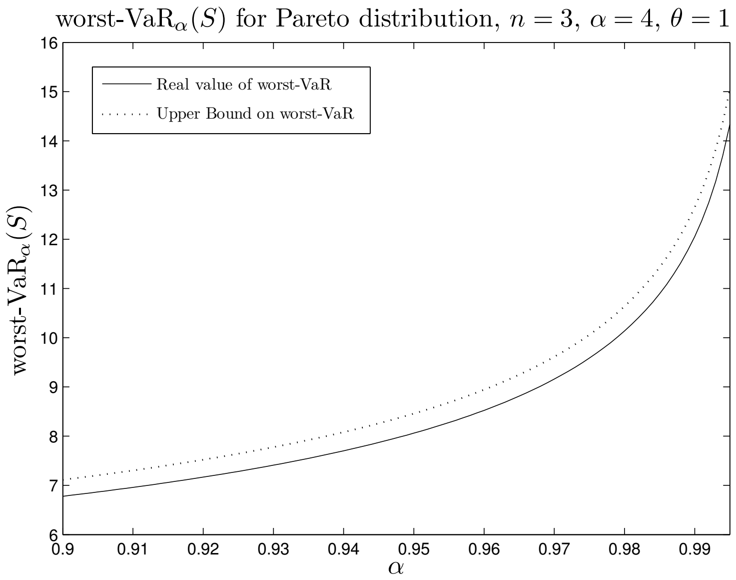

Fig. 3.3 Worst-VaR for a Pareto distribution

Example 3.10 Assume that $X \sim \text{Pareto}(\alpha, \theta), \alpha > 1, \theta > 0$ with density function

$$p(x) = \alpha\theta^\alpha x^{-\alpha-1}, \ x \geq \theta.$$

Then

$$F(x) = 1 - \left(\frac{x}{\theta}\right)^{-\alpha}, \ x \geq \theta, \ F^{-1}(t) = \theta(1-t)^{-1/\alpha}, \ t \in [0, 1].$$

Further we have that $c_n(a)$ is the smallest $c \in [0, \frac{1}{n}(1-a)]$ such that

$$\frac{\alpha}{\alpha - 1}((1-a-(n-1)c)^{1-1/\alpha} - c^{1-1/\alpha}) \geq (\frac{1}{n}(1-a)-c)((n-1)(1-a-(n-1)c)^{-1/\alpha} + c^{-1/\alpha}).$$

The numerical values of $m_+(s)$ for two Pareto distributions and $n = 3$ are plotted in Figure 3.4. A possible correlation structure is given in (3.13).

Example 3.11 Assume that $X \sim \text{Gamma}(\alpha, \lambda), \alpha \leq 1, \lambda > 0$ with density function

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

Then

$$F(x) = \gamma(\alpha, \lambda x), \ x > 0,$$

where $\gamma(\alpha, t) = \int_0^t \frac{1}{\Gamma(\alpha)} x^{\alpha-1} e^{-\lambda x} \mathrm{d}x$ is the lower incomplete Gamma function. Further $c_n(a)$ is the smallest $c \in [0, \frac{1}{n}(1-a)]$ such that

$$\frac{\alpha}{\lambda}(\gamma(\alpha+1, \lambda F^{-1}(1-c)) - \gamma(\alpha+1, \lambda F^{-1}(a+(n-1)c))) \geq (\frac{1}{n}(1-a)-c)H_a(c),$$

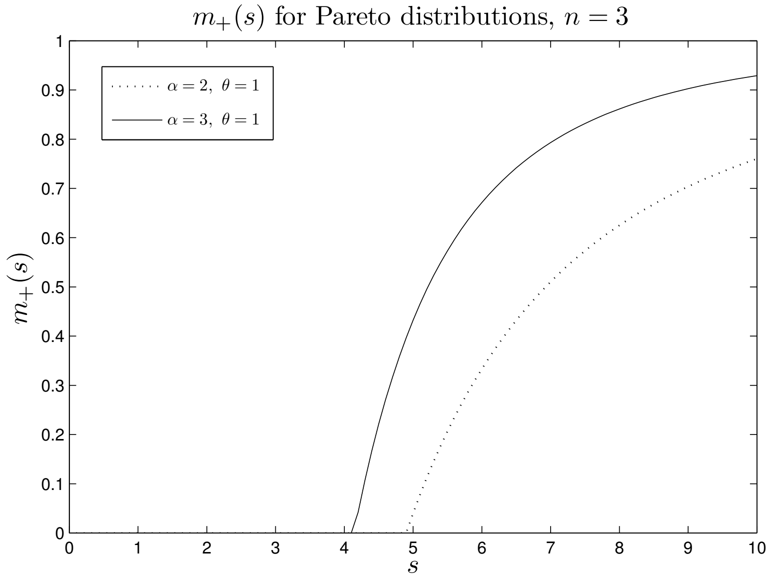

Fig. 3.4 $m_+(s)$ for Pareto distributions

which can be calculated numerically. The numerical values of $m_+(s)$ for two Gamma distributions and $n = 3$ are plotted in Figure 3.5. A possible correlation structure is given in (3.13).

4 Conclusions

In this paper, we provide a new lower bound for $m_+(s)$ with any given marginal distributions, and give a necessary and sufficient condition for its sharpness in terms of the joint mixability. When the marginal distributions have a common monotone density, the explicit value of $m_+(s)$ and the worst Value-at-Risk are obtained. We also extend these results to distributions with a common tail-monotone density.

Appendix

Proof of Proposition 2.4.

1. The case $n = 1$ is trivial. For $n \geq 2$, by the definition of JM distributions, there exist $X_1 \sim F_1, \cdots, X_n \sim F_n$ such that $\mathrm{Var}(X_1 + \cdots + X_n) = 0$. Since

$$\begin{aligned}\sqrt{\mathrm{Var}(X_1 + X_2 + \cdots + X_n)} &\geq \sqrt{\mathrm{Var}(X_1)} - \sqrt{\mathrm{Var}(X_2 + \cdots + X_n)} \\ &\geq \sigma_1 - \sum_{i=2}^{n} \sigma_i,\end{aligned}$$

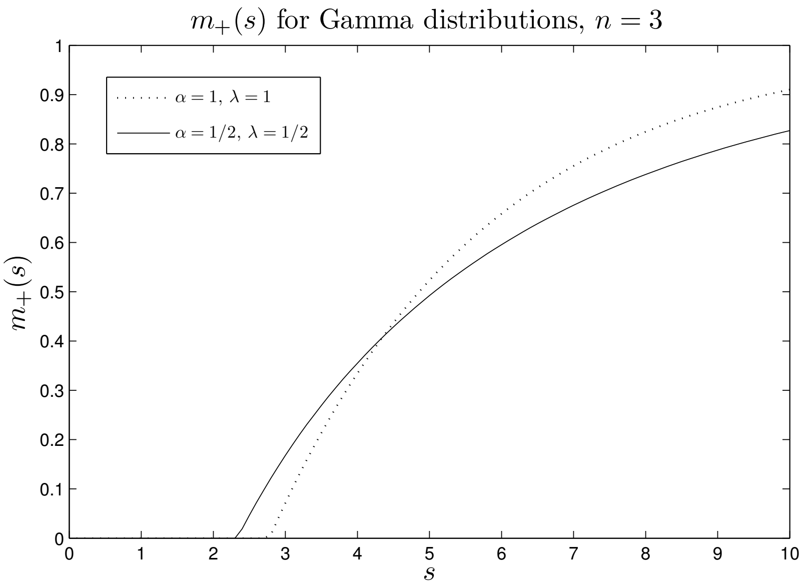

Fig. 3.5 $m_+(s)$ for Gamma distributions

we have $2\sigma_1 - \sum_{i=1}^n \sigma_i \le 0$. Similarly, we can show that $2\sigma_k - \sum_{i=1}^n \sigma_i \le 0$ for any $k = 1, \cdots, n$, i.e., (2.2) holds.

2. We only need to prove the "$\Leftarrow$" part for $n \ge 2$. Without loss of generality, we assume $\sigma_1 \ge \sigma_2 \ge \cdots \ge \sigma_n$. Let $\mathbf{X} = (X_1, \cdots, X_n)$ be a multivariate Gaussian random vector with known marginal distributions $F_1, \cdots, F_n$ and an unspecific correlation matrix Γ. We want to show there exists a correlation matrix Γ such that $\mathrm{Var}(X_1 + \cdots + X_n) = 0$.
Let T be the correlation matrix of $(X_2, \cdots, X_n)$ and $Y = X_2 + \cdots + X_n$. Define $f(T) = \sqrt{\mathrm{Var}(X_1)} - \sqrt{\mathrm{Var}(Y)}$. Obviously $f(T)$ is a continuous function of T with canonical distance measure. It is easy to check that $f(T) = \sigma_1 - \sum_{i=2}^n \sigma_i \le 0$ when $X_2 = \sigma_2 Z + \mu_2, \cdots, X_n = \sigma_n Z + \mu_n$ for some $Z \sim N(0, 1)$. Since $\sigma_1 \ge \sigma_2 \ge \cdots \ge \sigma_n$, we also have $f(T) = \sigma_1 - |\sum_{i=2}^n (-1)^i \sigma_i| \ge 0$ when $X_i = (-1)^i \sigma_i Z + \mu_i$ for $i = 2, \cdots, n$. Hence there exists a correlation matrix T_0 such that $f(T_0) = 0$. With the correlation matrix of $(X_2, \cdots, X_n)$ being T_0, we define $X_1 = -Y + \mathbb{E}(Y) + \mu_1$. Hence $X_1 \sim N(\mu_1, \sigma_1^2)$ and $\mathrm{Var}(X_1 + \cdots + X_n) = 0$, which imply that $F_1, \cdots, F_n$ are JM.

Proof of Proposition 2.9.

When $s \ge nF^{-1}(1)$, we have $m_+(s) = 1 = \psi^{-1}(s/n)$. Suppose $s < nF^{-1}(1)$. Obviously

$$1 - n \inf_{r \in [0, s/n)} \frac{\int_r^{s-(n-1)r} (1 - F(t))\mathrm{d}t}{s - nr} \ge 1 - n \inf_{r \in [0, s/n)} \frac{\int_r^{\infty} (1 - F(t))\mathrm{d}t}{s - nr}. \tag{4.1}$$

For $r \in [0, \frac{s}{n})$, from $\left(\frac{\int_r^\infty (1-F(t))\mathrm{d}t}{s-nr}\right)' = 0$, we have

$$g(r) := -(1 - F(r))(s - nr) + n \int_r^\infty (1 - F(t))\mathrm{d}t = 0. \tag{4.2}$$

Suppose $r = r^\star$ satisfies (4.2). Then

$$1 - F(r^\star) = \frac{n \int_{r^\star}^\infty \{1 - F(t)\}\mathrm{d}t}{s - nr^\star}. \tag{4.3}$$

Note that $r^\star$ always exists since g is continuous, $g(0) = -s + n\mu < 0$, $F(s/n) < 1$ and

$$g(s/n) = n \int_{s/n}^\infty \{1 - F(t)\}\mathrm{d}t > 0.$$

Integration by parts leads to

$$-\{1 - F(r^\star)\}(s - nr^\star) + n \int_{r^\star}^\infty \{1 - F(t)\}\mathrm{d}t = -s\{1 - F(r^\star)\} + n \int_{r^\star}^\infty t\mathrm{d}F(t) = 0,$$

and hence

$$s(1 - F(r^\star)) = n\mathbb{E}(X_1|X_1 > r^\star)(1 - F(r^\star)),$$

i.e.,

$$F(r^\star) = \psi^{-1}(s/n). \tag{4.4}$$

Therefore, the bound in (2.8) is greater than or equal to the bound in (2.6) by (4.1), (4.3) and (4.4). Note that the bound (2.8) is strictly greater if $F^{-1}(1) = \infty$.

For proving the second part of **Proposition 2.9**, we only need to consider the case of $s < nF^{-1}(1) < \infty$ since $m_+(s) = 1 = \psi^{-1}(s/n)$ when $nF^{-1}(1) \le s < \infty$.

Consider the problem (2.9) and

$$\inf_{r\in[0,s/n)} \frac{\int_r^{F^{-1}(1)} (1 - F(t))\mathrm{d}t}{s - nr}. \tag{4.5}$$

From the above proof, we can see that $r^\star$ is the unique solution to (4.5). Therefore, the bounds (2.6) and (2.8) are equal if and only if (2.9) and (4.5) are equal. Since

$$g\left(\frac{s - F^{-1}(1)}{n - 1}\right) = -\left(1 - F\left(\frac{s - F^{-1}(1)}{n - 1}\right)\right) \frac{nF^{-1}(1) - s}{n - 1} + n \int_{\frac{s-F^{-1}(1)}{n-1}}^1 (1 - F(t))\mathrm{d}t \ge 0,$$

we have $r^\star \in [0, \frac{s-F^{-1}(1)}{n-1}]$. Thus

$$\begin{aligned} \inf_{r\in[0,s/n)} \frac{\int_r^{F^{-1}(1)} (1 - F(t))\mathrm{d}t}{s - nr} &= \inf_{r\in[0,\frac{s-F^{-1}(1)}{n-1}]} \frac{\int_r^{F^{-1}(1)} (1 - F(t))\mathrm{d}t}{s - nr} \\ &= \inf_{r\in[0,\frac{s-F^{-1}(1)}{n-1}]} \frac{\int_r^{s-(n-1)r} (1 - F(t))\mathrm{d}t}{s - nr}. \end{aligned}$$

Therefore, the bounds (2.6) and (2.8) are equal if and only if a solution to (2.9) lies in $[0, \frac{s-F^{-1}(1)}{n-1}]$. □

Proof of Proposition 3.1.

1. By (a) in Section 3.1, for any $i \neq j$, $U_i \in [0, c] \Rightarrow U_j \in [1-(n-1)c, 1]$. Hence

$$A_i := \{U_i \in [0, c]\} \subseteq \{U_j \in [1-(n-1)c, 1]\} =: B_j$$

and $\mathbb{P}(A_i \cap A_j) = 0$. As a consequence, $\bigcup_{i \neq j} A_i \subseteq B_j$. Note that $\mathbb{P}(\bigcup_{i \neq j} A_i) = (n-1)c = \mathbb{P}(B_j)$. Thus $\mathbf{1}_{\bigcup_{i \neq j} A_i} = \mathbf{1}_{B_j}$ a.s. and

$$\mathbf{1}_{\bigcup_{i=1}^n A_i} = \mathbf{1}_{A_j \cup B_j} = \mathbf{1}_{\{U_j \in [0,c] \cup [1-(n-1)c, 1]\}} \quad \text{a.s.}$$

which imply that $\mathbf{1}_{\{U_j \in (c, 1-(n-1)c)\}} = \mathbf{1}_{(\bigcup_{i=1}^n A_i)^c}$ a.s. for $j = 1, \cdots, n$.

2. We only prove the case when F has an increasing density. When $c_n = 0$, (3.6) follows from the definition of Q_n^F. Next we assume $c_n > 0$. Write $D_j = A_j \cup B_j$ and $X_j = F^{-1}(U_j)$, $U_j \sim$ U[0,1] for $j = 1, \cdots, n$. First note that by condition (b) in Section 3.1, for any $j = 1, \cdots, n$, $F^{-1}(U_1) + \cdots + F^{-1}(U_n)$ is a constant on the set D_j^c. This constant equals its expectation, which is

$$\begin{aligned}
&\mathbb{E}(F^{-1}(U_1) + \cdots + F^{-1}(U_n) | D_j^c) \\
&= n\mathbb{E}(F^{-1}(U_1) | D_1^c) \\
&= \frac{n}{1-nc_n} \int_{c_n}^{1-(n-1)c_n} F^{-1}(x) \mathrm{d}x \\
&= \frac{n}{1-nc_n} \int_{c_n}^{\frac{1}{n}} F^{-1}(x) \mathrm{d}x + \frac{n}{1-nc_n} \int_{\frac{1}{n}}^{1-(n-1)c_n} F^{-1}(x) \mathrm{d}x \\
&= \frac{n}{1-nc_n} \int_{c_n}^{\frac{1}{n}} F^{-1}(x) \mathrm{d}x + \frac{n}{1-nc_n} \int_{c_n}^{\frac{1}{n}} F^{-1}(1-(n-1)t) \mathrm{d}(n-1)t \\
&= \frac{n}{1-nc_n} \int_{c_n}^{\frac{1}{n}} H(x) \mathrm{d}x \\
&= H(c_n).
\end{aligned}$$

The last equality holds because (3.2) and

$$\int_{c_n}^{\frac{1}{n}} H(x) \mathrm{d}x = (\frac{1}{n} - c_n) H(c_n) \quad \text{for} \quad c_n > 0.$$

Therefore, almost surely

$$\begin{aligned} S &= F^{-1}(U_1) + \cdots + F^{-1}(U_n) \\ &= \sum_{i=1}^{n} F^{-1}(U_i)\mathbf{1}_{D_1} + \sum_{i=1}^{n} F^{-1}(U_i)\mathbf{1}_{D_1^c} \\ &= \sum_{i=1}^{n} F^{-1}(U_i)\mathbf{1}_{\bigcup_{j=1}^{n} A_j} + H(c_n)\mathbf{1}_{D_1^c} \\ &= \sum_{i=1}^{n} F^{-1}(U_i)(\sum_{j=1}^{n}\mathbf{1}_{A_j}) + H(c_n)\mathbf{1}_{D_1^c} \\ &= \sum_{j=1}^{n}[F^{-1}(U_j) + (n-1)F^{-1}(1-(n-1)U_j)]\mathbf{1}_{A_j} + H(c_n)\mathbf{1}_{D_1^c} \\ &= \sum_{j=1}^{n} H(U_j)\mathbf{1}_{A_j} + H(c_n)\mathbf{1}_{D_1^c}. \end{aligned}$$

Since $c_n \le 1/n$ and the sets $A_1, \cdots, A_n$ and D_1^c are disjoint, we have

$$\begin{aligned} &\mathbb{P}(\sum_{j=1}^{n} H(U_j)\mathbf{1}_{A_j} + H(c_n)\mathbf{1}_{D_1^c} < t) \\ &= n\mathbb{P}(H(U_1)\mathbf{1}_{\{U_1\le c_n\}} < t) + \mathbb{P}(H(c_n)\mathbf{1}_{D_1^c} < t) \\ &= \mathbb{P}(H(U_1/n)\mathbf{1}_{\{U_1\le nc_n\}} < t) + \mathbb{P}(H(c_n)\mathbf{1}_{\{U_1>nc_n\}} < t) \\ &= \mathbb{P}(H(U_1/n)\mathbf{1}_{\{U_1\le nc_n\}} + H(c_n)\mathbf{1}_{\{U_1>nc_n\}} < t). \end{aligned}$$

Hence there exists a U ∼U[0,1] such that

$$\sum_{j=1}^{n} H(U_j)\mathbf{1}_{A_j} + H(c_n)\mathbf{1}_{D_j^c} = H(U/n)\mathbf{1}_{\{U\le nc_n\}} + H(c_n)\mathbf{1}_{\{U>nc_n\}}.$$

□

Proof of Lemma 3.3

(i) Under the assumption of F, $F^{-1}(x)$ is convex and differentiable. Thus $H_a(x)$ is convex and differentiable. The definition of $c_n(a)$ shows that the average of $H_a(x)$ on $[c_n(a), \frac{1}{n}(1-a)]$ is $H_a(c_n(a))$ if $0 < c_n(a) < \frac{1-a}{n}$, namely

$$\frac{1}{(1-a) - c_n(a)} \int_{c_n(a)}^{\frac{1}{n}(1-a)} H_a(t)\mathrm{d}t = H_a(c_n(a)).$$

With $H_a(x)$ being convex, we have $H_a'(c_n(a)) \le 0$ and so $H_a'(x) \le 0$ on $[0, c_n(a)]$. Here $H_a'(x)$ denotes $\partial H_a(x)/\partial x$. Note that for $n > 2$, $H_a'(\frac{1-a}{n}) = ((n-1)^2 - 1)(F^{-1})'(\frac{1-a}{n}) > 0$ implies

$$\int_{c}^{\frac{1}{n}(1-a)} H_a(t)\mathrm{d}t \ge (\frac{1}{n}(1-a) - c)H_a(c)$$

for some $c < \frac{1-a}{n}$, thus $c_n(a) < \frac{1-a}{n}$ always holds. For $n = 2$, $H_a'(x) \le 0$ on $[0, \frac{1-a}{n}]$ since $H_a'(\frac{1-a}{n}) = 0$ and H is convex.

(ii) It follows from similar arguments as in (i).

(iii) Suppose $c_n(a) > 0$. By the continuity of $H_a(x)$ w.r.t. x and (3.8), we know that $c_n(a)$ satisfies

$$\int_{c_n(a)}^{\frac{1}{n}(1-a)} H_a(t)\mathrm{d}t = (\frac{1}{n}(1-a) - c_n(a))H_a(c_n(a)).$$

Note that for any $c \in [0, \frac{1}{n}(1-a)]$,

$$\begin{aligned}\int_c^{\frac{1}{n}(1-a)} H_a(t)\mathrm{d}t &= \int_c^{\frac{1}{n}(1-a)} (n-1)F^{-1}(a+(n-1)t)\mathrm{d}t + \int_c^{\frac{1}{n}(1-a)} F^{-1}(1-t)\mathrm{d}t \\ &= \int_{a+(n-1)c}^{a+\frac{n-1}{n}(1-a)} F^{-1}(t)\mathrm{d}t + \int_{1-\frac{1}{n}(1-a)}^{1-c} F^{-1}(t)\mathrm{d}t \\ &= \int_{a+(n-1)c}^{1-c} F^{-1}(t)\mathrm{d}t.\end{aligned}$$

Thus it follows from the definition of $c_n(a)$ that $H_a(c_n(a)) = n\mathbb{E}[F^{-1}(V_a)]$. For the case $c_n(a) = 0$, it is obvious that $\psi(a) = n\phi(a) = n\mathbb{E}[F^{-1}(V_a)]$.

(iv) Note that in a given probability space, for any measurable set B with $\mathbb{P}(B) > 0$ and continuous random variable Z with cdf G, we have

$$\mathbb{E}(Z|B) \le \mathbb{E}[Z|Z \ge G^{-1}(1-\mathbb{P}(B))].$$

To see this, denote the conditional distribution of Z on B by G_1 and the conditional distribution on $\{Z \ge G^{-1}(1-\mathbb{P}(B))\}$ by G_2. Then we have

$$\begin{aligned}G_2(x) &= \frac{\mathbb{P}(Z \le x, G(Z) \ge 1-\mathbb{P}(B))}{\mathbb{P}(B)} \\ &= \frac{\max\{G(x) - 1 + \mathbb{P}(B), 0\}}{\mathbb{P}(B)} \\ &\le \frac{\mathbb{P}(Z \le x, B)}{\mathbb{P}(B)} = G_1(x),\ x \in \mathbb{R},\end{aligned} \tag{4.6}$$

which implies that for $U \sim$U[0,1],

$$\mathbb{E}(Z|B) = \mathbb{E}[G_1^{-1}(U)] \le \mathbb{E}[G_2^{-1}(U)] = \mathbb{E}[Z|Z \ge G^{-1}(1-\mathbb{P}(B))]. \tag{4.7}$$

Since $A = \bigcap_{i=1}^n \{U_i \in [a, 1-c_n(b)]\}$, we have $\mathbb{P}(A) \ge 1 - \frac{nc_n(b)}{1-a} > 0$ and $U_i \le 1 - c_n(b)$ on A. By defining $Z = F^{-1}(U_i)\mathbf{1}_{\{U_i \le 1-c_n(b)\}} + F^{-1}(a)\mathbf{1}_{\{U_i > 1-c_n(b)\}}$, it follows from (4.7) that

$$\begin{aligned}\mathbb{E}[F^{-1}(U_i)|A] &= \mathbb{E}[Z|A] \\ &\le \mathbb{E}[Z|Z \ge F^{-1}(1-c_n(b)-(1-a)\mathbb{P}(A))] \\ &\le \mathbb{E}[F^{-1}(U_i)|U_i \in [1-c_n(b)-(1-a)\mathbb{P}(A), 1-c_n(b)]] \\ &\le \mathbb{E}[F^{-1}(U_i)|U_i \in [a+(n-1)c_n(b), 1-c_n(b)]] \\ &< \mathbb{E}[F^{-1}(U_i)|U_i \in [b+(n-1)c_n(b), 1-c_n(b)]] \\ &= \mathbb{E}(F^{-1}(V_b)).\end{aligned} \tag{4.8}$$

(v) It follows from (i), (ii) and the arguments in *Remark* 3.2.

(vi) We first prove the case when F has a decreasing density. Since $H_a(x)$ is convex w.r.t. x and differentiable w.r.t. a, the definition of $c_n(a)$ implies that $c_n(a)$ is continuous. Hence $\phi(a) = n\mathbb{E}[F^{-1}(V_a)]$ is continuous.

Suppose $U_{a,1}, \cdots, U_{a,n} \sim \mathrm{U}[a, 1]$ with copula $Q_n^{\tilde{F}_a}$. Then $F^{-1}(U_{a,1}), \cdots, F^{-1}(U_{a,n}) \sim \tilde{F}_a$ and have copula $Q_n^{\tilde{F}_a}$ too. By (v), we have

$$F^{-1}(U_{a,1}) + \cdots + F^{-1}(U_{a,n}) \geq \phi(a). \tag{4.9}$$

Thus from (4.8) and (4.9) we have

$$\phi(a) \leq \mathbb{E}[\sum_{i=1}^{n} F^{-1}(U_{a,i})|A] < n\mathbb{E}(F^{-1}(V_b)) = \phi(b).$$

Next we prove the case when F has an increasing density. The continuity of $c_n(a)$ comes from the same arguments as above. By definition, $H_a(0)$ and $\psi(a)$ are continuous and increasing functions of a. So we only need to show that when $c_n(a)$ approaches 0, $H_a(0) - \psi(a)$ approaches 0 too. Suppose that as $a \nearrow a_0$, $c_n(a) \to 0$ and $c_n(a) \neq 0$ for $a_0 - \epsilon < a < a_0$ and $\epsilon > 0$. Then

$$\int_0^{\frac{1}{n}(1-a)} H_a(t)\mathrm{d}t \to \frac{1}{n}(1 - a_0)H_{a_0}(0),$$

which implies that

$$\psi(a) = \int_a^1 \frac{1}{1-a} F^{-1}(a + t)\mathrm{d}t = \frac{n}{1-a} \int_0^{\frac{1}{n}(1-a)} H_a(t)\mathrm{d}t \to H_{a_0}(0)$$

as $a \nearrow a_0$. Together with the continuity of $H_a(0)-\psi(a)$ we know $H_a(0)-\psi(a) \to 0$ as $a \to a_0$. □

Acknowledgment. We thank Co-editor Professor Kerry Back, an associate editor and two reviewers for their helpful comments which significantly improved this paper. Wang's research was partly supported by the Bob Price Fellowship at the Georgia Institute of Technology. Peng's research was supported by NSF Grant DMS-1005336. Yang's research was supported by the Key Program of National Natural Science Foundation of China (Grants No. 11131002).

References

1. Denuit, M., Genest, C., Marceau, É. (1999). *Stochastic bounds on sums of dependent risks*. Insurance: Mathematics and Economics **25**, 85-104.
2. Embrechts, P., Höing, A. and Juri, A. (2003). *Using copulae to bound the value-at-risk for functions of dependent risks*. Finance and Stochastics **7**, 145-167.
3. Embrechts, P and Höing, G. (2006). *Extreme VaR scenarios in higher dimensions*. Extremes **9**, 177-192.
4. Embrechts, P and Puccetti, G. (2006a). *Bounds for functions of dependent risks*. Finance and Stochastics **10**, 341-352.
5. Embrechts, P. and Puccetti, G. (2006b). *Bounds for functions of multivariate risks*. Journal of Multivariate Analysis **97**, 526-547.

6. Embrechts, P. and Puccetti, G. (2010). *Risk Aggregation*. In: Copula Theory and Its Applications, P.Jaworski et al. (eds.), Lecture Notes in Statistics **198**, Springer, Berlin/Heidelberg.
7. Gaffke, N. and Rüschendorf, L. (1981). *On a class of extremal problems in statistics*. Math. Operationsforschung Statist., Series Optimization **12**, 123-135.
8. Kaas, R., Laeven, R. and Nelsen, R. (2009). *Worst VaR scenarios with given marginals and measures of association*. Insurance: Mathematics and Economics **44**, 146-158.
9. Knott, M. and Smith, C.S. (2006). *Choosing joint distributions so that the variance of the sum is small*. Journal of Multivariate Analysis **97**, 1757-1765.
10. Puccetti, G. and Rüschendorf, L. (2012) *Bounds for joint portfolios of dependent risks.* J. Comp. Appl. Math. **236**(7), 1833-1840.
11. Rüschendorf, L. (1982). *Random variables with maximum sums*. Adv. Appl. Probab. **14**(3), 623-632.
12. Rüschendorf, L. (1983). *Solution of a statistical optimization problem by rearrangement methods*. Metrika **30**, 55-61.
13. Rüschendorf, L. and Uckelmann, L. (2002). *Variance minimization and random variables with constant sum,* in: Distributions with given marginals. Cuadras, et al. (Eds.), Kluwer, 211-222.
14. Wang, B. and Wang, R. (2011). *The complete mixability and convex minimization problems for monotone marginal distributions*. Journal of Multivariate Analysis **102**, 1344-1360.