Sharp bounds for sums of dependent risks

Giovanni Puccetti, Ludger Rüschendorf

December 14, 2011

Abstract

Sharp tail bounds for the sum of d random variables with given marginal distributions and arbitrary dependence structure are known from Makarov [4] and Rüschendorf [9] for $d = 2$ and, in some examples, for $d \geq 3$. In the homogeneous case $F_1 = \cdots = F_n$ with monotone density sharp bounds were found in Wang and Wang [11]. In this paper we derive sharp bounds for the tail risk of joint portfolios in the homogeneous case under general conditions which include in particular the case of monotone densities and concave densities. It turns out that the dual bounds of Embrechts and Puccetti [1] are sharp under our conditions.

Key words: Bounds for dependent risks, Fréchet bounds, mass transportation theory
AMS 2010 Subject Classification: 60E05, 91B30

1 Introduction

For a risk vector $\boldsymbol{X} = (X_1, \ldots, X_d)$, $d \geq 2$, we consider the problem to find sharp bounds for the tail probability of the sum $S = \sum_{i=1}^d X_i$ under the condition that the marginal distribution functions F_i of X_i are known but the dependence structure of $\boldsymbol{X}$ is completely unknown. Denoting by $\mathfrak{F}(F_1, \ldots, F_d)$ the Fréchet class of all joint distribution functions on $\mathbb{R}^d$ with marginal distribution functions F_i, we study the problem to determine

$$M(s) = \sup \{P(X_1 + \cdots + X_d \geq s); F_{\boldsymbol{X}} \in \mathfrak{F}(F_1, \ldots, F_d)\}. \tag{1.1}$$

The problem of obtaining tail bounds as in (1.1) is relevant in quantitative risk management since bounds for the distribution and for the tail risk of the joint portfolio are needed to compute bounds risk measures live the value-at-risk for regulatory purposes. For the motivation of this problem, we refer to [2]. A survey of the various approaches and literature of recent results on this problem is given in [5]. Sharp tail bounds for $d = 2$ were given independently in [4] and [9]. For any $s \in \mathbb{R}$, we have

$$\sup \{P(X_1 + X_2 \geq s) : X_i \sim F_i\} = \inf_{x \in \mathbb{R}} \left\{ \overline{F}_1(x-) + \overline{F}_2(s - x) \right\}, \tag{1.2}$$

where $\overline{F}_i(x) = 1 - F_i(x) = P(X_1 > x)$ and $\overline{F}_1(x-) = P(X_1 \geq x)$. For the case $d \geq 3$, [1] give an upper bound for the tail probability in the homogeneous case $F_1 = \cdots = F_d =$

F based on the following duality result (see [8, Theorem 5] and [3]). We denote by $\mathbf{1}(A)$ the indicator function of the set A. For notational simplicity, we write for instance $\mathbf{1}(x \geq 0)$ instead of $\mathbf{1}(\{x \geq 0\})$.

Theorem 1.1 (Duality theorem) *In the homogeneous case $F_i = F, 1 \leq i \leq d$, we have that:*

1. *Problem* (1.1) *has the following dual representation:*

$$M(s) = \inf\left\{d \int g(x)\, dF(x) : g \in \mathcal{D}(s)\right\}, \tag{1.3}$$

where

$$\mathcal{D}(s) = \left\{g : \mathbb{R} \to \mathbb{R}; g \text{ bounded}, \sum_{i=1}^{d} g(x_i) \geq \mathbf{1}\left(\sum_{i=1}^{d} x_i \geq s\right) \text{ for } x_1, \ldots, x_d \in \mathbb{R}\right\}.$$

An optimal dual solution $g^* \in \mathcal{D}(s)$ such that $M(s) = d \int g^*\, dF$ exists.

2. *A random vector $\boldsymbol{X}^*$ with distribution $F_{\boldsymbol{X}^*} \in \mathfrak{F}(F, \ldots, F)$ is a solution of $M(s) = P(\sum_{i=1}^{d} X_i^* \geq s)$ if and only if there exists an admissible function $g^* \in \mathcal{D}(s)$ such that*

$$P\left(\sum_{i=1}^{d} g^*(X_i^*) = \mathbf{1}\left(\sum_{i=1}^{d} X_i^* \geq s\right)\right) = 1. \tag{1.4}$$

A simple compactness argument shows that the sup in (1.1) *is attained and any solution $\boldsymbol{X}^*$ such that $M(s) = P(\sum_{i=1}^{d} X_i^* \geq s)$ is called an* optimal coupling.

[1] introduce the following class of piecewise-linear functions defined, for $t < s/d$, as

$$g_t(x) := \begin{cases} 0, & \text{if } x < t, \\ \frac{x-t}{s-dt} & \text{if } t \leq x \leq s - (d-1)t, \\ 1, & \text{otherwise.} \end{cases} \tag{1.5}$$

They establish that g_t are admissible, that is $g_t \in \mathcal{D}(s)$, and define the so-called *dual bound* $D(s)$ as

$$D(s) = \inf_{t<s/d}\left(d \int g_t\, dF\right) = d \inf_{t<s/d} \min\left\{\frac{\int_t^{s-(d-1)t} \overline{F}(x)\, dx}{s - dt}, 1\right\}. \tag{1.6}$$

In the homogeneous case $F_i = F, 1 \leq i \leq d$, the admissibility of the function g_t implies that

$$M(s) \leq D(s), \;\; s \in \mathbb{R}.$$

The dual bound $D(s)$ is numerically easy to evaluate independently of the size d of the portfolio $\boldsymbol{X}$. Based on the results of a numerical algorithm, sharpness of the dual bound ($M(s) = D(s)$) was conjectured in [6]. In a recent work of [11] and [10] based on the concept of *complete mixability*, optimal couplings $\boldsymbol{X}^*$ for problem (1.1) were found for the class of distribution functions F with monotone densities.

In this paper we derive sharp bounds for the tail of sums in the homogeneous case posing an attainment, a mixing and an ordering condition (see (**A1**)– (**A3**) below). Our main result implies sharpness of the dual bounds of [1] under these conditions. It implies in particular the results of [11], resp. [10] in the case of monotone densities and gives a strongly simplified proof. It also implies sharp bounds for further cases like the case of concave densities and for distributions which are typically used in quantitative risk management. In addition to the results stated in the above mentioned papers, we not only derive the optimal couplings but also give an effective method to calculate the sharp bounds.

The proofs of our main results (Proposition 2.3 and Proposition 2.4 below) are based on the complete mixability of the optimal dual function g^*, more precisely of $g^*(X_i)$. Therefore, we start with a summary of results on completely mixable distributions, used frequently in the remainder of the paper.

1.1 Some preliminaries on complete mixability

The following results on complete mixability can be found in [11] and references therein.

Definition 1.2 *A distribution function F on $\mathbb{R}$ is called d-*completely mixable *(d-CM) if there exist d random variables $X_1, \dots, X_d$ identically distributed as F such that*

$$P(X_1 + \cdots + X_d = d\mu) = 1, \tag{1.7}$$

*for some $\mu \in \mathbb{R}$. Any such μ is called a center of F and any vector $(X_1, \dots, X_d)$ satisfying (1.7) with $X_i \sim F, 1 \le i \le n$, is called a d-*complete mix.

If F is d-CM and has finite mean, then its center is unique and equal to its mean.

Definition 1.3 *If X has distribution F, we say that F is d-CM on the interval $A \subset \mathbb{R}$ if the conditional distribution of $(X|X \in A)$ is d-CM.*

Theorem 1.4 *The following statements hold.*

1. *The convex sum of d-CM distributions with center μ is d-CM with center μ.*

2. *Any linear transformation $L(x) = mx + q$ of a d-CM distribution with center μ is d-CM with center $m\mu + q$.*

3. *The Binomial distribution $B(n, p/q), p, q \in \mathbb{N}$ is q-CM.*

4. *Suppose F is a distribution on the real interval $[a, b]$, $a = F^{-1}(0)$ and $b = F^{-1}(1)$, having mean μ. A necessary condition for F to be d-CM is that*

$$a + (b - a)/d \le \mu \le b - (b - a)/d. \tag{1.8}$$

5. *If F is continuous with a monotone density on $[a, b]$, then condition (1.8) is sufficient for F to be d-CM.*

2 Sharpness of dual bounds

In our main result, we state some general conditions which imply that, if the infimum in (1.6) is attained at $t = a < s/d$, the dual bound

$$D(s) = \inf_{t<s/d} d \int g_t \, dF = d \int g_a \, dF$$

is sharp, that is $D(s) = M(s)$. The proof uses the following property of optimal couplings (see Proposition 3(c) in [9]).

Theorem 2.1 *For any marginal distribution F there exists an optimal coupling $\boldsymbol{X}^*$ with distribution $F_{\boldsymbol{X}^*} \in \mathfrak{F}(F, \ldots, F)$ such that $M(s) = P\left(\sum_{i=1}^d X_i^* \geq s\right)$ and for any such $\boldsymbol{X}^*$ we have*

$$\left\{X_i^* > F^{-1}(1 - M(s))\right\} \subset \left\{\sum_{i=1}^d X_i^* \geq s\right\} \subset \left\{X_i^* \geq F^{-1}(1 - M(s))\right\} \text{ a.s..} \tag{2.1}$$

In case F is continuous, one gets that

$$\left\{\sum_{i=1}^d X_i^* \geq s\right\} = \left\{X_i^* \geq F^{-1}(1 - M(s))\right\} \text{ a.s..}$$

Theorem 2.1 allows to reduce the class of admissible functions $\mathcal{D}(s)$ in Theorem 1.1. First, note that any $g \in \mathcal{D}(s)$ has to be nonnegative since

$$d\, g(x) \geq \mathbf{1}(x \geq s/d) \geq 0.$$

Then, combining point 2. in Theorem 1.1 with Theorem 2.1, we obtain that, if $g^* \in \mathcal{D}(s)$ is an optimal choice for (1.3), then

$$P\left(\sum_{i=1}^d g^*(X_i^*) = 0 \,\middle|\, \sum_{i=1}^d X_i^* < s\right) = P\left(\sum_{i=1}^d g^*(X_i^*) = 0 \,\middle|\, X_i^* < a^*\right) = 1, \tag{2.2}$$

where the second equality in the above equation follows from (2.1) with $a^* = F^{-1}(1 - M(s))$. Since g^* is non-negative, we conclude from (2.2) that

$$P\left(g^*(X_i^*) = 0 \,\middle|\, X_i^* < a^*\right) = 1, \ 1 \leq i \leq d. \tag{2.3}$$

As a consequence, any optimal dual choice g^* is a.s. zero on the interval $(-\infty, a^*)$. This means that, in order to solve problem (1.3), it is sufficient to determine the behavior of an optimal function g^* above the threshold a^*. This behavior is illustrated by the following theorem.

Theorem 2.2 *Let $a^* = F^{-1}(1 - M(s))$. A random vector $\boldsymbol{X}^*$ with distribution $F_{\boldsymbol{X}^*} \in \mathfrak{F}(F, \ldots, F)$ is a solution of $M(s) = P(\sum_{i=1}^d X_i^* \geq s)$ if and only if there exists an admissible function $g^* \in \mathcal{D}(s)$ such that the conditional distribution of*

$$\left(g^*(X_1) \,\middle|\, X_1 \geq a^*\right), \ 1 \leq i \leq d,$$

is d-CM with center $\mu = 1/d$.

Proof. Assume that $\boldsymbol{X}^*$ and g^* are an optimal coupling and, respectively, an optimal dual function for. By (2.3), we can assume that any $g^* \in \mathcal{D}(s)$ is zero below the threshold a^*. Using (1.4) and (2.1) similarly as in (2.2), we obtain that $g^* \in \mathcal{D}$ is an optimal choice for (1.3) if and only if

$$P\left(\sum_{i=1}^{d} g^*(X_i^*) = 1 \,\middle|\, \sum_{i=1}^{d} X_i^* \geq s\right) = P\left(\sum_{i=1}^{d} g^*(X_i^*) = 1 \,\middle|\, X_i^* \geq a^*\right) = 1, \tag{2.4}$$

for $1 \leq i \leq d$. □

We are now ready to prove the sharpness of the dual bound $D(s)$ defined in (1.6). We obtain this result in two steps. First, in Proposition 2.3 we state the complete mixability of the dual function g_t (see (1.5)) above a certain threshold a^* and for a suitable choice of the parameter $t = a$. Then, in Proposition 2.4, we show that $a^* = M^{-1}(s)$, hence obtaining the optimality of g_a.

Proposition 2.3 *In the homogeneous case $F_i = F, 1 \leq i \leq d$, with $d \geq 3$, let F be a continuous distribution and let X_1 have distribution F. For a real threshold s, suppose that it is possible to find a real value $a < s/d$ such that*

$$D(s) = \inf_{t<s/d} \frac{d\int_t^{s-(d-1)t} \overline{F}(x)dx}{(s-dt)} = \frac{d\int_a^b \overline{F}(x)dx}{(b-a)}, \tag{\textbf{A1}}$$

where $b = s - (d-1)a$, with $a^* = F^{-1}(1 - D(s)) \leq a$. Suppose also that

$$\text{the conditional distribution of } (X_1|X_1 \geq a^*) \text{ is } d\text{-CM on } (a, b). \tag{\textbf{A2}}$$

Then:

(i) The conditional distribution H of $(g_a(X_1)|X_1 \geq a^*)$ is d-CM with center $\mu = 1/d$.

(ii) We have that

$$M(s) \leq \overline{F}(a^*) = \overline{F}(a) + (d-1)\overline{F}(b). \tag{2.5}$$

Proof. (i) First order conditions on the argument of the infimum in (**A1**) at $t = a$ imply that

$$\frac{d\int_a^b \overline{F}(x)dx}{(b-a)} = \overline{F}(a) + (d-1)\overline{F}(b). \tag{2.6}$$

Therefore, $a^* \leq a$ satisfies

$$\overline{F}(a^*) = D(s) = \frac{d\int_a^b \overline{F}(x)dx}{(b-a)} = \overline{F}(a) + (d-1)\overline{F}(b). \tag{2.7}$$

Let $Y_{a^*} \stackrel{d}{=} (X_1|X_1 \geq a^*)$. We have to show that the distribution H of $g_a(Y_{a^*})$ is d-CM. From the definition (1.5) of the linear functions $g_t, t < s/d$, it follows that H

is the convex sum of a continuous distribution G_1 on $(0,1)$ and of a discrete distribution G_2 on $\{0,1\}$. Formally, if we denote by G_1 the conditional distribution of $\left(g_a(Y_{a^*}) \,\middle|\, Y_{a^*} \in (a,b)\right)$, and we define the distribution G_2 as

$$G_2(x) = \frac{p_1}{p_1+p_3}\mathbf{1}(x \geq 0) + \frac{p_3}{p_1+p_3}\mathbf{1}(x \geq 1),$$

we can write H as

$$H = p_2 G_1 + (p_1+p_3)G_2,$$

where $p_1 = P(Y_{a^*} \leq a)$, $p_2 = P(a < Y_{a^*} \leq b)$, and $p_3 = 1 - p_2 - p_1 = P(Y_{a^*} > b)$. Note that G_1 is the distribution of a linear transformation of the random variable Y_{a^*} on the interval (a,b). Using the assumption of complete mixability of the distribution of Y_{a^*} on (a,b) and point 5. in Theorem 1.4, it follows that G_1 is d-CM with center given by

$$\int x\, dG_1(x) = \int_a^b \frac{(x-a)dF(x)}{(b-a)(\overline{F}(a)-\overline{F}(b))}\, dx = \frac{\frac{\int_a^b \overline{F}(x)dx}{(b-a)} - \overline{F}(b)}{\overline{F}(a)-\overline{F}(b)} = 1/d. \tag{2.8}$$

Similarly, the mean of G_2 is given by

$$\int x\, dG_2(x) = \frac{p_3}{p_1+p_3} = \frac{\overline{F}(b)}{\overline{F}(b)+\overline{F}(a^*)-\overline{F}(a)} = 1/d. \tag{2.9}$$

In the above equations (2.8) and (2.9), the last equalities follow from (2.7). Note that the distribution G_2 is a Binomial $B(1, 1/d)$. By point 3. in Theorem 1.4, also G_2 is d-CM with center $1/d$. The distribution H is then the convex combination of two d-CM distributions with the same center. Thus, by point 1. in Theorem 1.4, H is d-CM with center $1/d$.

(ii) Inequality (2.5) is a direct consequence of the fact that $g_a \in \mathcal{D}(s)$. Thus

$$M(s) \leq d \int g_a\, dF = \frac{d\int_a^b \overline{F}(x)dx}{(b-a)} = \overline{F}(a) + (d-1)\overline{F}(b),$$

where, in the above equation, the last equality follows from (2.6). □

Postulating the optimality of the the dual function g_a, it is possible to find a candidate for the optimal coupling in (1.1). The complete mixability of the distribution of Y_{a^*} on the interval (a,b) implies that there exist random variables $Y_1, \ldots, Y_d$ identically distributed as Y_{a^*} such that their sum is constant when one of them lies in (a,b). Moreover, using the complete mixability of the distribution of the random variable $g_a(Y_{a^*})$ on the set $\{0,1\}$, it is possible to construct random variables $Y_1, \ldots, Y_d$ identically distributed as Y_{a^*} such that

$$P\left(\bigcap_{i \neq j} \{Y_i \leq a\} \,\middle|\, Y_j > b\right) = 1.$$

It turns out that a random vector satisfying the properties listed above is optimal under an extra ordering assumption.

Proposition 2.4 *Under the assumptions of Proposition 2.3, suppose that*

$$(d-1)\,(F(y) - F(b)) \le F(a) - F\left(\frac{s-y}{d-1}\right), \tag{\textbf{A3}}$$

for all $y \ge b$. Then, there exist a random vector $\boldsymbol{X}^*$ with distribution $F_{\boldsymbol{X}^*} \in \mathfrak{F}(F,\dots,F)$ for which

$$P\left(\sum_{i=1}^{d} X_i^* \ge s\right) = \overline{F}(a^*).$$

Proof. For a^* satisfying (**A1**), denote by $F_{a^*}(x) = (F(x) - F(a^*))/\overline{F}(a^*)$ the distribution of the random variable $Y_{a^*} \overset{\mathrm{d}}{=} (X_1 | X_1 \ge a^*)$. We show that there exist a random vector $\boldsymbol{Y} = (Y_1,\dots,Y_d)$ with distribution $F_{\boldsymbol{Y}} \in \mathfrak{F}(F_{a^*},\dots,F_{a^*})$ for which

$$P\left(\sum_{i=1}^{d} Y_i \ge s\right) = 1. \tag{2.10}$$

This will imply the existence of a vector $\boldsymbol{X}^*$ such that $P\left(\sum_{i=1}^{d} X_i^* \ge s\right) = \overline{F}(a^*)$. For instance, $\boldsymbol{X}^*$ can be defined as

$$\boldsymbol{X}^* = (X_1,\dots,X_1)\,\mathbf{1}\left(\cup_{i=1}^{d}\{X_i \le a^*\}\right) + \boldsymbol{Y}\,\mathbf{1}\left(\cap_{i=1}^{d}\{X_i > a^*\}\right).$$

We define the vector $\boldsymbol{Y} = (Y_1,\dots,Y_d)$ with distribution $F_{\boldsymbol{Y}} \in \mathfrak{F}(F_{a^*},\dots,F_{a^*})$ as follows:

(a) When one of the Y_i's lies in the interval (a,b), then all the Y_i's lie in (a,b) and

$$P\left(Y_1 + \cdots + Y_d = s \;\middle|\; Y_i \in (a,b)\right) = 1;$$

(b) For all $1 \le i \le d$, we have that

$$P\left(Y_j = F_{a^*}^{-1}\left((d-1)\overline{F}_{a^*}(Y_i)\right) \;\middle|\; Y_i \ge b\right) = 1, \text{ for all } j \ne i.$$

First, we note that a random vector $\boldsymbol{Y}$ with properties (a) and (b) exists. From the mixing condition (**A2**), the distribution F_{a^*} is completely mixable on the interval (a,b). Using linearity of the function g_a in the interval (a,b) and (2.8), it is easy to see that the conditional distribution of $(Y_{a^*} | Y_{a^*} \in (a,b))$ has mean

$$\frac{\int_a^b x\,dF(x)}{F(b) - F(a)} = s/d.$$

Therefore, there exists a vector $\boldsymbol{Y}$ having marginals F_{a^*} and satisfying property (a). From (2.7), it follows that $F_{a^*}^{-1}\left((d-1)\overline{F}_{a^*}(b)\right) = a$. From property (b), we obtain that

$$P\left(Y_j \le a \;\middle|\; Y_i \ge b\right) = 1, \text{ for all } j \ne i.$$

Consequently, the properties (a) and (b) describe the behavior of the vector $\boldsymbol{Y}$ in disjoint and complementary sets of $\mathbb{R}^d$. It is straightforward to see that property (b) is coherent with the fact that the Y_i's are identically distributed as F_{a^*}.

As $\sum_{i=1}^d Y_i = s$ a.s. when all the Y_i's lie in the interval (a, b), in order to prove (2.10) it remains to show that $\sum_{i=1}^d Y_i \geq s$ when one of the Y_i's is larger than b. To this aim, we define the function $\psi : [b, +\infty) \to \mathbb{R}$ as

$$\psi(y) = y + (d-1)F_{a^*}^{-1}\Big((d-1)\overline{F}_{a^*}(y)\Big). \tag{2.11}$$

Note that $\psi(y) \geq s$ if and only if

$$(d-1)\overline{F}_{a^*}(y) \geq F_{a^*}\left(\frac{s-y}{d-1}\right).$$

Expressing the above equation in terms of F, and using (2.5), we obtain that $\psi(y) \geq s$, $y \geq b$, is equivalent to condition (**A3**). □

As a corollary of Proposition 2.3 and Proposition 2.4, we now state the main result of our paper.

Theorem 2.5 (Sharpness of dual bounds) *Under the* attainment formula (**A1**)*, the* mixing condition (**A2**) *and the* ordering condition (**A3**)*, the dual bounds are sharp, that is*

$$M(s) = D(s) = \inf_{t<s/d} \frac{d\int_t^{s-(d-1)t} \overline{F}(x)dx}{(s-dt)} = \frac{d\int_a^b \overline{F}(x)dx}{(b-a)}.$$

Proof. From Proposition 2.3 and Proposition 2.4, we obtain that $M(s) = \overline{F}(a^*)$ and that the conditional distribution of $(g_a(X_1)|X_1 > a^*)$ is d-CM with center $\mu = 1/d$. By Theorem 2.2, the function g_a is then a solution of the dual problem in (1.3) and, therefore

$$M(s) = d\int g_a\, dF = \frac{d\int_a^b \overline{F}(x)dx}{(b-a)} = \inf_{t<s/d} \frac{d\int_t^{s-(d-1)t} \overline{F}(x)dx}{(s-dt)} = D(s). \quad \square$$

Remark 2.6 1. (*Monotone densities*) All continuous distribution functions F having a positive and decreasing density f on the unbounded interval $(a^*, +\infty)$ satisfy the assumption (**A2**) and (**A3**). In this case, the conditional distribution of $(Y_{a^*}|Y_{a^*} \in (a, b))$ inherits a decreasing density from F and has mean $\mu = s/d$. By point 5. in Theorem 1.4, the distribution of the random variable Y_{a^*} is then d-CM on (a, b). Moreover, if F is continuous with a decreasing density, then F is concave and F^{-1} is differentiable and convex. Then, the function ψ defined in (2.11) turns out to be convex and

$$\psi(b) = b + (d-1)F_{a^*}^{-1}\Big((d-1)\overline{F}_{a^*}(b)\Big) = b + (d-1)a = s - (d-1)a + (d-1)a = s. \tag{2.12}$$

Differentiating ψ on a right neighborhood of b, we obtain

$$\psi'_+(b) = 1 - (d-1)^2 \frac{f(b)}{f(a)}.$$

If F also satisfies the attainment condition (**A1**), second order conditions on the argument of the infimum in (**A1**) at $t = a$ imply that

$$f(a) - (d-1)^2 f(b) \geq 0, \tag{2.13}$$

that is $\psi'_+(1-c) \geq 0$. Convexity of ψ and (2.12) finally imply that $\psi(x) \geq s$ for all $x \geq b$. In consequence, Theorem 2.5 implies as particular case the results in [11] and [10] for the case of monotone decreasing densities. The couplings used in the proof of Proposition 2.4 are of a similar form as in [10] in this case. In our paper we obtain a motivation for the structure of the optimal coupling and for the mixing from the duality characterization of optimal couplings in Theorem 2.2. Also equation (2.7) gives us an useful clue to the calculation of the sharp bound.

2. (*Monotonicity in the tail*) As a consequence of the remark above, sharpness of the dual bound $D(s)$ can be stated, for s large enough, for all continuous, unbounded distribution functions which have a ultimately decreasing density. This is particularly useful in applications of quantitative risk management, where sharp bounds $M(s)$ are typically calculated for high thresholds s and positive, unbounded and continuous distributions F. In particular, the Pareto distribution $F(x) = 1 - (1+x)^{-\theta}, x > 0$, with tail parameter $\theta > 0$, satisfies the assumptions (**A1**)–(**A3**) for all $s \in \mathbb{R}$ at which $D(s) < 1$. As a consequence, the bounds in Section 5.2 in [1] are sharp. We will give some numerical examples regarding the Pareto and other types of distributions in Section 3.

3. (*Concave densities*) All continuous distribution functions F having a concave density f on the interval (a, b) satisfy the mixing assumption (**A2**). This result follows from Theorem 4.3 in [7]. In order to obtain sharpness of the dual bound $D(s)$ for these distributions, conditions (**A1**) and (**A3**) has to be checked numerically. In Proposition 2.7 below, we give an equivalent formulation of (**A3**) in terms of stochastic order.

4. ($d = 2$) When $d = 2$, condition (**A1**) is typically not satisfied at a point $a < s/d$. For the sum of two random variables, the sharp bound (1.2) is obtained by an optimal dual function which is the average of indicator functions. In some cases (see Section 4 in [1]) it is possible that the sharp bound is still given by the dual bound $D(s)$ for $d = 2$, but the infimum in (1.6) is not attained.

5. (*Lower tails*) The sharp bound $M(s)$ for the upper tail of the sum $S = \sum_{i=1}^d X_i$ can be used to get sharp bounds for the lower tail of S, i.e. for

$$m(s) = \sup\left\{P\left(\sum_{i=1}^d X_i \leq s\right); F_X \in \mathfrak{F}(F, \dots, F)\right\}, \tag{2.14}$$

by switching the sign of the X_i's.

We conclude this section by giving an equivalent formulation of condition (**A3**) in terms of stochastic ordering. Define the distribution functions F_1 and F_2 as

$$F_1(y) = \frac{\overline{F}\left(\frac{s-y}{d-1}\right) - \overline{F}(a)}{\overline{F}(a^*) - \overline{F}(a)} \text{ and } F_2(y) = \frac{\overline{F}(b) - \overline{F}(y)}{\overline{F}(b)}, \text{ for } y \geq b.$$

Proposition 2.7 *Under the assumption of Proposition 2.3, inequality* (**A3**) *holds if and only if $F_2(y) \leq F_1(y)$ for all $y \geq b$, that is if and only if F_2 is stochastically larger than F_1 ($F_1 \leq_{st} F_2$).*

Proof. Using (2.7), the proposition immediately follows by noting that $F_2(y) \leq F_1(y)$, $y \geq b$ is equivalent to

$$\frac{\overline{F}(b) - \overline{F}(y)}{\overline{F}(b)} \geq \frac{\overline{F}\left(\frac{s-y}{d-1}\right) - \overline{F}(a)}{\overline{F}(a^*) - \overline{F}(a)} = \frac{\overline{F}\left(\frac{s-y}{d-1}\right) - \overline{F}(a)}{(d-1)\overline{F}(b)},$$

which is equivalent to (**A3**). □

Remark 2.8 We remark the following points about Proposition 2.7.

1. If F has a density f, then by a well known condition stochastic ordering $F_1 \leq_{\mathrm{st}} F_2$ is implied by the monotone likelihood ratio criterion for their densities stating that that f_2/f_1 is increasing or, equivalently, that

$$\frac{f(y)}{f\left(\frac{s-y}{d-1}\right)} \text{ is increasing in } y, \text{ for } y \geq b. \tag{2.15}$$

   This condition allows to be checked in examples.

2. For distributions with monotone densities, condition (**A3**) holds true. In several examples of non-monotone densities we found that condition (**A3**), resp. (2.15), is satisfied; see Section 3.

3 Applications and numerical verifications

Equation (2.6) provides a clue to calculate the basic point a and, hence, the dual bound $D(s)$. Having calculated a, one can easily check the second order condition (2.13) which is necessary to guarantee that a is a point of minimum for (**A1**). At this point, the sharpness of the dual bound $D(s)$ can be obtained from different sets of assumptions:

- If F has a positive and decreasing density f on $(a^*, +\infty)$, then the mixing condition (**A2**) and the ordering condition (**A3**) are satisfied and the dual bound $D(s)$ is sharp; see point 1. in Remark 2.6.

- If F has a concave density f on (a, b), then the the mixing condition (**A2**) is satisfied (see point 3. in Remark 2.6) and one has only to check the ordering condition (**A3**) to get sharpness of dual bounds. This can be done numerically or by using the increasing densities quotient as indicated in point 1. in Remark 2.8.

These two sets of assumptions cover the distribution functions F typically used in applications of quantitative risk management. In the following, we provide some illustrative examples in which sharpness of dual bounds holds.

In Figure 1, we plot the dual bound $D(s)$ in (1.6) for a random vector $\boldsymbol{X}$ of $d = 3$ Pareto(2)-distributed risks. In the same figure, we also provide numerical values for the sharp bounds $M(s)$, at some thresholds s of interest. These values have been calculated using the rearrangement algorithm introduced in [6]. In Figure 2- 3, we plot the dual bound $D(s)$ in (1.6) for a random vector $\boldsymbol{X}$ of $d = 3$ LogNormal(2,1)- and, respectively, Gamma(3,1)-distributed risks, with numerical values for the sharp bounds $M(s)$. Finally, in Figure 4, we plot the dual bound $D(s)$ in (1.6) for a random vector $\boldsymbol{X}$ of $d = 1000$ Pareto(2)-distributed risks. For high dimensionality $d > 30$, the computation of the numerical values for $M(s)$ is not possible via the rearrangement algorithm introduced in [6]. In the homogeneous case, the dual bound methodology is the only way to obtain sharp bounds $M(s)$ for high dimensional vectors of risks. At this point, it is important to remark that the computation of the dual bound $M(s)$ is completely analytical and based on the solution of a one-dimensional equation. Therefore, all the analytical curves in the above figures can be obtained within seconds and this independently of the dimension d of the vector $\boldsymbol{X}$.

References

[1] Embrechts, P. and G. Puccetti (2006). Bounds for functions of dependent risks. *Finance Stoch. 10*(3), 341–352.

[2] Embrechts, P. and G. Puccetti (2006b). Aggregating risk capital, with an application to operational risk. *Geneva Risk Insur. Rev. 31*(2), 71–90.

[3] Gaffke, N. and L. Rüschendorf (1981). On a class of extremal problems in statistics. *Math. Operationsforsch. Statist. Ser. Optim. 12*(1), 123–135.

[4] Makarov, G. D. (1981). Estimates for the distribution function of the sum of two random variables with given marginal distributions. *Theory Probab. Appl. 26*, 803–806.

[5] Puccetti, G. and L. Rüschendorf (2011a). Bounds for joint portfolios of dependent risks. *Preprint, University of Freiburg*.

[6] Puccetti, G. and L. Rüschendorf (2011b). Computation of sharp bounds on the distribution of a function of dependent risks. Forthcoming in *J. Comp. Appl. Math.*.

[7] Puccetti, G., Wang, B., and R. Wang (2011). Advances in complete mixability. Preprint.

[8] Rüschendorf, L. (1981). Sharpness of Fréchet-bounds. *Z. Wahrsch. Verw. Gebiete 57*(2), 293–302.

[9] Rüschendorf, L. (1982). Random variables with maximum sums. *Adv. in Appl. Probab. 14*(3), 623–632.

[10] Wang, R., L. Peng, and J. Yang (2011). Bounds for the sum of dependent risks and worst Value-at-Risk with monotone marginal densities. Preprint.

[11] Wang, B. and R. Wang (2011). The complete mixability and convex minimization problems with monotone marginal densities. *J. Multivariate Anal.,* in press.

Giovanni Puccetti, Department of Mathematics for Decision Theory, University of Firenze, via delle Pandette, 50127 Firenze, Italy, Tel. +39-0554374656, Fax +39-0554374913,
giovanni.puccetti@unifi.it

Ludger Rüschendorf, Department of Mathematical Stochastics, University of Freiburg, Eckerstr. 1, 79104 Freiburg, Germany, Tel. +49-761-2035669, Fax +49-761-2035661,
ruschen@stochastik.uni-freiburg.de

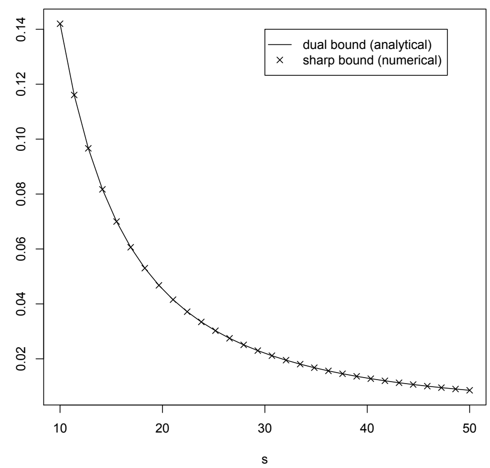

Figure 1: Dual bounds $D(s)$ (see (1.6)), for the sum of $d = 3$ Pareto(2)-distributed risks. Numerical values for the sharp bounds $M(s)$ are also provided at some thresholds s of interest.

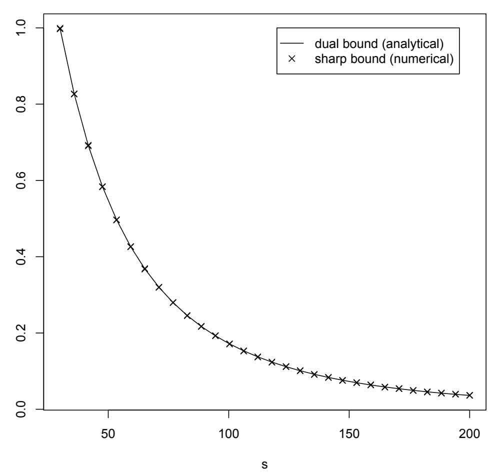

Figure 2: Dual bounds $D(s)$ (see (1.6)), for the sum of $d = 3$ LogNormal$(2, 1)$-distributed risks. Numerical values for the sharp bounds $M(s)$ are also provided at some thresholds s of interest.

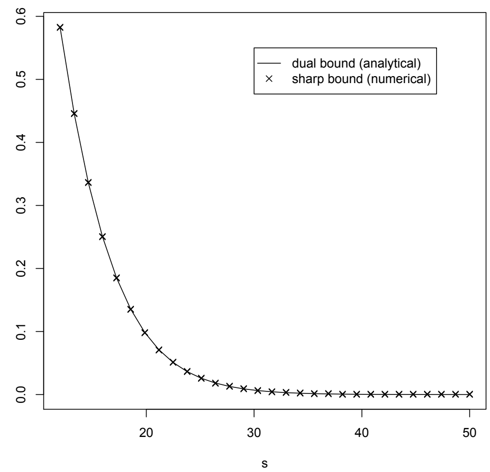

Figure 3: Dual bounds $D(s)$ (see (1.6)) for the sum of $d = 3$ Gamma(3, 1)-distributed risks. Numerical values for the sharp bounds $M(s)$ are also provided at some thresholds s of interest.

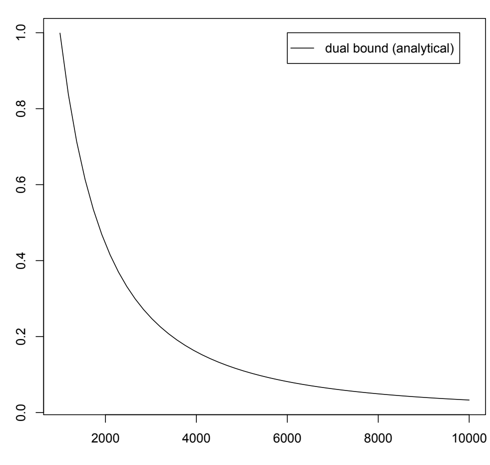

Figure 4: Dual bounds $D(s)$ (see (1.6)), for the sum of $d = 1000$ Pareto(2)-distributed risks.