J. Appl. Prob. **50,** 42–53 (2013)
Printed in England
© *Applied Probability Trust* 2013

SHARP BOUNDS FOR SUMS OF DEPENDENT RISKS

GIOVANNI PUCCETTI,* *University of Firenze*
LUDGER RÜSCHENDORF,** *University of Freiburg*

Abstract

Sharp tail bounds for the sum of d random variables with given marginal distributions and arbitrary dependence structure have been known since Makarov (1981) and Rüschendorf (1982) for $d = 2$ and, in some examples, for $d \geq 3$. Based on a duality result, dual bounds have been introduced in Embrechts and Puccetti (2006b). In the homogeneous case, $F_1 = \cdots = F_n$, with monotone density, sharp tail bounds were recently found in Wang and Wang (2011). In this paper we establish the sharpness of the dual bounds in the homogeneous case under general conditions which include, in particular, the case of monotone densities and concave densities. We derive the corresponding optimal couplings and also give an effective method to calculate the sharp bounds.

Keywords: Bounds for dependent risks; Fréchet bound; mass transportation theory

2010 Mathematics Subject Classification: Primary 60E05; 91B30

1. Introduction

For a risk vector $\boldsymbol{X} = (X_1, \ldots, X_d)$, $d \geq 2$, we consider the problem of finding 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 of determining

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

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 on risk measures, such as the value at risk for regulatory purposes. For the motivation of this problem, we refer the reader to Embrechts and Puccetti (2006a). A survey of the various approaches and literature of recent results on this problem is given in Puccetti and Rüschendorf (2012a). Sharp tail bounds for $d = 2$ were given independently in Makarov (1981) and Rüschendorf (1982). For any $s \in \mathbb{R}$, we have

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

where $\bar{F}_i(x) = 1 - F_i(x) = \mathbb{P}(X_1 > x)$ and $\bar{F}_1(x-) = \mathbb{P}(X_1 \geq x)$. For the case $d \geq 3$, Embrechts and Puccetti (2006b) gave an upper bound for the tail probability in the homogeneous

Received 3 January 2012.
* Postal address: Department of Mathematics for Decision Theory, University of Firenze, via delle Pandette, 50127 Firenze, Italy. Email address: giovanni.puccetti@unifi.it
** Postal address: Department of Mathematical Stochastics, University of Freiburg, Eckerstr. 1, 79104 Freiburg, Germany. Email address: ruschen@stochastik.uni-freiburg.de

https://doi.org/10.1239/jap/1363784423 Published online by Cambridge University Press

case, $F_1 = \cdots = F_d = F$, based on the following duality result (see Theorem 5 of Rüschendorf (1981) and Gaffke and Rüschendorf (1981)). 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, $F_i = F$, $1 \leq i \leq d$, the following statements hold.*

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

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

where

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

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

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

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

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

Embrechts and Puccetti (2006b) introduced the following class of piecewise-linear functions defined, for $t < s/d$, as

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

They established that the g_t are admissible, that is, $g_t \in \mathcal{D}(s)$, and defined the *dual bound* $D(s)$ as

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

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

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

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 Puccetti and Rüschendorf (2012b). In the recent works Wang and Wang (2011) and Wang *et al.* (2013), based on the concept of *complete mixability*,

https://doi.org/10.1239/jap/1363784423 Published online by Cambridge University Press

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 in (1.6) under these conditions. It implies in particular the results of Wang and Wang (2011) and Wang *et al.* (2013) in the case of monotone densities and gives a strongly simplified proof. It also implies sharp bounds for further cases, such as the case of concave densities and for distributions which are typically used in quantitative risk management. In addition to the results stated in the abovementioned 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 2.1 and Proposition 2.2 below) are based on complete mixability of the optimal dual function g^*, or, more precisely, of $g^*(X_i)$. Therefore, we start with a summary of results on completely mixable distributions which we use in the remainder of the paper.

1.1. Some preliminaries on complete mixability

Complete mixability is a concept of negative dependence of random vectors. This concept is a substitute of the notion of countermonotonic dependence in dimensions $d \geq 3$.

Definition 1.1. A distribution function F on $\mathbb{R}$ is called *d-completely mixable* (d-CM) if there exist d random variables $X_1, \ldots, X_d$ identically distributed as F such that

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

for some $\mu \in \mathbb{R}$. Any such μ is called a center of F and any vector $(X_1, \ldots, X_d)$ satisfying (1.7) with $X_i \sim F$, $1 \leq i \leq d$, 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 1.2. If X has distribution F, we say that F is d-CM on the interval $A \subset \mathbb{R}$ if the conditional distribution of $(X \mid X \in A)$ is d-CM.

The following results on complete mixability can be found in Wang and Wang (2011) and the references therein.

Theorem 1.2. *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 that 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*

$$a + \frac{b-a}{d} \leq \mu \leq b - \frac{b-a}{d}. \tag{1.8}$$

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

https://doi.org/10.1239/jap/1363784423 Published online by Cambridge University Press

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

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

is sharp, that is, $D(s) = M(s)$. The proof uses the following property of optimal couplings (see Proposition 3(c) of Rüschendorf (1982)).

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 $M(s) = \mathbb{P}(\sum_{i=1}^d X_i^* \geq s)$ and, for any such $\boldsymbol{X}^*$, we have*

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

In the case in which F is continuous, we obtain

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

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

$$dg(x) \geq \mathbf{1}\left(x \geq \frac{s}{d}\right) \geq 0.$$

Then it follows by combining point 2 of Theorem 1.1 with Theorem 2.1 that, if $g^* \in \mathfrak{D}(s)$ is an optimal choice for (1.3),

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

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

$$\mathbb{P}(g^*(X_i^*) = 0 \mid X_i^* < a^*) = 1, \qquad 1 \leq i \leq d. \tag{2.3}$$

As a consequence, any optimal dual choice g^* is a.s. 0 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 in 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 $M(s) = \mathbb{P}(\sum_{i=1}^d X_i^* \geq s)$ if and only if there exists an admissible function $g^* \in \mathfrak{D}(s)$ such that the conditional distribution of*

$$(g^*(X_1) \mid X_1 \geq a^*)$$

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

https://doi.org/10.1239/jap/1363784423 Published online by Cambridge University Press

Proof. Assume that $\boldsymbol{X}^*$ and g^* are respectively an optimal coupling and an optimal dual function for (1.3). By (2.3) we can assume that any $g^* \in \mathcal{D}(s)$ is 0 below the threshold a^*. Using (1.4) and (2.1) similarly as in (2.2), we find that $g^* \in \mathcal{D}$ is an optimal choice for (1.3) if and only if

$$\mathbb{P}\Big(\sum_{i=1}^{d} g^*(X_i^*) = 1 \,\Big|\, \sum_{i=1}^{d} X_i^* \geq s\Big) = \mathbb{P}\Big(\sum_{i=1}^{d} g^*(X_i^*) = 1 \,\Big|\, X_i^* \geq a^*\Big) = 1$$

for $1 \leq i \leq d$. This completes the proof.

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 2.1 below we state 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 2.2 below we show that $a^* = M^{-1}(s)$, hence obtaining the optimality of g_a.

Proposition 2.1. *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*

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

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

(A2) *X_1 is d-CM on (a, b).*

Then

1. *the conditional distribution H of $(g_a(X_1) \mid X_1 \geq a^*)$ is d-CM with center $\mu = 1/d$,*
2. *we have*
$$M(s) \leq \bar{F}(a^*) = \bar{F}(a) + (d-1)\bar{F}(b). \tag{2.4}$$

Proof. 1. First-order conditions on the argument of the infimum in (A1) at $t = a$ imply that

$$\frac{d\int_a^b \bar{F}(x)\,\mathrm{d}x}{b-a} = \bar{F}(a) + (d-1)\bar{F}(b). \tag{2.5}$$

Therefore, $a^* \leq a$ satisfies

$$\bar{F}(a^*) = D(s) = \frac{d\int_a^b \bar{F}(x)\,\mathrm{d}x}{b-a} = \bar{F}(a) + (d-1)\bar{F}(b). \tag{2.6}$$

Let $Y_{a^*} \overset{\mathrm{D}}{=} (X_1 \mid X_1 \geq a^*)$. We have to show that the distribution H of $g_a(Y_{a^*})$ is d-CM. From the definition of the linear functions g_t, $t < s/d$, given in (1.5), 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 $(g_a(Y_{a^*}) \mid Y_{a^*} \in (a, b))$, and we define the distribution G_2 as

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

https://doi.org/10.1239/jap/1363784423 Published online by Cambridge University Press

we can write H as

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

where $p_1 = \mathbb{P}(Y_{a^*} \leq a)$, $p_2 = \mathbb{P}(a < Y_{a^*} \leq b)$, and $p_3 = 1 - p_2 - p_1 = \mathbb{P}(Y_{a^*} > b)$. Note that G_1 is the distribution of a linear transformation of the random variable Y_{a^*} on the interval (a, b). From the mixing condition (A2), the distribution of Y_{a^*} is completely mixable on the interval (a, b). Using point 5 of Theorem 1.2, it follows that G_1 is d-CM with center given by

$$\int x \,\mathrm{d}G_1(x) = \int_a^b \frac{(x-a)\,\mathrm{d}F(x)}{(b-a)(\bar{F}(a) - \bar{F}(b))}\,\mathrm{d}x = \frac{\int_a^b \bar{F}(x)\,\mathrm{d}x/(b-a) - \bar{F}(b)}{\bar{F}(a) - \bar{F}(b)} = \frac{1}{d}. \quad (2.7)$$

Similarly, the mean of G_2 is given by

$$\int x\,\mathrm{d}G_2(x) = \frac{p_3}{p_1 + p_3} = \frac{\bar{F}(b)}{\bar{F}(b) + \bar{F}(a^*) - \bar{F}(a)} = \frac{1}{d}. \quad (2.8)$$

In (2.7) and (2.8), the right-hand equalities follow from (2.6). Note that the distribution G_2 is a binomial $B(1, 1/d)$. By point 3 of Theorem 1.2, G_2 is also 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 of Theorem 1.2, H is d-CM with center $1/d$.

2. Inequality (2.4) is a direct consequence of the fact that $g_a \in \mathcal{D}(s)$. Thus,

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

where the right-hand equality follows from (2.5).

Postulating the optimality of 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

$$\mathbb{P}\bigg(\bigcap_{i \neq j} \{Y_i \leq a\} \;\bigg|\; Y_j > b\bigg) = 1.$$

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

Proposition 2.2. *Under the assumptions of Proposition 2.1, suppose that*

(A3) $(d-1)(F(y) - F(b)) \leq F(a) - F\left(\dfrac{s-y}{d-1}\right)$

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

$$\mathbb{P}\bigg(\sum_{i=1}^d X_i^* \geq s\bigg) = \bar{F}(a^*).$$

Proof. For a^* satisfying (A1), denote by $F_{a^*}(x) = (F(x) - F(a^*))/\bar{F}(a^*)$ the distribution of the random variable $Y_{a^*} \overset{\mathrm{D}}{=} (X_1 \mid X_1 \geq a^*)$. We show that there exists a random vector

https://doi.org/10.1239/jap/1363784423 Published online by Cambridge University Press

$\boldsymbol{Y} = (Y_1, \dots, Y_d)$ with distribution $F_{\boldsymbol{Y}} \in \mathfrak{F}(F_{a^*}, \dots, F_{a^*})$ for which

$$\mathbb{P}\bigg(\sum_{i=1}^{d} Y_i \geq s\bigg) = 1. \tag{2.9}$$

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

$$\boldsymbol{X}^* = (X_1, \dots, X_1)\mathbf{1}\bigg(\bigcup_{i=1}^{d}\{X_i \leq a^*\}\bigg) + \boldsymbol{Y}\mathbf{1}\bigg(\bigcap_{i=1}^{d}\{X_i > a^*\}\bigg).$$

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_is lies in the interval (a, b) then all the Y_is lie in (a, b) and

$$\mathbb{P}(Y_1 + \cdots + Y_d = s \mid Y_i \in (a, b)) = 1.$$

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

$$\mathbb{P}(Y_j = F_{a^*}^{-1}((d-1)\bar{F}_{a^*}(Y_i)) \mid Y_i \geq b) = 1 \quad \text{for all } j \neq i.$$

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 (2.7), it is easy to see that the conditional distribution of $(Y_{a^*} \mid Y_{a^*} \in (a, b))$ has mean

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

Therefore, there exists a vector $\boldsymbol{Y}$ having marginals F_{a^*} and satisfying property (a). From (2.6), it follows that $F_{a^*}^{-1}((d-1)\bar{F}_{a^*}(b)) = a$. From property (b), we obtain

$$\mathbb{P}(Y_j \leq a \mid Y_i \geq b) = 1 \quad \text{for all } j \neq i.$$

Consequently, 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_is are identically distributed as F_{a^*}.

Note that $\sum_{i=1}^{d} Y_i = s$ a.s. when all the Y_is lie in the interval (a, b). Thus, in order to prove (2.9), it remains to show that $\sum_{i=1}^{d} Y_i \geq s$ when one of the Y_is is larger than b. To this end, we define the function $\psi\colon [b, +\infty) \to \mathbb{R}$ as

$$\psi(y) = y + (d-1)F_{a^*}^{-1}((d-1)\bar{F}_{a^*}(y)). \tag{2.10}$$

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

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

Expressing the above equation in terms of F, and using (2.4), reveals that $\psi(y) \geq s, \ y \geq b$, is equivalent to the ordering condition (A3). This completes the proof.

As a corollary of Proposition 2.1 and Proposition 2.2, we now state the main result of our paper.

https://doi.org/10.1239/jap/1363784423 Published online by Cambridge University Press

Theorem 2.3. (Sharpness of dual bounds.) *Under the attainment condition (A1), the mixing condition (A2), and the ordering condition (A3), the dual bounds are sharp, that is,*

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

Proof. From Proposition 2.1 and Proposition 2.2, $M(s) = \bar{F}(a^*)$ and the conditional distribution of $(g_a(X_1) \mid X_1 > a^*)$ 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,

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

Remark 2.1. *Monotone densities.* All continuous distribution functions F having a positive and decreasing density f on the unbounded interval $(a^*, +\infty)$ satisfy conditions (A2) and (A3). In this case, the conditional distribution of $(Y_{a^*} \mid Y_{a^*} \in (a,b))$ inherits a decreasing density from F and has mean $\mu = s/d$. By point 5 of Theorem 1.2, 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 (2.10) turns out to be convex and

$$\psi(b) = b+(d-1)F_{a^*}^{-1}((d-1)\bar{F}_{a^*}(b)) = b+(d-1)a = s-(d-1)a+(d-1)a = s. \quad (2.11)$$

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

$$f(a) - (d-1)^2 f(b) \geq 0, \quad (2.12)$$

that is, $\psi'_+(1-c) \geq 0$. Convexity of ψ and (2.11) finally imply that $\psi(x) \geq s$ for all $x \geq b$. Consequently, Theorem 2.3 implies as a particular case the results in Wang and Wang (2011) and Wang *et al.* (2013) for the case of monotone decreasing densities. The couplings used in the proof of Proposition 2.2 are of a similar form as in Wang *et al.* (2013) 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, (2.5) gives us a useful clue to the calculation of the sharp bound.

Monotonicity in the tail. As a consequence of the remark above, sharpness of the dual bound $D(s)$ can be stated, for large enough s, for all continuous, unbounded distribution functions which have an 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 assumptions (A1)–(A3) for all $s \in \mathbb{R}$ at which $D(s) < 1$. As a consequence, the bounds in Section 5.2 of Embrechts and Puccetti (2006b) are sharp. We will give some numerical examples regarding the Pareto and other types of distributions in Section 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 of Puccetti *et al.* (2012). In order to obtain sharpness of the dual bound $D(s)$ for these

https://doi.org/10.1239/jap/1363784423 Published online by Cambridge University Press

distributions, conditions (A1) and (A3) have to be checked analytically or numerically. In Proposition 2.3 below, we give an equivalent formulation of (A3) in terms of the stochastic order.

The $d = 2$ case. In this case 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 of Embrechts and Puccetti (2006b)) 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.

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

$$F_1(y) = \frac{\bar{F}((s-y)/(d-1)) - \bar{F}(a)}{\bar{F}(a^*) - \bar{F}(a)} \quad \text{and} \quad F_2(y) = \frac{\bar{F}(b) - \bar{F}(y)}{\bar{F}(b)} \quad \text{for } y \geq b.$$

Proposition 2.3. *Under the attainment condition (A1), the ordering condition (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 (written $F_1 \leq_{\mathrm{st}} F_2$).*

Proof. Using (2.6), the proposition immediately follows by noting that $F_2(y) \leq F_1(y)$, $y \geq b$, is equivalent to

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

which is equivalent to (A3).

Remark 2.2. We make the following remarks about Proposition 2.3.

1. By a well-known condition, the stochastic ordering $F_1 \leq_{\mathrm{st}} F_2$ is implied by the monotone likelihood ratio criterion for their densities stating that f_2/f_1 is increasing. In our case, the ordering condition (A3) is satisfied if F has a density f for which

$$\frac{f(y)}{f((s-y)/(d-1))} \text{ is increasing in } y \text{ for } y \geq b. \tag{2.13}$$

   This condition can be checked in examples.

2. For distributions with monotone densities, condition (A3) holds. In several examples of nonmonotone densities we found that condition (A3) or condition (2.13) is satisfied; see Section 3.

3. Applications and numerical verifications

Equation (2.5) provides a clue for how to calculate the basic point a and, hence, the dual bound $D(s)$. Having calculated a, one can easily check the second-order condition (2.12) 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 a different set 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 the comment on monotone densities in Remark 2.1.

https://doi.org/10.1239/jap/1363784423 Published online by Cambridge University Press

- If F has a concave density f on (a, b) then the mixing condition (A2) is satisfied (see the comment on concave densities in Remark 2.1) 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 of Remark 2.2.

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 the 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, as well as providing 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 Puccetti and Rüschendorf (2012b). In Figures 2–3, we respectively plot the dual bound $D(s)$ in (1.6) for a random vector $\boldsymbol{X}$ of $d = 3$ lognormal(2,1)- and 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.

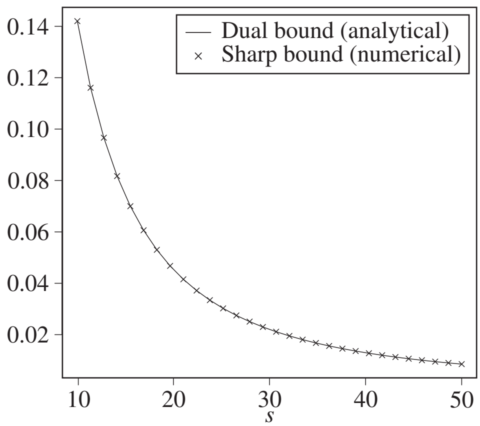

FIGURE 1: Dual bounds $D(s)$ (see (1.6)) for the sum of $d = 3$ Pareto(2)-distributed risks. Numerical values for the sharp bounds $M(s)$ are also provided at some thresholds s of interest.

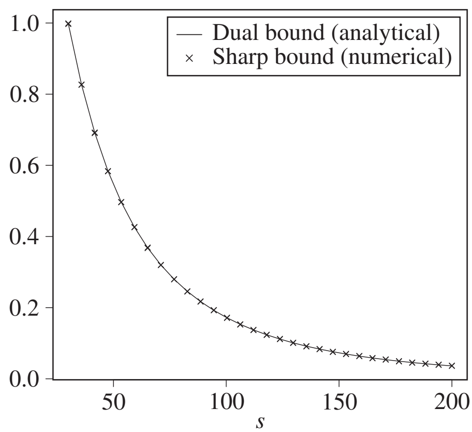

FIGURE 2: Dual bounds $D(s)$ (see (1.6)) for the sum of $d = 3$ lognormal(2, 1)-distributed risks. Numerical values for the sharp bounds $M(s)$ are also provided at some thresholds s of interest.

https://doi.org/10.1239/jap/1363784423 Published online by Cambridge University Press

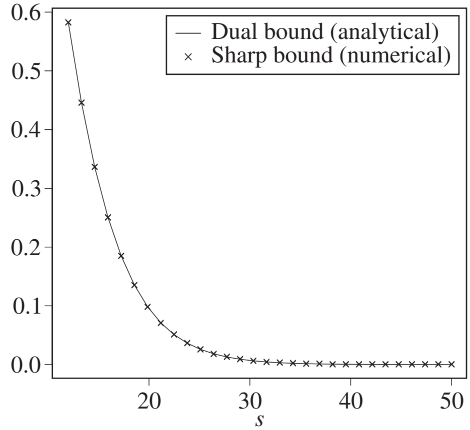

FIGURE 3: Dual bounds $D(s)$ (see (1.6)) for the sum of $d = 3$ gamma(3, 1)-distributed risks. Numerical values for the sharp bounds $M(s)$ are also provided at some thresholds s of interest.

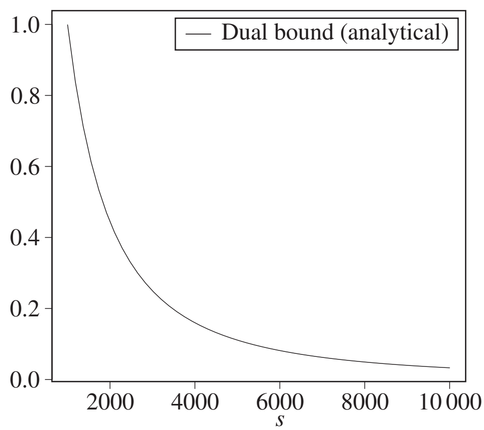

FIGURE 4: Dual bounds $D(s)$ (see (1.6)) for the sum of $d = 1000$ Pareto(2)-distributed risks.

In high dimensions, $d > 100$, the rearrangement algorithm introduced in Puccetti and Rüschendorf (2012b) cannot be practically used to compute $M(s)$. In the homogeneous case, the dual bound methodology can however lead to sharp bounds $M(s)$ for arbitrary large 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

EMBRECHTS, P. AND PUCCETTI, G. (2006a). Aggregating risk capital, with an application to operational risk. *Geneva Risk Insurance Rev.* **31,** 71–90.

EMBRECHTS, P. AND PUCCETTI, G. (2006b). Bounds for functions of dependent risks. *Finance Stoch.* **10,** 341–352.

GAFFKE, N. AND RÜSCHENDORF, L. (1981). On a class of extremal problems in statistics. *Math. Operationsforsch. Statist. Ser. Optimization* **12,** 123–135.

https://doi.org/10.1239/jap/1363784423 Published online by Cambridge University Press

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

Puccetti, G. and Rüschendorf, L. (2012a). Bounds for joint portfolios of dependent risks. *Statist. Risk Modeling* **29,** 107–132.

Puccetti, G. and Rüschendorf, L. (2012b). Computation of sharp bounds on the distribution of a function of dependent risks. *J. Comput. Appl. Math.* **236,** 1833–1840.

Puccetti, G., Wang, B. and Wang, R. (2012). Advances in complete mixability. *J. Appl. Prob.* **49,** 430–440.

Rüschendorf, L. (1981). Sharpness of Fréchet-bounds. *Z. Wahrscheinlichkeitsth.* **57,** 293–302.

Rüschendorf, L. (1982). Random variables with maximum sums. *Adv. Appl. Prob.* **14,** 623–632.

Wang, B. and Wang, R. (2011). The complete mixability and convex minimization problems with monotone marginal densities. *J. Multivariate Anal.* **102,** 1344–1360.

Wang, R., Peng, L. and Yang, J. (2013). Bounds for the sum of dependent risks and worst Value-at-Risk with monotone marginal densities. To appear in *Finance Stoch.*

https://doi.org/10.1239/jap/1363784423 Published online by Cambridge University Press