Model uncertainty and VaR aggregation

Paul Embrechts[a], Giovanni Puccetti[b,*], Ludger Rüschendorf[c]

[a] *Senior SFI Professor, RiskLab, Department of Mathematics, ETH Zurich, 8092 Zurich, Switzerland*
[b] *Department of Mathematics for Decision Theory, University of Firenze, 50127 Firenze, Italy*
[c] *Department of Mathematical Stochastics, University of Freiburg, 79104 Freiburg, Germany*

Abstract

Despite well-known shortcomings as a risk measure, Value-at-Risk (VaR) is still the industry and regulatory standard for the calculation of risk capital in banking and insurance. This paper is concerned with the numerical estimation of the VaR for a portfolio position as a function of different dependence scenarios on the factors of the portfolio. Besides summarizing the most relevant analytical bounds, including a discussion of their sharpness, we introduce a numerical algorithm which allows for the computation of reliable (sharp) bounds for the VaR of high-dimensional portfolios with dimensions d possibly in the several hundreds. We show that additional positive dependence information will typically not improve the upper bound substantially. In contrast higher order marginal information on the model, when available, may lead to strongly improved bounds. Several examples of practical relevance show how explicit VaR bounds can be obtained. These bounds can be interpreted as a measure of model uncertainty induced by possible dependence scenarios.

Keywords: Copula, Fréchet class, Model Uncertainty, Operational Risk, Positive Dependence, Rearrangement Algorithm, Risk Aggregation, Value-at-Risk, VaR-bounds.
2000 MSC: 62P05 (primary), 91G60, 91B30 (secondary)

1. Introduction

Since the early nineties, Value-at-Risk (VaR) has established itself as *a* (if not *the*) key metric for the calculation of regulatory capital within the financial industry. Furthermore, VaR is increasingly used as a risk management constraint within portfolio optimization. Whereas books like Jorion (2006) prize VaR as the industry standard, numerous papers have pointed out many of the (most obvious) shortcomings of VaR as a risk measure; see for instance McNeil et al. (2005) and the references therein, but also the recent Basel Committee on Banking Supervision (2012), already referred to as Basel 3.5. A very informative overview on the use of VaR technology within the banking industry is Pérignon and Smith (2010). As so often, a middle-of-the-road point of view is advisable: there is no doubt that the construction and understanding of the P&L distribution of a bank's trading book is of the utmost importance. The latter includes the availability of data warehouses, independent pricing tools and a complete risk factor mapping. And of course Corporate Governance decisions may have a major impact on the P&L, like for instance in the case of strategic decisions. In that sense, VaR, as a number, is just the peak of the risk management iceberg. Nonetheless, once the number leaves the IT system of the CRO, all too often it starts a life of its own and one often forgets the numerous warnings about its proper interpretation. Moreover, once several VaRs are involved, the temptation is there to calculate functions of them (like adding) forgetting the considerable model uncertainty underlying such constructions; see Basel Committee on Banking Supervision (2010) for a regulatory overview on risk aggregation. A typical such example is to be found in the realm of Operational Risk as defined under Basel II and III. Throughout the paper we will use the latter as a motivating example and consider the organization of an Operational Risk database in business lines and risk types; for a background to this and for further references, see for instance McNeil et al. (2005, Chapter 10). We want to stress however that the quantitative modeling of Operational Risk is just a motivating example where the techniques discussed in our paper can be applied naturally. The results obtained are applicable much more widely and related questions do occur frequently in banking and insurance.

To set the scene, consider the calculation of the VaR at a confidence level α for an aggregate loss random variable L^+ having the form

$$L^+ = \sum_{i=1}^{d} L_i,$$

where $L_1, \dots, L_d$, in the case of Operational Risk, correspond to the loss random variables for given business lines or risk types, over a fixed time period T. The VaR of the aggregate position L^+, calculated at a probability level $\alpha \in (0,1)$, is the α-quantile of its distribution, defined as

$$\mathrm{VaR}_\alpha(L^+) = F^{-1}_{L^+}(\alpha) = \inf\{x \in \mathbb{R} : F_{L^+}(x) > \alpha\}, \tag{1}$$

where $F_{L^+}(x) = P(L^+ \leq x)$ is the distribution function of L^+. As a statistical quantity and for α typically close to 1, $\mathrm{VaR}_\alpha(L^+)$ is

*Corresponding author
Email addresses: embrechts@math.ethz.ch (Paul Embrechts), giovanni.puccetti@unifi.it (Giovanni Puccetti), ruschen@stochastik.uni-freiburg.de (Ludger Rüschendorf)

a measure of extreme loss, i.e. $P(L^+ > \mathrm{VaR}_\alpha(L^+)) \leq 1 - \alpha$ is typically small.

The current regulatory framework for banking supervision, referred to as Basel II (becoming Basel III), allows large international banks to come up with internal models for the calculation of risk capital. For Operational Risk, under the so-called Loss Distribution Approach (LDA) within Basel II, financial institutions are given full freedom concerning the stochastic modeling assumptions used. The resulting risk capital must correspond to a 99.9%-quantile of the aggregated loss data over the period of a year; we leave out the specific details concerning internal, external and expert opinion data as they are less relevant for the results presented in this paper. Using the notation introduced above, the risk capital for the aggregate position L^+ is typically based on $\mathrm{VaR}_{0.999}(L^+)$. Concerning interdependence of risks, no specific rules are given beyond the statement that explicit and implicit correlation assumptions between loss random variables used have to be plausible and need to be well founded; in the case of Operational Risk, see Cope and Antonini (2008) and Cope et al. (2009). For the sequel of this paper, we leave out statistical (parameter) uncertainty.

In order to calculate $\mathrm{VaR}_\alpha(L^+)$, one needs a joint model for the random vector $(L_1, \ldots, L_d)'$. This would require an extensive d-variate dataset for the past occurred losses, which often is not available. Typically, only the marginal distribution functions F_i of L_i are known or statistically estimated, while the dependence structure between the L_i's is either completely or partially unknown. This situation also often occurs in the analysis of credit risk data; here the d could be viewed as the number of individual obligors, industry or geographic sectors, say.

In standard practice, the total capital charge C to be allocated is derived from the addition of the VaRs at probability level $\alpha = 0.999$ for the marginal random losses L_i, namely

$$\mathrm{VaR}_\alpha^+(L^+) = \sum_{i=1}^d \mathrm{VaR}_\alpha(L_i) = \sum_{i=1}^d F_i^{-1}(\alpha).$$

Indeed, industry typically reports

$$C = \delta \mathrm{VaR}_\alpha^+(L^+), \quad 0 < \delta \leq 1; \tag{2}$$

the value of δ is often in the range $(0.7, 0.9)$ and reflects so-called diversification effects. A capital charge based on (2) would imply a subadditive regime for VaR, i.e.

$$\mathrm{VaR}_\alpha(L^+) = \mathrm{VaR}_\alpha\left(\sum_{i=1}^d L_i\right) \leq \sum_{i=1}^d \mathrm{VaR}_\alpha(L_i) = \mathrm{VaR}_\alpha^+(L^+). \tag{3}$$

The case $\delta = 1$ (no diversification) in (2) can be mathematically justified by the assumption of perfect positive dependence (which implies maximal correlation) among marginal risks. Indeed, under this so-called *comonotonic dependence* scenario, $\mathrm{VaR}_\alpha(L^+) = \mathrm{VaR}_\alpha^+(L^+)$; see McNeil et al. (2005, Proposition 6.15). Practitioners criticize this assumption as not being realistic, and remark that random losses are not perfectly correlated in view of their heterogeneous nature. Though the $\delta = 1$ maximal-correlation scenario is often considered as highly conservative, the inequality in (3) is typically violated for either very heavy-tailed losses, very skewed losses, or losses exhibiting special dependencies. Such situations are no doubt present in Operational Risk data; see for instance Moscadelli (2004), Panjer (2006), Shevchenko (2011) and Bolancé et al. (2012). The three standard classes of examples violating (3) and mentioned above are to be found in McNeil et al. (2005, Examples 6.7, 6.22).

Based on the above example from the capital charge calculation of Operational Risk it is clear that there exists considerable model uncertainty underlying the *diversification* factor δ, which for practically relevant models could well take values above the additive case $\delta = 1$. It is exactly this kind of model uncertainty that the present paper addresses. In the discussion below, we will now abstract from the motivating Operational Risk example.

Recently, a number of numerical and analytical techniques have been developed in order to calculate *conservative* values for $\mathrm{VaR}_\alpha(L^+)$ under different dependence assumptions regarding the loss random variables L_i. In this paper we describe these methodologies and give insight in the worst-case dependence structure (copula) describing the worst-VaR scenario.

We summarize the main contributions of this paper:

- we introduce an algorithm which allows to calculate sharp bounds for the VaR of possibly high-dimensional portfolio positions allowing for inhomogeneous portfolios with dimension d in the several hundreds;
- we show that additional positive dependence information added on top of the marginal distributions does not improve the VaR bounds substantially;
- we show that additional information on higher dimensional sub-vectors of marginals leads to possibly much narrower VaR bounds, and
- we give the dependence structures (copulas) leading to worst-case scenarios.

The main message coming from our paper is that currently a whole toolkit of analytical and numerical techniques is available to better understand the aggregation and diversification properties of non-coherent risk measures such as Value-at-Risk.

We very much hope that our paper is both accessible to the academic researcher as well as to the more quantitative practitioner. With this goal in mind, we have strived at keeping the technical details to a minimum, stressing more the algorithmic, numerical aspects of the results discussed. Of course, we will direct the reader interested in more mathematical details to the relevant research papers. We strongly believe that the results and techniques summarized are sufficiently novel and will benefit the wider financial industry.

With financial/actuarial applications in mind, and without loss of generality, in almost all the examples contained in the paper we use power law models for the marginal distributions of the risks such as the Pareto distribution. In particular, we often use a Pareto distribution with tail parameter $\theta = 2$ in order to represent marginal risks with finite mean but infinite variance.

This choice is pedagogical and does not affect the computational properties of the methodologies discussed.

In Section 2, we study the case where the marginal distribution functions F_i of L_i are fixed while the dependence structure (copula) between the L_i's is completely unknown. In the *homogeneous* case where the risk factors L_i are identically distributed, a simple analytical formula allows to compute the worst-possible VaR for portfolios of arbitrary dimensions when the marginal distributions F_i are continuous. For *inhomogeneous* portfolios having arbitrary marginals, a new numerical algorithm, see Section 2.2, allows to compute best- and worst-possible VaR values in arbitrary dimensions; the main limiting factors are computer memory and numerical accuracy to be obtained. We test the algorithm in an example with $d = 648$.

Under the restriction of the dependence structure to positive dependence, possible improvements of the bounds are discussed in Section 3. Finally, in Section 4, we consider a more general case where extra information is known about sub-vectors of the marginal risks. In the Sub-sections 1.1–1.4 below we first gather some definitions, notation and basic methodological tools, together with some key references.

1.1. Fréchet classes

Denote $\boldsymbol{L} = (L_1, \ldots . L_d)'$. The Value-at-Risk for the aggregate position $L^+ = L_1 + \cdots + L_d$ is certainly not uniquely determined by the marginal distributions $F_1, \ldots, F_d$ of the risks L_i. In fact, there exist infinitely many joint distributions on $\mathbb{R}^d$ which are consistent with the choice of the marginals $F_1, \ldots, F_d$. We denote by $\mathfrak{F}(F_1, \ldots, F_d)$ the *Fréchet class* of all the possible joint distributions $F_{\boldsymbol{L}}$ on $\mathbb{R}^d$ having the given marginals $F_1, \ldots, F_d$. For $\alpha \in (0, 1)$, upper and lower bounds for the Value-at-Risk of L^+ are then defined as

$$\overline{\mathrm{VaR}}_\alpha(L^+) = \sup\left\{\mathrm{VaR}_\alpha(L_1 + \cdots + L_d) : F_{\boldsymbol{L}} \in \mathfrak{F}(F_1, \ldots, F_d)\right\}, \tag{4a}$$

$$\underline{\mathrm{VaR}}_\alpha(L^+) = \inf\left\{\mathrm{VaR}_\alpha(L_1 + \cdots + L_d) : F_{\boldsymbol{L}} \in \mathfrak{F}(F_1, \ldots, F_d)\right\}. \tag{4b}$$

The above definitions directly imply the VaR range for L^+ given by

$$\underline{\mathrm{VaR}}_\alpha(L^+) \le \mathrm{VaR}_\alpha(L_1 + \cdots + L_d) \le \overline{\mathrm{VaR}}_\alpha(L^+). \tag{5}$$

We refer to the bounds $\overline{\mathrm{VaR}}_\alpha(L^+)$ and $\underline{\mathrm{VaR}}_\alpha(L^+)$ as the worst-possible and, respectively, the best-possible VaR for the position L^+, at the probability level α. When attained, the upper and lower bounds in (4) are sharp (best-possible): they *cannot be improved* if further dependence information on $(L_1, \ldots, L_d)'$ is not available. We call any joint model for $(L_1^*, \ldots, L_d^*)'$ with prescribed marginals $F_1, \ldots, F_d$ such that

$$\overline{\mathrm{VaR}}_\alpha(L^+) = \mathrm{VaR}_\alpha(L_1^* + \cdots + L_d^*)$$

a *worst-case dependence* or *worst-case coupling*. Analogously, any joint model for $(L_1^*, \ldots, L_d^*)'$ with the prescribed marginals such that

$$\underline{\mathrm{VaR}}_\alpha(L^+) = \mathrm{VaR}_\alpha(L_1^* + \cdots + L_d^*)$$

is a *best-case dependence* or *best-case coupling*. Of course, the choice of wording *best* versus *worst* is arbitrarily and depends on the specific application at hand. Problems related to (4) with moment information have always been relevant in actuarial mathematics. One of the early contributors was De Vylder (1996); see also Hürlimann (2008a,b) for numerous examples from the realm of insurance.

1.2. Copulas

To make this paper self-contained, we give a brief introduction to some copula concepts that we will need in the following. The reader not familiar with the theory of copulas is referred to Nelsen (2006), McNeil et al. (2005, Chapter 5) and Durante and Sempi (2010).

A copula C is a d-dimensional distribution function (df) on $[0, 1]^d$ with uniform marginals. Given a copula C and d univariate marginals $F_1, \ldots, F_d$, one can always define a df F on $\mathbb{R}^d$ having these marginals by

$$F(x_1, \ldots, x_d) = C(F_1(x_1), \ldots, F_d(x_d)),\ x_1, \ldots, x_d \in \mathbb{R}. \tag{6}$$

Sklar's Theorem states conversely that we can always find a copula C coupling the marginals F_i of a fixed joint distribution F through the above expression (6). For continuous marginal dfs, this copula is unique. Hence Sklar's Theorem states that the copula C of a multivariate distribution F contains all the dependence information of F.

A first example of a copula is the *independence* copula

$$\Pi(u_1, \ldots, u_d) = \Pi_{i=1}^d u_i.$$

The name of this copula derives from the fact that the risk vector $(L_1, \ldots, L_d)'$ has copula Π if and only if its marginal risks L_i are independent. Under independence among the marginal risks, (6) reads as

$$F(x_1, \ldots, x_d) = \Pi(F_1(x_1), \ldots, F_d(x_d)) = F_1(x_1) \cdot \ldots \cdot F_d(x_d).$$

Any copula C satisfies the so-called Fréchet bounds

$$\max\left\{\sum_{i=1}^d u_i - d + 1, 0\right\} \le C(u_1, \ldots, u_d) \le \min\{u_1, \ldots, u_d\},$$

for all $u_1, \ldots, u_d \in [0, 1]$. The sharp upper Fréchet bound

$$M(u_1, \ldots, u_d) = \min\{u_1, \ldots, u_d\}$$

is the so-called *comonotonic* copula, which represents perfect positive dependence among the risks. In fact, a risk vector $(L_1, \ldots, L_d)'$ has copula M if and only if its marginal risks are all almost surely (a.s.) increasing functions of a common random factor. For a detailed discussion of the concept of comonotonicity within quantitative risk management we refer to Dhaene et al. (2002) and Dhaene et al. (2006); see also McNeil et al. (2005, Section 6.2.2). The lower Fréchet bound

$$W(u_1, \ldots, u_d) = [u_1 + \cdots + u_d - d + 1]^+$$

is also sharp. However, it is a well-defined copula only in dimension $d = 2$. In this case, it is called the *countermonotonic* copula and represents perfect negative dependence between two risks. A risk vector $(L_1, L_2)'$ has copula W if and only if its marginal risks are a.s. decreasing functions of each other.

The upper and lower Fréchet bounds are important for finding optimal couplings in many optimization problems of interest in quantitative risk management. For instance it is well known that the maximal variance for the sum of risks with given marginals is attained when the risks are comonotonic, that is when they have copula $C = M$. Analogously, the minimal variance for the sum of two risks with given marginals is attained when they are countermonotonic, $C = W$. These results derive from the classical Hoeffding-Fréchet bounds and can be seen as particular cases of a more general ordering theorem; see Corollary 3 in Rüschendorf (1983).

1.3. Worst and best VaR

Sklar's Theorem allows us to reformulate (4) as optimization problems over $\mathfrak{C}_d$, the set of all d-dimensional copulas:

$$\overline{\mathrm{VaR}}_\alpha(L^+) = \sup\left\{\mathrm{VaR}_\alpha(L_1^C + \cdots + L_d^C) : C \in \mathfrak{C}_d\right\}, \tag{7a}$$
$$\underline{\mathrm{VaR}}_\alpha(L^+) = \inf\left\{\mathrm{VaR}_\alpha(L_1^C + \cdots + L_d^C) : C \in \mathfrak{C}_d\right\}. \tag{7b}$$

Here the vector $(L_1^C, \dots, L_d^C)'$ has the same marginal distributions as $(L_1, \dots, L_d)'$ and copula C. In general, it is difficult to evaluate the bounds in (4) or in (7) in explicit form, especially when one has to deal with $d \geq 3$ risks. This is related to the fact that in general Value-at-Risk is *non-subadditive*. As a consequence, the comonotonic copula M is in general *not* a solution to the problem $\overline{\mathrm{VaR}}_\alpha(L^+)$ in (7a). Equivalently, the worst-VaR value $\overline{\mathrm{VaR}}_\alpha(L^+)$ in (4) is *not* attained when all the risks are perfectly positively dependent. Analogously, the countermonotonic copula W is in general *not* a solution to the problem $\underline{\mathrm{VaR}}_\alpha(L^+)$ in (7b) for $d = 2$.

As already stated above, in the comonotonic case $C = M$, we have that

$$\mathrm{VaR}_\alpha^+(L^+) = \mathrm{VaR}_\alpha(L_1^M + \cdots + L_d^M) = \sum_{i=1}^d \mathrm{VaR}_\alpha(L_i) = \sum_{i=1}^d F_i^{-1}(\alpha). \tag{8}$$

It is not difficult to provide examples of interest in quantitative risk management where, for a copula C, necessarily $C \neq M$, we have that

$$\mathrm{VaR}_\alpha(L_1^C + \cdots + L_d^C) > \sum_{i=1}^d \mathrm{VaR}_\alpha(L_i).$$

For instance, if the random losses $L_1, \dots, L_d$ are identically distributed like a symmetric θ-stable distribution with $\theta < 1$, we have that

$$\begin{aligned}&\mathrm{VaR}_\alpha(L^{\Pi,+}) = \mathrm{VaR}_\alpha(L_1^\Pi + \cdots + L_d^\Pi)\\ &= d^{1/\theta}\,\mathrm{VaR}_\alpha(L_1) > d\,\mathrm{VaR}_\alpha(L_1) = \mathrm{VaR}_\alpha(L_1^M + \cdots + L_d^M);\end{aligned}$$

see Mainik and Rüschendorf (2010). Other examples in which independence implies a larger VaR estimate than comonotonic dependence can be found in Embrechts and Puccetti (2010b, Section 5.3), Mainik et al. (2013) and Section 2.3 below.

1.4. Complete mixability

When dealing with extremal values for Value-at-Risk, the ideas of perfect positive and negative dependence, as represented by the Fréchet bounds M and W, can be deceiving. Handling non-subadditive risk measures requires the knowledge of alternative dependence concepts; *complete mixability* turns out to be such a concept. It turns out to be highly useful towards the calculation of VaR bounds.

Definition 1. *A distribution function F on $\mathbb{R}$ is d-*completely mixable *(d-CM) if there exist d random variables $X_1, \dots, X_d$, identically distributed as F, such that*

$$P(X_1 + \cdots + X_d = c) = 1, \tag{9}$$

for some constant $c \in \mathbb{R}$. Any vector $(X_1, \dots, X_d)'$ satisfying (9) *with $X_i \sim F, 1 \leq i \leq d$, is called a d-*complete mix*. If F has finite first moment μ, then $c = \mu d$.*

Complete mixability is a concept of negative dependence. In dimension $d = 2$ complete mixability implies countermonotonicity. Indeed, a risk vector $(L_1, L_2)'$ is a 2-complete mix if and only if $L_1 = k - L_2$ a.s, and this implies that its copula is the *lower* Fréchet bound W (the converse however does not hold). In higher dimensions, $d \geq 3$, a completely mixable dependence structure minimizes the variance of the sum of risks with given marginal distributions. In fact, a risk vector $(L_1, \dots, L_d)'$ with identically distributed marginals is a d-complete mix if and only if the variance of the sum of its components is equal to zero. Not all univariate distributions F are d-CM. As an example, it is sufficient to take F as the two-point distribution giving probability mass $p > 0$ to $x = 0$ and $1 - p$ to $x = 1$. Since the only way to make $L_1 + L_2$ a constant is to choose $L_2 = 1 - L_1$, F is not 2-CM for $p \neq 1/2$.

The structure of dependence (copula) corresponding to complete mixability is not so intuitive and, at the moment, does not have an easy mathematical formulation like in the case of the Fréchet bounds. We illustrate this with a discrete example. We choose F to give mass 1/5 to any of the first five integers. A 3-complete mix of F can be represented by the following matrix, in which any row is to be seen as a vector in $\mathbb{R}^3$ having probability mass 1/5:

$$\begin{bmatrix} 1 & 5 & 3 \\ 2 & 3 & 4 \\ 3 & 1 & 5 \\ 4 & 4 & 1 \\ 5 & 2 & 2 \end{bmatrix}.$$

Since the sum of each row in the above matrix is equal to $k = 9$ (note that the mean of F is equal to 3), F turns out to be 3-completely mixable. It is useful to compare the above matrix with the one representing comonotonicity among three F-distributed risks:

$$\begin{bmatrix} 1 & 1 & 1 \\ 2 & 2 & 2 \\ 3 & 3 & 3 \\ 4 & 4 & 4 \\ 5 & 5 & 5 \end{bmatrix}.$$

In this latter case, the variance of the row-wise sums is maximized. Some other examples of completely mixable distributions, as well as an insight into the theory of complete mixability, are given in Rüschendorf and Uckelmann (2002), Wang and Wang (2011) and Puccetti et al. (2012). Interesting cases where the concept of complete mixability plays an important role in the optimization problems (7) are the homogeneous case where the L_i's are identically distributed with a continuous distribution having an unbounded support and an ultimately decreasing density; see Puccetti and Rüschendorf (2013) .

2. Computing the VaR range with given marginal information

In this section, we consider the case when the risk vector $(L_1, \dots, L_d)'$ has given marginal distribution functions $F_1, \dots, F_d$ while its dependence structure is completely unknown. Recently, some new numerical and analytical tools have been developed to calculate the VaR range in (5) under these assumptions. First, we study the *homogeneous* case where the marginal risks are all identically distributed. Then, we will consider the more general *inhomogeneous* framework in which the marginal distributions are allowed to differ.

2.1. Identically distributed marginals

Throughout this section we assume that the marginal risks L_i are all identically distributed as F, that is $F_1 = \cdots = F_d = F$. In the case $d = 2$, the calculation of the sharp VaR bounds in (4) reduces to a simple formula if F satisfies some regularity conditions.

Proposition 2. *In the case $d = 2$ with $F_1 = F_2 = F$, let F be a continuous distribution concentrated on $[0, \infty)$ with an ultimately decreasing density on $(\overline{x}_F, \infty)$, for some $\overline{x}_F \geq 0$. Then*

$$\underline{\mathrm{VaR}}_\alpha(L^+) = F^{-1}(\alpha) \quad \text{and} \quad \overline{\mathrm{VaR}}_\alpha(L^+) = 2F^{-1}\left(\frac{1+\alpha}{2}\right), \tag{10}$$

for all $\alpha \in [F(\overline{x}_F), 1)$.

Remark 3. 1. If $\overline{x}_F = 0$, e.g. in the case F is Pareto distributed, that is

$$F(x) = 1 - (1+x)^{-\theta}, x > 0, \tag{11}$$

for some tail parameter $\theta > 0$, then the sharp bounds in (10) hold for any level of probability $\alpha \in (0, 1)$.

2. For $d = 2$, the sharp bounds $\overline{\mathrm{VaR}}_\alpha(L^+)$ and $\underline{\mathrm{VaR}}_\alpha(L^+)$ are known for *any* type of marginal distributions F_1,F_2. The slightly more complicated formulas to compute the bounds in the general case are given in Rüschendorf (1982, Proposition 1).

For a given α, a worst-case dependence vector (L_1^*, L_2^*) such that $\mathrm{VaR}_\alpha(L_1^* + L_2^*) = \overline{\mathrm{VaR}}_\alpha(L^+)$ is given by

$$\begin{cases} L_2^* = L_1^* & \text{a.s., when } L_1 < F^{-1}(\alpha), \\ L_2^* = F^{-1}\left(1 + \alpha - F(L_1^*)\right) & \text{a.s., when } L_1 \geq F^{-1}(\alpha). \end{cases}$$

In Figure 1, left, we show the copula of the risk vector $(L_1^*, L_2^*)'$. In the right part of the same figure, we show the support of the risk vector $(L_1^*, L_2^*)'$ when L_1^* and L_2^* are both Pareto(2)-distributed. The *support* of a random vector $\boldsymbol{X}$ is the smallest closed set A such that $P(\boldsymbol{X} \notin A) = 0$. It is interesting to note the interdependence of L_1^* and L_2^*. In the upper $(1-\alpha)$ part of their supports, the marginal risks L_1^* and L_2^* are countermonotonic. This means that the variance of the sum of the upper $(1-\alpha)$ parts of their supports is minimized. In the lower α-part of their supports, the marginal risks L_1^* and L_2^* are a.s. identical and hence comonotonic. This is however not relevant since the interdependence in this lower part of the joint distribution can be chosen arbitrarily; see Puccetti and Rüschendorf (2013, Theorem 2.1).

The case $d = 2$ is mainly pedagogical. The typical dimensions used in practice may vary from $d = 7$ or 8 to 56, say, for the aggregation of Operational Risk factors; see Moscadelli (2004), but may go up to d in the several hundreds or even thousands for hierarchical risk aggregation models; see for instance Arbenz et al. (2012). In the case $d > 2$, the sharp bound $\overline{\mathrm{VaR}}_\alpha(L^+)$ has been obtained only recently in the homogeneous case under different sets of assumptions. For a distribution function F, define the *dual bound* $D(s)$ as

$$D(s) = \inf_{t < s/d} \frac{d \int_t^{s-(d-1)t} \overline{F}(x)dx}{(s - dt)}, \tag{12}$$

where $\overline{F}(x) = 1 - F(x)$. The dual bound $D(s)$ in (12) is an upper bound on the tail function of L^+, that is

$$P(L_1 + \cdots + L_d > s) \leq D(s);$$

see for instance Puccetti and Rüschendorf (2013). This directly implies that

$$\overline{\mathrm{VaR}}_\alpha(L^+) \leq D^{-1}(1-\alpha) = \inf\{s \in \mathbb{R} : D(s) > 1 - \alpha\}. \tag{13}$$

The VaR bound $D^{-1}(1-\alpha)$ is numerically easy to evaluate independently of the size d of the portfolio $(L_1, \dots, L_d)'$. Under some extra assumptions, we have that the inequality in (13) becomes an equality.

Proposition 4 (Dual bound). *In the homogeneous case $F_i = F, 1 \leq i \leq d$, with $d \geq 3$, let F be a continuous distribution with an unbounded support and an ultimately decreasing density. Suppose that for any sufficiently large threshold s the infimum in (12) is attained at some $a < s/d$, that is assume that*

$$D(s) = \frac{d \int_a^b \overline{F}(x)dx}{(b-a)} = \overline{F}(a) + (d-1)\overline{F}(b), \tag{14}$$

where $b = s - (d-1)a$, with $F^{-1}(1 - D(s)) \leq a < s/d$. Then, for any sufficiently large threshold α we have that

$$\overline{\mathrm{VaR}}_\alpha(L^+) = D^{-1}(1-\alpha). \tag{15}$$

Remark 5. The above proposition is a particular case of Puccetti and Rüschendorf (2013, Theorem 2.5) and goes back to a conjecture made in Embrechts and Puccetti (2006b). We refer to the former paper and references therein for mathematical details in addition to the following points:

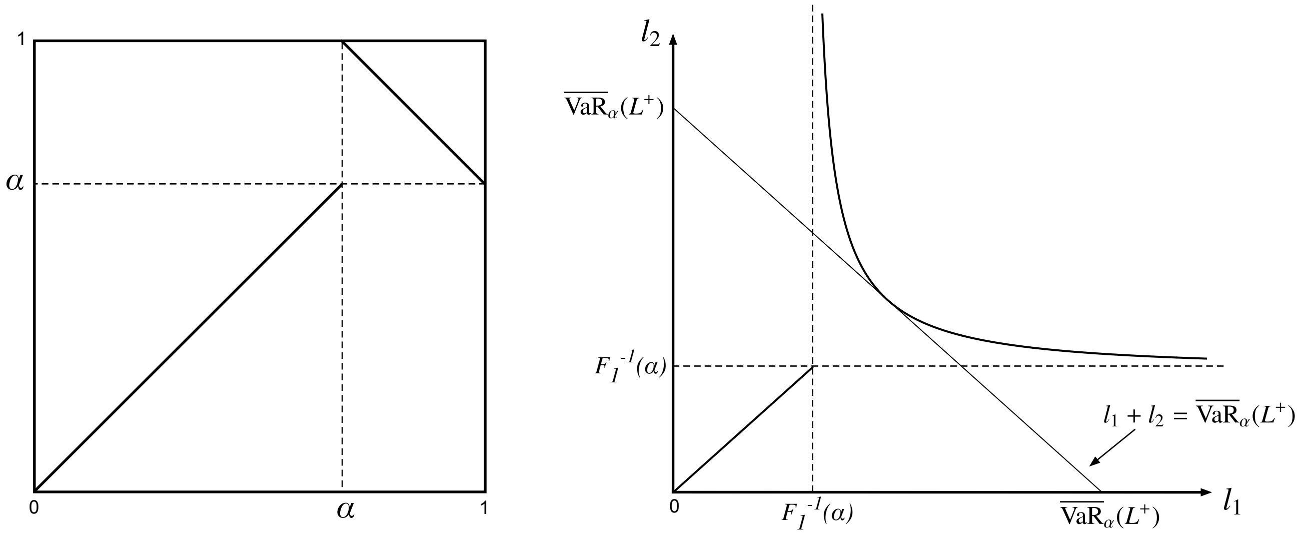

Figure 1: Bivariate copula (left) and support (right) of the vector $(L_1^*, L_2^*)'$ attaining the worst-possible VaR for $L_1 + L_2$ when L_1 and L_2 are both Pareto(2)-distributed.

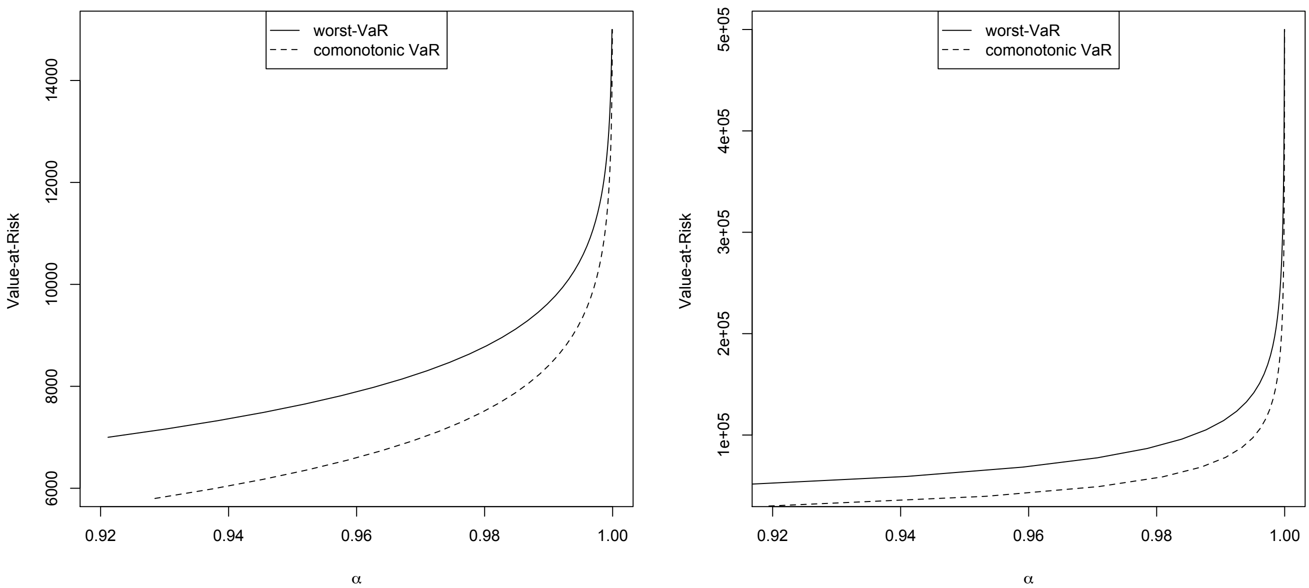

Figure 2: $\overline{\mathrm{VaR}}_\alpha(L^+)$ (see (15)) and $\mathrm{VaR}_\alpha^+(L^+)$ (see (8)) for the sum of $d = 1000$ Gamma(3, 1)- (left) and LogNormal(2, 1)- (right) distributed risks.

1. Under the assumptions of Proposition 4, the infimum in (12) is attained at $a < s/d$ if and only if the first order condition (14) holds. In order to calculate $\overline{\mathrm{VaR}}_\alpha(L^+)$ it is sufficient to compute the function $D(s)$ by solving numerically the univariate equation (14) and hence to compute numerically its inverse D^{-1} at the level $(1-\alpha)$. The treatment of an *arbitrary* number of identically distributed risks is then made possible; see Figure 2 and Table 4.

2. For the Pareto distribution (11) with tail parameter $\theta > 0$ we have that
$$\overline{\mathrm{VaR}}_\alpha(L^+) = D^{-1}(1-\alpha),$$
for any $\alpha \in (0,1)$. Portfolios of Pareto distributed risks are studied in Table 4.

3. The sharpness of the bound $D^{-1}(1-\alpha)$ in (15) can be stated under different sets of assumptions for the distribution function F. To cite a most useful case, sharpness holds for distributions F having a concave density on the interval (a,b). This allows for instance to compute the sharp bound $\overline{\mathrm{VaR}}_\alpha(L^+) = D^{-1}(1-\alpha)$ in case of Gamma and LogNormal distributions; see Figure 2.

4. The equation (15) typically holds for distributions F and confidence levels α standardly used in quantitative risk management, also in the case of heavy tailed, infinite-mean models.

5. So far, there does not exist a method which allows to compute $\underline{\mathrm{VaR}}_\alpha(L^+)$ analytically for $d \geq 3$.

When the distribution F satisfies the assumptions of Proposition 4, a worst-case dependence vector $(L_1^*, \dots, L_d^*)'$ such that $\overline{\mathrm{VaR}}_\alpha(L^+) = \mathrm{VaR}_\alpha(L_1^* + \cdots + L_d^*)'$ has been described in Wang et al. (2013) and Puccetti and Rüschendorf (2013). Here the concept of complete mixability is crucial. The random vector $(L_1^*, \dots, L_d^*)'$ satisfies the following two properties:

(a) When one of the L_i^*'s lies in the interval (a,b), then all the L_i^*'s lie in (a,b) and are a d-complete mix, i.e. for all $1 \leq i \leq d$,
$$P\left(L_1^* + \cdots + L_d^* = s \middle| L_i \in (a,b)\right) = 1;$$

(b) For all $1 \leq i \leq d$, we have that
$$P\left(L_j = F_{a^*}^{-1}\left((d-1)\overline{F}_{a^*}(L_i)\right) \middle| L_i \geq b\right) = 1, \text{ for all } j \neq i,$$

where $a^* = F^{-1}(1 - D(s))$ and $F_{a^*}(x) = (F(x) - F(a^*))/\overline{F}(a^*)$. $\overline{F}(a^*)$ is the distribution of the random variable $Y_{a^*} \overset{\mathrm{d}}{=} (L_1 | L_1 \geq a^*)$. The interdependence described by the two properties above can be summarized as:

if	$L_i \in [a^*, a]$	then	$L_j \geq b$	for some $j \neq i$,
if	$L_i \in (a,b)$	then	$\sum_{j=1}^d L_j = \overline{\mathrm{VaR}}_\alpha(L^+)$,	
if	$L_i \geq b$	then	$L_j \in [a^*, a]$	for all $j \neq i$.

The two properties (a) and (b) determine the behavior of the worst-case dependence only in the upper $(1-\alpha)$ parts of the marginal supports where $L_i \geq a^*, 1 \leq i \leq d$. Analogous to the case $d = 2$, the interdependence coupling in the α lower parts of the marginal supports can be set arbitrarily.

In Figure 3, left, we show a two-dimensional projection of the support of the d-variate copula merging the upper $(1-\alpha)$ parts of the optimal risks L_i^*. In practice, only two situations can occur: either one of the risks is large (above the threshold b) and all the others are small (below the threshold a), or all the risks are of medium size (they lie in the interval (a,b)) with their sum being equal to the threshold $\overline{\mathrm{VaR}}_\alpha(L^+)$. This is a negative dependence scenario analogous to the one underlying Figure 1. In fact the worst-VaR scenario contains a part where the risks are d-completely mixable, with the variance of their sum being equal to zero. In Figure 3, right, the two-dimensional is illustrated in case F is a Pareto(2) distribution and $\alpha = 99.9\%$. The interested reader can compare this figure with Figure 3.2 in Wang and Wang (2011).

For a risk vector $(L_1, \dots, L_d)'$ it is of interest to study the *superadditivity ratio*
$$\delta_\alpha(d) = \frac{\overline{\mathrm{VaR}}_\alpha(L^+)}{\mathrm{VaR}_\alpha^+(L^+)}$$
between the worst-possible VaR and the comonotonic VaR, at some given level of probability $\alpha \in (0,1)$. The value $\delta_\alpha(d)$ measures how much VaR can be superadditive as a function of the dimensionality d of the risk portfolio under study. For instance, for elliptically distributed risks it is well known that $\delta_\alpha(d) = 1$ for any $d \geq 1$; see McNeil et al. (2005, Theorem 6.8). A concept related to $\delta_\alpha(d)$ is the so-called diversification benefit discussed in Cope et al. (2009); for an earlier introduction of this concept, see Embrechts et al. (2002, Remark 2).

Using Proposition 4, in Figure 4 and Figure 5, left, we plot the function $\delta_\alpha(d)$ for a number of different homogeneous portfolios. In these cases, $\delta_\alpha(d)$ seems to settle down to a limit in d fairly fast. We denote
$$\delta_\alpha = \lim_{d \to +\infty} \delta_\alpha(d), \tag{16}$$
whenever this limit exists. For large dimensions d one can then approximate the worst-possible VaR value as
$$\overline{\mathrm{VaR}}_\alpha(L^+) \approx \delta_\alpha \mathrm{VaR}_\alpha^+(L^+) = d\delta_\alpha \mathrm{VaR}_\alpha(L_1). \tag{17}$$
The result (17) equivalently means that the VaR of a homogeneous risk portfolio can be δ_α times larger than the VaR under the assumption of comonotonicity. Below, we report numerical estimate for the superadditivity constant δ_α for some homogeneous risk portfolios of interest in finance and insurance. For portfolios of LogNormal(2,1)-distributed risks, we have $\delta_{0.99} \cong 1.49$ and $\delta_{0.999} \cong 1.37$; see Figure 4, left. For portfolios of Gamma(3,1)-distributed risks, we have $\delta_{0.99} \cong 1.15$ and $\delta_{0.999} \cong 1.11$; see Figure 4, right. For portfolios of Pareto(2)-distributed risks, we have $\delta_{0.99} \cong 2.11$ and $\delta_{0.999} \cong 2.03$; see Figure 5, left. In Figure 5, right, one can see how the limiting constant δ_α depends on the tail parameter θ of the Pareto marginals: the smaller the tail parameter θ, the more superadditive the VaR of the sums of the risks can be. It is also interesting that, in the

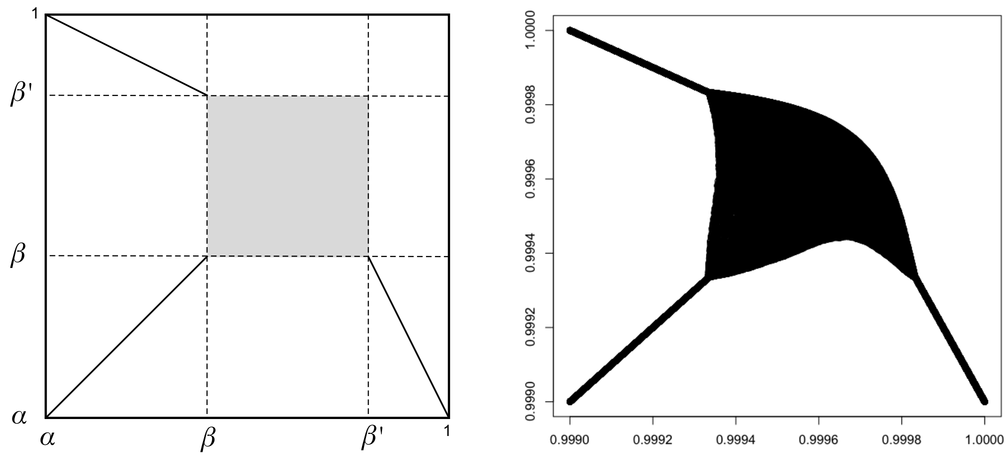

Figure 3: Left: one of the identical two-dimensional projections to $[\alpha, 1]^2$ of the support of the d-variate copula merging the upper $(1-\alpha)$ parts of the optimal risks L_i^*. In the figure, we have $\alpha = 1 - D(s), \beta = F(a)$ and $\beta' = F(b)$. The grey area represents a completely mixable part. Right: the same as in the left-part of the figure in the specific case of a portfolio of Pareto(2) distributed random variables and $\alpha = 0.999$.

examples studied, the superadditivity ratio is larger for smaller levels of α. A figure analogous to Figure 5 cannot be obtained analytically for the ratio $\mathrm{VaR}^+_\alpha(L^+)/\underline{\mathrm{VaR}}_\alpha(L^+)$; see point 5 in Remark 5. For non-homogeneous portfolios, the ratios above can be computed using the algorithm presented in Section 2.2. Motivated by the figures presented in a preliminary version of this paper, Puccetti and Rüschendorf (2012a) give an analytical proof of the limit in (16) under precise mathematical conditions. Other papers studying the superadditivity properties of risk portfolios are Mainik and Rüschendorf (2010), Mainik et al. (2013) and Mainik and Rüschendorf (2012).

2.2. *The Rearrangement Algorithm for VaR*

If one drops the assumption of identically distributed risks, the bounds given in (10) and (15) cannot be used. For $d = 2$, the sharp bounds $\underline{\mathrm{VaR}}_\alpha(L^+)$ and $\overline{\mathrm{VaR}}_\alpha(L^+)$ can be calculated easily, also in the inhomogeneous case, using Rüschendorf (1982, Proposition 1); see also Puccetti and Rüschendorf (2012b, Theorem 2.7). In higher dimension $d \geq 3$ the computation of the dual functional $D(s)$ with different marginal distributions may become numerically cumbersome. The numerical complexity of the dual bound $D(s)$ typically increases with the number of blocks of marginals with identical distributions. For instance, if all the d marginal distributions are different, the computation of dual bounds is manageable up to small dimension $d = 10$, say. An example with $d = 8$ is illustrated in Embrechts and Puccetti (2006a). However, it is possible to compute the dual bound $D(s)$ for relatively large dimensions d if the inhomogeneous risks L_i can be divided in n sub-groups having homogeneous marginals within. In this case, the numerical complexity of the dual bound $D(s)$ only depends on n, and is independent of the cardinality of each of the sub-groups of homogeneous marginals. It is also important to remark that the sharpness of dual bounds in dimension $d \geq 3$ has not been proved for inhomogeneous marginals.

For the computation of bounds on distribution functions Puccetti and Rüschendorf (2012c) introduced a *rearrangement algorithm* (RA) working well for dimension $d \leq 30$. In this paper we adapt and greatly improve this RA in order to compute the sharp bounds $\overline{\mathrm{VaR}}_\alpha(L^+)$ and $\underline{\mathrm{VaR}}_\alpha(L^+)$ in the inhomogeneous case. While the algorithm described in Puccetti and Rüschendorf (2012c) requires a time-consuming numeric inversion for the computation of VaR bounds, our modified version does not need any inversion and also decreases the number of iterations needed to obtain the final estimate by introducing a new termination condition based on the accuracy of the final estimate. Our modifications allow to apply the algorithm to high-dimensional inhomogeneous portfolios, even for dimensions $d \geq 1000$, say, which previously were well out of the range of numerical and analytical methods. Examples using dimensionality in the several hundreds are of particular interest in internal models built by financial institutions in order to fulfil the Basel and Solvency regulatory guidelines. An example where high dimensionality really occurs is to be found in the hierarchical aggregation model described in Section 5 in Arbenz et al. (2012), in use at SCOR, which determines the total solvency capital requirements of insurance companies using the standard model of QIS 5 by the European Insurance and Occupational Pensions Authority (EIOPA). We are also aware that some reinsurance companies have undisclosed internal models with d-values between 500 and 2000 marginal risks.

The RA can compute the worst and best VaR values in (4) with excellent accuracy for *any* set of marginals F_i and large dimensions d. In the following, we say that two vectors $\boldsymbol{a}, \boldsymbol{b} \in \mathbb{R}^N$ are *oppositely ordered* if $(a_j - a_k)(b_j - b_k) \leq 0$ holds for all $1 \leq j, k \leq N$. For a $(N \times d)$-matrix $\boldsymbol{X}$ define the operators $\underline{s}(\boldsymbol{X})$ and $\underline{t}(\boldsymbol{X})$ as

$$s(\boldsymbol{X}) = \min_{1\leq i\leq N} \sum_{1\leq j\leq d} x_{i,j}, \quad t(\boldsymbol{X}) = \max_{1\leq i\leq N} \sum_{1\leq j\leq d} x_{i,j},$$

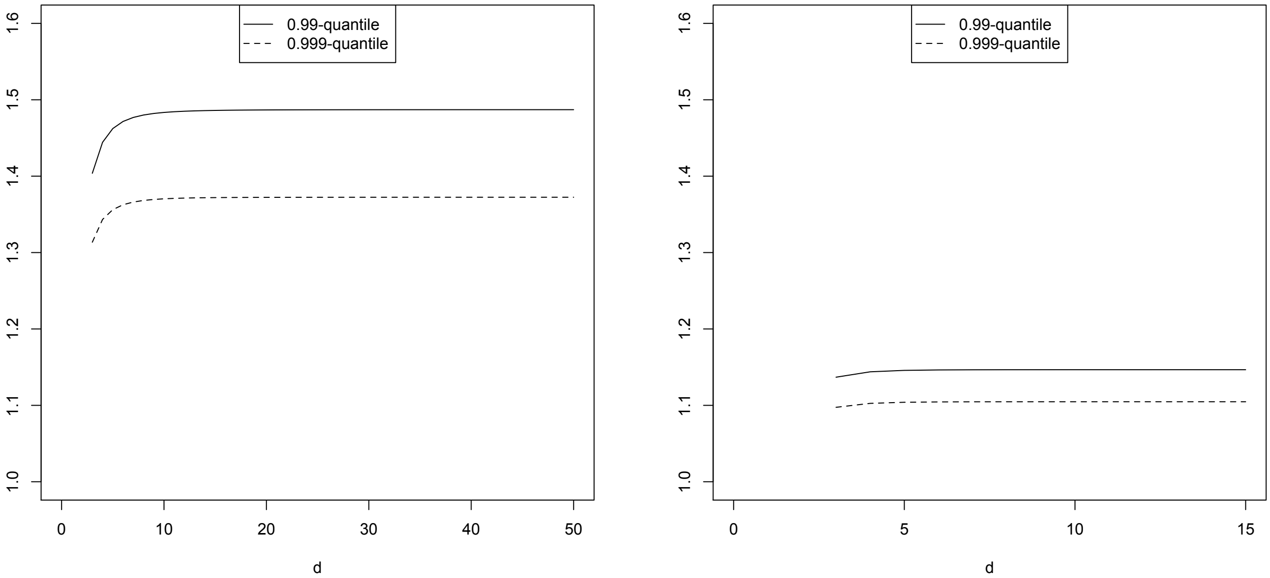

Figure 4: Plot of the function $\delta_\alpha(d)$ versus the dimensionality d of the portfolio for a risk vector of LogNormal(2,1)-distributed (left) and Gamma(3,1)-distributed (right) risks, for two different quantile levels.

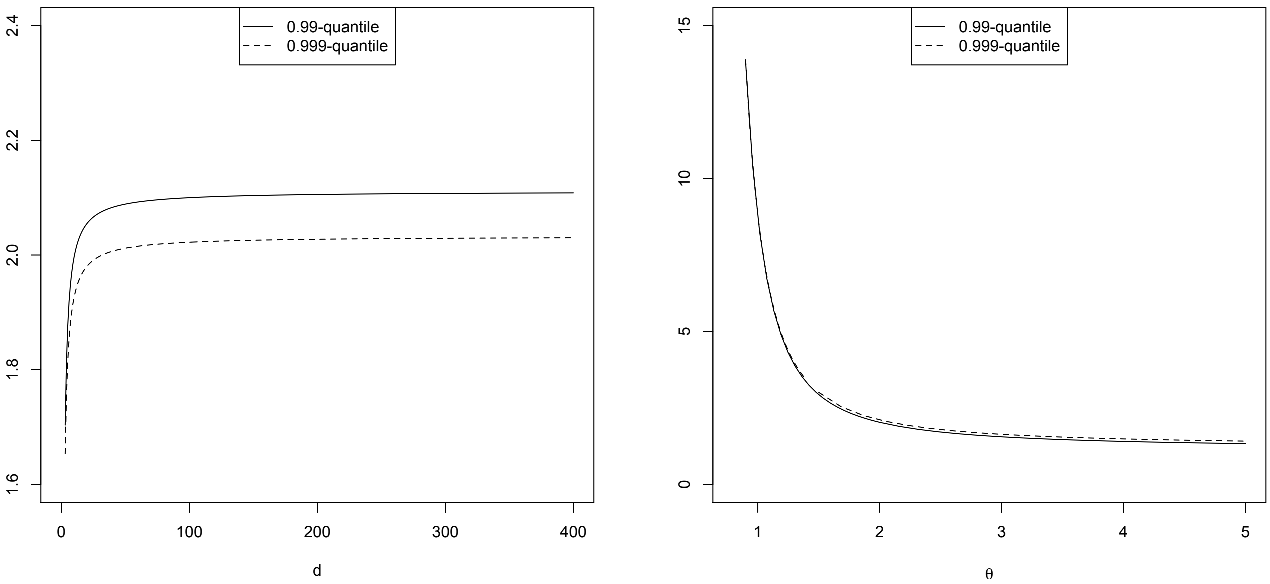

Figure 5: Left: plot of the function $\delta_\alpha(d)$ versus the dimensionality d of the portfolio for a risk vector of Pareto(θ)-distributed risks, for two different quantile levels and $\theta = 2$. Right: plot of the limiting constant δ_α versus the tail parameter θ of the Pareto distribution.

the row-wise minimum, respectively maximum, of the row-sums of $\boldsymbol{X}$.

Rearrangement Algorithm (RA) to compute $\overline{\mathbf{VaR}}_\alpha(L^+)$.

1. Fix a integer N and the desired level of accuracy $\epsilon > 0$.
2. Define the matrices $\underline{\boldsymbol{X}}^\alpha = (\underline{x}^\alpha_{i,j})$ and $\overline{\boldsymbol{X}}^\alpha = (\overline{x}^\alpha_{i,j})$ as
$$\underline{x}^\alpha_{i,j} = F_j^{-1}\left(\alpha + \frac{(1-\alpha)(i-1)}{N}\right), \quad \overline{x}^\alpha_{i,j} = F_j^{-1}\left(\alpha + \frac{(1-\alpha)i}{N}\right), \tag{18}$$
for $1 \le i \le N, 1 \le j \le d$.
3. Permute randomly the elements in each column of $\underline{\boldsymbol{X}}^\alpha$ and $\overline{\boldsymbol{X}}^\alpha$.
4. Iteratively rearrange the j–th column of the matrix $\underline{\boldsymbol{X}}^\alpha$ so that it becomes oppositely ordered to the sum of the other columns, for $1 \le j \le d$. A matrix $\underline{\boldsymbol{Y}}^\alpha$ is found.
5. Repeat Step 4. until
$$s(\underline{\boldsymbol{Y}}^\alpha) - s(\underline{\boldsymbol{X}}^\alpha) < \epsilon.$$
A matrix $\underline{\boldsymbol{X}}^*$ is found.
6. Apply Steps 4.–5. to the matrix $\overline{\boldsymbol{X}}^\alpha$ until a matrix $\overline{\boldsymbol{X}}^*$ is found.
7. Define
$$\underline{s}_N = s(\underline{\boldsymbol{X}}^*) \quad \text{and} \quad \overline{s}_N = s(\overline{\boldsymbol{X}}^*).$$
Then we have $\underline{s}_N \le \overline{s}_N$ and in practice we find that
$$\overline{s}_N \overset{N\to\infty}{\simeq} \underline{s}_N \overset{N\to\infty}{\simeq} \overline{\mathrm{VaR}}_\alpha(L^+). \tag{19}$$

Rearrangement Algorithm (RA) to compute $\underline{\mathbf{VaR}}_\alpha(L^+)$.

1. Fix a integer N and the desired level of accuracy $\epsilon > 0$.
2. Define the matrices $\underline{\boldsymbol{Z}}^\alpha = (\underline{z}^\alpha_{i,j})$ and $\overline{\boldsymbol{Z}}^\alpha = (\overline{z}^\alpha_{i,j})$ as
$$\underline{z}^\alpha_{i,j} = F_j^{-1}\left(\frac{\alpha(i-1)}{N}\right), \quad \overline{z}^\alpha_{i,j} = F_j^{-1}\left(\frac{\alpha i}{N}\right), \tag{20}$$
for $1 \le i \le N, 1 \le j \le d$.
3. Permute randomly the elements in each column of $\underline{\boldsymbol{Z}}^\alpha$ and $\overline{\boldsymbol{Z}}^\alpha$.
4. Iteratively rearrange the j–th column of the matrix $\underline{\boldsymbol{Z}}^\alpha$ so that it becomes oppositely ordered to the sum of the other columns, for $1 \le j \le d$. A matrix $\underline{\boldsymbol{W}}^\alpha$ is found.
5. Repeat Step 4. until
$$t(\underline{\boldsymbol{Z}}^\alpha) - t(\underline{\boldsymbol{W}}^\alpha) < \epsilon.$$
A matrix $\underline{\boldsymbol{Z}}^*$ is found.
6. Apply Steps 4.–5. to the matrix $\overline{\boldsymbol{Z}}^\alpha$ until a matrix $\overline{\boldsymbol{Z}}^*$ is found.
7. Define
$$\underline{t}_N = t(\underline{\boldsymbol{Z}}^*) \quad \text{and} \quad \overline{t}_N = t(\overline{\boldsymbol{Z}}^*).$$
Then we have $\underline{t}_N \le \overline{t}_N$ and in practice we find that
$$\overline{t}_N \overset{N\to\infty}{\simeq} \underline{t}_N \overset{N\to\infty}{\simeq} \underline{\mathrm{VaR}}_\alpha(L^+). \tag{21}$$

Remark 6. For mathematical details about the RA, we refer the reader to Puccetti and Rüschendorf (2012c). Here we limit our attention to the following, more practical points:

1. We call the interval $(\underline{s}_N, \overline{s}_N)$ the rearrangement range for $\overline{\mathrm{VaR}}_\alpha(L^+)$. The length $(\overline{s}_N - \underline{s}_N)$ of this interval depends on the dimensionality d of the risk portfolio under study and on N, the upper-tail discretization parameter. For sufficiently large N, we also have that $\underline{s}_N \le \overline{\mathrm{VaR}}_\alpha(L^+)$. Analogous considerations can be made for the rearrangement range $(\underline{t}_N, \overline{t}_N)$ for $\underline{\mathrm{VaR}}_\alpha(L^+)$. For sufficiently large N we have that $\overline{t}_N \ge \underline{\mathrm{VaR}}_\alpha(L^+)$.
2. There does not exist an analytic proof that results (19) and (21) hold for all initial configurations of the algorithm. Robert Weismantel provided examples with $F_j = \mathrm{U}(0,1)$, the uniform distribution, in which the sequences $(\overline{s}_N - \underline{s}_N)$ and $(\overline{t}_N - \underline{t}_N)$ do not converge to zero. These examples are however based upon a special choice of the starting matrix of the algorithm. Using the randomization Step 3. we found the algorithm to provide excellent approximations with moderately large values of N for all marginal distributions typically used in quantitative risk management. Also using this randomisation step, a proof of convergence of the RA remains an open problem.
3. In Table 4, we check the accuracy of the RA for some Pareto(2) risk portfolios for which we know, by Proposition 4, the exact value of $\overline{\mathrm{VaR}}_\alpha(L^+)$. This table also highlights the possibly large difference between the comonotonic $\mathrm{VaR}^+_\alpha(L^+)$ and the worst-possible $\overline{\mathrm{VaR}}_\alpha(L^+)$. In Table 4 we use different dimensions d as well as values of N which represent a good compromise between computational time used and accuracy obtained. In order to perform all the computations in the remainder of the paper we use an Apple MacBook Air (2 GHz Intel Core i7, 8 GB RAM). Computation times can no doubt be dramatically reduced on a more powerful machine.
4. As a numerical algorithm, the RA can be used with *any* type of marginal distributions, including empirical distribution functions. The figures in Table 4 are obtained for a homogeneous portfolio so as to be able to check the accuracy of the RA via the dual bound in Proposition 4. In general, if the evaluation of quantile functions F_j^{-1} in Step 2. is trivial both the accuracy and the computation time of the RA are *not* affected by the type of the marginal distributions used. If one instead has to evaluate the quantile functions by numerical root finding algorithms, this might be the most demanding point of the RA from a computational point of view, as we

show in Section 4.1. We thank Marius Hofert for this comment.

The probabilistic idea behind the RA is easy. For a fixed $\alpha \in [0,1]$, the j-th columns of the matrices $\underline{\boldsymbol{X}}^\alpha$ and $\overline{\boldsymbol{X}}^\alpha$ defined in (18) represent two stochastically ordered N-point discretizations of the $(1-\alpha)$ upper parts of the supports of the marginal risk L_j. The RA rearranges the columns of $\underline{\boldsymbol{X}}^\alpha$ into the matrix $\underline{\boldsymbol{X}}^*$ in order to find the maximal value $\underline{s}_N$ such that the componentwise sum of any row of $\underline{\boldsymbol{X}}^*$ is larger than $\underline{s}_N$. Analogously, the RA rearranges the columns of $\overline{\boldsymbol{X}}^\alpha$ into the matrix $\overline{\boldsymbol{X}}^*$ in order to find the maximal value $\overline{s}_N$ such that the componentwise sum of any row of $\overline{\boldsymbol{X}}^*$ is larger than $\overline{s}_N$. For N large enough we have that $\underline{s}_N \leq \overline{\mathrm{VaR}}_\alpha(L^+) \simeq \overline{s}_N$ as a consequence of Puccetti and Rüschendorf (2012c, Theorem 3.1). An analogous mechanism yields $\underline{\mathrm{VaR}}_\alpha(L^+)$.

We first illustrate the RA in an example with $d = 3$ Pareto marginals with identical tail parameters $\theta = 2$; the homogeneous case. Then, we set $N = 50$ and compute $\overline{\mathrm{VaR}}_\alpha(L^+)$ for $\alpha = 0.99$ via the RA. The initial matrix $\underline{\boldsymbol{X}}^\alpha$ defined in (18) for $\alpha = 0.99$ is shown in Table 1 (A). The j-th column of $\underline{\boldsymbol{X}}^\alpha$ represents a 50-point discretization of the upper 1% of the support of the j-th marginal distribution. In the same (A) part of the table, we also show the N-dimensional vector of the row-wise sums of $\underline{\boldsymbol{X}}^\alpha$, as well as the d-dimensional vector having as components the aggregate sums of the columns of $\underline{\boldsymbol{X}}^\alpha$.

During the iteration of the algorithm (Steps 3.–5.), the elements within each column of $\underline{\boldsymbol{X}}^\alpha$ are re-shuffled until a matrix $\underline{\boldsymbol{X}}^*$ is found with each column oppositely ordered to the sum of the others, see Table 1 (B); we re-ordered (B) in ascending order with respect to the row sums (final column). This rearrangement procedure of the columns of $\underline{\boldsymbol{X}}^\alpha$ aims at maximizing the minimal component of the vector of the row-wise sums of $\underline{\boldsymbol{X}}^*$. Indeed, note how the minimal component of the row-wise sums (27.0000) is increased (to 44.7671) when passing from $\underline{\boldsymbol{X}}^\alpha$ to $\underline{\boldsymbol{X}}^*$, while the column-wise sums remain unchanged (the marginals are still the same). Compared to $\underline{\boldsymbol{X}}^\alpha$, the matrix $\underline{\boldsymbol{X}}^*$ represents a different coupling (copula) of the same marginals in which the variance of the marginal numbers (rows) is reduced. The minimal component of the vector of the sums of the rows of $\underline{\boldsymbol{X}}^*$ is $\underline{s}_{50} = 44.7671$ and represents a lower bound on $\overline{\mathrm{VaR}}_\alpha(L^+)$. Performing an analogous rearrangement of the column of the matrix $\overline{\boldsymbol{X}}^\alpha$ one finds $\overline{s}_{50} = 46.4111$, which is instead approximately an upper bound on $\overline{\mathrm{VaR}}_\alpha(L^+)$. Note that the estimates $\underline{s}_{50}$ and $\overline{s}_{50}$ are actually random due to the randomization occurring in Step 3. This random uncertainty becomes negligible for values of N large enough. From the application of the RA described above for $N = 50$ one obtains $\overline{\mathrm{VaR}}_\alpha(L^+) \in [44.77, 46.41]$. It is sufficient to run the algorithm with $N = 1.0e05$ to obtain the first two decimals of $\overline{\mathrm{VaR}}_\alpha(L^+) = 45.99$ in less than one second. Of course, in this pedagogical case one could instantly obtain the exact value $\overline{\mathrm{VaR}}_\alpha(L^+) = 45.99$ from Proposition 4. The power of the RA is that it can be applied also to inhomogeneous portfolios of risks and is able to compute numerically also $\underline{\mathrm{VaR}}_\alpha(L^+)$.

It is interesting to see that already for $N = 50$, the final matrix $\underline{\boldsymbol{X}}^*$ in Table 1 (B) approximates the worst-case dependence for the sum of continuous homogeneous marginals shown in Figure 3. In Table 1 (B) we have ordered the final RA-output matrix in function of the last column. One can now easily check that basically two structures occur in the rows of $\underline{\boldsymbol{X}}^*$: either all the components of a row are close to each other, and sum up to a value which is just above the threshold $\underline{s}_{50} = 44.7671$ (rows 1–29), or one of them is large and all the others are small (rows 30–50). Of course, this structural dichotomy becomes much clearer when N increases and can also be made precise and proved mathematically; see Section 2.1 above.

2.3. Application to Operational Risk data

As a more realistic example stemming from Operational Risk, we study a risk portfolio where the marginal losses are distributed like a Generalized Pareto Distribution (GPD), that is we assume that

$$F_i(x) = 1 - \left(1 + \xi_i \frac{x}{\beta_i}\right)^{-1/\xi_i}, \quad x \geq 0, \ 1 \leq i \leq d. \tag{22}$$

For a GPD distribution, whenever $\xi_i \geq 1$, $E(L_i) = \infty$, and for $1/2 \leq \xi_i < 1$, $E(L_i) < \infty$ but $\mathrm{var}(L_i) = \infty$. Moscadelli (2004) contains an analysis of the Basel II data on Operational Risk coming out of the second Quantitative Impact Study (QIS); see also Chapter 10 in McNeil et al. (2005) for a discussion and further references. In this case $d = 8$ and for the parameters of the GPD distributions we take the values reported in Moscadelli (2004) for the losses in eight OR business lines. The values for the parameters in the different business lines are summarized in Table 2. Under these marginal assumptions, the risk vector $(L_1, \ldots, L_d)'$ exhibits very heavy-tailed behavior, with six out of eight losses L_i following an infinite mean marginal model. In the other two cases, where the mean is finite, the loss distributions do not have finite variance. Note that we use the parameter values of ξ_i and β_i from Moscadelli (2004) as a matter of example and do not consider here the remaining extra statistical issues underlying a full Operational Risk analysis for which we refer to Frachot et al. (2004), de Fontnouvelle et al. (2005) and the various references to the Operational Risk literature mentioned before.

In Table 3, we give the VaR range (5) as well as the estimates for $\mathrm{VaR}_\alpha^+(L^+)$ (VaR under comonotonicity) and $\mathrm{VaR}_\alpha(L^{\Pi,+})$ (VaR under independence) versus the confidence level α. $\mathrm{VaR}_\alpha(L^{\Pi,+})$ has been computed via the approximation

$$P(L^{\Pi,+} > x) \overset{x\to\infty}{\sim} P\left(\max_{1\leq i\leq d} L_i > x\right). \tag{23}$$

The above formula is valid for iid subexponential random variables, as explained in Embrechts et al. (1997, Section 1.3.2). It also holds more generally whenever the underlying random variables are independent with the heaviest tail being subexponential; see Lemma A3.28 in Embrechts et al. (1997). The resulting approximation goes under the name *the largest loss approximation* and has been used in the Operational Risk literature; see for instance Böcker and Klüppelberg (2010). In general, approximations of the type (23) are numerically bad,

(A)	1	2	3	∑
1	9.00000	9.00000	9.00000	27.0000
2	9.10153	9.10153	9.10153	27.3046
3	9.20621	9.20621	9.20621	27.6186
4	9.31421	9.31421	9.31421	27.9426
5	9.42572	9.42572	9.42572	28.2772
6	9.54093	9.54093	9.54093	28.6228
7	9.66004	9.66004	9.66004	28.9801
8	9.78328	9.78328	9.78328	29.3498
9	9.91089	9.91089	9.91089	29.7327
10	10.04315	10.04315	10.04315	30.1295
11	10.18034	10.18034	10.18034	30.5410
12	10.32277	10.32277	10.32277	30.9683
13	10.47079	10.47079	10.47079	31.4124
14	10.62476	10.62476	10.62476	31.8743
15	10.78511	10.78511	10.78511	32.3553
16	10.95229	10.95229	10.95229	32.8569
17	11.12678	11.12678	11.12678	33.3803
18	11.30915	11.30915	11.30915	33.9274
19	11.50000	11.50000	11.50000	34.5000
20	11.70001	11.70001	11.70001	35.1000
21	11.90994	11.90994	11.90994	35.7298
22	12.13064	12.13064	12.13064	36.3919
23	12.36306	12.36306	12.36306	37.0892
24	12.60828	12.60828	12.60828	37.8248
25	12.86750	12.86750	12.86750	38.6025
26	13.14214	13.14214	13.14214	39.4264
27	13.43376	13.43376	13.43376	40.3013
28	13.74420	13.74420	13.74420	41.2326
29	14.07557	14.07557	14.07557	42.2267
30	14.43033	14.43033	14.43033	43.2910
31	14.81139	14.81139	14.81139	44.4342
32	15.22214	15.22214	15.22214	45.6664
33	15.66667	15.66667	15.66667	47.0000
34	16.14986	16.14986	16.14986	48.4496
35	16.67767	16.67767	16.67767	50.0330
36	17.25742	17.25742	17.25742	51.7723
37	17.89822	17.89822	17.89822	53.6947
38	18.61161	18.61161	18.61161	55.8348
39	19.41241	19.41241	19.41241	58.2372
40	20.32007	20.32007	20.32007	60.9602
41	21.36068	21.36068	21.36068	64.0820
42	22.57023	22.57023	22.57023	67.7107
43	24.00000	24.00000	24.00000	72.0000
44	25.72612	25.72612	25.72612	77.1784
45	27.86751	27.86751	27.86751	83.6025
46	30.62278	30.62278	30.62278	91.8683
47	34.35534	34.35534	34.35534	103.0660
48	39.82483	39.82483	39.82483	119.4745
49	49.00000	49.00000	49.00000	147.0000
50	69.71068	69.71068	69.71068	209.1320
∑	851.72901	851.72901	851.72901	

(B)	1	2	3	∑
1	13.43376	15.66667	15.66667	44.7671
2	12.86750	14.81139	17.25742	44.9363
3	15.66667	13.14214	16.14986	44.9587
4	14.43033	16.14986	14.43033	45.0105
5	15.22214	11.90994	17.89822	45.0303
6	13.74420	17.89822	13.43376	45.0762
7	19.41241	12.86750	12.86750	45.1474
8	20.32007	11.70001	13.14214	45.1622
9	18.61161	14.43033	12.13064	45.1726
10	13.14214	13.43376	18.61161	45.1875
11	16.14986	12.36306	16.67767	45.1906
12	11.12678	13.74420	20.32007	45.1910
13	11.70001	12.13064	21.36068	45.1913
14	11.50000	22.57023	11.12678	45.1970
15	21.36068	11.50000	12.36306	45.2237
16	14.81139	15.22214	15.22214	45.2557
17	12.60828	21.36068	11.30915	45.2781
18	14.07557	18.61161	12.60828	45.2955
19	17.89822	12.60828	14.81139	45.3179
20	11.30915	20.32007	13.74420	45.3734
21	11.90994	14.07557	19.41241	45.3979
22	17.25742	16.67767	11.50000	45.4351
23	12.13064	19.41241	14.07557	45.6186
24	16.67767	17.25742	11.70001	45.6351
25	12.36306	10.78511	22.57023	45.7184
26	10.78511	10.95229	24.00000	45.7374
27	22.57023	11.30915	11.90994	45.7893
28	10.95229	24.00000	10.95229	45.9046
29	24.00000	11.12678	10.78511	45.9119
30	25.72612	10.62476	10.47079	46.8217
31	10.62476	10.47079	25.72612	46.8217
32	10.47079	25.72612	10.62476	46.8217
33	10.32277	27.86751	10.18034	48.3706
34	27.86751	10.18034	10.32277	48.3706
35	10.18034	10.32277	27.86751	48.3706
36	9.91089	10.04315	30.62278	50.5768
37	10.04315	30.62278	9.91089	50.5768
38	30.62278	9.91089	10.04315	50.5768
39	9.78328	34.35534	9.66004	53.7987
40	34.35534	9.66004	9.78328	53.7987
41	9.66004	9.78328	34.35534	53.7987
42	9.42572	9.54093	39.82483	58.7915
43	39.82483	9.42572	9.54093	58.7915
44	9.54093	39.82483	9.42572	58.7915
45	49.00000	9.31421	9.20621	67.5204
46	9.31421	9.20621	49.00000	67.5204
47	9.20621	49.00000	9.31421	67.5204
48	9.00000	9.10153	69.71068	87.8122
49	69.71068	9.00000	9.10153	87.8122
50	9.10153	69.71068	9.00000	87.8122
∑	851.72901	851.72901	851.72901	

Table 1: (A): The matrix $\underline{\boldsymbol{X}}^{\alpha}$ defined in (18) for $\alpha = 0.99$ and $N = 50$ (representing comonotonicity among the discrete marginals); (B): The matrix $\underline{\boldsymbol{X}}^{*}$ derived as an output of the iterative rearrangement of the columns of $\underline{\boldsymbol{X}}^{\alpha}$. The rows of $\underline{\boldsymbol{X}}^{*}$ are ordered accordingly to their sums. In this example we consider a discretization of $d = 3$ Pateto(2)-distributed risks.

Business line	i	ξ_i	β_i
Corporate Finance	1	1.19	774
Trading & Sales	2	1.17	254
Retail Banking	3	1.01	233
Commercial Banking	4	1.39	412
Payment & Settlement	5	1.23	107
Agency Services	6	1.22	243
Asset Management	7	0.85	314
Retail Brokerage	8	0.98	124

Table 2: Parameter values for the eight tail GPD-distributed risks following Moscadelli (2004). Note that Moscadelli (2004) uses tail GPD marginal models instead of pure GPD marginals as in (22).

α	$\underline{\mathrm{VaR}}_\alpha(L^+)$	$\mathrm{VaR}^+_\alpha(L^+)$	$\mathrm{VaR}_\alpha(L^{\Pi,+})$	$\overline{\mathrm{VaR}}_\alpha(L^+)$
0.99	1.78×10^5	5.14×10^5	7.08×10^5	2.56×10^6
0.995	4.68×10^5	1.22×10^6	1.68×10^6	5.96×10^6
0.999	4.38×10^6	9.33×10^6	1.28×10^7	4.34×10^7

Table 3: Estimates for $\mathrm{VaR}_\alpha(L^+)$ for a random vector of $d = 8$ GPD-distributed risks having the parameters in Table 2 and different dependence assumptions, i.e. (from left to right) best-case dependence, comonotonicity, independence, worst-case dependence. Each estimate for $\underline{\mathrm{VaR}}_\alpha(L^+)$ and $\overline{\mathrm{VaR}}_\alpha(L^+)$ has been obtained via the RA in about 9 mins using $N = 2\times10^6$ and $\epsilon = 0.1$.

except in the very heavy-tailed case, as we have here. From a more applied point of view, concerning Operational Risk, (23) does indeed occur more frequently as such cases like Nick Leeson (Barings Bank), Jérôme Kerviel (Société Générale) and Kweku Adoboli (UBS) show. The recent scandal around the LIBOR-fixing yields another example of the general idea behind (23).

VaR figures in Table 3 clearly show that the VaR estimate $\mathrm{VaR}^+_\alpha(L^+)$ is inadequate to capture the riskiness of the portfolio as

$$\mathrm{VaR}_\alpha(L^{\Pi,+}) > \mathrm{VaR}^+_\alpha(L^+),$$

a fact typically occurring when some of the marginal distributions have infinite mean. For practice, the wide VaR range for values of α typically used, that is $\alpha = 0.99, 0.995, 0.999$, should raise some concerns. For the dimension $d = 8$ in the Moscadelli example, the RA algorithm produces accurate estimate of $\overline{\mathrm{VaR}}_\alpha(L^+)$ and $\underline{\mathrm{VaR}}_\alpha(L^+)$ in about 9 mins with $N = 2\times10^6$. The results in these examples imply a considerable model uncertainty issue underlying VaR calculations for confidence levels close to 1.

3. Positive dependence information

The worst-VaR copulas given in Section 2, Figures 1 and 3, are often considered as unrealistic due to their minimal variance parts in which the risks are countermonotonic (for $d = 2$) or completely mixable (in the case $d \geq 3$). Of course, a positive dependence structure, as defined below, combined with the knowledge of the marginal distributions of $(L_1,\dots,L_d)'$ will *tighten* the interval of admissible VaRs in (5). However, assuming that the risks are positively dependent *does not* eliminate countermonotonicty and completely mixable parts from the worst-VaR scenarios and does not necessarily lower the estimate of $\overline{\mathrm{VaR}}_\alpha(L^+)$ by much. This latter point is the object of this section. We start by introducing a natural concept of positive dependence.

Definition 7. *The risk vector* $(L_1,\dots,L_d)'$ *is said to be* positively lower orthant dependent (PLOD) *if for all* $(x_1,\dots,x_d)' \in \mathbb{R}^d$

$$P(L_1 \leq x_1,\dots,L_d \leq x_d) \geq \prod_{i=1}^d P(X_i \leq x_i) = \prod_{i=1}^d F_i(x_i). \quad (24)$$

The risk vector $(L_1,\dots,L_d)'$ *is said to be* positively upper orthant dependent (PUOD) *if for all* $(x_1,\dots,x_d)' \in \mathbb{R}^d$

$$P(L_1 > x_1,\dots,L_d > x_d) \geq \prod_{i=1}^d P(X_i > x_i) = \prod_{i=1}^d \overline{F}_i(x_i). \quad (25)$$

Finally, the risk vector $(L_1,\dots,L_d)'$ *is said to be* positively orthant dependent (POD) *if it is both PLOD and PUOD.*

For $d = 2$, conditions (24) and (25) are equivalent. However, this is *not* the case for $d \geq 3$. In higher dimensions the PLOD and PUOD concepts are distinct; see for instance Nelsen (2006, Section 5.7). If $(L_1,\dots,L_d)'$ has copula C, condition (24) can be equivalently expressed as $C \geq \Pi$, the independence copula. Analogously, condition (25) can be written as $\overline{C} \geq \overline{\Pi}$, where $\overline{C}$ denotes the joint tail function of a copula C, also referred to as the survival copula; see Nelsen (2006, Section 2.6). Also note that POD implies positive correlations, given that the second moments exist.

Under the addition of a positive dependence restriction, VaR bounds for the sum of risks have been derived in Theorem 3.1 in Embrechts et al. (2003); see also Embrechts et al. (2005), Mesfioui and Quessy (2005), Rüschendorf (2005) and Puccetti and Rüschendorf (2012b). We state this result here in the case of identical marginals using the same notation as in the so-called unconstrained case, i.e. with no dependence information.

Proposition 8. *In the homogeneous case* $F_i = F$, $1 \leq i \leq d$, *let* F *be a distribution with decreasing density on its entire domain. If the risk vector* $(L_1,\dots,L_d)'$ *is PLOD then, for any fixed real threshold* s, *we have*

$$\overline{\mathrm{VaR}}_\alpha(L^+) \leq dF^{-1}\left((1-\alpha)^{\frac{1}{d}}\right). \quad (26)$$

Remark 9. In Embrechts et al. (2005), the bound (26) is given in a slightly more complicated form for any set of marginal distributions. In the same reference, an analogous bound for $\underline{\mathrm{VaR}}_\alpha(L^+)$ is given if the risk vector is assumed to be PUOD.

In the case $d = 2$ the inequality given in (26) is sharp. In Figure 6, left, we show the copula of a PLOD risk vector $(L_1^*, L_2^*)'$ for which $\mathrm{VaR}_\alpha(L_1^* + L_2^*) = \overline{\mathrm{VaR}}_\alpha(L^+)$. Even if the structure of dependence of this vector is PLOD, its geometry

d **= 8**	$N = 1.0e05$	*avg time*: 5 secs		
α	$\underline{\mathrm{VaR}}_\alpha(L^+)$ (RA range)	$\mathrm{VaR}^+_\alpha(L^+)$ (exact)	$\overline{\mathrm{VaR}}_\alpha(L^+)$ (exact)	$\overline{\mathrm{VaR}}_\alpha(L^+)$ (RA range)
0.99	9.00 – 9.00	72.00	141.67	141.66–141.67
0.995	13.13 – 13.14	105.14	203.66	203.65–203.66
0.999	30.47 – 30.62	244.98	465.29	465.28–465.30
d **= 56**	$N = 1.0e05$	*avg time*: 60 secs		
α	$\underline{\mathrm{VaR}}_\alpha(L^+)$ (RA range)	$\mathrm{VaR}^+_\alpha(L^+)$ (exact)	$\overline{\mathrm{VaR}}_\alpha(L^+)$ (exact)	$\overline{\mathrm{VaR}}_\alpha(L^+)$ (RA range)
0.99	45.82 – 45.82	504.00	1053.96	1053.80–1054.11
0.995	48.60 – 48.61	735.96	1513.71	1513.49–1513.93
0.999	52.56 – 52.58	1714.88	3453.99	3453.49–3454.48
d **= 648**	$N = 5.0e04$	*avg time*: 40 mins		
α	$\underline{\mathrm{VaR}}_\alpha(L^+)$ (RA range)	$\mathrm{VaR}^+_\alpha(L^+)$ (exact)	$\overline{\mathrm{VaR}}_\alpha(L^+)$ (exact)	$\overline{\mathrm{VaR}}_\alpha(L^+)$ (RA range)
0.99	530.12 – 530.24	5832.00	12302.00	12269.74-12354.00
0.995	562.33 – 562.50	8516.10	17666.06	17620.45-17739.60
0.999	608.08 – 608.47	19843.56	40303.48	40201.48-40467.92

Table 4: Estimates for $\overline{\mathrm{VaR}}_\alpha(L^+)$ and $\underline{\mathrm{VaR}}_\alpha(L^+)$ for random vectors of Pareto(2)-distributed risks. Computation times are for a single interval with $\epsilon = 10^{-3}$.

is not so different if compared to the optimal copula in the unconstrained case (Figure 1, left). Again, the copula of $(L_1^*, L_2^*)'$ contains a countermonotonic part, in which the risks are a.s. decreasing functions of each other. Thus, the assumption of positive dependence *does not* eliminate the possibility of such optimal copulas. The reason for this is not to be found in the concept of VaR but rather raises some questions about the appropriateness of PLOD (PUOD) as a concept of positive (negative) dependence.

Given the shape of the copula attaining the bound (26) under additional positive dependence restrictions, one cannot expect an essential improvement of the VaR bound given in the unconstrained case when only the marginals of the L_i's are known. Indeed, in Figure 6, right, we plot $\overline{\mathrm{VaR}}_\alpha(L^+)$ (see (26)) and $\mathrm{VaR}^+_\alpha(L^+)$ (see (8)), for the sum of two Pareto(2) distributions. The improvement of the bound given by the additional information is negligible.

The situation gets more involved in higher dimensions ($d \geq 3$), as the bound (26) fails to be sharp. The dual bound given in (15) for the unconstrained case actually turns out to be better than (26) with positive dependence information; this can be seen in Figure 7. This is not so suprising, as the dual bound given in (15) derives from a different methodology based on the powerful tool offered by the theory of mass transportation; see Embrechts and Puccetti (2006b) on this. As a matter of fact, the bound (26) is not useful for higher dimensions ($d \geq 3$) where the search for a sharp bound with marginal and positive dependence information is still open. However, we do not expect much improvement over the dual bounds even for optimal ones in the positive dependence case. Take for instance the problem of maximizing the covariance of $(L_1, L_2)'$ when $d = 2$ and the marginals F_1 and F_2 are given. By Hoeffding's covariance representation formula, see McNeil et al. (2005, Lemma 5.2.4), one has

$$\mathrm{Cov}(L_1, L_2) = \int (F(x_1, x_2) - F_1(x_1)F_2(x_2))\, dx_1\, dx_2,$$

where F is the joint distribution of $(L_1, L_2)'$. It is clear that here the PLOD constraint $F(x_1, x_2) \geq F_1(x_1)F_2(x_2)$ does not help to improve an upper bound on $\mathrm{Cov}(L_1, L_2)$.

Another example where positive additional information does not lead to improved bounds is the problem of maximizing the Expected Shortfall (ES) of a sum of risks with given marginal distributions. Since the worst ES is attained under comonotonic dependence, a restriction to PLOD/PUOD dependence will lead to the same solution. For a definition and more details on the maximization of ES, see Section 6.1 in McNeil et al. (2005).

4. Higher dimensional dependence information

For a vector $(L_1, \ldots, L_d)'$ for which one only knows the marginal distributions $F_1, \ldots, F_d$, we have (5). If one adds PLOD/PUOD information on top of the knowledge of the marginals, the worst VaR in (5) is only minimally affected. It is clear that in practice more dependence information on the vector $(L_1, \ldots, L_d)'$ may be available. Such a case would be when specific assumptions on sub-vectors of $(L_1, \ldots, L_d)'$ are made. One reason for this could be that the individual risk factors may be grouped in economically relevant sectors. This would lead to a narrowing of the range on $\mathrm{VaR}_\alpha(L_1 + \cdots + L_d)$ in (5) .

Thus, we consider the case that not only the one-dimensional marginal distributions of the risk vector are known, but also that for a class $\mathcal{E}$ of sets $J \subset \{1, \ldots, d\}$, the joint marginal distributions $F_J, J \in \mathcal{E}$ are fixed. In this case, we get the generalized Fréchet class

$$\mathfrak{F}_{\mathcal{E}} = \mathfrak{F}(F_J, J \in \mathcal{E})$$

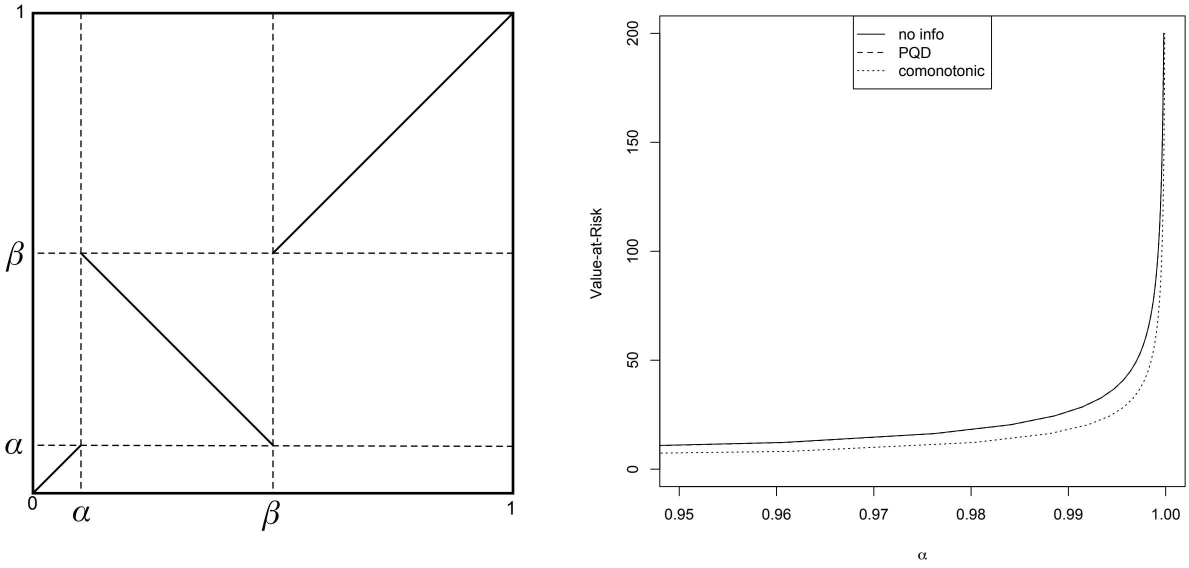

Figure 6: Bivariate copula of the vector $(L_1^*, L_2^*)'$ attaining the worst-VaR bound $M^{-1}(1-\alpha)$ under additional positive dependence restrictions (left). $\overline{\mathrm{VaR}}_\alpha(L^+)$ in the unconstrained case (no info), under additional positive dependence information (PQD), and $\mathrm{VaR}_\alpha^+(L^+)$ (see (8)), for the sum of two Pareto(2) marginals (right). Note that the first two curves are virtually identical.

of all probability measures on $\mathbb{R}^d$ having sub-vector models F_J on $\mathbb{R}^J$, for all $J \in \mathcal{E}$. W.l.o.g. we assume that $\bigcup_{J\in\mathcal{E}} J = \{1,\dots,d\}$. Thus, we have

$$\mathfrak{F}_\mathcal{E} \subset \mathfrak{F}(F_1,\dots,F_d),$$

that is $\mathfrak{F}_\mathcal{E}$ is a sub-class of the class of all possible joint distributions on $(L_1,\dots,L_d)'$. The knowledge of higher dimensional joint distributions is in general not sufficient to determine the joint model of $(L_1,\dots,L_d)'$. Nevertheless, having higher dimensional information restricts the class of possible dependence structures and thus leads to improved upper and lower bounds for the VaR of the joint portfolio.

In practice, loss event datasets often yield some insight into bivariate distributions. Therefore we consider, for d even, a class $\mathcal{E}$ of particular interest in actuarial applications: we set $\mathcal{E} = \{\{2j-1, 2j\} : j = 1,\dots,d/2\}$, defining the Fréchet class

$$\mathfrak{F}_\mathcal{E} = \mathfrak{F}(F_{12}, F_{34},\dots,F_{d-1d}).$$

Hence, in this case, risk estimates on the global position L^+ have to be obtained based on distributional information for all two-dimensional sub-vectors $(L_{2j-1}, L_{2j})'$. Other examples of marginals classes $\mathcal{E}$ have been treated in Puccetti and Rüschendorf (2012b) and Embrechts and Puccetti (2010a).

Our aim is to find bounds for the tail risks

$$\overline{\mathrm{VaR}}_\alpha^\mathcal{E}(L^+) = \sup\{\mathrm{VaR}_\alpha(L_1+\cdots+L_d) : F_L \in \mathfrak{F}_\mathcal{E}\}, \tag{27a}$$

$$\underline{\mathrm{VaR}}_\alpha^\mathcal{E}(L^+) = \inf\{\mathrm{VaR}_\alpha(L_1+\cdots+L_d) : F_L \in \mathfrak{F}_\mathcal{E}\}, \tag{27b}$$

which improve the corresponding bounds $\overline{\mathrm{VaR}}_\alpha(L^+)$ and $\underline{\mathrm{VaR}}_\alpha(L^+)$ defined in (4). If $F_L \in \mathfrak{F}_\mathcal{E}$, we have

$$\underline{\mathrm{VaR}}_\alpha(L^+) \le \underline{\mathrm{VaR}}_\alpha^\mathcal{E}(L^+) \le \mathrm{VaR}_\alpha(L^+) \le \overline{\mathrm{VaR}}_\alpha^\mathcal{E}(L^+) \le \overline{\mathrm{VaR}}_\alpha(L^+). \tag{28}$$

A *reduction method* introduced in Puccetti and Rüschendorf (2012b) allows to find *reduced bounds* $\overline{\mathrm{VaR}}_\alpha^\mathcal{E}(L^+)$ and $\underline{\mathrm{VaR}}_\alpha^\mathcal{E}(L^+)$ using Proposition 4 and the RA introduced in Section 2.2. The reduction method consists of associating to the risk vector $(L_1,\dots,L_d)'$ with $\mathfrak{F}_L \in \mathfrak{F}_\mathcal{E}$ the random vector $(Y_1,\dots,Y_n)'$ defined by

$$Y_j = L_{2j-1} + L_{2j},\ j = 1,\dots,n, \tag{29}$$

where $n = d/2$. If we also denote by H_j the distribution of Y_j, the risk vector $(Y_1,\dots,Y_n)'$ has fixed marginals $H_1,\dots,H_n$. Therefore, it is possible to apply the techniques introduced in Section 2 to compute the *reduced* VaR bounds:

$$\overline{\mathrm{VaR}}_\alpha^r(L^+) = \sup\{\mathrm{VaR}_\alpha(Y_1+\cdots+Y_n) : F_Y \in \mathfrak{F}(H_1,\dots,H_n)\}, \tag{30a}$$

$$\underline{\mathrm{VaR}}_\alpha^r(L^+) = \inf\{\mathrm{VaR}_\alpha(Y_1+\cdots+Y_n) : F_Y \in \mathfrak{F}(H_1,\dots,H_n)\}. \tag{30b}$$

Using the key fact that $L_1+\cdots+L_d = Y_1+\cdots+Y_n$, Proposition 3.3 in Puccetti and Rüschendorf (2012b) states that

$$\overline{\mathrm{VaR}}_\alpha^r(L^+) = \overline{\mathrm{VaR}}_\alpha^\mathcal{E}(L^+) \quad \text{and} \quad \underline{\mathrm{VaR}}_\alpha^r(L^+) = \underline{\mathrm{VaR}}_\alpha^\mathcal{E}(L^+)$$

for the particular class $\mathcal{E}$ introduced above. Therefore, we can rewrite (28) as

$$\underline{\mathrm{VaR}}_\alpha(L^+) \le \underline{\mathrm{VaR}}_\alpha^r(L^+) \le \mathrm{VaR}_\alpha(L^+) \le \overline{\mathrm{VaR}}_\alpha^r(L^+) \le \overline{\mathrm{VaR}}_\alpha(L^+).$$

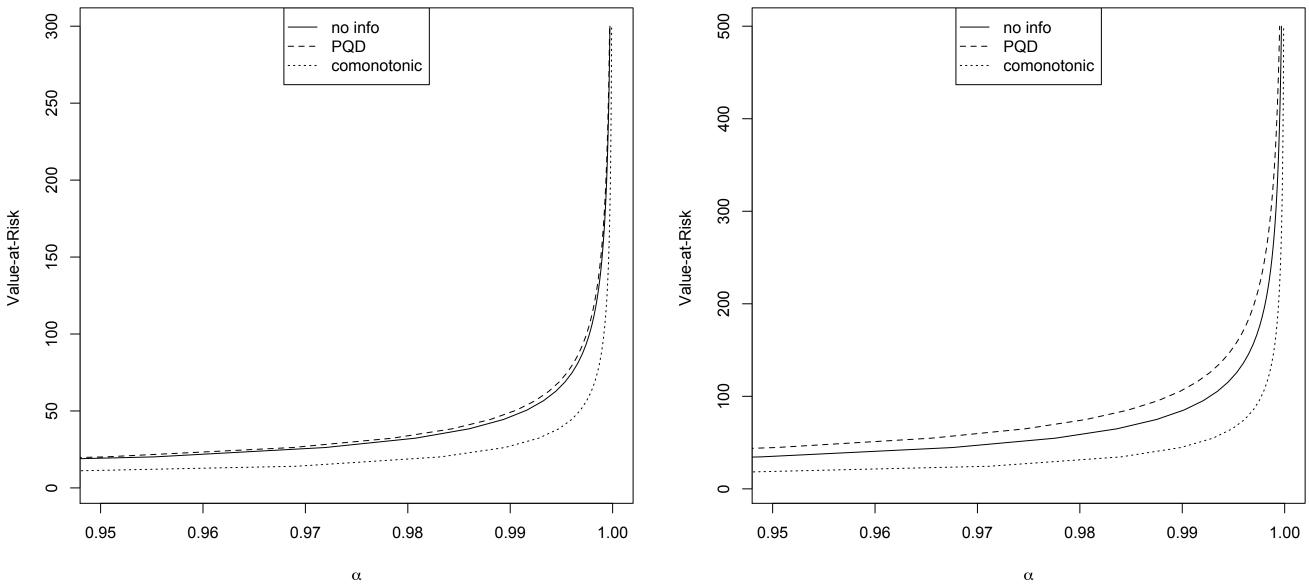

Figure 7: $\overline{\mathrm{VaR}}_\alpha(L^+)$ in the unconstrained case (no info), under additional positive dependence information (PQD), and $\mathrm{VaR}^+_\alpha(L^+)$ (see (8)) for the sum of $d = 3$ (left) and $d = 5$ (right) Pareto(2) marginals.

A reduction method similar to the one described above has also been given in Puccetti and Rüschendorf (2012b) in the case of a general marginal system $\mathcal{E}$. The corresponding reduced bounds $\overline{\mathrm{VaR}}^r_\alpha(L^+)$ and $\underline{\mathrm{VaR}}^r_\alpha(L^+)$ however may fail to be sharp.

We illustrate how to calculate the bounds in (30) in three examples. First we assume the bivariate distributions $F_{2j-1,2j}$, $1 \leq j \leq n$ to be identical and generated by coupling two Pareto marginals having tail parameter $\theta > 0$ by a Pareto copula with parameter $\gamma \neq 0$. The bivariate Pareto copula with parameter $\gamma > 0$ is given by

$$C^{Pa}_\gamma(u, v) = ((1-u)^{-1/\gamma} + (1-v)^{-1/\gamma} - 1)^{-\gamma} + u + v - 1.$$

Under these assumption, the bivariate distribution function F_{12} is given by

$$F_{12}(x_1, x_2) = 1+((1+x_1)^{\theta/\gamma}+(1+x_2)^{\theta/\gamma}-1)^{-\gamma}-(1+x_1)^{-\theta}-(1+x_2)^{-\theta}, \tag{31}$$

while the $n = d/2$ random variables Y_j defined in (29) are identically distributed as

$$H(x) = P\left(Y_j \leq x\right) = P\left(Y_1 \leq x\right) = P\left(L_1 + L_2 \leq x\right), j = 2, \dots, n.$$

Here, we have that

$$H(x) = \int_0^x F_{2|x_1}(x - x_1)\, dF_1(x_1), \tag{32}$$

where we denote by $F_{2|x_1}$ the conditional distribution of $(L_2|L_1 = x_1)$. For this example, the conditional distribution $F_{2|x_1}$ is available in closed form and

$$F_{2|x_1}(x) = 1 - (1 + x_1)^{\theta/\gamma+\theta}\left((1 + x)^{(\theta/\gamma)} + (1 + x_1)^{(\theta/\gamma)} - 1\right)^{-\gamma-1}.$$

Since the risk vector $(Y_1, \dots, Y_n)'$ is homogeneous, we can apply the dual bound methodology introduced in Proposition 4 to compute $\overline{\mathrm{VaR}}^r_\alpha(L^+)$ via (30a). In Proposition 4, we simply use $n = d/2$ (the number of the Y_r's) instead of d and set $F = H$.

In Figure 8, we plot the unconstrained sharp VaR bound $\overline{\mathrm{VaR}}_\alpha(L^+)$ and the reduced bound $\overline{\mathrm{VaR}}^r_\alpha(L^+)$ for a random vector of $d = 600$ Pareto(2)-distributed risks under the marginal system described above. In the left figure the parameter of the Pareto copula is set to $\gamma = 1.5$. This implies a strong positive dependence between consecutive marginals. In the right figure we assume instead that the marginals are pairwise independent. The higher dimensional information reduces the conservative estimate of VaR in both cases, the larger reduction occurring in the case of the bivariate independence constraints. Recall that the calculation of the bound $\overline{\mathrm{VaR}}^r_\alpha(L^+)$ in a homogeneous setting is independent of the dimensionality n of the risk vector $(Y_1, \dots, Y_n)'$, confirming that the dual bound methodology is very effective for homogeneous settings. In Table 5 we compare the estimates for $\mathrm{VaR}_\alpha(L_1 + \cdots + L_d)$ in the case of a homogeneous portfolio of Pareto(2) marginals and under different dependence scenarios.

In order to compute the improved bounds in (30) for *inhomogeneous* portfolios, one has to rely on the RA. We assume to have a portfolio of $d = 2n^2$ Pareto distributed risks, divided into n sub-groups of $2n$ risks. Risks within the same sub-group are assumed to be homogeneous, but risks in different sub-groups may have a different Pareto tail parameter. Within the i-th group, $1 \leq i \leq n$, we assume that each risk is Pareto(θ_i)-distributed and that the bivariate distributions $F_{2j-1,2j}$, $1 \leq j \leq n$ are of the form (31). A vector $\boldsymbol{\theta} = (\theta_1, \dots, \theta_n)'$ then gives

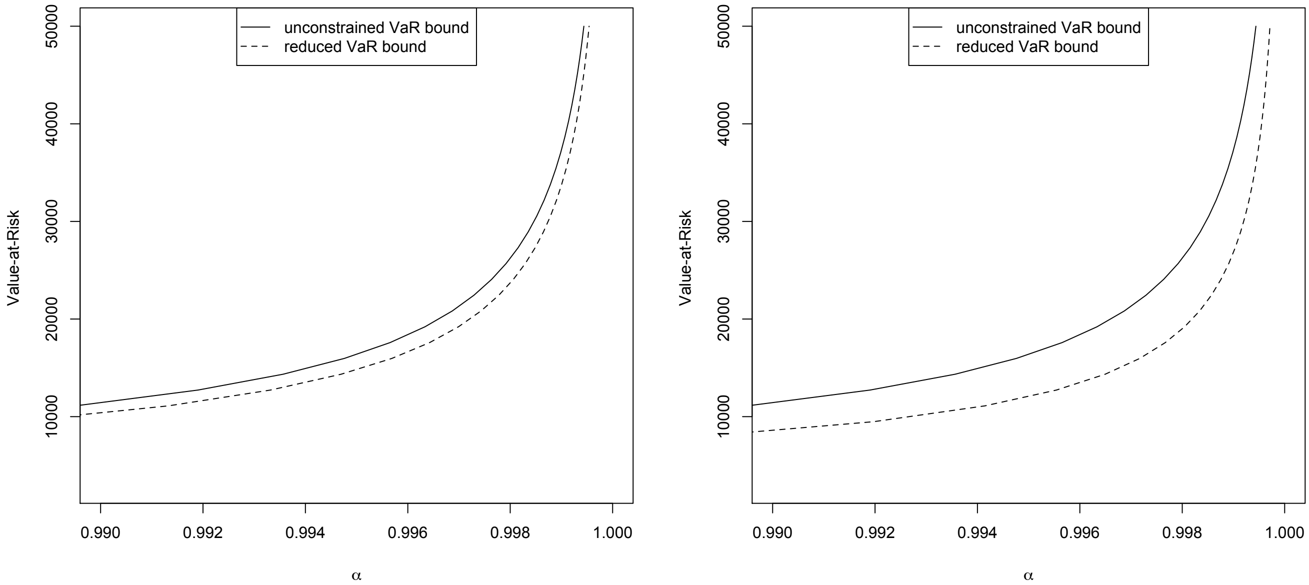

Figure 8: $\overline{\mathrm{VaR}}_\alpha(L^+)$ (see (5)) and $\overline{\mathrm{VaR}}^r_\alpha(L^+)$ (see (30a)) for a random vector of $d = 600$ Pareto(2)-distributed risks with fixed bivariate marginals $F_{2j-1,2j}$ generated by a Pareto copula with $\gamma = 1.5$ (left) and by the independence copula (right).

α	$\mathrm{VaR}^+_\alpha(L^+)$	$\overline{\mathrm{VaR}}^r_\alpha(L^+)$, (A)	$\overline{\mathrm{VaR}}^r_\alpha(L^+)$, (B)	$\overline{\mathrm{VaR}}_\alpha(L^+)$
0.99	5400.00	8496.13	10309.14	11390.00
0.995	7885.28	12015.04	14788.71	16356.42
0.999	18373.67	26832.20	33710.30	37315.70

Table 5: Estimates for $\mathrm{VaR}_\alpha(L^+)$ for a random vector of $d = 600$ Pareto(2)-distributed risks under different dependence scenarios: $\mathrm{VaR}^+_\alpha(L^+)$ ($(L_1,\dots,L_{600})'$ has copula $C = M$); $\overline{\mathrm{VaR}}^r_\alpha(L^+)$, (A): the bivariate marginals $F_{2j-1,2j}$ are independent; $\overline{\mathrm{VaR}}^r_\alpha(L^+)$, (B): the bivariate marginals $F_{2j-1,2j}$ have Pareto copula with $\gamma = 1.5$; $\overline{\mathrm{VaR}}_\alpha(L^+)$: no dependence assumptions are made.

a full description of the marginals of the risk portfolio. The copula parameter is set to $\gamma = 1.5$ in each of the sub-groups. In Table 6, we give RA ranges for $\overline{\mathrm{VaR}}^r_\alpha(L^+)$ and $\underline{\mathrm{VaR}}^r_\alpha(L^+)$, as well as for $\overline{\mathrm{VaR}}_\alpha(L^+)$ and $\underline{\mathrm{VaR}}_\alpha(L^+)$ for different values of n, and at the quantile level $\alpha = 0.999$. In Table 6, computation times are indicated for the computation of the reduced bounds $\overline{\mathrm{VaR}}^r_\alpha(L^+)$ and $\underline{\mathrm{VaR}}^r_\alpha(L^+)$. These times are in general larger when compared to the homogeneous case with the number of marginal distributions $d = n^2$. Indeed, in order to apply the RA to the marginals H_j, one has to compute the quantiles of the distribution H in (32) which in general is a more time consuming operation especially considering that one has to handle different tail parameters. If one has an efficient procedure to obtain these latter quantiles, then the RA computation times of the reduced VaR intervals are approximately the same as in the homogeneous case with $d = n^2$ marginal distributions.

4.1. Application to Operational Risk data

As an example we consider again the Operational Risk application introduced in Section 2.3. Thus, we assume to have a portfolio of $d = 8$ GPD distributed risks with the parameters as in Table 2. The random losses are here divided into 4 subgroups of 2 risks. Risks within the same sub-group are assumed to be independent, i.e. we assume that the bivariate distributions $F_{12}, F_{34}, F_{56}, F_{78}$ have copula Π. The subdivisions into subgroups is arbitrary and it is used here just to illustrate the narrowing of the worst-case, best-case VaR range. A related pair-copula construction is given in Hobæk Haff et al. (2010). In Table 7, we give RA ranges for $\overline{\mathrm{VaR}}^r_\alpha(L^+)$ and $\underline{\mathrm{VaR}}^r_\alpha(L^+)$, as well as for $\overline{\mathrm{VaR}}_\alpha(L^+)$ and $\underline{\mathrm{VaR}}_\alpha(L^+)$ for different quantile levels α. In this case, the computation of high quantiles of the convolution of two subexponential distributions is computationally demanding and a single reduced VaR estimate in Table 7 requires 2.5 hrs.

4.2. Conclusions

To summarize, the same techniques introduced in Section 2, where one only knows the marginal distributions of the risk vector $(L_1,\dots,L_d)'$ can be applied to the case where higher dimensional information is available. In order to use the reduction method one only needs to have the conditional distribution function $F_{i|x_1}$ available in closed form, for any $x_1 \in \mathbb{R}$. This conditional distribution is typically available for bivariate models derived from continuous marginals and a continuous copula, but it might be difficult to compute for higher dimensional subgroups of marginals.

n	d	comp. time	$\underline{\mathrm{VaR}}_{0.999}(L^+)$ (RA range)	$\underline{\mathrm{VaR}}^r_{0.999}(L^+)$ (RA range)	$\mathrm{VaR}^+_{0.999}(L^+)$ (exact)	$\overline{\mathrm{VaR}}^r_{0.999}(L^+)$ (RA range)	$\overline{\mathrm{VaR}}_{0.999}(L^+)$ (RA range)
2	8	8 mins	30.47 – 30.62	54.84 – 55.40	158.49	226.09 – 226.10	277.27 – 277.28
5	50	28 mins	30.47 – 30.62	54.87 – 55.44	652.92	1024.35 – 1024.58	1152.64 – 1152.90
18	648	1,8 hrs	339.90 – 339.97	341.09 – 341.22	7373.01	11415.23 – 11446.26	12643.78 – 12678.12

Table 6: Estimates for $\mathrm{VaR}_\alpha(L_1 + \cdots + L_d)$ for random vectors of Pareto-distributed risks with different tail parameters. The vector of tail parameters are $\boldsymbol{\theta} = (2, 3)'$ (first portfolio), $\boldsymbol{\theta} = (2, 2.5, 3, 3.5, 4)'$ (second portfolio) and $\boldsymbol{\theta} = (2, 2.125, \ldots, 4, 4.125)'$ (third portfolio). Under the additional dependence scenario, the bivariate marginals $F_{2j-1,2j}$ of the risk vector have Pareto copula with $\gamma = 1.5$. For the computation of each reduced bound we set $N = 5.0e04$, $\epsilon = 10^{-3}$.

α	$\underline{\mathrm{VaR}}_\alpha(L^+)$	$\underline{\mathrm{VaR}}^r_\alpha(L^+)$	$\mathrm{VaR}^+_\alpha(L^+)$	$\mathrm{VaR}_\alpha(L^{\Pi,+})$	$\overline{\mathrm{VaR}}^r_\alpha(L^+)$	$\overline{\mathrm{VaR}}_\alpha(L^+)$
0.99	1.78×10^5	2.26×10^5	5.14×10^5	7.08×10^5	2.06×10^6	2.56×10^6
0.995	4.68×10^5	5.36×10^5	1.22×10^6	1.68×10^6	4.82×10^6	5.96×10^6
0.999	4.38×10^6	4.72×10^6	9.33×10^6	1.28×10^7	3.56×10^7	4.34×10^7

Table 7: Estimates for $\mathrm{VaR}_\alpha(L^+)$ for a random vector of $d = 8$ GPD-distributed risks having the parameters in Table 2 and different dependence assumptions, i.e. (from left to right) best-case dependence, best-case under additional information, comonotonicity, independence, worst-case under additional information, worst-case dependence. Under the additional dependence scenarios, the random losses $L_{2j-1,2j}$ of the risk vector are assumed to be independent. Each estimate of $\underline{\mathrm{VaR}}^r_\alpha(L^+)$ and $\overline{\mathrm{VaR}}^r_\alpha(L^+)$ in this table has been obtained in 2.5 hours via the RA by setting $N = 10^5$ and $\epsilon = 10^{-1}$.

Worst-case dependence structures for the problems (27) are in general not available. However, some approximation results given in Embrechts and Puccetti (2010a, Section 5) indicate that they still contain a completely mixable component.

Our final message here is that additional constraints on the risk vector $(L_1, \ldots, L_d)'$ like positive or higher dimensional information knowledge added on top of the knowledge of the marginals will not help to avoid completely mixable dependence structures like the one illustrated in Figures 1, 3 and Table 1. Completely mixable dependence structures will always arise from (un)constrained optimisation problems having VaR as objective function.

Acknowledgments

The authours thank Oliver Chen for suggesting to use a termination condition for the rearrangement algorithm based on the accuracy of the final estimate. They are also grateful to Michael Fackler, Marius Hofert, Steven Vanduffel, Kjersti Aas and an anonymous referee for providing relevant comments on the paper. Part of this research was done at RiskLab, ETH Zurich. Giovanni Puccetti and Ludger Rüschendorf thank RiskLab for the financial support. Paul Embrechts acknowledges the Swiss Finance Institute for financial support.

References

Arbenz, P., C. Hummel, and G. Mainik (2012). Copula based hierarchical risk aggregation through sample reordering. *Insur. Math. Econ. 51*(1), 122 – 133.

Basel Committee on Banking Supervision (2010). *Developments in modelling risk aggregation.* Basel: Bank for International Settlements.

Basel Committee on Banking Supervision (2012). *Consultative Document May 2012. Fundamental review of the trading book.* Basel: Bank for International Settlements.

Böcker, K. and C. Klüppelberg (2010). Multivariate models for operational risk. *Quant. Finance 10*(8), 855–869.

Bolancé, C., M. Guillén, J. Gustafsson, and J. P. Nielson (2012). *Quantitative Operational Risk Models*. Boca Raton, FL: CRC Press.

Cope, E. and G. Antonini (2008). Observed correlations and dependencies among operational losses in the ORX consortium database. *J. Operational Risk 3*(4), 47–74.

Cope, E., G. Mignola, G. Antonini, and R. Ugoccioni (2009). Challenges and pitfalls in measuring operational risk from loss data. *J. Operational Risk 4*(4), 3–27.

de Fontnouvelle, P., V. DeJesus-Rueff, J. Jordan, and E. Rosengren (2005). Capital and risk: new evidence on implications of large operational losses. Working Paper 03-5, Federal Reserve Bank of Boston.

De Vylder, F. (1996). *Advanced Risk Theory*. Bruxelles: Edition de l'Université de Bruxelles.

Dhaene, J., M. Denuit, M. J. Goovaerts, R. Kaas, and D. Vyncke (2002). The concept of comonotonicity in actuarial science and finance: theory. *Insurance Math. Econom. 31*(1), 3–33.

Dhaene, J., S. Vanduffel, M. J. Goovaerts, R. Kaas, Q. Tang, and D. Vyncke (2006). Risk measures and comonotonicity: a review. *Stoch. Models 22*(4), 573–606.

Durante, F. and C. Sempi (2010). Copula theory: an introduction. In *Copula Theory and Its Applications*, Volume 198 of *Lecture Notes in Statistics*, pp. 3–31. Springer Berlin Heidelberg.

Embrechts, P., A. Höing, and A. Juri (2003). Using copulae to bound the Value-at-Risk for functions of dependent risks. *Finance Stoch. 7*(2), 145–167.

Embrechts, P., A. Höing, and G. Puccetti (2005). Worst VaR scenarios. *Insurance Math. Econom. 37*(1), 115–134.

Embrechts, P., C. Klüppelberg, and T. Mikosch (1997). *Modelling Extremal Events for Insurance and Finance*. Berlin: Springer.

Embrechts, P., A. J. McNeil, and D. Straumann (2002). Correlation and dependence in risk management: properties and pitfalls. In *Risk management: value at risk and beyond*, pp. 176–223. Cambridge: Cambridge Univ. Press.

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

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

Embrechts, P. and G. Puccetti (2010a). Bounds for the sum of dependent risks having overlapping marginals. *J. Multivariate Anal. 101*(1), 177 – 190.

Embrechts, P. and G. Puccetti (2010b). Risk aggregation. In *Copula Theory and Its Applications*, Volume 198 of *Lecture Notes in Statistics*, pp. 111–126. Springer Berlin Heidelberg.

Frachot, A., O. Moudoulaud, and T. Roncalli (2004). Loss distribution approach in practice. In M. K. Ong (Ed.), *The Basel Handbook: A Guide for Financial Practictioneers*, pp. 369–398. London: Risk Books.

Hobæk Haff, I., K. Aas, and A. Frigessi (2010). On the simplified pair-copula construction—simply useful or too simplistic? *J. Multivariate Anal. 101*(5), 1296–1310.

Hürlimann, W. (2008a). Extremal moment methods and stochastic orders: application in actuarial science. Chapters I, II and III. *Bol. Asoc. Mat. Venez. 15*(1), 5–110.

Hürlimann, W. (2008b). Extremal moment methods and stochastic orders: application in actuarial science. Chapters IV, V and VI. *Bol. Asoc. Mat. Venez. 15*(2), 153–301.

Jorion, P. (2006). *Value at Risk: The New Benchmark for Managing Financial Risk*. McGraw-Hill. Third edition.

Mainik, G., , and P. Embrechts (2013). Diversification in heavy-tailed portfolios: properties and pitfalls. *Annals Actuarial Science 7*(1), 26–45.

Mainik, G. and L. Rüschendorf (2010). On optimal portfolio diversification with respect to extreme risks. *Finance Stoch. 14*(4), 593–623.

Mainik, G. and L. Rüschendorf (2012). Ordering of multivariate risk models with respect to extreme portfolio losses. *Statistics & Risk Modeling 29*(1), 73–105.

McNeil, A. J., R. Frey, and P. Embrechts (2005). *Quantitative Risk Management: Concepts, Techiniques, Tools*. Princeton, NJ: Princeton University Press.

Mesfioui, M. and J.-F. Quessy (2005). Bounds on the value-at-risk for the sum of possibly dependent risks. *Insurance Math. Econom. 37*, 135–151.

Moscadelli, M. (2004). The modelling of operational risk: experience with the analysis of the data collected by the Basel Committee. Temi di discussione, Banca d'Italia.

Nelsen, R. B. (2006). *An Introduction to Copulas*. New York: Springer. Second edition.

Panjer, H. H. (2006). *Operational risk: modeling analytics*. Hoboken, NJ: Wiley-Interscience.

Pérignon, C. and D. R. Smith (2010). The level and quality of Value-at-Risk disclosure by commercial banks. *J. Banking Finance 34*(2), 362 – 377.

Puccetti, G. and L. Rüschendorf (2012a). Asymptotic equivalence of conservative VaR- and ES-based capital charges. Preprint, University of Freiburg.

Puccetti, G. and L. Rüschendorf (2012b). Bounds for joint portfolios of dependent risks. *Statistics & Risk Modeling 29*(2), 107–131.

Puccetti, G. and L. Rüschendorf (2012c). Computation of sharp bounds on the distribution of a function of dependent risks. *J. Comput. Appl. Math. 236*(7), 1833–1840.

Puccetti, G. and L. Rüschendorf (2013). Sharp bounds for sums of dependent risks. *J. Appl. Probab. 50(1),* 42–53.

Puccetti, G., B. Wang, and R. Wang (2012). Advances in complete mixability. *J. Appl. Probab. 49*(2), 430–440.

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

Rüschendorf, L. (1983). Solution of a statistical optimization problem by rearrangement methods. *Metrika 30*, 55–61.

Rüschendorf, L. (2005). Stochastic ordering of risks, influence of dependence and a.s. constructions. In *Advances on Models, Characterizations and Applications*, pp. 19–55. Chapman & Hall/CRC Press.

Rüschendorf, L. and L. Uckelmann (2002). Variance minimization and random variables with constant sum. In *Distributions with given marginals and statistical modelling*, pp. 211–222. Dordrecht: Kluwer Acad. Publ.

Shevchenko, P. V. (2011). *Modelling operational risk using Bayesian inference*. Heidelberg: Springer.

Wang, B. and R. Wang (2011). The complete mixability and convex minimization problems with monotone marginal densities. *J. Multivariate Anal. 102*(10), 1344–1360.

Wang, R., L. Peng, and J. Yang (2013). Bounds for the sum of dependent risks and worst Value-at-Risk with monotone marginal densities. *Finance Stoch. 17*(2), 395–417.