Value-at-Risk bounds
with variance constraints

C. Bernard∗  L. Rüschendorf†  S. Vanduffel‡§

July 10, 2015

Abstract

We study bounds on the Value-at-Risk (VaR) of a portfolio when besides the marginal distributions of the components also its variance is known, a situation that is of considerable interest in risk management. We discuss when the bounds are sharp (attainable) and also point out a new connection between the study of VaR bounds and the convex ordering of aggregate risk. This connection leads to the construction of an algorithm, called Extended Rearrangement Algorithm (ERA), that makes it possible to approximate sharp VaR bounds. We test the stability and the quality of the algorithm in several numerical examples. We apply the results to the case of credit risk portfolio models and verify that adding the variance constraint gives rise to significantly tighter bounds in all situations of interest.

Key-words: Value-at-Risk, Convex order, Comonotonicity, Model risk, Rearrangement algorithm.

∗Carole Bernard, Grenoble Ecole de Management, 12 Rue Pierre Sémard, 38003 Grenoble. (email: carole.bernard@grenoble-em.com).
†Ludger Rüschendorf, University of Freiburg, Eckerstraße 1, 79104 Freiburg, Germany. (email: ruschen@stochastik.uni-freiburg.de).
‡Corresponding author: Steven Vanduffel, Vrije Universiteit Brussel, Pleinlaan 2, 1050 Bruxelles, Belgium. (email: steven.vanduffel@vub.ac.be).
§Carole Bernard Bernard gratefully acknowledges support from NSERC, from the CAE research grant in Waterloo, from the project on Systemic Risk funded by the GRI in financial services, the Louis Bachelier Institute and from the Humboldt Research Foundation. S. Vanduffel acknowledges financial support of the BNP Paribas Fortis Chair in Banking and the Chair of Stewardship of Finance. We thank two reviewers and a co-editor for many useful suggestions. The paper was first presented at the International Cramér Symposium on Insurance Mathematics in Stockholm (June 2013).
1 Introduction

In this paper we study bounds on the Value-at-Risk (VaR) of sums of risks with known marginal distributions (describing the stand-alone risks) under the additional constraint that a bound on the variance of the sum is known. The variance of the sum partially describes the dependence among the risks, as it involves their correlations. This setting is of significant interest as in many practical situations it corresponds closely to the maximum information at hand when assessing the VaR of a portfolio. For example, in the context of credit risk portfolio models, one typically has knowledge regarding the marginal risks (through so-called PD, EAD and LGD models), and the variance of the aggregate risk (sum of the individual losses) is also often available, as obtained from default correlation models or through statistical analysis of observed credit losses. The same setting also appears in the context of risk aggregation and solvency calculations (Basel III and Solvency II). Indeed, banks and insurers typically have models to estimate risk distributions and VaR per risk type (credit risk, market risk, operational risk, etc.) and per business line, after which they rely on a correlation matrix to obtain the VaR of the aggregated portfolio. Taking this information into account, we assess in this paper the extent of the model risk that remains in the computation of VaR.

Note that in this paper we do not enter into the debate on the merits and drawbacks of VaR. Rather, our motivation is practical and driven by the observation that, subsequent to its inception in the early nineties, VaR rapidly became the preferred reference measure for risk quantification and (regulatory) solvency assessment for banks and insurers. Its main challenger is Tail-Value-at-Risk (TVaR), also known as Expected Shortfall (ES), which addresses several of the shortcomings of VaR but which has its own deficiencies. Interestingly, in the presence of uncertainty the worst possible VaR and the worst possible TVaR of a given portfolio may coincide, a fact that hints at a similarity between the two measures and puts the discussion of the appropriateness of either risk measure in new perspective.

VaR bounds in the unconstrained case (i.e., in which marginal distributions are fixed) have been studied by Rüschendorf (1982) (in the case of two risks) and, more recently, by Denuit et al. (1999), Puccetti and Rüschendorf (2012a, 2013), and Embrechts et al. (2013). In particular, there exist several explicit results on VaR bounds when the marginals are distributed identically. In the inhomogeneous case, however, the analysis rapidly becomes more involved and explicit results are scarce. Puccetti and Rüschendorf (2012a) propose the rearrangement algorithm (RA) as a practical way to approximate VaR bounds of a sum of individual risks when the marginal distributions are known; see also Embrechts et al. (2013). So far, numerical experiments have shown that the RA presents very good accuracy. However, the gap between upper and lower VaR bounds is wide, a feature that can only be explained by the non-use of dependence information.

The problem of finding VaR bounds in the presence of partial dependence information appears to be challenging, as there are few theoretical results available in the literature; see, however, Rüschendorf (1991), Embrechts and Puccetti (2009) and Embrechts et al. (2013) for results when some of the higher
dimensional distributions are known. However, the bounds that are proposed in these papers are often hard to compute numerically, and in practical situations higher dimensional distributions are typically not known. By contrast, the variance of the entire portfolio sum can often be estimated statistically with a sufficient degree of accuracy, or its value can be implied by the availability of the correlations among risks. Intuitively, as the variance measures the average spread of the aggregate portfolio loss around the mean, one could expect that knowledge of this feature would have a significant impact on the maximum possible VaR.

Hence, in this paper we study bounds on the VaR of a sum with fixed marginals and known upper bound on the variance. This setting is related to, but different from, the setting that is assumed in the literature on so-called moment bounds for VaR (as well as related risk characteristics such as survival probabilities and stop-loss premiums). In this stream of the literature, the bounds are derived under the assumption that some of the moments of the portfolio sum are known; see Kaas and Goovaerts (1986a, 1986b), H{"u}rlimann (1998, 2002) and De Schepper and Heijnen (2010) for treatments in the context of actuarial science and Grundy (1991), De Schepper and Heijnen (2007) and Lo (1987) for related results in finance. The main difference between our approach and that of these papers is thus that we assume that the marginal distributions as primary source of portfolio information are available, whereas this information is not used in the literature on moment bounds. It is clear that, as compared to assuming knowledge of the marginal means or variances, knowledge of the marginal distributions makes it possible to improve the VaR bounds. However, when supplementing the information on the marginal distributions with information on the portfolio variance (as a source of dependence information), it turns out that, when the variance constraint is “low” enough, our analytical bounds coincide with moment bounds. As we also show that in this case our analytical bounds are nearly sharp for large portfolios, it follows that the (older) literature on moment bounds for risk characteristics is of interest in this case. However, note that in the context of smaller portfolios (or when the portfolio depicts significant concentration) one cannot expect the bounds to be sharp. To accommodate this situation we develop an algorithm that makes it possible to approximate sharp bounds. As a by-product, this algorithm also sheds light on the composition of extreme portfolios.

In Section 2 we provide simple upper and lower VaR bounds in terms of the Tail Value-at-Risk (TVaR) when only marginal information is available (unconstrained bounds). These unconstrained bounds are also valid for the case of heterogeneous portfolios. We find that the upper and lower VaR bounds are sharp when it is possible to construct random variables that are mixing (i.e., when their sum is constant) on the upper and lower parts of the distribution, respectively. We also show that approximate sharpness of the unconstrained bounds can be expected for large portfolios (asymptotic sharpness).

In this section we also establish a new connection between VaR bounds and results on convex ordering. This connection leads to further improved bounds in certain cases. In this regard, we point out that convex order has been used in the literature as a tool to approximate various risk characteristics of portfolios. Primarily, in the context of lognormal risks, Kaas et al. (2000) show
that portfolio sums can be closely “approximated” by taking the appropriate conditional expectation. The sum obtained by conditioning is smaller than the original one in the sense of convex order, and may be easier to deal with in the sense that closed form expressions for the risk measure at hand (VaR, TVaR, stop-loss premium, etc.) can often be obtained. For applications of this idea in various contexts, such as the modeling of annuities, option pricing and portfolio selection problems, see Curran (1994), Darkiewicz et al. (2009), Vanduffel et al. (2008) and Dhaene et al. (2005), among others. For extensions to situations that go beyond risks that are lognormally distributed, see Valdez et al. (2009) and Landsman (2011). However, in this stream of the literature one merely replaces the original sum by another one that is easier to deal with (in certain contexts), and one uses this new sum to approximate the VaR or TVaR of the original portfolio sum. The results on convex order then make it possible to conclude that the true TVaR or stop-loss premium is underestimated, but nothing meaningful can be said on the relative degree of under or overestimation of the VaR. By contrast, in this paper we show that the problem of finding convex lower bounds for sums (with given marginal distributions of its components) is intimately connected to the problem of VaR bounds.

In Section 3 we consider an additional constraint on the variance of the joint portfolio and provide analytical bounds in this case. We show that these (constrained) VaR bounds can be significantly tighter than the bounds obtained in the unconstrained case. To obtain sharp VaR bounds for the constrained case, we show that it is necessary to make the sum as flat as possible in the upper as well as in the lower part while considering at the same time the variance constraint.

This basic insight yields the intuition to develop, in Section 4, a new algorithm to determine approximate sharp bounds in the constrained case. This algorithm extends the rearrangement algorithm, which was proposed by Embrechts et al. (2013) to approximate sharp VaR bounds in the unconstrained case. We simultaneously rearrange the upper and lower parts of the distribution of the sum and move systematically through the domain of the random sum in order to fulfill the variance constraint. A series of examples shows that the extended rearrangement algorithm (ERA) works well and that the additional variance constraint typically leads to improved bounds.

Finally, in Section 5 we apply the results to the case of a credit risk portfolio. We show that the variance constraint gives rise to VaR bounds that significantly improve upon the unconstrained ones. We also provide evidence that models that are used in the industry and regulatory frameworks may underestimate risk, and believe that some caution regarding the use of portfolio VaRs that are computed at high confidence levels (e.g., 99.5%) as the basis for setting capital requirements is justified. Final remarks are presented in Section 6. Most of the proofs are relegated to the Appendix.

1.1 Problem description

Consider a portfolio containing \( n \) individual risks \( X_j \) (\( j = 1, \ldots, n \)) with finite mean and variance. Assume that the marginal distributions \( F_j \) of the \( X_j \) are also given: we write \( X_j \sim F_j \). Since the
marginal distributions are fixed, the mean of the aggregate portfolio loss, \( S = X_1 + X_2 + \cdots + X_n \), is known, and we denote
\[
\mu := E(X_1 + X_2 + \cdots + X_n).
\] (1.1)

In the main part of this paper we derive bounds on the Value-at-Risk of the sum \( S \) when its variance is required to stay below some level \( s^2 \). The dependence between the different \( X_i \) that attain the VaR bounds in the absence of a variance constraint may indeed give rise to a too high variance of the sum (as compared to the observed variance of the portfolio). This feature will later be confirmed with several examples.

Let us denote the VaR of the portfolio sum \( S \) at confidence level \( q \in (0,1) \) by \( \text{VaR}_q(S) \),
\[
\text{VaR}_q(S) = \inf \{ x \in \mathbb{R} \mid F_S(x) \geq q \}, \quad q \in (0,1),
\]
where \( F_S(x) \) is the distribution function of \( S \). The VaR is thus defined as the left inverse of the distribution function, and we may also write that \( \text{VaR}_q(S) = F_S^{-1}(q) \). Similarly, we define the upper VaR as an upper \( q \)-quantile, i.e.,
\[
\text{VaR}_q^+(S) = \sup \{ x \in \mathbb{R} \mid F_S(x) \leq q \}.
\]

In this paper we are interested in finding the minimum possible VaR of the sum
\[
m(s^2) = \inf \text{VaR}_q(S)
\]
subject to \( X_j \sim F_j, \operatorname{var}(S) \leq s^2 \) (1.2)

and the maximum possible VaR (in terms of the upper quantile)
\[
M(s^2) = \sup \text{VaR}_q^+(S)
\]
subject to \( X_j \sim F_j, \operatorname{var}(S) \leq s^2 \). (1.3)

Note that we allow for \( s^2 = \infty \) to include the absence of variance constraint. In Section 2 we consider the unconstrained case (thus without the constraint on the variance) and derive an upper bound for \( M(\infty) \) and a lower bound for \( m(\infty) \). We also discuss conditions that ensure sharpness of these bounds. In Section 3 we derive bounds in the constrained case and describe when they are sharp. In Section 4 we introduce an algorithm, called ERA, that makes it possible to obtain approximations to sharp bounds in all circumstances. The VaR bounds that we propose are based on two other classic risk measures, namely Tail Value-at-Risk (TVaR) and Left Tail Value-at-Risk (LTVaR). Formally, for \( q \in (0,1) \) we denote by \( \text{TVaR}_q(X_i) \) the Tail Value-at-Risk at level \( q \),
\[
\text{TVaR}_q(X_i) = \frac{1}{1-q} \int_q^1 \text{VaR}_u^+(X_i) du,
\]
and by $\text{LTVaR}_q(X_i)$ the left Tail Value-at-Risk,

$$\text{LTVaR}_q(X_i) = \frac{1}{q} \int_0^q \text{VaR}_u(X_i) du.$$ 

So, $\text{TVaR}_q$ is the average of all upper VaRs from level $q$ onward. Likewise, $\text{LTVaR}_q$ is the average of all lower VaRs. 

## 2 Portfolio VaR bounds with fixed marginal distributions

In this section we derive explicit analytical VaR bounds for the portfolio sum and discuss their sharpness properties. We also establish the close connection between the study of VaR bounds and the so-called convex order. This connection makes it possible to obtain improved bounds in certain cases (i.e., improving on $A$ and $B$, respectively). Moreover, it also yields the proper intuition for constructing an algorithm that is useful for approximating sharp VaR bounds. All bounds in this section require finite first moments of the risks $X_i$ but not necessarily finite second moments.

### 2.1 Unconstrained bounds

We are interested in finding the dependence among the risks $X_i$ that maximizes the VaR of the portfolio sum $S = X_1 + X_2 + \cdots + X_n$. It appears intuitive (but incorrect, as we clarify below) that the maximum will be obtained when the risks $X_i$ have maximum correlation, a situation that occurs when all risks $X_i$ are increasing in each other, i.e., when they are comonotonic. In this case, we can express the risks as $X_i = F_i^{-1}(U)$ ($1 \leq i \leq n$), where $U$ is a standard uniformly distributed random variable. When the risks $X_i$ are comonotonic we denote them by $X^c_i$; $(X^c_1, \ldots, X^c_n)$ is called a comonotonic vector and $S^c = \sum_{i=1}^n X^c_i$ is the corresponding comonotonic portfolio sum. To show that the VaR is not necessarily maximum in the comonotonic case (maximum dependence), let us note that, per the definition of $\text{TVaR}_q$, we have the inequality

$$\text{VaR}^+_q(S) \leq \text{TVaR}_q(S).$$

Since $\text{TVaR}_q$ is well-known to be a subadditive risk measure and is clearly additive for comonotonic risks, we find that

$$\text{TVaR}_q(S) \leq B := \text{TVaR}_q(S^c) = \sum_{i=1}^n \text{TVaR}_q(X_i).$$

Thus, for every portfolio sum $S = X_1 + X_2 + \cdots + X_n$, we find that $\text{VaR}^+_q(S)$ and $\text{TVaR}_q(S)$ are each bounded by $B$. In Figure 2.1, we provide a geometric interpretation of the upper bound $B$.

---

1. Note that in the expressions for TVaR and LTVaR one can also integrate (without impact on the result) over $\text{VaR}_u(X)$ and $\text{VaR}^+_u(X)$, respectively.
Figure 2.1 The graph depicts VaRs of the comonotonic portfolio sum $S^c = X_1^c + X_2^c + \cdots + X_n^c$ as a function of the level $u \in (0, 1)$. We present $\text{VaR}_q(S^c)$, $A := \text{LTVaR}_q(S^c)$ and $B := \text{TVaR}_q(S^c)$.

The bound $B$, as an average of the upper VaRs (from level $q$ onward) in the comonotonic case, clearly dominates the $\text{VaR}_q^+$ of the comonotonic sum. It is then not possible to find a portfolio $S$ with an (upper) quantile function that is strictly larger than $B$ on the whole interval $[q, 1]$, as this would imply that $\text{TVaR}_q(S)$ would be strictly larger than $B$, which is not possible by (2.2). Hence, $B$ is a sharp bound if and only if the (upper) quantile function of $S$ is taking the constant value $B$ on $[q, 1]$. Similar reasoning shows that $\text{VaR}_q(S)$ is bounded below by $A := \text{LTVaR}_q(S^c)$, and the bound $A$ is sharp (attainable) if and only if the quantile function of $S$ is taking the constant value $A$ on $[0, q)$. The sharpness of the bounds will be discussed further in Section 2.2.

In summary, we obtain the following theorem that provides simple upper and lower bounds for the VaR of a sum of risks with given marginals without considering a variance constraint on the portfolio sum (formally, dealing with Problems (1.3) and (1.4) with $s^2 = \infty$).

**Theorem 2.1** (Unconstrained bounds). Let $q \in (0, 1)$, $X_i \sim F_i$ ($i = 1, 2, \ldots, n$), $S = \sum_{i=1}^{n} X_i$ and $S^c = \sum_{i=1}^{n} X_i^c$. Then,

$$A := \sum_{i=1}^{n} \text{LTVaR}_q(X_i) = \text{LTVaR}_q(S^c) \leq \text{VaR}_q(S) \leq \text{VaR}_q^+(S) \leq B := \sum_{i=1}^{n} \text{TVaR}_q(X_i) = \text{TVaR}_q(S^c).$$

The proof for the lower bound can be found in the Appendix. Note that the upper bound $B$ can essentially be found in the literature (see e.g. Puccetti and Rüschendorf (2012b), Theorem 2.3)) but has not been stated in explicit form. It is important to note that the bounds $A$ and $B$ are given in explicit form and can be computed directly from the marginal distributions. Note that they are also...
valid (and as easy to compute) for heterogeneous portfolios (i.e., when individual risks do not have the same distribution). This feature contrasts with earlier results in the literature (see e.g., Puccetti and Rüschendorf, 2012a, 2014 and Embrechts et al., 2013), in which the bounds are typically more difficult to compute and are not always available when portfolios exhibit heterogeneity.

Remark 2.2.

1. Connection with moment bounds:
   If we replace “the information on marginal distributions” by “the information on the portfolio mean,” we return to the basic moment problem:

   \[ K = \sup \text{VaR}_q^+(S) \]
   \[ \text{subject to } E(S) = \mu. \]

   It is clear that this problem is not well posed when \( S \) is allowed to be unbounded. If we restrict to non-negative variables, we find that \( K = \frac{\mu}{1-q} \), which will typically be worse than the upper bound \( B \) in Theorem 2.1.

2. Worst case dependence:
   The upper bound \( B \) is sharp and attained if and only if the quantile function of \( S \) is taking the constant value \( B \) on \([q, 1]\). Hence, the worst case dependence structure for the VaR depends on the probability level \( q \) and involves some negative dependence in the upper part of the portfolio (to render the quantile function of \( S \) constant). By contrast, for the TVaR (or equivalently, ES), there is no unique worst case dependence structure. One possible TVaR worst case dependence structure – independently of the probability level \( q \) – appears when the risks are comonotonic. However, there are many other worst case dependence structures. For example, any dependence structure among the risks such that the quantile function of the portfolio sum coincides with the quantile function of the comonotonic sum on \([0, q]\) (but not necessarily on \([q, 1]\)) yields the same (maximum) TVaR. In particular, the worst case dependence for \( \text{VaR}_q \) (i.e., when \( \text{VaR}_q \) becomes maximum) is also a worst case structure for TVaR. It is a priori not so clear which worst case dependence structure is the most realistic. For instance, if risks display strong positive tail dependence in the upper tail region, then one could restrict to comonotonic dependence. However, in many applications this feature is difficult to confirm or does not readily hold. Note that similar conclusions regarding worst case dependence can be drawn for the lower bound \( A \).

3. Tail risk probability and expected shortfall bounds:
   By inverting the VaR-bounds in Theorem 2.1 we obtain, as a consequence, also tail risk bounds, i.e., upper and lower bounds for \( P(\sum_{i=1}^{n} X_i \geq t) \) for \( t \in \mathbb{R} \). Note that by construction the upper and lower bounds are also valid for TVaR\(_q\)(\(S\)).

---

\(^2\)The same observation holds for a stop-loss premium of a portfolio that becomes maximized when the risks are comonotonic, but this solution is not unique (for a given level of retention). Papers that have studied maximization of stop-loss premia in the presence of (partial) information on the marginal distributions but without assuming knowledge of the dependence structure among the risks include Genest et al. (2002) and Lo (1987). The techniques employed in the present paper make it possible to improve these stop-loss bounds when the variance of the portfolio sum is given.
In Section 2.2, we elaborate on the sharpness properties of the bounds in detail. Section 2.3 examines bounds for large portfolios. Both sections are somewhat technical and some readers may prefer to skip over them in their first reading of the paper.

### 2.2 Sharpness of unconstrained bounds

In discussing sharpness, it is useful to represent without loss of generality the risks $X_i$ as $X_i = f_i(U)$ ($i = 1, 2, \ldots, n$) for some $U \sim U(0, 1)$ such that $\{S \geq \text{VaR}_q(S)\} = \{U \in [q, 1]\}$, i.e., the upper $q$-part of the distribution of $S$ is given by the upper $q$-part $\{U \geq q\}$ of the random variable $U$; see Rüschendorf (Lemma 1, 1983) and Puccetti and Rüschendorf (2013). As $f_i(U)$ and $F_i^{-1}(U)$ have the same distribution, we say that $f_i$ is a “rearrangement” of $F_i^{-1}$ on $[0, 1]$. Furthermore, if $f_i(V)$ and $F_i^{-1}(V)$ have the same distribution for some random variable $V$ that is uniformly distributed on a subset $T$ of $[0, 1]$, then we say that $f_i$ is a rearrangement of $F_i^{-1}$ on $T$. We formulate the following theorem, the proof of which can be found in the Appendix.

**Theorem 2.3** (Sharpness of the unconstrained bounds). Let $X_i = f_i(U) \sim F_i$, $1 \leq i \leq n$ and let $S = \sum_{i=1}^{n} X_i$ be the portfolio sum. Then:

a) The upper bound $B$ in Theorem 2.1 is attained by $S$ if and only if the two following conditions are satisfied:

1) The $f_i$ are rearrangements of $F_i^{-1}$ on $[q, 1]$, $1 \leq i \leq n$.

2) $X_1, X_2, \ldots, X_n$ are mixing on $\{U \geq q\}$, i.e., $\sum_{i=1}^{n} f_i(u) = c$, almost surely on $[q, 1]$ for some $c \in \mathbb{R}$.

b) The lower bound $A$ in Theorem 2.1 is attained by $S$ if and only if the two following conditions are satisfied:

1) The $f_i$ are rearrangements of $F_i^{-1}$ on $[0, q]$, $1 \leq i \leq n$.

2) $X_1, X_2, \ldots, X_n$ are mixing on $\{U < q\}$, i.e., $\sum_{i=1}^{n} f_i(u) = c$, almost surely on $[0, q]$ for some $c \in \mathbb{R}$.

The mixing properties in Theorem 2.3 may be too strong to achieve, but a weakened form of asymptotic mixability for large sample sizes (formally, for $n \to \infty$) may hold true in great generality. This condition then implies that our unconstrained upper bound is asymptotically sharp. The idea of this result can be found in Theorems 2.3 and 2.5 of Puccetti and Rüschendorf (2012b). In the homogeneous case, an asymptotic sharpness result has been provided in Puccetti and Rüschendorf (2014); for some extensions, see Puccetti, Wang and Wang (2013). The following result states that one can expect approximative sharpness of the unconstrained bound $B$ (for $M(\infty)$) in the case of large portfolios.
Theorem 2.4 (Asymptotic sharpness of unconstrained bound). Let \( X_i \) be integrable random variables, \( X_i \sim F_i \). Assume that there exists \( X_i^n = f_i^n(U) \sim F_i \), \( 1 \leq i \leq n \) such that for some sequences \( \alpha_n \downarrow q \), \( \beta_n \uparrow 1 \), the asymptotic mixability condition (AM) holds

\[
\sum_{i=1}^{n} X_i^n = c_n + R_n \quad \text{on } [\alpha_n, \beta_n],
\]

where \( c_n \) are constants and \( R_n = r_n(U) \) with \( \mathbb{E} r_n(U_{[q,1]}) \xrightarrow{n \to \infty} 0 \), i.e. \( S^n := \sum_{i=1}^{n} X_i^n \) is ‘asymptotically mixing’ on \([q,1]\).

Then

\[
\frac{\text{VaR}_q^+(S^n)}{\text{TVaR}_q(S^n)} \to 1,
\]

i.e., the upper Tail Value-at-Risk bound is asymptotically sharp.

The proof of this result is similar to the proof of the sharpness under the mixing condition in Theorem 2.2 and is therefore omitted. The asymptotic mixability condition (AM) can be verified in particular in the homogeneous case with decreasing densities (Wang and Wang [2011]), but also in more general cases (Puccetti, Wang and Wang [2013]). Note that a similar result for the approximate sharpness of the unconstrained lower bound \( A \) (for \( m(\infty) \)) can also be formulated.

2.3 VaR bounds and convex order

Sharpness of the unconstrained upper and lower bounds requires that the portfolio sum is constant on the upper and lower parts of the distribution, respectively. This feature suggests that there is a connection between studying “the variability” of sums and finding VaR bounds. Specifically, we now describe that the problem of obtaining good VaR bounds is closely related to the problem of determining lower minimal elements with respect to convex order, for the portfolio sum. This connection is useful, as it leads to improved bounds (i.e., improving on \( A \) respectively \( B \)) in certain cases. Moreover, it also yields the intuition to construct an algorithm that makes it possible to approximate sharp VaR bounds.

In this regard, we recall that the convex ordering, \( \leq_{\text{cx}} \), between random variables \( X \) and \( Y \) is defined by

\[
X \leq_{\text{cx}} Y \quad \text{if } E(f(X)) \leq E(f(Y))
\]

for all convex functions \( f(\cdot) \) such that the expectation exists. Note that \( E(X) \leq_{\text{cx}} X \) (Jensen’s inequality; see Rudin [1987]) and that \( X \leq_{\text{cx}} Y \) implies, in particular, that \( Y \) has the same mean as \( X \) but a larger variance. The convex order is thus a device that allows for comparing the variability of random variables. For details of this ordering see Müller and Stoyan [2002] or Denuit et al. [2001]. Furthermore, in what follows we always denote by \( F_i^q \) the distribution of \( F_i \) when restricted to the
upper $q$-part of $F_i$, i.e., formally, $F_i^q$ is the distribution of $F_i^{-1}(U)$ where $U$ is uniformly distributed on $[q, 1]$. The VaR upper bound problem can indeed be equivalently described by restricting to the upper $q$-part of the distributions (see also Puccetti and Rüschendorf (2013)).

The following theorem states that the problem of obtaining a sharp upper bound for the Value-at-Risk of a sum is closely related to determining minimal sums in convex order (for given marginal distributions). More precisely, an improvement of a sum with respect to convex order leads to an improvement of the Value-at-Risk. Thus, sharp upper bounds for the Value-at-Risk of a sum are given by a sum minimal in convex order.

**Theorem 2.5 (VaR bounds and convex order).** Let $F_i^q$ denote the restriction of $F_i$ to the upper $q$-part of $F_i$ (as defined in the above paragraph). Then

$$a) \quad M = \sup_{X_i \sim F_i} \text{VaR}_q^+ \left( \sum_{i=1}^n X_i \right) = \sup_{Y_i^q \sim F_i^q} \text{VaR}_0^+ \left( \sum_{i=1}^n Y_i^q \right).$$

$$b) \quad \text{If } X_i^q, Y_i^q \sim F_i^q \text{ and }$$

$$S^q = \sum_{i=1}^n Y_i^q \leq_{cx} \sum_{i=1}^n X_i^q,$$

then

$$\text{VaR}_0^+ \left( \sum_{i=1}^n X_i^q \right) \leq \text{VaR}_0^+ (S^q) \leq B.$$ (2.5)

For the proof, see the Appendix. As a consequence, sharpness is obtained, in particular, when the distributions are mixing on the upper part. Specifically,

**Corollary 2.6.** Assume that there exist $Y_i^q \sim F_i^q$ with $S^q = \sum_{i=1}^n Y_i^q = c$. Then for all $X_i \sim F_i$ it holds that

$$\text{VaR}_q^+ \left( \sum_{i=1}^n X_i \right) \leq \text{VaR}_0^+ (S^q) = B.$$ (2.6)

Here are some additional observations:

(i) In general, as stated in (2.5), the worst value for $\text{VaR}_q(S)$ is attained for some minimal element in convex order in the class $\mathcal{F}^q$ containing sums of the components of random vectors $(Y_1, \ldots, Y_n)$ such that $Y_i \sim F_i^q$. In dimension $n = 2$, a smallest element in convex order on the upper part exists; it is given by

$$Y_1^q = (F_1)^{-1}(U), \quad Y_2^q = (F_2)^{-1}(1 - U),$$

where $U$ is uniformly distributed on $[q, 1]$. The resulting $\text{VaR}_q^+(Y_1^q + Y_2^q)$ is, by Theorem 2.5, a sharp upper bound and is identical to the solution of this case in Rüschendorf (1982). When
It is typically not possible to find a smallest element in convex order (see, e.g., Bernard et al. (2013) for an example in which it fails to exist). However, under a rather weak notion of positive dependence among the risks (called cumulative positive dependence) a convex minimum is obtained when the risks are independent (see e.g., Denuit et al. (2001)).

(ii) It is well-known that the comonotonic sum is the largest possible element in convex order (Meilijson and Nadas (1979)). Hence, in $\mathcal{F}^q$,

$$S \leq_{cx} S^c \quad (2.7)$$

where $S = \sum_{i=1}^{n} Z_i$, $S^c = \sum_{i=1}^{n} Z^c_i$ and $Z_i \sim F^q_i$. Hence, according to (2.6), the comonotonic dependence does not yield a valid solution of our VaR upper bound but gives just the most optimistic value of all candidates in Theorem 2.5.

(iii) Similarly to Theorem 2.5, we also obtain an improved lower bound for $\text{VaR}_q(\sum_{i=1}^{n} X_i)$. Denote by $F_{i,q}$ the distribution of $F_{i}^{-1}(U)$ where $U$ is uniformly distributed on $(0, q)$, i.e., the $F_{i,q}$ are the restrictions of $F_i$ to the lower $q$-part. Let $X_i, Y_i \sim F_i$ and $X_{i,q}, Y_{i,q} \sim F_{i,q}$. Then, we get:

$$A \leq \text{VaR}_1(S_q) \leq \text{VaR}_q \left( \sum_{i=1}^{n} X_i \right). \quad (2.8)$$

Hence, also in this case one obtains better VaR bounds by decreasing the sum $\sum_{i=1}^{n} X_{i,q}$ in convex order. If the $Y_{i,q} (i = 1, 2, \ldots, n)$ can be chosen mixing on $[0, q]$, then obviously $S_q = \sum_{i=1}^{n} Y_{i,q} \leq_{cx} \sum_{i=1}^{n} X_{i,q}$ and we obtain as a consequence sharp lower bounds (Theorem 2.3).

3 VaR bounds with a variance constraint on the portfolio sum

3.1 Variance-constrained bounds

The bounds $A$ and $B$ have been derived without considering a constraint on the variance of the portfolio and are thus also bounds in the presence of such constraint. Hence, from Theorem 2.4 it follows immediately that $A \leq m \leq M \leq B$, where we write $m = m(s^2)$ and $M = M(s^2)$ for ease of exposition. We now propose to improve the bounds by constraining the variance of the sum $X_1 + X_2 + \cdots + X_n$ to be below a maximum level $s^2$. From the introduction we recall that this setting is highly relevant, as in many practical situations historical data on observed portfolio losses can be used to estimate (a bound on) the variance of the portfolio sum. We first provide an example that provides some intuition as to how we can deal with the constrained problem.
Example 3.1. Let $X_1$ and $X_2$ be both uniformly distributed on the unit interval and assume that $s^2 \geq \frac{3}{16}$ (e.g., $s^2 = \infty$). We are interested in finding VaR bounds for $X_1 + X_2$ with $q = 0.75$. From Theorem 2.1, $A = \frac{2}{3}, B = \frac{7}{4}$. Next, consider $Y_1$ and $Y_2$, uniformly distributed random variables over $(0, 1)$ such that $Y_2 = \frac{3}{4} - Y_1$ if $Y_1 < \frac{3}{4}$ and $Y_2 = \frac{7}{4} - Y_1$ if $Y_1 \geq \frac{3}{4}$. Note that $Y_1 + Y_2 = \frac{3}{4} = A$ for $Y_1 < \frac{3}{4}$ and $Y_1 + Y_2 = \frac{7}{4} = B$ for $Y_1 \geq \frac{3}{4}$. Thus, $\text{VaR}_q(Y_1 + Y_2) = A$ and $\text{VaR}_q^+(Y_1 + Y_2) = B$. Furthermore, $\text{var}(Y_1 + Y_2) = q(A - 1)^2 + (1 - q)(B - 1)^2 = \frac{3}{16} \leq s^2$; thus, the variance constraint is satisfied. Hence, for $s^2 \geq \frac{3}{16}$, $m = A$ and $M = B$ and both $m$ and $M$ are sharp, i.e., attained by $(Y_1, Y_2)$.

This example shows that the unconstrained bounds $A$ and $B$ can also be sharp for the constrained problems (1.3) and (1.4). More surprisingly, the dependence structures to attain the upper and lower bounds are identical. When $s^2 < \frac{3}{16}$, the bounds $A$ and $B$ will not be sharp for the respective problems (1.3) and (1.4).

Inspired by Example 3.1 we define a random variable $X^*$ that takes two possible values, corresponding to the bounds $A$ and $B$ that we derived in Theorem 2.1:

$$X^* = \begin{cases} A \text{ with probability } q, \\ B \text{ with probability } 1 - q. \end{cases} \quad (3.1)$$

Then, the cdf $F$ of $X^*$ verifies $F(x) = 0$ if $x < A$, $F(x) = q$ if $A \leq x < B$ and $F(x) = 1$ if $x \geq B$. Note that $E(X^*) = \mu$, whereas its variance is given as $\text{var}(X^*) = q(A - \mu)^2 + (1 - q)(B - \mu)^2$. This distribution will play a key role in solving the constrained problems (1.3) and (1.4).

In the presence of an additional variance constraint on the portfolio sum, $A$ and $B$ (as in Theorem 2.1) are still bounds for $\text{VaR}_q(X_1 + X_2 + \cdots + X_n)$, and they may still be attained, in which case they are the best possible. For example, assume that the lowest value that $X_1 + X_2 + \cdots + X_n$ takes is $A$ with probability $q$. In this case, $X_1 + X_2 + \cdots + X_n$ has minimum variance if $X_1 + X_2 + \cdots + X_n \equiv_d X^*$ and thus $A$ may be attained depending on the value $s^2$. Similarly, the upper bound for the Value-at-Risk (with confidence $q$) is reached when $B$ is the largest value that $X_1 + X_2 + \cdots + X_n$ can take (with probability $1 - q$). Thus, when the variable $X^*$ satisfies the variance constraint, i.e., $\text{var}(X^*) \leq s^2$, then the bounds $A$ and $B$ cannot be readily improved. However, if $\text{var}(X^*) > s^2$, then $A$ and $B$ are generally too wide, and better bounds can be constructed. It is then intuitive that better bounds can be found by constructing a variable $Y$ taking two values $a$ (larger than $A$) and $b$ (smaller than $B$) with respective probabilities $q$ and $1 - q$ in such a way that the variance constraint of the portfolio sum is satisfied. Note that a two-point distribution taking value $\mu - s \sqrt{\frac{1 - q}{q}}$ with probability $q$ and value $\mu + s \sqrt{\frac{q}{1 - q}}$ with probability $1 - q$ is the only two-point distribution with mean $\mu$ and variance $s^2$.

The following theorem shows that the construction as outlined above gives bounds on Value-at-

\footnote{It is easy to prove that $X^*$ is of minimum variance among all the random variables $Z$ that take the value $A$ with probability $q$ and that satisfy $E(Z) = \mu$.}
Risk in the presence of a variance constraint.

**Theorem 3.2** (constrained bounds). Let $q \in (0, 1)$, $X_i \sim F_i \ (i = 1, 2, \ldots, n)$, and $S = \sum_{i=1}^{n} X_i$ satisfy $\text{var}(S) \leq s^2$. Then, we have

$$a := \max \left( \mu - s \sqrt{\frac{1-q}{q}}, A \right) \leq \mu \leq \text{VaR}_q(S) \leq \text{VaR}_q^+(S) \leq M \leq b := \min \left( \mu + s \sqrt{\frac{q}{1-q}}, B \right).$$

In particular, if $s^2 \geq q(A - \mu)^2 + (1-q)(B - \mu)^2$, then $a = A$ and $b = B$ (and the unconstrained bounds are not improved by the presence of the constraint on variance).

A detailed proof of Theorem 3.2 is provided in the Appendix. The presence of the variance constraint does not always make it possible to strengthen the bounds $A$ and $B$. Indeed, the variable $X^*$ taking the values $A$ and $B$ may also satisfy the variance constraint, i.e., $\text{var}(X^*) \leq s^2$ and, in this case, $a = A$ and $b = B$. Hence, we conclude that if $s^2$ is not too large (i.e., when $s^2 \leq q(A - \mu)^2 + (1-q)(B - \mu)^2$), then the bounds $a$ and $b$ that are obtained for the constrained case strictly outperform the bounds in the unconstrained case.

The question thus arises: What is meant by “too large”? This aspect pertains to the characteristics of the problem and the data at hand. However, a few observations are of interest. When all risks are distributed identically, then the bounds $A$ and $B$ grow linearly with the size of the portfolio; however, as the standard deviation of a portfolio is subadditive, the condition $s^2 \geq q(A - \mu)^2 + (1-q)(B - \mu)^2$ becomes more difficult to satisfy, meaning that it becomes more likely that bounds $a$ and $b$ are better than $A$ and $B$. For example, when the risks (e.g., in a life insurance context) are approximately independent, then bounds $a$ and $b$ will strictly improve upon $A$ and $B$ for moderate portfolio sizes. On the other hand, when all risks are positively and equally correlated, then, when $n$ becomes large, a new risk $X_{n+1}$, which is added to an existing portfolio loss $S$, becomes perfectly correlated with this portfolio, thus resulting in standard deviations that add up, meaning that $a$ and $b$ may be identical to $A$ and $B$ again.

Two observations follow:

(i) When no information on the dependence is available, we are still able to conclude that

$$\mu - \sum_{i=1}^{n} \sigma_i \sqrt{\frac{q}{1-q}} \leq \text{VaR}_q \left( \sum_{i=1}^{n} X_i \right) \leq \text{VaR}_q^+ \left( \sum_{i=1}^{n} X_i \right) \leq \mu + \sum_{i=1}^{n} \sigma_i \sqrt{\frac{q}{1-q}} \quad (3.2)$$

where $\sigma_i^2 = \text{var}(X_i)$, $1 \leq i \leq n$. Indeed, as the standard deviation is a subadditive risk measure, $\sum_{i=1}^{n} \sigma_i$ is an upper bound for the standard deviation of the portfolio so that (3.2) becomes a consequence of Theorem 3.2. Note that, in fact, the unconstrained bound, $B$, is also a bound for the constrained case and improves upon $\mu + \sum_{i=1}^{n} \sigma_i \sqrt{\frac{q}{1-q}}$.

(ii) Sometimes the correlations between some of the risks $X_i$ are known. This partial information can
then be used to provide an upper bound on the variance of the portfolio sum, which could sharpen
the unconstrained Value-at-Risk bounds that we derived in Theorem 3.2. For example, assume that
\[ \text{var}(X_1 + X_2 + \cdots + X_i) \leq s_1^2 \] and
\[ \text{var}(X_{i+1} + X_2 + \cdots + X_n) \leq s_2^2. \] Then, invoking subadditivity of
the standard deviation again, and using a similar rearrangement as for the proof of Theorem 3.2,
\[ \mu - (s_1 + s_2) \sqrt{\frac{q}{1-q}} \leq \text{VaR}_q \left( \sum_{i=1}^{n} X_i \right) \leq \text{VaR}_q^+ \left( \sum_{i=1}^{n} X_i \right) \leq \mu + (s_1 + s_2) \sqrt{\frac{q}{1-q}}. \quad (3.3) \]

Remark 3.3 (connection with moment bounds). If we replace “the information on marginal distribu-
tions” with “the information on the portfolio mean,” we can formulate the following moment problem:

\[ K = \sup \text{VaR}_q^+(S) \quad (3.4) \]
subject to \( E(S) = \mu \) and \( \text{var}(S) \leq s^2 \).

It is not difficult to show that \( K = \mu + s \sqrt{\frac{q}{1-q}} \). In fact, this bound corresponds to the so-called Cantelli bound (see also Barrieu and Scandolo (2015)). Thus, if the variance constraint is not “too
large,” we find that the moment bound \( K \) coincides with the analytical bound \( b \). Note that Hürlimann
(2002) extends these results to the case of bounded variables. Given that the bound \( b \) turns out to
be nearly sharp for sufficiently large portfolios, our paper provides evidence that the literature on
moment bounds is of significant interest when analyzing VaR bounds. For smaller (or concentrated)
portfolios, however, one cannot expect sharpness of the bound \( b \). In Section 4 we develop an algorithm
that makes it possible to approximate the sharp bounds for any size \( n \) and variance bound \( s^2 \).

In the following subsection we discuss sharpness of the analytical bounds and provide a precise
connection with convex order.

3.2 Sharpness and convex order

By Theorem 3.2b), the bounds in the variance-constrained case are given by a two-point distribution
in \( a, b \). Similarly, as in Theorem 2.3 for the unconstrained case, we show that the constrained bounds
are sharp if and only if the risks \( X_1, X_2, \ldots, X_n \) have a mixing property.

Theorem 3.4 (Sharpness of variance-constrained bounds). Let \( X_i = f_i(U) \sim F_i, \ 1 \leq i \leq n \) where \( U \)
is uniformly distributed on \( (0, 1) \), and \( S = \sum_{i=1}^{n} X_i \) satisfies the variance constraint, i.e., \( \text{var}(S) \leq s^2 \).
Without loss of generality, let \( \{ S \geq \text{VaR}_q(S) \} = \{ U \geq q \} \). Then, the upper bound \( b \) in Theorem 3.2
is attained if and only if the lower bound \( a \) is attained and, equivalently, if

\[ S = b \text{ on } \{ U \geq q \} \text{ and } S = a \text{ on } \{ U < q \}, \quad (3.5) \]
i.e., \( S \) is mixing on the upper \( q \)-part \( \{ U \geq q \} \) and on the lower \( q \)-part \( \{ U < q \} \) with mixing constants \( b \) and \( a \).
In general, the bounds proposed in Theorem 3.2 are not sharp. Theorem 3.4, however, suggests how to obtain sharp VaR bounds when there is a constraint on the variance of the sum. Roughly speaking, the outcomes of the variables should be rearranged to produce a dependence between the risks such that the outcomes for the sum are as concentrated as possible around the two values \(a\) and \(b\) that occur with respective probabilities \(1 - q\) and \(q\). This idea is concordant with the aim of finding convex order bounds. Indeed, the improved result in Theorem 2.5 based on convex order that was valid for the unconstrained case extends in a similar way to the variance-constrained case.

Let \(Y_i = f_i(U) \sim F_i, i = 1, 2, \ldots, n\), and let \(S = \sum_{i=1}^{n} Y_i\) with \(\text{var}(S) \leq s^2\) be an admissible solution for the constrained VaR\(_q\) upper bound problem. Let the upper \(q\)-part of the distribution of \(S\), \(\{S \geq \text{VaR}\_q(S)\}\), be, without loss of generality, identical to \(\{U \geq q\}\).

In what follows, for given random variables \(X, Y\) and subset \(T\) of \(\Omega\), \(X|_T \leq_{cx} Y|_T\) denotes that the conditional distribution of the restriction \(X|_T\) is smaller in convex order than \(Y|_T\).

**Theorem 3.5** (Variance-constrained bounds and convex order). If \(X_i \sim F_i\) and \(S_n = \sum_{i=1}^{n} X_i\) satisfy

\[
S_n|_{U \geq q} \leq_{cx} S|_{U \geq q} \quad \text{and} \quad S_n|_{U < q} \leq_{cx} S|_{U < q},
\]  

(3.6)

then \(S_n\) is admissible, \(\text{var}(S_n) \leq s^2\) and \(S_n\) is an improvement of \(S\) in the sense that

\[
\text{VaR}_q^+(S_n) \geq \text{VaR}_q^+(S) \quad \text{and} \quad \text{VaR}_q(S_n) \leq \text{VaR}_q(S).
\]

The proof of Theorem 3.5 is provided in the Appendix. Theorem 3.5 says that in order to obtain sharp upper bounds of \(\text{VaR}_q^+(S)\) and sharp lower bounds for \(\text{VaR}_q(S)\) one should try to rearrange the random variables \(Y_i\) on the upper \(q\)-part \(T = \{U \geq q\}\) of \(S\) and on their lower \(q\)-part \(T^c = \{U < q\}\) such that the distribution is as flat as possible (in convex order) on \(T\) and on \(T^c\). In particular, this holds true for an optimal solution of the variance-constrained problem. The two flattenings can be performed separately or simultaneously.

In the following section, this basic idea is developed into an algorithm that makes it possible to approximate sharp \(\text{VaR}_q\) bounds in the variance-constrained case.

**4 The extended rearrangement algorithm (ERA)**

This section extends the RA for finding approximate sharp bounds on the distribution of a function of \(n\) dependent random variables having fixed marginal distributions. The fundamental idea of the ERA algorithm is based on Section 3.2 and aims at making the distribution of \(S\) as flat as possible on the upper and lower part by applying the RA algorithm on both parts and by moving through the domains in a systematic way in order to satisfy the variance constraint.
The algorithm requires random variables that are discretely distributed, but note that continuous distributions can be approximated to any degree of accuracy by discrete ones. Let \( d \) be the number of points used to discretize the random variables and assume that each risk \( X_j \) \((j = 1, 2, \ldots, n)\) is sampled into \( d \) equiprobable values.

To avoid confusion, when we compute the bounds \( A, B \) and \( a, b \), respectively, for the discretized (sampled) risks, then we use the notations \( A_d, B_d \) and \( a_d, b_d \), respectively. Hence,

\[
B_d = \sum_{j=1}^{n} \frac{1}{d-k} \sum_{i=k+1}^{d} x_{ij}, \quad A_d = \frac{1}{k} \sum_{j=1}^{n} \sum_{i=1}^{d} x_{ij} - \frac{d-k}{k} B_d. \tag{4.1}
\]

Denote by \( \mu_d \) the mean of the sum of the discretized risks. We first describe the RA briefly after which we describe in detail the extended rearrangement algorithm that we propose as a suitable way to compute numerical VaR bounds of portfolios in the presence of a variance constraint. We point out that, unlike the theoretical VaR bounds, the algorithm does not require that the risks have finite mean (unconstrained bounds) or variance (constrained bounds).

### 4.1 The Rearrangement Algorithm (RA)

**Description:** The rearrangement algorithm (RA) can be seen as a method by which to construct dependence between variables \( X_j \) \((j = 1, 2, \ldots, n)\) such that the distribution of \( S \) becomes as small as possible in convex order. For each \( j \), define \( x_{ij} := F_j^{-1}(\frac{i}{d+1}) \) \((i = 1, \ldots, d; j = 1, 2, \ldots, n)\) to obtain a \( d \times n \) matrix \( X = (x_{ij}) \), corresponding to a multivariate vector \((X_1, X_2, \ldots, X_n)\) that is comonotonic. Let us denote the matrix after rearrangement by \( X^* = (x^*_{ij}) \). In order to make the distribution of \( S^* = X^*_1 + X^*_2 + \cdots + X^*_n \) as small as possible in convex order, one needs to rearrange the \( d \) elements \( x_{ij} \) \((i = 1, \ldots, d)\) within each column \( j \) \((j = 1, 2, \ldots, n)\) such that the function \( i \rightarrow \sum_{j=1}^{n} x^*_{ij} \) is “as flat as possible” corresponding to the objective of making the distribution of \( S^* = X^*_1 + X^*_2 + \cdots + X^*_n \) as small as possible in convex order. Note that as the rearrangements are performed only within columns, \( X_1 \) and \( X^*_1 \) will have the same distribution. Of course, when the distribution of \( S \) is the smallest possible in the sense of convex order, it must also have minimum variance. Hence, it must hold that for all \( \ell = 1, 2, \ldots, n \), \( X_\ell \) is countermonotonic with \( \sum_{k=1, k\neq \ell}^{n} X^*_k \) (see Puccetti and Rüschendorf (2012a, Theorem 2.1)); this observation lies at the core of their rearrangement method. The subsequent columns of the matrix are rearranged such that they become countermonotonic with the sum of all other columns until convergence is reached.

**Remark 4.1** (Block rearrangements). If the distribution of \( S^* = X^*_1 + X^*_2 + \cdots + X^*_n \) is minimal in convex order, it must satisfy the following necessary condition: For any decomposition of \( \{1, 2, \ldots, n\} = I_1 \cup I_2 \) into two disjoint sets \( I_1 \) and \( I_2 \), the sums \( \sum_{i \in I_1} X^*_i \) and \( \sum_{i \in I_2} X^*_i \) must be countermonotonic. From this observation, it is possible to improve the algorithm by rearranging “blocks of columns” instead of one column at a time. The idea is simple. Split the number of columns into two disjoint sets.
Then, sort the rows of each set according to their sums (in the first set, one arranges the rows in increasing order with their sum; in the other set, they are arranged in decreasing order with their sum). It is expected that the number of rearrangements needed to approximate sharp bounds will be smaller as the reduction in the variance caused by moving a block is larger than that caused by moving only one column at a time. This improvement can be useful in dealing with large matrices or when attempting to avoid a local minimum such as can be encountered in the situation when all columns are countermonotonic with the sum of the others.

**Application to approximate sharp VaR bounds:** The standard rearrangement algorithm that involves operating on individual columns has been used successfully to compute (approximate) VaR bounds on the sum of \( n \) dependent risks with given marginal distributions. Indeed, to maximize the Value-at-Risk of a sum of dependent risks, Embrechts et al. (2013) applied the RA to the last rows of the matrix (corresponding to the highest values for each risk), accounting for a probability \( 1 - q \).

To fix the ideas, let us assume that there exists \( k \in \{0, 1, \ldots, d\} \) such that

\[
q = \frac{k}{d}.
\]

Embrechts et al. (2013) then apply the RA to a \((d-k) \times n\) sub-matrix consisting of the rows \(k+1, \ldots, d\) of the original matrix \(X\) and account for a probability \(1 - q\). In doing so, the quantile function of \(S^*\) becomes “as flat as possible” on \([q, 1]\) while preserving the distributions of \(X_j\). Then, in the ideal situation, in which the quantile function is constant on \([q, 1]\), one obtains a sum \(S^*\) with a VaR (at confidence level \(q\)) that is equal to the TVaR of the comonotonic sum, and thus it is the maximum possible. It is clear that the minimum VaR can be obtained approximately by applying the RA to a \(k \times n\) sub-matrix consisting of the rows \(1, \ldots, k\) of the original matrix \(X\). It can be easily seen that the RA converges, in general, to a local minimum and maximum, respectively (“a candidate solution”). Indeed, when a column is rearranged, one strictly decreases the variance of the portfolio sum. Then, as long as the variance does not become constant, one can continue rearranging subsequent columns, and each time a new matrix is obtained. As there is a finite number of rearranged matrices, there exists a situation in which all columns exhibit the countermonotonicity property and the algorithm stops. It has been shown that a few iterations of the algorithm lead, in all examples considered so far, to a good approximation of the global solution.

As suggested by Theorem 3.5, for the solution of the constrained VaR bound problem it is necessary to determine minimal elements of the distributions of \(S\) in convex order on the upper and on the lower \(q\)-part while satisfying the variance constraint. The following extension of the RA algorithm is consistent with this idea in that it chooses, in a suitable way, rearrangements that are admissible with respect to the variance constraint and applies the RA algorithm seperately to the corresponding upper and lower \(q\)-parts.
4.2 Extended Rearrangement Algorithm

Based on Theorems 3.4 and 3.5, it is natural to modify the rearrangement algorithm used by Embrechts et al. (2013) and Puccetti and Rüschendorf (2012a) to construct numerically the minimum and maximum bounds. We simultaneously rearrange the upper and the lower part of the distribution of the sum and move through the domains in a systematic way in order to satisfy the variance constraint.

- Step 1: For each $j \in \{1, \ldots, n\}$ and $i \in \{1, \ldots, d\}$, define $x_{ij} := F_j^{-1}(\frac{i}{d+1})$. Let $\mu_d = \frac{1}{d} \sum_{i=1}^{d} \sum_{j=1}^{n} x_{ij}$.

- Step 2: Calculate $A_d$ and $B_d$ as in (1.1) (Theorem 2.1) and $b_d$ (Theorem 3.2) for the sampled variables.

- Step 3: If $q(A_d - \mu_d)^2 + (1 - q)(B_d - \mu_d)^2 \leq s^2$, then go directly to step 4 and the mixing constants are $A_d$ and $B_d$. Otherwise:
  - Step 3a: for all $m = 1, 2, \ldots, k$ compute $b_d(m) := \frac{\sum_{i=k+1}^{d-m} \sum_{j=1}^{n} x_{ij}}{d-k}$.
  - Step 3b: $m^* = \min\{m | b_d(m) \leq b_d\}$.
  - Step 3c: replace $X$ with a new matrix:

$$
M := (m_{ij})_{ij} \leftarrow \begin{pmatrix}
x_{(d-m^*+2)1} & x_{(d-m^*+2)2} & \cdots & x_{(d-m^*+2)n} \\
x_{(d-m^*+3)1} & x_{(d-m^*+3)2} & \cdots & x_{(d-m^*+3)n} \\
\vdots & \vdots & \ddots & \vdots \\
x_{(d-m^*+1)1} & x_{(d-m^*+1)2} & \cdots & x_{(d-m^*+1)n}
\end{pmatrix}.
$$

Note that when $m^* = 1$, then $m_{ij} = x_{ij}$ for all $i = 1, \ldots, d$ and $j = 1, \ldots, n$. Concordant with the calculation of $m^*$, the idea is to consider the lower part of the matrix, i.e., $\{(m_{ij})_{i=(k+1),\ldots,d,j=1,\ldots,n}\}$ in order to approximate the upper VaR bound, whereas the upper part, i.e., $\{(m_{ij})_{i=1,\ldots,k,j=1,\ldots,n}\}$, will be used to approximate the lower VaR bound.

- Step 4: Initiate $r \leftarrow 0$ and $v_{-1} \leftarrow +\infty$.

- Step 5: Apply the RA to $(I) := \{(m_{ij})_{i=1,\ldots,k,j=1,\ldots,n}\}$ and obtain $(x_{ij}^*)_{i=1,\ldots,k,j=1,\ldots,n}$.

\[\text{In this case, we know that } A_d \text{ and } B_d \text{ will also be the constrained bounds in this case and that the variance constraint is satisfied automatically (see the second part of Theorem 3.2).}\]
• Step 6: Apply the RA to \((II) := \{(m_{ij})_{i=(k+1),\ldots,d,j=1,\ldots,n}\}\) and obtain \((x^*_{ij})_{i=(k+1),\ldots,d,j=1,\ldots,n}\).

• Step 7: Compute \(S^*_i = \sum_j x^*_{ij}\) for each \(i = 1,\ldots,d\).

• Step 8: Compute \(v_r := \frac{1}{d} \sum_i (S^*_i)^2 - (\sum_i S^*_i)^2\).

- If \(v_r < s^2\), then the approximate solutions to Problems (1.3) and (1.4) have been found; the lower bound is given by \(m_d := \max_{i \leq k} (x^*_{i1} + \cdots + x^*_{in})\) and the upper bound by \(M_d := \min_{i \geq k+1} (x^*_{i1} + \cdots + x^*_{in})\). The algorithm must stop here.

- If \(v_r > v_{r-1}\), the variance starts to increase and the algorithm will not converge (as the variance bound can not be satisfied). The algorithm must stop here.

- Otherwise, \(r \leftarrow r + 1\) and go back to step 5 with the following new matrix \(M:\)

\[
M := (m_{ij})_{ij} \leftarrow \\
\begin{bmatrix}
  m_{d1} & m_{d2} & \cdots & m_{dn} \\
  m_{11} & m_{12} & \cdots & m_{1n} \\
  \vdots & \vdots & \ddots & \vdots \\
  m_{(d-1)1} & m_{(d-1)2} & \cdots & m_{(d-1)n}
\end{bmatrix}.
\]

At each step 5 and 6, the algorithm makes the quantile function of the sum as flat as possible to the left and the right, respectively, of the cutoff point corresponding to desired probability level \(q\), used to assess the VaR. By construction, one expects the first value obtained for \(v\) to be slightly larger than \(s^2\). As \(r\) increases, the average value on the left increases but the average value on the right decreases. Hence, the new value \(v\) will decrease and may become smaller than \(s^2\). If \(s^2\) is a “feasible” variance constraint, then the algorithm will stop as soon as the variance constraint is met. However, due to the fact that there is no formal guarantee that the algorithm provides the true solution, it may happen that \(v\) never becomes smaller than \(s^2\) or that the values for \(v\) begin to increase again. The algorithm stops as soon as one of these two events happens and no solution is provided. This can only occur when the variance constraint \(s^2\) is very close to the minimum that is possible for the given portfolio (with fixed marginal distributions).

To further improve the quality of the algorithm, it appears useful (especially when risks are heavy tailed) to compute an upper bound (as an approximation for \(b_d\)) and a lower bound (as an approximation for \(a_d\)) separately. To this end, one can use the property that for any random variable \(X\) it holds that

\[-\text{VaR}_{1-q}^+(-X) = \text{VaR}_q(X).\]

Hence, one runs the extended RA twice: once as it is described above and once with \(X\) replaced by \(-X\) and \(q\) replaced by \(1 - q\). Next, one takes the best upper and lower bounds among the two algorithms. Typically, the algorithm as described above works better for obtaining an upper bound, whereas replacing the role of \(X\) by \(-X\) and \(q\) by \(1 - q\) facilitates a better approximation for the lower
bound. Finally, note that by setting the constraint \( s > \sum_{i=1}^{n} \sigma_i \) the ERA behaves as the RA and thus allows one to determine unconstrained numerical bounds.

The bounds that are obtained by running the algorithm are within the interval \((a_d, b_d)\). If the obtained values are close to the boundary values \( a_d \) and \( b_d \), then this means that the algorithm is indeed able to identify a dependence that gives rise to (almost) sharp VaR bounds for the portfolio sum. In the following section we assess the performance of the ERA using examples and find that the algorithm that we propose for computing approximate VaR bounds performs very well.

Remark 4.2 (Adapted ERA algorithm). The ERA ultimately amounts to applying the RA of Puccetti and Rüschendorf (2012) on two submatrices, namely (I) and (II). While the procedure that we use to (iteratively) select these submatrices is intuitive and appears to provide very good results, it is lacking a strong theoretical foundation. In fact, there might exist another partition of the matrix \( X \) that allows for better approximations of the sharp bounds. In this respect, the following procedure appears very useful in selecting these two submatrices. Specifically, we add an extra column to the initial matrix \( X \) representing a variable \(-X_{n+1}\) that takes the values \(-a\) and \(-b\) with probabilities \( q \) and \( 1-q \), respectively. The idea is to then apply the basic RA to the full matrix; after applying the RA, one obtains the candidate approximate sharp lower bound \( m_d := \max_{x \in A} (x_1^* + \cdots + x_n^*) \) as well as the candidate approximate sharp upper bound \( M_d := \min_{x \in B} (x_1^* + \cdots + x_n^*) \), in which \( A \) and \( B \) are defined as \( A = \{ i \in \{1, 2, \ldots, d\} : x_{i,n+1}^* = a \} \) and \( B = \{1, 2, \ldots, d\} \setminus A \). In the ideal situation, we observe that after running the RA the row sums are all equal to zero, and thus \( m_d \) and \( M_d \) are sharp. However, in general, the row sums will not be equal to zero and it is not a priori clear whether the sum \( X_1^* + X_2^* + \cdots + X_n^* \) meets the variance constraint. If it does, then \( m_d \) and \( M_d \) are our approximations for the sharp bounds. Otherwise, one considers the variable \(-X_{n+1}^\varepsilon (\varepsilon > 0)\), taking the values \(-a_\varepsilon\) and \(-b_\varepsilon\) (with respective probabilities \( q \) and \( 1-q \)), in which case \( a_\varepsilon = \max \left( \mu - (s - \varepsilon) \sqrt{\frac{1-q}{q}} , A \right) \) and \( b_\varepsilon = \min \left( \mu + (s - \varepsilon) \sqrt{\frac{1-q}{q}} , B \right) \), and one repeats the above procedure to yield new candidate approximations for the sharp bounds. By gradually increasing \( \varepsilon \), one obtains a situation in which \( X_1^* + X_2^* + \cdots + X_n^* \) meets the variance constraint and approximate VaR bounds are obtained. The primary advantages of this version of ERA is that it effectively determines a partition of the matrix \( X \) in a natural way, and that it is more computationally efficient as it involves fewer runs of the RA. A more detailed analysis of the ERA and the use of the block rearrangements (as in Remark 4.1) will be presented in a subsequent study.

4.3 Example with normally distributed risks

Assume that \((X_1, X_2, \ldots, X_n)\) is a vector of dependent standard normally distributed random variables with a correlation matrix \((\varrho_{ij})\) such that \( \varrho_{ii} = 1 \) for \( i = 1, \ldots, n \) and \( \varrho_{ij} = \varrho \) for all \( i \neq j \). Note that the dependence structure of \((X_1, X_2, \ldots, X_n)\) is only partially specified (through the knowledge of the
pairwise correlations), so that the variance of the sum is equal to

$$\text{var}(X_1 + X_2 + \cdots + X_n) = \sum_{i=1}^{n} \sum_{j=1}^{n} \rho_{ij}. $$

However, we cannot compute the VaR of the portfolio sum $S = X_1 + X_2 + \cdots + X_n$ precisely. Hence, we apply the ERA to compute numerical upper and lower bounds on VaR satisfying the variance constraint. We then compare these bounds with the theoretical bounds of Theorem 2.1 and Theorem 3.2.

In Table 4.1 we assess the VaRs at 95%, 99% and 99.5% for three levels of correlation $\rho = 0$, $\rho = 0.15$ and $\rho = 0.3$ and for portfolio sizes $n = 10$ and $n = 100$. Panel A shows the constrained lower and upper VaR bounds using the ERA with discretization level $d = 10,000$. We denote these constrained lower and upper bounds by $m_d$ and $M_d$, respectively. Panel B shows the corresponding constrained bounds $a_d$ and $b_d$ as an application of Theorem 3.2, whereas Panel C shows the unconstrained bounds $A_d$ and $B_d$ using Theorem 2.1. The last rows in Panels B and C show the values for $a_\infty := a$ and $b_\infty := b$, respectively: $A_\infty := A$ and $B_\infty := B$. These bounds are thus based on the original (non-discretized) distributions and are explicitly given as

$$A = -n \frac{\phi(\Phi^{-1}(q))}{q}, \quad B = n \frac{\phi(\Phi^{-1}(q))}{1-q}$$

and

$$a = \max \left( -s \sqrt{\frac{1-q}{q}}, A \right), \quad b = \min \left( s \sqrt{\frac{1-q}{q}}, B \right),$$

where $\phi$ and $\Phi$ denote the standard normal density and distribution function and where $s^2 = n + n(n-1)\rho$ is the variance of $S$. Finally, in Panel D, $(X_1, X_2, \ldots, X_n)$ is multivariate normally distributed (the dependence is thus assumed to be Gaussian), in which case the VaRs of the portfolio sum can be computed exactly.

There are several observations. First, when comparing the results of Panel A and Panel B, we observe that the ERA is performing remarkably well. The obtained numerical bounds $m_d$ and $M_d$ are very close to their theoretical counterparts $a_d$ and $b_d$, showing that the ERA is able to construct the dependence between the risks such that the sum is (almost exactly) concentrated on two values: $a_d$ and $b_d$.

Second, the distance between the upper and lower bounds, as reported in Panels A, B and C, is typically significant. For example, Panel B shows that the true 95%-VaR of a portfolio of 100 uncorrelated (but not independent) normally distributed risks lies in the interval $(-2.292; 43.55)$. Considering that the given portfolio has zero mean and a standard deviation of 10, this interval appears to be rather wide. Note indeed that when the risks are independent, then the exact 95%-VaR can be computed and is given by 16.449. In other words, when the risks are known to be independent, the 95%-VaR is approximately three times smaller than the reported upper bound (i.e., 43.59) that
is valid when we only know that the correlations are equal to zero. When we ignore the variance constraint, then the upper bound is as high as 206.3. We also observe that the distance between the bounds becomes wider when the level of the probability $q$ used to assess VaR increases. These observations already suggest that misspecification of models is a significant concern, especially when the VaRs are assessed at high probability levels (which is the case in solvency frameworks such as Solvency II and Basel III, where $q = 99.5\%$).

**Third**, when comparing the results shown in Panel B and Panel C, we observe that adding a variance constraint may have a significant impact on the unconstrained bounds. When the portfolio exhibits low to moderate correlation, then the constrained bounds $a$ and $b$ that we propose improve upon the unconstrained ones, $A$ and $B$. It is straightforward to show that

$$ a > A, \quad b < B \iff q \leq \hat{\rho}_1(n, q) := \frac{n (\phi(\Phi^{-1}(q)))^2}{q(1-q)(n-1)} - \frac{1}{n-1} \quad (4.2) $$

The relation (4.2) allows us to derive the critical correlation values $\hat{\rho}_1(n, q)$ as a function of portfolio size $n$ and probability level $q$. If the correlation $q$ of the portfolio is lower than the critical value,
then the constrained bounds improve upon the unconstrained ones. We report some values here: 
\( \hat{\rho}_1(10, 0.95) = 0.138, \hat{\rho}_1(10, 0.995) = -0.0644, \hat{\rho}_1(100, 0.95) = 0.216 \) and \( \hat{\rho}_1(100, 0.995) = 0.0324 \). Note that the critical correlation levels decrease when the probability level that is used for VaR assessment increases, indicating that adding dependence information does not readily allow for improving the unconstrained bounds when going “deep in the tail.” In this regard, it is also of interest to compare the constrained upper bound \( b \) with the portfolio VaR that one obtains in the case in which all risks are perfectly dependent (comonotonic), i.e., when the portfolio VaR is equal to \( n \Phi^{-1}(q) \).

We find that \( b < n \Phi^{-1}(q) \iff \rho \leq \hat{\rho}_2(n, q) := \frac{n(1 - q)(\Phi^{-1}(q))^2}{q(n - 1)} - \frac{1}{n - 1}. \) (4.3)

As critical values for the correlation parameter we report \( \hat{\rho}_2(10, 0.95) = 0.0471, \hat{\rho}_2(10, 0.995) = -0.0741, \hat{\rho}_2(100, 0.95) = 0.1337, \) and \( \hat{\rho}_2(100, 0.995) = 0.0236 \), which shows that adding the variance constraint may give rise to a VaR bound that strictly improves upon the comonotonic VaR. When the variance constraint \( s^2 \) is “too high” or the probability level \( q \) is “too high,” there is no improvement.5

4.4 Examples with Pareto distributed risks

We assume that \((X_1, X_2, \ldots, X_n)\) is an homogeneous portfolio of dependent Pareto distributed random variables (of type II). Hence, \( F_i(x) = 1 - (1+x)^{-\theta} (i = 1, 2, \ldots, n) \) with \( x > 0 \) and with a tail parameter \( \theta > 0 \). The correlation matrix \((\rho_{ij})\) is such that \( \rho_{ii} = 1 \) for \( i = 1, \ldots, n \) and \( \rho_{ij} = \rho \) for all \( i \neq j \).

We first consider the case \( \theta = 3 \) so that the first two moments exist, which allows us to compute the VaR bounds that we discussed in the previous sections. We first calculate

\[
E(X_i) = \frac{1}{\theta - 1}, \quad \text{var}(X_i) = \frac{2}{(\theta - 1)(\theta - 2)} - \frac{1}{(\theta - 1)^2}, \quad F_{X_i}^{-1}(q) = (1 - q)^{-1/\theta} - 1, \\
TVaR_q(X_i) = \frac{(1 - q)^{-1/\theta}}{1 - \frac{1}{\theta}} - 1, \quad \text{LTVaR}_q(X_i) = \frac{1}{q} (E(X_i) - (1 - q)TVaR_q(X_i)).
\]

When applying Theorem 2.1 we find that the absolute unconstrained bounds are

\[
B = n \frac{(1 - q)^{-1/\theta}}{1 - \frac{1}{\theta}} - n, \quad A = \frac{n}{q} \frac{1}{\theta - 1} - B \frac{(1 - q)}{q}. 
\]

And, from Theorem 3.2

\[
a = \max \left( -s \sqrt{\frac{1 - q}{q}}, A \right), \quad b = \min \left( s \sqrt{\frac{1 - q}{q}}, B \right),
\]

5This feature is consistent with the work of Cheung and Vanduffel [2013]. These authors show that in the presence of a variance constraint one can always construct a dependence structure among the risks \( X_i \) (with fixed marginal distributions) such that the sum \( S \) exhibits upper-comonotonicity (i.e., behaves as a comonotonic sum “in the tail”). In this case, for \( q \) large enough, one has that \( \text{VaR}_q(S) = \sum_{i=1}^n \text{VaR}_q(X_i) \).
where \( s^2 = \left( n + n(n-1) \rho \right) \left( \frac{2}{(\theta-1)(\theta-2)} \right) - \frac{1}{(\theta-1)^2} \).

We present the results of our calculations in Table 4.2 in a similar way as in the previous example. Panel A shows the numerical sharp lower and upper bounds obtained using the ERA (note that the discretization involves the computation of \( x_{ij} = (1 - i/(d + 1))^{-1/\theta} - 1 \) for \( i = 1, \ldots, d; \ j = 1, 2, \ldots, n \). Panel B shows the corresponding absolute constrained bounds, and Panel C shows the absolute bounds in the unconstrained case. The results are in line with those that we obtain in the case of normally distributed risks. Also, in this case the ERA gives rise to numerical bounds that are close to those that were obtained theoretically. In other words, the absolute bounds are “nearly sharp” in this case. The Pareto distribution has heavy tails, and hence one observes a significant difference between \( A_d \) (given by (4.1)) and \( A \) (respectively, \( B_d \) and \( B \)) as it appears in Table 4.2. The difference between the upper and lower bounds is again significant, confirming that in the case in which there is little or no information on the dependence, the model risk that goes along with a particular model is an issue. Note also that the impact of the variance constraint is more significant than in the normal case. For example, when \( \theta = 99.5\% \), \( n = 100 \) and \( \rho = 0.15 \), we find a numerical sharp bound \( m_{10,000} = 499.1 \) (close to the absolute bound \( b_{10,000} = 500.01 \)), whereas the unconstrained bound amounts to \( B_{10,000} = 741.1 \). Similar conclusions can be reached for other discretization levels, such as \( d = 1,000 \) and \( d = 100,000 \).

Panel A: Approximate sharp bounds obtained by the ERA as presented in Section 5.2

<table>
<thead>
<tr>
<th>((m_d, M_d))</th>
<th>(n = 10)</th>
<th>(n = 100)</th>
</tr>
</thead>
<tbody>
<tr>
<td>(\rho = 0)</td>
<td>(\rho = 0.15)</td>
<td>(\rho = 0.3)</td>
</tr>
<tr>
<td>(d = 10,000)</td>
<td>VaR(_{95%})</td>
<td>(4.401; 15.72)</td>
</tr>
<tr>
<td>VaR(_{99%})</td>
<td>(5.486; 28.69)</td>
<td>(4.591; 43.45)</td>
</tr>
<tr>
<td>VaR(_{99.5%})</td>
<td>(6.820; 39.48)</td>
<td>(5.471; 59.60)</td>
</tr>
</tbody>
</table>

Panel B: Variance-constrained bounds as obtained in Theorem 6.2

<table>
<thead>
<tr>
<th>((a_d, b_d))</th>
<th>(n = 10)</th>
<th>(n = 100)</th>
</tr>
</thead>
<tbody>
<tr>
<td>(\rho = 0)</td>
<td>(\rho = 0.15)</td>
<td>(\rho = 0.3)</td>
</tr>
<tr>
<td>(d = 10,000)</td>
<td>VaR(_{95%})</td>
<td>(4.398; 16.03)</td>
</tr>
<tr>
<td>VaR(_{99%})</td>
<td>(4.725; 30.20)</td>
<td>(4.589; 43.64)</td>
</tr>
<tr>
<td>VaR(_{99.5%})</td>
<td>(4.800; 40.74)</td>
<td>(4.705; 59.80)</td>
</tr>
<tr>
<td>(d = \infty)</td>
<td>VaR(_{95%})</td>
<td>(4.372; 16.94)</td>
</tr>
<tr>
<td>VaR(_{99%})</td>
<td>(4.725; 32.25)</td>
<td>(4.578; 46.77)</td>
</tr>
<tr>
<td>VaR(_{99.5%})</td>
<td>(4.806; 43.63)</td>
<td>(4.702; 64.22)</td>
</tr>
</tbody>
</table>

Panel C: Unconstrained bounds as obtained in Theorem 7.2 independent of \( \rho \)

<table>
<thead>
<tr>
<th>((A_d, B_d))</th>
<th>(n = 10)</th>
<th>(n = 100)</th>
</tr>
</thead>
<tbody>
<tr>
<td>(d = 10,000)</td>
<td>VaR(_{95%})</td>
<td>(3.646; 30.33)</td>
</tr>
<tr>
<td>VaR(_{99%})</td>
<td>(4.447; 57.76)</td>
<td>(4.4; 577.6)</td>
</tr>
<tr>
<td>VaR(_{99.5%})</td>
<td>(4.633; 74.11)</td>
<td>(46.33; 741.1)</td>
</tr>
<tr>
<td>(d = \infty)</td>
<td>VaR(_{95%})</td>
<td>(3.647; 30.72)</td>
</tr>
<tr>
<td>VaR(_{99%})</td>
<td>(4.448; 59.62)</td>
<td>(44.48; 596.2)</td>
</tr>
<tr>
<td>VaR(_{99.5%})</td>
<td>(4.635; 77.72)</td>
<td>(46.35; 777.2)</td>
</tr>
</tbody>
</table>

Table 4.2 Bounds on Value-at-Risk of sums of Pareto distributed risks (\( \theta = 3 \))

Also in this case we find that the constrained bounds \( a \) and \( b \) that we propose may improve upon the unconstrained ones, \( A \) and \( B \). Using the same notation as in the previous example, we find that \( \hat{\theta}_1(10, 0.95) = 0.2756 \), \( \hat{\theta}_1(10, 0.995) = 0.1842 \), \( \hat{\theta}_1(100, 0.95) = 0.3415 \) and \( \hat{\theta}_1(100, 0.995) = \)
Finally, the constrained upper bound $b$ may also sharpen the comonotonic VaR bound $n(1-q)^{-1/\theta} - n$. As critical values for the correlation parameter that make it possible to improve upon the comonotonic VaR bounds, we find $\hat{\rho}_2(10, 0.95) = 0.0608$, $\hat{\rho}_2(10, 0.995) = 0.0201$, $\hat{\rho}_2(100, 0.95) = 0.1462$ and $\hat{\rho}_2(100, 0.995) = 0.1092$.

5 VaR bounds of credit risk portfolios

In this section, we apply the theory regarding VaR bounds to the case of credit risk portfolios and discuss the results in the context of model risk assessment. Here, we merely define model risk as the risk that the computed portfolio-VaR is incorrect as a result of using a misspecified model. Note that a related issue concerns the risk of using model parameters that are estimated, a concern that arises because of the statistical uncertainty of these estimations; we refer to Bignozzi and Tsanakas (2015) for a study on the impact of parameter uncertainty on capital adequacy.

We consider a portfolio $(X_1, X_2, \ldots, X_n)$ containing non-negative risks $X_i = v_i \mathcal{B}(p_i) \sim F_i, (i = 1, 2, \ldots, n)$. Here, each $X_i$ can be seen as a representation of the bank’s risk exposure when providing a loan to company “$i$.“ Specifically, $p_i$ is the probability that the $i$-th company defaults and, in the case of a default, the loss incurred is equal to $v_i$. We denote by $p_{ij}$ the pairwise default probability that both company $i$ and company $j$ default. The pairwise default correlation $\varrho^D_{ij} (i, j = 1, 2, \ldots, n)$ is then given as

$$\varrho^D_{ij} = \frac{p_{ij} - p_i p_j}{\sqrt{p_i(1-p_i)} \sqrt{p_j(1-p_j)}}. \quad (5.1)$$

The variance of the portfolio sum $S = X_1 + X_2 + \cdots + X_n$ thus depends on the exposures $v_i$ (net of recoveries), the single default probabilities $p_i$ and the pairwise default probabilities $p_{ij} (i, j = 1, 2, \ldots, n)$. As there is an intrinsic lack of sufficient default statistics (joint defaults are inherently very rare events), it becomes clear that in practice the knowledge of the above mentioned parameters is the maximum amount of information available when building models. In other words, all models that compute risk measures for credit risk portfolios require some further ad-hoc assumptions for describing the full dependence (e.g., the specification of the probabilities that three or more loans default together). This line of reasoning shows that the problem setting discussed in this paper is particularly relevant to credit risk.

To compare our VaR bounds with the VaRs calculated from various standard models, we consider a homogeneous portfolio of credit risks with net exposures $v_i = 1 (i = 1, 2, \ldots, n)$. Let $p$ and $\varrho^D$ denote the default probability and pairwise default correlation between two risks $X_i$ and $X_j$ for $i \neq j$.

\footnote{In fact, it is already ambitious to have all pairwise correlations available. It is more realistic to assume knowledge of the portfolio variance (for example, based on an analysis of the aggregate default statistics).}
5.1 Credit Risk Portfolios as mixtures

Many industry credit risk portfolio models rely on “Merton’s model of the firm” when computing the VaR of a portfolio (see the survey conducted by Baer et al. (2009)). The Basel III and Solvency II regulatory frameworks rely on this model when setting their VaR-based capital requirements. The basic idea of the Merton approach is to model a default as the event when the asset value drops below a threshold value. Formally, after normalization, a default for the $i$-th risk occurs for the event $\{N_i < c\}$ where $N_i$ is the normalized asset return and $c$ is the constant threshold value such that $p = p_i = P(N_i < c)$. Merton’s model further assumes that the joint asset (log-)returns are multivariate normally distributed. Hence, for an homogeneous portfolio one can express the standardized asset return for the $i$-th obligor $N_i$ conveniently as

$$N_i = \sqrt{\varrho^A} Z + \sqrt{1 - \varrho^A} \varepsilon_i.$$  \hspace{1cm} (5.2)

Here, $Z$ and $\varepsilon_i$ ($i = 1, 2, \ldots, n$) are independent standard normally distributed random variables and $\varrho^A \geq 0$ is the (asset) correlation coefficient. The variable $Z$ corresponds to a systemic factor (describing the global economy) and the $\varepsilon_i$ reflect the idiosyncratic (individual) risks. For a given realization for $Z = z$ the conditional default probability $p(z)$ is given as

$$p(z) = P(N_i < c | Z = z) = \Phi \left( \frac{\Phi^{-1}(p) - \sqrt{\varrho^A} z}{\sqrt{1 - \varrho^A}} \right).$$  \hspace{1cm} (5.3)

where $\Phi$ is the distribution of the standard normal random variable. The VaRs of the portfolio sum $S$ can now be easily estimated, for example, using Monte-Carlo simulations. In fact, the model as described above is an example of a one-factor mixture model in which the default event of the obligor is assumed to be driven by a common economic factor $Z$. It can also be seen as the one-factor version of the KMV model widely used in the industry and enshrined in regulatory frameworks such as the Basel III and Solvency II framework.

It is clear that other distributions for $p(Z)$ can also be used and other choices that have been made in the literature include a logit-normal mixing distribution (one obtains the one-factor version of the CreditMetrics approach) and a Beta distribution. Note that this last model is intimately related to the actuarial approach to credit risk portfolio modelling, which is also known as the one-factor CreditRisk+ model; see Vandendorpe et al. (2008) for an analysis. A detailed study of these mixtures models can be found in McNeil et al. (2005).

**Parameterization:** The natural parameterization of Merton’s model consists in the knowledge of the default probability $p$ and the asset correlation $\varrho^A$. The latter parameter can be estimated using asset value data, but unfortunately these values are not readily observable. One way to deal with this issue is to generate pseudo-asset values that are based on equity value data or some other data series that can be used (after suitable transformation) as a proxy for asset values; see e.g., Duellmann et al.
Another approach consists in estimating the default correlation $\rho^D$ using default statistics and then inferring from $\rho^D$ and $p$ the implied asset correlation $\rho^A$ (using the KMV model$^7$). Papers relevant to the latter category include Gordy (2000), Frey and McNeil (2003) and Dietsch and Petey (2004).

We remark that it is common to use the asset correlations (and not the default correlations) as the basis for comparison. The above mentioned studies report different values for the asset correlations, depending on such parameters as loan quality (level of default probability), the sector of activity of the obligor and so on. As a conclusion, the reported values range between 1% and 30%, with a mean in the area of 8 – 10%. This range for asset correlations is also consistent with the values that are used in regulatory frameworks (for example, Basel III uses asset correlations up to 30%). In Table 5.3 we report the default correlation as a function of the default probability and the asset correlation.

<table>
<thead>
<tr>
<th>$\rho^A$</th>
<th>$p = 0.05%$</th>
<th>$p = 0.50%$</th>
<th>$p = 1%$</th>
<th>$p = 5%$</th>
</tr>
</thead>
<tbody>
<tr>
<td>0%</td>
<td>0.00%</td>
<td>0.00%</td>
<td>0.00%</td>
<td>0.00%</td>
</tr>
<tr>
<td>4%</td>
<td>0.03%</td>
<td>0.19%</td>
<td>0.32%</td>
<td>0.94%</td>
</tr>
<tr>
<td>8%</td>
<td>0.08%</td>
<td>0.44%</td>
<td>0.71%</td>
<td>1.99%</td>
</tr>
<tr>
<td>12%</td>
<td>0.15%</td>
<td>0.75%</td>
<td>1.18%</td>
<td>3.14%</td>
</tr>
<tr>
<td>16%</td>
<td>0.24%</td>
<td>1.13%</td>
<td>1.75%</td>
<td>4.40%</td>
</tr>
<tr>
<td>20%</td>
<td>0.38%</td>
<td>1.60%</td>
<td>2.41%</td>
<td>5.78%</td>
</tr>
<tr>
<td>24%</td>
<td>0.56%</td>
<td>2.17%</td>
<td>3.19%</td>
<td>7.28%</td>
</tr>
</tbody>
</table>

Table 5.3 Default correlations $\rho^D$ for given levels of the default probability $p$ and the asset correlation $\rho^A$.

Whereas in the case of the KMV model the natural parameterization consists in a direct specification of the default probability $p$ and the asset correlation $\rho^A$, in the other models the input parameters are the default probability $p$ and the default correlation $\rho^D$ (obtained though an historical data analysis or inferred from the asset correlation $\rho^A$). Hence, the mean and the variance of $p(Z)$ are specified since $E(p(Z)) = p$ and $\text{var}(p(Z)) = \rho^D p(1-p)$. Next, one infers the two parameters of the mixing distribution at hand. In the case of a Beta distribution these parameters can be obtained analytically; for the CreditMetrics approach one can use numerical techniques. We provide further details in Appendix A.6.

5.2 Comparing the VaRs and assessing the impact of the variance constraint

We first consider a portfolio of 10,000 loans. Using Table 8.6 on page 365 of McNeil et al. (2005), we fix the default probability $p = 0.049$ and $\rho^D = 0.0157$. The variance $s^2$ of the portfolio sum of $n$
correlated Bernoulli risks is thus equal to

\[ s^2 = np(1-p) + n(n-1)p(1-p)q^D. \]

Consider for instance \( d = 1,000 \). Each \( X_j \) (\( j = 1, 2, \ldots, n \); \( n = 10,000 \)) takes the value 0 in 951 states and 1 in 49 states, so that, effectively, each \( X_j \) has a Bernoulli distribution with parameter 0.049. We thus consider a matrix \((x_{ij})_{i=1,d;j=1,...,n}^{d \times n}\) of \( d \times n \) entries. Let \( q \) denote the probability level that is used to compute the Value-at-Risk. We then apply Theorem 2.1 to find \( A_d \) and \( B_d \):

\[ A_d = n \left( \frac{p - \min(p, (1-q))}{q} \right), \quad B_d = n \min(p/(1-q), 1). \]

Furthermore, using Theorem 3.2 we derive \( a_d \) and \( b_d \) as

\[ a_d = \max \left( np - s \sqrt{\frac{1-q}{q}}, A_d \right), \quad b_d = \min \left( np + s \sqrt{\frac{q}{1-q}}, B_d \right). \]

Similarly to the cases involving normally and Pareto distributed risks, we apply the extended RA to find numerical VaR bounds. Finally, we also report the results that are obtained using different mixture models that we apply asymptotically. The results are reported in Table 5.4. All numbers are normalized as percentage of the maximum possible total loss, i.e., \( n \). In other words, in Table 5.4 the outputs \( A_d, B_d, a_d, b_d \) and the approximate sharp bounds that are obtained by applying the extended RA are divided by \( n \) and multiplied by 100. The example shows that adding a variance constraint has a significant impact on the level of the VaR bounds. For example, the unconstrained upper bound of the 95%-VaR is 98%, but the constrained one is only 16.73%. As expected, the difference between the upper and lower bounds is increasing significantly when the probability level used for VaR assessments increases. Clearly, when using \( q = 99.5\% \) as the basis for calculating VaR and capital requirements, the results of the models typically fall within the broad range of possible values of VaR; however, as these models only use information on the default probability \( p \) and the default correlation \( q^D \), the results are difficult to justify. Model risk appears more limited when using lower probability levels to assess the VaR. For example, when using the 90%-VaR, the distance between the upper bound \( b_d \) and the lower bound \( a_d \) is smaller and all industry models provide outcomes that are nearly in the middle of the interval \((a_d, b_d)\).

<table>
<thead>
<tr>
<th>VaR</th>
<th>((A_d, B_d))</th>
<th>((a_d, b_d))</th>
<th>((m_d, M_d))</th>
<th>KMV</th>
<th>Beta</th>
<th>CreditMetrics</th>
</tr>
</thead>
<tbody>
<tr>
<td>VaR0.8</td>
<td>(0%; 24.50%)</td>
<td>(3.543%; 10.327%)</td>
<td>(3.555%; 10.32%)</td>
<td>6.84%</td>
<td>6.95%</td>
<td>6.71%</td>
</tr>
<tr>
<td>VaR0.9</td>
<td>(0%; 49.00%)</td>
<td>(3.996%; 13.039%)</td>
<td>(4.00%; 13.03%)</td>
<td>8.51%</td>
<td>8.54%</td>
<td>8.41%</td>
</tr>
<tr>
<td>VaR0.95</td>
<td>(0%; 98.00%)</td>
<td>(4.278%; 16.727%)</td>
<td>(4.28%; 16.72%)</td>
<td>10.10%</td>
<td>10.01%</td>
<td>10.11%</td>
</tr>
<tr>
<td>VaR0.995</td>
<td>(4.42%; 100.00%)</td>
<td>(4.708%; 43.176%)</td>
<td>(4.71%; 43.17%)</td>
<td>15.15%</td>
<td>14.34%</td>
<td>15.87%</td>
</tr>
</tbody>
</table>

Table 5.4 VaR bounds and VaR estimates computed in different models (KVV, Beta, CreditMetrics). Here, \((A_d, B_d)\) are the unconstrained theoretical bounds, \((a_d, b_d)\) are the theoretical constrained bounds and \((m_d, M_d)\) are the approximate bounds. The approximate bounds have been obtained following Remark 4.2 as doing so provided the best results.
Finally, we compute the bounds $A$, $a$, $b$ and $B$ as well as the VaRs in a KMV framework for an infinitely large homogenous portfolio assuming a relevant range of default probabilities and asset correlations. The results are reported in Table 5.5 and confirm the findings of the previous example.

<table>
<thead>
<tr>
<th>$\varphi^4$</th>
<th>$(A, B)$</th>
<th>$p = 0.25%$</th>
<th>KMV</th>
<th>$(A, B)$</th>
<th>$p = 1%$</th>
<th>KMV</th>
</tr>
</thead>
<tbody>
<tr>
<td>0%</td>
<td>(0%; 50%)</td>
<td>(0.25%; 0.25%)</td>
<td>0.25%</td>
<td>(0.50%; 100%)</td>
<td>(1.00%; 1.00%)</td>
<td>1.0%</td>
</tr>
<tr>
<td>6%</td>
<td>(0%; 50%)</td>
<td>(0.23%; 3.27%)</td>
<td>1.2%</td>
<td>(0.50%; 100%)</td>
<td>(0.95%; 10.98%)</td>
<td>4.0%</td>
</tr>
<tr>
<td>12%</td>
<td>(0%; 50%)</td>
<td>(0.23%; 5.05%)</td>
<td>2.1%</td>
<td>(0.50%; 100%)</td>
<td>(0.92%; 16.27%)</td>
<td>6.3%</td>
</tr>
<tr>
<td>18%</td>
<td>(0%; 50%)</td>
<td>(0.23%; 6.84%)</td>
<td>2.9%</td>
<td>(0.50%; 100%)</td>
<td>(0.90%; 21.18%)</td>
<td>8.7%</td>
</tr>
<tr>
<td>24%</td>
<td>(0%; 50%)</td>
<td>(0.21%; 8.76%)</td>
<td>3.8%</td>
<td>(0.50%; 100%)</td>
<td>(0.87%; 26.09%)</td>
<td>11.1%</td>
</tr>
<tr>
<td>30%</td>
<td>(0%; 50%)</td>
<td>(0.20%; 10.85%)</td>
<td>4.8%</td>
<td>(0.50%; 100%)</td>
<td>(0.85%; 31.13%)</td>
<td>13.7%</td>
</tr>
</tbody>
</table>

Table 5.5 Unconstrained and constrained upper and lower 0.995-VaR bounds for several combinations of default probability and correlation and VaR in the (one-factor) KMV model

In particular, the results shown in Table 5.5 demonstrate the significant impact of the variance constraint on the VaR bounds. For example, when the asset correlation $\varphi^4 = 6\%$ and $p = 1\\%$, one has that the unconstrained upper bound for the 99.5%-VaR is 100\%, whereas the constrained bound is only 11.1\%. These findings also confirm that computing capital requirements based on a 99.5\% VaR is prone to significant model error as the distance between the upper and lower VaR bounds is too large. For example, for an asset correlation $\varphi^4 = 6\%$ and a default probability $p = 0.25\%$, the most optimistic model would report a 99.5\%-VaR that is equal to 0.2\%, whereas the most pessimistic one would provide in this instance a 99.5\%-VaR equal to 3.3\%. Note that extreme distributions having mass points in $A$ and $B$ (unconstrained case) and in $a$ and $b$ (constrained case), respectively, can be obtained by taking $p(Z)$ as a two point distribution with the same mass points. All in all, the results show that knowledge of the marginal distributions and the correlations is not a sufficient basis on which to estimate true portfolio VaRs with confidence. If one has additional dependence information at hand, such as a more precise description of the interaction among risks in times of stress, then the quality of VaR assessments increases. It might then be possible to obtain more acceptable risk bounds even when the confidence level used is high.

6 Conclusions

Recent literature has dealt with the problem of finding sharp bounds on the Value-at-Risk of risky portfolios when the distributions of the risky components are known. This problem is challenging and few theoretical results are as yet available that contribute to our capacity to deal with situations in which some information on the dependence is known.

\[ \Phi^{-1}(p) + \sqrt{\varphi^4} \Phi^{-1}(q) \]

Note that since the default events are conditionally independent, it follows that $\lim_{n \to \infty} \text{VaR}_q \left( \frac{n}{m} \right) = \Phi \left( \frac{\Phi^{-1}(p) + \sqrt{\varphi^4} \Phi^{-1}(q)}{\sqrt{1 - \varphi^4}} \right)$, see Vasicek (2001).
In this paper we consider a variance constraint for the portfolio sum as a source of information on the dependence, and we propose simple bounds that are easy to compute. These bounds are typically not sharp, but their construction as well as some of the theoretical results regarding convex ordering and mixability provide us with the intuition to propose a new algorithm that allows us to approximate the sharp bounds. Several numerical examples show that the algorithm performs well and confirm that a variance constraint can significantly improve the unconstrained bounds of Embrechts et al. (2013). This algorithm can thus be seen as a practical way to deal with a problem that is otherwise difficult to solve theoretically.

We believe that our results are useful for studying model risk. In the paper we touch on this issue by discussing model risk of credit risk portfolio models. We show that the VaR computed in typical credit models reported by financial institutions do not necessarily reflect the true risk and are difficult to compare. In this respect, we note that under the internal model approach of Basel III and Solvency II, financial institutions are allowed to use their own models to set their capital requirements. However, it is difficult, if not impossible, to show that a particular model is better than all others, as all might be consistent with available information (namely, default probabilities and default correlations in the credit risk context). The bounds are then useful for discussing adequacy of internal models used to set capital requirements and regulators may use them when considering policy. For example, in the context of credit risk portfolios they may want to ask financial institutions to also report VaRs numbers using a common model (e.g., KMV) so that the risk could be readily compared across institutions.

It is expected that adding more dependence information will reduce model uncertainty and one may obtain acceptable risk bounds even when the confidence level used is high. Recent attempts in that direction are Bernard and Vanduffel (2015) who add information on the joint distribution in some states and Bernard, Denuit and Vanduffel (2015) who include information on higher moments as source of dependence information.

A Appendix

A.1 Proof of Theorem 2.1

The result for the upper bound has already been proven. For the lower bound we make use of the fact that for any random variable $X$

$$q \quad \text{LTVaR}_q(X) + (1 - q) \quad \text{TVaR}_q(X) = E(X).$$

This implies as in the first part that $\text{VaR}_q(S) \geq \text{LTVaR}_q(S) \geq \sum_{i=1}^{n} \text{LTVaR}_q(X_i) = A.$

$\square$
A.2 Proof of Theorem 2.3

The upper bound \( B \) is sharp if and only if the inequalities in (2.1) and (2.2) are in fact equalities. Equality in (2.2) is by definition of TVaR equivalent to the fact that the \( f_i \) are rearrangements of \( F_{i}^{-1} \) on \([q, 1]\) (see for example Puccetti and Rüschendorf \(2013\)). Furthermore, equality holds in (2.1) if and only if \( S = \sum_{i=1}^{n} f_i(u) = c \) almost surely on \([q, 1]\), i.e. if and only if the random variables \( X_i = f_i(U) \) are mixing on \([q, 1]\) (see also Figure 2.1). The argument for the sharpness in b) is similar.

\[ \square \]

A.3 Proof of Theorem 2.5

Let us first remark that using the introduction to Section 2.2,

\[
M = \sup_{X_i \sim F_i} \text{VaR}_q^+ \left( \sum_{i=1}^{n} X_i \right) = \sup_{Y_i^q \sim F_i^q} \text{VaR}_0^+ \left( \sum_{i=1}^{n} Y_i^q \right). \tag{A.1}
\]

Moreover,

\[
\text{VaR}_0^+ \left( \sum_{i=1}^{n} Y_i^q \right) = \essinf \left( \sum_{i=1}^{n} Y_i^q \right) \tag{A.2}
\]

is the minimal support of the distribution of \( \sum_{i=1}^{n} Y_i^q \). Then, the problem of maximizing the VaR is equivalent to maximizing the minimal support of \( \sum_{i=1}^{n} Y_i^q \) over all possible \( Y_i^q \sim F_i^q \). This problem in turn is closely related to convex ordering. Let \( Y = \sum_{i=1}^{n} Y_i^q \) and \( Z = \sum_{i=1}^{n} Z_i^q \) have cdf \( F \) and \( G \), respectively (with \( Y_i^q, Z_i^q \sim F_i^q \)). Then,

\[
Y \leq_{\text{cx}} Z \iff \int_0^u F^{-1}(p)dp \geq \int_0^u G^{-1}(p)dp \quad \text{for all } 0 < u < 1 \tag{A.3}
\]

\[
\iff F^{-1} \lesssim_S G^{-1},
\]

where \( \lesssim_S \) is the Schur order (see Rüschendorf (1983b; 2013, Corollary 3.26 page 430) and Dhaene et al. \(2006\)). In particular, (A.3) implies that

\[
F^{-1}(0) = \essinf \left( \sum_{i=1}^{n} Y_i^q \right) \geq G^{-1}(0) = \essinf \left( \sum_{i=1}^{n} Z_i^q \right) \tag{A.4}
\]

and \( F^{-1}(1) \leq G^{-1}(1) \). The minimal support of \( Y \) is larger than the minimal support of \( Z \) and the maximal support of \( Y \) is smaller than the maximal support of \( Z \). As a consequence, the convex ordering in (A.3) implies ordering of the VaRs.

\[ \square \]
A.4 Proof of Theorem 3.2

Let us define the function

\[ B(\alpha) := \frac{1}{1 - q} \int_{q}^{1 - \alpha} \text{VaR}_u(S^c) \, du \]  

(A.5)
on the interval \([0, q]\) and note that \(B(0) = B\). Let us also define the variable \(X^*_\alpha (\alpha \in [0, q])\) that takes the values \(A(\alpha)\) and \(B(\alpha)\),

\[ X^*_\alpha = \begin{cases}  
A(\alpha) & \text{with probability } q, \\
B(\alpha) & \text{with probability } 1 - q,
\end{cases} \]  

(A.6)
in which \(A(\alpha) := \frac{\mu - B(\alpha)(1 - q)}{q}\). Note that \(A(0) = A\) and that \(X^*_0 =_{d} X^*\). Furthermore, as per construction, \(E(X^*_\alpha) = \mu\). The variance function \(\alpha \to \text{var}(X^*_\alpha)\) has the following monotonicity property:

**Monotonicity property of \(\text{var}(X^*_\alpha)\).** The variance of \(X^*_\alpha\) given by \(\text{var}(X^*_\alpha) = q(A(\alpha) - \mu)^2 + (1 - q)(B(\alpha) - \mu)^2\) is maximum when \(\alpha = 0\).

To prove this property, note that by increasing \(\alpha\) we decrease \(B(\alpha)\) and thus increase \(A(\alpha)\). Hence, there exists \(0 < \beta < q\) such that \(\alpha \to \text{var}(X^*_\alpha) := q(A(\alpha) - \mu)^2 + (1 - q)(B(\alpha) - \mu)^2\) is continuously decreasing on \([0, \beta]\) with minimum value given by \(\text{var}(X^*_\beta) = 0\) and with maximum value given by \(\text{var}(X^*_0) = q(A - \mu)^2 + (1 - q)(B - \mu)^2\).

Next, we first prove that

\[ a := A(\alpha^*) \leq m \leq \text{VaR}_q(S) \leq \text{VaR}_q^+(S) \leq M \leq b := B(\alpha^*), \]  

(A.7)
in which \(\alpha^*\) is defined as

\[ \alpha^* := \min \{ \alpha \mid 0 \leq \alpha \leq q, \text{var}(X^*_\alpha) \leq s^2 \}. \]  

(A.8)

**Proof of (A.7):** If \(\text{var}(X^*_\alpha) < s^2\), this means that \(\alpha^* = 0\). Hence, \(A(\alpha^*)\) and \(B(\alpha^*)\) correspond to the absolute bounds and there is nothing to prove. We further assume that \(\text{var}(X^*_\alpha) = s^2\) and denote by \(G\) the distribution of \(X^*_\alpha\). We first prove that \(b\) is an absolute upper bound for feasible solutions of (1.2). Hence, assume there exist \((X_1, X_2, \ldots, X_n)\) such that \(\text{VaR}_q^+(X_1 + X_2 + \cdots + X_n) > b\). One has that \(\forall a \leq x < b, F_{X_1 + X_2 + \cdots + X_n}(x) \leq G(x) = q\). When \(b \leq x\), \(F_{X_1 + X_2 + \cdots + X_n}(x) \leq G(x) = 1\). Since \(G(x) = 0\) when \(x < a\), this implies that,

\[ \forall x < a, F_{X_1 + X_2 + \cdots + X_n}(x) \geq G(x), \]
\[ \forall x \geq a, F_{X_1 + X_2 + \cdots + X_n}(x) \leq G(x). \]  

(A.9)

In other words, the distribution function \(F_{X_1 + X_2 + \cdots + X_n}\) crosses \(G\) once from above. Since \(E(X_1 + X_2 + \cdots + X_n) = \mu\), this implies that \(X^*_\alpha \leq_{cx} X_1 + X_2 + \cdots + X_n\) (see Karlin and Novikoff (1963)).
Müller and Stoyan (2002)). Since \( \text{var}(X^*_\alpha) = s^2 \), the feasibility of \((X_1, X_2, \ldots, X_n)\) requires that \( \text{var}(X_1 + X_2 + \cdots + X_n) = \text{var}(X^*_\alpha) \). In view of the convex ordering between \( X_1 + X_2 + \cdots + X_n \) and \( X^*_\alpha \), this is only possible when \( X_1 + X_2 + \cdots + X_n \overset{d}{=} X^*_\alpha \) (here, \( \overset{d}{=} \) means that there is equality in distribution), which is a contradiction.

The proof that \( a \) is an absolute lower bound can be given in a similar way. Let now \((X_1, X_2, \ldots, X_n)\) be such that \( \text{VaR}_q(X_1 + X_2 + \cdots + X_n) < a \). One has that \forall x \leq a, F_{X_1+X_2+\cdots+X_n}(x) \geq G(x) = 0 \). When \( a \leq x < b \), \( F_{X_1+X_2+\cdots+X_n}(x) \geq G(x) = q \). Since \( G(x) = 1 \) when \( x \geq b \), this implies that,

\[
\begin{align*}
\forall x < b, & \quad F_{X_1+X_2+\cdots+X_n}(x) \geq G(x), \\
\forall x \geq b, & \quad F_{X_1+X_2+\cdots+X_n}(x) \leq G(x).
\end{align*}
\] (A.10)

In other words, the distribution function \( F_{X_1+X_2+\cdots+X_n} \) crosses \( G \) once from above. By symmetry of the argument the result follows from the first part of the proof.

From the expression (A.7), we can finish the proof of Theorem 3.2. If \( s^2 \geq q(A-\mu)^2+(1-q)(B-\mu)^2 \) then the result is obvious from Theorem 2.1 and the monotonicity property stated above. In the other case, the proposition implies that there exists \( \alpha^* \) such that \( \text{var}(X^*_\alpha) = s^2 \). Hence, \( a \) and \( b \) can be seen as the mass points from a 2-point distribution satisfying the mean constraint \( \mu \) and the variance constraint \( s^2 \). This yields the desired expressions for \( a \) and \( b \) immediately.

\[\square\]

A.5 Proof of Theorem 3.5

We note that by the convex ordering assumption in (3.6) we obtain for the upper \( q \)-part \( T = \{U \geq q\} \) of \( S \)

\[
E(S)^2 = E((S)^2 \mid T)P(T) + E((S)^2 \mid T^c)P(T^c)
\leq E(S^2 \mid T)P(T) + E(S^2 \mid T^c)P(T^c)
= E(S^2)
\]

and thus \( \text{var}(\overline{S}) \leq \text{var}(S) \leq s^2 \). The argument for the increase of \( \text{VaR}_q^+(\overline{S}) \) (resp. decrease of \( \text{VaR}_q^+(S) \)) compared to \( \text{VaR}_q^+(S) \) (resp. \( \text{VaR}_q(S) \)) is similar as in Theorem 2.5.

\[\square\]

A.6 Mixing Distributions useful for Section 5.1

Note that all mixture models that we discuss require 2 parameters only and are thus completely specified once the default probability and the default correlation (or the asset correlation) are known.

**Beta distribution.** The mixing variable \( p(Z) \) has a Beta distribution with parameters \( a, b > 0 \) if it has
a density given as

$$f(x) = \frac{\Gamma(a+b)}{\Gamma(a)\Gamma(b)}x^{a-1}(1-x)^{b-1} \text{ for } 0 < x < 1,$$

(A.11)

where $\Gamma(x)$ is the Gamma function. It readily shows that $p = \frac{a}{a+b}$ and $g^D = \frac{1}{a+b+1}$, which allows us to obtain $a$ and $b$ as a function of $g^D$ and $p$ in explicit form:

$$a = p \left(\frac{1}{g^D} - 1\right), \quad b = (1-p) \left(\frac{1}{g^D} - 1\right).$$

Probit-Norm distribution (KMV). The mixing variable $p(Z)$ is said to have a Probit-Norm distribution if it writes as $p(Z) = \Phi(\mu + \sigma Z)$ where $\Phi$ is the distribution function of a standard normal random variable. In this case,

$$\mu = \frac{\Phi^{-1}(p)}{\sqrt{1 - g^A}}, \quad \sigma^2 = \frac{g^A}{1 - g^A}.$$

If the default correlation is provided and not the asset correlation then one first needs to back out the asset correlation.

Logit-Norm distribution. Finally, in the case of the Logit-Normal mixing distribution, $p(Z) = F(\mu + \sigma Z)$ with $F(x) = \frac{1}{1 + \exp(-x)}$. As the moments of $p(Z)$ are not known analytically, parameters $\mu$ and $\sigma$ can only be obtained numerically.
References


