Bounds on Aggregate Assets

by
Xiao Jiang

A thesis
presented to the University of Waterloo
in fulllment of the
thesis requirement for the degree of
Master of Quantitative Finance

Waterloo, Ontario, Canada, 2013

I hereby declare that I am the sole author of this thesis. This is a true copy of the thesis, including any required final revisions, as accepted by my examiners.

I understand that my thesis may be made electronically available to the public.

Abstract

Aggregating financial assets together to form a portfolio, commonly referred to as "asset pooling", is a standard practice in the banking and insurance industries. Determining a suitable probability distribution for this portfolio with each underlying asset is a challenging task unless several distributional assumptions are made. On the other hand, imposing assumptions on the distribution inhibits its ability to capture various idiosyncratic behaviours. It limits the model's usefulness in its ability to provide realistic risk metrics of the true portfolio distribution. In order to conquer this limitation, we propose two methods to model a pool of assets with much less assumptions on the correlation structure by way of finding analytical bounds.

Our first method uses the Fréchet-Hoeffding copula bounds to calculate model-free upper and lower bounds for aggregate assets evaluation. For the copulas with specific constraints, we improve the Fréchet-Hoeffding copula bounds by providing bounds with narrower range. The improvements proposed are very robust for different types of constraints on the copula function. However, the lower copula bound does not exist for dimension three and above.

Our second method tackles the open problem of finding lower bounds for higher dimensions by introducing the concept of Complete Mixability property. With such technique, we are able to find the lower bounds with specified constraints. Three theorems are proposed. The first theorem deals with the case where all marginal distributions are identical. The lower bound defined by the first theorem is sharp under some technical assumptions. The second theorem gives the lower bound in a more general setup without any restriction on the marginal distributions. However the bound achieved in this context is not sharp. The third theorem gives the sharp lower bound on Conditional *VaR*. Numerical results are provided for each method to demonstrate sharpness of the bounds.

Finally, we point out some possible future research directions, such as looking for a general sharp lower bound for high dimensional correlation structures.

Acknowledgments

Foremost, I would like to express my sincere gratitude to my supervisors Prof. Carole Bernard and Prof. Ruodu Wang for their continuous support of my Master study. They are my mentors in quantitative finance and in professional life in general. I would also like to thank Stephen Szaura, Yilan Gu and my second readers, Professor David Saunders and Professor Chengguo Weng, for their constructive inputs. Without their patient guidance and enthusiasm I cannot accomplish this thesis.

Contents

List of Tables

1 Introduction

Asset pooling is the practice of combining many different assets into one portfolio. This is a common idea in finance and one of its major motivations is to manage the overall risk. By combining assets that have either small or negative correlations, the entire portfolio becomes less volatile. Some examples include portfolio of mortgages that have similar time to maturity but are in different geographic regions. The combined portfolio manages to diversify and even eliminate some risk. More detailed examples will be discussed in section 2 to demonstrate this concept.

Theory in asset pooling is well established. Markowitz addresses this practice systematically in Modern Portfolio Theory (Markowitz (1952)). Development in statistical models for correlation structure is mature. For instance, a copula is a robust function to model the correlation structure of pooled assets. Nelsen (2006) gives a comprehensive review on copulas. Fréchet (1951) and Hoeffding (1940) together proposed the Fréchet-Hoeffding point-wise copula bounds that lay the foundation for finding bounds on aggregate assets. Tankov improved the standard Fréchet-Hoeffding copula bounds in Tankov (2011). In this thesis, we first focus on improving Tankov's work on copula bounds, which is presented in section 3.

Studies on copulas are relatively advanced in dealing with bivariate dependence structures. On the other hand, there are computational and convergence issues with statistical inference of multidimensional data, and the choice of multivariate distributions is rather limited compared with the modeling of marginal distributions. An inappropriate dependence assumption can have significant risk management consequences. For example, the abuse of the Gaussian multivariate copula can severely underestimate probability of simultaneous default in a large basket of firms (McNeil et al. (2005)). In section 4, we focus on another model based on the Complete Mixability property introduced by Wang and Wang (2011). We prove three theorems based on the Complete Mixability property to improve point-wise lower bounds for dependence structures with specific marginal distributions. Some numerical illustrations are provided to demonstrate the improvements of the lower bounds over old models.

In section 5, we conclude the research covered in the thesis and propose some possible extensions.

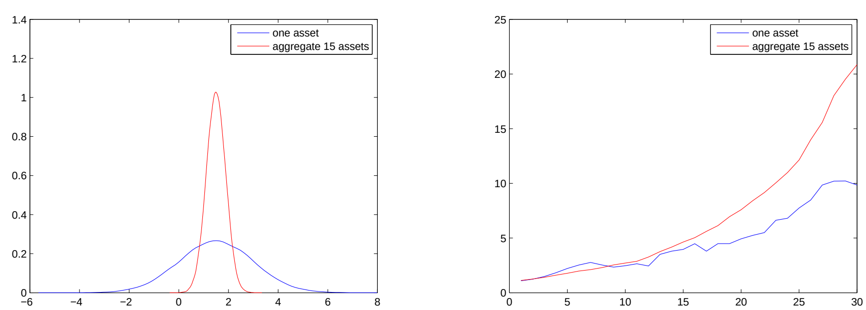
2 Background Knowledge

2.1 Motivation for asset pooling

In order to show how risk can be reduced with diversification, we use a simplified numerical example. Diversification can be defined as spreading a fixed amount of funding on a variety of assets to reduce risk. Its benefit can be justified by two arguments. The convexity of risk measures is the first argument. Assuming we have 15 independently distributed assets each worth \$1, with yearly return distributions modeled by Gaussian distributions of mean 0.1 and standard deviation 0.1. Consider two portfolios A and B: portfolio A consists of \$15 dollars of the first asset, and portfolio B consists of one dollar of each asset. Using Monte Carlo simulation, the distribution of return of the two portfolios can be generated as shown in Panel I in figure 1.

Figure 1: Return distribution of two different portfolios (Panel I)
40 periods cumulative return simulation of two different portfolios (Panel II)

Panel I Panel II



Note that from Figure 1 Panel I, both density distributions center on 1.5. This is because the two portfolios have the same expected return. However, the return distribution of the one asset portfolio (i.e., Portfolio A) is much more spread out than that of the multi-asset portfolio (i.e., Portfolio B). If risk is considered as a measure of dispersion, we can say that the multi-asset portfolio has less risk. Over a longer period of time, the multi-asset portfolio will observe a much smoother growth in value as observed in Panel II of Figure 1. The diversification effect illustrated in this example is widely observed in finance. This describes an incentive for a financial institution or a corporation to construct a portfolio of diversified assets.

This argument explains the benefit of asset pooling, diversification of assets effectively limits risk while keeping expected return the same. This is why insurance companies can establish a business by collecting a fee from each individual and managing the collective risk. Since asset pooling explains the incentive of large financial institutions, the modeling problem on the aggregate assets is a very important issue. However, it is a complex exercise to model the distribution of a pool of different assets, each of which might behave differently. The following sections will further explore the models in more depth. Models in the industry often face the challenge of either making too many assumptions or calibrating too many parameters. The two models proposed in the following sections can solve this problem without the trade-off.

2.2 Modeling aggregate assets

In order to model pools of assets, it is necessary to understand the distribution of each individual asset, and calibrate their dependence structure. The problem with calibrating a high dimensional distribution is that without too many assumptions, many parameters need to be estimated, which decreases the accuracy of the results. Different models have been proposed. The study of copulas is one widely used tool for multidimensional dependence structure. A copula does not depend on the marginal distribution; it is a function of the quantiles of each variable. This claim can be asserted by the following theorem named after the mathematician Sklar (1959). A more formal treatment of copulas is hereafter presented in section 3.

Theorem 2.1 (Sklar's theorem) *(i) A cumulative distribution function,*

$$H(x_1, \ldots, x_d) = \mathbb{P}[X_1 \leqslant x_1, \ldots, X_d \leqslant x_d]$$

of a random vector $(X_1, X_2, \ldots, X_d)$ *with marginals* $F_i(x) = \mathbb{P}[X_i \leqslant x]$ *can be written as*

$$H(x_1, \ldots, x_d) = C\left(F_1(x_1), \ldots, F_d(x_d)\right),$$

where C *is a copula.*

(ii) Given a cumulative distribution function H*, the copula is unique on* $Range(F_1) \times \cdots \times Range(F_d)$*, which is the cartesian product of the Range of*

the marginal cdf's. This implies that the copula is unique if the marginals F_i are continuous.

Sklar's theorem shows that we can divide the modelling multidimensional cumulative distribution into two parts, finding the marginal distribution and defining the copula. In practice, there exist many accurate statistical techniques to estimate the respective marginal distributions of $X_1, ..., X_n$, while the joint dependence structure of $(X_1, X_2, \ldots, X_d)$ is often much more difficult to capture. Therefore, deeper understanding of models defined with copula functions can help us model joint distributions with more sophisticated dependence structure.

In order to avoid calibration of too many parameters, assumptions have to be made about the copula. For 2-dimensional problems, the Gaussian copula used to be a popular tool to price aggregate assets. The advantage of the Gaussian copula is that it is a one parameter model which is easy to calibrate. The convenience of the model made it a popular choice for pricing in the industry. However, the assumption of a Gaussian copula is very specific, and the abuse of this model caused dire consequences in the 2008 financial crisis.

Among all the useful properties copulas have, one property is that there exists a point-wise upper bound for copula in any dimension. This is a special copula named the co-monotonic copula or the Fréchet-Hoeffding upper bound defined as the following,

$$C\left(u_1, u_2, u_3, ..., u_n\right) = \min\left(u_1, u_2, u_3, ..., u_n\right).$$

In two dimensions, copulas also have a Fréchet-Hoeffding lower bound. Studies have been done on improving Fréchet-Hoeffding bounds under specific conditions on the copula function. The existence of such point-wise bounds can provide price bounds on measures of aggregate assets without making any assumptions on the dependence structure. The Fréchet-Hoeffding bounds are convenient but they are too general since they are bounds on the whole copula space. In reality, the dependence relationship between real assets have many idiosyncratic structures; Fréchet-Hoeffding bounds are too wide to accurately capture the distribution of the asset pool. The bounds offer a limit on the asset pool by giving the best and the worst case scenarios of the measure. By understanding and specifying conditions on assets, further improvements on the bounds can be achieved. Section 3 reformats the paper Bernard, Jiang and Vanduffel (2012) discussing different scenarios where the copula bounds can be improved. This study is based on the paper Tankov (2011), whereas the conditions on which the bounds can be improved are broadened.

The copula is a very useful tool in modeling dependence structure. One problem with copula is that there does not exist a Fréchet-Hoeffding lower bound in any space with dimension higher than three. One way to solve the minimization problem in higher dimensions is to use the complete mixability property. Complete mixability property is concerned with the sum distribution of n random variables $X_1,\ X_2,\ X_3, ...,\ X_n$, where each random variable follows defined marginal distributions, $X_1 \sim F_1,\ X_2 \sim F_2,\ X_3 \sim F_3, ...,\ X_n \sim F_n$. If there exist a set of $X_1,\ X_2,\ X_3, ...,\ X_n$ such that the sum is a constant, this set is called a complete mixable set. Notice that as the sum of $X_i's$ is a constant in this setup,

$$S = \sum_{i=1}^{n} X_i = Constant.$$

Notice that since S is constant, the standard deviation $\sigma(S) = 0$ is at its minimum. The Fréchet-Hoeffding lower bound achieves the same minimum in dimension two. In section 3, we prove that for any convex function f, the complete mixable distribution has the constant $S = C$ as the solution to the minimization problem

$$\begin{aligned} &\min \mathbb{E}[f(S)] \\ s.t.\, &X_1 \sim F_1,\ X_2 \sim F_2,\ ...,\ X_n \sim F_n \\ &S = \sum_{i=1}^{n} X_i. \end{aligned}$$

With a bit of careful examination, we can use the complete mixability property on a much wider range of distributions to find lower bounds. We propose and prove three theorems that give the lower convex ordering bound. The first theorem proves the existence of a bound in the case where all marginal distributions are identical. The lower bound provided in this theorem is sharp under some technical conditions. The second theorem gives the lower bound without restrictions on the marginal distributions. As a tradeoff between sharpness and generality, it can be shown that the bound in theorem two, although more general, is not a sharp bound. The third theorem gives the sharp lower bound on *TVaR*. Using the Rearrangement Algorithm as described by Puccetti and Rüchendorf (2012), the bounds are verified numerically for sharpness.

3 Improved Copula Bounds

A copula is a mathematical function that describes all of the information in a correlation structure in a high dimensional distribution. It is widely used in risk and portfolio management for pooled assets and pricing of basket asset derivatives. The method using copula for pricing of collateralized debt obligation (CDOs) was popularized in the early 2000s for its simplicity. It was believed by some that the Gaussian copula model for pricing CDOs was partially responsible for the global financial crisis in 2008-2009. It is true that before and after the crisis, Gaussian copula was recognized to have limitations on extreme tail events (the event where all single random variables have values in their respective lower range). It is a known phenomena that during a market shock and a financial crisis, higher level market co-movement is observed. By assuming Gaussian copula, the probability of tail events decays exponentially, thus it cannot capture this type of systematic risk. This problem can be solved by giving the Fréchet-Hoeffding bound on copulas without making more assumptions or introducing more parameters.

3.1 Definition of Copula

In probability terms, a n-dimensional copula $C : [0,1]^n \to [0,1]$ is a cumulative distribution function on the unit hypercube $[0,1]^n$ with each marginal distribution as the uniform distribution.

Definition 3.1 *An n-dimensional copula is a function* $C : [0,1]^n \to [0,1]$ *with the following properties:*

1. $C(u_1, u_2, ..., u_{i-1}, 0, u_{i+1}, ..., u_n) = 0$, for each $i = 1, 2, ..., n$.
2. $C(1, 1, ..., 1, u, 1, ..., 1) = u$, for each $i = 1, 2, ..., n$.
3. C is $d-$*increasing*, i.e., for each hyper-rectangle

$$B = \prod_{i=1}^{d}[x_i, y_i] \subseteq [0,1]^d$$

its C-volume is non-negative:

$$\int_B \mathrm{d}C(u) = \sum_{\mathbf{z} \in \times_{i=1}^{d}\{x_i, y_i\}} (-1)^{N(\mathbf{z})} C(\mathbf{z}) \geqslant 0,$$

where

$$N(\mathbf{z}) = \#\{k : z_k = x_k\}.$$

This section will only discuss the two dimensional case of the copula, and therefore the third condition of the copula function can be simplified to conditions 3 and 4 in the following definition:

Definition 3.2 *A bivariate copula is a function* $C : [0,1]^2 \to [0,1]$ *with the following properties:*

1. $C(0,u) = C(u,0) = 0 \ \forall u \in (0,1)$.

2. $C(u,1) = C(1,u) = u \ \forall u \in (0,1)$.

3. $C(u,v)$ is non-decreasing in each variable, i.e., $C(u,v_0)$ is non-decreasing function in u and $C(u_0,v)$ a non-decreasing function in v.

4. For each pair of $(u_1,v_1),(u_2,v_2) \in [0,1]^2$, with $u_1 \leqslant u_2$ and $v_1 \leqslant v_2$, we have the following inequality $C(u_2,v_2) - C(u_1,v_2) - C(u_2,v_1) + C(u_1,v_1) \geqslant 0$

As a generalization of copula, we can replace the third property by a weaker assumption to define a quasi-copula as follows,

Definition 3.3 *A bivariate quasi-copula is a function* $C : [0,1]^2 \to [0,1]$ *with the following properties:*

1. $C(0,v) = C(u,0) = 0$, for each $i = 1,2,...,n$.

2. $C(u,1) = C(1,v) = 1$, for each $i = 1,2,...,n$.

3. $C(u,v)$ is non-decreasing in each variable, i.e., for each $u_0, v_0 \in (0,1)$, $C(u,v_0)$ is non-decreasing function in u and $C(u_0,v)$ a non-decreasing function in v.

4. Lipschitz property: $|C(u_2,v_2) - C(u_1,v_1)| \leqslant |u_2 - u_1| + |v_2 - v_1|$ for all $u_1,u_2,v_1,v_2 \in [0,1]$.

3.2 Properties of Copula

Copula is a popular object in the study of dependence structure because it has very nice properties. Here is a list of well-known properties of copulas. Part (1) can be found in Nelsen (2006), the rest of the properties and proofs can be found in Deheuvel (1981). The best summary of the properties can be found from the paper Durrleman et al. (2001),

1. (Existence) Sklar's theorem: Let H be a joint distribution function with margins F and G. Then there exists a copula C such that for all x, y in $\mathbb{R}$,
$$H(x, y) = C(F(x), G(y)).$$

2. (Convexity) The set of all copulas, $\mathcal{C}$ is convex, namely if $A,\ B \in \mathcal{C}$, then $\forall \lambda \in (0,1)\,,\ \lambda A + \lambda B \in \mathcal{C}$. This property can help us generate new copulas via the known copulas.

3. (Compactness) The set of all copulas, $\mathcal{C}$ is compact with any of the following topologies: point-wise convergence, uniform convergence on $[0,1]$, weak convergence of the associated probability measure (Deheuvel (1978)).

4. (Scale-Invariant)If $h_1, ..., h_n$ are strictly monotonic and non-decreasing mappings of $\mathbb{R}$ to itself, any copula function of $(X_1, ..., X_N)$ is also a copula function of $(h_1\left(X_1\right), ..., h_N\left(X_N\right))$.

5. (Convergence in distribution) If $\left\{\mathbf{F}^{(m)}, m \geqslant 1\right\}$ is a sequence of cumulative distribution functions in $\mathbb{R}^N$, the convergence of $\mathbf{F}^{(m)}$ to a distribution function F with continuous margins $\mathbf{F}_n$, as $m \to \infty$, is equivalent to the following two conditions:

   (a) $\forall 1 \leqslant n \leqslant N,\ \mathbf{F}_n^{(m)} \to \mathbf{F}_n$ pointwise.

   (b) if $\mathbf{C}$ is the unique copula function associated to $\mathbf{F}$, and if $\mathbf{C}^{(m)}$ is a copula function associated with $\mathbf{F}_n^{(m)},\ \mathbf{C}^{(m)} \to \mathbf{C}$ (with the weak topology of $\mathcal{C}$).

6. (Lipschitz condition) Every copula C is continuous and satisfies the following inequality
$$\left|C\left(u_1, u_2, u_3, ..., u_n\right) - C\left(v_1, v_2, v_3, ..., v_n\right)\right| \leqslant \sum_{i=1}^{n} |u_i - v_i|$$

7. (Fréchet-Hoeffding bounds) The classical Fréchet-Hoeffding bounds are defined as follows: The Fréchet-Hoeffding lower bound $W(u,v) := \max(0, u+v-1)$. The Fréchet-Hoeffding upper bound $M(u,v) := \min(u,v)$. The classical Fréchet-Hoeffding bounds are absolute point-wise bounds, meaning for any copula C in two dimensional space, the following is always true

$$W(u,v) \geqslant C(u,v) \geqslant M(u,v). \ \ \forall (u,v) \in [0,1]^2.$$

Remark 3.1

(1) The existence property shows that the copula theory can be applied to correlation structures for all possible joint distributions, thus ensuring the wide practicality of this study. **(2)** The convexity of the copula space can help to generate new copula from the existing classes of copulas, i.e. for any two known different copulas where $A,\ B,\ C = \lambda A + (1-\lambda)B,\ for\ \lambda \in (0,1)$, the convex combinations of these copulas can create new classes of copulas for calibration. **(3)** Compactness implies that any continuous function on copula will attain its optima. **(4)** Scale invariance means copula only describes different random variables' ranking relative to each other. Notice that Kendall's τ and Spearman's ρ can be defined as a function of copula inheriting these properties. For example, consider two random X, Y, where $Y = e^X$, since exponent is an increasing function (for any two events ω_1, ω_2, if $X(\omega_1) \leqslant X(\omega_2)$ then $Y(\omega_1) \leqslant Y(\omega_2)$), the copula of the joint distribution (X, Y) is the Fréchet-Hoeffding upper bound. Also as a result, both Kendall's τ and Spearman's ρ on (X, Y) are equal to 1. **(5)** The convergence in distribution is equivalent to the convergence in copulas, this means, the studies of the topology in the copula space can be used to study topological space of joint distributions. **(6)** The Lipschitz condition is a regularity condition on copula analytically, it is a function with many properties such as uniform continuity. **(7)** Fréchet-Hoeffding bounds are point-wise bounds, it could provide solutions to many optimizations on copulas.

3.3 Tankov's improved bound

Tankov (2011) introduced an improved bound on copula with fixed values $Q(a,b)$ defined on a set $(a,b) \in \mathcal{S}$.

We will use the following notations consistent with relevant literature. Let us denote by $\mathcal{S}$ a compact subset of the unit square $[0,1]^2$ and let Q

be a quasi-copula. Let $Q_{\mathcal{S}}$ denote the set of all quasi-copula satisfying the following condition:

$$Q_{\mathcal{S}} = \{C \ a \ quasi-copula | C(u,v) = Q(u,v) \ \ \forall (u,v) \in \mathcal{S}\}.$$

Denote $A^{\mathcal{S},Q}$ as the upper bound and $B^{\mathcal{S},Q}$ the lower bound,

$$\begin{aligned} A^{\mathcal{S},Q}(u,v) &= \min\left\{u, v, \min_{(a,b)\in\mathcal{S}}\{Q(a,b) + (u-a)^{+} + (v-b)^{+}\}\right\}, \quad (3.1) \\ B^{\mathcal{S},Q}(u,v) &= \max\left\{0, u+v-1, \max_{(a,b)\in\mathcal{S}}\{Q(a,b) - (a-u)^{+} - (b-v)^{+}\}\right\}, \end{aligned}$$

where $(u,v) \in [0,1]^2$.

We will say $\mathcal{S}$ is increasing if for any two points $(a_0, b_0), (a_1, b_1) \in \mathcal{S}$, $a_0 \leqslant a_1 \implies b_0 \leqslant b_1$. Symmetrically, if for any two points $(a_0, b_0), (a_1, b_1) \in \mathcal{S}$, $a_0 \leqslant a_1 \implies b_0 \geqslant b_1$, then we will call $\mathcal{S}$ decreasing. Theorem 1 of Tankov (2011) states the following properties.

Theorem 3.1 .

i. *$A_{\mathcal{S},Q}$ and $B_{\mathcal{S},Q}$ are quasi-copulas satisfying*

$$B^{\mathcal{S},Q}(u,v) \leqslant Q'(u,v) \leqslant A^{\mathcal{S},Q}(u,v) \quad \forall (u,v) \in [0,1]^2$$

for every $Q' \in Q_{\mathcal{S}}$ and

$$A^{\mathcal{S},Q}(a,b) = B^{\mathcal{S},Q}(a,b) = Q(a,b)$$

for all $(a,b) \in \mathcal{S}$.

ii. *If the set S is increasing, then $B^{\mathcal{S},Q}$ is a copula; if the set $\mathcal{S}$ is decreasing, then $A^{\mathcal{S},Q}$ is a copula.*

In other words, the above theorem shows that amongst all quasi-copulas Q' coinciding with Q on $\mathcal{S}$, $A^{\mathcal{S},Q}$ (resp. $B^{\mathcal{S},Q}$) is the best possible upper (resp. lower) bound, and thus improves the Fréchet-Hoeffding bounds in particular. Similar improved bounds have been provided in the paper of Rachev and Rüchendorf (1994) and it is also discussed in Section 7.3 in Rachev and Rüchendorf (1998). Furthermore, Tankov (2011) also showed that $A^{\mathcal{S},Q}$ (resp. $B^{\mathcal{S},Q}$) are quasi-copulas and demonstrates that a sufficient condition for $A^{\mathcal{S},Q}$ (resp. $B^{\mathcal{S},Q}$) to be a copula is to suppose that $\mathcal{S}$ is non-increasing (resp. non-decreasing). In this section, we extend this result by showing that when Q is a copula, $A^{\mathcal{S},Q}$ (resp. $B^{\mathcal{S},Q}$) is a copula when $\mathcal{S}$ is a compact set satisfying some additional conditions, namely a "non-increasingness" (resp. "non decreasingness") and a "connectivity" property. For instance, when S is a rectangle then both $A^{\mathcal{S},Q}$ and $B^{\mathcal{S},Q}$ are copulas.

Theorem 1 of Tankov (2011) and our additions to it are of interest in finance. Tankov already demonstrated how his results are instrumental in finding model-free bounds for the prices of some two-asset derivatives. He shows how information embedded in the financial market (such as the price of another two-asset option) translates into extra information about dependence, and thus allows to sharpen the traditional bounds for prices. These improvements on traditional bounds are based on Fréchet-Hoeffding bounds on copulas (where information on dependence is ignored). In this section, we show that the study of optimal investment strategies is intimately connected to finding bounds on their dependence with the so-called stochastic discount factor (pricing kernel or state-price process). In particular, knowing that $B^{\mathcal{S},Q}$ is a copula is useful to determine investment strategies that are optimal for investors with state-dependent constraints, i.e. when they not only care about the distribution of final wealth but also about the states where cash-flows are received. More details are given in Section 3.6. Both mentioned applications make clear that it is of interest to know more situations for which the bounds appearing in Tankov (2011) are copulas (Theorems 3.3, 3.4 and 3.6 in this section). The main part of the next chapter is largely quoted from a paper by Bernard, Jiang and Vanduffel (2012). Some new developments are presented at the end of the chapter with a new theorem and a brief discussion of a numerical method to generate improved Fréchet-Hoeffding bounds.

3.4 Extensions of Theorem 1 of Tankov (2011)

In this section, we extend Theorem 1 of Tankov (2011). To this end, we need the following lemma.

Lemma 3.2 *Assume $f : [0,1]^2 \mapsto \mathbb{R}$ is two-increasing, non-decreasing in each argument and satisfies the Lipschitz property. Define function $g : [0,1]^2 \mapsto$*

$\mathbb{R}$ *as*

$$g = \max\{f,\ W\},$$

where $W(u,w) = \max\{u+v-1, 0\}$ *is the anti-monotonic copula. Then* g *is also two-increasing, non-decreasing in each argument and satisfies the Lipschitz property.*

Proof. Note that W is Lipschitz continuous. Hence g, as the maximum of two functions with the Lipschitz property, also satisfies the Lipschitz property (as shown in Part (i) of the proof of Theorem 1 in Tankov (2011)). It is obvious that g is also non-decreasing in each argument. In order to prove that g is two-increasing, let us consider any rectangle $R = [u_1, u_2] \times [v_1, v_2]$ [1] in the unit square. We identify the following three cases:

Case 1: Assume that either $\max\{f(u_2, v_1), u_2 + v_1 - 1\} \leqslant 0$ or $\max\{f(u_1, v_2), u_1 + v_2 - 1\} \leqslant 0$. Since both functions are non-decreasing in each argument we find that

$$\max\{f(u_1, v_1), u_1 + v_1 - 1\} \leqslant 0.$$

Without loss of generality we can take $\max\{f(u_2, v_1), u_2 + v_1 - 1\} \leqslant 0$ (the other case is similar). Then the $g-$volume of the rectangle R is given by

$$\begin{aligned} V_g(R) &= g(u_2, v_2) - g(u_2, v_1) - g(u_1, v_2) + g(u_1, v_1) \\ &\geqslant g(u_2, v_2) - g(u_1, v_2) \\ &= \max\{f(u_2, v_2),\ W(u_2, v_2)\} - \max\{f(u_1, v_2),\ W(u_1, v_2)\} \\ &\geqslant 0, \end{aligned}$$

where the last inequality follows from $f(u_2, v_2) \geqslant f(u_1, v_2)$ and $W(u_2, v_2) \geqslant W(u_1, v_2)$.

For cases 2 and 3, we can now assume that both $\max\{f(u_2, v_1), u_2 + v_1 - 1\} > 0$ and $\max\{f(u_1, v_2), u_1 + v_2 - 1\} > 0$.

Case 2: Assume that $f(u_2, v_1) \geqslant u_2 + v_1 - 1$ and $f(u_1, v_2) \geqslant u_1 + v_2 - 1$. This implies that $g(u_2, v_1) = f(u_2, v_1)$ and $g(u_1, v_2) = f(u_1, v_2)$. Hence the $g-$volume of the rectangle R satisfies

$$\begin{aligned} V_g(R) &= g(u_2, v_2) - g(u_2, v_1) - g(u_1, v_2) + g(u_1, v_1) \\ &\geqslant f(u_2, v_2) - f(u_2, v_1) - f(u_1, v_2) + f(u_1, v_1) \\ &\geqslant 0, \end{aligned}$$

[1] For any rectangle $R = [u_1, u_2] \times [v_1, v_2]$, we conventionally assume $u_1 < u_2$ and $v_1 < v_2$.

where the last inequality follows from the two-increasing property for f.

Case 3: Assume that $f(u_2, v_1) < u_2 + v_1 - 1$ or $f(u_1, v_2) < u_1 + v_2 - 1$, without loss of generality we take $f(u_2, v_1) < u_2 + v_1 - 1$ (the other case is similar). Since $\max\{f(u_2, v_1), u_2 + v_1 - 1\} > 0$ (assumption in Case 1), it follows that $u_2 + v_1 - 1 > 0$ and thus also $u_2 + v_2 - 1 > 0$. Furthermore, the Lipschitz property for f then also implies that $f(u_2, v_2) < u_2 + v_2 - 1$. Therefore

$$\begin{aligned}
V_g(R) &= g(u_2, v_2) - g(u_2, v_1) - g(u_1, v_2) + g(u_1, v_1) \\
&= (u_2 + v_2 - 1) - (u_2 + v_1 - 1) - g(u_1, v_2) + g(u_1, v_1) \\
&= (v_2 - v_1) - (g(u_1, v_2) - g(u_1, v_1)) \\
&\geqslant 0,
\end{aligned}$$

where the last inequality follows from the Lipschitz property for g. ■

Let us denote by $\mathcal{S}_1$ the set obtained by the first variable projection of the compact set $\mathcal{S}$, namely $u \in \mathcal{S}_1$ if and only if there exists $v \in (0, 1)$ such that $(u, v) \in \mathcal{S}$. Similarly, we define $\mathcal{S}_2$ as the second variable projection. Define the two following functions

$$\begin{aligned}
\gamma_1 : \mathcal{S}_1 &\to \mathcal{S}_2 \\
u &\mapsto \min\{v | (u, v) \in \mathcal{S}\}
\end{aligned} \tag{3.2}$$

and

$$\begin{aligned}
\gamma_2 : \mathcal{S}_1 &\to \mathcal{S}_2 \\
u &\mapsto \max\{v | (u, v) \in \mathcal{S}\},
\end{aligned} \tag{3.3}$$

The existence of the above maxima and minima is guaranteed because of the compactness of $\mathcal{S}$. The points $(u, \gamma_1(u))$ are the "lower" boundary points of $\mathcal{S}$. Similarly, $(u, \gamma_2(u))$ are the "upper" boundary points. We are now ready to prove the following result.

Theorem 3.3 *Let Q be a copula and $\mathcal{S} \subseteq [0, 1]^2$ be a compact set with both γ_1 and γ_2 as non-decreasing functions, and satisfying the following property:*

$$\forall (u, v_0), (u, v_1) \in \mathcal{S}, \ \left(u, \frac{v_0 + v_1}{2}\right) \in \mathcal{S}. \tag{3.4}$$

Then $B^{\mathcal{S},Q}$ is a copula.

Proof. Tankov (2011) already showed that $B^{\mathcal{S},Q}$ is a quasi-copula. Thus we only need to show that $B^{\mathcal{S},Q}$ is two-increasing. Let us write $B^{\mathcal{S},Q}$ as $\max\left\{f^{\mathcal{S},Q}, W\right\}$, where $f^{\mathcal{S},Q}$ is the function

$$f^{\mathcal{S},Q}(u,v) := \max_{(a,b)\in\mathcal{S}}\{Q(a,b) - (a-u)^+ - (b-v)^+\}.$$

Tankov (2011) proved that $f^{\mathcal{S},Q}$ satisfies the Lipschitz condition. Since $f^{\mathcal{S},Q}$ is also non-decreasing in each argument, it remains to prove that it is also two-increasing. Then, Lemma 3.2 implies that $B^{\mathcal{S},Q}$ is a two-increasing quasi-copula and therefore a copula. Let us consider any rectangular area $R = [u_1,u_2]\times[v_1,v_2]$. We want to prove that $V_{f^{\mathcal{S},Q}}([u_1,v_1]\times[u_2,v_2]) \geqslant 0$.

By compactness of $\mathcal{S}$, there exist $(u_1^*,v_2^*)\in\mathcal{S}$ and $(u_2^*,v_1^*)\in\mathcal{S}$, such that

$$\begin{aligned} f^{\mathcal{S},Q}(u_1,v_2) &= \max_{(a,b)\in\mathcal{S}}\left\{Q(a,b) - (a-u_1)^+ - (b-v_2)^+\right\} \\ &= Q(u_1^*,v_2^*) - (u_1^*-u_1)^+ - (v_2^*-v_2)^+, \end{aligned}$$

and

$$\begin{aligned} f^{\mathcal{S},Q}(u_2,v_1) &= \max_{(a,b)\in\mathcal{S}}\left\{Q(a,b) - (a-u_2)^+ - (b-v_1)^+\right\} \\ &= Q(u_2^*,v_1^*) - (u_2^*-u_2)^+ - (v_1^*-v_1)^+. \end{aligned}$$

Case 1: First, we assume that (u_1^*,v_2^*) and (u_2^*,v_1^*) form a non-decreasing set. Observe that

$$\begin{aligned} f^{\mathcal{S},Q}(u_2,v_2) &= \max_{(a,b)\in\mathcal{S}}\left\{Q(a,b) - (a-u_2)^+ - (b-v_2)^+\right\} \\ &\geqslant Q(u_1^*,v_2^*) - (u_1^*-u_2)^+ - (v_2^*-v_2)^+, \end{aligned}$$

and

$$\begin{aligned} f^{\mathcal{S},Q}(u_1,v_1) &= \max_{(a,b)\in\mathcal{S}}\left\{Q(a,b) - (a-u_1)^+ - (b-v_1)^+\right\} \\ &\geqslant Q(u_2^*,v_1^*) - (u_2^*-u_1)^+ - (v_1^*-v_1)^+. \end{aligned}$$

Then we bound the volume of the rectangle $[u_1,v_1]\times[u_2,v_2]$ as follows:

$$\begin{aligned} & f^{\mathcal{S},Q}(u_2,v_2) - f^{\mathcal{S},Q}(u_1,v_2) - f^{\mathcal{S},Q}(u_2,v_1) + f^{\mathcal{S},Q}(u_1,v_1) \\ &\geqslant \left[Q(u_1^*,v_2^*) - (u_1^*-u_2)^+ - (v_2^*-v_2)^+\right] - \left[Q(u_1^*,v_2^*) - (u_1^*-u_1)^+ - (v_2^*-v_2)^+\right] \\ &\qquad - \left[Q(u_2^*,v_1^*) - (u_2^*-u_2)^+ - (v_1^*-v_1)^+\right] + \left[Q(u_2^*,v_1^*) - (u_2^*-u_1)^+ - (v_1^*-v_1)^+\right] \\ &= Q(u_1^*,v_2^*) - Q(u_1^*,v_2^*) - Q(u_2^*,v_1^*) + Q(u_2^*,v_1^*) \\ &\qquad + \left[(u_1^*-u_1)^+ - (u_1^*-u_2)^+\right] - \left[(u_2^*-u_1)^+ - (u_2^*-u_2)^+\right] \\ &\geqslant \left[(u_1^*-u_1)^+ - (u_1^*-u_2)^+\right] - \left[(u_2^*-u_1)^+ - (u_2^*-u_2)^+\right] \qquad (3.5) \end{aligned}$$

where the last inequality (3.5) holds because Q is two-increasing and (u_1^*, v_2^*) and (u_2^*, v_1^*) form a non-decreasing set. Hence, if $u_1^* \leqslant u_2^*$ in (3.5), then the volume $V_{f^{\mathcal{S},Q}}([u_1, v_1] \times [u_2, v_2]) \geqslant 0$ holds true. In the opposite case ($u_2^* > u_1^*$), we proceed similarly. Indeed, it also holds that

$$f^{\mathcal{S},Q}(u_2, v_2) \geqslant Q(u_2^*, v_1^*) - (u_2^* - u_2)^+ - (v_1^* - v_2)^+,$$

$$f^{\mathcal{S},Q}(u_1, v_1) \geqslant Q(u_1^*, v_2^*) - (u_1^* - u_1)^+ - (v_2^* - v_1)^+.$$

Therefore using the same proof as above, we obtain

$$\begin{aligned} & f^{\mathcal{S},Q}(u_2, v_2) - f^{\mathcal{S},Q}(u_1, v_2) - f^{\mathcal{S},Q}(u_2, v_1) + f^{\mathcal{S},Q}(u_1, v_1) \\ & \geqslant \left[(u_2^* - u_1)^+ - (u_2^* - u_2)^+\right] - \left[(u_1^* - u_1)^+ - (u_1^* - u_2)^+\right] \\ & \geqslant 0. \end{aligned}$$

Case 2: Second, we assume (u_1^*, v_2^*) and (u_2^*, v_1^*) form a non-increasing set.

When $u_1^* \leqslant u_2^*$, then $v_2^* \geqslant v_1^*$. By compactness of $\mathcal{S}$, property (3.4) implies that for each $u \in \mathcal{S}_1$, $\mathcal{S}$ contains the vertical segment connecting $(u, \gamma_1(u))$ and $(u, \gamma_2(u))$. Thus $\gamma_1(u_1^*) \leqslant v_2^* \leqslant \gamma_2(u_1^*)$ and $\gamma_1(u_2^*) \leqslant v_1^* \leqslant \gamma_2(u_2^*)$. Moreover, by the non-decreasing property of γ_1 and γ_2, we have $\gamma_1(u_1^*) \leqslant \gamma_1(u_2^*)$ and $\gamma_2(u_1^*) \leqslant \gamma_2(u_2^*)$. Therefore $\gamma_1(u_1^*) \leqslant v_1^* \leqslant \gamma_2(u_1^*)$ and $\gamma_1(u_2^*) \leqslant v_2^* \leqslant \gamma_2(u_2^*)$. Hence $(u_1^*, v_1^*) \in \mathcal{S}$ and $(u_2^*, v_2^*) \in \mathcal{S}$. Similarly, we can prove that when $u_2^* < u_1^*$, $(u_1^*, v_1^*) \in \mathcal{S}$ and $(u_2^*, v_2^*) \in \mathcal{S}$.

We obtain that, for (u_1, v_1) and (u_2, v_2),

$$f^{\mathcal{S},Q}(u_1, v_1) \geqslant Q(u_1^*, v_1^*) - (u_1^* - u_1)^+ - (v_1^* - v_1)^+,$$

and

$$f^{\mathcal{S},Q}(u_2, v_2) \geqslant Q(u_2^*, v_2^*) - (u_2^* - u_2)^+ - (v_2^* - v_2)^+.$$

We can then conclude that the volume of the rectangle $[u_1, v_1] \times [u_2, v_2]$ is non-negative because

$$\begin{aligned} & f^{\mathcal{S},Q}(u_2, v_2) - f^{\mathcal{S},Q}(u_1, v_2) - f^{\mathcal{S},Q}(u_2, v_1) + f^{\mathcal{S},Q}(u_1, v_1) \\ & \geqslant \left[Q(u_2^*, v_2^*) - (u_2^* - u_2)^+ - (v_2^* - v_2)^+\right] - \left[Q(u_1^*, v_2^*) - (u_1^* - u_1)^+ - (v_2^* - v_2)^+\right] \\ & \qquad - \left[Q(u_2^*, v_1^*) - (u_2^* - u_2)^+ - (v_1^* - v_1)^+\right] + \left[Q(u_1^*, v_1^*) - (u_1^* - u_1)^+ - (v_1^* - v_1)^+\right] \\ & = Q(u_2^*, v_2^*) - Q(u_1^*, v_2^*) - Q(u_2^*, v_1^*) + Q(u_1^*, v_1^*) \\ & \geqslant 0. \end{aligned}$$

We have proved that $f^{\mathcal{S},Q}$ is two-increasing. Lemma 3.2 implies that $B^{\mathcal{S},Q}$ is a copula. ■

Define the two following functions which define the "left" and "right" boundary points of $\mathcal{S}$.

$$\begin{aligned}\gamma_3 : \mathcal{S}_2 &\to \mathcal{S}_1 \\ v &\mapsto \min\{u|(u,v) \in \mathcal{S}\}\end{aligned} \tag{3.6}$$

and

$$\begin{aligned}\gamma_4 : \mathcal{S}_2 &\to \mathcal{S}_1 \\ v &\mapsto \max\{u|(u,v) \in \mathcal{S}\}\end{aligned} \tag{3.7}$$

The following result is dual to Theorem 3.3. The proof is obtained by symmetry.

Theorem 3.4 *Let Q be a copula and $\mathcal{S} \subseteq [0,1]^2$ be a compact set with both γ_3 and γ_4 being non-decreasing functions and satisfying the following property:*

$$\forall (u_0, v), (u_1, v) \in \mathcal{S}, \ \left(\frac{u_0 + u_1}{2}, v\right) \in \mathcal{S}. \tag{3.8}$$

Then $B^{\mathcal{S},Q}$ is a copula.

Remark 3.2

The conditions in Theorems 3.3 and 3.4 cannot be readily relaxed. Indeed consider $\mathcal{S} = \{A, B, C, D\}$, where $A = (1/3, 0)$, $B = (1/3, 2/3)$, $C = (2/3, 1/3)$, $D = (2/3, 1)$ and let $Q = \min\{u, v\}$. Note that property (3.4) is not satisfied and also that γ_3 (as well as γ_4) is not non-decreasing, so that neither Theorem 3.3 nor Theorem 3.4 can be invoked to show that $B^{\mathcal{S},Q}$ is a copula. We observe that $B^{\mathcal{S},Q}$ is not a copula indeed, because

$$\begin{aligned}&B^{\mathcal{S},Q}\left(\frac{2}{3},\frac{2}{3}\right) - B^{\mathcal{S},Q}\left(\frac{2}{3},\frac{1}{3}\right) - B^{\mathcal{S},Q}\left(\frac{1}{3},\frac{2}{3}\right) + B^{\mathcal{S},Q}\left(\frac{1}{3},\frac{1}{3}\right) \\ =&\frac{1}{3} - \frac{1}{3} - \frac{1}{3} + 0 \\ =&-\frac{1}{3}.\end{aligned}$$

Remark 3.3

At first glance, part(ii) of Theorem 1 in Tankov (2011) does not appear to always follow from Theorem 3.3 or Theorem 3.4. For example, let us consider the compact set $\mathcal{S} = \{A, B, C\}$ where $A = (x_1, y_1)$, $B = (x_2, y_1)$, $C = (x_2, y_2)$, and where $x_1 < x_2$ and $y_1 < y_2$. Then neither property (3.4) nor property (3.8) is satisfied. Nevertheless, $\mathcal{S}$ is a non-decreasing compact set for which part(ii) of Theorem 1 in Tankov (2011) can be applied implying that $B^{\mathcal{S},Q}$ is a copula. However, we can also use our results combined with a limiting argument to obtain the same result. Consider $\mathcal{S}_n = \{A, B_n, C\}$, where $A = (x_1, y_1)$, $B_n = (x_2 - \frac{x_2 - x_1}{n}, y_1)$ and $C = (x_2, y_2)$. Then using Theorem 3.3, we have that for all positive $n \in \mathbb{N}$, $B^{\mathcal{S}_n,Q}$ is a copula. Moreover,

$$\begin{aligned} B^{\mathcal{S}_n,Q}(u,v) = \max \Big\{ & 0, u+v-1, Q(x_1,y_1) - (x_1-u)^+ - (y_1-v)^+, \\ & Q(x_2,y_2) - (x_2-u)^+ - (y_2-v)^+, \\ & Q\left(x_2 - \frac{x_2-x_1}{n}, y_1\right) - \left(x_2 - \frac{x_2-x_1}{n} - u\right)^+ - (y_1-v)^+ \Big\} \end{aligned}$$

converges point-wise for all $(u, v) \in [0, 1]^2$ to $B^{\mathcal{S},Q}(u, v)$. Finally, to prove that $B^{\mathcal{S},Q}$ is a copula we need to verify the boundary conditions and the two-increasing property. Both elements are satisfied when the sequence converges point-wise[2]. Therefore, $B^{\mathcal{S},Q}$ is a copula. The same limiting arguments can be used to show that $B^{\mathcal{S},Q}$ is a copula when S is a non-decreasing compact that contains a vertical part and a horizontal part that are both disconnected (so that both (3.4) and (3.8) are not satisfied). In this sense, the results in Tankov (2011) appear as a special case of ours.

Corollary 3.5 *Let Q be a copula and $\mathcal{S} \subseteq [0, 1]^2$ be a compact convex set satisfying*

$$\exists\,(a_0, b_0) \in \mathcal{S}, \exists\,(a_1, b_1) \in \mathcal{S}, \forall (u, v) \in \mathcal{S},\ a_0 \leqslant u \leqslant a_1,\ b_0 \leqslant v \leqslant b_1. \quad (3.9)$$

Then $B^{\mathcal{S},Q}$ is a copula.

[2] In fact, if the point-wise limit of a sequence of copulas exists at each point of $[0, 1]^2$ then the limit must be a copula (see comment of Nelsen (2006) page 97 after definition 3.3.4).

Proof. We prove that γ_1 and γ_2 are non-decreasing on (a_0, a_1) and apply Theorem 3.3 (since property (3.4) is satisfied). Indeed, by convexity of $\mathcal{S}$, for any two points $(x_1, \gamma_2(x_1))$ and $(x_2, \gamma_2(x_2))$, we have $\left(\frac{x_1+x_2}{2}, \frac{\gamma_2(x_1)+\gamma_2(x_2)}{2}\right) \in \mathcal{S}$. We can conclude that

$$\left(\frac{\gamma_2(x_1)+\gamma_2(x_2)}{2}\right) \leqslant \max\left\{v \,\middle|\, \left(\frac{x_1+x_2}{2}, v\right) \in \mathcal{S}\right\} = \gamma_2\left(\frac{x_1+x_2}{2}\right),$$

thus γ_2 is concave. Similarly, γ_1 is convex. Finally, since γ_2 is concave, $R(x_1, x_2) = \frac{\gamma_2(x_2)-\gamma_2(x_1)}{x_2-x_1}$ is non-increasing in x_1 for x_2 fixed, and in x_2 for x_1 fixed. Therefore γ_2 is non-decreasing on $[a_0, a_1]$ because of property (3.9). A similar reasoning shows that γ_1 is also non-decreasing. ■

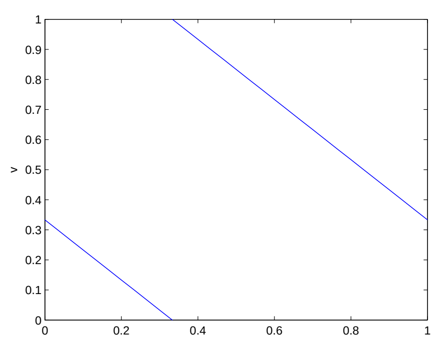
Note that Corollary 3.5 is not valid when the compact $\mathcal{S}$ is simply convex and compact as shown by the following example.

Example 3.1 *Let $\mathcal{S}$ be the line connecting $\left(\frac{1}{3}, \frac{2}{3}\right)$ to $\left(\frac{2}{3}, \frac{1}{3}\right)$. Let Q be the copula defined by the support in Figure 2, namely,*

$$Q(u,v) = \begin{cases} \max\left\{u + \min\left\{v, \frac{1}{3}\right\} - \frac{1}{3}, 0\right\}, & v \in \left[0, \frac{1}{3}\right] \\ \max\left\{u + v - 1, \min\left\{v, \frac{1}{3}\right\}\right\}, & v \in \left[\frac{1}{3}, 1\right]. \end{cases}$$

It can be easily shown that Q takes the constant value $\frac{1}{3}$ on $\mathcal{S}$. Observe that $B^{\mathcal{S},Q}\left(\frac{2}{3}, \frac{2}{3}\right) = \frac{1}{3}$, $B^{\mathcal{S},Q}\left(\frac{1}{3}, \frac{2}{3}\right) = \frac{1}{3}$, $B^{\mathcal{S},Q}\left(\frac{2}{3}, \frac{1}{3}\right) = \frac{1}{3}$, $B^{\mathcal{S},Q}\left(\frac{1}{3}, \frac{1}{3}\right) = 0$. Therefore, on the rectangle $\left[\frac{1}{3}, \frac{2}{3}\right]^2$, $B^{\mathcal{S},Q}$ is not two-increasing.

Figure 2: Support of the copula in Example 3.1



However Theorems 3.3 and 3.4 show that there is a wide class of convex compact sets such that $B^{\mathcal{S},Q}$ given by (3.1) is a copula. Similar results can be obtained for $A^{\mathcal{S},Q}$.

Theorem 3.6 *Let Q be a copula and $\mathcal{S} \subseteq [0,1]^2$ be a compact set. (i) If γ_1 and γ_2 are non-increasing functions and $\mathcal{S}$ satisfies (3.4) then $A^{\mathcal{S},Q}$ is a copula. (ii) If γ_3 and γ_4 are non-increasing functions and $\mathcal{S}$ satisfies (3.8) then $A^{\mathcal{S},Q}$ is a copula.*

Proof. (i) Similarly as in the proof of Theorem 1 of Tankov (2011) we note that $A^{\mathcal{S},Q}(u,v) = \overline{B^{\overline{\mathcal{S}},\overline{Q}}}(u,v)$ where $\overline{\mathcal{S}}$ is defined as $\overline{\mathcal{S}} = \{(a,b) \mid (a,1-b) \in \mathcal{S}\}$ and $\overline{Q}(u,v) = u - Q(u,1-v)$. The non-increasing property of γ_1 and γ_2 (defined on $\mathcal{S}$) implies that $\overline{\gamma}_1$ and $\overline{\gamma}_2$ (defined in an obvious way on $\overline{\mathcal{S}}$) are non-decreasing. Since $\overline{Q}$ is a copula the first part of the proof implies that $B^{\overline{\mathcal{S}},\overline{Q}}(u,v)$ is copula, hence $A^{\mathcal{S},Q}(u,v) = \overline{B^{\overline{\mathcal{S}},\overline{Q}}}(u,v)$ is also a copula. The proof for (ii) is similar. ■

As an immediate result of Theorems 3.3 and 3.6, we have also the following corollary.

Corollary 3.7 *For any copula Q and any rectangle $\mathcal{S} = [u_1,u_2] \times [v_1,v_2]$ in the unit square, $A^{\mathcal{S},Q}$ and $B^{\mathcal{S},Q}$ are both copulas.*

Proof. For a rectangle $\mathcal{S}$, γ_1 and γ_2 (as defined by (3.2) and (3.3)) are clearly non-decreasing and non-increasing and property (3.4) is obviously satisfied. Therefore using Theorems 3.3 and 3.6, $A^{\mathcal{S},Q}$ and $B^{\mathcal{S},Q}$ are both copulas. ■

3.5 New development on Improved Bounds

The above chapter found a few conditions on $\mathcal{S}$ that imply $A^{\mathcal{S},Q}$ and $B^{\mathcal{S},Q}$ are copulas. Different conditions on $\mathcal{S}$ with the same implication have been discovered since the publication of Bernard, Jiang and Vanduffel (2012). Bernard et. al. (2012) also expanded on conditions with only constraints on the four points (a_0,b_0), (a_0,b_1), (a_1,b_0), (a_1,b_1).

The following paragraphs also introduce new conditions on $\mathcal{S}$ that can define $A^{\mathcal{S},Q}$ and $B^{\mathcal{S},Q}$ as copulas. There seems to be duality in terms of the region $\mathcal{S}$ that allows $A^{\mathcal{S},Q}$ and $B^{\mathcal{S},Q}$ to be copulas. Let us denote$\mathcal{S}^{\complement} = [0,1]^2 \setminus \mathcal{S}$. Numerical results seem to suggest that the class of $A^{\mathcal{S},Q}$ and $B^{\mathcal{S},Q}$ we identified as copulas, $A^{\mathcal{S}^{\complement},Q}$ and $B^{\mathcal{S}^{\complement},Q}$ are copula as well.

Theorem 3.8 (reshuffle theorem) *For any copula Q and any rectangle $\mathcal{S} = [u_1, u_2] \times [v_1, v_2]$ in the unit square, $A^{\mathcal{S}^{\complement},Q}$ and $B^{\mathcal{S}^{\complement},Q}$ are both copulas.*

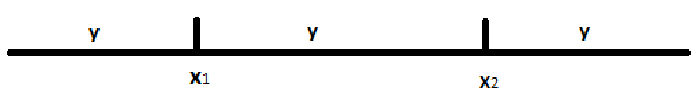
Proof. Again, to show that $A^{\mathcal{S}^{\complement},Q}$ is a copula, we only have to discuss one case where $(a_1, b_1), (a_2, b_2) \in [u_1, u_2] \times [v_1, v_2]$. All other cases can be reduced to the first case. Heuristically, for any two points $(a_1, b_1), (a_2, b_2) \in [u_1, u_2] \times [v_1, v_2]$ with $a_1 < a_2$ and $b_1 < b_2$, $A^{\mathcal{S}^{\complement},Q}(a_1, b_1) = \min\{Q(a_1, v_2), Q(u_2, b_1)\}$, therefore

$$\begin{aligned}
&A^{\mathcal{S}^{\complement},Q}(a_2, b_2) - A^{\mathcal{S}^{\complement},Q}(a_1, b_2) - A^{\mathcal{S}^{\complement},Q}(a_2, b_1) + A^{\mathcal{S}^{\complement},Q}(a_1, b_1) \\
=&\min\{Q(a_2, v_2), Q(u_2, b_2)\} - \min\{Q(a_1, v_2), Q(u_2, b_2)\} \\
&\quad - \min\{Q(a_2, v_2), Q(u_2, b_1)\} + \min\{Q(a_1, v_2), Q(u_2, b_1)\} \\
=&\min\{Q(a_2, v_2), Q(u_2, b_2)\} - \min\{Q(a_1, v_2), Q(u_2, b_2)\} \\
&\quad - (\min\{Q(a_2, v_2), Q(u_2, b_1)\} - \min\{Q(a_1, v_2), Q(u_2, b_1)\}) \\
=&(\min\{x_2, y_2\} - \min\{x_1, y_2\}) - (\min\{x_2, y_1\} - \min\{x_1, y_1\}).
\end{aligned}$$

Using the figure below, we can show that the function $f(y) = \min\{x_2, y\} - \min\{x_1, y\}$ is an increasing function. The function $f(y)$ can be redefined as follows,

$$f(y) = \begin{cases} x_2 - x_1 & for\ y > x_2 \\ x_2 - y & for\ y \in [x_1, x_2] \\ 0 & for\ y < x_1 \end{cases}$$

Figure 3: Interval Proof



Therefore, we have the following equation:

$$\min\{x_2, y_2\} - \min\{x_1, y_2\} \geqslant \min\{x_2, y_1\} - \min\{x_1, y_1\}.$$

We can conclude that

$$A^{\mathcal{S}^{\complement},Q}(a_2, b_2) - A^{\mathcal{S}^{\complement},Q}(a_1, b_2) - A^{\mathcal{S}^{\complement},Q}(a_2, b_1) + A^{\mathcal{S}^{\complement},Q}(a_1, b_1) \geqslant 0.$$

By symmetry, we can also show $B^{\mathcal{S}^{\complement},Q}$ is a copula. ■

Remark 3.4 (numerical solutions for improved bounds)

Numerically, the idea of this theorem is very easy to understand. Imagine we numerically generate a copula Q, we can proceed to do the following to find the improved Fréchet-Hoeffding bounds:

1. we select all points $\{(a_i, b_i)\}_i \in [u_1, u_2] \times [v_1, v_2]$.
2. replace $\{(a_i, b_i)\}_i$ with $\{(a_{\sigma_1(i)}, b_{\sigma_2(i)})\}_i$ where $\sigma_1(i), \sigma_2(i)$ are both sorting algorithms, meaning,

$$\sigma_1(i) < \sigma_1(j) \implies a_{\sigma_1(i)} \leqslant a_{\sigma_1(j)}.$$

This implies that the new copula generated with $\{(a_{\sigma_1(i)}, b_{\sigma_2(i)})\}_i$ follows the co-monotonic structure on the rectangle $[u_1, u_2] \times [v_1, v_2]$ and everywhere else, it still has the copula value of Q. This numerical solution can offer new ways to generate improved Fréchet-Hoeffding bounds with easy visualization aspect. An example is shown in figure 4, the first copula is generated by two random variables X, Y, where X follows the Pareto(1,1) distribution and Y follows the uniform distribution on $[0, 1]$. Define $\mathcal{S} = [0.4, 0.6] \times [0, 0.2]$, Panel(B) gives the Fréchet-Hoeffding upper bound $A^{\mathcal{S},Q}$, and Panel(C) gives the Fréchet-Hoeffding lower bound $B^{\mathcal{S},Q}$. Based on the numerical method, we can also conjecture the following statement:

Conjecture 3.9 *Let Q be a copula and $\mathcal{S} \subseteq [0,1]^2$ be a compact set with both γ_1 and γ_2 being non-decreasing functions. Then $A^{\mathcal{S}^{\complement},Q}$ is a copula.*

Remark 3.5 (Ideas on numerical proof of the conjecture)

Numerically, for any two points $(u_1, v_2), (u_2, v_1) \in \mathcal{S}$ with $u_1 < u_2,\ v_1 < v_2$, since γ_1 and γ_2 are non-decreasing functions, we have $\gamma_2(u_2) \geqslant \gamma_2(u_1) \geqslant v_2$, $\gamma_1(u_2) \leqslant v_1 < v_2$. We can therefore conclude that $(u_2, v_2) \in \mathcal{S}$. Similarly the inequalities $\gamma_1(u_1) \leqslant \gamma_1(u_2) \leqslant v_1$, $\gamma_2(u_1) \geqslant v_2 > v_2$ show that $(u_1, v_1) \in \mathcal{S}$. The rearrangement solution in the remark 3.5 can be decomposed into a sequence of rearrangement switching the above pair $(u_1, v_2), (u_2, v_1) \to (u_1, v_1), (u_2, v_2)$. Therefore, the result of the rearrangement algorithm will stay in $\mathcal{S}$, thereby giving a new copula C with $C(u, v) = Q(u, v)\ \forall\ (u, v) \in \mathcal{S}^{\complement}$ and $C(u, v)$ maximized over the $\mathcal{S}$ region. This is identical to $A^{\mathcal{S},Q}$.

Example 3.2 (Numerical Solution)

In the following figure 4, the Q distribution is generated using two independent random variables X, Y and the correlation coefficient ρ. X follows a Pareto(1,1) distribution and Y follows a uniform distribution. Let us denote the new random variable $Y^{'} = \rho X + \sqrt{1-\rho^2}Y$. The multidimensional distribution is described by

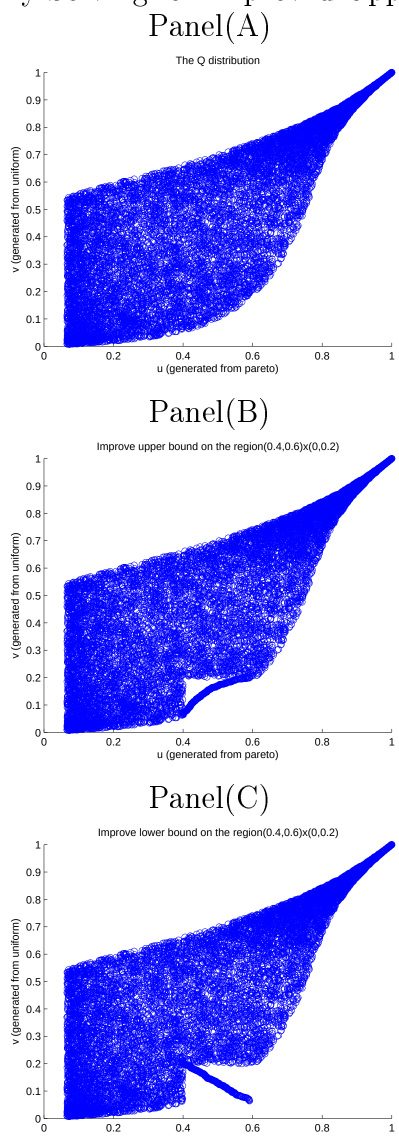
$$\begin{aligned} F(x,y) &= \mathbb{P}\left[X \leqslant x,\ \rho X + \sqrt{1-\rho^2}Y \leqslant y\right] \\ &= \mathbb{P}\left[X \leqslant x,\ Y^{'} \leqslant y\right]. \end{aligned}$$

Denote the cumulative distribution function of X as F_X and the cumulative distribution function of $Y^{'}$ as $F_{Y'}$. From the multidimensional distribution, its corresponding $Q(u,v)$ can be generated as follows:

$$Q(u,v) = \mathbb{P}\left[F_X(X) \leqslant u,\ F_{Y'}\left(Y^{'}\right) \leqslant v\right].$$

Let us define $\mathcal{S} = [0.4, 0.6] \times [0, 0.2]$. If the numerical remark 3.5 is correct, then in the following figure 4, Panel(B) and Panel(C) offer a solution to the problem of finding the improved bounds.

Figure 4: Numerically Solving for Improved Upper and Lower Bound



3.6 Application to optimal investment strategies and derivative pricing

Tankov (2011) also described the broad application of copula bounds. Consider a derivative payoff scheme $f(X, Y)$ based the final return of two assets

X *and* Y, if the marginal distribution of the two asset returns are given, then the price of the derivative $\pi = \mathbb{E}_{\mathbb{Q}}\left[f\left(X, Y\right)\right]$ is a function of the copula C of X and Y.

$$\begin{aligned}\pi(C) &= \int_0^{\infty}\int_0^{\infty} f(x,y)dC\left(F_X(x), F_Y(y)\right)\\ &= \int_0^{\infty}\int_0^{\infty} f(F_X^{-1}\left(x\right), F_Y^{-1}\left(y\right))dC\left(x,y\right).\end{aligned}$$

Assume f is always two-increasing, then π is non-decreasing with respect to the concordance order of copulas, i.e., if for copulas A and B, $A\left(u_1, u_2\right) < B\left(u_1, u_2\right)\ \forall u_1, u_2 \in \left[0,1\right]^2$, then $\pi(A) \leqslant \pi(B)$. Therefore, once we find the concordance Fréchet-Hoeffding bound on the copulas then we can decide the upper and lower bounds on the price of the derivative.

Let $(\Omega, \mathcal{F}, \mathbb{P})$ be a probability space describing a financial market. Using a suitable equilibrium model or no-arbitrage arguments, financial theory shows that the price of a strategy with terminal payoff X_T (paid at time $T > 0$) can be written as

$$c(X_T) = \mathbb{E}[\xi_T X_T], \tag{3.10}$$

where ξ_T is some given stochastic discount factor (also called state-price process at T). In fact, for $\omega \in \Omega$, $\xi_T(\omega)$ can be interpreted as the price of consuming one unit in state ω and zero in all others. It is high in the worst states of the economy, that is when the "market" is at its lowest levels.

Example 3.3

Let $Q(u,v) = uv$ be the independence copula. And

$$\mathcal{S} = \left\{(a,b) \in [0,1]^2 \mid a \geqslant u_0,\ b \in [0,1]\right\}.$$

Applying Corollary 3.7, we find after some calculations that the maximum copula $A^{\mathcal{S},Q}$ satisfying

$$\forall u \in [u_0, 1],\ v \in [0,1] \quad A^{\mathcal{S},Q}(u,v) = uv, \tag{3.11}$$

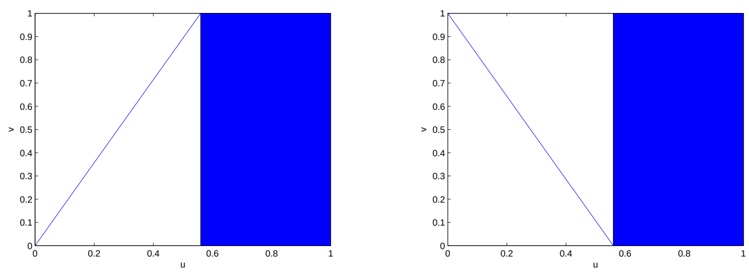
is given by $A^{\mathcal{S},Q}(u,v) = \min(u, u_0 v)1_{u\leqslant u_0} + uv1_{u>u_0}$. Similarly, we find that

$$B^{\mathcal{S},Q}(u,v) = \max\left(0, u_0(v-1) + u\right)1_{u\leqslant u_0} + uv1_{u>u_0}.$$

is the minimum copula. The supports of $A^{\mathcal{S},Q}$ and $B^{\mathcal{S},Q}$ are represented graphically in Panel A and B of Figure 5.

Figure 5: Supports for the copulas $A^{\mathcal{S},Q}$ (Panel A) and $B^{\mathcal{S},Q}$ (Panel B) of Example 3.3 with $u_0 = 0.56$. To simulate from $A^{\mathcal{S},Q}$ or $B^{\mathcal{S},Q}$, we use $\left(A_u^{\mathcal{S},Q}\right)^{-1}(y) = \left(\frac{u}{u_0}\right) 1_{u \leqslant u_0} + y 1_{u > u_0}$, and $\left(B_u^{\mathcal{S},Q}\right)^{-1}(y) = \left(1 - \frac{u}{u_0}\right) 1_{u \leqslant u_0} + y 1_{u > u_0}$.

Panel A Panel B



The minimum copula obtained in this example allows us to construct a strategy that provides at the lowest possible cost, the desired distribution which also exhibits independence with the market when the latter is low (high states for ξ_T). This is thus a very useful strategy for investors who seek for diversification (i.e. some degree of protection) in times of crisis.

Example 3.4

Similar to Tankov's example, use the Black-Scholes model to price the following basket option:

$$f(X, Y) = (\alpha X + \beta Y - K)^{+}.$$

The copula assumed by the model is the Gaussian copula, namely for a given correlation matrix $\Sigma \in \mathbb{R}^{2\times 2}$. The Gaussian copula with parameter matrix Σ can be written as :

$$Q_{\Sigma}^{Gauss}(u, v) = \Phi_{\Sigma}\left(\Phi^{-1}(u), \Phi^{-1}(v)\right),$$

where Φ is the cumulative distribution function of a standard normal distribution and Φ_{Σ} is the joint cumulative distribution function of a multivariate normal distribution with mean vector zero and covariance matrix equal to the correlation matrix Σ.

However, as we mentioned in the introduction, during a financial crisis, it is a known phenomenon that the market exhibits a higher level of co-movement. In order to account for the systematic risk, we assume that the usual Gaussian copula dependence structure is destroyed in an extreme event of a crisis. In the copula, we define this event to be in the rectangle $[0, 0.2] \times [0, 0.2]$. We consider $\mathcal{Q}_{\mathcal{S}}$ the set of all copulas C such that $C(u, v) = Q_{\Sigma}^{Gauss}(u, v)$, for all $(u, v) \in \mathcal{S}$, where $\mathcal{S}$ is the set $\{u \geqslant 0.2\} \cup \{v \geqslant 0.2\}$.

From Corollary 3.8, we know that the maximum copula $A^{\mathcal{S},Q}$ satisfying

$$\forall u \in [u_0, 1],\ v \in [0, 1] \quad A^{\mathcal{S},Q}(u, v) = Q_{\Sigma}^{Gauss}(u, v), \tag{3.12}$$

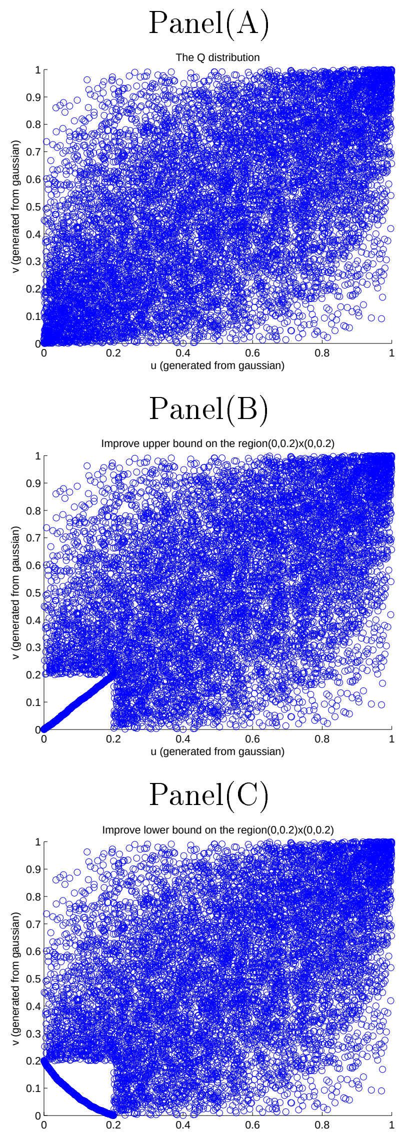
is given by $A^{\mathcal{S},Q}(u, v) = \min(u, u_0 v) 1_{u \leqslant u_0} + Q_{\Sigma}^{Gauss}(u, v) 1_{u > u_0}$. Similarly, we find that

$$B^{\mathcal{S},Q}(u, v) = \max\left(0, u_0(v - 1) + u\right) 1_{u \leqslant u_0} + Q_{\Sigma}^{Gauss}(u, v) 1_{u > u_0}.$$

is the minimum copula.

Using these copula to generate the price of the derivative numerically and setting $\alpha = \beta = 1$ we have the following figure, using a numerical method similar to the one in remark 3.5 to solve for the improved bounds.

Figure 6: Numerically Solving for Improved Upper and Lower Bound for the Gaussian Copula



4 Convex Ordering Bounds on Risk Aggregation

A copula is a very powerful tool with bivariate distributions. However, in higher dimensional space, the Fréchet-Hoeffding lower bound of copula does not exist, i.e. we cannot readily find a lower bound by convex order. This lower bound can give us the lowest price of a basket asset derivative or a lower convex risk measure of a portfolio. Therefore, minimization becomes a challenging problem. To solve this problem, it is necessary to use a completely different approach. In this thesis, we demonstrate that a certain type of minimization is solvable by looking for the lower bound of the convex order. The rest of the section will cite results from **?**. Major theorems that improved the lower bound are introduced and discussed. A more in-depth discussion based on interesting observations of the numerical results will be provided. This section will also propose some possible new directions.

We formalize the new problem with the following definitions.

Definition 4.1 (Convex order) *Let X and Y be two random variables with finite mean. X is larger than Y* in convex order*, denoted by $X \prec Y$, if $\forall$ convex functions f,*

$$\mathbb{E}[f(X)] \leqslant \mathbb{E}[f(Y)].$$

To translate convex order optimization into copula optimization problems, we have the co-monotonic copula as the convex upper bound

$$S \prec_{CX} F_1^{-1}(U) + F_2^{-1}(U) + F_3^{-1}(U) + ... + F_n^{-1}(U),$$

and in two variable case, we have the anti-comonotonic bound (Fréchet-Hoeffding lower bound) expressed as the convex lower bound

$$F_1^{-1}(U) + F_2^{-1}(1-U) \prec_{CX} S,$$

where $U \sim \mathcal{U}[0,1]$. Proofs for this assertion can be found in Meilijson and Nadas (1979), Tchen (1980) and Rüchendorf (1980, 1983). As stated above, there does not exist a general solution for the lower bound for over dimension 3. However, partial solutions exist for certain cases. Wang and Wang (2011) obtained the sharp lower bound for $n \geqslant 3$ in the special case when marginal distributions are identical with a monotone density function.

To introduce the complete mixability property, we will introduce a few notations. To keep the notation consistent, we will consider this from a risk manager point of view. However, the bound obtained in these calculations can be applied to derivative pricing as well.

Aggregate risk can be defined as the following problem. Assume we have n assets displaying return distributions of $F_1, F_2, F_3, ..., F_n$ along with random variables $X_1, X_2, X_3, ..., X_n$. Let us denote this space $\mathcal{F}_n$,

$$\mathcal{F}_n := \{X_i \sim F_i,\ i = 1, 2, 3, ..., n\}$$

For ease of notation, denote $\boldsymbol{X} = \{X_1, X_2, X_3, ..., X_n\}$, and $\boldsymbol{F} = \{F_1, F_2, F_3, ..., F_n\}$. Complete mixability property concerns itself on the distribution property of the sum of the random variables. Let us define the distribution of this type of sum as the admissible risk class.

Assume that all random variables live in a general atomless probability space $(\Omega, \mathcal{A}, \mathbb{P})$. This means that for all $A \subset \Omega$ with $\mathbb{P}(A) > 0$, there exists $B \subset A$ such that $\mathbb{P}(B) > 0$. The atomless assumption is very weak: in our context, it is equivalent to saying that there exists at least one continuously distributed random variable in this space (roughly, $(\Omega, \mathcal{A}, \mathbb{P})$ is not a finite space). In particular, it does not prevent discrete variables from coming into existence. In such a probability space, we can generate independent random vectors with any distribution. We denote by $L^0(\Omega, \mathcal{A}, \mathbb{P})$ the set of all random variables defined in the atomless probability space $(\Omega, \mathcal{A}, \mathbb{P})$. See Delbaen (2002) for details on risk measures defined in an atomless probability space.

Definition 4.2 (Admissible risk) *An aggregate risk S is called an* admissible risk *of marginal distributions $F_1, \cdots, F_n$ if it can be written as $S = X_1 + \cdots + X_n$ where $X_i \sim F_i$ for $i = 1, \cdots, n$. The* admissible risk class *is defined by the set of admissible risks of given marginal distributions:*

$$\begin{aligned} \mathfrak{S}_n(F_1, \cdots, F_n) &= \{\textit{admissible risk of marginal distributions } F_1, \cdots, F_n\} \\ &= \{X_1 + \cdots + X_n : X_i \sim F_i,\ i = 1, \cdots, n\}. \end{aligned}$$

The admissible risk class has many nice properties. To state the theorem on its properties, we need to introduce a few notations. $\mathbf{I}_A$ is the indicator function for the set $A \in \mathcal{A}$, and $T_{a,b}$ is an affine operator on univariate distributions such that for $a, b \in \mathbb{R}$,

$$T_{a,b}(\text{distribution of } X) = \text{distribution of } aX + b.$$

We also use $F \otimes G$ to denote the distribution of $X+Y$ where $X \sim F$ and $Y \sim G$ are independent, i.e. $(F \otimes G)(x) = \int_{-\infty}^{x} F(x-y)dG(y)$, and use $\stackrel{d}{=}$ and $\stackrel{d}{\to}$ to denote equality and convergence in law, respectively.

Theorem 4.1 (Properties of the admissible risk class)

1. *(convexity) If* $S_1 \in \mathfrak{S}_n(\mathbf{F})$, $S_2 \in \mathfrak{S}_n(\mathbf{G})$, *then* $\mathbf{I}_A S_1 + (1-\mathbf{I}_A)S_2 \in \mathfrak{S}_n(\mathbb{P}(A)\mathbf{F} + (1-\mathbb{P}(A))\mathbf{G})$ *for* $A \in \mathcal{A}$ *independent of* S_1 *and* S_2. *In particular,*

   (a) *if* $S_1, S_2 \in \mathfrak{S}_n(\mathbf{F})$, *then* $\mathbf{I}_A S_1 + (1-\mathbf{I}_A)S_2 \in \mathfrak{S}_n(\mathbf{F})$ *for* $A \in \mathcal{A}$ *independent of* S_1 *and* S_2;

   (b) *if* $S \in \mathfrak{S}_n(\mathbf{F}) \cap \mathfrak{S}_n(\mathbf{G})$, *then* $S \in \mathfrak{S}_n(\lambda\mathbf{F} + (1-\lambda)\mathbf{G})$ *for* $\lambda \in [0,1]$. *That is,* $\mathfrak{S}_n(\mathbf{F}) \cap \mathfrak{S}_n(\mathbf{G}) \subset \mathfrak{S}_n(\lambda\mathbf{F} + (1-\lambda)\mathbf{G})$ *for* $\lambda \in [0,1]$.

2. *(independent sum) If* $S_1 \in \mathfrak{S}_n(\mathbf{F})$ *and* $S_2 \in \mathfrak{S}_n(\mathbf{G})$ *are independent, then* $S_1 + S_2 \in \mathfrak{S}_n(F_1 \otimes G_1, \cdots, F_n \otimes G_n)$.

3. *(dependent sum) If* $S_1 \in \mathfrak{S}_n(\mathbf{F})$ *and* $S_2 \in \mathfrak{S}_m(\mathbf{G})$, *then* $S_1 + S_2 \in \mathfrak{S}_{n+m}(F_1, \cdots, F_n, G_1, \cdots, G_m)$.

4. *(affine invariance)* $S \in \mathfrak{S}_n(\mathbf{F}) \Leftrightarrow aS + b \in \mathfrak{S}_n(T_{a,b_1}F_1, \cdots, T_{a,b_n}F_n)$ *for* $a, b_i \in \mathbb{R}$, $i = 1, \cdots, n$ *and* $b = \sum_{i=1}^{n} b_i$.

5. *(permutation invariance) Let* σ *be an* n*-permutation, then* $\mathfrak{S}_n(\mathbf{F}) = \mathfrak{S}_n(\sigma(\mathbf{F}))$.

6. *(completeness) If* $S_k \in \mathfrak{S}_n(\mathbf{F})$, $k = 1, 2, \cdots$, *and* $S_k \stackrel{d}{\to} S$, *then* $S \in \mathfrak{S}_n(\mathbf{F})$.

7. *(continuity) If* $F_i^{(k)} \to F_i$ *point-wise when* $k \to +\infty$ *and for* $i = 1, \cdots, n$, *then*

   (a) *each* $S \in \mathfrak{S}_n(\mathbf{F})$ *is the weak limit of a sequence* $S_k \in \mathfrak{S}_n(F_1^{(k)}, \cdots, F_n^{(k)})$.

   (b) *each weakly convergent sequence* $S_k \in \mathfrak{S}_n(F_1^{(k)}, \cdots, F_n^{(k)})$ *has its weak limit* $S \in \mathfrak{S}_n(\mathbf{F})$.

Proof. In the proof, we first recall that the definition of admissible risks only concerns the distribution. That is, if $S_1 \stackrel{d}{=} S_2$, then $S_1 \in \mathfrak{S}_n(\mathbf{F}) \Leftrightarrow S_2 \in \mathfrak{S}_n(\mathbf{F})$.

1. Write $S_1 = \mathbf{X}\mathbf{1}_n$ and $S_2 = \mathbf{Y}\mathbf{1}_n$ where $\mathbf{X} \in \mathfrak{F}_n(\mathbf{F})$ and $\mathbf{Y} \in \mathfrak{F}_n(\mathbf{G})$. Let $B \in \mathcal{A}$ be independent of $\mathbf{X}$ and $\mathbf{Y}$, and $\mathbb{P}(B) = \mathbb{P}(A)$. It is easy to check that $\mathrm{I}_A S_1 + (1 - \mathrm{I}_A) S_2 \stackrel{d}{=} \mathrm{I}_B S_1 + (1 - \mathrm{I}_B) S_2$. Note that $\mathrm{I}_B S_1 + (1 - \mathrm{I}_B) S_2 = (\mathrm{I}_B X_1 + (1 - \mathrm{I}_B) Y_1) + \cdots + (\mathrm{I}_B X_n + (1 - \mathrm{I}_B) Y_n) \in \mathfrak{S}_n(\mathbb{P}(B)\mathbf{F} + (1 - \mathbb{P}(B))\mathbf{G})$. It follows that $\mathrm{I}_A S_1 + (1 - \mathrm{I}_A) S_2 \in \mathfrak{S}_n(\mathbb{P}(A)\mathbf{F} + (1 - \mathbb{P}(A))\mathbf{G})$.

2. Write $S_1 = \mathbf{X}\mathbf{1}_n$ and $S_2 = \mathbf{Y}\mathbf{1}_n$ where $\mathbf{X} \in \mathfrak{F}_n(\mathbf{F})$ and $\mathbf{Y} \in \mathfrak{F}_n(\mathbf{G})$. Let $\mathbf{Z} \in \mathfrak{F}_n(\mathbf{G})$ be independent of $\mathbf{X}$ and $\mathbf{Z} \stackrel{d}{=} \mathbf{Y}$. It is easy to check that $S_1 + S_2 \stackrel{d}{=} \mathbf{X}\mathbf{1}_n + \mathbf{Z}\mathbf{1}_n$. Note that $\mathbf{X}\mathbf{1}_n + \mathbf{Z}\mathbf{1}_n = (\mathbf{X} + \mathbf{Z})\mathbf{1}_n \in \mathfrak{S}_n(F_1 \otimes G_1, \cdots, F_n \otimes G_n)$. It follows that $S_1 + S_2 \in \mathfrak{S}_n(F_1 \otimes G_1, \cdots, F_n \otimes G_n)$.

3. (iii)-(v) Trivial.

4. (vi) This is a special case of (vii)(b) below.

5. (vii)

   (a) Write $S = \mathbf{X}\mathbf{1}_n$, where $\mathbf{X} \in \mathfrak{F}_n(\mathbf{F})$ and let C be the copula of $\mathbf{X}$. Let $S_k = \mathbf{X}_k\mathbf{1}_n$, where $\mathbf{X}_k \in \mathfrak{F}_n(F_1^{(k)}, \cdots, F_n^{(k)})$ with copula C. It is obvious that $S_k \stackrel{d}{\to} S$.

   (b) Write $S_k = \mathbf{X}_k\mathbf{1}_n$, where $\mathbf{X}_k \in \mathfrak{F}_n(F_1^{(k)}, \cdots, F_n^{(k)})$ with copula C_k. Note that the space of n-copulas is a compact space. Hence, there is a subsequence C_{k_i} of C_k such that C_{k_i} has a limit. Then the subsequence $S_{k_i} \stackrel{d}{\to} \mathbf{X}\mathbf{1}_n$ where $\mathbf{X} \in \mathfrak{F}_n(\mathbf{F})$ with copula C as the limit of C_{k_i}. Since $S_k \stackrel{d}{\to} S$, we have $S \stackrel{d}{=} \mathbf{X}\mathbf{1}_n \in \mathfrak{S}_n(\mathbf{F})$.

■

Remark 4.1

Part (B) of the convexity of the admissible risk class is equivalent to the convexity of the set $\mathcal{S}$ of all copulas. Similarly the convexity of the admissible risk class can help to generate new copulas from the existing classes. The completeness of $\mathfrak{S}_n$ means that any optimization problem with the entire admissible risk class as feasible region can always reach a solution. The continuity can help to find properties on S by approximating using discrete distributions numerically.

4.1 Convex Ordering Bounds on Admissible Risks

4.1.1 Convex ordering bounds

Let F be a distribution on $\mathbb{R}^+$ with finite mean. We first consider the homogeneous case and give a lower convex ordering bound on $\mathfrak{S}_n(F, \cdots, F)$ for $n \geqslant 3$ in Theorems 4.2 and 4.3. Let us define $H(\cdot)$ and $D(\cdot)$ as follows.

$$\forall x \in \left[0, \frac{1}{n}\right], \quad H(x) = (n-1)F^{-1}\left((n-1)\,x\right) + F^{-1}\left(1-x\right),$$

$$\forall a \in \left[0, \frac{1}{n}\right), \quad D(a) = \frac{n}{1-na}\int_a^{\frac{1}{n}} H(x)dx = n\frac{\int_{(n-1)a}^{1-a} F^{-1}(y)dy}{1-na}, \qquad (4.1)$$

and $H(0) = +\infty$ when the support of F is unbounded. The possible infiniteness of $H(0)$ is for convenience only and will not be problematic in what follows. Note also that $D(a)$ is always finite since $\int_a^{\frac{1}{n}} H(x)\mathrm{d}x \leqslant \int_0^{\frac{1}{n}} H(x) = \mathbb{E}\left[X_1\right]$ is finite (as F is a distribution with finite mean). Let us give some intuition about these two quantities. From the last expression of $D(a)$, it is clear that $D(a)$ is directly related to the average sum when its components $(X_1, \cdots, X_n)$ are all in the *middle* of the distribution (also called *body* of the distribution). Precisely,

$$D(a) = \sum_{j=1}^{n} \mathbb{E}\left[X_j \,\middle|\, X_j \in \left[F^{-1}((n-1)a), F^{-1}(1-a)\right]\right] \qquad (4.2)$$

because $\mathbb{P}\left[X_j \in [F^{-1}((n-1)a), F^{-1}(1-a)]\right] = 1 - na$ and $X_1,\, X_2, \cdots, X_n$ all have the same distribution. It is also clear that $H(x)$ and $D(a)$ can be easily calculated for a given distribution F.

Intuitively, the dependence scenario to attain the convex ordering lower bound is constructed such that when one of the X_i is large then all the others are small (all X_i are in the *tails* of the distribution; the pair (X_i, X_j) is counter-monotonic for X_i large and $j \neq i$) and when one of the X_i is of medium size (in the *body* of the distribution) we treat the sum $\sum_i X_i$ as a constant equal to its conditional expectation as in (4.2). Precisely, the lower bound in the coming theorem corresponds exactly to the following dependence structure. The probability space is split into two parts: the *tails* (with probability na for a small value of $a \in [0, 1/n]$) and the *body* (with probability $1 - na$). $H(\cdot)$ gives the values of S in the tails and $D(a)$ is the value of S in the body of the distribution. To this end, for $a \in \left[0, \frac{1}{n}\right]$, we introduce a random variable

$$T_a = H(U/n)\mathrm{I}_{\{U \in [0,na]\}} + D(a)\mathrm{I}_{\{U \in (na,1]\}}, \tag{4.3}$$

where $U \sim \mathcal{U}[0,1]$. The atomless assumption of the probability space $(\Omega, \mathcal{A}, \mathbb{P})$ allows us to generate such U, and since we only care about distributions to prove convex order, we do not specify the random variable U. In Theorem 4.2, we prove that T_a is a convex ordering lower bound given that $H(\cdot)$ satisfies a monotonicity property. In the proof of Theorem 4.3, we find the best convex ordering bound and exhibit the worst dependence structure explicitly.

Theorem 4.2 (Convex ordering lower bound for homogeneous risks) *Suppose condition (A) holds:*
(A) for some $a \in \left[0, \frac{1}{n}\right]$, $H(x)$ is non-increasing on the interval $[0,a]$ and $\lim_{x\to a-} H(x) \geqslant D(a)$,
then,

1. $T_a \prec_{CX} S$ *for all* $S \in \mathfrak{S}_n(F, \cdots, F)$;
2. $T_u \prec_{CX} T_v$ *for all* $0 \leqslant u \leqslant v \leqslant \frac{1}{n}$. *Thus, the most accurate lower bound is obtained by the largest a such that (A) holds.*

Proof.

1. Let $\mathbf{X} \in \mathfrak{F}_n(F, \cdots, F)$, $S = \mathbf{X}\mathbf{1}_n \in \mathfrak{S}_n(F, \cdots, F)$ and T_a be defined in (4.3). It is straightforward to check $\mathbb{E}[T_a]$. Let F_S and F_{T_a} be the cdf of S and T_a respectively, and further let $U_1, \cdots, U_n$ be $\mathcal{U}[0,1]$ random variables such that $F^{-1}(U_i) = X_i$ for $i = 1, \cdots, n$. Such $U_1, \cdots, U_n$ always exist in an atomless probability space. Our goal is to show that

$$\forall c \in [0,1], \quad \int_c^1 F_{T_a}^{-1}(t)dt \geqslant \int_c^1 F_S^{-1}(t)dt. \tag{4.4}$$

   Property (4.4) together with $\mathbb{E}[T_a] = \mathbb{E}[S]$ is equivalent to $T_a \prec_{\mathrm{cx}} S$ (for example, see Theorem 2.5 of Bäuerle and Müller (2006)).

   To obtain this, denote $A_S(u) = \bigcup_i \{U_i > 1-u\}$ and let $W(u) = \mathbb{P}(A_S(u))$. Obviously $u \leqslant W(u) \leqslant nu$ and W is non-decreasing. For $c \in [0, na]$, let $u^\star = W^{-1}(c)$, it then follows that $c \geqslant u^\star \geqslant c/n$

and $\{U_i \in [1-c/n, 1]\} \subset \{U_i \in [1-u^\star, 1]\} \subset A_S(u^\star)$. Note that $\mathbb{P}(A_S(u^\star)) = c$, therefore $\mathbb{P}\{A_S(u^\star) \setminus U_i \in [1-c/n, 1]\} = c - c/n = \mathbb{P}\{U_i \in [0, (n-1)c/n]\}$. Since $X_i = F^{-1}(U_i)$ is non-decreasing in U_i and the above two sets have the same measure, we have

$$\mathbb{E}\left[\mathrm{I}_{\{U_i \in [0,(n-1)c/n]\}} X_i\right] \leqslant \mathbb{E}\left[\mathrm{I}_{A_S(u^\star)\setminus\{U_i\in[1-c/n,1]\}} X_i\right]. \tag{4.5}$$

It follows that

$$\begin{aligned}
\mathbb{E}\left[\mathrm{I}_{\{U\leqslant c\}} T_a\right] &= \mathbb{E}\left[\mathrm{I}_{\{U\leqslant c\}} H(U/n)\right] \\
&= n\int_0^{c/n} \big((n-1)F^{-1}((n-1)x) + F^{-1}(1-x)\big)\, dx \\
&= n\int_0^{\frac{(n-1)c}{n}} F^{-1}(t)dt + n\int_{1-c/n}^1 F^{-1}(t)dt \\
&= n\mathbb{E}\left[\left(\mathrm{I}_{\{U_i\in(n-1)c/n\}} + \mathrm{I}_{\{U_i\in[1-c/n,1]\}}\right) X_i\right] \\
&\leqslant n\mathbb{E}\left[\left(\mathrm{I}_{A_S(u^\star)\setminus\{U_i\in[1-c/n,1]\}} + \mathrm{I}_{\{U_i\in[1-c/n,1]\}}\right) X_i\right]
\end{aligned}$$

where the inequality follows from (4.5). We then find that $\mathbb{E}\left[\mathrm{I}_{\{U\leqslant c\}} T_a\right] \leqslant n\mathbb{E}[\mathrm{I}_{A_S(u^\star)} X_i] = \mathbb{E}[\mathrm{I}_{A_S(u^\star)} S]$. Thus we have

$$\mathbb{E}\left[\mathrm{I}_{\{U\leqslant c\}} T_a\right] \leqslant \mathbb{E}[\mathrm{I}_{A_S(u^\star)} S]. \tag{4.6}$$

Note that $H(x)$ is non-increasing on $[0, a]$ and $\lim_{x\to a-} H(x) \geqslant D(a)$. Thus for $c \in [0, na]$,

$$\mathbb{E}\left[\mathrm{I}_{\{U\leqslant c\}} T_a\right] = \mathbb{E}[\mathrm{I}_{\{U\leqslant c\}} H(U/n)] = \int_{1-c}^1 F_{T_a}^{-1}(t)\mathrm{d}t. \tag{4.7}$$

Also note that

$$\mathbb{E}\left[\mathrm{I}_{A_S(u^\star)} S\right] \leqslant \int_{1-c}^1 F_S^{-1}(t)\mathrm{d}t \tag{4.8}$$

since $\mathbb{P}(A_S(u^\star)) = c$. It follows from (4.6), (4.7) and (4.8) that for any $c \in [0, na]$,

$$\int_{1-c}^1 F_{T_a}^{-1}(t)\mathrm{d}t \leqslant \int_{1-c}^1 F_S^{-1}(t)\mathrm{d}t. \tag{4.9}$$

For $x \in [0, 1-na]$, let $G(x) = \int_x^1 F_S^{-1}(t)\mathrm{d}t - \int_x^1 F_{T_a}^{-1}(t)\mathrm{d}t$. Note that $\int_x^1 F_S^{-1}(t)\mathrm{d}t$ is concave, and $F_{T_a}^{-1}(t) = D(a)$ is a constant when $t \in$

$[0, 1-na)$, hence $G(x)$ is concave over $[0, 1-na)$. Since G is concave, $G(0) = \mathbb{E}[S] - \mathbb{E}[T_a] = 0$, and $G(1-na) \geqslant 0$ by (4.9), we have $G(x) \geqslant 0$ over $[0, 1-na]$. Thus

$$\int_c^1 F_{T_a}^{-1}(t)\mathrm{d}t \leqslant \int_c^1 F_S^{-1}(t)\mathrm{d}t \tag{4.10}$$

for any $c \in [0,1]$. This implies $T_a \prec_{\mathrm{cx}} S$.

2. For $0 \leqslant u \leqslant v \leqslant \frac{1}{n}$, it can be easily checked that the distribution of T_u is a fusion of the distribution of T_v, and thus $T_u \prec_{\mathrm{cx}} T_v$ (see Theorem 2.8 of Bäuerle and Müller (2006) for the definition of a fusion and a proof of this assertion).

■

Definition 4.3 (Complete Mixability) *A distribution function F on $\mathbb{R}$ is* n-completely mixable *(n-CM) if there exist n random variables $X_1, \ldots, X_n$ identically distributed as F such that*

$$X_1 + \cdots + X_n = n\mu \tag{4.11}$$

for some $\mu \in \mathbb{R}$ referred as a center *of F. A distribution function F on $\mathbb{R}$ is called* n-CM on an interval *I (finite or infinite) if the conditional distribution of F on I is n-CM.*

As F has finite mean, if F is n-CM, then its center is unique and equal to the mean. Note that F is n-CM equivalent to $n\mathbb{E}[X] \in \mathfrak{S}_n(F, \cdots, F)$, where $X \sim F$. Some straightforward examples and properties of completely mixable distributions are given in Wang and Wang (2011) and Puccetti et al. (2012). By Theorem 4.2, one needs to find the largest possible a to get the most accurate lower bound. This motivates us to define c_n by

$$c_n = \inf c \in \left(0, \frac{1}{n}\right) : H(c) \leqslant D(c) \tag{4.12}$$

Note that c_n is the largest possible a satisfying $\lim_{x \to a-} H(x) \geqslant D(a)$. When F is a continuous distribution, $H(c_n) = D(c_n)$. On the other hand, c_n is exactly the smallest possible a such that F on $I = [F^{-1}((n-1)a), F^{-1}(1-a)]$

satisfies the mean condition necessary for the CM property. See, for example, (7) in Proposition 2.1 of Wang and Wang (2011) for more details on this condition.

? improved the aforementioned convex ordering lower bounds under two broader assumptions. The first bound is proposed by assuming $F_1 = ... = F_n$, namely all random variables being added up follow the same distribution. Under this assumption, a sharp lower bound can be obtained. This case significantly reduces the complexity of the problem but is relevant in practice. For example, it is useful for an insurer who has a portfolio of identically distributed policyholders' individual risks. In another context, it can be used to find bounds on prices of variance options when subsequent stocks' log-returns are identically distributed. More information on these examples is given in Section 4.2.3. In the last part of this section, we then generalize to the case when the distributions F_i can be different, called *heterogeneous* risks.

Theorem 4.3 (Sharp convex ordering lower bound for homogeneous risks) *Suppose*
(A) $H(x)$ is non-increasing on the interval $[0, c_n]$, where c_n is given by (4.12)
then $T_{c_n} \prec_{CX} S$ for all $S \in \mathfrak{S}_n(F, \cdots, F)$. Moreover, $T_{c_n} \in \mathfrak{S}_n(F, \cdots, F)$ that is T_{c_n} is sharp if (B) holds:
(B) F is n-CM on the interval $I = [F^{-1}((n-1)c_n), F^{-1}(1-c_n)]$.

Proof. $T_{c_n} \prec_{\text{cx}} S$ follows from Theorem 4.2 by noting that $\lim_{x\to c_n-} H(x) \geqslant D(c_n)$ from the definition of c_n in (4.12). Let us prove the second half of the theorem. When condition (B) holds, that is F is n-CM on I, there exist random variables $Y_1, \cdots, Y_n$ from the conditional distribution F on I such that $Y_1 + \cdots + Y_n$ is a constant. Thus, as Y has finite mean (because F has finite mean), $Y_1 + \cdots + Y_n = n\mathbb{E}(Y_1) = D(c_n)$ by (4.1) and (4.2). Now we construct $S \in \mathfrak{S}_n(F_1, \cdots, F_n)$ which has the same distribution as T_{c_n}, by imposing a special dependence structure. For each i, when $X_i \in I$ (the *body* part), we let $X_i = Y_i$ and when $X_i \notin I$ (the *tail* part), we let (X_i, X_j) be counter-monotonic for each $j \neq i$. That is,

$$X_i = \mathrm{I}_{\{U > nc_n\}} Y_i + \mathrm{I}_{\{U \leqslant nc_n\}} F^{-1}(V_i), \tag{4.13}$$

where $U \sim \mathcal{U}[0,1]$, $(V_1, \cdots, V_n)$ is independent of U and uniformly distributed on the line segments

$$\mathcal{O} = \bigcup_{k=1}^{n} \{(v_1, \cdots, v_n) : v_j = (n-1)(1-v_k),\ v_k \in [1-c_n, 1]\,,\ j = 1, \cdots, n,\ j \neq k\}. \tag{4.14}$$

We can check that V_i is uniformly distributed on $[0, (n-1)c_n] \cup [1-c_n, 1]$, and thus the distribution of $F^{-1}(V_i)$ is the conditional distribution of F on $\mathbb{R}^+ \setminus I$. Moreover by construction, Y_i has the conditional distribution of F on I. It follows that $X_i \sim F$. Then

$$\begin{aligned} S &= \sum_{i=1}^{n} \left(\mathrm{I}_{\{U > nc_n\}} Y_i + \mathrm{I}_{\{U \leqslant nc_n\}} F^{-1}(V_i) \right) \\ &= \mathrm{I}_{\{U > nc_n\}} D(c_n) + \mathrm{I}_{\{U \leqslant nc_n\}} \sum_{i=1}^{n} F^{-1}(V_i). \end{aligned}$$

Note that

$$\sum_{i=1}^{n} F^{-1}(V_i) = \sum_{i=1}^{n} \mathrm{I}_{\{V_i \geqslant 1-c_n\}} (F^{-1}((n-1)(1-V_i)) + F^{-1}(V_i)) = \sum_{i=1}^{n} \mathrm{I}_{\{V_i \geqslant 1-c_n\}} H(1-V_i),$$

and for $t > 0$,

$$\begin{aligned} \mathbb{P}\left(\sum_{i=1}^{n} F^{-1}(V_i) \leqslant t \right) &= \mathbb{P}\left(\sum_{i=1}^{n} \mathrm{I}_{\{V_i \geqslant 1-c_n\}} H(1-V_i) \leqslant t \right) \\ &= \mathbb{E}\left(\sum_{i=1}^{n} \mathrm{I}_{\{V_i \geqslant 1-c_n\}} \mathbb{P}(H(1-V_i) \leqslant t | V_i \geqslant 1-c_n) \right) \\ &= \mathbb{P}(H(1-V_1) \leqslant t | V_1 \geqslant 1-c_n) \\ &= \mathbb{P}\left(H(V) \leqslant t \right) \end{aligned}$$

for some $V \sim \mathcal{U}\,[0, c_n]$, independent of U. Note that the second equality holds because $\{V_i \geqslant 1-c_n\}$ are mutually exclusive. Therefore, $S \stackrel{d}{=} \mathrm{I}_{\{U > nc_n\}} D(c_n) + \mathrm{I}_{\{U \leqslant nc_n\}} H(V) \stackrel{d}{=} T_{c_n}$, and thus $T_{c_n} \in \mathfrak{S}_n(F, \cdots, F)$. ■

Theorem 4.4 *Suppose $H(x)$ is strictly decreasing on $[0, c_n]$. Then,*

1. *$T_{c_n} \in \mathfrak{S}_n(F, \cdots, F)$ if and only if (B) holds;*

2. $T_a \notin \mathfrak{S}_n(F, \cdots, F)$ *for all* $a < c_n$.

Proof.

1. The "$\Leftarrow$" part follows directly from Theorem 4.3. Let us show the "$\Rightarrow$" part. We begin by showing this assertion in the discrete case. Let F be any continuous distribution on $\mathbb{R}^+$, with F^{-1} strictly increasing. Let G be the distribution of $F^{-1}(V)$ where V is a discrete uniform distribution on $\{0, \frac{1}{K}, \cdots, \frac{K-1}{K}\}$ for some large number $K > n$ and let $\hat{T}_{c_n}$ be defined as T_{c_n} with F replaced by G:

$$\hat{T}_{c_n} = \hat{H}(U/n)\mathrm{I}_{\{U \in 0, nc_n\}} + \hat{D}\mathrm{I}_{\{U \in (nc_n, 1]\}}, \tag{4.15}$$

where $\hat{H}(x) = (n-1)G^{-1}((n-1)x) + G^{-1}(1-x)$, $U \sim \mathcal{U}[0,1]$,

$$c_n = \inf\left\{c \in \left\{0, \frac{1}{K}, \cdots, \left\lfloor \frac{K}{n} \right\rfloor \frac{1}{K}\right\} : \hat{H}(c) \leqslant \frac{n}{1-nc}\int_c^{\frac{1}{n}} \hat{H}(x)\mathrm{d}x\right\},$$

and $\hat{D} := \frac{n}{1-nc_n}\int_{c_n}^{\frac{1}{n}} \hat{H}(x)\mathrm{d}x$ is a constant. Note that $G^{-1}(t) = F^{-1}(t)$ for $t = 0, \frac{1}{K}, \cdots, \frac{K-1}{K}$, and $G^{-1}(x) = F^{-1}(\frac{\lfloor xK \rfloor}{K})$ for $x \in [0,1)$. Thus, $H(t) = \hat{H}(t)$ for $t = 0, \frac{1}{K}, \cdots, \frac{K-1}{K}$, and the interval $I = G^{-1}((n-1)c_n), G^{-1}(1-c_n)$. Note that since G is discrete, this function $\hat{H}$ is not non-increasing, but this would not hurt our proof since we are not using the results in convex order. To simulate the strict decreasing property, we assume

$$\min_{i \leqslant x < i+1} \hat{H}\left(\frac{x}{K}\right) > \max_{i+1 \leqslant x < i+2} \hat{H}\left(\frac{x}{K}\right) \quad \text{for } i = 0, \cdots, Kc_n - 2. \tag{4.16}$$

Suppose $\hat{T}_{c_n} = \mathbf{X}\mathbf{1}_n \in \mathfrak{S}_n(G, \cdots, G)$ for $\mathbf{X} \in \mathfrak{F}_n(G, \cdots, G)$. Let us show that this implies G is n-CM on I. Note that by definition of $\hat{T}_{c_n}$ and (4.16),

$$\mathbb{P}\left[\hat{T}_{c_n} - G^{-1}\left(1 - \frac{1}{K}\right) \in \left((n-1)G^{-1}(0), (n-1)G^{-1}(n-1)\frac{1}{K}\right)\right] = \frac{n}{K},$$

and

$$\mathbb{P}\left[\hat{T}_{c_n} > (n-1)G^{-1}(n-2)\frac{1}{K} + G^{-1}1 - \frac{1}{K}\right] = 0.$$

This implies that when one of X_i takes the value $G^{-1}\left(1-\frac{1}{K}\right)$, all the others must take values in $\left[G^{-1}(0), G^{-1}\left((n-1)\frac{1}{K}\right)\right)$, by observing that $G^{-1}(x)$ is strictly increasing. . Using this argument again,we obtain that when one of X_i, takes the value $G^{-1}\left(1-\frac{2}{K}\right)$, all the others must take values in

$$\left[G^{-1}(n-1)\frac{1}{K}, G^{-1}(n-1)\frac{2}{K}\right).$$

Eventually, we have that for all $1 \leqslant j < Kc_n$, when X_i takes the value $G^{-1}\left(1-\frac{j}{K}\right)$, all the others must take values in $\left[G^{-1}\left((n-1)\frac{j-1}{K}\right), G^{-1}\left((n-1)\frac{j}{K}\right)\right)$. The remaining part is

$$\mathrm{P}\left[\hat{T}_{c_n} = \hat{D}\right] = 1 - nc_n.$$

Let $A = \{\hat{T}_{c_n} = \hat{D}\}$. The conditional distribution of X_i on A is exactly the conditional distribution G on I, since $\{X_i \notin I\}$ has been contained in the set A^c. Since $\hat{T}_{c_n}$ is a constant on A, we have G is n-CM on I. The above proof shows that for a discrete distribution G, if G^{-1} is strictly increasing and $\hat{H}$ satisfies (4.16), then T_{c_n} is admissible implies that the conditional distribution is n-CM on I. To prove the case of F being continuous, we can simply replace $\frac{1}{K}$ by an infinitesimal $\mathrm{d}t$, and the condition (4.16) is equivalent to H being strictly decreasing. Note that H being strictly increasing is sufficient for F^{-1} to be strictly increasing on $[1-nc_n, 1]$, which is sufficient for our proof.

2. By (4.1), we know $D(a)$ is a strictly decreasing function of a. Suppose $a < c_n$ and let $c = \frac{1}{2}a + \frac{1}{2}c_n$, then $c < \frac{1}{n}$ and $D(a) > D(c)$. It is straightforward to check that

$$\mathbb{E}[(T_a - D(a))^+] = \mathbb{E}[T_a] - D(a) = \mathbb{E}[T_c] - D(a) < \mathbb{E}[(T_c - D(a))^+]$$

since $\mathbb{P}(T_c < D(a)) \geqslant \mathbb{P}(T_c = D(c)) \geqslant 1 - nc > 0$. This shows $T_c \not\prec_{\mathrm{cx}} T_a$ by the definition of convex order. Since $c < c_n$, we have $H(c) \geqslant D(c)$, and by Theorem 4.2 $T_c \prec_{\mathrm{cx}} S$ for any $S \in \mathfrak{S}_n(F, \cdots, F)$. Thus we conclude that $T_a \notin \mathfrak{S}_n(F, \cdots, F)$ for $a < c_n$.

■

We have the following theorem as a generalization of Theorem 4.2.

Theorem 4.5 (Convex ordering lower bound for heterogeneous risks)

1. *$\mathfrak{S}_n(F_1, \cdots, F_n) \subset \mathfrak{S}_n(F, \cdots, F)$.*
2. *Suppose (A) holds, then $T_a \prec_{CX} S$ for all $S \in \mathfrak{S}_n(F_1, \cdots, F_n)$.*

Proof.

1. Let σ_k, $k = 1, 2, \cdots, n!$ be all different n-permutations. By Theorem 4.1 (i)(b) and (iv), we have

$$\mathfrak{S}_n(F_1, \cdots, F_n) = \bigcap_{k=1}^{n!} \mathfrak{S}_n(\sigma_k(F_1, \cdots, F_n)) \subset \mathfrak{S}_n\left(\sum_{k=1}^{n!} \lambda_k \sigma(F_1, \cdots, F_n)\right),$$

   where $\lambda_k \geqslant 0$, $k = 1, 2, \cdots, n!$ and $\sum_{k=1}^{n!} \lambda_k = 1$. Take $\lambda_k = \frac{1}{n!}$ for all k then we get $\mathfrak{S}_n(F_1, \cdots, F_n) \subset \mathfrak{S}_n(F, \cdots, F)$.

2. By Theorem 4.2 and (i), $T_a \prec_{\text{cx}} S$ for all $S \in \mathfrak{S}_n(F, \cdots, F)$, and hence $T_a \prec_{\text{cx}} S$ for all $S \in \mathfrak{S}_n(F_1, \cdots, F_n)$.

■

4.2 Applications to Risk Measures, Finance and Insurance

We now present several applications of the theorems of the previous section on convex order for risk aggregation. Our results on convex order apply naturally on bounds on convex risk measures and in particular on coherent risk measures as well as on convex expectations. The first paragraph recalls definitions and properties of risk measures. The second paragraph contains the main result on convex risk measures, a third paragraph is dedicated to bounds on *TVaR* and the last paragraph describes a series of applications on convex expectations.

Throughout the applications, we use the conditions (A), (A') and (B) introduced in Section 4.1:

1. (A) For $a \in \left[0, \frac{1}{n}\right]$, $H(x)$ is non-increasing on $[0, a]$ and $\lim_{x \to a-} H(x) \geqslant D(a)$.

2. (A') $H(x)$ is non-increasing on the interval $[0, c_n]$.

3. (B) The distribution F is n-CM on the interval

$$I = \left[F^{-1}((n-1)c_n), F^{-1}(1-c_n)\right]$$

.

Here, for consistency, $H(x)$ and $D(a)$ are defined as in Section 4.5 for marginal distributions $F_1, \cdots, F_n$ (this definition coincides with the one in Section 4.1.1 when $F = F_1 = \cdots = F_n$), and c_n is defined by (4.12). (A) is used for both homogeneous and heterogeneous risks, while (A') and (B) are used only for homogeneous risks.

4.2.1 Convex and coherent risk measures

A risk measure is a mapping from random variables to real numbers, which can be used as a capital requirement to regulate risk assumed by market participants. For a detailed introduction on risk measures and more specifically on coherent risk measures, we refer to Artzner et al. (1999). Consider a *risk measure* as $\rho : L^0(\Omega, \mathcal{A}, \mathbb{P}) \to \mathbb{R} \cup \{\infty\}$. Most discussions focus on risk measures on $L^p(\Omega, \mathcal{A}, \mathbb{P})$ for $p \in [1, \infty]$. Delbaen (2009) studied the case of non-integrable random variables, and proved that there exist no finite convex risk measures defined on $L^p(\Omega, \mathcal{A}, \mathbb{P})$ for $p \in [0, 1)$. Since convex order is defined for L^1 random variables, we restrict our discussion on $\rho : L^1(\Omega, \mathcal{A}, \mathbb{P}) \to \mathbb{R}$. Let $X, X_1, X_2, \cdots \in L^1(\Omega, \mathcal{A}, \mathbb{P})$. Recall the following properties of a risk measure $\rho(\cdot)$

1. Monotonicity: if $X_1 \leqslant X_2$ then $\rho(X_1) \leqslant \rho(X_2)$.

2. Translation invariance: $\rho(X + m) = \rho(X) + m$ for $m \in \mathbb{R}$.

3. Subadditivity: $\rho(X_1 + X_2) \leqslant \rho(X_1) + \rho(X_2)$.

4. Positive homogeneity: $\rho(\lambda X) = \lambda \rho(X)$ for $\lambda > 0$.

5. Convexity: $\rho(\lambda X_1 + (1-\lambda)X_2) \leqslant \lambda \rho(X_1) + (1-\lambda)\rho(X_2)$ for $\lambda \in [0, 1]$.

6. Law invariance: if $X_1 \overset{d}{=} X_2$, then $\rho(X_1) = \rho(X_2)$.

7. L^1-Fatou property: if $X_n \to X$ in L^1, then $\rho(X) \leqslant \liminf \rho(X_n)$.

A risk measure is *coherent* if it satisfies properties (1-4). It is immediate that a coherent risk measure satisfies also (5). Recall that a coherent risk measure has the typical dual representation

$$\rho(X) = \sup_{Q \in \mathcal{Q}} \mathbb{E}_Q[X]$$

where $\mathcal{Q}$ is some family of probability measures on Ω. This was introduced in Artzner et al. (1999) in a finite state probability space and discussed in Delbaen (2002) in a more general probability space.

A risk measure on $L^\infty(\Omega, \mathcal{A}, \mathbb{P})$ is called a *convex risk measure*, defined in Föllmer and Schied (2002), if it satisfies properties (1,2,5). A dual representation is also given in the same paper. The concept was later studied in Svindland (2008) and Kaina and Rüschendorf (2009), for more general probability spaces. A recent review of convex and coherent measures can be found in Föllmer and Schied (2010).

- The commonly used risk measure Value-at-Risk (*VaR*), defined as

  $$VaR_p(X) = \inf\{x : \mathbb{P}(X \leqslant x) \geqslant p\}, \qquad p \in (0,1),$$

  satisfies (1,2,4,6). It is often criticized for not being subadditive (and thus it is neither convex nor coherent).

- Another commonly used risk measure is the Tail Value-at-Risk (*TVaR*; it has other names and variations such as *CTE*, *AVaR*, *CVaR* and *ESF* in different contexts). It is defined as

  $$\mathrm{TVaR}_p(X) = \frac{1}{1-p} \int_p^1 \mathrm{VaR}_\alpha(X) d\alpha, \qquad p \in [0,1).$$

  As it satisfies (1-7), it is a coherent risk measure. Furthermore, any risk measures on $L^1(\Omega, \mathcal{A}, \mathbb{P})$ satisfying (1-7) has a representation of

  $$\rho(X) = \sup_{\mu \in P_0} \int_0^1 TVaR_p(X) \mu(dp) \tag{4.17}$$

  where P_0 is a compact, convex set of probability measures on $[0,1]$ (for this result, see Bäuerle and Müller (2006); Kusuoka (2009)).

- The standard deviation principle, defined as $\rho(X) = \mathbb{E}(X) + k\sqrt{var(X)}$ for some constant k, satisfies (2-7): it is neither coherent nor convex.

- A distortion risk measure, defined as

$$\rho(X) = \int_0^1 F^{-1}(t) g'(1-t) dt$$

for an increasing function g with $g(0) = 0$, $g(1) = 1$ is coherent if g is concave on $[0,1]$. In particular it satisfies convex order.

- The entropic risk measure, defined as

$$\rho(X) = \frac{1}{\theta} \log \mathbb{E}[e^{\theta X}],$$

satisfies (1-2) and (5-7): it is an example of a non-coherent convex risk measure.

Due to the increasing importance of $TVaR$ in risk management according to recent industrial regulations (see e.g. Panjer (2006) and Basel Committee on Banking Supervision (2010); **?**) and the representation (4.17) of law-invariant coherent risk measures, bounds for $TVaR_p(S)$ are of practical interest.

Theorem 4.6 (Bounds on TVaR of admissible risk)

1. For $p \in [0,1]$, if (A) holds, then

$$\inf_{S \in \mathfrak{S}_n(F_1, \cdots, F_n)} \mathrm{TVaR}_p(S) \geqslant \begin{cases} \frac{1}{1-p}[\mathbb{E}[S] - pD(a)] & p \leqslant 1 - na \\ \frac{n}{1-p} \int_0^{(1-p)/n} H(x)dx & p > 1 - na \end{cases} \tag{4.18}$$

2. In the homogeneous case $F_1 = \cdots = F_n = F$, the bound (4.18) is sharp for $a = c_n$ if (A') and (B) hold.

3. In the homogeneous case $F_1 = \cdots = F_n = F$, if (A) holds for $a \geqslant \frac{1-p}{n}$, then

$$\inf_{S \in \mathfrak{S}_n(F, \cdots, F)} \mathrm{TVaR}_p(S) = \frac{n}{1-p} \int_0^{(1-p)/n} H(x) \mathrm{d}x \tag{4.19}$$

if

$$\inf_{S \in \mathfrak{S}_n(F_J, \cdots, F_J)} \mathbb{P}\left[S > H\left(\frac{1-p}{n}\right)\right] = 0, \tag{4.20}$$

where F_J is the conditional distribution of F on $J = F^{-1}\frac{(n-1)(1-p)}{n}, F^{-1}1 - \frac{1-p}{n}$.

Corollary 4.7 *For a convex function f, if (A) holds, then*

$$\inf_{S \in \mathfrak{S}_n(F_1,\cdots,F_n)} \mathbb{E}\left[f(S)\right] \geqslant n \int_0^a f(H(x))dx + (1-na)f(D(a)). \tag{4.21}$$

Specifically, in the homogeneous case

$$\inf_{S \in \mathfrak{S}_n(F,\cdots,F)} \mathbb{E}\left[f(S)\right] \geqslant n \int_0^a f(H(x))dx + (1-na)f(D(a)), \tag{4.22}$$

and moreover, the equality in (4.26) *holds for $a = c_n$ if (A') and (B) hold.*

Remark 4.2

We can always use discrete distributions to approximate the marginal distributions $F_1, \cdots, F_n$. When using a discrete approximation, the optimization over all possible dependence structures becomes a finite-state problem, and hence it can be solved numerically. For example, Puccetti (2013) used the Rearrangement Algorithm (RA) to calculate the bounds on $TVaR$ over the admissible risk class. There are three notable facts about the merits of our theoretical results compared to the RA approximation. First, our result gives an explicit form and a sharpness condition, while the RA only gives a numerical approximation. Second, although being easy to implement, there is yet no proof that the RA approximation converges to the sharp lower bound on the $TVaR$ as the number of discretization steps m goes to infinity. Third, the RA becomes slow when the dimension n or the number of discretization steps m is large. Our method only requires to numerically find c_n and the complexity does not depend on n. We provide some numerical examples in Section 4.3.

4.2.2 Bounds on convex risk measures of aggregate risk

In practice, information about dependence is limited. Bounds on a convex (or coherent) risk measure $\rho(S)$ over the admissible risk class $\mathfrak{S}_n(F_1, \cdots, F_n)$ are thus of much importance in risk management. The consistency of convex order and convex risk measures is given in Theorem 4.3 of Bäuerle and Müller (2006). Since it is well-known that the convex ordering upper bound of $\mathfrak{S}_n(F_1, \cdots, F_n)$ is given by the co-monotonic scenario of $\mathbf{X}$, a sharp upper bound on $\rho(S)$ over $S \in \mathfrak{S}_n(F_1, \cdots, F_n)$ is $\rho(nF^{-1}(U))$ where $U \sim \mathcal{U}[0,1]$ and it is well-discussed in the literature (for a review, see

Dhaene et al. (2006)). On the other hand, the lower bound on $\rho(S)$ over $S \in \mathfrak{S}_n(F_1, \cdots, F_n)$ is unknown in the literature except for $n = 2$. Using the results in Section 4.1, we are able to give a lower bound on $\rho(S)$, as follows:

Corollary 4.8 *For any risk measure ρ satisfying (5-7), i.e. law-invariant, convex, L^1-Fatou, if (A) holds, then*

$$\inf_{S \in \mathfrak{S}_n(F_1, \cdots, F_n)} \rho(S) \geqslant \rho(T_a), \tag{4.23}$$

where T_a is defined by (4.3). *Moreover, in the homogeneous case $F_1 = \cdots = F_n = F$, if (A') and (B) hold, then the above bound is sharp for $a = c_n$ and*

$$\{\rho(S) : S \in \mathfrak{S}_n(F, \cdots, F)\} = [\rho(T_{c_n}), \rho(nF^{-1}(U))], \tag{4.24}$$

where $U \sim \mathcal{U}[0,1]$.

Proof. The inequality (4.23) is a corollary of Theorem 4.5 in this paper and Theorem 4.3 of Bäuerle and Müller (2006). The sharpness in the homogeneous case is implied by Theorem 4.3. The property (4.24) is guaranteed by Theorem 4.1 (i). ■

Remark 4.3 *Note that we assume finite means for $F, F_1, \cdots, F_n$, thus only the behavior of ρ on $L^1(\Omega, \mathcal{A}, \mathbb{P})$ matters. In Corollary 4.8, we do not require ρ to satisfy (1,2), and thus ρ is not necessarily a convex risk measure as defined in Föllmer and Schied (2002) and does not necessarily have a financial interpretation. A law-invariant coherent risk measure with the Fatou property is thus only a special case in this corollary.*

Remark 4.4 *Since the VaR does not satisfy the convexity (5), Corollary 4.8 does not provide its lower bounds. However, similar ideas based on completely mixable distributions can be used to find sharp bounds on VaR over the admissible risk class $\mathfrak{S}_n(F_1, \cdots, F_n)$. This is not the focus of this paper. The readers are referred to Theorem 2.3 and Corollaries 3.5-3.6 of Wang et al. (2013) for some special cases of sharp bounds on the VaR based on the idea of completely mixable distributions.*

4.2.3 Convex expectation and applications in finance and insurance

A *convex (concave) expectation* of a random variable X is defined as $\mathbb{E}[f(X)]$ where $f : \mathbb{R} \to \mathbb{R}$ is a convex (concave) function. If f is convex and bounded, then $\mathbb{E}[f(X)]$ satisfies (5-7) and thus is a risk measure as described in Corollary 4.8. Theoretically, $\mathbb{E}[f(X)]$ can be infinity. By definition of convex order, we have a straightforward corollary about the lower bound on a convex expectation (or upper bound on a concave expectation) over the admissible risk class $\mathfrak{S}_n(F_1, \cdots, F_n)$,

$$\mathbb{E}\left[f(S)\right] = \mathbb{E}\left[f(X_1 + X_2 + \cdots + X_n)\right] \tag{4.25}$$

regardless of $\mathbb{E}[f(S)]$ being finite or infinite. Recall that when f is convex, the upper bound can be computed explicitly with the co-monotonic dependence scenario.

Corollary 4.9 *For a convex function f, if (A) holds, then*

$$\inf_{S \in \mathfrak{S}_n(F_1, \cdots, F_n)} \mathbb{E}\left[f(S)\right] \geqslant n \int_0^a f(H(x))dx + (1 - na)f(D(a)). \tag{4.26}$$

Specifically, in the homogeneous case

$$\inf_{S \in \mathfrak{S}_n(F, \cdots, F)} \mathbb{E}\left[f(S)\right] \geqslant n \int_0^a f(H(x))dx + (1 - na)f(D(a)), \tag{4.27}$$

and moreover, the equality in (4.27) *holds for $a = c_n$ if (A') and (B) hold.*

Remark 4.5 *Corollary 4.9 can be seen as a generalization of Jensen's inequality as* (4.26) *is simply Jensen's inequality when $a = 0$. It can also be seen as a generalization of Theorem 3.5 of Wang and Wang (2011), where monotone densities were assumed.*

Although finite convex expectations can be viewed mathematically as a special case of law-invariant risk measures, the application and financial interpretation of convex expectations are different from those of risk measures. Some quantities of interest that can be viewed as a convex or concave expectation of the aggregate risk S include:

1. the variance of a joint portfolio S with dependent assets because $\mathbb{E}[S]$ is a constant and $f(S) = (S - \mathbb{E}[S])^2$ is convex.

2. the price of a European basket option written on a joint portfolio of assets with values $X_1, \cdots, X_n$ at a future time T. Precisely, a European

basket call option (respectively a European basket put option) with strike K and maturity T has the following price

$$\mathbb{E}_Q\left[D_T(S-K)^+\right] \quad \text{resp.} \quad \mathbb{E}_Q\left[D_T(K-S)^+\right]$$

where Q is a risk-neutral measure and $S = \sum_i X_i$. If interest rates are deterministic, this is a convex expectation as D_T can be factored out of the expectation. When interest rates are stochastic, we can use a change of numéraire with the zero-coupon bond and introduce the forward neutral risk measure Q_T. The basket call price becomes $P(0,T)E_{Q_T}(S-K)^+$ (where $P(0,T)$ is the price of the zero-coupon bond at time 0), which is a convex expectation.

3. the expected utility of a joint portfolio S for risk-avoiding or risk-seeking utilities. An investor or portfolio manager can be concerned about the expected utility $\mathbb{E}[u(S)]$ of the portfolio. Her utility function $u(\cdot)$ is typically a concave (or convex) function (for instance the exponential utility function is given by $u(x) = 1 - e^{-kx}$ for $k \geqslant 0$). When the dependence of $\mathbf{X}$ is unknown, the upper bound on $\mathbb{E}[u(S)]$ given by Corollary 4.9 can be useful to investors to make decisions.

4. the stop-loss premium of an aggregate loss S with dependent risks. Consider for instance an insurance company with n customers: X_i denotes the potential loss for policyholder i and S denotes the insurer's aggregate risk exposure. The insurer is interested for example in the variance of S, or in $\mathbb{E}[(S-K)^+]$ for some level K. The latter quantity is the stop-loss (net) premium, which is important for stop-loss reinsurance with retention K on the aggregate loss S.

5. the price of a European option on realized variance of an asset price process S_t with partition $t_0, \cdots, t_n$. In this case $X_i = (\ln(S_{t_i}) - \ln(S_{t_{i-1}}))^2$, for $i = 1, \cdots, n$; the price of a call option on the realized level of variance associated with the partition $0 = t_0 < t_1 < \cdots < t_n = T$ of the time interval $[0,T]$ is

$$\mathbb{E}_Q\left[\left(D_T \sum_{i=0}^{n-1} \ln \frac{S_{t_{i+1}}}{S_{t_i}}^2 - K\right)^+\right], \tag{4.28}$$

where the underlying asset price is denoted by S_t (see Carr and Lee (2009)). We assume that the distribution of log increments $(\ln \frac{S_{t_{i+1}}}{S_{t_i}})$ is known but that their dependence is not perfectly known (in particular, they are not necessarily independent). These bounds can be useful to detect arbitrage (see for example Tankov (2011)).

6. the expected n-period return $\mathbb{E}[S_n/S_0] = \mathbb{E}[\exp\{X_1 + \cdots + X_n\}]$ with dependent single-period return rates;

7. some convex risk measures of an aggregate risk S, such as the entropic risk measure (as defined in Section 4.2.1).

Bounds on convex or concave expectations help to analyze risks under best or worst case scenarios when the information on dependence is unreliable. The last section gives some further illustration and proposes a method to check property (B) numerically.

4.3 Numerical Illustrations

Considering the conditions (A), (A') and (B) are sometimes difficult to check, we give some numerical illustrations in this section. As mentioned in Remark 4.2, a natural idea is to construct a discretization of the marginal distributions $F_1, \cdots, F_n$, then the optimization over all possible dependence structures becomes a finite-state problem and is always solvable. For each discretization, we find the optimal discrete structure with respect to minimal convex order, and compare some quantities such as variance and $TVaR$ with our theoretical results.

4.3.1 Rearrangement algorithm

The Rearrangement Algorithm (RA) introduced in Puccetti and Rüchendorf (2012) and also used in Embrechts et al. (2013) and Puccetti (2013) is a quick algorithm to provide discrete numerical approximations for the optimal structure with respect to minimal convex order. In the following, the RA is used to approximate the lower bound on $\mathbb{E}[f(S)]$ for some convex functions f and for $TVaR_p(S)$ when $p = 0.95$. We compare the RA approximation with the lower bound suggested by Theorem 4.6 and Corollary 4.7. The numerical results suggest that the lower bound for homogeneous risks is very likely to be sharp (and thus that (B) is satisfied thanks to Theorem 4.4).

4.3.2 Homogeneous case

In this section, we compare the RA approximation with the lower bound suggested by Corollary 4.7 for different settings for homogeneous risks. We take the number of discretization steps in the RA as $m = 10^5$.

Table 1: RA results vs theoretical bounds, homogeneous case

We consider 6 different settings: Pareto(1,3), Pareto(1,4) with $n = 4$, Gamma(2,0.5), Gamma(3,1) with $n = 3$, Log-Normal(0,1) with $n = 3$ and $n = 10$. Four quantities are calculated: the variance: $f(S) = (S - n\mu)^2$ where $\mu = \mathbb{E}[X_1]$, the European call option prices $f(S) = (S-K)^+$ when $K = n\mu$ and when $K = n\mu - \frac{n\sigma}{4}$ where $\sigma = \sum_{i=1}^{n} \sqrt{var(X_i)}$, and the $TVaR$ of S at level 95%.

	Pareto(θ, α); n		Gamma(α, β); n		Log-Normal(μ, σ^2); n	
	(1,3); 4	(1,4); 4	(2,0.5); 3	(3,1); 3	(0,1); 3	(0,1); 10
	Variance					
RA	1.2903	0.2567	0.3014	0.0236	2.5220	1.6649
Corollary 4.7	1.2904	0.2562	0.3002	0.0235	2.5071	1.6668
	Option Price when $K = n\mu$					
RA	0.2318	0.1111	0.1866	0.0510	0.6230	0.1615
Corollary 4.7	0.2317	0.1112	0.1865	0.0510	0.6227	0.1614
	Option Price when $K = n\mu - n\sigma/4$					
RA	0.8482	0.4695	2.0268	1.2461	1.3301	5.2285
Corollary 4.7	0.8481	0.4694	2.0073	1.2342	1.3186	5.2234
	$TVaR$ at level 0.95					
RA	9.4729	6.9996	15.1148	10.0058	13.0479	20.3635
Theorem 4.6	9.4748	6.9999	15.1148	10.0058	13.0483	20.3623
Independent	11.0538	8.1348	24.2688	16.2819	16.4913	35.4328
Co-monotonic	16.2100	11.2968	35.5736	22.7693	25.6970	85.6566

Numerical results are given in Table 1 in 6 different settings. For each setting, we also give the $TVaR$ under the assumption of independence and co-monotonicity to show the impact of various dependence assumptions. Note that the Gamma and Lognormal distributions above do not have a decreasing density and therefore theoretically we do not know whether they satisfy (B), while for the Pareto distributions we know the bounds given in Theorem 4.6 and Corollary 4.7 are sharp.

From Table 1, we conclude that the bounds obtained for homogeneous risks in Theorems 4.3 and 4.6 are very likely to be sharp for all above distributions, and the structure described in Theorem 4.3 is likely to be optimal.

4.3.3 Heterogeneous Case

In the heterogeneous case, we give a numerical example with 3 different Pareto distributions and n Log-Normal distributions in Table 2.

From Table 2, it appears that the bounds given in Theorem 4.5 are not sharp in general. Note also that the theoretical bounds tend to be more precise when the distributions are similar. This confirms the intuition provided when deriving the lower bound for heterogeneous risks in Section 4.5.

4.3.4 Checking condition (B)

Recall that Condition (B) in Theorem 4.17 corresponds to checking that F is n-CM on the interval $I = [F^{-1}((n-1)c_n), F^{-1}(1-c_n)]$. This is equivalent to

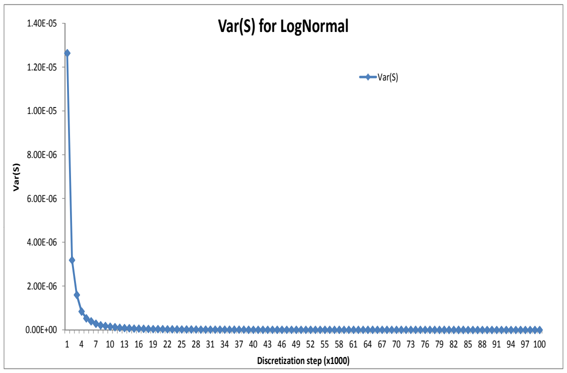
$$\underline{var}(S) := \inf_{S \in \mathfrak{S}_n(F_I, \cdots, F_I)} var(S) = 0 \tag{4.29}$$

where F_I is the conditional distribution of F on I. Since the RA gives a discrete approximation of the optimal dependence structure, (4.29) holds if the RA approximation of $\underline{var}(S)$, denoted by $\underline{var}(S)_m$, goes to zero when the number of discretization steps m goes to infinity (however, in the opposite direction, (4.29) does not imply that $\underline{var}(S)_m \to 0$ since the convergence of the RA approximation is not proved). To illustrate this convergence of the rearrangement algorithm, we represent in Figures 7 and 8 the variance of the sum of n risks for different distributions as a function of the discretization step m.

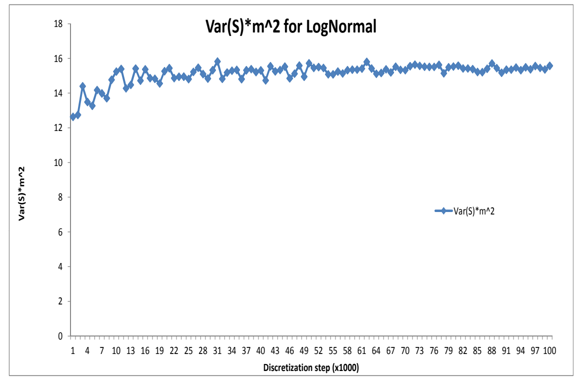
Table 2: RA results vs theoretical bounds, heterogeneous case

Four quantities are calculated: the variance: $f(S) = (S - n\mu)^2$ where $\mu = \mathbb{E}[X_1]$, the European call option prices $f(S) = (S-K)^+$ when $K = n\mu$ and when $K = n\mu - \frac{n\sigma}{4}$ where $\sigma = \sum_{i=1}^{n} \sqrt{var(X_i)}$, and the $TVaR$ of S at level 95%.

	Pareto(θ,α), $n=3$ $S = X_1+X_2+X_3$		$X_i \sim$Log-Normal$(i/10,1), i=1,\cdots,n$		
			$n=3$	$n=5$	$n=10$
	(1,3); (1,4); (1,5)	(1,3.5); (1,4); (1,4.5)	$\sum_{i=1}^{3} X_i$	$\sum_{i=1}^{5} X_i$	$\sum_{i=1}^{10} X_i$
	Variance				
RA	0.3649	0.3298	9.0348	10.2393	14.9596
Corollary 4.7	0.2996	0.2985	8.9000	9.6031	10.5260
	Option price with strike $K = n\mu$				
RA	0.1504	0.1384	0.7699	0.6415	0.4656
Corollary 4.7	0.1438	0.1347	0.7635	0.6155	0.3560
	Option price with strike $K = n\mu - n\sigma/4$				
RA	0.6361	0.4545	1.7878	2.9796	5.9592
Corollary 4.7	0.6361	0.4545	1.7876	2.9792	5.9579
	$TVaR$ at level 0.95				
RA	6.4234	6.1456	16.0726	21.7753	39.0012
Corollary 4.7	5.8742	5.8117	15.9904	21.4194	23.8094
Independent	7.0898	6.6018	19.9349	30.9346	67.0175
Comonotonic	9.1574	8.6407	31.4487	58.8250	154.8750

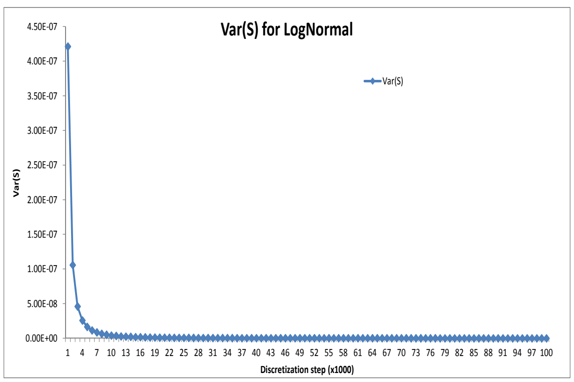


Panel A: $\underline{var}(S)_m$ w.r.t. m

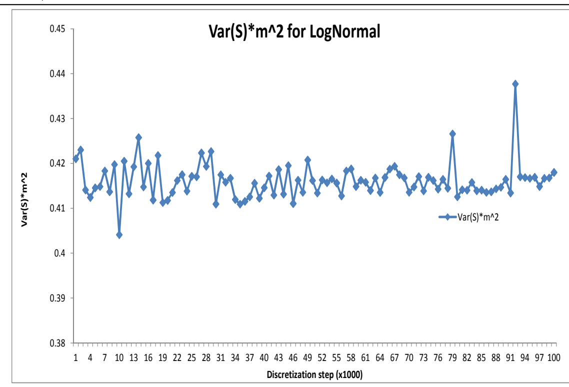


Panel B: $\underline{var}(S)_m \times m^2$ w.r.t. m

Log-Normal(0,1), $n = 3$.



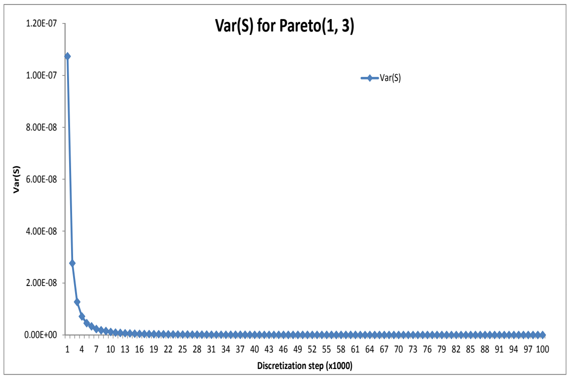
Panel C: $\underline{var}(S)_m$ w.r.t. m

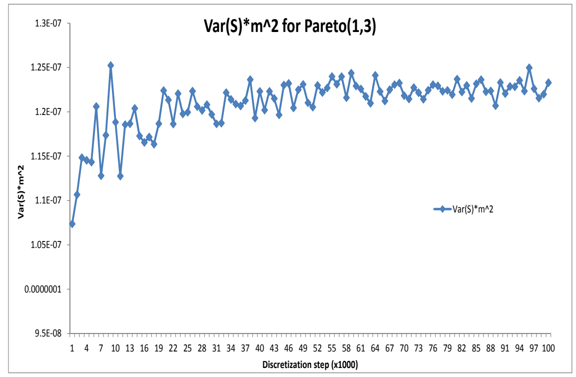


Panel D: $\underline{var}(S)_m \times m^2$ w.r.t. m

Log-Normal(0,1), $n = 10$.

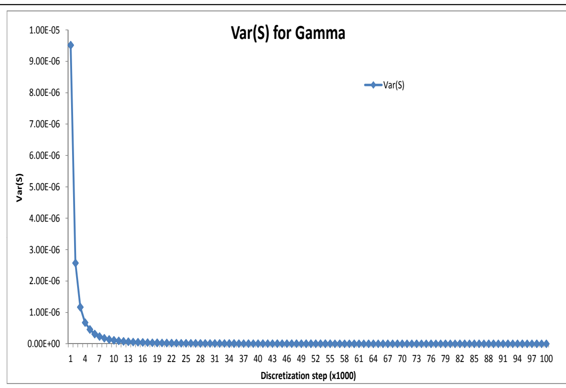
Figure 7: Panels A and C display $\underline{var}(S)_m$ as a function of m for a Pareto distribution and Panels B and D illustrate the speed of convergence in $1/m^2$.

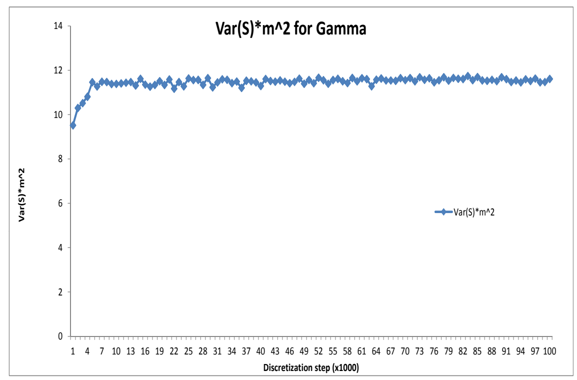




Panel A: $\underline{var}(S)_m$ w.r.t. m Panel B: $\underline{var}(S)_m \times m$ w.r.t. m

Pareto(1,3), $n = 4$





Panel C: $\underline{var}(S)_m$ w.r.t. m Panel D: $\underline{var}(S)_m \times m^2$

Gamma(2,0.5), $n = 3$

Figure 8: Panels A and C display $\underline{var}(S)_m$ as a function of m for a Pareto distribution and a Gamma distribution and Panels B and D illustrate the speed of convergence in $1/m^2$.

From Figures 7 and 8, the RA approximations $\underline{var}(S)_m$ clearly converge to zero, at a rate of m^{-2}. Based on all the observations in Section 5, we have the following conjecture.

Conjecture 4.10 *A Gamma or Log-Normal distribution F is n-CM on the interval $I = [F^{-1}((n-1)c_n), F^{-1}(1-c_n)]$ for any integer n, and the convex ordering bounds in Theorems 4.3 and 4.5 are sharp.*

Even if we are not able to prove this conjecture at this moment, the numerical results clearly show that the lower bounds on convex risk measures and convex expectations are sharp enough to apply in practice, for identical or almost identical marginal distributions.

5 Conclusions and Future Work

This thesis develops different models to measure the distribution of a pool of assets. Explanations from diversification and Modern Portfolio theory are provided to demonstrate the economical reasoning behind the asset pooling. Two different methods in the financial literature are introduced, and improvements on models required in both methods are presented.

The first method, introduced in the second section, improves Fréchet-Hoeffding copula bounds to calculate a model free upper and lower bounds for aggregate assets evaluation. $A^{\mathcal{S},Q}$ and $B^{\mathcal{S},Q}$ are defined as improved point-wise quasi-copula bounds for the class of quasi-copulas with $C(u,v) = Q(u,v),\ \forall(u,v) \in \mathcal{S}$. The rest of the section focuses on conditions on the set $\mathcal{S}$, such that $A^{\mathcal{S},Q}$ and $B^{\mathcal{S},Q}$ are also copulas. As a result, the two functions improve point-wise bounds for the class of copulas with the same constraint as stated above. The major theorem in the section proved that for $\mathcal{S} \subseteq [0,1]^2$, a compact set with both γ_1 and γ_2 non-decreasing functions, satisfying the following property $\forall\,(u,v_0),(u,v_1) \in \mathcal{S},\ \left(u, \frac{v_0+v_1}{2}\right) \in \mathcal{S}$, then $B^{\mathcal{S},Q}$ is a copula.

Numerical methods are suggested in solving improved Fréchet-Hoeffding copula bounds. Duality in terms of the region $\mathcal{S}$ are proposed that allows $A^{\mathcal{S},Q}$ and $B^{\mathcal{S},Q}$ to be copulas. Denote $\mathcal{S}^{\complement} = [0,1]^2 \setminus \mathcal{S}$, numerical results are described to suggest that the class of $A^{\mathcal{S},Q}$ and $B^{\mathcal{S},Q}$ we identified as copulas, $A^{\mathcal{S}^{\complement},Q}$ and $B^{\mathcal{S}^{\complement},Q}$, are copulas as well. Partial results have been shown in theorem 3.8 and Conjecture 3.9. If the hypothesis is true, the class of improved Fréchet-Hoeffding copula bounds can be almost doubled.

In section three, we introduce and investigate the admissible risk class $\mathfrak{S}_n(F_1,\cdots,F_n) = \{X_1 + \cdots + X_n : X_i \sim F_i,\ i = 1,\cdots,n\}$ for given marginal risk distributions $F_1,\cdots,F_n$. We give a new lower bound over $\mathfrak{S}_n(F_1,\cdots,F_n)$. In the homogeneous case, $F_1 = \cdots = F_n$, we give a sufficient condition for the new lower bound to be sharp. The results can be used to find sharp bounds on convex risk measures and other quantities in finance when the dependence information among individual risks is missing. Numerical illustrations suggest that the new lower bound is likely to be sharp for most risk distributions and the conditions used in our main results are usually satisfied.

Some future directions related to this topic include proving Conjecture 4.10. More generally, we expect Conjecture 4.10 to hold for all unimodal densities given some smooth conditions and also for heterogeneous risks under some additional conditions. Recall that the heterogeneous analog of complete mixability is called *joint mixability* and is introduced in Wang et al. (2013). Note that proving Conjecture 4.10 for heterogeneous risks is an open problem even in the case of decreasing densities. Finally, it is of interest to determine conditions under which convex ordering bounds for heterogeneous risks (over $\mathfrak{S}_n(F_1, \cdots, F_n)$) are sharp. We believe that these research directions are all technically challenging and relevant to quantitative risk management.

Bibliography

ALBRECHER, H., MAYER, P., SCHOUTENS, W. (2008). General lower bounds for arithmetic Asian option prices. *Applied Mathematical Finance*, **15**(2), 123–149.

ARTZNER, P., DELBAEN, F., EBER, J.-M., HEATH, D. (1999). Coherent measures of risk. *Mathematical Finance*, **9**(3), 203–228.

D'ASPREMONT, A., EL GHAOUI, L. (2006). Static arbitrage bounds on basket option prices. *Mathematical Programming: Series A and B*, **106**(3), 467–489.

BASEL COMMITTEE ON BANKING SUPERVISION (2010). *Developments in modeling risk aggregation.* Basel: Bank for International Settlements.

BASEL COMMITTEE ON BANKING SUPERVISION (2012). *Consultative Document May 2012. Fundamental review of the trading book.* Basel: Bank for International Settlements. BÄUERLE, N. AND MÜLLER, A. (2006). Stochastic orders and risk measures: Consistency and bounds. *Insurance Mathematics and Economics*, **38**(1), 132–148.

BERNARD, C., BOYLE, P., VANDUFFEL, S. (2011). Explicit Representation of Cost-Efficient Strategies, *Working Paper available at SSRN.*

BERNARD, C., LIU Y., MACGILLIVRAY N., ZHANG J. (2012). Bounds on Capital Requirements For Bivariate Risk with Given Marginals and Partial Information on the Dependence. Working Paper.

BERNARD, C., JIANG, X., VANDUFFEL, S. (2012). Note on Improved Fréchet Bounds. *Journal of Applied Probability*, **49**(3), 866–875.

BERNARD, C., JIANG, X., WANG, R. (2012). Risk Aggregation with Dependence Uncertainty. *Insurance: Mathematics and Economics.*

CARR, P. AND LEE, R. (2009). Volatility derivatives. *Annual Review of Financial Economics*, **1**, 319–339.

CHEUNG, K.C. AND VANDUFFEL, S. (2011). Bounds for sums of random variables when the marginal distributions and the variance of the sum are given. *Scandinavian Actuarial Journal*, DOI:10.1080/03461238.2011.558186.

DEELSTRA, G., DHAENE J. AND VANMAELE, M. (2011). An overview of comonotonicity and its applications in finance and insurance, in: *Advanced Mathematical Methods for Finance,* Di Nunno, G. and Øksendal B. (Eds.), Springer, 2011.

DELBAEN, F. (2002). Coherent measures of risk on general probability spaces. In: *Advances in Finance and Stochastics. Essays in Honour of Dieter Sondermann*, Springer-Verlag, 1–37.

DELBAEN, F. (2009). Risk measures for non-integrable random variables. *Mathematical Finance,* **19**(2), 329–333.

DEHEUVEL, P. (1978). Caractérisation complete des lois extrêmes multivariés et de la convergence des types extrêmes, *Publications de l'Institut de Statistique de l'Université de Paris*, 23, 1-36

DEHEUVEL, P. (1981). A non parametric test for independence, Publications de l'Institut de Statistique de *l'Université de Paris*, 26, 29-50

DENUIT, M., GENEST, C. AND MARCEAU, É. (1999). Stochastic bounds on sums of dependent risks. *Insurance: Mathematics and Economics*, **25**, 85–104.

DHAENE, J., DENUIT, M., GOOVAERTS, M.J., KAAS, R. AND VYNCHE, D. (2002). The concept of comonotonicity in actuarial science and finance: Theory. *Insurance: Mathematics and Economics* 31(1), 3-33.

DHAENE, J., DENUIT, M., GOOVAERTS, M.J., KAAS, R., TANG, Q. AND VYNCHE, D. (2006). Risk measures and comonotonicity: a review. *Stochastic models*, **22**, 573–606.

DHAENE, J., LINDERS, D., SCHOUTENS, W. AND VYNCKE, D. (2012). The Herd Behavior Index: A new measure for the implied degree of comovement in stock markets *Insurance: Mathematics and Economics,* **50**, 357-370.

DURRLEMAN, V. NIKEGHBALI A., AND RONCALLI T. (2001). Which copula is the right one? *Working paper*, Credit Lyonnais.

EMBRECHTS, P., HÖING, A. AND JURI, A. (2003). Using copulas to bound the value-at-risk for functions of dependent risks. *Finance and Stochastics*, **7**, 145–167.

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

Embrechts, P. and Puccetti, G. (2010). Risk aggregation. In: *Copula Theory and Its Applications*, P. Jaworski et al. (eds.), Lecture Notes in Statistics **198**, Springer, Berlin/Heidelberg.

Embrechts, P., Puccetti, G. and Rüchendorf, L. (2013) Model uncertainty and *VaR* aggregation. *Preprint, ETHZ.*

Frechet M. (1951). "Sur les tableaux de corrélation dont les marges sont données,"*Ann. Univ. Lyon Sect.A,* Series 3, 14, 53-77.

Föllmer, H. and Schied, A. (2002). Convex measures of risk and trading constraints. *Finance and Stochastics*, **6**, 429–447.

Föllmer, H., Schied, A. (2010). Coherent and convex risk measures, in *Encyclopedia of Quantitative Finance*, Cont, R. (ed.) New York: John Wiley & Sons, 355-363.

Hoeffding, W. (1940). "Masstabinvariante Korrelationstheorie," *Schriften des mathematischen Instituts und des Instituts für angewandte Mathematik der Universität Berlin* 5, 179-233.

Hobson, D., Laurence, P. and Wang, T.-H. (2005). Static-arbitrage upper and lower bounds for the prices of basket options. *Quantitative Finance*, **5**(4), 329–342.

Kaas, R., Laeven, R. and Nelsen, R. (2009). Worst *VaR* scenarios with given marginals and measures of association. *Insurance: Mathematics and Economics*, **44**, 146–158.

Kaas, R., Dhaene, J. and Goovaerts, M. (2000). Upper and lower bounds for sums of random variables. *Insurance: Mathematics and Economics*, **27**(2), 151–168.

Kaina, M. and L. Rüchendorf, L. (2009). On convex risk measures on L^p-spaces. *Mathematical Methods of Operations Research*, **69**, 475–495.

Keller-Ressel, M. and G. Clauss (2012). Convex order of discrete, continuous and predictable quadratic variation & Applications to options on variance. *Working paper available on ArXiv.*

KUSUOKA, S. (2001). On law invariant coherent risk measures. *Advances in Mathematical Economics*, **3**, 83–95.

Markowitz, H.M. (1952). Portfolio Selection. *The Journal of Finance,* 77–91.

MCNEIL, A. J., FREY, R. AND EMBRECHTS, P. (2005). *Quantitative Risk Management: Concepts, Techniques, Tools*. Princeton, NJ: Princeton University Press.

MEILIJSON, I., NADAS, A. (1979). Convex majorization with an application to the length of critical paths. *Journal of Applied Probability*, **16**, 671–677.

MÜLLER, A. (1997). Stop-loss order for portfolios of dependent risks. *Insurance: Mathematics and Economics*, **21**, 219-223.

MÜLLER, A. (1997). Stochastic orderings generated by integrals: A unified study. *Advances in Applied Probability*, **29**, 414-428.

NELSEN, R. (2006). "An introduction to copulas", 2nd edition, Springer series in Statistics.

PANJER, H. (2006). *Operational Risk: Modeling Analytics*, John Wiley & Sons.

PUCCETTI, G. (2013). Sharp bounds on the expected shortfall for a sum of dependent random variables. *Statistics and Probability Letters,* **83**(4), 1227-1232.

PUCCETTI, G. AND RÜCHENDORF L. (2012). Computation of sharp bounds on the distribution of a function of dependent risks. *Journal of Computational and Applied Maths*, **236** (7), 1833-1840.

PUCCETTI, G., WANG, B. AND WANG, R. (2012). Advances in complete mixability. *Journal of Applied Probability,* **49**(2), 430–440.

RACHEV, S.T. AND RÜCHENDORF, L. (1994), "Solution of some transportation problems with relaxed or additional constraints", *SIAM J. Control and Optimization*, 32, pp. 673-689.

RACHEV, S.T. AND RÜCHENDORF, L. (1998), "Mass Transportation Problems", Springer.

ROGERS, L.C.G. AND SHI, Z. (1995). The value of an Asian option. *Journal of Applied Probability,* **32**(1), 1077–1088.

RÜCHENDORF, L. (1980). Inequalities for the expectation of Δ-monotone functions, *Zeitschrift f Wahrscheinlichkeitstheorie*, **54**, 341–349.

RÜCHENDORF, L. (1982). Random variables with maximum sums. *Advances in Applied Probability*, **14**(3), 623–632.

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

SVINDLAND, G. (2008). Convex Risk Measures Beyond Bounded Risk. *Ph.D. Thesis,* Ludwig Maximilian Université Müchen.

SKLAR, A. (1959), Fonctions de répartition à n dimensions et leurs marges, Publ. Inst. Statist. Univ. Paris 8: 229–231

TANKOV, P. (2011). Improved Fréhet bounds and model-free pricing of multi-asset options. *Journal of Applied Probability*, **48**, 389–403.

TCHEN, A. (1980). Inequalities for distributions with given marginals. *The Annals of Probability*, **8**(4): 814–827.

VALDEZ, E.A., DHAENE, J., MAJ, M. AND VANDUFFEL, S. (2009). Bounds and approximations for sums of dependent log-elliptical random variables. *Insurance: Mathematics and Economics*, **44**(3): 385–397.

WANG, R., PENG, L. AND YANG, J. (2013). Bounds for the sum of dependent risks and worst Value-at-Risk with monotone marginal densities. *Finance and Stochastics*, **17**(2), 395-417.

WANG, B. AND WANG, R. (2011). The complete mixability and convex minimization problems for monotone marginal distributions. *Journal of Multivariate Analysis*, **102**, 1344–1360.

WANG, S. (2000). A class of distortion operations for pricing financial and insurance risks, *Journal of Risk and Insurance*, **67**, 15–36