On Aggregation Sets and Lower-Convex Sets

Tiantian Mao*† and Ruodu Wang†

November 24, 2014

Abstract

It has been a challenge to characterize the set of all possible sums of random variables with given marginal distributions, referred to as an aggregation set in this paper. We study the aggregation set via its connection to the corresponding lower-convex set, which is the set of all sums of random variables that are smaller than the respective marginal distributions in convex order. Theoretical properties of the two sets are discussed, assuming that all marginal distributions have finite mean. In particular, an aggregation set is always a subset of its corresponding lower-convex set, and the two sets are identical in the asymptotic sense after scaling. We also show that a lower-convex set is identical to the set of comonotonic sums with the same marginal constraint. The main theoretical results contribute to the field of multivariate distributions with fixed margins.

Key-words: aggregation set; convex order; comonotonicity; dependence uncertainty; Fréchet classes.

1 Introduction

The study of probability measures with given margins has been an active field in multivariate probability theory for a long time; see for instance Strassen (1965). One of the challenging questions in this field is to determine all possible distributions of $S_n = X_1 + \cdots + X_n$ for given distributions $F_1, \ldots, F_n$, where $X_i \sim F_i$, $i = 1, \ldots, n$, are random variables in a standard probability space $(\Omega, \mathcal{F}, \mathbb{P})$, assumed to be atomless unless otherwise specified.

This question has raised a lot of attention in the recent research of *dependence uncertainty* in quantitative risk management. To be more precise, in the modeling of an aggregate risk S_n, model uncertainty lies at both the level of the marginal distributions $F_1, \ldots, F_n$, and at the

*Department of Statistics and Finance, University of Science and Technology of China, Hefei, Anhui 230026, China.

†Department of Statistics and Actuarial Science, University of Waterloo, Waterloo, ON N2L3G1, Canada. Emails: `tmao@ustc.edu.cn` (T. Mao), `wang@uwaterloo.ca` (R. Wang).

level of the joint distribution of $(X_1, \dots, X_n)$ (i.e. dependence uncertainty). In the practice of quantitative risk management, one often has reliable information on the marginal distributions, but very little information on the joint distribution; see Embrechts et al. (2013) for examples in the context of operational risk. With dependence uncertainty, one has to find bounds on quantities of interest over all possible models of S_n in the set

$$\mathcal{D}_n = \mathcal{D}_n(F_1, \dots, F_n) = \{X_1 + \dots + X_n,\ X_i \sim F_i,\ i = 1, \dots, n\}, \tag{1.1}$$

which we call an *aggregation set* in this paper. For instance, the calculation of $\sup\{\rho(S_n) : S_n \in \mathcal{D}_n\}$, where ρ is a risk measure, is useful in obtaining conservative values of ρ, a practical concern in risk management with model uncertainty. This problem and its corresponding infimum problem $\inf\{\rho(S_n) : S_n \in \mathcal{D}_n\}$ have recently been studied in Wang et al. (2013); Embrechts et al. (2013, 2014b); Puccetti et al. (2013); Bernard et al. (2013, 2014) for the most popular regulatory risk measures Value-at-Risk and Expected Shortfall. We refer to the survey paper Embrechts et al. (2014a) for an overview and a history of this topic.

The core question is to characterize the aggregation set $\mathcal{D}_n$. It is well known that even in the case of $n = 2$, the characterization of $\mathcal{D}_2$ is generally an open question; see Bernard et al. (2014). In the latter paper $\mathcal{D}_n$ is called an *admissible risk class* from a risk management perspective, and some properties of $\mathcal{D}_n$ are discussed. A frequently studied question in recent research is to determine whether $\mathcal{D}_n$ contains a constant random variable, in which case we call that $F_1, \dots, F_n$ are *jointly mixable* (see Wang et al., 2013). Apparently the characterization of $\mathcal{D}_n$ is a more ambitious aim than the determination of joint mixability; even the latter is a challenging open question, only solved for some specific classes of marginal distributions (see for instance Wang and Wang, 2011). Many contributions to the research on $\mathcal{D}_n$ are made by using copula and mass-transportation techniques; we refer the interested reader to Rüschendorf (2013) for a comprehensive overview. The study of $\mathcal{D}_n$ generally belongs to the field of research on *Fréchet classes* and *distributions with marginal constraints*; see for instance Joe (1997) from a copula perspective.

In this paper, we study $\mathcal{D}_n$ by connecting it with the following set

$$\mathcal{C}_n = \mathcal{C}_n(F_1, \dots, F_n) = \{X_1 + \dots + X_n,\ X_i \leqslant_{\mathrm{cx}} Y_i,\ Y_i \sim F_i,\ i = 1, \dots, n\}, \tag{1.2}$$

which we call a *lower-convex set* (a lower set with respect to convex order). Here, $\leqslant_{\mathrm{cx}}$ represents convex order. When $n = 1$, we use the notation

$$\mathcal{C}(F) = \mathcal{C}_1(F) = \{X : X \leqslant_{\mathrm{cx}} Y,\ Y \sim F\}. \tag{1.3}$$

We assume distributions $F_1, \dots, F_n$ have finite mean in this paper. Convex order is consistent with risk preferences in economic decision theory; see for instance Yaari (1987). As such, $\mathcal{C}_n$

contains all aggregate risks $\sum_{i=1}^{n} X_i$ such that for each $i = 1, \dots, n$, X_i is preferred compared to an F_i-distributed risk via convex order; in quantitative risk management it can be interpreted as a set of acceptable risks with marginal constraints.

We investigate properties of the sets $\mathcal{D}_n$ and $\mathcal{C}_n$, and in particular, $\mathcal{C}_n$ is closed with respect to L^1-convergence, and it can be fully characterized as the set of random variables smaller than a comonotonic sum in convex order. One of the main contributions of this paper is to show that in the homogeneous setting when $F_1 = \cdots = F_n$, $\mathcal{D}_n$ has an upper limit $\mathcal{C}$ after scaling by $1/n$, as $n \to \infty$. This result is a complement to the laws of large numbers. It presents all the possible limits of $(X_1 + \cdots + X_n)/n$ as $n \to \infty$ by removing the assumption of independence, that is, allowing arbitrary dependence among the sequence of random variables. Another main contribution is to show that all the elements in $\mathcal{C}_n(F_1, \dots, F_n)$ can be written as a *comonotonic* sum (see Dhaene et al., 2002, for comonotonicity) of random variables $X_1, \dots, X_n$ which are smaller than $F_1, \dots, F_n$ in convex order, respectively. We also give some direct implications of our main results in the theory of risk measures.

The rest of the paper is organized as follows. Some preliminaries on convex order are given in Section 2. In Section 3, we study some theoretical properties and the asymptotic behavior of $\mathcal{D}_n$, and identify its limit in the homogeneous setting. In Section 4, we show the equivalence between $\mathcal{C}_n$ and the set of corresponding comonotonic sums. A conclusion is drawn in Section 5.

2 Preliminaries on convex order

Recall that a random variable X is called smaller than another random variable Y in *convex order*, denoted by $X \leqslant_{\mathrm{cx}} Y$, if

$$\mathbb{E}[\phi(X)] \leqslant \mathbb{E}[\phi(Y)] \quad \text{for all convex} \quad \phi : \mathbb{R} \to \mathbb{R},$$

provided that both expectations exist. We also write $F \leqslant_{\mathrm{cx}} G$ if $X \leqslant_{\mathrm{cx}} Y$, $X \sim F$ and $Y \sim G$. Standard references for convex order can be found in Müller and Stoyan (2002) and Shaked and Shanthikumar (2007). Throughout, we say that a distribution or a random variable is integrable if it has finite mean, and we use L^1 for the set of integrable random variables. In the paper, we mainly focus on integrable distributions, which are the main subject in the study of convex order; for instance, integrability is required in the definition of convex order in Müller and Stoyan (2002, Definition 1.5.1).

There is a martingale characterization about the convex order which is useful for understanding convex order and will be used several times later.

Lemma 2.1. (Theorem 3.A.4, Shaked and Shanthikumar (2007)) *The L^1 random variables X*

and Y satisfy $X \leqslant_{\mathrm{cx}} Y$ if, and only if, there exist two random variables $\hat{X}$ and $\hat{Y}$, defined on the same probability space, such that

$$\hat{X} \stackrel{d}{=} X, \quad \hat{Y} \stackrel{d}{=} Y \quad \text{and} \quad \mathbb{E}[\hat{Y}|\hat{X}] = \hat{X} \quad a.s.$$

Convex order is a stochastic order to compare the variability of random variables. There is extensive research on transfers of mass between two random variables that are ordered by $\leqslant_{\mathrm{cx}}$. In the following we state a result established by Rothschild and Stiglitza (1970); see also Theorem 1.5.29 of Müller and Stoyan (2002) and Theorem 2.5.4 of Müller (2013). We need the following definition of *mean preserving spreads*, see Rothschild and Stiglitza (1970) or Definition 1.5.28 of Müller and Stoyan (2002) for more details.

Definition 2.1. Let F and G be distribution functions of discrete distributions whose union support is a finite set of points $x_1 < x_2 < \cdots < x_n$ with probability mass functions f and g respectively. Then G is said to be a *mean preserving spread* of F, if they have the same mean and there exists $i \in \{2, \dots, n-1\}$ such that

$$g(x_{i-1}) \geqslant f(x_{i-1}), \quad g(x_i) \leqslant f(x_i), \quad g(x_{i+1}) \geqslant f(x_{i+1}),$$

and

$$g(x_j) = f(x_j), \quad j \notin \{i-1, i, i+1\}.$$

Lemma 2.2. *Suppose that F and G are two distribution functions supported in finite sets. Then $F \leqslant_{\mathrm{cx}} G$ is equivalent to that there is a finite sequence $F_1, \dots, F_k$ with $F_1 = F$ and $F_k = G$, such that F_{i+1} is a mean preserving spread of F_i for $i = 1, \dots, k-1$, i.e., G differs from F by finitely many mean preserving spreads.*

In the following sections, we denote $F^{-1}(t) = \inf\{x : F(x) \geqslant t\}$, $t \in (0, 1]$ for any distribution function F. Two random variables X and Y are said to be *comonotonic*, if there exists a random variable U and two non-decreasing functions f, g such that $X = f(U)$ and $Y = g(U)$ almost surely. Such U can be chosen as $\mathrm{U}[0, 1]$ distributed, and f and g can be chosen as the inverse distribution functions of X and Y. For any distributions F and G, we denote by $F \oplus G$ the distribution of the sum of comonotonic random variables with respective distributions F and G. In other words, $F \oplus G$ is the distribution of $F^{-1}(U) + G^{-1}(U)$ where $U \sim \mathrm{U}[0, 1]$.

3 Aggregation sets and lower-convex sets

3.1 Basic properties

First, the inclusion of $\mathcal{D}_n$ in $\mathcal{C}_n$ follows directly from the definitions of $\mathcal{D}_n$ and $\mathcal{C}_n$ in (1.1) and (1.2). This simple fact will be used repeatedly, and hence we state it here as a proposition.

Proposition 3.1. *$\mathcal{D}_n(F_1,\dots,F_n) \subset \mathcal{C}_n(F_1,\dots,F_n)$ for any integrable distributions $F_1,\dots,F_n$.*

We aim to investigate the aggregation set $\mathcal{D}_n$ by its superset $\mathcal{C}_n$. We first give a closer look at $\mathcal{C}_n$. Recall its definition:

$$\mathcal{C}_n(F_1,\dots,F_n) = \{X_1 + \dots + X_n,\ X_i \leqslant_{\mathrm{cx}} Y_i,\ Y_i \sim F_i,\ i = 1,\dots,n\},$$

where $F_1,\dots,F_n \in \mathcal{F}^1$. We define another set

$$\begin{aligned}\mathcal{C}'_n(F_1,\dots,F_n) &= \{S:\ S \leqslant_{\mathrm{cx}} X_1^c + \dots + X_n^c,\ X_i^c \sim F_i,\ X_i^c, i = 1,\dots,n,\ \text{comonotonic,}\}\\ &= \mathcal{C}(F_1 \oplus \dots \oplus F_n).\end{aligned}$$

We will first show that the two sets $\mathcal{C}_n$ and $\mathcal{C}'_n$ are identical; this result will become very useful in the later analysis. Note that the definition of $\mathcal{C}_n$ involves *arbitrary dependence* as in $\mathcal{D}_n$ (hence it is not straightforward to characterize), whereas $\mathcal{C}'_n$ only concerns a single inequality of convex order and is a fully characterized set.

Proposition 3.2. *$\mathcal{C}'_n(F_1,\dots,F_n) = \mathcal{C}_n(F_1,\dots,F_n)$ for any integrable distributions $F_1,\dots,F_n$.*

Proof. It suffices to prove $\mathcal{C}'_n \subset \mathcal{C}_n$, since the converse $\mathcal{C}_n \subset \mathcal{C}'_n$ follows from Corollary 1 in Dhaene et al. (2002). For any $S \in \mathcal{C}'_n$, by Lemma 2.1, there exist $\hat{S} \overset{d}{=} S$ and $\hat{Y} \overset{d}{=} \sum_{i=1}^n X_i^c$ with $X_i^c \sim F_i$, $i = 1,\dots,n$, comonotonic such that $\mathbb{E}[\hat{Y}|\hat{S}] = \hat{S}$ a.s. Let F_{S^c} denote the distribution function of $\sum_{i=1}^n X_i^c$ and $X_i = \mathbb{E}[F_i^{-1}(F_{S^c}(\hat{Y}))|\hat{S}]$, $i = 1,\dots,n$. Then $X_i \leqslant_{\mathrm{cx}} X_i^c$, since $F_i^{-1}(F_{S^c}(\hat{Y})) \sim F_i$, $i = 1,\dots,n$. Then we have $\sum_{i=1}^n X_i = \mathbb{E}[\sum_{i=1}^n F_i^{-1}(F_{S^c}(\hat{Y}))|\hat{S}] = \mathbb{E}[\hat{Y}|\hat{S}] = S$ a.s. $\square$

From Proposition 3.2, it is straightforward to determine whether $S \in \mathcal{C}_n$ for a given random variable S via checking convex order, whereas $\mathcal{D}_n$ is yet open to characterize. Another property of $\mathcal{C}_n$ that we will use later is the closure property under the weak convergence.

Proposition 3.3. *For any integrable distributions $F_1,\dots,F_n$,*

(i) $\mathcal{C}_n(F_1,\dots,F_n)$ is uniformly integrable;

(ii) $\mathcal{C}_n(F_1,\dots,F_n)$ is closed with respect to the topology induced by weak convergence.

Proof. By Proposition 3.2, we only need to prove that the theorem holds for $n = 1$, since $\mathcal{C}_n(F_1,\dots,F_n) = \mathcal{C}'_n(F_1,\dots,F_n) = \mathcal{C}(F_1 \oplus \dots \oplus F_n)$. (i) follows directly from Elton and Hill (1992, Theorem 4.2). It suffices to prove (ii). Let $X_n \in \mathcal{C}(F)$, $n \in \mathbb{N}$ satisfying that $X_n \xrightarrow{d} X$ as $n \to \infty$. By Theorem 3.2.2 of Durrett (2010), there exist X'_n, $n \in \mathbb{N}$ and X' on the same probability space such that $X'_n \overset{d}{=} X_n$, $n \in \mathbb{N}$, $X' \overset{d}{=} X$ and

$$X'_n \xrightarrow{a.s.} X' \quad \text{as } n \to \infty.$$

It follows that, by the uniform integrability of $\{X_n',\ n \in \mathbb{N}\}$ from (i),

$$\mathbb{E}[Y] = \lim_{n\to\infty} \mathbb{E}[X_n'] = \mathbb{E}[X].$$

For any $t \in \mathbb{R}$, applying Fatou's lemma to the sequence $\{(X_n' - t)_+,\ n \geqslant 1\}$ which converges a.s. to $(X - t)_+$, we have that

$$\mathbb{E}[(X-t)_+] \leqslant \liminf_{n\to\infty} \mathbb{E}[(X_n' - t)_+] \leqslant \mathbb{E}[(Y-t)_+]$$

where the last inequality follows from $X_n' \leqslant_{\mathrm{cx}} Y$ for all $n \in \mathbb{N}$. Thus, we have $X \leqslant_{\mathrm{cx}} Y$. ☐

Remark 3.1. (i) By noting that $\frac{1}{n}(X_1 + \ldots + X_n) \leqslant_{\mathrm{cx}} X$ holds for $X_i \stackrel{d}{=} X$, $i = 1, \ldots, n$, from Theorem 3.3 (i) we can directly obtain the following: suppose that $\{Y_n,\ n \in \mathbb{N}\}$ is a sequence defined as

$$Y_n = \frac{1}{n}\left(X_{n1} + \cdots + X_{nn}\right), \quad n \in \mathbb{N},$$

where $\{X_{ni}\}$ is any triangular array such that $X_{ni} \stackrel{d}{=} X$, $i = 1, \ldots, n$, $n \in \mathbb{N}$, for some integrable random variable X. Then $\{Y_n,\ n \in \mathbb{N}\}$ is uniformly integrable.

(ii) From Theorem 3.3 (ii), the set $\mathcal{C}_n(F_1, \ldots, F_n)$ is also closed with respect to a.s.-convergence and L^1-convergence since the latter two types of convergence are stronger than weak convergence. Moreover, $\mathcal{D}_n$ is closed under the same topology as shown in Bernard et al. (2014).

(iii) If some of the distribution functions $F_1, \ldots, F_n$ are not integrable, then the result in Proposition 3.3 (i) fails to hold since there exists element in $\mathcal{C}_n(F_1, \ldots, F_n)$ which is not integrable. Note that the set $\mathcal{C}_n(F_1, \ldots, F_n)$ is not closed with respect to a.s. convergence, implying that Proposition 3.3 (ii) also fails; see Shaked and Shanthikumar (2007, Theorem 4.A.8).

3.2 Motivating examples

Now that we have $\mathcal{D}_n \subset \mathcal{C}_n$, one naturally wonders about the difference between the two sets. The following example of Bernoulli distributions motivates us to believe that the difference between the two sets can be, in some sense, very small.

Example 3.1. Suppose that $F = \mathrm{Bern}(p)$ for some $p \in [0, 1]$, i.e., for $X \sim F$,

$$\mathbb{P}(X = 0) = 1 - p, \quad \mathbb{P}(X = 1) = p.$$

Denote by $L(\mathbb{N})$ the set of random variables which take values in $\mathbb{N}$. We have that

$$\mathcal{D}_n(F, \ldots, F) = \mathcal{C}_n(F, \ldots, F) \cap L(\mathbb{N}).$$

Proof. Denote $\mathcal{D}_n' = \mathcal{C}_n(F, \dots, F) \cap L(\mathbb{N})$. It is obvious that $\mathcal{D}_n \subset \mathcal{D}_n'$ by Proposition 3.1. For the converse, let $X \in \mathcal{D}_n'$. Then by Proposition 3.2, $X \leqslant_{\mathrm{cx}} nY$ with $Y \sim F$ and X only takes values in $\{1, \dots, n\}$. Suppose that $\mathbb{P}(X = i) = p_i \geqslant 0$, $i = 0, \dots, n$, with $\sum_{i=0}^n p_i = 1$ and $\sum_{i=1}^n i p_i = np$. Define exchangeable random variables $\mathbf{X} = (X_1, \dots, X_n)$ by

$$\mathbb{P}(\mathbf{X} = \sigma_i) = p_i \Big/ \binom{n}{i}, \quad i = 1, \dots, n,$$

where σ_i denotes any permutation of n-dimensional vector $\mathbf{u} = (0, \dots, 0, 1, \dots, 1)$ with $||\mathbf{u}||_1 = i$, $i = 1, \dots, n$, where $||\cdot||_1$ is the L_1-norm defined by $||\mathbf{x}||_1 = \sum_{i=1}^n |x_i|$ for $\mathbf{x} = (x_1, \dots, x_n)$. Then

$$\mathbb{P}(X_i = 1) = \sum_{i=1}^n \frac{i}{n} p_i = \frac{1}{n} np = p, \quad i = 1, \dots, n.$$

and

$$\mathbb{P}\left(\sum_{i=1}^n X_i = j\right) = \sum_{\sigma_j} \mathbb{P}(\mathbf{X} = \sigma_j) = p_j, \quad j = 0, \dots, n.$$

This means $X = \sum_{i=1}^n X_i \in \mathcal{D}_n$. □

Motivated by Example 3.1, one may wonder whether the two sets $\mathcal{D}_n$ and $\mathcal{C}_n \cap L$ are identical, where L is the set of random variables with the proper range. However, this is not true in general, even for some very simple choices of marginal distributions. The following two examples show that $\mathcal{D}_n$ is strictly smaller than $\mathcal{C}_n \cap L$ for the case of tri-atomic distributions and uniform distributions. We hope those simple examples help the reader to understand challenges arising in problems related to $\mathcal{D}_n$. We remark that the only fully characterized aggregation sets $\mathcal{D}_n$ so far are those of Bernoulli distributions.

Example 3.2. Suppose that F is a tri-atomic distribution, i.e, for $X \sim F$,

$$\mathbb{P}(X = 0) = p, \quad \mathbb{P}(X = 1) = q, \quad \mathbb{P}(X = 2) = 1 - p - q,$$

for some $p, q \geqslant 0$ and $p + q \leqslant 1$. Then, it generally holds that

$$\mathcal{D}_n(F, \dots, F) \neq \mathcal{C}_n(F, \dots, F) \cap L(\mathbb{N}).$$

Proof. We show by providing a counter-example in the case of $n = 2$. Let $p = q = 1/3$ and Y be a random variable defined by

$$\mathbb{P}(Y = 1) = \mathbb{P}(Y = 3) = 1/2.$$

It is easy to see that $Y \leqslant_{\mathrm{cx}} 2X$; hence $Y \in \mathcal{C}_2(F, F) \cap L(\mathbb{N})$. We show that one cannot find $X_1, X_2 \sim F$ such that $Y = X_1 + X_2$. Suppose that $Y = X_1 + X_2$ for some random variables $X_1, X_2 \sim F$. Note that $\mathbb{P}(Y = 1) = \mathbb{P}(X_1 = 0, X_2 = 1) + \mathbb{P}(X_1 = 1, X_2 = 0)$, and

$\mathbb{P}(Y=3)=\mathbb{P}(X_1=1,X_2=2)+\mathbb{P}(X_1=2,X_2=1)$. Since $\mathbb{P}(Y=1)+\mathbb{P}(Y=3)=1$, we have that

$$\mathbb{P}(X_1=0,X_2=1)+\mathbb{P}(X_1=1,X_2=0)+\mathbb{P}(X_1=1,X_2=2)+\mathbb{P}(X_1=2,X_2=1)=1.$$

It follows that $\{X_1=0\}\cup\{X_1=2\}=\{X_2=1\}$ a.s. However $\mathbb{P}(\{X_1=0\}\cup\{X_1=2\})=2/3>\mathbb{P}(X_2=1)$. The contradiction shows that $Y\notin\mathcal{D}_2(F,F)$. $\square$

Example 3.3. Suppose that $F=\mathrm{U}[0,1]$. Then,

$$\mathcal{D}_2(F,F)\neq\mathcal{C}_2(F,F).$$

Proof. Let X be a random variable such that

$$\mathbb{P}\left(X=\frac{6}{5}\right)=\mathbb{P}\left(X=\frac{4}{5}\right)=\frac{1}{2}.$$

It is straightforward to check that $X\in\mathcal{C}_2(F,F)$. To see that $X\notin\mathcal{D}_2(F,F)$, assume that there exist $U_1,U_2\sim\mathrm{U}(0,1)$ such that $X=U_1+U_2$. Let $A_i=\{U_1\in[(i-1)/5,i/5)\}$, $B_i=\{U_2\in[(i-1)/5,i/5)\}$, $i=1,\ldots,5$, and $C=\{X=4/5\}$. In the following, sets are considered as identical if their indicator functions are almost surely equal.

Note that $X=4/5$ implies that $U_i\leqslant 4/5$, $i=1,2$, and $X=6/5$ implies that $U_i\geqslant 1/5$, $i=1,2$, that is, $C\subset A_5^c\cap B_5^c$ and $C^c\subset A_1^c\cap B_1^c$. Then it follows that

$$A_5\cup B_5\subset C^c \quad\text{and}\quad A_1\cup B_1\subset C,$$

and from $U_1+U_2=X$ we further have that

$$A_5=B_2,\quad A_2=B_5,\quad A_1=B_4 \quad\text{and}\quad A_4=B_1.$$

Now, $A_3=(A_1\cup A_2\cup A_4\cup A_5)^c=(B_4\cup B_5\cup B_1\cup B_2)^c=B_3$. It follows that $\mathbb{P}(A_3)=\mathbb{P}(C\cap A_3)+\mathbb{P}(C^c\cap A_3)=\mathbb{P}(C\cap A_3\cap B_3)+\mathbb{P}(C^c\cap A_3\cap B_3)=0$. This contradiction shows that $X\notin\mathcal{D}_2(F,F)$. $\square$

The above examples reveal some substantial challenges to determine the set $\mathcal{D}_n$ even in some very simple homogeneous settings. In the next section, we will investigate the asymptotic properties of $\mathcal{D}_n$ as $n\to\infty$ in homogeneous settings.

3.3 Asymptotic behavior of aggregation sets

In this section we investigate the asymptotic behavior of sets $\mathcal{D}_n(F_1,\ldots,F_n)$ when $F_1=\cdots=F_n$. To analyze the asymptotic behavior, one needs to normalize $\mathcal{D}_n$ by a constant $1/n$. We denote

$$\mathcal{B}_n(F)=\left\{\tfrac{1}{n}(X_1+\cdots+X_n),\ X_i\sim F,\ i=1,\ldots,n\right\}=\left\{\tfrac{1}{n}S: S\in\mathcal{D}_n(F,\ldots,F)\right\}. \quad (3.1)$$

The following lemma helps to justify our motivation for an asymptotic analysis of the set $\mathcal{B}_n$. We use the standard definition of upper limit for a sequence of sets $\{A_n,\ n \geqslant 1\}$, that is, $\limsup_{n\to\infty} A_n = \cap_{n\geqslant 1} \cup_{k\geqslant n} A_k$.

Lemma 3.4. *$\mathcal{B}_n(F) \subset \mathcal{B}_{nk}(F) \subset \limsup\limits_{m\to\infty} \mathcal{B}_m(F) \subset \mathcal{C}(F)$ for any $n, k \in \mathbb{N}$ and any integrable distribution F.*

Proof. For any $X \in \mathcal{B}_n(F)$, there exist $X_1, \dots, X_n \sim F$ such that

$$\frac{1}{n}(X_1 + \cdots + X_n) = X.$$

Then for each $k \in \mathbb{N}$, define $X_{i,j} = X_i$, $i = 1, \dots, n$, $j = 1, \dots, k$. Then

$$\frac{1}{nk}\sum_{j=1}^{k}\sum_{i=1}^{n} X_{i,j} = \frac{1}{nk}\sum_{i=1}^{n} kX_i = X.$$

This implies that $X \in \mathcal{B}_{nk}(F)$. Moreover, since k is arbitrary, we have that $X \in \limsup_{m\to\infty} \mathcal{B}_m(F)$; thus $\mathcal{B}_n(F) \subset \mathcal{B}_{nk}(F) \subset \limsup_{m\to\infty} \mathcal{B}_m(F)$. By Proposition 3.1, we have that for each $n \in \mathbb{N}$, $\mathcal{D}_n(F, \dots, F) \subset \mathcal{C}_n(F, \dots, F) = \{X : X \leqslant_{\mathrm{cx}} nY,\ Y \sim F\}$, thus $\mathcal{B}_n(F) \subset \mathcal{C}(F)$. Therefore,

$$\cup_{m\geqslant 1} \mathcal{B}_m(F) \subset \limsup_{m\to\infty} \mathcal{B}_m(F) \subset \cup_{m\geqslant 1}\mathcal{B}_m(F) \subset \mathcal{C}(F). \tag{3.2}$$

$\square$

We are now ready to present the main result on the asymptotic behavior of $\mathcal{B}_n$.

Theorem 3.5. *For any integrable distribution F, let $\mathcal{B}_n(F)$ and $\mathcal{C}(F)$ be given by (3.1) and (1.3), respectively. Then*

$$\overline{\limsup_{n\to\infty} \mathcal{B}_n(F)} = \mathcal{C}(F), \tag{3.3}$$

where $\overline{A}$ denotes the closure of A with respect to the topology induced by L^1-convergence.

Proof. By Lemma 3.4, $\cup_{n\geqslant 1}\mathcal{B}_n(F) \subset \mathcal{C}(F)$. By Proposition 3.3, we know that $\mathcal{C}(F)$ is closed with respect to the topology induced by L^1-convergence. Thus, we have $\overline{\limsup_{n\to\infty} \mathcal{B}_n(F)} \subset \mathcal{C}(F)$. For the converse, we show by the following two steps.

Step 1. Denote by L^* the set of random variables taking values in a finite set. First, we show that

$$\mathcal{C}(F) \cap L^* \subset \overline{\cup_{n\geqslant 1}\mathcal{B}_n(F)}. \tag{3.4}$$

Note that $\mathcal{C}(F) \cap L^*$ is not empty. For any $X \in \mathcal{C}(F) \cap L^*$, without loss of generality, denote the support of distribution of X by $\mathrm{supp}(X) = \{x_1, \dots, x_k\}$. By Lemma 2.1, there exists a random variable $Y \sim F$ such that

$$\mathbb{E}[Y|X = x_i] = x_i, \quad \text{for } i = 1, \dots, k.$$

One can define a sequence $\{Y_n,\ n \in \mathbb{N}\}$ such that given $X = x_i$, they are independent, and $[Y_n|X = x_i] \stackrel{d}{=} [Y|X = x_i]$ for $i = 1, \dots, k$. To see this, by the Kolmogorov consistency theorem, there exist sequences of independent random variables $\{Y_{ni},\ n \in \mathbb{N}\}$ defined on probability space $(A_i, \mathcal{F}_i, \mathbb{P}|_{A_i})$, where $A_i = \{\omega \in \Omega :\ X(\omega) = x_i\}$, $\mathcal{F}_i = \{A \cap A_i, A \in \mathcal{F}\}$ and $\mathbb{P}|_{A_i}$ is the probability measure on A_i given by $\mathbb{P}|_{A_i}(A) = \mathbb{P}(A)/\mathbb{P}(A_i)$ for all $A \in \mathcal{F}_i$, $i = 1, \dots, n$, such that with $Y_{ni} \stackrel{d}{=} [Y|X = x_i]$, $n \in \mathbb{N}$. Then we define $\{Y_n,\ n \in \mathbb{N}\}$ on Ω by

$$Y_n(\omega) = \begin{cases} Y_{n1}(\omega), & \omega \in A_1, \\ \dots, & \dots \\ Y_{nk}(\omega), & \omega \in A_k, \end{cases} \qquad n \in \mathbb{N}. \tag{3.5}$$

Then by the law of large numbers, we have

$$\frac{1}{n}[Y_1 + \cdots + Y_n|X = x_i] \xrightarrow{a.s.} x_i \quad \text{as } n \to \infty,$$

which means

$$Y^n := \frac{1}{n}(Y_1 + \cdots + Y_n) \xrightarrow{a.s.} X \quad \text{as } n \to \infty. \tag{3.6}$$

It should be noted that $Y_n,\ n \in \mathbb{N}$ are not independent unconditionally, and it can be easily verify that $Y_n \stackrel{d}{=} X$ for all $n \in \mathbb{N}$. Then by Remark 3.1, we have that $\{Y^n,\ n \in \mathbb{N}\}$ is uniformly integrable, which combined with (3.6) implies that $Y^n \xrightarrow{L^1} X$ as $n \to \infty$. Since $Y^n \in \mathcal{B}_n(F)$, $n \in \mathbb{N}$ and $\overline{\cup_{n \geqslant 1} \mathcal{B}_n(F)}$ is closed with respect to the topology induced by L^1-convergence, then it follows that $X \in \overline{\cup_{n \geqslant 1} \mathcal{B}_n(F)}$. Hence, we obtain (3.4).

Step 2. Second, we show that

$$\mathcal{C}(F) \subset \overline{\cup_{n \geqslant 1} \mathcal{B}_n(F)}.$$

Based on Step 1, it suffices to prove that

$$\mathcal{C}(F) \subset \overline{\mathcal{C}(F) \cap L^*}. \tag{3.7}$$

As we know that $\mathcal{C}(F)$ is a closed set, and L^* is dense in $\mathcal{C}(F)$ in the sense of L^1-convergence, (3.7) is naturally expected to hold; in the following we show this by construction. For any $X \in \mathcal{C}(F)$, by Lemma 2.1, without loss of generality, assume that there exists $Y \sim F$ such that $\mathbb{E}[Y|X] = X,\ a.s.$ Define

$$X_n = \sum_{i=-n2^n+1}^{n2^n} \mu_i I_{\left\{\frac{i-1}{2^n} \leqslant X < \frac{i}{2^n}\right\}} + \mu_{n2^n+1} I_{\{X \geqslant n\}} + \mu_{-n2^n} I_{\{X < -n\}}, \tag{3.8}$$

with $\mu_i = \mathbb{E}\left[X \middle| \frac{i-1}{2^n} \leqslant X < \frac{i}{2^n}\right]$, $i = -n2^n + 1, \dots, n2^n$, $\mu_{n2^n+1} = \mathbb{E}[X|X \geqslant n]$ and $\mu_{-n2^n} = \mathbb{E}[X|X < -n]$. It is easy to see that $X_n \xrightarrow{a.s.} X$ as $n \to \infty$. By Remark 3.1, we have $\{X_n,\ n \in \mathbb{N}\}$ is uniformly integrable. It follows immediately that

$$X_n \xrightarrow{L^1} X \quad \text{as} \quad n \to \infty. \tag{3.9}$$

Moreover, note that for all $n \in \mathbb{N}$

$$\mathbb{E}[Y|X_n] = \mathbb{E}\left[\mathbb{E}[Y|X]|X_n\right] = \mathbb{E}[X|X_n] = X_n,\ a.s.$$

which means that $X_n \leqslant_{\mathrm{cx}} Y$, i.e., $X_n \in \mathcal{C}(F) \cap L^*$ for all $n \in \mathbb{N}$. This, combined with (3.9), implies (3.7).

Finally, it follows from (3.2) that

$$\bigcup_{n \geqslant 1} \mathcal{B}_n(F) = \limsup_{n\to\infty} \mathcal{B}_n(F).$$

Combining with Steps 1-2, we complete the proof of the theorem. □

Remark 3.2. (i) Since $\mathcal{C}(F)$ is closed with respect to weak or a.s. convergence, we can see that $\overline{\limsup_{n\to\infty} \mathcal{B}_n(F)}$ is also the closure of $\limsup_{n\to\infty} \mathcal{B}_n(F)$ with respect to weak or a.s. convergence. Indeed, in the case of bounded random variables, $\overline{\limsup_{n\to\infty} \mathcal{B}_n(F)}$ is also the closure with respect to the topology induced by L^∞-convergence, as summarized in Proposition 3.6 below.

(ii) If the distribution F is not integrable, then the result in Theorem 3.5 cannot be obtained using the same proof. Note that if $\mathbb{E}[X^+] = \infty$ and $\mathbb{E}[X^-] < \infty$ for $X \sim F$, then $\mathcal{C}(F) \cap L^*$ is an empty set (by checking with the convex function $\phi(x) = -x$), not to mention that our proof requires $\mathcal{C}(F) \cap L^*$ to be dense in $\mathcal{C}(F)$ in the sense of L^1-convergence.

Proposition 3.6. *Suppose that F has bounded support, then*

$$\overline{\limsup_{n\to\infty} \mathcal{B}_n(F)}^* = \mathcal{C}(F), \tag{3.10}$$

where $\overline{A}^*$ denotes the closure of A with respect to the topology induced by L^∞-convergence.

Proof. It suffices to show the corollary by modifying some details in the proof of Theorem 3.5. In Step 1, using Corollary A.2 in Embrechts et al. (2014b), there exist $Y_{ni} \stackrel{d}{=} [Y|X = x_i]$, $i = 1, \dots, n$, $n \in \mathbb{N}$, such that

$$\left| \frac{1}{n} \sum_{k=1}^{n} Y_{ki} - x_i \right| < \frac{2}{n} \|X\|_\infty, \quad n \in \mathbb{N},$$

where $\|X\|_\infty = \text{ess-sup}|X| < \infty$. Then it follows that the Y^n, $n \in \mathbb{N}$ defined by (3.5) satisfy that $Y^n \xrightarrow{L^\infty} X$ as $n \to \infty$. In Step 2, since X bounded, it is easy to see that the X_n, $n \in \mathbb{N}$ defined by (3.8) satisfy that $X_n \xrightarrow{L^\infty} X$ as $n \to \infty$. Combining the above arguments yields that

$$\mathcal{C}(F) \subset \overline{\limsup_{n\to\infty} \mathcal{B}_n(F)}^*.$$

This completes the proof. ∎

In the following we reveal an important connection between Proposition 3.6 and a recently established result in risk management and dependence uncertainty: the asymptotic equivalence between the worst scenarios of Value-at-Risk (VaR) and Expected Shortfall (ES). We use the standard definitions of VaR and ES:

$$\mathrm{VaR}_p(X) = F^{-1}(p), \quad X \sim F, \quad p \in (0,1),$$

and

$$\mathrm{ES}_p(X) = \frac{1}{1-p}\int_p^1 \mathrm{VaR}_q(X)\mathrm{d}q, \quad p \in (0,1),$$

respectively. The asymptotic equivalence was established in Puccetti and Rüschendorf (2014), Puccetti et al. (2013) and Wang (2014) under different extra conditions based on the theory of *complete mixability*; see also Embrechts et al. (2014a, Section 3) for a history of this problem. Using Proposition 3.6, we obtain a substantially shorter and less technical proof of this result for bounded random variables. The complete version of this result for unbounded random variables is given recently in Wang and Wang (2014).

Corollary 3.7. *Let $X \sim F$ be a bounded random variable, then for $p \in (0,1)$,*

$$\lim_{n\to\infty} \frac{1}{n} \sup\{\mathrm{VaR}_p(X_1 + \cdots + X_n),\ X_i \sim F,\ i = 1\ldots,n\} = \mathrm{ES}_p(X).$$

Proof. Note that $\frac{1}{n}\sup\{\mathrm{VaR}_p(X_1 + \cdots + X_n),\ X_i \sim F,\ i = 1\ldots,n\} = \sup_{Y\in\mathcal{B}_n(F)} \mathrm{VaR}_p(Y)$. It can be easily verified that as $n \to \infty$, the limit of $\sup_{Y\in\mathcal{B}_n(F)} \mathrm{VaR}_p(Y)$ exists; see Wang et al. (2014, Proposition 2.1). Since ES_p preserves the convex order and $\mathcal{B}_n(F) \subset \mathcal{C}(F)$, we have $\sup_{Y\in\mathcal{B}_n(F)} \mathrm{ES}_p(Y) \leqslant \sup_{Y\in\mathcal{C}(F)} \mathrm{ES}_p(Y) = \mathrm{ES}_p(X)$. Thus,

$$\lim_{n\to\infty} \sup_{Y\in\mathcal{B}_n(F)} \mathrm{VaR}_p(Y) \leqslant \lim_{n\to\infty} \sup_{Y\in\mathcal{B}_n(F)} \mathrm{ES}_p(Y) \leqslant \mathrm{ES}_p(X).$$

To prove the converse inequality, take any $Y \in \mathcal{C}(F)$. By Proposition 3.6, there exists a sequence of random variables $X_k \in \mathcal{B}_{n_k}(F)$, $k \in \mathbb{N}$, such that $X_k \overset{L^\infty}{\to} Y$ as $k \to \infty$. This implies $\mathrm{VaR}_p(X_k) \to \mathrm{VaR}_p(Y)$ as $k \to \infty$. It follows that

$$\lim_{n\to\infty} \sup_{Y\in\mathcal{B}_n(F)} \mathrm{VaR}_p(Y) = \limsup_{n\to\infty} \sup_{Y\in\mathcal{B}_n(F)} \mathrm{VaR}_p(Y) \geqslant \sup_{Y\in\mathcal{C}(F)} \mathrm{VaR}_p(Y).$$

It remains to prove that $\sup_{Y\in\mathcal{C}(F)} \mathrm{VaR}_p(Y) \geqslant \mathrm{ES}_p(X)$. Take $\overline{Y} = F^{-1}(U)I_{\{0\leqslant U\leqslant p\}} + \mathrm{ES}_p(X)I_{\{U>p\}}$ with $U \sim \mathrm{U}[0,1]$. Then $\mathbb{E}[F^{-1}(U)|\overline{Y}] = \overline{Y}$ a.s., which implies that $\overline{Y} \leqslant_{\mathrm{cx}} X$, i.e., $\overline{Y} \in \mathcal{C}(F)$. Since $\mathrm{VaR}_p(\overline{Y}) = \mathrm{ES}_p(X)$, it follows that $\sup_{Y\in\mathcal{C}(F)} \mathrm{VaR}_p(Y) \geqslant \mathrm{VaR}_p(\overline{Y}) = \mathrm{ES}_p(X)$. This completes the proof. ∎

4 Lower-convex sets and comonotonic sums

For any integrable distributions $F_1, \dots, F_n$, recall that

$$\mathcal{C}_n(F_1, \dots, F_n) = \left\{ \sum_{i=1}^n X_i : X_i \leqslant_{\mathrm{cx}} Y_i,\ Y_i \sim F_i,\ i = 1, \dots, n \right\}. \tag{4.1}$$

Consider a lower-convex set represented by comonotonic random variables:

$$\mathcal{C}_n^*(F_1, \dots, F_n) = \left\{ \sum_{i=1}^n X_i^c : X_i^c \leqslant_{\mathrm{cx}} Y_i,\ Y_i \sim F_i,\ X_i^c \text{ comonotonic},\ i = 1, \dots, n \right\}. \tag{4.2}$$

It is obvious that $\mathcal{C}_n^*(F_1, \dots, F_n) \subset \mathcal{C}_n(F_1, \dots, F_n)$. The main result in this section states that $\mathcal{C}_n(F_1, \dots, F_n) \subset \mathcal{C}_n^*(F_1, \dots, F_n)$ also holds, i.e. the above two sets are actually identical. This is equivalent to say that

$$\begin{aligned} \mathcal{C}_n(F_1, \dots, F_n) &= \left\{ \sum_{i=1}^n G_i^{-1}(U_i) : G_i \leqslant_{\mathrm{cx}} F_i,\ U_i \sim \mathrm{U}[0,1],\ i = 1, \dots, n \right\} \\ &= \left\{ \sum_{i=1}^n G_i^{-1}(U) : G_i \leqslant_{\mathrm{cx}} F_i,\ U \sim \mathrm{U}[0,1],\ i = 1, \dots, n \right\}; \end{aligned}$$

hence elements in $\mathcal{C}_n$ has a much simpler form, driven by one single random source. Note the difference between the definition of $\mathcal{C}_n^*$ and the other set $\mathcal{C}_n'$. We first need the following lemma.

Lemma 4.1. *Let X and Y be two random variables on $(\Omega, \mathcal{F}, \mathbb{P})$ with $\Omega = \{1, \dots, n\}$, $\mathcal{F} = 2^\Omega$ and $\mathbb{P}(\{i\}) = p_i$, $i = 1, \dots, n$, given by*

$$X(i) = x_i,\ Y(i) = y_i,\ i = 1, \dots, n.$$

Then there exist comonotonic random variables X^c and Y^c on $(\Omega, \mathcal{F}, \mathbb{P})$ such that

$$X^c \leqslant_{\mathrm{cx}} X, \quad Y^c \leqslant_{\mathrm{cx}} Y \quad \text{and} \quad X^c + Y^c \overset{d}{=} X + Y.$$

Proof. We prove the result by induction. The result holds trivially for $n = 1$. Assume that it also holds for $n \leqslant k$. We aim to show that it holds for $n = k + 1$. Define the probability space $(\Omega_k, \mathcal{F}_k, \mathbb{P}|_{\Omega_k})$ by $\Omega_k = \Omega \setminus \{k+1\}$, $\mathcal{F}_k = 2^{\Omega_k}$ and $\mathbb{P}|_{\Omega_k}(A) = \mathbb{P}(A)/\mathbb{P}(\Omega_k)$ for any $A \in \mathcal{F}_k$, and define random variables X^k and Y^k on $(\Omega_k, \mathcal{F}_k, \mathbb{P}|_{\Omega_k})$ given by

$$X^k = [X|\Omega_k] \quad \text{and} \quad Y^k = [Y|\Omega_k].$$

By induction, there exist comonotonic random variables X_k^c and Y_k^c on Ω_k such that $X_k^c \leqslant_{\mathrm{cx}} X^k$, $Y_k^c \leqslant_{\mathrm{cx}} Y^k$ and

$$X^k + Y^k \overset{d}{=} X_k^c + Y_k^c.$$

Without loss of generality, assume that

$$X_k^c(i) = x_i^* \quad \text{and} \quad Y_k^c(i) = y_i^*, \quad i = 1, \dots, k,$$

and $x_1^* \leqslant \ldots \leqslant x_k^*$ and $y_1^* \leqslant \ldots \leqslant y_k^*$. Define the extension to Ω of random variables X_k^c and Y_k^c, still denoted by X_k^c and Y_k^c for simplicity, given by

$$X_k^c(i) = x_i^*, \quad i = 1, \ldots, k, \quad X_k^c(k+1) = x_{k+1},$$

and

$$Y_k^c(i) = y_i^*, \quad i = 1, \ldots, k, \; Y_k^c(k+1) = y_{k+1},$$

respectively. Then it is obvious that $X_k^c + Y_k^c \overset{d}{=} X + Y$, $X_k^c \leqslant_{\mathrm{cx}} X$ and $Y_k^c \leqslant_{\mathrm{cx}} Y$. However, generally X_k^c and Y_k^c are not comonotonic. To complete the proof, consider the following cases:

(1) When $x_{k+1} < x_1^*$ and $y_{k+1} > y_k^*$: let

$$\delta = \min\left\{p_n(x_1^* - x_{k+1}),\; p_n(y_{k+1} - y_k^*)\right\}.$$

We only deal with the case when $x_1^* - x_{k+1} \leqslant y_{k+1} - y_k^*$. The other case is symmetric. Define random variables X_{k+1} and Y_{k+1} on Ω as (note that $n = k+1$)

$$X_{k+1}(i) = x_i^* - \delta, \; i = 1, \ldots, k, \quad X_{k+1}(k+1) = x_{k+1} + \delta\frac{1-p_n}{p_n} = X_{k+1}(1),$$

and

$$Y_{k+1}(i) = y_i^* + \delta, \; i = 1, \ldots, k, \quad Y_{k+1}(k+1) = y_{k+1} - \delta\frac{1-p_n}{p_n} \geqslant Y_{k+1}(k).$$

It is easy to verify that $X_k^c + Y_k^c = X_{k+1} + Y_{k+1}$. X_k^c and Y_k^c differ from X_{k+1} and Y_{k+1} by k mean preserving spreads, respectively, which, by Lemma 2.2, implies that $X_{k+1} \leqslant_{\mathrm{cx}} X$ and $Y_{k+1} \leqslant_{\mathrm{cx}} Y$. Now X_{k+1} and Y_{k+1} both take the smallest value at $\omega = 1$, and by the following steps we can reduce to the case of k.

Define random variables X^{k+1} and Y^{k+1} on probability space $(\Omega_1, \mathcal{F}_1, \mathbb{P}|_{\Omega_1})$ with $\Omega_1 = \Omega \setminus \{1\}$, $\mathcal{F}_1 = 2^{\Omega_1}$ and $\mathbb{P}|_{\Omega_1}(A) = \mathbb{P}(A)/\mathbb{P}(\Omega_1)$ for any $A \in \mathcal{F}_1$, given by

$$X^{k+1} = [X_{k+1}|\Omega_1] \;\text{ and }\; Y^{k+1} = [Y_{k+1}|\Omega_1].$$

By induction, there exist comonotonic random variables X_{k+1}^c and Y_{k+1}^c on Ω_1 such that $X_{k+1}^c \leqslant_{\mathrm{cx}} X^{k+1}$, $Y_{k+1}^c \leqslant_{\mathrm{cx}} Y^{k+1}$ and

$$X_{k+1}^c + Y_{k+1}^c \overset{d}{=} X^{k+1} + Y^{k+1}.$$

Repeating the extension procedure, we get the their extension versions X_{k+1}^c and Y_{k+1}^c on Ω by defining its value on $\{1\}$ as

$$X_{k+1}^c(1) = X_{k+1}(1) \;\text{ and }\; Y_{k+1}^c(1) = Y_{k+1}(1).$$

By noting that

$$X_{k+1}^c(1) = \min\{X_{k+1}(\omega) : \omega \in \Omega_1\} \leqslant \min\{X_{k+1}^c(\omega) : \omega \in \Omega_1\},$$

we have that

$$X_{k+1}^c(1) = \min\{X_{k+1}^c(\omega) : \omega \in \Omega\},$$

and similarly,

$$Y_{k+1}^c(1) = \min\{Y_{k+1}^c(\omega) : \omega \in \Omega\}.$$

It follows that X_{k+1}^c and Y_{k+1}^c are comonotonic on Ω. Also, it is easy to see that

$$X_{k+1}^c + Y_{k+1}^c \stackrel{d}{=} X_{k+1} + Y_{k+1} \stackrel{d}{=} X + Y,$$

and $X_{k+1}^c \leqslant_{\mathrm{cx}} X$ and $Y_{k+1}^c \leqslant_{\mathrm{cx}} Y$.

(2) When $y_{k+1} > x_1^*$ and $y_{k+1} < y_k^*$: it is similar to the first case.

(3) In all the remaining cases X_k^c and Y_k^c both take the smallest value at $\omega = 1$, and using the argument in Step 1 we can reduce to the case of k.

The proof of the lemma is complete. □

With Lemma 4.1, we can show the main result of this section.

Theorem 4.2. *For any integrable distributions $F_1, \ldots, F_n$, let $\mathcal{C}_n(F_1, \ldots, F_n)$ and $\mathcal{C}_n^*(F_1, \ldots, F_n)$ be defined by (4.1) and (4.2), respectively. Then*

$$\mathcal{C}_n(F_1, \ldots, F_n) = \mathcal{C}_n^*(F_1, \ldots, F_n).$$

Proof. It suffices to show $\mathcal{C}_n(F_1, \ldots, F_n) \subset \mathcal{C}_n^*(F_1, \ldots, F_n)$ since the converse is obvious. For any $S \in \mathcal{C}_n(F_1, \ldots, F_n)$, there exist $X_i \leqslant_{\mathrm{cx}} Y_i$, $Y_i \sim F_i$, $i = 1, \ldots, n$ such that $S \stackrel{d}{=} \sum_{i=1}^n X_i$.

Step 1. We first show the result for $n = 2$ when $X_1, X_2 \in L^*$, where L^* is the set of random variables taking values in a finite set. By Lemma 4.1, there exist comonotonic random variables $X_i^c \leqslant_{\mathrm{cx}} X_i \leqslant_{\mathrm{cx}} Y_i$, $i = 1,\ 2$, such that $S \stackrel{d}{=} \sum_{i=1}^n X_i^c$.

Step 2. Consider the case that $n = 2$ and X_1 and X_2 are general random variables. There exist $X_1^k \in L^*$ and $X_2^k \in L^*$ which are increasing in convex order, $k = 1, 2, \ldots$, such that

$$X_1^k \xrightarrow{a.s.} X_1 \quad \text{and} \quad X_2^k \xrightarrow{a.s.} X_2 \ \text{ as } k \to \infty.$$

By Step 1, for each $k \in \mathbb{N}$, there exist comonotonic random variables $X_1^{k,c} \in \mathcal{C}(F_1)$ and $X_2^{k,c} \in \mathcal{C}(F_2)$ such that

$$X_1^{k,c} + X_2^{k,c} \stackrel{d}{=} X_1^k + X_2^k, \quad k \in \mathbb{N}.$$

Let μ_k and ν_k be the probability measures on $\mathbb{R}$ induced by $X_1^{k,c}$ and $X_2^{k,c}$, respectively. By Helly theorem, there exist subsequences μ_{n_k} and ν_{n_k} such that

$$\mu_{n_k} \xrightarrow{v} \mu \quad \text{and} \quad \nu_{n_k} \xrightarrow{v} \nu \ \text{ as } k \to \infty,$$

where $\xrightarrow{v}$ represents vague convergence. We claim that μ and ν are both probability measures. To see this, for a real number $M > 0$,

$$\mu(\mathbb{R}) \geqslant \lim_{k\to\infty} \mu_{n_k}([-M, M]) \geqslant 1 - \lim_{k\to\infty} \frac{1}{M}\mathbb{E}[|X_1^{n_k,c}|] \geqslant 1 - \frac{\mathbb{E}[|X|]}{M}, \tag{4.3}$$

where the last inequality follows from that $X_1^{n_k,c} \leqslant_{\mathrm{cx}} X$ and $\phi(\cdot) = |\cdot|$ is a convex function. Letting $M \to \infty$ yields that $\mu(\mathbb{R}) = 1$. Similarly, we can show that $\nu(\mathbb{R}) = 1$. Therefore

$$\mu_{n_k} \xrightarrow{w} \mu \quad \text{and} \quad \nu_{n_k} \xrightarrow{w} \nu \ \text{ as } k \to \infty,$$

where $\xrightarrow{w}$ represents the weak convergence. Let X_1^c and X_2^c be comonotonic random variables such that $\mathbb{P}(X_1^c \in \cdot) = \mu(\cdot)$ and $\mathbb{P}(X_2^c \in \cdot) = \nu(\cdot)$. Then we have

$$X_1^c + X_2^c \overset{d}{=} X_1 + X_2.$$

On the other hand, by Theorem 3.3,

$$X_i^c \leqslant_{\mathrm{cx}} X_i, \quad i = 1, 2.$$

This completes the proof for $n = 2$ and general X_1 and X_2.

Step 3. For general $n \geqslant 3$, we prove it by induction. Denote

$$S = \sum_{i=1}^{n} X_i =: S_{n-1} + X_n.$$

Denote by F^{n-1} the distribution function of S_{n-1} and consider $\mathcal{C}_2(F^{n-1}, F_n)$. By induction, there exist $S_{n-1}^c \leqslant_{\mathrm{cx}} S_{n-1}$ and $X_n^c \leqslant_{\mathrm{cx}} X_n$ such that S_{n-1}^c and X_n^c are comonotonic and

$$S_n \overset{d}{=} S_{n-1}^c + X_n^c.$$

Note that $S_{n-1}^c \in \mathcal{C}_{n-1}(F_1, \ldots, F_{n-1})$ by Proposition 3.2. By induction again, there exist $X_i^c \in \mathcal{C}(F_i)$, $i = 1, \ldots, n-1$, comonotonic, such that

$$S_{n-1}^c \overset{d}{=} \sum_{i=1}^{n-1} X_i^c.$$

Define $X_i = F_{X_i^c}^{-1}(U)$, $i = 1, \ldots, n$ for some U$[0, 1]$ random variable U. Then $X_i \in \mathcal{C}(F_i)$, $i = 1, \ldots, n$ are comonotonic and $S_n \overset{d}{=} \sum_{i=1}^n X_i$. This completes the proof of the theorem.

$\square$

Remark 4.1. One may also consider the difference between two random variables instead of the sum of them. Since $X \leqslant_{\mathrm{cx}} Y$ is equivalent to $-X \leqslant_{\mathrm{cx}} -Y$, one can see that

$$
\begin{aligned}
&\{X - Y : X \in \mathcal{C}(F),\ Y \in \mathcal{C}(G)\} \\
&= \{X + Y : X \in \mathcal{C}(F),\ Y \in \mathcal{C}(G^*),\} \\
&= \{X^c + Y^c : X^c \in \mathcal{C}(F),\ Y^c \in \mathcal{C}(G^*),\ X^c,\ Y^c \text{ comonotonic}\} \\
&= \{X^c - Y^c : X^c \in \mathcal{C}(F),\ Y^c \in \mathcal{C}(G),\ X^c,\ Y^c \text{ counter-comonotonic}\},
\end{aligned}
$$

where $G^*(\cdot) = 1 - G(\cdot-)$ and $G(x-)$ denotes the left limit of G at $x \in \mathbb{R}$.

Remark 4.2. If some of $F_1, \dots, F_n$ are not integrable, it remains open whether $\mathcal{C}_n(F_1, \dots, F_n) = \mathcal{C}_n^*(F_1, \dots, F_n)$ still holds. The main difficulty is that (4.3) generally fails to hold as $\mathbb{E}[|X_1^{n_k,c}|]$ might be unbounded, so the same logic in the proof could not be applied directly.

Below we discuss an interesting consequence of Theorem 4.2 in the theory of risk measures. A risk measure is a mapping from a set (typically, a convex cone) of random variables $\mathcal{X}$ to $\mathbb{R}$. A classic interpretation of $\rho(X)$ is the capital requirement for a risk $X \in \mathcal{X}$ held by a financial institution. Most commonly-used risk measures are *law-determined*, i.e. $\rho(X)$ only depends on the distribution of X. We refer to Föllmer and Schied (2011, Section 4) for more on risk measures.

One important property for risk measures is the *comonotonic additivity* (see Kusuoka, 2001): for comonotonic random variables $X, Y \in \mathcal{X}$, $\rho(X + Y) = \rho(X) + \rho(Y)$. This interprets into that the capital requirement principle ρ does not allow *diversification benefit* for comonotonic risks. Another important property for risk measures is *preserving convex order*: for $X, Y \in \mathcal{X}$, $X \leqslant_{\mathrm{cx}} Y$ implies that $\rho(X) \leqslant \rho(Y)$. This interprets into that the capital requirement principle ρ penalizes on the more volatile risk Y compared to the more stable risk X; see for instance Föllmer and Schied (2011, Section 4.5). VaR and ES defined in Section 3 are both law-determined and comonotonic additive, and ES also preserves convex order. The following corollary builds up a bridge between those two concepts.

Corollary 4.3. *Let ρ be a comonotonic additive risk measure. Define risk measure $\hat{\rho}(X)$ for $X \sim F$ as*

$$\hat{\rho}(X) = \sup_{Y \in \mathcal{C}(F)} \rho(Y), \quad X \in L^1.$$

Then $\hat{\rho}$ is comonotonic additive and preserves convex order.

Proof. That $\hat{\rho}$ preserves convex order follows from that the set $\mathcal{C}(F)$ is increasing as F is increasing in convex order. In the following we show that $\hat{\rho}$ is comonotonic additive. Let $X \sim F$

and $Y \sim G$ be comonotonic random variables and $H = F \oplus G$. By Theorem 4.2, we have that $\mathcal{C}(H) = \{X_1 + Y_1 : X_1 \in \mathcal{C}(F),\ Y_1 \in \mathcal{C}(G)\} = \{X_1^c + Y_1^c : (X_1^c,\ Y_1^c) \in \mathcal{C}_{(F,G)}\}$ where $\mathcal{C}_{(F,G)} := \{(X_1^c, Y_1^c) : X_1^c \in \mathcal{C}(F),\ Y_1^c \in \mathcal{C}(G),\ X_1^c,\ Y_1^c \text{ comonotonic}\}$. Hence,

$$\begin{aligned}
\hat{\rho}(X+Y) &= \sup_{Z \in \mathcal{C}(H)} \rho(Z) = \sup_{(X_1^c, Y_1^c) \in \mathcal{C}_{(F,G)}} \rho(X_1^c + Y_1^c) \\
&= \sup_{(X_1^c, Y_1^c) \in \mathcal{C}_{(F,G)}} \rho(X_1^c) + \rho(Y_1^c) \\
&= \sup_{X_1 \in \mathcal{C}(F), Y_1 \in \mathcal{C}(G)} \rho(X_1) + \rho(Y_1) \\
&= \sup_{X_1 \in \mathcal{C}(F)} \rho(X_1) + \sup_{Y_1 \in \mathcal{C}(G)} \rho(Y_1) \\
&= \hat{\rho}(X) + \hat{\rho}(Y).
\end{aligned}$$

This completes the proof. □

Remark 4.3. If a *monetary risk measure* ρ is comonotonic additive and preserves convex order, then ρ must be a *spectral risk measure*; see Yaari (1987) and Acerbi (2002).

5 Conclusion

In this paper, for integrable distributions $F_1, \ldots, F_n$, we studied the set $\mathcal{D}_n$ of the sums of n random variables with given respective distributions $F_1, \ldots, F_n$, and the set $\mathcal{C}_n$ of the sums of random variables that are smaller than $F_1, \ldots, F_n$ in convex order. We obtained some theoretical properties of $\mathcal{D}_n \subset \mathcal{C}_n$, and showed that $\mathcal{D}_n$ has a limit $\mathcal{C}_1$ after scaling by $1/n$, as $n \to \infty$. It was also shown that random variables in $\mathcal{C}_n$ can be represented by comonotonic sums of random variables smaller than the corresponding marginal distributions in convex order. The techniques provided in this paper are directly related to open questions regarding dependence uncertainty in quantitative risk management. We remark that a characterization of $\mathcal{D}_n$ is still generally not yet clear.

Acknowledgement

We thank Carole Bernard, Giovanni Puccetti, Bin Wang, two referees, an Associate Editor, and an Editor for helpful comments and discussions on an earlier version of this paper. This work was carried out during the period of T. Mao's postdoctoral fellowship supported by the Department of Statistics and Actuarial Science, University of Waterloo. T. Mao was supported by the Fundamental Research Funds for the Central Universities and the NNSF of China (Nos. 11301500, 11371340, 11271347). R. Wang acknowledges support from the Natural Sciences and Engineering Research Council of Canada (NSERC).

References

Acerbi, C. (2002). Spectral measures of risk: A coherent representation of subjective risk aversion. *Journal of Banking and Finance*, **26**(7), 1505–1518.

Bernard, C., Jiang, X. and Wang, R. (2014). Risk aggregation with dependence uncertainty. *Insurance: Mathematics and Economics*, **54**, 93–108.

Bernard, C., Rüschendorf, L. and Vanduffel, S. (2013). Value-at-Risk bounds with variance constraints. *Preprint*, University of Freiburg.

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.

Durrett, R. (2010). *Probability: theory and examples.* 4th Edition. Cambridge University Press.

Elton J. and Hill T. P. (1992). Fusions of a probability distribution. *The Annals of Probability*, **20**(1), 421–454.

Embrechts, P., Puccetti, G. and Rüschendorf, L. (2013). Model uncertainty and VaR aggregation. *Journal of Banking and Finance*, **37**(8), 2750–2764.

Embrechts, P., Puccetti, G., Rüschendorf, L., Wang, R. and Beleraj, A. (2014a). An academic response to Basel 3.5. *Risks*, **2**(1), 25–48.

Embrechts, P., Wang, B. and Wang, R. (2014b). Aggregation-robustness and model uncertainty of regulatory risk measures. *Finance and Stochastics*, to appear.

Föllmer, H. and Schied, A. (2011). *Stochastic Finance: An Introduction in Discrete Time.* Walter de Gruyter, Third Edition.

Joe, H. (1997). *Multivariate models and dependence concepts.* London: Chapman & Hall.

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

Müller, A. (2013). Duality theory and transfers for stochastic order relations. *Stochastic orders in reliability and risk.* Springer. **208**, 41–57.

Müller, A. and Stoyan, D. (2002). *Comparison methods for statistical models and risks.* Wiley, England.

Puccetti, G. and Rüschendorf, L. (2014). Asymptotic equivalence of conservative value-at-risk- and expected shortfall-based capital charges. *Journal of Risk*, **16**(3), 3–22.

Puccetti, G., Wang, B. and Wang, R. (2013). Complete mixability and asymptotic equivalence of worst-possible VaR and ES estimates. *Insurance: Mathematics and Economics*, **53**(3), 821–828.

Rothschild, M. and Stiglitza, J. E. (1970). Increasing risk: a definition. *Journal of Economic Theory*, **2**(3), 225–243.

Rüschendorf, L. (2013). *Mathematical risk analysis. Dependence, risk bounds, optimal allocations and portfolios.* Springer, Heidelberg.

Shaked, M. and Shanthikumar, J. G. (2007). *Stochastic orders.* Springer Series in Statistics.

Strassen, V. (1965). The existence of probability measures with given marginals. *Annals of Mathematical Statistics*, **36**(2), 423–439.

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, B. and Wang, R. (2014). Extreme negative dependence and risk aggregation. *Preprint available at* http://arxiv.org/abs/1407.6848, version 25 Jul 2014.

Wang, R. (2014). Asymptotic bounds for the distribution of the sum of dependent random variables. *Journal of Applied Probability*, **51**(3), 780–798.

Wang, R., Bignozzi, V. and Tsanakas, A. (2014). How superadditive can a risk measure be? *Available at SSRN* http://ssrn.com/abstract=2373149.

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.

Yaari, M. E. (1987). The dual theory of choice under risk. *Econometrica*, **55**(1), 95–115.