

## Chapter 19

# Dependence Comparison of Multivariate Extremes via Stochastic Tail Orders

Haijun Li

**Abstract:** A stochastic tail order is introduced to compare right tails of distributions and related closure properties are established. The stochastic tail order is then used to compare the dependence structure of multivariate extreme value distributions in terms of upper tail behaviors of their underlying samples.

### 19.1 Introduction

Let  $\mathbf{X}_n = (X_{1,n}, \dots, X_{d,n})$ ,  $n = 1, 2, \dots$  be independent and identically distributed (i.i.d.) random vectors with common distribution function (df)  $F$ . Define component-wise maxima  $M_{i,n} := \bigvee_{j=1}^n X_{i,j}$  and minima  $m_{i,n} := \bigwedge_{j=1}^n X_{i,j}$ ,  $1 \leq i \leq d$ . Here and hereafter  $\bigvee$  ( $\bigwedge$ ) denotes the maximum (minimum). This paper focuses on

---

H. Li (✉)

Department of Mathematics, Washington State University,  
Pullman, WA 99164, USA

e-mail: [lih@math.wsu.edu](mailto:lih@math.wsu.edu)

dependence comparison of the limiting distributions of properly normalized vectors of component-wise maxima  $\mathbf{M}_n := (M_{1,n}, \dots, M_{d,n})$  and of component-wise maxima  $\mathbf{m}_n := (m_{1,n}, \dots, m_{d,n})$ , as  $n \rightarrow \infty$ . The comparison method is based on asymptotic comparisons of upper tails of  $F$  of the underlying sample  $(\mathbf{X}_n, n \geq 1)$ .

For any two vectors  $\mathbf{a}, \mathbf{b} \in \mathbb{R}^d$ , the sum  $\mathbf{a} + \mathbf{b}$ , product  $\mathbf{a}\mathbf{b}$ , quotient  $\mathbf{a}/\mathbf{b}$ , and vector inequalities such as  $\mathbf{a} \leq \mathbf{b}$  are all operated component-wise. Let  $G$  and  $H$  be dfs defined on  $\mathbb{R}^d$  with nondegenerate margins. A df  $F$  is said to be in the domain of attraction of  $G$  for the maxima, denoted as  $F \in \text{DA}_\vee(G)$ , if there exist  $\mathbb{R}^d$ -valued sequences  $\mathbf{a}_n = (a_{1,n}, \dots, a_{d,n})$  with  $a_{i,n} > 0$ ,  $1 \leq i \leq d$ , and  $\mathbf{b}_n = (b_{1,n}, \dots, b_{d,n})$ ,  $n = 1, 2, \dots$ , such that for any  $\mathbf{x} = (x_1, \dots, x_d)$ , as  $n \rightarrow \infty$ ,

$$\begin{aligned} & \text{P} \left\{ \frac{M_{1,n} - b_{1,n}}{a_{1,n}} \leq x_1, \dots, \frac{M_{d,n} - b_{d,n}}{a_{d,n}} \leq x_d \right\} \\ &= F^n(\mathbf{a}_n \mathbf{x} + \mathbf{b}_n) \rightarrow G(\mathbf{x}), \end{aligned} \tag{19.1.1}$$

and in this case,  $G$  is called a max multivariate extreme value (MEV) distribution. Similar definitions for min MEV distributions and their domain of attraction can be made. For minima, Eq.(19.1.1) is replaced by

$$\begin{aligned} & \text{P} \left\{ \frac{m_{1,n} - b_{1,n}}{a_{1,n}} > x_1, \dots, \frac{m_{d,n} - b_{d,n}}{a_{d,n}} > x_d \right\} \\ &= \overline{F}^n(\mathbf{a}_n \mathbf{x} + \mathbf{b}_n) \rightarrow \overline{H}(\mathbf{x}), \end{aligned} \tag{19.1.2}$$

which is denoted by  $F \in \text{DA}_\wedge(H)$ . Here and hereafter bars on the top of dfs denote (joint) survival functions. A key property of an MEV distribution  $G$  is that all positive powers of  $G$  are also distributions, and max MEV distributions coincide with the max-stable distributions, which form a subclass of max-infinitely divisible distributions. Similarly min MEV distributions coincide with the min-stable distributions, which form a subclass of min-infinitely divisible distributions. One needs only to study the case of maxima as the theory for minima is similar.

Let  $\mathbf{X} = (X_1, \dots, X_d)$  be a generic random vector with distribution  $F$  and continuous, univariate margins  $F_1, \dots, F_d$ . If  $F \in \text{DA}_\vee(G)$ , then  $G$  is closely related to the upper tail distribution of  $X$ , which often possesses the heavy tail property of *regular variation*. Without loss of generality, we may assume that  $\mathbf{X}$  is nonnegative component-wise.

Consider the standard case in which the survival functions  $\overline{F}_i(x) := 1 - F_i(x)$ ,  $1 \leq i \leq d$  of the margins are right tail equivalent; that is,

$$\frac{\overline{F}_i(x)}{\overline{F}_1(x)} = \frac{1 - F_i(x)}{1 - F_1(x)} \rightarrow 1, \text{ as } x \rightarrow \infty, \quad 1 \leq i \leq d. \quad (19.1.3)$$

The distribution  $F$  or random vector  $\mathbf{X}$  is said to be *multivariate regularly varying* (MRV) at  $\infty$  with intensity measure  $\nu$  if there exists a scaling function  $b(t) \rightarrow \infty$  and a nonzero Radon measure  $\nu(\cdot)$  such that as  $t \rightarrow \infty$ ,

$$t \mathbb{P} \left\{ \frac{\mathbf{X}}{b(t)} \in B \right\} \rightarrow \nu(B), \forall \text{ relatively compact sets } B \subset \overline{\mathbb{R}}_+^d \setminus \{0\},$$

with  $\nu(\partial B) = 0$ , (19.1.4)

where  $\overline{\mathbb{R}}_+^d := [0, \infty]^d$ . The extremal dependence information of  $\mathbf{X}$  is encoded in the intensity measure  $\nu$  that satisfies that  $\nu(tB) = t^{-\alpha}\nu(B)$ , for all relatively compact subsets  $B$  that are bounded away from the origin, where  $\alpha > 0$  is known as the *tail index*. Since the set  $B_1 = \{\mathbf{x} \in \mathbb{R}^d : x_1 > 1\}$  is relatively compact within the cone  $\overline{\mathbb{R}}_+^d \setminus \{0\}$  and  $\nu(B_1) > 0$  under Eq. (19.1.3), it follows from Eq. (19.1.4) that the scaling function  $b(t)$  can be chosen to satisfy that  $\overline{F}_1(b(t)) = t^{-1}$ ,  $t > 0$ , after appropriately normalizing the intensity measure by  $\nu(B_1)$ . That is,  $b(t)$  can be chosen as  $b(t) = \overline{F}^{-1}(t^{-1}) = F_1^{-1}(1 - t^{-1})$  under the condition (19.1.3), and thus, Eq. (19.1.4) can be expressed equivalently as

$$\lim_{t \rightarrow \infty} \frac{\mathbb{P}\{\mathbf{X} \in tB\}}{\mathbb{P}\{X_1 > t\}} = \nu(B), \quad \forall \text{ relatively compact sets } B \subset \overline{\mathbb{R}}_+^d \setminus \{0\}, \quad (19.1.5)$$

satisfying that  $\nu(\partial B) = 0$ . It follows from Eqs. (19.1.5) and (19.1.3) that for  $1 \leq i \leq d$ ,

$$\lim_{t \rightarrow \infty} \frac{\mathbb{P}\{X_i > ts\}}{\mathbb{P}\{X_i > t\}} = \nu((s, \infty] \times \overline{\mathbb{R}}^{d-1}) = s^{-\alpha}\nu((1, \infty] \times \overline{\mathbb{R}}^{d-1}), \quad \forall s > 0.$$

That is, univariate margins have regularly varying right tails. In general, a Borel-measurable function  $g : \mathbb{R}_+ \rightarrow \mathbb{R}_+$  is regularly varying with exponent  $\rho \in \mathbb{R}$ , denoted as  $g \in \text{RV}_\rho$ , if and only if

$$g(t) = t^\rho \ell(t), \text{ with } \ell(\cdot) \geq 0 \text{ satisfying that } \lim_{t \rightarrow \infty} \frac{\ell(ts)}{\ell(t)} = 1, \text{ for } s > 0. \quad (19.1.6)$$

The function  $\ell(\cdot)$  is known as a slowly varying function and denoted as  $\ell \in \text{RV}_0$ . Since  $\bar{F}_1 \in \text{RV}_{-\alpha}$ ,  $1/\bar{F}_1 \in \text{RV}_\alpha$ , and thus, by Proposition 2.6 (v) of [387], the scaling function  $b \in \text{RV}_{\alpha-1}$ .

Since all the margins are tail equivalent as assumed in Eq. (19.1.3), one has

$$\bar{F}_i(t) = t^{-\alpha} \ell_i(t), \text{ where } \ell_i \in \text{RV}_0, \text{ and } \ell_i(t)/\ell_j(t) \rightarrow 1 \text{ as } t \rightarrow \infty, \text{ for any } i \neq j, \tag{19.1.7}$$

which, together with  $\bar{F}_1(b(t)) = t^{-1}$ , imply that

$$\lim_{t \rightarrow \infty} t \text{P}\{X_i > b(t)s\} = \lim_{t \rightarrow \infty} \frac{\text{P}\{X_i > b(t)s\} \bar{F}_i(b(t))}{\bar{F}_i(b(t)) \bar{F}_1(b(t))} = s^{-\alpha}, \quad s > 0, \quad 1 \leq i \leq d. \tag{19.1.8}$$

The detailed discussions on univariate and multivariate regular variations can be found in [61, 387]. The extension of MRV beyond the nonnegative orthant can be done by using the tail probability of  $\|X\|$ , where  $\|\cdot\|$  denotes a norm on  $\mathbb{R}^d$ , in place of the marginal tail probability in Eq. (19.1.5) (see [387], Sect. 6.5.5). The case that the limit in Eq. (19.1.3) is any nonzero constant can be easily converted into the standard tail equivalent case by properly rescaling margins. If the limit in Eq. (19.1.3) is zero or infinity, then some margins have heavier tails than others. One way to overcome this problem is to standardize the margins via marginal monotone transforms (see Theorem 6.5 in [387]) or to use the copula method [281].

**Theorem 19.1.1** (Marshall and Olkin [309]). *Assume that Eq. (19.1.3) holds. Then there exist normalization vectors  $\mathbf{a}_n > \mathbf{0}$  and  $\mathbf{b}_n$  such that, as  $n \rightarrow \infty$ ,*

$$\text{P}\left\{ \frac{\mathbf{M}_n - \mathbf{b}_n}{\mathbf{a}_n} \leq \mathbf{x} \right\} \rightarrow G(\mathbf{x}), \quad \forall \mathbf{x} \in \mathbb{R}_+^d,$$

where  $G$  is a  $d$ -dimensional distribution with Fréchet margins  $G_i(s) = \exp\{-s^{-\alpha}\}$ ,  $1 \leq i \leq d$ , if and only if  $F$  is MRV with intensity measure  $\nu([0, \mathbf{x}]^c) := -\log G(\mathbf{x})$ .

In other words,  $F \in \text{DA}_\vee(G)$  where  $G$  has Fréchet margins with tail index  $\alpha$  if and only if  $F$  is MRV with intensity measure  $\nu([0, \mathbf{x}]^c) = -\log G(\mathbf{x})$ .

**Remark 19.1.2.**

1. The normalization vectors  $\mathbf{a}_n > \mathbf{0}$  and  $\mathbf{b}_n$  in Theorem 19.1.1 can be made precise so that  $\mathbf{b}_n = \mathbf{0}$  and  $\mathbf{a}_n = (\bar{F}_1^{-1}(1/n), \dots, \bar{F}_d^{-1}(1/n))$  that depend only on the margins of  $F$ .

2. If Eq. (19.1.3) does not hold, Theorem 19.1.1 can still be established, but the nonstandard global regular variation with different scaling functions among various margins needs to be used in place of Eq. (19.1.5), which uses the same scaling function among different margins.
3. One-dimensional version of Theorem 19.1.1 is due to Gnedenko [184]. Note that the parametric feature enjoyed by univariate extremes is lost in the multivariate context.
4. Let  $\mathbb{S}_+^{d-1} = \{\mathbf{a} : \mathbf{a} = (a_1, \dots, a_d) \in \mathbb{R}_+^d, \|\mathbf{a}\| = 1\}$ , where  $\|\cdot\|$  is a norm defined in  $\mathbb{R}^d$ . Using the polar coordinates,  $G$  can be expressed as follows:

$$G(\mathbf{x}) = \exp \left\{ -c \int_{\mathbb{S}_+^d} \max_{1 \leq i \leq d} \{(a_i/x_i)^\alpha\} \mathbb{Q}(d\mathbf{a}) \right\},$$

where  $c > 0$  and  $\mathbb{Q}$  is a probability measure defined on  $\mathbb{S}_+^{d-1}$ . This is known as the Pickands representation [374], and  $c\mathbb{Q}(\cdot)$  is known as the spectral or angular measure.

5. Note that the spectral measure is a finite measure that can be approximated by a sequence of discrete measures. Using this idea, Marshall and Olkin [309] showed that the MEV distribution  $G$  is positively associated. This implies that as  $n$  is sufficiently large, we have asymptotically,

$$\mathbb{E} \left[ f(\mathbf{M}_n) g(\mathbf{M}_n) \right] \geq \mathbb{E} \left[ f(\mathbf{M}_n) \right] \mathbb{E} \left[ g(\mathbf{M}_n) \right]$$

for all nondecreasing functions  $f, g : \mathbb{R}^d \mapsto \mathbb{R}$ . Observe that the sample vector  $\mathbf{X}_n$  could have any dependence structure, but the strong positive dependence emerges among multivariate extremes.

6. Since  $G$  is max-infinitely divisible, all bivariate margins of  $G$  are  $\text{TP}_2$ , a positive dependence property that is even stronger than the positive association of bivariate margins (see Theorem 2.6 in [211]).

Since the normalization vectors  $\mathbf{a}_n > 0$  and  $\mathbf{b}_n$  in Theorem 19.1.1 depend only on the margins, dependence comparison of  $G$  can be easily established using the *orthant dependence order* on sample vectors.

Recall that a  $d$ -dimensional random vector  $\mathbf{X} = (X_1, \dots, X_d)$  with df  $F$  is said to be smaller than another  $d$ -dimensional random vector  $\mathbf{X}' = (X'_1, \dots, X'_d)$  with df  $F'$  in the upper (lower) orthant order, denoted as  $\mathbf{X} \leq_{\text{uo}} \mathbf{X}'$  or  $F \leq_{\text{uo}} F'$  ( $\mathbf{X} \leq_{\text{lo}} \mathbf{X}'$  or  $F \leq_{\text{lo}} F'$ ), if

$$P\{X_1 > x_1, \dots, X_d > x_d\} \leq P\{X'_1 > x_1, \dots, X'_d > x_d\}, \quad (19.1.9)$$

$$P\{X_1 \leq x_1, \dots, X_d \leq x_d\} \leq P\{X'_1 \leq x_1, \dots, X'_d \leq x_d\}, \quad (19.1.10)$$

for all  $(x_1, \dots, x_d) \in \mathbb{R}^d$ . If, in addition, their corresponding univariate margins are identical, then  $\mathbf{X}$  is said to be smaller than  $\mathbf{X}'$  in the upper (lower) orthant dependence order, denoted as  $\mathbf{X} \leq_{\text{uod}} \mathbf{X}'$  or  $F \leq_{\text{uod}} F'$  ( $\mathbf{X} \leq_{\text{lod}} \mathbf{X}'$  or  $F \leq_{\text{lod}} F'$ ). Clearly  $\mathbf{X} \leq_{\text{uod}} \mathbf{X}'$  implies that  $\mathbf{X} \leq_{\text{uo}} \mathbf{X}'$ , but the order  $\leq_{\text{uod}}$  focuses on comparing scale-invariant dependence among components. The detailed discussions on these orders can be found in [335, 426]. The following result is immediate due to the fact that the orthant order is closed under weak convergence.

**Proposition 19.1.3.** *Let  $(\mathbf{X}_n, n \geq 1)$  and  $(\mathbf{X}'_n, n \geq 1)$  be two i.i.d. samples with dfs  $F$  and  $F'$ , respectively. If  $F \in DA_{\vee}(G)$  and  $F' \in DA_{\vee}(G')$  with Fréchet margins, then  $\mathbf{X}_n \leq_{\text{lod}} \mathbf{X}'_n$  implies that  $G \leq_{\text{lod}} G'$ .*

Note, however, that the ordering  $\mathbf{X}_n \leq_{\text{lod}} \mathbf{X}'_n$  is strongly affected by the behavior at the center and often too strong to be valid. The fact that MRV is a tail property motivates us to focus on comparing only upper tails of  $\mathbf{X}_n$  and  $\mathbf{X}'_n$ , leading to weaker notions of *stochastic tail orders*. In Sect. 19.2, we introduce a notion of stochastic tail order for random variables and establish related closure properties and discuss its relation with other asymptotic orders that are already available in the literature. In Sect. 19.3, we extend the stochastic tail order to random vectors and show that the stochastic tail order of sample vectors sufficiently implies the orthant dependence order of the corresponding MEV distributions.

## 19.2 Stochastic Tail Orders

Let  $X$  and  $Y$  be two  $\mathbb{R}_+$ -valued random variables.  $X$  is said to be smaller than  $Y$  in the sense of stochastic tail order, denoted as  $X \leq_{\text{sto}} Y$ , if there exists a threshold constant  $t_0 > 0$  (usually large) such that

$$P\{X > t\} \leq P\{Y > t\}, \quad \forall t > t_0. \quad (19.2.1)$$

**Remark 19.2.1.**

1. The stochastic tail order  $\leq_{sto}$  is reflexive and transitive.  $\leq_{sto}$  is antisymmetric if tail identically distributed random variables are considered to be equivalent.
2. If  $X$  is smaller than  $Y$  in the usual stochastic order (denoted as  $X \leq_{st} Y$ ; see Sect. 1.A in [426]); that is,  $P\{X > t\} \leq P\{Y > t\}$  for all  $t$ , then  $X \leq_{sto} Y$ .
3.  $X \leq_{sto} Y$  if and only if there exists a small open neighborhood of  $\infty$  within which  $X$  is stochastically smaller than  $Y$ .
4.  $X \leq_{sto} Y$  implies

$$\limsup_{t \rightarrow \infty} \frac{P\{X > t\}}{P\{Y > t\}} \leq 1. \tag{19.2.2}$$

The stochastic tail orders using limiting inequalities such as Eq. (19.2.2) have been introduced and studied in [43, 242, 243, 389–391] and more recently in [300]. Most of these tail orders, however, are based on limiting approaches rather than stochastic comparison theory.

1. Mainik and Rüschendorf studied in [300] the following weak tail order: A random variable  $X$  is said to be smaller than another random variable  $Y$  in the asymptotic portfolio loss order, denoted as  $X \leq_{apl} Y$ , if the limiting inequality (19.2.2) holds. Observe that  $\sup_{s>t} \frac{P\{X>s\}}{P\{Y>s\}}$  is decreasing in  $t$ , and as such, in the case of  $X \leq_{apl} Y$  with

$$\limsup_{t \rightarrow \infty} \frac{P\{X > t\}}{P\{Y > t\}} = \lim_{t \rightarrow \infty} \left[ \sup_{s>t} \frac{P\{X > s\}}{P\{Y > s\}} \right] = 1,$$

one can find in any open neighborhood  $(c, \infty]$  of  $\infty$  that  $P(X > s) \geq P(Y > s)$  for some  $s > c$ . That is, neither  $X$  nor  $Y$  could dominate the other in any open neighborhood  $(c, \infty]$  of  $\infty$ , but asymptotically, the right tail of  $X$  decays at the rate that is bounded from above by the tail decay rate of  $Y$ .

2. Rojo introduced in [391] a stronger version of tail orders: Define  $X <_{sq} Y$  if

$$\limsup_{u \rightarrow 1} \frac{F^{-1}(u)}{G^{-1}(u)} < 1$$

where  $F^{-1}(\cdot)$  and  $G^{-1}(\cdot)$  denote the left-continuous inverses of dfs of  $X$  and  $Y$ , respectively. Obviously,  $X <_{\text{sq}} Y$  implies that  $X \leq_{\text{sto}} Y$ . Note, however, that  $<_{\text{sq}}$  is not a partial ordering.

The stochastic tail orders via limiting inequalities resemble the idea of comparing the asymptotic decay rates that is often employed in theory of large (and small) deviations [284]. In contrast, the notion Eq. (19.2.1) compares stochastically random variables in a small open neighborhood of  $\infty$  within which theory of stochastic orders retains its full power. For example, coupling remains valid in a small open neighborhood of  $\infty$ .

**Theorem 19.2.2.** *Let  $X$  and  $Y$  be two positive random variables with support  $[0, \infty)$ .  $X \leq_{\text{sto}} Y$  if and only if there exists a random variable  $Z$  defined on the probability space  $(\Omega, \mathcal{F}, \mathbb{P})$  with support  $[a, b]$ , and nondecreasing functions  $\psi_1$  and  $\psi_2$  with  $\lim_{z \rightarrow b} \psi_i(z) = \infty$ ,  $i = 1, 2$ , such that  $X \stackrel{d}{=} \psi_1(Z)$ ,  $Y \stackrel{d}{=} \psi_2(Z)$  and  $\mathbb{P}\{\psi_1(Z) \leq \psi_2(Z) \mid Z \geq z_0\} = 1$  for some  $z_0 > 0$ .*

**Proof:** Let  $X$  and  $Y$  have distributions  $F$  and  $G$  with support  $[0, \infty)$ , respectively, and let  $F^{-1}(\cdot)$  and  $G^{-1}(\cdot)$  denote the corresponding left-continuous inverses. Recall that for any df  $H$  on  $\mathbb{R}$ , the left-continuous inverse of  $H$  is defined as

$$H^{-1}(u) := \inf\{s : H(s) \geq u\}, \quad 0 \leq u \leq 1.$$

The left-continuous inverse has the following desirable properties:

1.  $H(H^{-1}(u)) \geq u$  for all  $0 \leq u \leq 1$ , and  $H^{-1}(H(x)) \leq x$  for all  $x \in \mathbb{R}$ .
2.  $H^{-1}(u) \leq x$  if and only if  $u \leq H(x)$ .
3. The set  $\{s : H(s) \geq u\}$  is closed for each  $0 \leq u \leq 1$ .

Necessity: Using Properties 1 and 2,  $X \leq_{\text{sto}} Y$  implies that  $F^{-1}(u) \leq G^{-1}(u)$ ,  $\forall u > u_0$  for some  $0 < u_0 < 1$ . Let  $U$  be a random variable with standard uniform distribution, and thus  $\mathbb{P}\{F^{-1}(U) \leq G^{-1}(U) \mid U \geq u_0\} = 1$ . Using Property 2,  $\mathbb{P}\{F^{-1}(U) \leq x\} = \mathbb{P}\{U \leq F(x)\} = F(x)$ . Similarly,  $\mathbb{P}\{G^{-1}(U) \leq x\} = G(x)$ .



Sufficiency: For all  $t \geq c$ , a constant with  $c > \psi_1(z_0)$ ,

$$\begin{aligned} \mathbb{P}\{X > t\} &= \mathbb{P}\{Z \geq z_0\} \mathbb{P}\{\psi_1(Z) > t \mid Z \geq z_0\} \\ &\leq \mathbb{P}\{Z \geq z_0\} \mathbb{P}\{\psi_2(Z) > t \mid Z \geq z_0\} \\ &\leq \mathbb{P}\{\psi_2(Z) > t\} \\ &= \mathbb{P}\{Y > t\}. \end{aligned}$$

■

The tail coupling presented in Theorem 19.2.2 enables us to establish desirable closure properties for the stochastic tail order. A Borel measurable function  $\psi : \mathbb{R}^d \rightarrow \mathbb{R}$  is called a Radon function if  $\psi$  is bounded on every compact subset of  $\mathbb{R}^d$ . Obviously, any nondecreasing function and any continuous function defined on  $\mathbb{R}^d$  are Radon functions.

**Definition 19.2.3.** A Borel measurable function  $\psi : \mathbb{R}_+^d \rightarrow \mathbb{R}$  is said to be eventually increasing if there exists a compact subset  $S \subset \mathbb{R}_+^d$  such that  $\psi$  is component-wise nondecreasing on  $S^c$  with  $\lim_{x_i \rightarrow \infty} \psi(x_1, \dots, x_i, \dots, x_d) = \infty$ .

**Proposition 19.2.4.** Let  $X$  and  $Y$  be two positive random variables with support  $[0, \infty)$ .

1.  $X \leq_{\text{sto}} Y$  implies  $g(X) \leq_{\text{sto}} g(Y)$  for any Radon function  $g$  that is eventually increasing.
2. If  $X_1, X_2$  are independent, and  $X'_1, X'_2$  are independent, then  $X_1 \leq_{\text{sto}} X'_1$  and  $X_2 \leq_{\text{sto}} X'_2$  imply  $g(X_1, X_2) \leq_{\text{sto}} g(X'_1, X'_2)$  for any Radon function  $g$  that is eventually increasing.

**Proof:**

- (1) Since  $g$  is a Radon function that is eventually increasing, there exists a threshold  $x_0 > 0$  such that  $g(\cdot)$  is increasing to  $\infty$  on  $[x_0, \infty)$ . By Theorem 19.2.2, there exists a random variable  $Z$  defined on the probability space  $(\Omega, \mathcal{F}, \mathbb{P})$  and nondecreasing functions  $\psi_1$  and  $\psi_2$  with  $\lim_{z \rightarrow b} \psi_i(z) = \infty$ ,  $i = 1, 2$ , such that  $X \stackrel{d}{=} \psi_1(Z)$  and  $Y \stackrel{d}{=} \psi_2(Z)$  and

$$\mathbb{P}\{\psi_1(Z) \leq \psi_2(Z) \mid Z \geq z_0\} = 1 \quad \text{for some } z_0 > 0 \text{ with } \psi_1(z_0) > x_0.$$

Thus,

$$P\{g(\psi_1(Z)) \leq g(\psi_2(Z)) \mid Z \geq z_0\} = 1 \quad \text{for } z_0 > 0.$$

Clearly,  $g(X) \stackrel{d}{=} g(\psi_1(Z))$  and  $g(Y) \stackrel{d}{=} g(\psi_2(Z))$ , and thus,

$$\begin{aligned} P\{g(X) > t\} &= P\{Z \geq z_0\}P\{g(\psi_1(Z)) > t \mid Z \geq z_0\} \\ &\leq P\{Z \geq z_0\}P\{g(\psi_2(Z)) > t \mid Z \geq z_0\} \\ &\leq P\{g(\psi_2(Z)) > t\} = P\{g(Y) > t\} \end{aligned}$$

for any  $t \geq c$  where  $c$  is a constant with  $c > g(\psi_1(z_0))$ .

- (2) Without loss of generality, assume that  $(X_1, X_2)$  and  $(X'_1, X'_2)$  are independent. We only need to show that  $g(X_1, X_2) \leq_{sto} g(X'_1, X_2)$ . Since  $g(\cdot)$  is a Radon function that is eventually increasing, there exists a  $(x_1, x_2)$  such that  $g(\cdot)$  is bounded on  $[0, x_1] \times [0, x_2]$  and increasing on  $([0, x_1] \times [0, x_2])^c$ .

1. By Theorem 19.2.2, there exists a random variable  $Z_1$  defined on the probability space  $(\Omega_1, \mathcal{F}_1, P_1)$  and nondecreasing functions  $\psi_1$  and  $\psi'_1$  such that  $X_1 \stackrel{d}{=} \psi_1(Z_1)$  and  $X'_1 \stackrel{d}{=} \psi'_1(Z_1)$  and  $P_1\{\psi_1(Z_1) \leq \psi'_1(Z_1) \mid Z_1 \geq z_1\} = 1$  for some  $z_1 > 0$  with  $\psi_1(z_1) > x_1$ .
2. Let  $(\Omega_2, \mathcal{F}_2, P_2)$  denote the underlying probability space of  $X_2$ .

Construct a product probability space  $(\Omega, \mathcal{F}, P) = (\Omega_1 \times \Omega_2, \mathcal{F}, P_1 \times P_2)$ , where  $\mathcal{F}$  is the  $\sigma$ -field generated by  $\mathcal{F}_1 \times \mathcal{F}_2$ . On this enlarged product probability space, since  $P\{\psi_1(Z_1) \leq \psi'_1(Z_1) \mid Z_1 \geq z_1\} = 1$ , and  $g(\cdot)$  is increasing on  $([0, x_1] \times [0, x_2])^c$ , we have that  $P\{g(\psi_1(Z_1), X_2) \leq g(\psi'_1(Z_1), X_2) \mid Z_1 \geq z_1 \text{ or } X_2 > x_2\} = 1$ . Clearly,  $g(X_1, X_2) \stackrel{d}{=} g(\psi_1(Z_1), X_2)$  and  $g(X'_1, X_2) \stackrel{d}{=} g(\psi'_1(Z_1), X_2)$ , and thus

$$\begin{aligned} &P\{g(X_1, X_2) > t\} \\ &= P\{Z_1 \geq z_1 \text{ or } X_2 > x_2\}P\{g(\psi_1(Z_1), X_2) > t \mid Z_1 \geq z_1 \text{ or } X_2 > x_2\} \\ &\leq P\{Z_1 \geq z_1 \text{ or } X_2 > x_2\}P\{g(\psi'_1(Z_1), X_2) > t \mid Z_1 \geq z_1 \text{ or } X_2 > x_2\} \\ &\leq P\{g(\psi'_1(Z_1), X_2) > t\} \\ &= P\{g(X'_1, X_2) > t\} \end{aligned}$$

for any  $t \geq c$  where  $c$  is a constant with  $c > g(\psi_1(z_1), x_2)$ . That is,  $g(X_1, X_2) \leq_{sto} g(X'_1, X_2)$ . Similarly,  $g(X'_1, X_2) \leq_{sto} g(X'_1, X'_2)$ . ■

**Corollary 19.2.5.** *If  $X_1, X_2$  are independent and  $Y_1, Y_2$  are independent, then  $X_1 \leq_{\text{sto}} Y_1$  and  $X_2 \leq_{\text{sto}} Y_2$  imply*

$$\begin{aligned} X_1 X_2 &\leq_{\text{sto}} Y_1 Y_2, & X_1 + X_2 &\leq_{\text{sto}} Y_1 + Y_2, \\ X_1 \vee X_2 &\leq_{\text{sto}} Y_1 \vee Y_2, & X_1 \wedge X_2 &\leq_{\text{sto}} Y_1 \wedge Y_2. \end{aligned}$$

In particular,  $R_1 \leq_{\text{sto}} R_2$  implies that  $R_1 V \leq_{\text{sto}} R_2 V$  for any non-negative random variable  $V$  that is independent of  $R_1, R_2$ . Mainik and Rüschendorf obtained this inequality in [300] for the random variable  $V$  that is bounded using the ordering  $\leq_{\text{apl}}$ , and their proof is based on the method of mixture.

**Proposition 19.2.6.** *Let  $X$  and  $Y$  be two positive random variables with support  $[0, \infty)$  and  $\Theta$  be a random variable with bounded support  $[\theta_L, \theta_U]$ . Assume that:*

1.  $\Theta$  is a random variable with finite masses.
2.  $\Theta$  is a continuous random variable such that  $\text{P}\{X > t \mid \Theta = \theta\}$  and  $\text{P}\{Y > t \mid \Theta = \theta\}$  are continuous in  $\theta$ .

*If  $[X \mid \Theta = \theta] \leq_{\text{sto}} [Y \mid \Theta = \theta]$  for all  $\theta$  in the support of  $\Theta$ , then  $X \leq_{\text{sto}} Y$ .*

**Proof:** Since  $[X \mid \Theta = \theta] \leq_{\text{sto}} [Y \mid \Theta = \theta]$  for all  $\theta \in [\theta_L, \theta_U]$ , there exists a threshold  $t_\theta$  that is given by

$$t_\theta := \sup \{s : \text{P}\{X > s \mid \Theta = \theta\} > \text{P}\{Y > s \mid \Theta = \theta\}\}, \quad (19.2.3)$$

such that

$$\text{P}\{X > t \mid \Theta = \theta\} \leq \text{P}\{Y > t \mid \Theta = \theta\}, \quad \forall t > t_\theta.$$

Notice that the threshold  $t_\theta$  depends on the mixing value  $\theta$ . Consider the following two cases. Construct the threshold  $t_{[\theta_L, \theta_U]}$  as follows:

1. If  $\Theta$  is discrete with finite masses, then define

$$t_{[\theta_L, \theta_U]} := \max\{t_\theta : \theta \in [\theta_L, \theta_U]\} < \infty.$$

2. If  $\Theta$  is continuous, then  $t_\theta$  is continuous in  $\theta$  due to the assumption that  $\text{P}\{X > t \mid \Theta = \theta\}$  and  $\text{P}\{Y > t \mid \Theta = \theta\}$  are continuous in  $\theta$ . Define

$$t_{[\theta_L, \theta_U]} := \sup\{t_\theta : \theta \in [\theta_L, \theta_U]\},$$

which is finite because of the continuity of  $t_\theta$  and the compactness of  $[\theta_L, \theta_U]$ .

In any case, for any  $\theta \in [\theta_L, \theta_U]$ , any  $t > t_{[\theta_L, \theta_U]}$ ,

$$P\{X > t \mid \Theta = \theta\} \leq P\{Y > t \mid \Theta = \theta\}.$$

Taking the integrations on both sides from  $\theta_L$  to  $\theta_U$ , we obtain  $P\{X > t\} \leq P\{Y > t\}$  for any  $t > t_{[\theta_L, \theta_U]}$ . ■

**Remark 19.2.7.** The closure property under mixture when the mixing variable has unbounded support, say  $[0, \infty)$ , becomes more subtle. This is because the threshold  $t_\theta$  defined in Eq. (19.2.3) can approach to infinity as  $\theta$  goes to infinity. Our conjecture is that in the case of unbounded support,  $[X \mid \Theta = \theta] \leq_{sto} [Y \mid \Theta = \theta]$  for all  $\theta$  in the support of  $\Theta$  implies that  $\limsup_{t \rightarrow \infty} \frac{P\{X > t\}}{P\{Y > t\}} \leq 1$ ; that is,  $X \leq_{apl} Y$ .

In the examples to be discussed below, all involved random variables fail to satisfy the usual stochastic order.

**Example 19.2.8.**

1. Let  $X$  have the Weibull distribution with unit scale parameter and shape parameter  $k$  and  $Y$  have the exponential distribution with unit (scale) parameter. If the shape parameter  $k > 1$  (i.e., increasing hazard rate), then  $X \leq_{sto} Y$ . If the shape parameter  $k < 1$  (i.e., decreasing hazard rate), then  $X \geq_{sto} Y$ . Note that both  $X$  and  $Y$  have exponentially decayed right tails.
2. Let  $X$  have the exponential distribution with unit (scale) parameter and  $Y$  have the distribution of Pareto Type II with tail index  $\alpha = 2$ ; that is,

$$P\{Y > t\} = (1 + t)^{-2}, \quad t \geq 0. \tag{19.2.4}$$

Then  $X \leq_{sto} Y$ . Note that  $Y$  has regularly varying right tail as described in Eq. (19.1.6).

3. If  $X$  has the Fréchet distribution with tail index  $\alpha = 3$  (see Theorem 19.1.1) and  $Y$  has the distribution Eq. (19.2.4) of Pareto Type II with tail index  $\alpha = 2$ , then  $X \leq_{sto} Y$ . Note that  $X$  and  $Y$  are regularly varying with respective tail indexes 3 and 2, but  $Y$  has a heavier tail than that of  $X$ .

4. Let  $X$  have the survival function of Pareto Type I as defined as follows:

$$P\{X > t\} = \left(\frac{t}{0.5}\right)^{-1}, \quad t \geq 0.5.$$

Let  $Y$  have the survival function of Pareto Type II with tail index 1; that is,

$$P\{Y > t\} = (1 + t)^{-1}, \quad t \geq 0.$$

Then  $X \leq_{\text{sto}} Y$ . Note that both  $X$  and  $Y$  have regularly varying right tails with same tail index 1.

5. Let  $R_1$  and  $R_2$  have regularly varying distributions with tail indexes  $\alpha_1$  and  $\alpha_2$ , respectively. If  $\alpha_1 > \alpha_2$ , then  $R_1 \leq_{\text{sto}} R_2$ . That is, the random variable with heavier regularly varying right tail is larger stochastically in the tail.
6. Let  $R$  be regularly varying with tail index  $\alpha$ . If  $V_1$  and  $V_2$  are random variables with finite moments of any order, independent of  $R$ , such that  $E[V_1^\alpha] < E[V_2^\alpha]$ . By Breiman's theorem (see [387], p. 232),

$$\lim_{t \rightarrow \infty} \frac{P\{RV_1 > t\}}{P\{R > t\}} = E[V_1^\alpha] < E[V_2^\alpha] = \lim_{t \rightarrow \infty} \frac{P\{RV_2 > t\}}{P\{R > t\}}$$

Thus, for  $t > t_0$  where  $t_0$  is sufficiently large,  $P\{RV_1 > t\} < P\{RV_2 > t\}$ , implying that  $RV_1 \leq_{\text{sto}} RV_2$ . ■

A multivariate extension of scale mixtures discussed in Example 19.2.8 (6) includes the multivariate elliptical distribution. A random vector  $\mathbf{X} \in \mathbb{R}^d$  is called elliptically distributed if  $\mathbf{X}$  has the representation:

$$\mathbf{X} \stackrel{d}{=} \boldsymbol{\mu} + RA\mathbf{U} \tag{19.2.5}$$

where  $\boldsymbol{\mu} \in \mathbb{R}^d$ ,  $A \in \mathbb{R}^{d \times d}$  and  $\mathbf{U}$  is uniformly distributed on  $\mathbb{S}_2^{d-1} := \{\mathbf{x} \in \mathbb{R}^d : \|\mathbf{x}\|_2 = 1\}$  and  $R \geq 0$  is independent of  $\mathbf{U}$ . We denote this by  $\mathbf{X} \sim \mathcal{E}(\boldsymbol{\mu}, \Sigma, R)$  where  $\Sigma = AA^\top$ .

**Proposition 19.2.9.** *Let  $\mathbf{X} \sim \mathcal{E}(\boldsymbol{\mu}_1, \Sigma_1, R_1)$  and  $\mathbf{Y} \sim \mathcal{E}(\boldsymbol{\mu}_2, \Sigma_2, R_2)$ . If*

*$\boldsymbol{\mu}_1 \leq \boldsymbol{\mu}_2$ ,  $R_1 \leq_{\text{sto}} R_2$ , and  $\boldsymbol{\xi}^\top \Sigma_1 \boldsymbol{\xi} \leq \boldsymbol{\xi}^\top \Sigma_2 \boldsymbol{\xi}$ , for fixed  $\boldsymbol{\xi} \in \mathbb{S}_1^{d-1} := \{\mathbf{x} \in \mathbb{R}^d : \|\mathbf{x}\|_1 = 1\}$ ,*

*then  $|\boldsymbol{\xi}^\top \mathbf{X}| \leq_{\text{sto}} |\boldsymbol{\xi}^\top \mathbf{Y}|$ .*

**Proof:** Let  $a_i := \boldsymbol{\xi}^\top \Sigma_i \boldsymbol{\xi}$ ,  $i = 1, 2$ . Without loss of generality, we can assume that  $\boldsymbol{\mu}_1 = \boldsymbol{\mu}_2 = \mathbf{0}$  and  $a_1 > 0$ . Let

$$\mathbf{v}_i := \frac{A^\top \boldsymbol{\xi}}{\boldsymbol{\xi}^\top \Sigma_i \boldsymbol{\xi}} = \frac{A^\top \boldsymbol{\xi}}{a_i}, \quad i = 1, 2.$$

Clearly  $\mathbf{v}_i^\top \mathbf{v}_i = 1$ ,  $i = 1, 2$ , and thus, by symmetry,  $\mathbf{v}_1^\top \mathbf{U}$  and  $\mathbf{v}_2^\top \mathbf{U}$  have the same distribution. Let  $\Theta := \mathbf{v}_1^\top \mathbf{U}$ , and we have

$$|\boldsymbol{\xi}^\top \mathbf{X}| = R_1 a_1 |\mathbf{v}_1^\top \mathbf{U}| \stackrel{d}{=} R_1 a_1 |\Theta|, \quad |\boldsymbol{\xi}^\top \mathbf{Y}| = R_2 a_2 |\mathbf{v}_2^\top \mathbf{U}| \stackrel{d}{=} R_2 a_2 |\Theta|.$$

The inequality then follows from Corollary 19.2.5 immediately. ■

This is our  $\leq_{\text{sto}}$ -version of a similar result that is obtained in [300] using the  $\leq_{\text{apl}}$  order.

**Remark 19.2.10.** Anderson in [12], Fefferman, Jodeit, and Perlman in [159] show that if  $\boldsymbol{\mu}_1 = \boldsymbol{\mu}_2$ ,  $R_1 \stackrel{d}{=} R_2$ , and

$$\boldsymbol{\xi}^\top \Sigma_1 \boldsymbol{\xi} \leq \boldsymbol{\xi}^\top \Sigma_2 \boldsymbol{\xi}, \quad \forall \boldsymbol{\xi} \in \mathbb{R}^d,$$

then  $E(\psi(\mathbf{X})) \leq E(\psi(\mathbf{Y}))$  for all symmetric and convex functions  $\psi : \mathbb{R}^d \mapsto \mathbb{R}$ , such that the expectations exist. This is known as *dilatation*, which can be defined on any locally convex topological linear space  $\mathbb{V}$  (traced back to Karamata’s work in 1932; see [308], pp. 16–17). Dilatation provides various versions of continuous majorization [308].

### 19.3 Tail Orthant Orders

Let  $\mathbf{X} = (X_1, \dots, X_d)$  and  $\mathbf{X}' = (X'_1, \dots, X'_d)$  be nonnegative random vectors with dfs  $F$  and  $F'$ , respectively. Observe that  $\mathbf{X} \leq_{\text{lo}} \mathbf{X}'$  is equivalent to that  $\max_{1 \leq i \leq d} \{X_i/w_i\} \geq_{\text{st}} \max_{1 \leq i \leq d} \{X'_i/w_i\}$ , and  $\mathbf{X} \leq_{\text{uo}} \mathbf{X}'$  is equivalent to that  $\min_{1 \leq i \leq d} \{X_i/w_i\} \leq_{\text{st}} \min_{1 \leq i \leq d} \{X'_i/w_i\}$ . In comparing orthant tails of these random vectors, we focus on the cones  $\mathbb{R}_+^d$  and  $\bar{\mathbb{R}}_+^d := \{(x_1, \dots, x_d) : x_i > 0, 1 \leq i \leq d\}$ . Note that  $\mathbf{x} \in \bar{\mathbb{R}}_+^d$  can have some components taking  $+\infty$ .

**Definition 19.3.1.**

1.  $\mathbf{X}$  is said to be smaller than  $\mathbf{X}'$  in the sense of tail lower orthant order, denoted as  $\mathbf{X} \leq_{\text{tlo}} \mathbf{X}'$ , if for all  $\mathbf{w} = (w_1, \dots, w_d) \in \bar{\mathbb{R}}_+^d$ ,  $\max_{1 \leq i \leq d} \{X_i/w_i\} \geq_{\text{sto}} \max_{1 \leq i \leq d} \{X'_i/w_i\}$ .

2.  $\mathbf{X}$  is said to be smaller than  $\mathbf{X}'$  in the sense of tail upper orthant order, denoted as  $\mathbf{X} \leq_{\text{tuo}} \mathbf{X}'$ , if for all  $\mathbf{w} = (w_1, \dots, w_d) \in \mathbb{R}_+^d$ ,
 
$$\min_{1 \leq i \leq d} \{X_i/w_i\} \leq_{\text{sto}} \min_{1 \leq i \leq d} \{X'_i/w_i\}.$$

It follows from Eq. (19.2.1) that  $\mathbf{X} \leq_{\text{tlo}} \mathbf{X}'$  is equivalent to for  $\mathbf{w} = (w_1, \dots, w_d) \in \overline{\mathbb{R}}_+^d$ ,

$$P\{X_1 \leq tw_1, \dots, X_d \leq tw_d\} \leq P\{X'_1 \leq tw_1, \dots, X'_d \leq tw_d\} \quad (19.3.1)$$

for all  $t > t_{\mathbf{w}}$  for some  $t_{\mathbf{w}} > 0$  that may depend on  $\mathbf{w}$ . Similarly,  $\mathbf{X} \leq_{\text{tuo}} \mathbf{X}'$  is equivalent to for  $\mathbf{w} = (w_1, \dots, w_d) \in \mathbb{R}_+^d$ ,

$$P\{X_1 > tw_1, \dots, X_d > tw_d\} \leq P\{X'_1 > tw_1, \dots, X'_d > tw_d\} \quad (19.3.2)$$

for all  $t > t_{\mathbf{w}}$  for some  $t_{\mathbf{w}} > 0$  that may depend on  $\mathbf{w}$ .

In comparing tail dependence, however, we assume that all the margins of  $F$  and  $F'$  are tail equivalent. Since we need to compare upper interior orthant tails given fixed marginal tails, consider the two smaller cones:

1.  $\mathbb{C}_l := \overline{\mathbb{R}}_+^d \setminus \cup_{j=1}^d \{t\mathbf{e}_j^{-1}, t \geq 0\}$ , where  $\mathbf{e}_j^{-1}$ ,  $1 \leq j \leq d$ , denotes the vector with the  $j$ -th component being 1 and infinity otherwise.
2.  $\mathbb{C}_u := \mathbb{R}_+^d \setminus \cup_{j=1}^d \{t\mathbf{e}_j, t \geq 0\}$ , where  $\mathbf{e}_j$ ,  $1 \leq j \leq d$ , denotes the vector with the  $j$ -th component being 1 and zero otherwise.

That is,  $\mathbb{C}_l$  and  $\mathbb{C}_u$  are the subsets of  $\overline{\mathbb{R}}_+^d$  after eliminating all the axes that correspond to the margins of a distribution. Note that  $\mathbf{w} = (w_1, \dots, w_d) \in \mathbb{C}_l$  if and only if at least two components of  $w$  are finite, and  $\mathbf{w} = (w_1, \dots, w_d) \in \mathbb{C}_u$  if and only if at least two components of  $w$  are positive.

**Definition 19.3.2.**

1.  $\mathbf{X}$  is said to be smaller than  $\mathbf{X}'$  in the sense of tail lower orthant dependence order, denoted as  $\mathbf{X} \leq_{\text{tlod}} \mathbf{X}'$ , if all the margins of  $F$  and  $F'$  are tail equivalent, and for all  $\mathbf{w} = (w_1, \dots, w_d) \in \mathbb{C}_l$ ,

$$\max_{1 \leq i \leq d} \{X_i/w_i\} \geq_{\text{sto}} \max_{1 \leq i \leq d} \{X'_i/w_i\}.$$

2.  $\mathbf{X}$  is said to be smaller than  $\mathbf{X}'$  in the sense of tail upper orthant dependence order, denoted as  $\mathbf{X} \leq_{\text{tuod}} \mathbf{X}'$ , if all the margins of  $F$  and  $F'$  are tail equivalent, and for all  $\mathbf{w} = (w_1, \dots, w_d) \in \mathbb{C}_u$ ,

$$\min_{1 \leq i \leq d} \{X_i/w_i\} \leq_{\text{sto}} \min_{1 \leq i \leq d} \{X'_i/w_i\}.$$

It follows from Eq. (19.2.1) that if all the margins of  $F$  and  $F'$  are tail equivalent, then  $\mathbf{X} \leq_{\text{tlod}} \mathbf{X}'$  is equivalent to that Eq. (19.3.1) holds for  $\mathbf{w} = (w_1, \dots, w_d) \in \mathbb{C}_l$ , and  $\mathbf{X} \leq_{\text{tuod}} \mathbf{X}'$  is equivalent to that Eq. (19.3.2) holds for  $\mathbf{w} = (w_1, \dots, w_d) \in \mathbb{C}_u$ .

**Remark 19.3.3.**

1. If

$$(w_1, \dots, w_d) \in \bigcup_{j=1}^d \{t\mathbf{e}_j^{-1}, t \geq 0\} \quad \text{or} \quad (w_1, \dots, w_d) \in \bigcup_{j=1}^d \{t\mathbf{e}_j, t \geq 0\},$$

then the inequalities in Definition 19.3.2 reduce to the stochastic tail orders of the marginal distributions. Since the margins are assumed to be tail equivalent, which may not satisfy stochastic tail comparison, we need to eliminate the margins from consideration. On the other hand, with given fixed marginal tails, what really matters in dependence comparison is various interior orthant subsets of  $\mathbb{C}_l$  or  $\mathbb{C}_u$ .

2. If some corresponding margins of  $F$  and  $F'$  are not tail equivalent, one can still define the tail orthant orders  $\leq_{\text{tlo}}$  and  $\leq_{\text{tuo}}$  to compare their tail behaviors in orthants. But all corresponding margins of  $F$  and  $F'$  have to be tail equivalent in order to compare their tail dependence.
3. If some margins of  $F$  (or  $F'$ ) are not tail equivalent, then one can still define the tail orthant dependence order, but scaling functions would be different among the components.
4. Another alternative is to convert all the margins of  $F$  and  $F'$  to standard Pareto margins, resulting in Pareto copulas [245], and then compare their Pareto copulas using the  $\leq_{\text{tlod}}$  and  $\leq_{\text{tuod}}$  orders.

The preservation properties under the  $\leq_{\text{tlod}}$  and  $\leq_{\text{tuod}}$  orders can be easily established using Definitions 19.3.1 and 19.3.2, and Propositions 19.2.4 and 19.2.6. In particular, we have the following.

**Proposition 19.3.4.** *Let  $\mathbf{X} = (X_1, \dots, X_d)$ ,  $\mathbf{X}' = (X'_1, \dots, X'_d)$  and  $\mathbf{Y} = (Y_1, \dots, Y_d)$ ,  $\mathbf{Y}' = (Y'_1, \dots, Y'_d)$  be positive random vectors with support  $\mathbb{R}_+^d$ , and  $\Theta$  be a random variable with bounded support. Assume that  $(\mathbf{X}, \mathbf{X}')$  and  $(\mathbf{Y}, \mathbf{Y}')$  are independent, and the regularity conditions of Proposition 19.2.6 are satisfied:*



1.  $\mathbf{X} \leq_{\text{tlo}} \mathbf{X}'$  and  $\mathbf{Y} \leq_{\text{tlo}} \mathbf{Y}'$  imply that  $\mathbf{X} \vee \mathbf{Y} \leq_{\text{tlo}} \mathbf{X}' \vee \mathbf{Y}'$ .  
 $\mathbf{X} \leq_{\text{tlod}} \mathbf{X}'$  and  $\mathbf{Y} \leq_{\text{tlod}} \mathbf{Y}'$  imply that  $\mathbf{X} \vee \mathbf{Y} \leq_{\text{tlod}} \mathbf{X}' \vee \mathbf{Y}'$ .
2.  $\mathbf{X} \leq_{\text{tuo}} \mathbf{X}'$  and  $\mathbf{Y} \leq_{\text{tuo}} \mathbf{Y}'$  imply that  $\mathbf{X} \wedge \mathbf{Y} \leq_{\text{tuo}} \mathbf{X}' \wedge \mathbf{Y}'$ .  
 $\mathbf{X} \leq_{\text{tuod}} \mathbf{X}'$  and  $\mathbf{Y} \leq_{\text{tuod}} \mathbf{Y}'$  imply that  $\mathbf{X} \wedge \mathbf{Y} \leq_{\text{tuod}} \mathbf{X}' \wedge \mathbf{Y}'$ .
3. If  $[\mathbf{X} \mid \Theta = \theta] \leq_{\text{tlo}} [\mathbf{X}' \mid \Theta = \theta]$  for all  $\theta$  in the bounded support of  $\Theta$ , then  $\mathbf{X} \leq_{\text{tlo}} \mathbf{X}'$ . If  $[\mathbf{X} \mid \Theta = \theta] \leq_{\text{tlod}} [\mathbf{X}' \mid \Theta = \theta]$  for all  $\theta$  in the bounded support of  $\Theta$ , then  $\mathbf{X} \leq_{\text{tlod}} \mathbf{X}'$ .
4. If  $[\mathbf{X} \mid \Theta = \theta] \leq_{\text{tuo}} [\mathbf{X}' \mid \Theta = \theta]$  for all  $\theta$  in the bounded support of  $\Theta$ , then  $\mathbf{X} \leq_{\text{tuo}} \mathbf{X}'$ . If  $[\mathbf{X} \mid \Theta = \theta] \leq_{\text{tuod}} [\mathbf{X}' \mid \Theta = \theta]$  for all  $\theta$  in the bounded support of  $\Theta$ , then  $\mathbf{X} \leq_{\text{tuod}} \mathbf{X}'$ .

**Example 19.3.5.** Let  $\mathbf{X} \sim \mathcal{E}(\mathbf{0}, \Sigma_1, R_1)$  and  $\mathbf{X}' \sim \mathcal{E}(\mathbf{0}, \Sigma_2, R_2)$  [see Eq. (19.2.5)], where  $\Sigma_1 = A_1 A_1^\top = (\sigma_{ij})$  and  $\Sigma_2 = A_2 A_2^\top = (\lambda_{ij})$ . Consider  $\mathbf{X}_+ = \mathbf{X} \vee \mathbf{0}$  and  $\mathbf{X}'_+ = \mathbf{X}' \vee \mathbf{0}$ :

1. Suppose that

$$R_1 \leq_{\text{sto}} R_2, \Sigma_1 \leq \Sigma_2 \text{ component-wise with } \sigma_{ii} = \lambda_{ii}, i = 1, \dots, d.$$

It follows from Example 9.A.8 in [426] that  $\mathbf{X}_+ \leq_{\text{uo}} R_1(A_2 \mathbf{U} \vee \mathbf{0})$ , which implies that  $\mathbf{X}_+ \leq_{\text{tuo}} R_1(A_2 \mathbf{U} \vee \mathbf{0})$ . Clearly,

$$\underbrace{(R_1, \dots, R_1)}_d \leq_{\text{tuo}} \underbrace{(R_2, \dots, R_2)}_d,$$

which, together with Proposition 19.3.4 (4) and the fact that  $A_2 \mathbf{U}$  has a bounded support, imply that  $\mathbf{X}_+ \leq_{\text{tuo}} R_1(A_2 \mathbf{U} \vee \mathbf{0}) \leq_{\text{tuo}} R_2(A_2 \mathbf{U} \vee \mathbf{0})$ . Thus  $\mathbf{X}_+ \leq_{\text{tuo}} \mathbf{X}'_+$ .

2. Suppose that

$$R_1 \geq_{\text{sto}} R_2, \Sigma_1 \leq \Sigma_2 \text{ component-wise with } \sigma_{ii} = \lambda_{ii}, i = 1, \dots, d.$$

It follows from Example 9.A.8 in [426] that  $\mathbf{X}_+ \leq_{\text{lo}} R_1(A_2 \mathbf{U} \vee \mathbf{0})$ , which implies that  $\mathbf{X}_+ \leq_{\text{tlo}} R_1(A_2 \mathbf{U} \vee \mathbf{0})$ . Clearly,

$$\underbrace{(R_1, \dots, R_1)}_d \leq_{\text{tlo}} \underbrace{(R_2, \dots, R_2)}_d,$$

which, together with Proposition 19.3.4 (3) and the fact that  $A_2 \mathbf{U}$  has a bounded support, imply that  $\mathbf{X}_+ \leq_{\text{tlo}} R_1(A_2 \mathbf{U} \vee \mathbf{0}) \leq_{\text{tlo}} R_2(A_2 \mathbf{U} \vee \mathbf{0})$ . Thus  $\mathbf{X}_+ \leq_{\text{tlo}} \mathbf{X}'_+$ . ■

To construct a wide class of examples involving the  $\leq_{\text{tlod}}$  and  $\leq_{\text{tuod}}$  orders, we employ the copula approach. A copula  $C$  is a multivariate distribution with standard uniformly distributed margins on  $[0, 1]$ . Sklar’s theorem (see, e.g., [211], Sect.1.6) states that every multivariate distribution  $F$  with margins  $F_1, \dots, F_d$  can be written as  $F(x_1, \dots, x_d) = C(F_1(x_1), \dots, F_d(x_d))$  for some  $d$ -dimensional copula  $C$ . In fact, in the case of continuous margins,  $C$  is unique and

$$C(u_1, \dots, u_d) = F(F_1^{-1}(u_1), \dots, F_d^{-1}(u_d))$$

where  $F_i^{-1}(u_i)$  are the quantile functions of the  $i$ -th margin,  $1 \leq i \leq d$ . Let  $(U_1, \dots, U_d)$  denote a random vector with  $U_i, 1 \leq i \leq d$ , being uniformly distributed on  $[0, 1]$ . The survival copula  $\widehat{C}$  is defined as follows:

$$\widehat{C}(u_1, \dots, u_n) = P\{1-U_1 \leq u_1, \dots, 1-U_n \leq u_n\} = \overline{C}(1-u_1, \dots, 1-u_n) \tag{19.3.3}$$

where  $\overline{C}$  is the joint survival function of  $C$ . The upper exponent and upper tail dependence functions (see [207, 213, 244, 360]) are defined as follows,

$$a(\mathbf{w}; C) := \lim_{u \rightarrow 0^+} \frac{P\{\cup_{i=1}^d \{U_i > 1 - uw_i\}\}}{u}, \quad \forall \mathbf{w} = (w_1, \dots, w_d) \in \mathbb{R}_+^d \setminus \{0\} \tag{19.3.4}$$

$$b(\mathbf{w}; C) := \lim_{u \rightarrow 0^+} \frac{P\{\cap_{i=1}^d \{U_i > 1 - uw_i\}\}}{u}, \quad \forall \mathbf{w} = (w_1, \dots, w_d) \in \mathbb{R}_+^d \setminus \{0\} \tag{19.3.5}$$

provided that the limits exist. Note that both  $a(\mathbf{w}; C)$  and  $b(\mathbf{w}; C)$  are homogeneous of order 1 in the sense that  $a(c\mathbf{w}; C) = ca(\mathbf{w}; C)$  and  $b(c\mathbf{w}; C) = cb(\mathbf{w}; C)$  for any  $c > 0$ .

**Theorem 19.3.6.** *Let  $\mathbf{X} = (X_1, \dots, X_d)$  and  $\mathbf{X}' = (X'_1, \dots, X'_d)$  be two  $d$ -dimensional random vectors with respective copulas  $C$  and  $C'$  and their respective continuous margins  $F_1, \dots, F_d$  and  $F'_1, \dots, F'_d$ .*

1. *If  $C = C'$  and  $F_i \leq_{\text{sto}} F'_i, 1 \leq i \leq d$ , then  $\mathbf{X} \leq_{\text{tuo}} \mathbf{X}'$ .*
2. *Assume that the upper tail dependence functions  $b(\cdot; C)$  and  $b(\cdot; C')$  exist, and  $\overline{F}_i \in RV_{-\alpha_i}$  and  $\overline{F}'_i \in RV_{-\alpha'_i}, i = 1, \dots, d$ . If  $F_i$  and  $F'_i, 1 \leq i \leq d$ , are all tail equivalent, and  $b(\mathbf{w}; C) < b(\mathbf{w}; C')$  for all  $\mathbf{w} = (w_1, \dots, w_d)$  with  $w_i > 0, 1 \leq i \leq d$ , then  $\mathbf{X} \leq_{\text{tuod}} \mathbf{X}'$ .*

3. Assume that the upper tail dependence functions  $b(\cdot; C)$  and  $b(\cdot; C')$  exist, and  $\bar{F}_i \in RV_{-\alpha_i}$  and  $\bar{F}'_i \in RV_{-\alpha'_i}$ ,  $i = 1, \dots, d$ . If  $F_i \leq_{\text{sto}} F'_i$ ,  $1 \leq i \leq d$ , and  $b(\mathbf{w}; C) < b(\mathbf{w}; C')$  for all  $\mathbf{w} = (w_1, \dots, w_d)$  with  $w_i > 0$ ,  $1 \leq i \leq d$ , then  $\mathbf{X} \leq_{\text{tuo}} \mathbf{X}'$ .

**Proof:**

- (1) Since  $F_i(tw_i) \geq F'_i(tw_i)$ ,  $1 \leq i \leq d$ , for all  $t > t_{w_i}$ , we have, for all  $t > \max\{t_{w_i}, 1 \leq i \leq d\}$

$$\begin{aligned} & \mathbb{P}\{X_1 > tw_1, \dots, X_d > tw_d\} \\ &= \mathbb{P}\{F_1(X_1) > F_1(tw_1), \dots, F_d(X_d) > F_d(tw_d)\} \\ &\leq \mathbb{P}\{F_1(X_1) > F'_1(tw_1), \dots, F_d(X_d) > F'_d(tw_d)\} \\ &= \mathbb{P}\{F'_1(X'_1) > F'_1(tw_1), \dots, F'_d(X'_d) > F'_d(tw_d)\} \\ &= \mathbb{P}\{X'_1 > tw_1, \dots, X'_d > tw_d\}. \end{aligned}$$

- (2) Write  $\bar{F}_i(t) = L_i(t)t^{-\alpha_i}$  and  $\bar{F}'_i(t) = L'_i(t)t^{-\alpha'_i}$ ,  $1 \leq i \leq d$ . Since all the margins are tail equivalent, we have

$$\alpha_i = \alpha'_i =: \alpha, \text{ and } \lim_{t \rightarrow \infty} \frac{L_i(t)}{L_1(t)} = \lim_{t \rightarrow \infty} \frac{L'_i(t)}{L_1(t)} = 1, \quad 1 \leq i \leq d.$$

In addition, since the functions  $L_i(\cdot)$ s and  $L'_i(\cdot)$ s are slowly varying, then for all  $\mathbf{w} = (w_1, \dots, w_d)$  with  $w_i > 0$ ,  $1 \leq i \leq d$ ,

$$\lim_{t \rightarrow \infty} \frac{L_i(tw_i)}{L_1(t)} = \lim_{t \rightarrow \infty} \frac{L'_i(tw_i)}{L_1(t)} = 1, \quad 1 \leq i \leq d.$$

That is, for any  $\varepsilon > 0$ , there exists  $t_w$  that is sufficiently large, such that, for  $1 \leq i \leq d$  and all  $t > t_w$ ,

$$(1 - \varepsilon)L_1(t) \leq L_i(tw_i) \leq (1 + \varepsilon)L_1(t),$$

$$(1 - \varepsilon)L_1(t) \leq L'_i(tw_i) \leq (1 + \varepsilon)L_1(t).$$

Observe that

$$\begin{aligned} & \mathbb{P}\{X_1 > tw_1, \dots, X_d > tw_d\} \\ &= \mathbb{P}\{F_i(X_i) > 1 - L_i(tw_i)t^{-\alpha}w_i^{-\alpha}, 1 \leq i \leq d\} \\ &\leq \mathbb{P}\{F_i(X_i) > 1 - L_1(t)t^{-\alpha}(1 + \varepsilon)w_i^{-\alpha}, 1 \leq i \leq d\}, \end{aligned}$$

and thus

$$\begin{aligned} & \limsup_{t \rightarrow \infty} \frac{\mathbb{P}\{X_1 > tw_1, \dots, X_d > tw_d\}}{\overline{F}_1(t)} \\ & \leq \lim_{t \rightarrow \infty} \frac{\mathbb{P}\{F_i(X_i) > 1 - \overline{F}_1(t)(1 + \varepsilon)w_i^{-\alpha}, 1 \leq i \leq d\}}{\overline{F}_1(t)} \\ & = b((1 + \varepsilon)\mathbf{w}^{-\alpha}; C) = (1 + \varepsilon)b(\mathbf{w}^{-\alpha}; C). \end{aligned}$$

Similarly,

$$\liminf_{t \rightarrow \infty} \frac{\mathbb{P}\{X_1 > tw_1, \dots, X_d > tw_d\}}{\overline{F}_1(t)} \geq (1 - \varepsilon)b(\mathbf{w}^{-\alpha}; C).$$

Let  $\varepsilon \rightarrow 0$ , we have

$$\lim_{t \rightarrow \infty} \frac{\mathbb{P}\{X_1 > tw_1, \dots, X_d > tw_d\}}{\overline{F}_1(t)} = b(\mathbf{w}^{-\alpha}; C).$$

For  $\mathbf{X}'$  with copula  $C'$ , we have

$$\lim_{t \rightarrow \infty} \frac{\mathbb{P}\{X'_1 > tw_1, \dots, X'_d > tw_d\}}{\overline{F}_1(t)} = b(\mathbf{w}^{-\alpha}; C').$$

Since  $b(\mathbf{w}^{-\alpha}; C) < b(\mathbf{w}^{-\alpha}; C')$  for each  $\mathbf{w} = (w_1, \dots, w_d)$  with  $w_i > 0, 1 \leq i \leq d$ , there exists  $t_{\mathbf{w}}$  such that, for all  $t > t_{\mathbf{w}}$ ,

$$\mathbb{P}\{X_1 > tw_1, \dots, X_d > tw_d\} \leq \mathbb{P}\{X'_1 > tw_1, \dots, X'_d > tw_d\}.$$

(3) The stochastic tail order follows from (1) and (2). ■

For the  $\leq_{\text{tlo}}$  order, we can establish a similar result.

**Theorem 19.3.7.** *Let  $\mathbf{X} = (X_1, \dots, X_d)$  and  $\mathbf{X}' = (X'_1, \dots, X'_d)$  be  $d$ -dimensional random vectors with respective copulas  $C$  and  $C'$  and respective continuous margins  $F_1, \dots, F_d$  and  $F'_1, \dots, F'_d$ .*

1. *If  $C = C'$  and  $F_i \geq_{\text{sto}} F'_i, 1 \leq i \leq d$ , then  $\mathbf{X} \leq_{\text{tlo}} \mathbf{X}'$ .*
2. *Assume that the exponent functions  $a(\cdot; C)$  and  $a(\cdot; C')$  exist, and  $\overline{F}_i \in RV_{-\alpha_i}$  and  $\overline{F}'_i \in RV_{-\alpha'_i}, i = 1, \dots, d$ . If  $F_i$  and  $F'_i, 1 \leq i \leq d$ , are all tail equivalent, and  $a(\mathbf{w}; C) > a(\mathbf{w}; C')$  for all  $\mathbf{w} = (w_1, \dots, w_d)$  with  $w_i > 0, 1 \leq i \leq d$ , then  $\mathbf{X} \leq_{\text{tlo}} \mathbf{X}'$ .*
3. *Assume that the upper tail dependence functions  $a(\cdot; C)$  and  $a(\cdot; C')$  exist, and  $\overline{F}_i \in RV_{-\alpha_i}$  and  $\overline{F}'_i \in RV_{-\alpha'_i}, i = 1, \dots, d$ . If  $F_i \geq_{\text{sto}} F'_i, 1 \leq i \leq d$ , and  $a(\mathbf{w}; C) > a(\mathbf{w}; C')$  for all  $\mathbf{w} = (w_1, \dots, w_d)$  with  $w_i > 0, 1 \leq i \leq d$ , then  $\mathbf{X} \leq_{\text{tlo}} \mathbf{X}'$ .*

**Example 19.3.8.** The tail dependence functions of Archimedean copulas were derived in [34, 84] (also see Propositions 2.5 and 3.3 in [213]). Let  $C_{\text{Arch}}(\mathbf{u}; \phi) = \phi(\sum_{i=1}^d \phi^{-1}(u_i))$  be an Archimedean copula with strict generator  $\phi^{-1}$ , where  $\phi$  is regularly varying at  $\infty$  with tail index  $\theta > 0$ . The upper tail dependence function of the survival copula  $\widehat{C}_{\text{Arch}}$  is given by

$$b(\mathbf{w}; \widehat{C}_{\text{Arch}}) = \left( \sum_{j=1}^d w_j^{-1/\theta} \right)^{-\theta}.$$

Observe that  $b(\mathbf{w}; \widehat{C}_{\text{Arch}})$  is strictly increasing in  $\theta$ . For  $\theta < \theta'$ , and  $C$  and  $C'$  be two copulas with df  $\widehat{C}_{\text{Arch}}$  having parameters  $\theta$  and  $\theta'$  respectively. Thus  $b(\mathbf{w}; C) < b(\mathbf{w}; C')$  for all  $\mathbf{w} > \mathbf{0}$ . For  $1 \leq i \leq d$ , let  $F_i$  have the Fréchet df with tail index  $\alpha$  (i.e.,  $F_i(s) = \exp\{-s^{-\alpha}\}$ ) and  $F'_i$  have the distribution of Pareto Type II with tail index  $\alpha$  (i.e.,  $F'_i(s) = 1 - (1 + s)^{-\alpha}$ ). Clearly,  $F_i$  and  $F'_i$  are tail equivalent. Let  $\mathbf{X}$  and  $\mathbf{X}'$  have the dfs of

$$C(F_1(x_1), \dots, F_d(x_d)), \text{ and } C'(F'_1(x_1), \dots, F'_d(x_d)),$$

respectively, and by Theorem 19.3.6,  $\mathbf{X} \leq_{\text{tuod}} \mathbf{X}'$ . ■

**Example 19.3.9.** The exponent functions of Archimedean copulas were derived in [34, 177] (also see Propositions 2.5 and 3.3 in [213]). Let  $C_{\text{Arch}}(\mathbf{u}; \phi) = \phi(\sum_{i=1}^d \phi^{-1}(u_i))$  be an Archimedean copula, where the generator  $\phi^{-1}$  is regularly varying at 1 with tail index  $\beta > 1$ . The upper exponent function of  $C_{\text{Arch}}$  is given by

$$a(\mathbf{w}; C_{\text{Arch}}) = \left( \sum_{j=1}^d w_j^\beta \right)^{1/\beta}.$$

Observe that  $a(\mathbf{w}; C_{\text{Arch}})$  is strictly decreasing in  $\beta$ . For  $\beta < \beta'$ , and  $C$  and  $C'$  be two copulas with df  $C_{\text{Arch}}$  having parameters  $\beta$  and  $\beta'$ , respectively. Thus  $a(\mathbf{w}; C) > a(\mathbf{w}; C')$  for all  $w > 0$ . For  $1 \leq i \leq d$ , let  $F_i$  have the Fréchet df with tail index  $\alpha$  (i.e.,  $F_i(s) = \exp\{-s^{-\alpha}\}$ ) and  $F'_i$  have the distribution of Pareto Type II with tail index  $\alpha$  (i.e.,  $F'_i(s) = 1 - (1 + s)^{-\alpha}$ ), such that  $F_i$  and  $F'_i$  are tail equivalent. Let  $\mathbf{X}$  and  $\mathbf{X}'$  have the dfs of

$$C(F_1(x_1), \dots, F_d(x_d)) \text{ and } C'(F'_1(x_1), \dots, F'_d(x_d)),$$

respectively, and by Theorem 19.3.7,  $\mathbf{X} \leq_{\text{t lod}} \mathbf{X}'$ . ■

**Remark 19.3.10.** Due to the homogeneity property, the conditions on tail dependence and exponent functions used in Theorem 19.3.6 (2) and (3) and in Theorem 19.3.7 (2) and (3) can be simplified. For example, it is sufficient in Theorem 19.3.6 (2) and (3) to verify that  $b(\mathbf{w}; C) < b(\mathbf{w}; C')$  for all  $\mathbf{w} = (w_1, \dots, w_d)$  with  $w_i > 0$ ,  $1 \leq i \leq d$ , and  $\|\mathbf{w}\| = 1$ , where  $\|\cdot\|$  denotes any norm on  $\mathbb{R}_+^d$ .

**Theorem 19.3.11.** Let  $(\mathbf{X}_n, n \geq 1)$  and  $(\mathbf{X}'_n, n \geq 1)$  be two i.i.d. samples with dfs  $F$  and  $F'$ , respectively. Assume that  $F \in DA_\vee(G)$  and  $F' \in DA_\vee(G')$  with tail equivalent Fréchet margins.

1. If  $\mathbf{X}_n \leq_{\text{t lod}} \mathbf{X}'_n$ , then  $G \leq_{\text{lod}} G'$ .
2.  $G \leq_{\text{lod}} G'$  if and only if  $G \leq_{\text{t lod}} G'$ .

**Proof:** Let  $\mathbf{Y} = (Y_1, \dots, Y_d)$  and  $\mathbf{Y}' = (Y'_1, \dots, Y'_d)$  denote two random vectors that have the same distributions as these of  $\mathbf{X}_n$  and  $\mathbf{X}'_n$ , respectively.

(1) It follows from Theorem 19.1.1 and Remark 19.1.2 (1) that

$$[\mathbb{P}\{\mathbf{X}_k \leq \mathbf{a}_n \mathbf{x}\}]^n \rightarrow G(\mathbf{x}), \quad \forall \mathbf{x} = (x_1, \dots, x_d) \in \mathbb{R}_+^d,$$

where  $\mathbf{a}_n = (a_{1,n}, \dots, a_{d,n}) = (\bar{F}_1^{-1}(1/n), \dots, \bar{F}_d^{-1}(1/n))$ . Taking the logarithm on both sides, we have, as  $n \rightarrow \infty$ ,

$$n \log \mathbb{P}\{\mathbf{X}_k \leq \mathbf{a}_n \mathbf{x}\} \approx -n \mathbb{P}\{\cup_{i=1}^d \{Y_i > a_{i,n} x_i\}\} \rightarrow \log G(\mathbf{x}). \tag{19.3.6}$$

Since the margins are tail equivalent,  $a_{i,n}/a_{1,n} \rightarrow 1$  as  $n \rightarrow \infty$ . Thus, for any small  $\varepsilon > 0$ , when  $n$  is sufficiently large,

$$(1 - \varepsilon)a_{1,n} \leq a_{i,n} \leq (1 + \varepsilon)a_{1,n}, \quad i = 1, \dots, n, \tag{19.3.7}$$

which imply that

$$\begin{aligned} -n \mathbb{P}\{\cup_{i=1}^d \{Y_i > a_{1,n} x_i\}\} &\geq -n \mathbb{P}\{\cup_{i=1}^d \{Y_i > a_{i,n}(1+\varepsilon)^{-1} x_i\}\} \\ &\rightarrow \log G((1 + \varepsilon)^{-1} \mathbf{x}). \end{aligned}$$

Observing that  $\log G(\mathbf{x})$  is homogeneous of order  $-\alpha$ , we have

$$\liminf_{n \rightarrow \infty} [-n \mathbb{P}\{\cup_{i=1}^d \{Y_i > a_{1,n} x_i\}\}] \geq (1 + \varepsilon)^\alpha \log G(\mathbf{x}).$$

Similarly,

$$\limsup_{n \rightarrow \infty} \left[ -n\mathbf{P} \left\{ \cup_{i=1}^d \{Y_i > a_{1,n}x_i\} \right\} \right] \leq (1 - \varepsilon)^\alpha \log G(\mathbf{x}).$$

Let  $\varepsilon \rightarrow 0$ , we obtain that

$$\lim_{n \rightarrow \infty} \left[ -n\mathbf{P} \left\{ \cup_{i=1}^d \{Y_i > a_{1,n}x_i\} \right\} \right] = \log G(\mathbf{x}). \quad (19.3.8)$$

That is, using the tail equivalence Eq.(19.3.7), we can rewrite the limit Eq.(19.3.6) in the form of Eq.(19.3.8), in which the scaling  $a_{1,n}$  is the same for all margins. Working on  $\mathbf{X}'_n$  in the same way, we also obtain that

$$\lim_{n \rightarrow \infty} \left[ -n\mathbf{P} \left\{ \cup_{i=1}^d \{Y'_i > a'_{1,n}x_i\} \right\} \right] = \log G'(\mathbf{x}), \quad (19.3.9)$$

where  $\mathbf{a}'_n = (a'_{1,n}, \dots, a'_{d,n}) = (\overline{F'_1}^{-1}(1/n), \dots, \overline{F'_d}^{-1}(1/n))$ . Again, since the margins are tail equivalent,  $a'_{1,n}/a_{1,n} \rightarrow 1$  as  $n \rightarrow \infty$ . Using the same idea as that of Eq.(19.3.7), the limit Eq.(19.3.9) is equivalent to

$$\lim_{n \rightarrow \infty} \left[ -n\mathbf{P} \left\{ \cup_{i=1}^d \{Y'_i > a_{1,n}x_i\} \right\} \right] = \log G'(\mathbf{x}). \quad (19.3.10)$$

Since  $\mathbf{X}_n \leq_{\text{tlo}} \mathbf{X}'_n$ , via Eq.(19.3.1), we have

$$\mathbf{P} \left\{ \cup_{i=1}^d \{Y_i > a_{1,n}x_i\} \right\} \geq \mathbf{P} \left\{ \cup_{i=1}^d \{Y'_i > a_{1,n}x_i\} \right\}.$$

It follows from Eqs.(19.3.8) and (19.3.10) that  $G(\mathbf{x}) \leq G'(\mathbf{x})$  for all  $\mathbf{x} \in \mathbb{R}_+^d$ .

- (2) It follows from the Pickands representation [see Remark 19.1.2 (4)] that

$$G(\mathbf{x}) = \exp \left\{ -c \int_{\mathbb{S}_+^d} \max_{1 \leq i \leq d} \{(a_i/x_i)^\alpha\} \mathbb{Q}(\mathbf{d}\mathbf{a}) \right\}, \quad (19.3.11)$$

$$G'(\mathbf{x}) = \exp \left\{ -c' \int_{\mathbb{S}_+^d} \max_{1 \leq i \leq d} \{(a_i/x_i)^\alpha\} \mathbb{Q}'(\mathbf{d}\mathbf{a}) \right\}, \quad (19.3.12)$$

where  $c > 0, c' > 0$ , and  $\mathbb{Q}$  and  $\mathbb{Q}'$  are probability measures defined on  $\mathbb{S}_+^{d-1}$ . Taking the scaling function  $t$  for both dfs, we have

$$G(t\mathbf{x}) = \exp \left\{ -\frac{c}{t^\alpha} \int_{\mathbb{S}_+^d} \max_{1 \leq i \leq d} \{(a_i/x_i)^\alpha\} \mathbb{Q}(\mathbf{d}\mathbf{a}) \right\},$$

$$G'(t\mathbf{x}) = \exp \left\{ -\frac{c'}{t^\alpha} \int_{\mathbb{S}_+^d} \max_{1 \leq i \leq d} \{(a_i/x_i)^\alpha\} Q'(\mathbf{d}\mathbf{a}) \right\}.$$

For each fixed  $\mathbf{x}$ , when  $t$  is sufficiently large,

$$\frac{1 - G(t\mathbf{x})}{1 - G'(t\mathbf{x})} \sim \frac{c \int_{\mathbb{S}_+^d} \max_{1 \leq i \leq d} \{(a_i/x_i)^\alpha\} Q(\mathbf{d}\mathbf{a})}{c' \int_{\mathbb{S}_+^d} \max_{1 \leq i \leq d} \{(a_i/x_i)^\alpha\} Q'(\mathbf{d}\mathbf{a})}.$$

If  $G \leq_{\text{t lod}} G'$ , then  $1 - G(t\mathbf{x}) \geq 1 - G'(t\mathbf{x})$  for  $t > t_{\mathbf{x}}$ , where  $t_{\mathbf{x}}$  is sufficiently large. That is,

$$c \int_{\mathbb{S}_+^d} \max_{1 \leq i \leq d} \{(a_i/x_i)^\alpha\} Q(\mathbf{d}\mathbf{a}) \geq c' \int_{\mathbb{S}_+^d} \max_{1 \leq i \leq d} \{(a_i/x_i)^\alpha\} Q'(\mathbf{d}\mathbf{a}),$$

which, together with Eqs.(19.3.11) and (19.3.12), imply that  $G(\mathbf{x}) \leq G'(\mathbf{x})$  for all  $\mathbf{x}$ .

Conversely, it is trivial that  $G \leq_{\text{lod}} G'$  implies that  $G \leq_{\text{t lod}} G'$ . ■

Using similar arguments, we can also establish the upper orthant dependence comparisons for MEV distributions.

**Theorem 19.3.12.** *Let  $(\mathbf{X}_n, n \geq 1)$  and  $(\mathbf{X}'_n, n \geq 1)$  be two i.i.d. samples with dfs  $F$  and  $F'$  respectively. Assume that  $F \in DA_\wedge(H)$  and  $F' \in DA_\wedge(H')$  with tail equivalent, negative Fréchet margins (i.e.,  $F_i(x) = F'_i(x) = 1 - \exp\{-(-x)^{-\theta}\}$ ).*

1. *If  $X_n \leq_{\text{tuod}} X'_n$ , then  $H \leq_{\text{uod}} H'$ .*
2.  *$H \leq_{\text{uod}} H'$  if and only if  $H \leq_{\text{tuod}} H'$ .*

As illustrated in Theorems 19.3.11 (2) and 19.3.12 (2), upper tail comparisons of MEV dfs trickle down to comparisons of entire distributions due to the scalable property of homogeneity.

We conclude this paper with an example to illustrate an idea for obtaining asymptotic Fréchet bounds.

**Example 19.3.13.** Let  $(T_1, T_2)$  denote a random vector with a Marshall-Olkin distribution [307] as defined as follows:

$$T_1 = E_1 \wedge E_{12}, \quad T_2 = E_2 \wedge E_{12},$$

where  $E_1, E_2, E_{12}$  are i.i.d. exponentially distributed with unit mean. Clearly, for  $t_1 > 0$ , or  $t_2 > 0$ ,



$$P\{T_1 > t_1, T_2 > t_2\} = e^{-(t_1+t_2)-t_1 \vee t_2} < e^{-t_1 \vee t_2} = P\{T_1 > t_1, T_1 > t_2\}.$$

Hence  $(T_1, T_2) \leq_{\text{uod}} (T_1, T_1)$ , which is known as the Fréchet upper bound for the class of dfs with fixed exponential margins.

Let  $R_1$  and  $R_2$  denote two nonnegative random variables that are independent of  $(T_1, T_2)$ . Assume that the survival functions of  $R_1^{-1}$  and  $R_2^{-1}$  are tail equivalent, regularly varying with tail index  $-\alpha$  [see Eq. (19.1.6)]. Since  $R_1^{-1}$  and  $R_2^{-1}$  are tail equivalent, the margins of  $(R_2^{-1}T_1, R_2^{-1}T_2)$  and  $(R_1^{-1}T_1, R_1^{-1}T_1)$  are all tail equivalent. It follows from Theorem 3.2 of [280] that the upper tail dependence functions of  $(R_2^{-1}T_1, R_2^{-1}T_2)$  and  $(R_1^{-1}T_1, R_1^{-1}T_1)$  are given by

$$\begin{aligned} b(w_1, w_2) &= 2^\alpha(w_1 + w_2 + w_1 \vee w_2)^{-\alpha} \\ b'(w_1, w_2) &= 2^\alpha(w_1 \vee w_2)^{-\alpha}. \end{aligned}$$

Clearly,  $b(w_1, w_2) < b'(w_1, w_2)$ , and thus by Theorem 19.3.6 we have  $(R_2^{-1}T_1, R_2^{-1}T_2) \leq_{\text{tuod}} (R_1^{-1}T_1, R_1^{-1}T_1)$ . Note that the df of  $(R_1^{-1}T_1, R_1^{-1}T_1)$  is viewed as an asymptotic Fréchet upper bound in the sense of tail upper orthant order, because the respective margins of  $(R_2^{-1}T_1, R_2^{-1}T_2)$  and  $(R_1^{-1}T_1, R_1^{-1}T_1)$  are only tail equivalent, rather than being identical as required in the case of Fréchet bounds. ■

**Acknowledgments:** The author would like to thank Alfred Müller for his comments on this paper during IWAP 2012 in Jerusalem, Israel, and especially for his comment on several references that are closely related to this work. Haijun Li is supported by NSF grants CMMI 0825960 and DMS 1007556.