Tail Dependence of Multivariate Pareto Distributions

Haijun Li
Department of Mathematics
and Department of Statistics
Washington State University
Pullman, WA 99164, U.S.A.
lih@math.wsu.edu
June 2006

Abstract

Various multivariate Pareto distributions are known to exhibit the heavy tail behaviors. This paper examines the tail dependence properties of a general class of multivariate Pareto distributions with the Pareto index and some common scale parameters. The multivariate tail dependence describes the amount of dependence in the upper-orthant tail or lower-orthant tail of a multivariate distribution and can be used in the study of dependence among extreme values. We derive the explicit expressions of tail dependencies of the multivariate Pareto distributions and related copulas of Archimedean type. Properties of the tail dependence coefficients are discussed and some examples are presented to illustrate our results.

Key words and phrases: Multivariate Pareto distribution, Marshall-Olkin distribution, tail dependence, heavy tails, copula, Archimedean copula.

1 Introduction

The Pareto distribution and its multivariate versions have been studied extensively in the literature (Arnold 1983), and have been applied to various practical situations in which an equilibrium is found in the distribution of many “small ones” and a few “large ones”, such as the claim size distribution in an insurance portfolio or the length distribution in jobs assigned
to supercomputers. The multivariate Pareto distributions are known to possess power laws in their upper-orthant tails, and this paper focuses on extremal dependence exhibited among multivariate heavy tails of a general class of multivariate Pareto distributions.

There are several basic methods of constructing multivariate Pareto distributions (Arnold 1983, Kotz, Balakrishnan and Johnson 2000), and one of them yields the following stochastic representation for the multivariate Pareto distributions of our interest:

\[
(Y_1, \ldots, Y_n) = \left( \mu_1 + \sigma_1 \left( \frac{T_1}{Z} \right)^{\gamma_1}, \ldots, \mu_n + \sigma_n \left( \frac{T_n}{Z} \right)^{\gamma_n} \right),
\]  

(1.1)

where \(\gamma_i, 1 \leq i \leq n\), is known as the (marginal) Gini index, \(Z\) has a gamma distribution with shape parameter \(a > 0\) (Pareto index) and scale parameter 1, and random vector \((T_1, \ldots, T_n)\), independent of \(Z\), has a certain multivariate exponential distribution. Obviously, the dependence structure of a multivariate Pareto distribution of (1.1) depends not only on the common \(Z\) but also on the dependence structure of \((T_1, \ldots, T_n)\), and various choices of multivariate exponential distributions lead to a versatile class of multivariate Pareto distributions.

The random vector \((T_1, \ldots, T_n)\) to be considered in this paper has the multivariate exponential distribution of Marshall-Olkin type (Marshall and Olkin 1967). The Marshall-Olkin distribution with rate (inverse scale) parameters \(\{\lambda_S, S \subseteq \{1, \ldots, n\}\}\) is the joint distribution of

\[
T_i = \min\{E_S : S \ni i\}, \ i = 1, \ldots, n,
\]

(1.2)

where \(\{E_S, S \subseteq \{1, \ldots, n\}\}\) is a sequence of independent, exponentially distributed random variables, with \(E_S\) having mean \(1/\lambda_S\). In the reliability context, \(T_1, \ldots, T_n\) can be viewed as the lifetimes of \(n\) components operating in a random shock environment where a fatal shock governed by Poisson process \(\{N_S(t), t \geq 0\}\) with rate \(\lambda_S\) destroys all the components with indexes in \(S \subseteq \{1, \ldots, n\}\) simultaneously. In credit-risk modeling, \(T_1, \ldots, T_n\) can be viewed as the times to default of various different counterparties or types of counterparty, for which the Poisson shocks might be a variety of underlying economic events (Embrechts, Lindskog and McNeil 2003).

To characterize the extremal dependence of a multivariate Pareto distribution (1.1) with \((T_1, \ldots, T_m)\) given by (1.2), we utilize the notions of upper-orthant and lower-orthant tail dependence, which can be described in terms of copulas of a multivariate distribution. The copula is a useful tool for handling multivariate distributions with given univariate marginals. Formally, a copula \(C\) is a distribution function, defined on the unit cube \([0, 1]^n\), with uniform one-dimensional marginals. Given a copula \(C\), if one defines

\[
F(t_1, \ldots, t_n) = C(F_1(t_1), \ldots, F_n(t_n)), \ (t_1, \ldots, t_n) \in \mathbb{R}^n,
\]

(1.3)
then $F$ is a multivariate distribution with univariate marginal distributions $F_1, \cdots, F_n$. Given a distribution $F$ of a random vector $(X_1, \cdots, X_n)$ with marginals $F_1, \cdots, F_n$, there exists a copula $C$ such that (1.3) holds. If $F_1, \cdots, F_n$ are all continuous, then the corresponding copula $C$ is unique, and can be written as

$$C(u_1, \cdots, u_n) = F(F_1^{-1}(u_1), \cdots, F_n^{-1}(u_n)), \quad (u_1, \cdots, u_n) \in [0, 1]^n.$$ 

Thus, for continuous multivariate distribution functions, the univariate marginals and multivariate dependence structure can be separated, and the dependence structure can be represented by a copula. The copula was first developed in Sklar (1959), and the copula theory and its applications can be found, for example, in Nelsen (1999).

The survival copula can be defined similarly. Consider a random vector $(X_1, \cdots, X_n)$ with continuous marginals $F_1, \cdots, F_n$ and copula $C$. Observe that $\bar{F}_i(X_i) = 1 - F_i(X_i)$, $1 \leq i \leq n$, is also uniformly distributed over $[0, 1]$, and thus

$$\hat{C}(u_1, \cdots, u_n) \triangleq \Pr\{\bar{F}_1(X_1) \leq u_1, \cdots, \bar{F}_n(X_n) \leq u_n\} \quad (1.4)$$

is a copula, and called the survival copula of $(X_1, \cdots, X_n)$. The survival function of random vector $(X_1, \cdots, X_n)$ can be expressed as

$$\bar{F}(t_1, \cdots, t_n) = \Pr\{X_1 > t_1, \cdots, X_n > t_n\} = \hat{C}(\bar{F}_1(t_1), \cdots, \bar{F}_n(t_n)), \quad (t_1, \cdots, t_n) \in \mathbb{R}^n.$$ 

It also follows that for any $(u_1, \cdots, u_n) \in [0, 1]^n$,

$$\tilde{C}(u_1, \cdots, u_n) \triangleq \Pr\{F_1(X_1) > u_1, \cdots, F_n(X_n) > u_n\} = \hat{C}(1 - u_1, \cdots, 1 - u_n),$$

where $\tilde{C}$ is the joint survival function of copula $C$.

The tail dependence of a bivariate distribution has been discussed extensively in statistics literature (Joe 1997), but the tail dependence of the general case has not been adequately addressed. Schmidt (2002) and Li (2006) proposed the following multivariate extension.

**Definition 1.1.** Let $X = (X_1, \cdots, X_n)$ be a random vector with continuous marginals $F_1, \cdots, F_n$ and copula $C$.

1. $X$ is said to be upper-orthant tail dependent if for some subset $\emptyset \neq J \subset \{1, \cdots, n\}$, the following limit exists and is positive.

   $$\tau^C_J = \lim_{u \uparrow 1} \Pr\{F_j(X_j) > u, \forall j \notin J \mid F_i(X_i) > u, \forall i \in J\} > 0. \quad (1.5)$$

   If for all $\emptyset \neq J \subset \{1, \cdots, n\}$, $\tau^C_J = 0$, then we say $X$ is upper-orthant tail independent.
2. **\( \mathbf{X} \)** is said to be lower-orthant tail dependent if for some subset \( \emptyset \neq J \subset \{1, \cdots, n\} \), the following limit exists and is positive.

\[
\zeta_C J = \lim_{u \downarrow 0} \Pr \{ F_j(X_j) \leq u, \forall j \notin J \mid F_i(X_i) \leq u, \forall i \in J \} > 0.
\]  

(1.6)

If for all \( \emptyset \neq J \subset \{1, \cdots, n\} \), \( \zeta_C J = 0 \), then we say \( \mathbf{X} \) is lower-orthant tail independent.

The limits \( \tau^C_J \)'s (\( \zeta_C J \)'s) are called the upper (lower) tail dependence coefficients. Obviously, the tail dependence is a copula property, and does not depend on the marginal distributions. Since

\[
\Pr \{ F_j(X_j) > u, \forall j \notin J \mid F_i(X_i) > u, \forall i \in J \} = \Pr \{ \bar{F}_j(X_j) \leq 1 - u, \forall j \notin J \mid \bar{F}_i(X_i) \leq 1 - u, \forall i \in J \},
\]

we obtain a duality property for continuous multivariate distributions,

\[
\tau^C_J = \zeta_C J, \quad \text{for all} \; \emptyset \neq J \subset \{1, \cdots, n\}.
\]

(1.7)

Similarly, \( \zeta_C J = \tau^C_J \). That is, the copula \( C \) is upper-orthant (lower-orthant) tail dependent if and only if the survival copula \( \hat{C} \) is lower-orthant (upper-orthant) tail dependent.

It is well-known that the bivariate normal distribution is asymptotically tail independent if its correlation coefficient \( \rho < 1 \). Schmidt (2002) showed that bivariate elliptical distributions possess the tail dependence property if the tail of their generating random variable is regularly varying. Elliptical copulas do not have closed form expressions and are restricted to have radial symmetry (\( C = \hat{C} \)). In engineering and financial applications, there is often a stronger dependence among big losses than among big gains (Embrechts, Lindskog and McNeil 2003). Such asymmetries cannot be modeled with elliptical copulas. In this paper, we discuss the multivariate Pareto distribution (1.1) that are asymmetric and have a closed form survival function. We obtain the explicit expressions of the tail dependence coefficients of (1.1) with \( (T_1, \cdots, T_n) \) distributed according to a Marshall-Olkin distribution (1.2). We also illustrate that the tail dependence of the multivariate Pareto distribution can be decreased not only by increasing the Pareto index \( a \) but also by coordinate the inverse scale parameters \( \lambda_S \)'s in certain fashion without modifying the marginal distributions.

The paper is organized as follows. In Section 2, we derive the explicit expressions of tail dependence of the multivariate Pareto distributions and discuss their properties. Motivated by the copulas of these multivariate Pareto distributions, we discuss in Section 3 a general class of copulas of Archimedean type and their tail dependence. Finally, some comments in Section 4 conclude the paper.
2 Tail Dependence of Multivariate Pareto Distributions

Since tail dependence is a copula property, we only need to consider the Pareto distributed random vector

$$X = (X_1, \ldots, X_n) = \left( \frac{T_1}{Z}, \ldots, \frac{T_n}{Z} \right),$$

(2.1)

where, as in (1.1), $Z$ has a gamma distribution with shape parameter $a > 0$ (Pareto index) and scale parameter 1, and random vector $(T_1, \ldots, T_n)$, independent of $Z$, has a multivariate Marshall-Olkin distribution with rate (inverse scale) parameters $\{\lambda_S, S \subseteq \{1, \ldots, n\}\}$.

It follows from (1.2) that the survival function of $(T_1, \ldots, T_n)$ can be written as

$$\Pr\{T_1 > t_1, \ldots, T_n > t_n\} = \exp\left[ - \sum_{i=1}^{n} \lambda_i t_i - \sum_{i<j} \lambda_{ij} \max\{t_i, t_j\} \ldots - \lambda_{12\ldots n} \max\{t_1, \ldots, t_n\} \right].$$

We introduce the following notations, for any $S \subseteq \{1, \ldots, n\}$,

$$\wedge_{i \in S} t_i = \min\{t_i, i \in S\}, \ \vee_{i \in S} t_i = \max\{t_i, i \in S\}.$$

The survival function of $(T_1, \ldots, T_n)$ can be expressed in a more compact form as

$$\Pr\{T_1 > t_1, \ldots, T_n > t_n\} = \exp\left[ - \sum_{\emptyset \neq I \subseteq \{1, \ldots, n\}} \lambda_I \vee_{i \in I} t_i \right].$$

(2.2)

It follows from (2.2) that the survival function of the $i$-component of Marshall-Olkin vector $(T_1, \ldots, T_n)$ is given by

$$\Pr\{T_i > t_i\} = \exp\left[ - \left( \sum_{i \in I} \lambda_I \right) t_i \right], \ 1 \leq i \leq n.$$

(2.3)

The survival function of $(X_1, \ldots, X_n)$ in (2.1) can be easily derived from (2.2). Consider, for any $x_1 \geq 0, \ldots, x_n \geq 0$,

$$F(x_1, \ldots, x_n) = \Pr\{X_1 > x_1, \ldots, X_n > x_n\}$$

$$= \int \Pr\{T_1 > x_1z, \ldots, T_n > x_nz\} dF_Z(z)$$

$$= E \left\{ \exp\left[ - \sum_{\emptyset \neq I \subseteq \{1, \ldots, n\}} \lambda_I \vee_{i \in I} x_i \right] Z \right\}.$$
Observe that the last expression is the moment generating function of $Z$ which has a gamma distribution, then we obtain that

$$
\bar{F}(x_1, \cdots, x_n) = \left(1 + \sum_{\emptyset \neq I \subseteq \{1, \cdots, n\}} \lambda_I \vee_{i \in I} x_i\right)^{-a}, \quad x_1 \geq 0, \cdots, x_n \geq 0. \quad (2.4)
$$

For any vector $x = (x_1, \cdots, x_n)$ and any $\emptyset \neq J \subseteq \{1, \cdots, n\}$, let $x_J$ denote the $|J|$-dimensional sub-vector of the entries of $x$ with indexes in $J$. Let $\bar{F}_J(x_J)$ denote the multivariate marginal survival function of $X_J$, then we have

$$
\bar{F}_J(x_J) = \left(1 + \sum_{\emptyset \neq I \subseteq \{1, \cdots, n\}} \lambda_I \vee_{i \in I \cap J} x_i\right)^{-a} = \left(1 + \sum_{\emptyset \neq K \subseteq J} \left(\sum_{K = I \cap J} \lambda_I \vee_{i \in K} x_i\right)^{-a}\right). \quad (2.5)
$$

In particular, the marginal survival function of $X_i$ is given by

$$
\bar{F}_i(x_i) = \left(1 + \sum_{i \in I} \lambda_I \right)^{-a} x_i.
$$

The distribution (2.4) includes some well-known special cases. For example, if $\lambda_I = 0$ for all $|I| > 1$, (2.4) becomes

$$
\bar{F}(x_1, \cdots, x_n) = \left(1 + \sum_{i=1}^n \lambda_i x_i\right)^{-a}, \quad x_1 \geq 0, \cdots, x_n \geq 0,
$$

which is known as the multivariate Pareto distribution of type II, or $\text{MP}^{(2)}(0, \{1/\lambda_i\}, a)$ (Arnold 1983).

To derive the tail dependence coefficients of (2.1), we introduce the sums of relative rates:

$$
\alpha_J = \sum_{\emptyset \neq K \subseteq J} \sum_{i \in K} \lambda_{K \cup L} \wedge_{i \in K} \sum_{i \in S} \lambda_S, \quad \emptyset \neq J \subseteq \{1, \cdots, n\}. \quad (2.6)
$$

Here and in the sequel, $J^c$ denotes the complement of set $J$. The sums of relative rates have the following properties.

**Lemma 2.1.** Let $\{\lambda_S, S \subseteq \{1, \cdots, n\}\}$ be a set of rate parameters.

1. $\alpha_i = 1$ for any $1 \leq i \leq n$, and $\alpha_{\{1, \cdots, n\}} = \sum_{\emptyset \neq K \subseteq \{1, \cdots, n\}} \frac{\lambda_K}{\wedge_{i \in K} \sum_{i \in S} \lambda_S}$.
2. If \( \lambda_S = 0 \) for any \(|S| > 1\), then \( \alpha_J = |J| \) for all \( \emptyset \neq J \subseteq \{1, \ldots, n\} \).

3. If \( \lambda_S = 0 \) for any \(|S| < n\), then \( \alpha_J = 1 \) for all \( \emptyset \neq J \subseteq \{1, \ldots, n\} \).

4. For any index \( j_c \notin J \), \( \alpha_J \leq \alpha_{J \cup \{j_c\}} \). In particular, \( \alpha_J \leq \alpha_{\{1, \ldots, n\}} \) for all \( \emptyset \neq J \subseteq \{1, \ldots, n\} \).

Proof. (1) follows easily from (2.6). If \( \lambda_S = 0 \) for any \(|S| > 1\), then

\[
\alpha_J = \sum_{k \in J} \frac{\sum_{L \subseteq J: \{k\} \cup L} \lambda_{(k) \cup L}}{\Lambda_i \sum_{i \in S} \lambda_S} = \sum_{k \in J} \frac{\lambda_{(k)}}{\Lambda(\{k\})} = |J|,
\]

and (2) holds. If \( \lambda_S = 0 \) for any \(|S| < n\), then \( \cap_{i \in K} \sum_{i \in S} \lambda_S = \lambda_{\{1, \ldots, n\}} \). Thus, \( \alpha_J \) in (3) can be simplified to

\[
\alpha_J = \frac{\lambda_{J \cup J_c}}{\lambda_{\{1, \ldots, n\}}} = 1.
\]

To prove (4), consider

\[
\alpha_{J \cup \{j_c\}} = \sum_{\emptyset \neq K \subseteq J \cup \{j_c\}} \frac{\sum_{L \subseteq (J \cup \{j_c\}): \{k\} \cup L} \lambda_{(k) \cup L}}{\Lambda_i \sum_{i \in S} \lambda_S} \\
\geq \sum_{\emptyset \neq K \subseteq J} \frac{\sum_{L \subseteq (J \cup \{j_c\}): \{k\} \cup L} \lambda_{(k) \cup L}}{\Lambda_i \sum_{i \in S} \lambda_S} + \sum_{\emptyset \neq K \subseteq J} \frac{\sum_{L \subseteq (J \cup \{j_c\}): \{k\} \cup L} \lambda_{(k) \cup \{j_c\} \cup L}}{\Lambda_i \sum_{i \in S} \lambda_S} \\
\geq \sum_{\emptyset \neq K \subseteq J} \frac{\sum_{L \subseteq (J \cup \{j_c\}): \{k\} \cup L} \lambda_{(k) \cup L}}{\Lambda_i \sum_{i \in S} \lambda_S} + \sum_{\emptyset \neq K \subseteq J} \frac{\sum_{L \subseteq (J \cup \{j_c\}): \{k\} \cup L} \lambda_{(k) \cup \{j_c\} \cup L}}{\Lambda_i \sum_{i \in S} \lambda_S} \\
= \sum_{\emptyset \neq K \subseteq J} \frac{\sum_{L \subseteq J \cup \{j_c\} \cup L} \lambda_{\{k\} \cup L}}{\Lambda_i \sum_{i \in S} \lambda_S} = \alpha_J.
\]

(4) follows as claimed. \( \square \)

**Theorem 2.2.** Let \((X_1, \ldots, X_n)\) have a multivariate Pareto distribution with the Pareto index \(a\) and inverse scale parameters \(\{\lambda_S, S \subseteq \{1, \ldots, n\}\}\).

1. The upper-orthant tail dependence coefficients are given by

\[
\tau_J = \left( \frac{\alpha_J}{\alpha_{\{1, \ldots, n\}}} \right)^a, \quad \emptyset \neq J \subseteq \{1, \ldots, n\}. \quad (2.7)
\]

2. The lower-orthant tail dependence coefficients are given by

\[
\zeta_J = \frac{\sum_{\emptyset \neq I \subseteq \{1, \ldots, n\}} (-1)^{|I|-1} \alpha_I}{\sum_{\emptyset \neq I \subseteq J} (-1)^{|I|-1} \alpha_I}, \quad (2.8)
\]

provided that \(\sum_{\emptyset \neq I \subseteq J} (-1)^{|I|-1} \alpha_I \neq 0\).
Proof. (1) First, it follows from Lemma 2.1 (4) that the right-hand side expression of (2.7) is always less than or equal to 1. Set

\[ \bar{F}_i(x_i) = \left(1 + \left(\sum_{i \in S} \lambda_S\right)x_i\right)^{-a} = 1 - u_i, \]

and we obtain that

\[ F_i^{-1}(u_i) = \frac{(1 - u_i)^{-\frac{1}{a}} - 1}{\sum_{i \in S} \lambda_S}, \quad 0 \leq u_i \leq 1. \]

It follows from (2.4) that the survival function of the multivariate Pareto copula \( C \) is given by

\[ \bar{C}(u_1, \cdots, u_n) = \left(1 + \sum_{\emptyset \neq I \subseteq \{1, \cdots, n\}} \lambda_I \sqrt{\frac{(1 - u_I)^{-\frac{1}{a}} - 1}{\sum_{i \in S} \lambda_S}}\right)^{-a}, \quad 0 \leq u_i \leq 1, 1 \leq i \leq n. \tag{2.9} \]

Hence

\[ \bar{C}(u, \cdots, u) = \left(1 + \sum_{\emptyset \neq K \subseteq \{1, \cdots, n\}} \lambda_K \left(\frac{(1 - u)^{-\frac{1}{a}} - 1}{\sum_{i \in S} \lambda_S}\right)^{-a} \right)^{-a}, \quad 0 \leq u \leq 1. \]

Similarly, it follows from (2.5) that

\[ \bar{C}_J(u, \cdots, u) = \left(1 + \sum_{\emptyset \neq K \subseteq J} \left(\sum_{K = I \cap J} \lambda_I \frac{(1 - u)^{-\frac{1}{a}} - 1}{\sum_{i \in S} \lambda_S}\right)^{-a} \right)^{-a}, \quad 0 \leq u \leq 1. \tag{2.10} \]

Notice that \( u \to 1 \) if and only if \( (1 - u)^{-\frac{1}{a}} - 1 \to \infty \). From Definition 1.1 (1), we can now obtain the upper-orthant tail dependence.

\[ \tau_J = \lim_{u \uparrow 1} \frac{\bar{C}(u, \cdots, u)}{\bar{C}_J(u, \cdots, u)} = \lim_{u \uparrow 1} \left(\frac{1}{(1-u)^{-\frac{1}{a}}-1} + \frac{\alpha_{\{1, \cdots, n\}}}{\alpha_J}\right)^{-a} = \left(\frac{\alpha_J}{\alpha_{\{1, \cdots, n\}}}\right)^a. \]

(2) Since \( C(u_1, u_2, \cdots, u_n) = 1 + \sum_{\emptyset \neq I \subseteq \{1, \cdots, n\}} (-1)^{|I|} \bar{C}_I(u_I), \quad 0 \leq u_i \leq 1, 1 \leq i \leq n, \) we have

\[ C(u, u, \cdots, u) = 1 + \sum_{\emptyset \neq I \subseteq \{1, \cdots, n\}} (-1)^{|I|} \left(1 + \alpha_I [(1 - u)^{-\frac{1}{a}} - 1]\right)^{-a}. \tag{2.10} \]
Similarly, we have

\[
C_J(u, \cdots, u) = 1 + \sum_{\emptyset \neq I \subseteq J} (-1)^{|I|} \left( 1 + \alpha_I [(1 - u)^{-\frac{a}{J}} - 1] \right)^{-a}.
\] (2.11)

Some direct calculations yield the following derivatives:

\[
\left[ \frac{dC(u, u, \cdots, u)}{du} \right]_{u=0} = \sum_{\emptyset \neq I \subseteq \{1, \cdots, n\}} (-1)^{|I|-1} \alpha_I
\]

\[
\left[ \frac{dC_J(u, \cdots, u)}{du} \right]_{u=0} = \sum_{\emptyset \neq I \subseteq J} (-1)^{|I|-1} \alpha_I.
\]

Suppose that \( \sum_{\emptyset \neq I \subseteq J} (-1)^{|I|-1} \alpha_I \neq 0 \). It then follows from L’Hospital’s rule that

\[
\zeta_J = \lim_{u \to 0} \frac{C(u, \cdots, u)}{C_J(u, \cdots, u)} = \left[ \frac{dC(u, u, \cdots, u)}{du} \right]_{u=0} \left[ \frac{dC_J(u, \cdots, u)}{du} \right]_{u=0}^{-1}
\]

\[
= \frac{\sum_{\emptyset \neq I \subseteq \{1, \cdots, n\}} (-1)^{|I|-1} \alpha_I}{\sum_{\emptyset \neq I \subseteq J} (-1)^{|I|-1} \alpha_I},
\]

for any \( \emptyset \neq J \subseteq \{1, \cdots, n\} \). \( \square \)

If \( \sum_{\emptyset \neq I \subseteq J} (-1)^{|I|-1} \alpha_I = 0 \), such as in the case of the multivariate Pareto distributions of type II, then the lower-orthant tail dependence can be obtained by successive use of L’Hospital’s rule for the ratio of (2.10) to (2.11).

Any strictly increasing (marginal) transformations of a random vector preserve its copula, and thus the distribution of (1.1) also has the copula (2.10). Since tail dependence is a copula property, the tail dependence coefficients of (1.1) are also given by (2.7) and (2.8), which do not depend on the marginal parameters, such as the Gini indexes \( \gamma_i \)’s.

Li (2006) showed that the upper-orthant tail dependence coefficient of a Marshall-Olkin distributed random vector \( (T_1, \cdots, T_n) \) is either zero or one. That is, the Marshall-Olkin distribution is either perfectly upper-tail dependent or upper-tail independent. In contrast, the multivariate Pareto distributions demonstrate a variety of upper-tail dependence, that depends not only on the Pareto index \( a \) but also on the common rate parameters. This is quite reasonable because the multivariate Pareto distribution (1.1) is obtained from multiplying a Marshall-Olkin distributed vector by an inverse gamma distributed random variable \( 1/Z \), which is known to have a heavy tail distribution. This heavy tail phenomenon slows down exponential decay of the dependence of the Marshall-Olkin distribution.

**Example 2.3.** The following special cases follow from Theorem 2.2 immediately.
1. If \( \lambda_S = 0 \) for any \(|S| < n\), then, by Lemma 2.1 (3), \( \alpha_J = 1 \) for all \( \emptyset \neq J \subseteq \{1, \ldots, n\} \). Thus, \( \tau_J = 1 \) and \( \zeta_J = 1 \) for all \( \emptyset \neq J \subseteq \{1, \ldots, n\} \). Note that in this case, \( X_1 = \cdots = X_n \) in (2.1) almost surely.

2. If \( \lambda_S = 0 \) for any \(|S| > 1\), we have a multivariate Pareto distribution of type II. By Lemma 2.1 (2), \( \alpha_J = |J| \), and thus \( \tau_J = \left( \frac{|J|}{n} \right)^a \) for all \( \emptyset \neq J \subseteq \{1, \ldots, n\} \). Observe that for any \( n \geq 2 \), we have
   \[
   \sum_{\emptyset \neq I \subseteq \{1, \ldots, n\}} (-1)^{|I|-1} \alpha_I = \sum_{k=1}^{n} (-1)^{k-1} k \binom{n}{k} = n \sum_{k=0}^{n-1} (-1)^{k} \binom{n-1}{k} = 0.
   \]
   Thus, \( \zeta_J = 0 \) if \( |J| = 1 \) and \( n \geq 2 \). Note that if \( |J| > 1 \), then \( \zeta_J \) cannot be determined by (2.8) because of \( \sum_{\emptyset \neq I \subseteq J} (-1)^{|I|-1} \alpha_I \) is also zero in this case.

3. Consider the case of equal marginals in which \( \sum_{i \in S} \lambda_S = \sum_{j \in S} \lambda_S \) for any \( i \neq j \). Then
   \[
   \tau_J = \left( \frac{\Lambda - \sum_{I \subseteq \emptyset \neq I \subseteq \{1, \ldots, n\}} \lambda_I}{\Lambda} \right)^a, \tag{2.12}
   \]
   where \( \Lambda = \sum_{\emptyset \neq I \subseteq \{1, \ldots, n\}} \lambda_I \). If, furthermore, \( \lambda_S = \lambda \) for all \( S \subseteq \{1, \ldots, n\} \), then in this case,
   \[
   \tau_J = \left( \frac{2^n - 2|J|}{2^n - 1} \right)^a, \quad \zeta_J = \frac{2^n + \sum_{\emptyset \neq I \subseteq \{1, \ldots, n\}} (-1)^{|I|} 2^{|I|}}{2^n + \sum_{\emptyset \neq I \subseteq J} (-1)^{|I|} 2^{|I|}}.
   \]
   For example, when \( n = 3 \) and \( J = \{1, 2\} \), we have \( \tau_J = \left( \frac{6}{7} \right)^a \) and \( \zeta_J = \frac{1}{2} \). \( \square \)

It follows from Theorem 2.2 that the upper-orthant tail dependence coefficients of a multivariate Pareto distribution is non-increasing as the Pareto index \( a \) is increasing. The tail dependence can be also decreased by properly adjusting rate parameters.

**Example 2.4.** Let \((X_1, X_2, X_3)\) be distributed according to (2.1) where \((T_1, T_2, T_3)\) has a Marshall-Olkin distribution with rate parameters \( \{\lambda_S, S \subseteq \{1, 2, 3\}\} \). Consider the following three cases.

1. \( \lambda_S = 1 \) for all \( S \subseteq \{1, 2, 3\} \).
2. \( \lambda_{\{1,2,3\}} = \lambda_{\{3\}} = 2, \lambda_{\{1,3\}} = \lambda_{\{2,3\}} = 0, \) and \( \lambda_S = 1 \) for all other \( S \).
3. \( \lambda_{\{2,3\}} = \lambda_{\emptyset} = 2, \lambda_{\{2\}} = \lambda_{\{3\}} = 0, \) and \( \lambda_S = 1 \) for all other \( S \).
It is easy to verify that \( \sum_{i \in S} \lambda_S = 4 \) for all \( i = 1, 2, 3 \), and thus the marginal distributions are the same for all three cases. It follows from a result in Li and Xu (2000) that \((T_1, T_2, T_3)\) in Cases 2 and 3 are more dependent than \((T_1, T_2, T_3)\) in Case 1 in the sense of supermodular order. Thus, \((X_1, X_2, X_3)\) in Cases 2 and 3 are also more dependent than \((X_1, X_2, X_3)\) in Case 1 in the sense of supermodular order. Let \( \tau_{(1,2)}^i \) denote the upper-orthant tail dependence coefficient of Case \( i, i = 1, 2, 3 \). It follows from (2.12) that

\[
\tau_{(1,2)}^1 = \left( \frac{6}{7} \right)^a, \tau_{(1,2)}^2 = \left( \frac{5}{7} \right)^a, \tau_{(1,2)}^3 = 1
\]

Note that \( \tau_{(1,2)}^1 > \tau_{(1,2)}^2 \), even though Case 2 is more dependent than Case 1.

\[
\begin{align*}
\text{3 Copulas of Archimedean Type}
\end{align*}
\]

The results from Theorem 2.2 and Example 2.3 indicate that the expressions (2.7) and (2.8) can be simplified considerably if random variables \( T_1, \cdots, T_n \) have the same marginal distribution. This motivates us to concentrate in this section on the class of multivariate Pareto distributed random vectors (2.1) where \((T_1, \cdots, T_n)\) is distributed according to a Marshall-Olkin distribution with standard exponential univariate marginals. That is, the marginal rates satisfy

\[
\sum_{i \in S} \lambda_S = 1, \text{ for all } i = 1, \cdots, n.
\]  \( (3.1) \)

Note that the Marshall-Olkin distribution with unequal marginals cannot be converted into the one with equal marginals under marginal scalings, the distributions of (2.1) with (3.1) constitute only a special class of the distributions studied in Section 2. The survival functions of the copulas for those multivariate Pareto distributions with (3.1) are simplified from (2.9) to the following:

\[
\tilde{C}(u_1, \cdots, u_n) = \left( 1 + \sum_{\emptyset \neq I \subseteq \{1, \cdots, n\}} \lambda_I \vee_{i \in I} ((1 - u_i)^{-\frac{1}{a}} - 1) \right)^{-a}, \quad 0 \leq u_i \leq 1, 1 \leq i \leq n.
\]

Thus their corresponding survival copulas are given by

\[
\tilde{C}(u_1, \cdots, u_n) = \left( 1 + \sum_{\emptyset \neq I \subseteq \{1, \cdots, n\}} \lambda_I \vee_{i \in I} (u_i^{-\frac{1}{a}} - 1) \right)^{-a}, \quad 0 \leq u_i \leq 1, 1 \leq i \leq n. \quad (3.2)
\]

These copulas yield a versatile class of the copulas of Archimedean type. Let \( \phi(u) = u^{-\frac{1}{a}} - 1 \),
and thus $\phi^{-1}(x) = (1 + x)^{-a}$ and (3.2) becomes:

$$
\hat{C}(u_1, \cdots, u_n) = \phi^{-1} \left( \sum_{\emptyset \neq I \subseteq \{1, \cdots, n\}} \lambda_I \bigvee_{i \in I} \phi(u_i) \right).
$$

(3.3)

Note that $\phi^{-1}$ is completely monotone over $[0, \infty)$; that is, $(-1)^k \frac{d^k}{dx^k} \phi^{-1}(x) \geq 0$ for all $k = 0, 1, 2, \cdots$. In fact, even if $\phi^{-1}$ is an arbitrary completely monotone function, then the expression of (3.3) is still a copula.

**Proposition 3.1.** Let $\phi$ is a strictly decreasing function from $[0, 1]$ to $[0, \infty)$ with continuous derivative and $\phi(1) = 0$, and $\{\lambda_S, S \subseteq \{1, \cdots, n\}\}$ is a family of rate parameters with $\sum_{i \in S} \lambda_S = 1$ for all $1 \leq i \leq n$. If $\phi^{-1}$ is completely monotone, then

$$
C^{\phi, \{\lambda_S\}}(u_1, \cdots, u_n) = \phi^{-1} \left( \sum_{\emptyset \neq I \subseteq \{1, \cdots, n\}} \lambda_I \bigvee_{i \in I} \phi(u_i) \right)
$$

(3.4)

is a copula.

**Proof.** First, observe that for any $u_k$,

$$
C^{\phi, \{\lambda_S\}}(1, \cdots, 1, u_k, 1, \cdots, 1) = \phi^{-1} \left( \sum_{k \in I} \lambda_I \phi(u_k) \right) = \phi^{-1}(\phi(u_k)) = u_k,
$$

and in particular, $C^{\phi, \{\lambda_S\}}(1, \cdots, 1) = 1$. If $u_k = 0$, then

$$
C^{\phi, \{\lambda_S\}}(1, \cdots, u_k, \cdots, 1) \leq C^{\phi, \{\lambda_S\}}(1, \cdots, 1, u_k, \cdots, 1) = 0,
$$

since $C^{\phi, \{\lambda_S\}}(u_1, \cdots, u_n)$ is non-decreasing with respect to the component-wise order of $(u_1, \cdots, u_n)$. To show that $C^{\phi, \{\lambda_S\}}(u_1, \cdots, u_n)$ is $n$-increasing, let $[I] = \min\{i : i \in I\}$ for any $\emptyset \neq I \subseteq \{1, \cdots, n\}$. For any $u_1 < u_2 < \cdots < u_n$,

$$
\frac{\partial}{\partial u_k} \left( \sum_{\emptyset \neq I \subseteq \{1, \cdots, n\}} \lambda_I \bigvee_{i \in I} \phi(u_i) \right) = \frac{\partial}{\partial u_k} \sum_{\emptyset \neq I \subseteq \{1, \cdots, n\}} \lambda_I \phi(u_{[I]}) = \sum_{k=[I]} \lambda_I \frac{\partial}{\partial u_k} \phi(u_k)
$$

is a non-positive function of $u_k$ only. Therefore,

$$
\frac{\partial^n C^{\phi, \{\lambda_S\}}(u_1, \cdots, u_n)}{\partial u_1 \cdots \partial u_n} = \left( \frac{d^n \phi^{-1}(x)}{dx^n} \right)_{x=\sum_{\emptyset \neq I \subseteq \{1, \cdots, n\}} \lambda_I \bigvee_{i \in I} \phi(u_i)} \prod_{k=1}^{n} \sum_{\emptyset \neq I \subseteq \{1, \cdots, n\}} \lambda_I \frac{\partial}{\partial u_k} \phi(u_k)
$$

is a continuous, non-negative function for all $u_1 < u_2 < \cdots < u_n$. Note that this derivative function may not be defined at the boundary points where $u_{i_1} = u_{i_2} = \cdots = u_{i_t}$ for some
1 ≤ i_1 < i_2 < \cdots < i_l ≤ n. But, by taking limits, the limits of this derivative at the boundary points are still non-negative. Thus, the derivative is non-negative almost everywhere, which implies that \( C^{\phi,\{\lambda_S\}}(u_1, \ldots, u_n) \) is n-increasing. Thus, by the definition of copulas, \( C^{\phi,\{\lambda_S\}}(u_1, \ldots, u_n) \) is a copula.

The copula (3.4) is called a copula of Archimedean type with generator \( \phi^{-1} \) and rate parameters \( \{\lambda_S, S \subseteq \{1, \ldots, n\}\} \). If \( \lambda_S = 0 \) for any \(|S| > 1\), then \( C^{\phi,\{\lambda_S\}}(u_1, \ldots, u_n) = \phi^{-1}(\sum_{i=1}^{n} \phi(u_i)) \), which is a standard Archimedean copula with generator \( \phi^{-1} \) (Nelson 1999).

If, moreover, \( \phi^{-1}(x) = (1 + x)^{-a} \), then we have a Clayton copula family.

To discuss the tail dependence of Archimedean copulas, certain conditions on \( \phi^{-1} \) are usually needed. A function \( g : [0, \infty) \rightarrow [0, \infty) \) is called regularly varying at \( \infty \) with index \( a \in \mathbb{R} \) if for any \( c > 0 \),

\[
\lim_{x \to \infty} \frac{g(cx)}{g(x)} = c^a.
\]

Regularly varying functions behave asymptotically like power functions (Resnick 1987). For example, \( \phi^{-1}(x) = (1 + x)^{-a} \) is a regularly varying function with index \(-a\).

**Theorem 3.2.** Let \( C^{\phi,\{\lambda_S\}}(u_1, \ldots, u_n) \) be an Archimedean copula (3.4) as described in Proposition 3.1. If \( \phi(0) = \infty \), and generator \( \phi^{-1} \) is regularly varying at \( \infty \) with index \(-a\) where \( a \geq 0 \), then its lower-orthant tail dependence coefficients are given by

\[
\zeta^{C^{\phi,\{\lambda_S\}}}_J = \left( \frac{\sum_{\emptyset \neq K \subseteq J} \sum_{L \subseteq J^c} \lambda_{K \cup L}}{\sum_{\emptyset \neq K \subseteq \{1, \ldots, n\}} \lambda_K} \right)^a.
\]  

\( (3.5) \)

**Proof.** We have for any \( 0 \leq u \leq 1 \),

\[
C^{\phi,\{\lambda_S\}}(u, \ldots, u) = \phi^{-1}\left( \sum_{\emptyset \neq K \subseteq \{1, \ldots, n\}} \lambda_K \phi(u) \right)
\]

\[
C^{\phi,\{\lambda_S\}}_J(u, \ldots, u) = \phi^{-1}\left( \sum_{\emptyset \neq K \subseteq J} \sum_{L \subseteq J^c} \lambda_{K \cup L} \phi(u) \right)
\]

As \( u \downarrow 0 \), \( \phi(u) \uparrow \infty \). Since \( \phi^{-1} \) is regularly varying at \( \infty \) with index \(-a \leq 0\), we have

\[
\zeta^{C^{\phi,\{\lambda_S\}}}_J = \lim_{u \downarrow 0} \frac{C^{\phi,\{\lambda_S\}}_J(u, \ldots, u)}{C^{\phi,\{\lambda_S\}}(u, \ldots, u)} = \left( \frac{\sum_{\emptyset \neq K \subseteq J} \sum_{L \subseteq J^c} \lambda_{K \cup L}}{\sum_{\emptyset \neq K \subseteq \{1, \ldots, n\}} \lambda_K} \right)^a,
\]

provided that \( \sum_{\emptyset \neq K \subseteq J} \sum_{L \subseteq J^c} \lambda_{K \cup L} > 0 \). \( \square \)
Note that if \( \phi^{-1}(x) = (1 + x)^{-a} \), then it follows from (1.7) that \( \zeta^a_J \) in (3.5) is the upper-orthant tail dependence coefficient given in (2.7). For the upper-orthant tail dependence, we only discuss the standard Archimedean copula

\[
C^\phi(u_1, \cdots, u_n) = \phi^{-1} \left( \sum_{i=1}^{n} \phi(u_i) \right),
\]

with generator \( \phi^{-1} \), where \( \phi(1) = 0 \) and \( \phi(0) = \infty \).

**Theorem 3.3.** Let \( C^\phi(u_1, \cdots, u_n) \) be an Archimedean copula (3.6). If the \( k \)-th order derivative of \( \phi^{-1} \) at 0 is finite and non-zero for any \( 1 \leq k \leq n - 1 \), then all its upper-orthant tail dependence coefficients are zero.

**Proof.** Observe that \( \tau^\phi_J \leq \tau^\phi_{J'} \) for any \( J \subseteq J' \), and thus, by rearranging the indexes, we only need to show that \( \tau^\phi_{\{1, \ldots, n-1\}} = 0 \). Since \( C^\phi(u, \cdots, u) = \phi^{-1}(|J|\phi(u)) \), we have

\[
\tilde{C}^\phi(u, \cdots, u) = 1 + \sum_{k=1}^{n} (-1)^k \binom{n}{k} \phi^{-1}(k\phi(u))
\]

and

\[
\tilde{C}^\phi_{\{1, \ldots, n-1\}}(u, \cdots, u) = 1 + \sum_{k=1}^{n-1} (-1)^k \binom{n-1}{k} \phi^{-1}(k\phi(u)).
\]

Let \( \delta = \phi(u) \) and \( g = \phi^{-1} \). Notice that as \( u \to 1 \), \( \delta \to 0 \). Since \( \frac{d\phi^{-1}(k\phi(u))}{du} = \frac{kg'(k\delta)}{g(\delta)} \), we have

\[
\frac{d}{du} \tilde{C}^\phi(u, \cdots, u) = \sum_{k=1}^{n} (-1)^k \binom{n}{k} \frac{kg'(k\delta)}{g(\delta)}
\]

and

\[
= \frac{n}{g(\delta)} \sum_{k=1}^{n} (-1)^k \binom{n-1}{k-1} g'(k\delta).
\]

Similarly,

\[
\frac{d}{du} \tilde{C}^\phi_{\{1, \ldots, n-1\}}(u, \cdots, u) = \frac{n-1}{g(\delta)} \sum_{k=1}^{n-1} (-1)^k \binom{n-2}{k-1} g'(k\delta).
\]

By L’hospital’s rule, the upper-orthant tail dependence can be expressed as follows

\[
\tau^\phi_{\{1, \ldots, n-1\}} = \lim_{u \to 1} \frac{d}{du} \tilde{C}^\phi(u, \cdots, u) = \frac{n}{n-1} \lim_{\delta \to 0} \frac{\sum_{k=1}^{n} (-1)^k \binom{n-1}{k-1} g'(k\delta)}{\sum_{k=1}^{n-1} (-1)^k \binom{n-2}{k-1} g'(k\delta)}.
\]
Since \( \binom{n-1}{k-1} = \binom{n-2}{k-1} + \binom{n-2}{k-2} \), we have

\[
\sum_{k=1}^{n} (-1)^k \binom{n-1}{k-1} g'(k\delta) = \sum_{k=1}^{n-1} (-1)^k \binom{n-2}{k-1} g'(k\delta) + \sum_{k=2}^{n} (-1)^k \binom{n-2}{k-2} g'(k\delta)
\]

and thus

\[
\tau_{\{1, \ldots, n-1\}} = \frac{n}{n-1} (1 - G_n' g')
\]

where

\[
G_n' = \lim_{\delta \downarrow 0} \frac{\sum_{k=1}^{n-1} (-1)^k \binom{n-2}{k-1} g'((k+1)\delta)}{\sum_{k=1}^{n-1} (-1)^k \binom{n-2}{k-1} g'(k\delta)}.
\]

Again since \( \binom{n-2}{k-1} = \binom{n-3}{k-1} + \binom{n-3}{k-2} \), we have

\[
\sum_{k=1}^{n-1} (-1)^k \binom{n-2}{k-1} g'((k+1)\delta)
\]

\[
= \sum_{k=2}^{n-2} (-1)^k \binom{n-3}{k-2} g'((k+1)\delta) + \sum_{k=1}^{n-2} (-1)^k \binom{n-3}{k-1} g'((k+1)\delta)
\]

\[
= -\sum_{k=1}^{n-2} (-1)^k \binom{n-3}{k-1} g'((k+2)\delta) + \sum_{k=1}^{n-2} (-1)^k \binom{n-3}{k-1} g'((k+1)\delta)
\]

\[
= -\sum_{k=1}^{n-2} (-1)^k \binom{n-3}{k-1} \left( g'((k+2)\delta) - g'((k+1)\delta) \right).
\]

Similarly,

\[
\sum_{k=1}^{n-1} (-1)^k \binom{n-2}{k-1} g'(k\delta) = -\sum_{k=1}^{n-2} (-1)^k \binom{n-3}{k-1} \left( g'((k+1)\delta) - g'(k\delta) \right).
\]
Since $g'(x)$ is continuous and differentiable, then

$$G_n^{g'} = \lim_{\delta \to 0} \frac{\sum_{k=1}^{n-2} (-1)^k \binom{n-3}{k-1} (g'((k+2)\delta) - g'((k+1)\delta))}{\sum_{k=1}^{n-2} (-1)^k \binom{n-3}{k-1} (g'((k+1)\delta) - g'(k\delta))}$$

$$= \lim_{\delta \to 0} \lim_{\delta' \to 0} \frac{\sum_{k=1}^{n-2} (-1)^k \binom{n-3}{k-1} \left( g'((k+1)\delta+\delta') - g'((k+1)\delta) \right)}{\sum_{k=1}^{n-2} (-1)^k \binom{n-3}{k-1} \left( g''(k\delta) \right)} = G_{n-1}^{g''}. \tag{3.7}$$

Use (3.7) successively, we obtain that $G_n^{g'} = G_2^{g[n-1]}$, where $g[n-1] = \frac{d^{n-1} g}{dx^{n-1}}$. Thus,

$$G_n^{g'} = \lim_{\delta \to 0} \frac{-g[n-1](2\delta)}{-g[n-1](\delta)} = 1,$$

which implies that $\tau_{(1,\ldots,n-1)}^{G_{\phi}} = 0$. \(\square\)

The duality (1.7) and Theorem 3.3 imply that the lower-orthant tail dependence $\zeta_J$ in Example 2.3 (2) is zero for any $J \subset \{1, \ldots, n\}$. It now follows from Theorem 3.3 that if an Archimedean copula is upper-tail dependent, then this can only occur when the derivatives of generator $\phi^{-1}$ at 0 are zero or do not exist. For example, for a bivariate Gumbel copula with parameter $\theta > 1$, $\phi^{-1}(x) = \exp(-x^{1/\theta})$, and thus $\phi^{-1}(x) = -x^{1/\theta-1} \exp(-x^{1/\theta})/\theta$ is infinity at 0. From the direct calculation, the upper-tail dependence coefficient of the Gumbel copula is $2 - 2^{1/\theta}$. This example and the tail dependence of bivariate Archimedean copulas have been discussed in Embrechts, Lindskog and McNeil (2003).

4 Concluding Remarks

A multivariate Pareto distribution is a scale mixture of multivariate exponential distributions with heavy tail mixing random variable, and such a heavy tail property thickens the exponential tails, that yields the heavy tail behavior of the multivariate Pareto distribution.
Tail dependence emerges from the multivariate heavy tail phenomenon, and thus the multivariate Pareto distributions are among the distributions which possess the multivariate tail dependence.

We have derived the explicit expressions of tail dependence coefficients of a general class of the multivariate Pareto distributions and related copulas, and our results demonstrate that the tail dependence depends not only on the Pareto index but also on common scale parameters. The tail dependence of a multivariate Pareto distribution can be adjusted by properly coordinate the scale parameters without modifying the marginal distributions. These tail dependence coefficients can be easily calculated or estimated from data.

References


