EXTREME BEHAVIOR OF
MULTIVARIATE PHASE-TYPE DISTRIBUTIONS

Alexandru V. Asimit[1]

Bruce L. Jones[2]

Department of Statistical and Actuarial Sciences
University of Western Ontario
London, Ontario
Canada N6A 5B7

February 16, 2006

Abstract

This paper investigates the limiting distributions of the componentwise maxima and minima of suitably normalized iid multivariate phase-type random vectors. In the case of maxima, a large parametric class of multivariate extreme value (MEV) distributions is obtained. The flexibility of this new class is exemplified in the bivariate setup. For minima, it is shown that the dependence structure of the Marshall-Olkin class arises in the limit.

Keywords: Componentwise maxima (minima); Copula; Marshall-Olkin exponential distribution; Multivariate extreme value distribution; Pickands' representation

[1]E-mail: aasimit@stats.uwo.ca

[2]Corresponding Author. Telephone: 519-661-3149; Fax: 519-661-3813; E-mail: jones@stats.uwo.ca

1 Introduction

Extreme value theory has received increasing attention in the actuarial literature in recent years. The severe financial implications of extreme events justify the need for such quantitative tools. Since many insurance portfolios include several (or many) dependent risks, multivariate extreme value theory is needed to properly quantify the overall risk.

The limiting distribution of the normalized componentwise maxima (minima) of a sequence of iid random vectors is a fundamental and thoroughly studied topic in the area of multivariate extreme value theory. The possible limit distributions are known as max (min) multivariate extreme value (MEV) distributions. One of the key features of these distributions is that they cannot be specified in terms of a function involving a finite number of parameters (see Beirlant *et al.*, 2004). A number of parametric families of multivariate extreme value distributions have been discussed in the literature. However, none is sufficiently broad to widely cover the entire class, and most simple families are quite restricted in their behavior.

In this paper we establish the limit distribution for the normalized componentwise maxima and minima of a sequence of random vectors with multivariate phase-type (MPH) distributions. Introduced by Assaf *et al.* (1984), multivariate phase-type random vectors can be viewed as representing the times until absorption into overlapping non-empty subsets of the state space of a finite-state continuous-time Markov chain. MPH distributions have been used in reliability theory (see Assaf *et al.*, 1984), queueing theory (see Li and Xu, 2000) and ruin theory (see Cai and Li, 2005a).

The collection of limiting distributions forms a rich subclass of the max extreme value distributions. We provide some examples of bivariate phase-type distributions and explore the behavior of the Pickands' function corresponding to the limiting distribution of componentwise maxima.

In Section 2, we present some preliminaries on MEV distributions and establish some of the notation that will be used throughout the paper. This is continued in Section 3 where we discuss the basics of univariate phase-type distributions, including the limiting distributions of normalized maxima and minima. Section 4 introduces the multivariate phase-type distribution and the bivariate special case, and gives the main results of the paper - the limiting distributions of normalized componentwise maxima and minima along with the norming constants. Some examples illustrating the flexibility of this class of distributions are provided in Section 5. Conclusions are given in Section 6.

2 Preliminaries

Let $\mathbf{Z}_1 = (X_1, Y_1)$, $\mathbf{Z}_2 = (X_2, Y_2), \ldots$ be a sequence of independent random vectors with common distribution F, and let

$$\mathbf{U}_n = \left(\max_{i=1,\ldots,n} \{X_i\}, \max_{i=1,\ldots,n} \{Y_i\} \right).$$

That is, $\mathbf{U}_n$ is the vector of componentwise maxima of $\mathbf{Z}_1, \ldots, \mathbf{Z}_n$. If there exist sequences of vectors of constants $\mathbf{a}_n$, $\mathbf{b}_n \in \Re^2$ and a random vector $\mathbf{U}$ with distribution G and nondegenerate marginals such that $\mathbf{a}_n\mathbf{U}_n + \mathbf{b}_n$ converges weakly to $\mathbf{U}$, then G, the limit distribution of normalized componentwise maxima, is said to be a *max extreme value distribution*. We then say that F is in the *max domain of attraction* of G with *normalizing vectors of constants* $\mathbf{a}_n$ and $\mathbf{b}_n$ and write $F \in \mathrm{MaxDA}(G)$. It follows that

$$\lim_{n\to\infty} n[1 - F(\mathbf{a}_n\mathbf{x} + \mathbf{b}_n)] = -\log G(\mathbf{x}), \tag{1}$$

for all $\mathbf{x}$ such that $G(\mathbf{x}) > 0$. This relation is useful in verifying the limit distribution of the normalized componentwise maxima.

Analogous to $\mathbf{U}_n$, define $\mathbf{L}_n$ to be the vector of componentwise minima of $\mathbf{Z}_1, \ldots, \mathbf{Z}_n$. If there exist sequences of vectors of constants $\mathbf{a}_n$, $\mathbf{b}_n \in \Re^2$ and a random vector $\mathbf{L}$ with distribution G and nondegenerate marginals such that $\mathbf{a}_n\mathbf{L}_n + \mathbf{b}_n$ converges weakly to $\mathbf{L}$, then G, the limit distribution of normalized componentwise minima, is said to be a *min extreme value distribution*, and F is said to be in the *min domain of attraction* of G. We write $F \in \mathrm{MinDA}(G)$. It follows that

$$\lim_{n\to\infty} n[1 - \bar{F}(\mathbf{a}_n\mathbf{x} + \mathbf{b}_n)] = -\log \bar{G}(\mathbf{x}), \tag{2}$$

for all $\mathbf{x}$ such that $\bar{G}(\mathbf{x}) > 0$.

Necessary and sufficient conditions for multivariate distributions to be in the max or min domains of attraction of multivariate extreme value distributions are given by Marshall and Olkin (1983).

Necessary conditions for $F \in \mathrm{MaxDA}(G)$, respectively $F \in \mathrm{MinDA}(G)$, are that each marginal F_i of F is in the (univariate) MaxDA, respectively MinDA, of the corresponding marginal G_i of G. Classical results concerning max and min extreme value distributions in the univariate case are provided by Gnedenko (1943). In particular, if $F_i \in \mathrm{MDA}(G_i)$ then, by the Fisher-Tippett theorem, G_i belongs to the type of the distribution

$$H_\xi(x) = \begin{cases} \exp\left\{-(1+\xi x)^{-1/\xi}\right\}, & 1+\xi x > 0, \qquad \xi \neq 0 \\ \exp\{-e^{-x}\}, & -\infty < x < \infty, \quad \xi = 0 \end{cases} . \tag{3}$$

H_ξ is known as the *generalized extreme value distribution* (see Beirlant, *et al.*, 2004). For $\alpha > 0$, $\Phi_\alpha(x) := H_{1/\alpha}(\alpha(x-1))$ is the standard Fréchet distribution, $\Psi_\alpha(x) := H_{-1/\alpha}(\alpha(x+1))$ is the standard Weibull distribution, and $\Lambda(x) := H_0(x)$ is the standard Gumbel distribution. This is the well-known Fisher-Tippett theorem. Analogously, if $F_i \in \mathrm{MinDA}(G_i)$ for some non-degenerate df G_i, then G_i belongs to the type of the distribution $H^*_\xi(x) := 1 - H_\xi(-x)$. For $\alpha > 0$, $\Phi^*_\alpha(x) := 1 - \Phi_\alpha(-x)$ is of *type I*, $\Psi^*_\alpha(x) := 1 - \Psi_\alpha(-x)$ is of *type II*, and $\Lambda^*(x) := 1 - \Lambda(-x)$ is of *type III*.

The dependence structure associated with the distribution of a random vector can be characterized in terms of a *copula*. A two-dimensional copula is a bivariate distribution function defined on $[0,1]^2$ with uniformly distributed marginals. Due to Sklar's Theorem (see Sklar, 1959), if F is a joint distribution function with continuous marginals F_1 and F_2 respectively, then there exists a unique copula, C, given by

$$C(u,v) = F(F_1^{\leftarrow}(u), F_2^{\leftarrow}(v)), \tag{4}$$

where $h^{\leftarrow}(u) = \inf\{x : h(x) \geq u\}$ is the generalized inverse function. Similarly, the *survival copula* is defined as the copula relative to the joint survival function and is given by

$$\hat{C}(u,v) = u + v - 1 + C(1-u, 1-v). \tag{5}$$

A more formal definition, properties and examples of copulas are given in Nelsen (1999). For max extreme value distributions, the copula has the form

$$C(u,v) = \exp\left\{\log(uv) A\left(\frac{\log u}{\log(uv)}\right)\right\}, \tag{6}$$

where A is the Pickands' representation function, which is a convex function on $[0,1]$ such that $\max(t, 1-t) \le A(t) \le 1$ (see Pickands, 1981). Note that, for $A(t) \equiv 1$, we have independence, and, for $A(t) = \max(t, 1-t)$, we have perfect positive dependence. For higher dimensional max extreme value distributions, representation functions for the dependence structure are given in, for example, Beirlant *et al.* (2004) and Resnick (1987). Since we focus our attention on the bivariate case, we need not discuss other representations. In the case of min extreme value distributions, the survival copula has the form (6).

3 Univariate Phase-Type Distributions

Let $\{Y(t),\, t \ge 0\}$ be a right-continuous, continuous-time Markov Chain (CTMC) with state space $\xi = \{\Delta, 1, \ldots, d\}$, and initial distribution $\boldsymbol{\beta} = (0, \boldsymbol{\alpha})$. That is, it is assumed that the process starts in $\xi \setminus \{\Delta\}$. Suppose that the CTMC has infinitesimal generator

$$\mathbf{Q} = \begin{pmatrix} 0 & \mathbf{0} \\ \textbf{-Ae} & \mathbf{A} \end{pmatrix}, \tag{7}$$

where the subgenerator $\mathbf{A} = (a_{i,j})$ is a $d \times d$ matrix, $\mathbf{0} = (0, \ldots, 0)$ is a row vector of zeroes and $\mathbf{e} = (1, \ldots, 1)'$ is a column vector of ones. Then the nonnegative random variable X of the time until absorption into state Δ is said to be *phase-type (PH)* distributed with representation $(\boldsymbol{\alpha}, \mathbf{A}, d)$. We assume that absorption into state Δ is certain, or equivalently, that the matrix $\mathbf{A}$ is nonsingular. The survival function of X, denoted by $\bar{F}$, can be expressed as follows:

$$\bar{F}(x) = \Pr(Y(x) \notin \{\Delta\}) = \boldsymbol{\alpha} e^{\mathbf{A}x}\mathbf{e}, \quad x \ge 0. \tag{8}$$

For other properties of PH distributions, see Rolski *et al.* (1999). PH distributions have been used in reliability theory (see Neuts, 1994), queueing theory (see Asmussen, 1992) and ruin theory (see Drekic *et al.*, 2004).

All of the eigenvalues of the subgenerator $\mathbf{A}$ have negative real parts. Also, the matrix $\mathbf{A}$ is of Metzler type. That is, all of its off-diagonal entries are nonnegative. Therefore, it has a real dominant eigenvalue $-\eta$, not necessarily unique, such that for all complex eigenvalues λ, $\mathrm{Re}(\lambda) < -\eta$ (see MacCluer, 2000). If the matrix $\mathbf{A}$ is irreducible, then the dominant eigenvalue $-\eta$ is unique. By expressing $\mathbf{A}$ in Jordan canonical form, one can conclude that there exists a nonnegative matrix of constants $\mathbf{M}$ that satisfies:

1. if $-\eta$ is a simple eigenvalue of $\mathbf{A}$ then

$$e^{\mathbf{A}x} = e^{-\eta x}(\mathbf{M} + \boldsymbol{O}(1)), \tag{9}$$

2. if $-\eta$ has algebric multiplicity l, then there exists an integer k $(0 \leq k \leq l-1)$ such that

$$e^{\mathbf{A}x} = x^k e^{-\eta x}(\mathbf{M} + \boldsymbol{O}(1)), \tag{10}$$

where $\boldsymbol{O}(1)$ is a matrix with entries that are $o(1)$ as $x \to \infty$, and $k+1$ is the maximal order of Jordan blocks corresponding to $-\eta$, called the index of $-\eta$ (see Perko, 2001 ch. 1, or Horn and Johnson, 1985, ch. 3).

This suggests the following approach to finding the matrix $\mathbf{M}$. First, determine the eigenvalues of $\mathbf{A}$. Let $-\eta$ be the largest real eigenvalue. If $-\eta$ has algebraic multiplicity 1, then let

$$\mathbf{M} = \lim_{x \to \infty} e^{\eta x} e^{\mathbf{A}x}.$$

If $-\eta$ has algebraic multiplicity $l > 1$, then calculate the matrix

$$\lim_{x \to \infty} x^{-k} e^{\eta x} e^{\mathbf{A}x},$$

for $k = 0, \ldots, l-1$, and let $\mathbf{M}$ be the matrix obtained using the largest value of k such that the above limit does not give the zero matrix.

This approach adapts and gives a more general way of finding the matrix $\mathbf{M}$ than that of Theorem 9 from Kang and Serfozo (1999). When all of the eigenvalues are real, the Putzer algorithm (see Theorem 8.2.2., Rolski *et al.*, 1999) leads to a simpler alternative than the method described above. These results are sufficient to find the limiting distribution of the normalized maxima of a sequence of iid PH-distributed random variables. It is well-known that this distribution must be one of the three distributions in the class of generalized extreme value (GEV) distributions – the Fréchet, Weibull and Gumbel distributions given in (3). The following proposition indicates that it is the Gumbel distribution and gives the corresponding norming constants. The norming constants for the case in which (9) holds are also given in Theorem 9 of Kang and Serfozo (1999).

Proposition 1 *Let X be a PH distributed random variable with representation $(\boldsymbol{\alpha}, \boldsymbol{A}, d)$. Then its distribution is in the MaxDA(Λ). If (9) holds, then the norming constants are*

$$a_n = \frac{1}{\eta}, \quad b_n = \frac{\log nc}{\eta}, \tag{11}$$

and if (10) holds, then the norming constants are

$$a_n = \frac{1}{\eta}, \quad b_n = \frac{\log nc + k\ \log\log n - k\ \log\eta}{\eta}, \tag{12}$$

where $c = \boldsymbol{\alpha M e}$ is assumed to be positive.

Proof. Since (9) is the special case of (10) with $k = 0$, it is sufficient to check that the convergence criterion in (1) is satisfied using the norming constants in (12). From (8) and (10), we have

$$\begin{aligned} n\bar{F}(a_n x + b_n) &= n\boldsymbol{\alpha}[\mathbf{M} + \boldsymbol{O}(1)]\mathbf{e}(a_n x + b_n)^k e^{-\eta(a_n x + b_n)} \\ &= n[c + o(1)]\left(\frac{\log n + o(\log n)}{\eta}\right)^k e^{-x-\log nc - k\log\log n + k\log\eta} \\ &= (1 + o(1))e^{-x}\left(\frac{\log n + o(\log n)}{\log n}\right)^k \to e^{-x}, \quad \text{as} \quad n \to \infty, \end{aligned} \tag{13}$$

which completes the proof. ■

The limiting distribution of the normalized minima of a sequence of iid PH-distributed random variables along with the norming constants is given by Proposition 2. We first provide a lemma which will be used in proving the proposition.

Lemma 1 *If the random variable X is PH distributed with representation $(\boldsymbol{\alpha}, \boldsymbol{A}, d)$, then m is the minimum number of transitions needed for the underlying CTMC to be absorbed if and only if*

$$-\boldsymbol{\alpha}\boldsymbol{A}^m\boldsymbol{e} > 0 \text{ and when } m \geq 2,\ -\boldsymbol{\alpha}\boldsymbol{A}^\ell\boldsymbol{e} = 0,\ \ell = 1, \ldots m-1. \tag{14}$$

Proof. Let $I_0 = \{i|\ \alpha_i > 0\}$ be a subset of the state space ξ and $\mathbf{a}_\Delta = -\mathbf{Ae}$, with ith component $a_{i,\Delta}$, be the exit rate vector from the CTMC. Then

$$-\boldsymbol{\alpha}\mathbf{Ae} = \sum_{i \in I_0} \alpha_i\ a_{i,\Delta}. \tag{15}$$

If $m = 1$, then the right-hand side of (15) is positive since there exists at least one transient state with positive probability of being the initial state for which direct absorption is possible.

When $m \geq 2$, then, for $i \in I_0$, $i_1, i_2, \ldots \in \xi \setminus \{\Delta\}$, and all $\ell = 1, \ldots, m-1$,

$$a_{i,i_1}a_{i_1,i_2}, \ldots, a_{i_{\ell-1},\Delta} = 0. \tag{16}$$

Also,

$$a_{i,i_1}a_{i_1,i_2}, \ldots, a_{i_{m-1},\Delta} > 0 \tag{17}$$

for some $\{i_1, \ldots, i_{m-1}\}$ since absorption is possible on the mth transition. Furthermore, whenever the left-hand side of (17) is not positive, it must be 0. We see this by noting that the product can be negative only if an odd number of terms are negative. However, from (16), the product of the remaining terms must be 0. Now

$$-\boldsymbol{\alpha}\mathbf{A}^\ell\mathbf{e} = \sum_{i \in I_0} \alpha_i \sum_{i_1, \ldots, i_{\ell-1}} a_{i,i_1}a_{i_1,i_2}, \ldots, a_{i_{\ell-1},\Delta}, \text{ for } \ell \geq 2. \tag{18}$$

For $\ell < m$, each term on the right-hand side of (18) vanishes due to (16), and for $\ell = m$, the right-hand side of (18) must be positive due to (17). This completes the proof of necessity. The sufficiency part of the proof follows from the same arguments. ∎

Proposition 2 *Let X be a PH distributed random variable with representation $(\boldsymbol{\alpha}, \boldsymbol{A}, d)$. Then its distribution is in the $MinDA(\Psi^*_m)$ with norming constants*

$$a_n = \left(\frac{m!}{nc}\right)^{\frac{1}{m}}, \quad b_n = 0, \tag{19}$$

where the constant $c = -\boldsymbol{\alpha}\boldsymbol{A}^m\boldsymbol{e}$, and m is the minimum number of transitions needed for the CTMC to be absorbed.

Proof. Let F be the distribution function of X. It is sufficient to check the convergence criterion (2). Using Lemma 1 and the fact that $e^{\mathbf{A}x} = \mathbf{I} + \sum_{i=1}^{m} \frac{\mathbf{A}^i x^i}{i!} + \boldsymbol{O}(x^m)$ as $x \downarrow 0$ we have

$$
\begin{aligned}
n[1 - \bar{F}(a_n x + b_n)] &= n\left[1 - \boldsymbol{\alpha}\left\{\mathbf{I} + \sum_{i=1}^{m} \frac{\mathbf{A}^i x^i}{i!}\left(\frac{m!}{nc}\right)^{\frac{i}{m}} + \boldsymbol{O}(n^{-1})\right\}\mathbf{e}\right] \qquad (20)\\
&\to x^m, \text{ as } n \to \infty,
\end{aligned}
$$

where $\mathbf{I}$ is the identity matrix. Thus, F is in the *type II* class. ■

4 Multivariate Phase-Type Distributions

Let $\{Y(t),\, t \geq 0\}$ be a continuous-time Markov Chain (CTMC) with finite state space $\xi = \{\Delta, 1, \ldots, d\}$ and infinitesimal generator $\mathbf{Q}$ defined as in (7). Let ξ_i, $i = 1, \ldots, p$, be nonempty stochastically closed subsets of the state space ξ such that $\bigcap_{i=1}^{p} \xi_i$ is a proper subset of ξ. A subset of the state space is said to be stochastically closed if, once the process $\{Y(t),\, t \geq 0\}$ enters the subset, it never leaves. We assume that absorption into $\bigcap_{i=1}^{p} \xi_i$ is certain. Since we are interested in the process only until it is absorbed into $\bigcap_{i=1}^{p} \xi_i$, we may assume that $\bigcap_{i=1}^{p} \xi_i$ consists of one state denoted by Δ. We may write $\xi = \left(\bigcup_{i=1}^{p} \xi_i\right) \bigcup \xi_0$ for some subset $\xi_0 \subset \xi$ with $\xi_0 \bigcap \xi_i = \emptyset$ for $i = 1, \ldots, p$. Let $\boldsymbol{\beta} = (0, \boldsymbol{\alpha})$ be the initial distribution, with each component representing the probability that the process starts in a particular state in ξ.

We define $X_i = \inf\{t \geq 0 : Y(t) \in \xi_i\}$ for $i = 1, \ldots, p$. For simplicity, we may assume that $\Pr\{X_1 > 0, \ldots, X_p > 0\} = 1$, which means that the CTMC starts within ξ_0. The joint distribution of $(X_1, \ldots, X_p)$ is called a *multivariate phase-type* (MPH) distribution with representation $(\boldsymbol{\alpha}, \mathbf{A}, \xi, \xi_1, \ldots, \xi_p)$, and $(X_1, \ldots, X_p)$ is called a *phase-type* random vector (see Assaf *et al.*, 1984). Thus, a MPH distribution is a joint distribution of first passage times to various subsets of the state space ξ of a CTMC.

Examples of MPH distributions include, among many others, the multivariate exponential distributions of Marshall and Olkin (1967). The set of p-dimensional MPH distributions is dense in the set of all distributions on $[0, \infty)^p$. For further details and for discussions of the closure properties of these distributions, see Assaf *et al.* (1984) or Cai and Li (2005a). Some results on order statistics of MPH random vectors are given in Cai and Li (2005b).

Let $\bar{F}$ denote the joint survival function of a MPH distribution. Then by Assaf *et al.* (1984) we have for $0 \le x_p \le \cdots \le x_1$ that

$$\bar{F}(x_1, \ldots, x_p) = \Pr\left(\bigcap_{i=1}^{p}\{X_i > x_i\}\right) = \boldsymbol{\alpha}\, e^{\mathbf{A}x_p}\, \mathbf{g}_p\, e^{\mathbf{A}(x_{p-1}-x_p)}\, \mathbf{g}_{p-1} \cdots e^{\mathbf{A}(x_1-x_2)}\, \mathbf{g}_1\mathbf{e}, \quad (21)$$

where, for $k = 1, \ldots, p$, $\mathbf{g}_k$ is a $d \times d$ diagonal matrix whose ith diagonal entry, for $i = 1, \ldots, d$, equals 1 if $i \in \xi \setminus \xi_k$ and 0 otherwise.

The random variable X_i represents the first passage time of the CTMC into ξ_i. This implies that X_i is univariate PH distributed with representation $(\boldsymbol{\alpha}_{\xi\setminus\xi_i}, \mathbf{A}_{\xi\setminus\xi_i}, d+1-|\xi_i|)$, where $\boldsymbol{\alpha}_{\xi\setminus\xi_i}$ and $\mathbf{A}_{\xi\setminus\xi_i}$ are the probability entry distribution and subgenerator matrix restricted to the state space $\xi \setminus \xi_i$. As in Section 2, η_i, k_i, and $\mathbf{M}_i$ are defined for the i-th marginal of the MPH random vector. The matrix $\mathbf{M}_i$ is extended to have dimension d, by padding it with zeroes. In order to avoid an abuse of notation, this padded matrix is denoted by $\mathbf{M}_i$.

In the bivariate case, the subgenerator has the special form

$$\mathbf{A} = \begin{pmatrix} \mathbf{A_0} & \mathbf{B_1} & \mathbf{B_2} \\ 0 & \mathbf{A_1} & 0 \\ 0 & 0 & \mathbf{A_2} \end{pmatrix}, \quad (22)$$

where, for $i = 0, 1, 2$, $\mathbf{A}_i$ represents the subgenerator for states in $\xi_i \setminus \{\Delta\}$, and for $i = 1, 2$, $\mathbf{B}_i$ represents the matrix of transition intensities from states in ξ_0 to states in $\xi_i \setminus \{\Delta\}$.

The following theorem establishes the limiting distribution for bivariate PH distributions. The extension to higher dimensions will be outlined later.

Theorem 1 *Let F be the distribution function of a bivariate PH distribution with representation $(\boldsymbol{\alpha}, \boldsymbol{A}, \xi, \xi_1, \xi_2)$. Then there exist sequences of constants $\boldsymbol{a}_n$, $\boldsymbol{b}_n \in \Re^2$ such* *that (1) holds with G given by*

$$G(x_1,x_2)=\begin{cases} e^{-e^{-x_1}}e^{-e^{-x_2}}\exp\left\{\frac{e^{-x_1}}{c_1}\boldsymbol{\alpha}\boldsymbol{M}_1 e^{\boldsymbol{A}\left(x_2+\log c_2-x_1-\log c_1\right)\eta^{-1}}\boldsymbol{g}_2\boldsymbol{e}\right\}, \\ \qquad\qquad \text{if } x_1+\log c_1\le x_2+\log c_2 \\ e^{-e^{-x_1}}e^{-e^{-x_2}}\exp\left\{\frac{e^{-x_2}}{c_2}\boldsymbol{\alpha}\boldsymbol{M}_2 e^{\boldsymbol{A}\left(x_1+\log c_1-x_2-\log c_2\right)\eta^{-1}}\boldsymbol{g}_1\boldsymbol{e}\right\}, \\ \qquad\qquad \text{if } x_2+\log c_2\le x_1+\log c_1 \end{cases} \tag{23}$$

whenever $\eta_1=\eta_2=\eta$ *and* $k_1=k_2=k$*, where* $c_i=\boldsymbol{\alpha}\boldsymbol{M}_i\boldsymbol{e}$ *is assumed to be positive for* $i=1,2$*. For any other case we have independence, and* $G(x_1,x_2)=\exp(-e^{-x_1})\exp(-e^{-x_2})$.

Proof. Since X_1 and X_2 are PH distributed, both are in the MaxDA of the Gumbel distribution with respective normalizing constants $a_{n,1}$, $b_{n,1}$, and $a_{n,2}$, $b_{n,2}$. From basic probability we have

$$\begin{aligned} &n[1-\Pr(X_1\le a_{n,1}x_1+b_{n,1},X_2\le a_{n,2}x_2+b_{n,2})] \\ &\quad= n\Pr(X_1>a_{n,1}x_1+b_{n,1})+n\Pr(X_2>a_{n,2}x_2+b_{n,2}) \\ &\qquad\qquad -n\Pr(X_1>a_{n,1}x_1+b_{n,1},X_2>a_{n,2}x_2+b_{n,2}). \end{aligned}\tag{24}$$

From Proposition 1, the first two terms on the right hand side of (24) have limits e^{-x_1} and e^{-x_2}, respectively. If $\eta_1=\eta_2=\eta$ and $k_1=k_2=k$, for x_1 and x_2 such that $x_1+\log c_1\le x_2+\log c_2$, from (9 or 10) and (21) we obtain

$$\begin{aligned} &n\Pr(X_1>a_{n,1}x_1+b_{n,1},X_2>a_{n,2}x_2+b_{n,2}) \\ &\quad= \frac{e^{-x_1}}{c_1}\left(\frac{\log n+o(\log n)}{\log n}\right)^k\boldsymbol{\alpha}(\mathbf{M}_1+\boldsymbol{O}(1))e^{\mathbf{A}\left(x_2+\log c_2-x_1-\log c_1\right)\eta^{-1}}\mathbf{g}_2\mathbf{e} \\ &\quad\to \frac{e^{-x_1}}{c_1}\boldsymbol{\alpha}\mathbf{M}_1 e^{\mathbf{A}\left(x_2+\log c_2-x_1-\log c_1\right)\eta^{-1}}\mathbf{g}_2\mathbf{e},\quad \text{as } n\to\infty, \end{aligned}$$

which completes the proof for this case.

If $\eta_1>\eta_2$, then for n sufficiently large, $a_{n,1}x_1+b_{n,1}<a_{n,2}x_2+b_{n,2}$, and

$$\begin{aligned} a_{n,2}x_2+b_{n,2}-a_{n,1}x_1-b_{n,1} &= \left(\frac{1}{\eta_2}-\frac{1}{\eta_1}\right)\log n+o(\log n) \\ &\to \infty,\ \text{as } n\to\infty. \end{aligned}$$

Therefore,

$$\begin{aligned} & n \Pr(X_1 > a_{n,1}x_1 + b_{n,1}, X_2 > a_{n,2}x_2 + b_{n,2}) \\ = & \frac{e^{-x_1}}{c_1} \boldsymbol{\alpha}(\mathbf{M}_1 + \boldsymbol{O}(1)) \boldsymbol{O}(1) \mathbf{e} \\ \to & 0, \text{ as } n \to \infty. \end{aligned}$$

This implies that we are in the independence case. In a similar way, the remaining cases yield the same result. ■

Starting with (6), simple algebraic computations show that, if F is a bivariate PH distribution function, then (1) holds, where G is a BEV distribution with Gumbel marginals and dependence structure given by the Pickands' representation function

$$A(t) = \begin{cases} 1 - \frac{1-t}{c_2} \boldsymbol{\alpha} \mathbf{M}_2 e^{\mathbf{A} \frac{1}{\eta} \log \frac{c_1}{c_2} \frac{1-t}{t}} \mathbf{g}_1 \mathbf{e}, & \text{if } 0 \le t \le \frac{c_1}{c_1+c_2} \\ 1 - \frac{t}{c_1} \boldsymbol{\alpha} \mathbf{M}_1 e^{\mathbf{A} \frac{1}{\eta} \log \frac{c_2}{c_1} \frac{t}{1-t}} \mathbf{g}_1 \mathbf{e}, & \text{if } \frac{c_1}{c_1+c_2} \le t \le 1 \end{cases}. \tag{25}$$

A number of other characterizations of MEV distributions have been proposed (see, for example, Balkema and Resnick, 1977, de Haan and Resnick, 1977). For further discussion of the different representations, see de Haan and de Ronde (1998) or Beirlant *et al.* (2004). For ease of presentation of our examples in Section **??**, we consider only Pickands' representation.

By using the same logic as in Theorem 1 and the identity

$$\begin{aligned} & \Pr\Big(\bigcap_{i=1}^{p} \{X_i \le x_i\}\Big) \\ = & 1 - \sum_{i=1}^{p} \Pr(X_i > x_i) + \sum_{i<j} \Pr(X_i > x_i, X_j > x_j) - \ldots (-1)^p \Pr\Big(\bigcap_{i=1}^{p} \{X_i > x_i\}\Big), \end{aligned}$$

we can obtain the limit distribution for higher dimensional MPH distributions. In order to take more advantage of the structure of $\mathbf{A}$, it is convenient to rearrange the state space. As in Cai and Li (2005a), ξ is partitioned as follows:

$$
\begin{aligned}
\Gamma^{p}_{\emptyset} &= \xi_0, \\
\Gamma^{p-1}_{\{i\}} &= \xi_i \setminus \big(\bigcup_{k\neq i}(\xi_i \cap \xi_k)\big),\ i = 1,\ldots,p \\
\Gamma^{p-2}_{\{i,j\}} &= \big(\xi_i \cap \xi_j\big) \setminus \big(\bigcup_{k\neq i, k\neq j}(\xi_i \cap \xi_j \cap \xi_k)\big),\ i,j = 1,\ldots,p,\, i \neq j \\
&\vdots \\
\Gamma^{p-|\mathcal{D}|}_{\mathcal{D}} &= \big(\bigcap_{i\in\mathcal{D}} \xi_i\big) \setminus \big(\bigcup_{k\notin\mathcal{D}} \big(\bigcap_{i\in\mathcal{D}} \xi_i \cap \xi_k\big)\big),\ \mathcal{D} \subset \{1,2,\ldots,p\} \\
&\vdots \\
\Gamma^{0}_{\xi\setminus\{\Delta\}} &= \{\Delta\},
\end{aligned}
$$

where $|\cdot|$ denotes set cardinality. Notice that, by partitioning the state space in this fashion and reordering the states so that $i < j$ whenever $i \in \Gamma^{p-|\mathcal{D}_1|}_{\mathcal{D}_1}$, $j \in \Gamma^{p-|\mathcal{D}_2|}_{\mathcal{D}_2}$ and $|\mathcal{D}_1| < |\mathcal{D}_2|$, the subgenerator $\mathbf{A}$ becomes a block upper triangular matrix. Therefore, its eigenvalue set coincides with the union of eigenvalue sets of diagonal blocks, which simplifies the problem of finding the eigenvalues for high cardinality state spaces.

The following theorem provides the analogous limit distribution for normalized componentwise minima of bivariate PH distributed random vectors.

Theorem 2 *Let F be the distribution function of a bivariate PH distribution with representation $(\boldsymbol{\alpha}, \boldsymbol{A}, \xi, \xi_1, \xi_2)$. Then there exist sequences of constants $\boldsymbol{a}_n$, $\boldsymbol{b}_n \in \Re^2$ such that (2) holds with $\bar{G}$ given by*

$$\bar{G}(x_1, x_2) = \exp\left\{-x_1^m - x_2^m + c\min\left(\frac{x_1^m}{c_1}, \frac{x_2^m}{c_2}\right)\right\}, \tag{26}$$

where $c_i = -\boldsymbol{\alpha}\boldsymbol{A}^{m_i}\boldsymbol{g}_i\boldsymbol{e}$, $i = 1, 2$, and $c = -\boldsymbol{\alpha}\boldsymbol{A}^m\boldsymbol{e}$, provided that $m_1 = m_2 = m$, where m_i is the minimum number of transitions required in order to enter ξ_i. Otherwise, we are in the independence case and $\bar{G}(x_1, x_2) = \exp(-x^{m_1} - x^{m_2})$.

Remark: If $m_1 = m_2 = m$ then the limiting distribution has the Marshall-Olkin dependence structure

$$\hat{C}(u, v) = \min(u^{1-a}v, uv^{1-b}),\ 0 < a, b < 1,$$

where $a = \dfrac{c}{c_1}$ and $b = \dfrac{c}{c_2}$.

Proof. Let $a_{n,i} = \left(\frac{m_i!}{nc_i}\right)^{\frac{1}{m_i}}$ and $b_{n,i} = 0$ be the norming constants defined as in (19). It is sufficient to verify the convergence criterion (2). Throughout the proof, we will make use of the fact that $\boldsymbol{\alpha}\mathbf{g}_i = \boldsymbol{\alpha}$ for $i = 1, 2$, since the underlying CTMC starts in ξ_0.

In the case that $m_1 = m_2 = m$, if $0 \leq \frac{x_1^{m_1}}{c_1} \leq \frac{x_2^{m_2}}{c_2}$, then similar to the proof of Proposition 2, we have

$$\begin{aligned}
& n[1 - \bar{F}(\mathbf{a}^{(n)}\mathbf{x} + \mathbf{b}^{(n)})] \\
= \ & n\Bigg\{1 - \boldsymbol{\alpha}\Bigg[\mathbf{I} + \sum_{i=1}^{m} \mathbf{A}^i \frac{x_1^i}{i!}\left(\frac{m!}{nc_1}\right)^{\frac{i}{m}} + \boldsymbol{O}(n^{-1})\Bigg]\mathbf{g}_1 \\
& \qquad \Bigg[\mathbf{I} + \sum_{j=1}^{m} \mathbf{A}^j \frac{1}{j!}\left(\frac{m!}{n}\right)^{\frac{j}{m}} \left[x_2 c_2^{-\frac{1}{m}} - x_1 c_1^{-\frac{1}{m}}\right]^j + \boldsymbol{O}(n^{-1})\Bigg]\mathbf{g}_2\mathbf{e}\Bigg\} \\
= \ & n\Bigg\{\sum_{i=1}^{m}(-\boldsymbol{\alpha}\mathbf{A}^i\mathbf{g}_1\mathbf{g}_2\mathbf{e})\left(\frac{m!}{nc_1}\right)^{\frac{i}{m}}\frac{x_1^i}{i!} \\
& \qquad + \sum_{j=1}^{m}(-\boldsymbol{\alpha}\mathbf{A}^j\mathbf{g}_2\mathbf{e})\left(\frac{m!}{n}\right)^{\frac{j}{m}}\left[x_2 c_2^{-\frac{1}{m}} - x_1 c_1^{-\frac{1}{m}}\right]^j \frac{1}{j!} \\
& \qquad + \sum_{i=1}^{m-1}\sum_{j=1}^{m-i}(-\boldsymbol{\alpha}\mathbf{A}^i\mathbf{g}_1\mathbf{A}^j\mathbf{g}_2\mathbf{e})\left(\frac{m!}{n}\right)^{\frac{i+j}{m}} x_1^i c_1^{-i/m} \\
& \qquad\qquad \frac{1}{i!j!}\left[x_2 c_2^{-\frac{1}{m}} - x_1 c_1^{-\frac{1}{m}}\right]^j + \boldsymbol{O}(n^{-1})\Bigg\} \\
\to \ & x_1^m + x_2^m - c\frac{x_1^m}{c_1}, \text{ as } n \to \infty,
\end{aligned} \tag{27}$$

where (27) follows from Lemma 1 and the fact that

$$-\boldsymbol{\alpha}\mathbf{A}^i\mathbf{g}_1\mathbf{A}^{m-i}\mathbf{g}_2\mathbf{e} = -\boldsymbol{\alpha}\mathbf{A}^m\mathbf{g}_2\mathbf{e},$$

and

$$c = c_1 + c_2 - (-\boldsymbol{\alpha}\mathbf{A}^m\mathbf{g}_1\mathbf{g}_2\mathbf{e}).$$

This completes the proof in the $m_1 = m_2 = m$ case. Similar matrix manipulations are used to prove the Theorem in the $m_1 \neq m_2$ case. ∎

5 Examples

In this section, we present some simple examples of bivariate PH distributions. We find that, even in simple cases, we are able to achieve a wide variety of dependence structures within the BEV class. We explore this by examining the Pickands' representation function which is given by (25).

Example 1

In this example we consider a bivariate PH distribution with representation $(\boldsymbol{\alpha}, \mathbf{A}, \xi, \xi_1, \xi_2)$, where

$$\boldsymbol{\alpha} = (1,0,0), \quad \mathbf{A} = \begin{pmatrix} -a & p & q \\ 0 & -b & 0 \\ 0 & 0 & -c \end{pmatrix}, \quad a < \min(b,c),\ p+q \leq a,$$

$$\xi = \{\Delta, 1, 2, 3\},\ \ \xi_1 = \{\Delta, 2\},\ \ \xi_2 = \{\Delta, 3\}.$$

Then one gets $\eta = a$, $k = 1$, $c_1 = 1 + \frac{q}{c-a}$, $c_2 = 1 + \frac{p}{b-a}$, and from (25), the Pickands' representation function is given by

$$A(t) = \begin{cases} 1 - t + \left(\frac{b-a}{(c-a)(b+p-a)}\right)^{1-\frac{c}{a}} q(c+q-a)^{-\frac{c}{a}} t^{\frac{c}{a}} (1-t)^{1-\frac{c}{a}}, & 0 \leq t \leq \frac{c_1}{c_1+c_2} \\ t + \left(\frac{c-a}{(b-a)(q+c-a)}\right)^{1-\frac{b}{a}} p(p+b-a)^{-\frac{b}{a}} t^{1-\frac{b}{a}} (1-t)^{\frac{b}{a}}, & \frac{c_1}{c_1+c_2} \leq t \leq 1 \end{cases}.$$

In the special case where $p = q = 0$, we have $A(t) = \max(t, 1-t)$, which corresponds to the perfect positive dependence case. In this case the underlying CTMC is certain to make a direct transition from state 1 to state Δ. Therefore, $X_1 = X_2$ with probability 1 and the componentwise maxima must also be equal.

Figure **??** shows the Pickands' function obtained using three different sets of parameters. In all three cases, we have assumed that $b = c$ and $p = q$. The resulting symmetry leads to symmetric $A(t)$ functions. Notice that as b and c approach a, we move closer to the independence case: $A(t) = 1$. Also, note that by choosing p and q so that $p + q = a$, the underlying CTMC cannot make a direct transition from state 1 to state Δ. Therefore, X_1 and X_2 are different with probability 1.

Figure 1: Plots of Pickands' $A(t)$ function for Example 1 with $(a, b, c, p, q) = (2, 3, 3, 0, 0) \rightarrow$ solid line, $(a, b, c, p, q) = (2, 3, 3, 1, 1) \rightarrow$ long-dashed line, $(a, b, c, p, q) = (2, 2.1, 2.1, 1, 1) \rightarrow$ short-dashed line.

Figure **??** shows the Pickands' function for three different sets of parameters. Each of these functions is asymmetric.

Figure 2: Plots of Pickands' $A(t)$ function for Example 1 with $(a, b, c, p, q) = (2, 2.1, 3, 1, 1) \rightarrow$ solid line, $(a, b, c, p, q) = (2, 3, 2.5, 0.1, 1) \rightarrow$ long-dashed line, $(a, b, c, p, q) = (2, 3, 3, 1, 0.1) \rightarrow$ short-dashed line.

Example 2

In this example, we consider the same setup as Example 1, except that we assume $a = b = c$. We also require that $a, p, q > 0$. Thus, we have

$$\mathbf{A} = \begin{pmatrix} -a & p & q \\ 0 & -a & 0 \\ 0 & 0 & -a \end{pmatrix},$$

which implies that $\eta = a$, $k = 2$, $c_1 = q$, $c_2 = p$, and $A(t) = 1$. Notice that we are in the independence case, even though we have satisfied all the conditions of Theorem 1 necessary for G to be given by (23).

Example 3

In this example we consider a bivariate PH distribution with representation $(\boldsymbol{\alpha}, \mathbf{A}, \xi, \xi_1, \xi_2)$, where

$$\boldsymbol{\alpha} = (p, 1-p, 0, 0),\ 0 \le p \le 1, \quad \mathbf{A} = \begin{pmatrix} -5 & 0 & 1 & 2 \\ 0 & -5 & 2 & 0 \\ 0 & 0 & -7 & 0 \\ 0 & 0 & 0 & -6 \end{pmatrix},$$

$$\xi = \{\Delta, 1, 2, 3, 4\}, \;\; \xi_1 = \{\Delta, 3\}, \;\; \xi_2 = \{\Delta, 4\}.$$

then $\eta = 5$, $k = 1$, $c_1 = 1 + 2p$, $c_2 = 2 - \frac{p}{2}$, which implies that

$$A(t) = \begin{cases} 1 - t + 2^{\frac{4}{5}} p(1+2p)^{-\frac{6}{5}}(4-p)^{\frac{1}{5}} t^{\frac{6}{5}}(1-t)^{-\frac{1}{5}}, & 0 \le t \le \frac{2+4p}{6+3p} \\ t + 2^{\frac{2}{5}}(2-p)(4-p)^{-\frac{7}{5}}(1+2p)^{\frac{2}{5}} t^{-\frac{2}{5}}(1-t)^{\frac{7}{5}}, & \frac{2+4p}{6+3p} \le t \le 1 \end{cases}.$$

Figure **??** shows the Pickands' function for three different values of the parameter p.

Figure 3: Plots of Pickands' $A(t)$ function for Example 3 with $p = 0 \to$ solid line, $p = 0.5 \to$ long-dashed line, $p = 1 \to$ short-dashed line.

6 Summary and Conclusions

In this paper, we establish the set of attractors for the componentwise minima and maxima of an iid sequence of random vectors from a fairly general class of multivariate distributions known as the multivariate phase-type (MPH) distributions. The norming constants and corresponding MEV distributions are explicitly given. For the sake of simplicity, we focus on the bivariate case.

The limiting distribution of the componentwise maxima of bivariate phase-type random vectors has a complicated form. In order to investigate its behavior, the Pickands' representation is chosen. Our examples illustrate that the limit distribution allows considerable flexibility within the BEV class. This suggests that the MPH distribution functions are well suited for statistical inference of multivariate data.

It is shown that the dependence structure of the limit distribution of componentwise minima of MPH random vectors coincides with that of the multivariate exponential (Marshall-Olkin) distributions.

Acknowledgments

The authors wish to thank an anonymous referee for helpful comments which improved the paper. This work was supported by a Discovery Grant from the Natural Sciences and Engineering Research Council of Canada.

References

Asmussen, S. 1992. Phase-Type Representations in Random Walk and Queueing Problems. *Annals of Probability*, 20(2), 772-789.

Assaf, D., Langberg, N.A., Savits, T.H. and Shaked, M. 1984. Multivariate Phase-Type Distributions, *Operations Research*, 32(3), 688-702.

Balkema, A.A. and Resnick, S.I. 1977. Max-Infinite Divisibility, *Journal of Applied Probability*, 14(2), 309-319.

Beirlant, J., Goegebeur, Y., Segers, J. and Teugels, J. 2004. *Statistics of Extremes: Theory and Applications*. Wiley, Chichester.

Cai, J. and Li, H. 2005a. Multivariate Risk Model of Phase Type, *Insurance: Mathematics and Economics*, 36(2), 137-152.

Cai, J. and Li, H. 2005b. Conditional Tail Expectations for Multivariate Phase-Type Distributions, *Journal of Applied Probability*, 42(3), 810-825.

Drekic, S., Dickson, D., Stanford, D., and Willmot, G.E. 2004. On the Distribution of the Deficit at Ruin when Claims are Phase-Type. *Scandinavian Actuarial Journal*, 2, 105-120.

de Haan, L. and Resnick, S.I. 1977, Limit Theory for Multivariate Sample Extremes, *Zeitschrift für Wahrscheinlichkeitstheorie und verwandte Gebiete*, 40, 317-337.

de Haan, L. and de Ronde, J. 1998. Sea and Wind: Multivariate Extremes at Work, *Extremes*, 1(1), 7-45.

Gnedenko, B.V. 1943. Sur la distribution limité du terme maximum d'une série aléatoaire, *Annals of Mathematics*, 44(3), 423-453.

Horn, R.A. and Johnson, C.R. 1985. *Matrix Analysis*, CUP, Cambridge.

Kang, S. and Serfozo, R.F. 1999. Extreme Values of Phase-Type and Mixed Random Variables with Parallel-Processing Examples, *Journal of Applied Probability*, 36(1), 194-210.

Li, H. and Xu, S.H. 2000. On the Dependence Structure and Bounds of Correlated Parallel Queues and their Applications to Synchronize Stochastic Systems, *Journal of Applied Probability* 37(4), 1020-1043.

MacCluer, C.R. 2000. The Many Proofs and Applications of Perron's Theorem, *SIAM Review*, 42(3), 487-498.

Marshall, A.W., Olkin, I. 1967. A Multivariate Exponential Distribution, *Journal of the American Statistical Association*, 62, 30-44.

Marshall, A.W., Olkin, I. 1983. Domains of Attraction of Multivariate Extreme Value Distributions, *Annals of Probability*, 11(1), 168-177.

Nelsen, R. B. 1999. *An Introduction to Copulas.* Springer-Verlag, New York.

Neuts, M.F. 1994. *Matrix-Geometric Solutions in Stochastic Models: An Algorithmic Approach*, Dover Publications, New York.

Perko, L. 2001. *Differential Equations and Dynamical Systems*, Springer, New York.

Pickands, J. 1981. Multivariate Extreme Value Distributions, *Bulletin of the International Statistical Institute, Proceedings of the 43rd Session, Buenos Aires*, 49, 859-878.

Resnick, S.I. 1987. *Extreme Values, Regular Variation and Point Processes.* Springer-Verlag, New York.

Rolski, T., Schmidli, H., Schmidt, V. and Teugels, J. 1999. *Stochastic Processes for Finance and Insurance*, Wiley, New York.

Sklar, A. 1959. Fonctions de répartion à n dimensions et leurs marges, Publications de l'Institut de Statistique de l'Université de Paris, 8, 229-231.