

CONDITIONAL TAIL EXPECTATIONS FOR MULTIVARIATE PHASE-TYPE DISTRIBUTIONS

JUN CAI,* *University of Waterloo*

HAIJUN LI,** *Washington State University*

Abstract

The conditional tail expectation in risk analysis describes the expected amount of risk that can be experienced given that a potential risk exceeds a threshold value, and provides an important measure of right-tail risk. In this paper, we study the convolution and extreme values of dependent risks that follow a multivariate phase-type distribution, and derive explicit formulae for several conditional tail expectations of the convolution and extreme values for such dependent risks. Utilizing the underlying Markovian property of these distributions, our method not only provides structural insight, but also yields some new distributional properties of multivariate phase-type distributions.

Keywords: Univariate phase-type distribution; multivariate phase-type distribution; continuous-time Markov chain; convolution; extreme value; Marshall–Olkin distribution; conditional tail expectation; excess loss; residual lifetime

2000 Mathematics Subject Classification: Primary 62E15

Secondary 62P05; 91B30

1. Introduction

Let risk X be a nonnegative random variable with cumulative distribution function F , where X may refer to a claim for an insurance company or a loss on an investment portfolio. The conditional expectation of X given that $X > t$, denoted by $\text{CTE}_X(t) = E(X | X > t)$, is called the *conditional tail expectation* (CTE) of X at t . Observe that

$$\text{CTE}_X(t) = t + E(X - t | X > t),$$

where the random variable $(X - t | X > t)$ is known as the *residual lifetime* in reliability (Shaked and Shanthikumar (1994)) and the *excess loss* or *excess risk* in insurance and finance (Embrechts *et al.* (1997)). Since $(d/dt)(t + E(X - t | X > t)) \geq 0$ for any continuous risk X (see Shaked and Shanthikumar (1994, p. 45)), the CTE function $\text{CTE}_X(t)$ is increasing in t , $t \geq 0$.

The CTE is an important measure of right-tail risk which is frequently encountered in the fields of insurance and financial investment. It is known that the CTE for continuous risks satisfies all the desirable properties of a *coherent risk measure* (Artzner *et al.* (1999)) and that it provides a more conservative measure of risk than the *value-at-risk* at the same confidence level (Landsman and Valdez (2003)). Hence, the CTE is preferable in many applications, and has recently received increased attention in the insurance and finance literature.

Received 4 March 2005; revision received 18 April 2005.

* Postal address: Department of Statistics and Actuarial Science, University of Waterloo, Waterloo, ON N2L 3G1, Canada. Email address: jcai@math.uwaterloo.ca

** Postal address: Department of Mathematics, Washington State University, Pullman, WA 99164, USA. Email address: lih@math.wsu.edu

To analyze dependent risks, several CTE measures emerge, and the most popular ones are the CTEs of the total risk and extreme risks. Let $X = (X_1, \dots, X_n)$ be a risk vector, where X_i denotes risk (claim or loss) in subportfolio i for $i = 1, \dots, n$. Then $S = X_1 + \dots + X_n$ is the aggregate risk or the total risk in the portfolio consisting of the n subportfolios, and $X_{(1)} = \min\{X_1, \dots, X_n\}$ and $X_{(n)} = \max\{X_1, \dots, X_n\}$ are extreme risks in the portfolio. In the risk analysis of such a portfolio, we are not only interested in the CTE of each risk X_i , but also in the CTEs of the statistics S , $X_{(1)}$, and $X_{(n)}$, which are respectively denoted by

$$\begin{aligned} \text{CTE}_S(t) &= E(S \mid S > t), \\ \text{CTE}_{X_{(1)}}(t) &= E(X_{(1)} \mid X_{(1)} > t), \end{aligned} \tag{1.1}$$

$$\text{CTE}_{X_{(n)}}(t) = E(X_{(n)} \mid X_{(n)} > t). \tag{1.2}$$

The risk measures related to these statistics also include, among others,

$$\text{CTE}_{X_i \mid S}(t) = E(X_i \mid S > t), \tag{1.3}$$

$$\text{CTE}_{X_i \mid X_{(1)}}(t) = E(X_i \mid X_{(1)} > t), \tag{1.4}$$

$$\text{CTE}_{X_i \mid X_{(n)}}(t) = E(X_i \mid X_{(n)} > t), \tag{1.5}$$

for $i = 1, 2, \dots, n$.

All these risk measures have useful interpretations in insurance, finance, and other fields. For instance, $\text{CTE}_{X_i \mid S}(t)$ represents the contribution of the i th risk X_i to the aggregate risk S , since $\text{CTE}_S(t) = \sum_{i=1}^n \text{CTE}_{X_i \mid S}(t)$. The CTEs $E(X_{(1)} \mid X_{(1)} > t)$ and $E(X_{(n)} \mid X_{(n)} > t)$ respectively describe the expected minimal and maximal risk in all the subportfolios, given that the minimal or maximal risk exceeds some threshold t . More interestingly, $E(X_i \mid X_{(1)} > t)$ represents the average contribution of the i th risk given that all the risks exceed some value t , whereas $E(X_i \mid X_{(n)} > t)$ represents the average contribution of the i th risk given that at least one risk exceeds a certain value t . Besides their interpretations in risk analysis, the CTEs in (1.1)–(1.5) also have interpretations in life insurance. For instance, in a group life insurance, let X_i , $1 \leq i \leq n$, be the lifetime of the i th member in a group that consists of the n members. Then $X_{(1)}$ is the joint-life status and $X_{(n)}$ is the last-survivor status (Bowers *et al.* (1997)). As such, $E(X_i \mid X_{(1)} > t)$ is the expected lifetime of member i given that all members are alive at time t , and $E(X_i \mid X_{(n)} > t)$ is the expected lifetime of member i given that at least one member is still alive at time t .

Landsman and Valdez (2003) obtained explicit formulae for $\text{CTE}_S(t)$ and $\text{CTE}_{X_i \mid S}(t)$ for the multivariate elliptical distributions, which include distributions such as the multivariate normal, stable, Student t , etc. The focus of this paper is to derive explicit formulae for various CTEs, such as $\text{CTE}_S(t)$, $\text{CTE}_{X_{(1)}}(t)$, $\text{CTE}_{X_{(n)}}(t)$, $E(X_{(n)} \mid X_{(1)} > t)$, and $E(X_{(n)} \mid X_i > t)$, $i = 1, 2, \dots, n$, for the multivariate *phase-type* distributions.

Univariate phase-type distributions (see Section 2 for the definition) have been widely used in queueing and reliability modeling (Asmussen (2003), Neuts (1981)), and in risk management and finance (Asmussen (2000), Rolski *et al.* (1999)). Multivariate phase-type distributions were introduced and studied in Assaf *et al.* (1984). The multivariate phase-type distributions include, as special cases, many well-known multivariate distributions, such as the Marshall–Olkin distribution (Marshall and Olkin (1967)), and retain many desirable properties similar to those in the univariate case. For example, the set of n -dimensional phase-type distributions is dense in the set of all distributions on $[0, \infty)^n$ and, hence, any nonnegative n -dimensional distribution can be approximated by n -dimensional phase-type distributions. Furthermore,

Kulkarni (1989) showed that the sum of random variables that have a multivariate phase-type distribution follows a (univariate) phase-type distribution.

Owing to their complex structure, however, the applications of multivariate phase-type distributions have been limited. Cai and Li (2005) employed Kulkarni's method and derived the explicit phase-type representation for the convolution, and applied the multivariate phase-type distributions to ruin theory in a multidimensional risk model. In this paper, we further utilize the underlying Markovian structure to explore the right-tail distributional properties of phase-type distributions that are relevant to explicit calculations of the CTE risk measures. Our method yields, in a unified fashion, explicit expressions for most of the above-mentioned CTE functions for phase-type distributions, and also gives some new distributional properties for multivariate phase-type distributions.

The paper is organized as follows. After giving a brief introduction to phase-type distributions, in Section 2 we discuss the CTE for univariate phase-type distributions. In Section 3, we detail various CTE measures that involve the multivariate phase-type distributions. In Section 4, we conclude the paper with some illustrative examples. Throughout the paper, we denote by $X \stackrel{=st}{=} Y$ the fact that two random variables X and Y are identically distributed. The vector \mathbf{e} denotes a column vector of 1s of appropriate dimension, and the vector $\mathbf{0}$ denotes a row vector of 0s of appropriate dimension. Note that the entries of all the probability vectors (subvectors) and matrices are indexed according to the state space of a Markov chain.

2. The CTE for univariate phase-type distributions

A nonnegative random variable X , or its distribution function F , is said to be of *phase type* (PH) with representation $(\boldsymbol{\alpha}, \mathbf{A}, d)$ if X is the time to absorption into the absorbing state 0 in a finite Markov chain $\{X(t), t \geq 0\}$ with state space $\{0, 1, \dots, d\}$, initial distribution $\boldsymbol{\beta} = (0, \boldsymbol{\alpha})$, and infinitesimal generator

$$\mathbf{Q} = \begin{bmatrix} 0 & \mathbf{0} \\ -\mathbf{A}\mathbf{e} & \mathbf{A} \end{bmatrix},$$

where subgenerator \mathbf{A} is a $d \times d$ nonsingular matrix and $\boldsymbol{\alpha} = (\alpha_1, \dots, \alpha_d)$. Thus, a nonnegative random variable X is of phase type with representation $(\boldsymbol{\alpha}, \mathbf{A}, d)$ if $X = \inf\{t \geq 0: X(t) = 0\}$, where $\{X(t), t \geq 0\}$ is the underlying Markov chain for X .

Let $\bar{F}(x) = 1 - F(x)$ denote the survival function. Then the random variable X is of phase type with representation $(\boldsymbol{\alpha}, \mathbf{A}, d)$ if and only if

$$\bar{F}(x) = \Pr\{X(x) \in \{1, \dots, d\}\} = \boldsymbol{\alpha}\mathbf{e}^{x\mathbf{A}}, \quad x \geq 0. \quad (2.1)$$

Thus,

$$\mathbb{E}X^k = \int_0^\infty x^k dF(x) = (-1)^k k! \boldsymbol{\alpha}\mathbf{A}^{-k}\mathbf{e}, \quad k = 1, 2, \dots \quad (2.2)$$

See, for example, Rolski *et al.* (1999) for details.

The CTE for a univariate phase-type distribution has an explicit expression, which follows from the following proposition.

Proposition 2.1. *If X has a PH distribution with representation $(\boldsymbol{\alpha}, \mathbf{A}, d)$ then, for any $t > 0$, the excess risk $(X - t \mid X > t)$ has a PH distribution with representation $(\boldsymbol{\alpha}_t, \mathbf{A}, d)$, where*

$$\boldsymbol{\alpha}_t = \frac{\boldsymbol{\alpha}\mathbf{e}^{t\mathbf{A}}}{\boldsymbol{\alpha}\mathbf{e}^{t\mathbf{A}}\mathbf{e}}. \quad (2.3)$$

Proof. The survival function of $(X - t \mid X > t)$ is given by

$$\Pr\{X - t > x \mid X > t\} = \Pr\{X > t + x \mid X > t\} = \frac{\bar{F}(t + x)}{\bar{F}(t)}.$$

It follows from (2.1) that

$$\Pr\{X - t > x \mid X > t\} = \frac{\alpha e^{(t+x)A} \mathbf{e}}{\alpha e^{tA} \mathbf{e}} = \alpha_t e^{xA} \mathbf{e}.$$

Hence, $(X - t \mid X > t)$ has a PH distribution with representation (α_t, A, d) .

Corollary 2.1. *If random variable X has a PH distribution with representation (α, A, d) then, for any $t > 0$,*

$$\text{CTE}_X(t) = t - \frac{\alpha A^{-1} e^{tA} \mathbf{e}}{\alpha e^{tA} \mathbf{e}}. \tag{2.4}$$

Proof. It follows from Proposition 2.1 and (2.2) that

$$\text{CTE}_X(t) = t + E(X - t \mid X > t) = t - \alpha_t A^{-1} \mathbf{e} = t - \frac{\alpha e^{tA} A^{-1} \mathbf{e}}{\alpha e^{tA} \mathbf{e}} = t - \frac{\alpha A^{-1} e^{tA} \mathbf{e}}{\alpha e^{tA} \mathbf{e}},$$

where the last equality holds due to the fact that $A^{-1} e^{tA} = e^{tA} A^{-1}$.

As one immediate application of Corollary 2.1, we can obtain an explicit expression for the conditional expectation $E(X \mid X \leq t)$, as follows.

Proposition 2.2. *If X has a PH distribution with representation (α, A, d) then, for any $t > 0$,*

$$E(X \mid X \leq t) = \frac{-\alpha A^{-1} \mathbf{e} - t \alpha e^{tA} \mathbf{e} + \alpha A^{-1} e^{tA} \mathbf{e}}{1 - \alpha e^{tA} \mathbf{e}}, \tag{2.5}$$

$$E(t - X \mid X \leq t) = \frac{t + \alpha A^{-1} (\mathbf{I} - e^{tA}) \mathbf{e}}{1 - \alpha e^{tA} \mathbf{e}}, \tag{2.6}$$

where \mathbf{I} is the $d \times d$ identity matrix.

Proof. Equation (2.5) follows from

$$E(X) = E(X \mid X > t) \Pr\{X > t\} + E(X \mid X \leq t) \Pr\{X \leq t\},$$

$E(X) = -\alpha A^{-1} \mathbf{e}$, $\Pr\{X \leq t\} = 1 - \alpha e^{tA} \mathbf{e}$, $\Pr\{X > t\} = \alpha e^{tA} \mathbf{e}$, and (2.4). Equation (2.6) is then obtained from $E(t - X \mid X \leq t) = t - E(X \mid X \leq t)$ and (2.5).

The conditional expectation $E(X \mid X \leq t)$ will be used later in the paper. In risk analysis, $E(t - X \mid X \leq t)$ describes the surplus beyond the risk that has been experienced. It is called the expected inactivity time in reliability theory.

3. The CTE for multivariate phase-type distributions

Let $\{X(t), t \geq 0\}$ be a right-continuous, continuous-time Markov chain on a finite state space \mathcal{E} with generator \mathbf{Q} . Let $\mathcal{E}_i, i = 1, \dots, n$, be n nonempty, stochastically closed subsets of \mathcal{E} such that $\bigcap_{i=1}^n \mathcal{E}_i$ is a proper subset of \mathcal{E} . (A subset of the state space is said to be stochastically closed if, once the process $\{X(t), t \geq 0\}$ enters it, it never leaves.)

We assume that absorption into $\bigcap_{i=1}^n \mathcal{E}_i$ is certain. Since we are interested in the process only until it is absorbed into $\bigcap_{i=1}^n \mathcal{E}_i$, we may assume, without loss of generality, that $\bigcap_{i=1}^n \mathcal{E}_i$ consists of one state, which we shall denote by Δ . Thus, without loss of generality, we may write $\mathcal{E} = \bigcup_{i=1}^n \mathcal{E}_i \cup \mathcal{E}_0$ for some subset $\mathcal{E}_0 \subset \mathcal{E}$ with $\mathcal{E}_0 \cap \mathcal{E}_j = \emptyset$ for $1 \leq j \leq n$. The states in \mathcal{E} are enumerated in such a way that Δ is the first element of \mathcal{E} . Thus, the generator of the chain has the form

$$Q = \begin{bmatrix} 0 & \mathbf{0} \\ -A\mathbf{e} & A \end{bmatrix}, \tag{3.1}$$

where subgenerator A is a $d \times d$ nonsingular matrix (as before), and $d = |\mathcal{E}| - 1$. Let $\boldsymbol{\beta} = (0, \boldsymbol{\alpha})$ be an initial probability vector on \mathcal{E} such that $\boldsymbol{\beta}(\Delta) = 0$.

We define

$$X_i = \inf\{t \geq 0: X(t) \in \mathcal{E}_i\}, \quad i = 1, \dots, n. \tag{3.2}$$

As in Assaf *et al.* (1984), for simplicity we shall assume that $\Pr\{X_1 > 0, \dots, X_n > 0\} = 1$, which means that the underlying Markov chain $\{X(t), t \geq 0\}$ starts within \mathcal{E}_0 almost surely. The joint distribution of (X_1, \dots, X_n) is called a *multivariate phase-type* (MPH) distribution with representation $(\boldsymbol{\alpha}, A, \mathcal{E}, \mathcal{E}_1, \dots, \mathcal{E}_n)$, and (X_1, \dots, X_n) is called a phase-type random vector.

When $n = 1$, the distribution of (3.2) reduces to the univariate PH distribution introduced in Neuts (1981) (see our Section 2). Examples of MPH distributions include, among many others, the well-known Marshall–Olkin distribution (Marshall and Olkin (1967)). The MPH distributions, their properties, and some related applications in reliability theory were discussed in Assaf *et al.* (1984). As in the univariate case, those MPH distributions (and their densities, Laplace transforms, and moments) can be written in closed form. The set of n -dimensional MPH distributions is dense in the set of all distributions on $[0, \infty)^n$. It was also shown in Assaf *et al.* (1984) and in Kulkarni (1989) that MPH distributions are closed under marginalization, finite mixture, convolution, and the formation of coherent systems.

The sum $S = X_1 + \dots + X_n$ and the extreme values $X_{(1)} = \min\{X_1, \dots, X_n\}$ and $X_{(n)} = \max\{X_1, \dots, X_n\}$ are all of phase type if (X_1, \dots, X_n) has a multivariate phase-type distribution. Thus, Corollary 2.1 will yield explicit expressions for the CTEs of $S, X_{(1)},$ and $X_{(n)}$ if we can find the phase-type representations of $S, X_{(1)},$ and $X_{(n)}$. In the following subsections, we will discuss these representations. We will also discuss the phase-type representation for the random vector $(X_{(1)}, X_i, X_{(n)}), i = 1, 2, \dots, n,$ and then obtain the related CTEs.

3.1. The CTE of sums

Cai and Li (2005) derived an explicit representation for the convolution distribution of $S = X_1 + \dots + X_n$. To state this result, we partition the state space as follows:

$$\begin{aligned} \Gamma_{\emptyset}^n &= \mathcal{E}_0, & \Gamma_i^{n-1} &= \mathcal{E}_i \setminus \bigcup_{k \neq i} (\mathcal{E}_i \cap \mathcal{E}_k), \quad i = 1, \dots, n, \\ \Gamma_{ij}^{n-2} &= (\mathcal{E}_i \cap \mathcal{E}_j) \setminus \bigcup_{\substack{k \neq i, \\ k \neq j}} (\mathcal{E}_i \cap \mathcal{E}_j \cap \mathcal{E}_k), \quad i \neq j, \end{aligned}$$

and so on.

For any $\mathcal{D} \subseteq \{1, \dots, n\}$, we also define

$$\Gamma_{\mathcal{D}}^{n-|\mathcal{D}|} = \left(\bigcap_{i \in \mathcal{D}} \mathcal{E}_i \right) \setminus \bigcup_{k \notin \mathcal{D}} \left(\bigcap_{i \in \mathcal{D}} \mathcal{E}_i \cap \mathcal{E}_k \right), \dots, \Gamma_{12\dots n}^0 = \{\Delta\},$$

where $|\cdot|$ denotes set cardinality.

In other words, $\Gamma_{\mathcal{D}}^{n-|\mathcal{D}|}$ contains the states only in \mathcal{E}_i for all $i \in \mathcal{D}$, but not in any other \mathcal{E}_j , $j \notin \mathcal{D}$. Note that these ‘ Γ ’s form a partition of \mathcal{E} . For each state $i \in \mathcal{E}$, define

$$k(i) = |\{j : i \notin \mathcal{E}_j, 1 \leq j \leq n\}|.$$

For example, $k(i) = n$ for all $i \in \Gamma_{\emptyset}^n$, $k(i) = 0$ for all $i \in \Gamma_{12\dots n}^0$, and, in general, $k(i) = n - |\mathcal{D}|$ for all $i \in \Gamma_{\mathcal{D}}^{n-|\mathcal{D}|}$.

Lemma 3.1. (Cai and Li (2005).) *Let (X_1, \dots, X_n) be a phase-type vector whose distribution has representation $(\alpha, A, \mathcal{E}, \mathcal{E}_1, \dots, \mathcal{E}_n)$, where $A = (a_{i,j})$. Then $\sum_{i=1}^n X_i$ has a phase-type distribution with representation $(\alpha, T, |\mathcal{E}| - 1)$, where $T = (t_{i,j})$ is given by*

$$t_{i,j} = \frac{a_{i,j}}{k(i)}; \tag{3.3}$$

that is, $t_{i,j} = a_{i,j}/m$ if $i \in \Gamma_{\mathcal{D}}^m$, for some $\mathcal{D} \subset \{1, \dots, n\}$.

Theorem 3.1. *Let (X_1, \dots, X_n) be a phase-type vector whose distribution has representation $(\alpha, A, \mathcal{E}, \mathcal{E}_1, \dots, \mathcal{E}_n)$. Then the CTE of $S = X_1 + \dots + X_n$ is given, for any $t > 0$, by*

$$\text{CTE}_S(t) = t - \frac{\alpha T^{-1} e^{tT} \mathbf{e}}{\alpha e^{tT} \mathbf{e}}, \tag{3.4}$$

where T is defined by (3.3).

Proof. Equation (3.4) follows from the phase-type representation $(\alpha, T, |\mathcal{E}| - 1)$ of S in Lemma 3.1 and (2.4).

3.2. The CTE of order statistics

Let $\bar{F}(x_1, \dots, x_n)$ and $F(x_1, \dots, x_n)$ respectively denote the joint survival and distribution functions of a phase-type random vector (X_1, \dots, X_n) . Then (see Assaf *et al.* (1984)), for $x_1 \geq x_2 \geq \dots \geq x_n \geq 0$, we have

$$\begin{aligned} \bar{F}(x_1, \dots, x_n) &= \Pr\{X_1 > x_1, \dots, X_n > x_n\} \\ &= \alpha e^{x_n A} \mathbf{g}_n e^{(x_{n-1} - x_n) A} \mathbf{g}_{n-1} \dots e^{(x_1 - x_2) A} \mathbf{g}_1 \mathbf{e}, \end{aligned} \tag{3.5}$$

$$\begin{aligned} F(x_1, \dots, x_n) &= \Pr\{X_1 \leq x_1, \dots, X_n \leq x_n\} \\ &= \beta e^{x_n Q} \mathbf{h}_n e^{(x_{n-1} - x_n) Q} \mathbf{h}_{n-1} \dots e^{(x_1 - x_2) Q} \mathbf{h}_1 \mathbf{e}, \end{aligned} \tag{3.6}$$

where, for $k = 1, \dots, n$, \mathbf{g}_k is defined as a diagonal $d \times d$ matrix whose i th diagonal element, for $i = 1, \dots, d$, equals 1 if $i \in \mathcal{E} \setminus \mathcal{E}_k$ and equals 0 otherwise, and \mathbf{h}_k is defined as a diagonal $(d + 1) \times (d + 1)$ matrix whose i th diagonal element, for $i = 1, \dots, d + 1$, equals 1 if $i \in \mathcal{E}_k$ and equals 0 otherwise.

For the matrix A in (3.1), we now introduce two Markov chains. Let $\mathcal{E} \setminus \{\Delta\} = S \cup S^*$, where $S \cap S^* = \emptyset$. The matrix Q in (3.1) can be partitioned as follows:

$$Q = \begin{bmatrix} 0 & \mathbf{0} & \mathbf{0} \\ -(A_S e + A_{SS^*} e) & A_S & A_{SS^*} \\ -(A_{S^*S} e + A_{S^*S^*} e) & A_{S^*S} & A_{S^*S^*} \end{bmatrix}. \tag{3.7}$$

Here A_S and A_{S^*} are the submatrices of A gained by removing the s th row and the s th column of A for all $s \notin S$ and $s \notin S^*$, respectively, and A_{SS^*} and A_{S^*S} are the submatrices of A gained by removing the s th row and the s^* th column of A for all $s \in S^*$, $s^* \in S$ and $s \in S$, $s^* \in S^*$, respectively. In particular, we have $A_{\mathcal{E} \setminus \{\Delta\}} = A$.

(I) The matrix

$$Q_S = \begin{bmatrix} 0 & \mathbf{0} \\ -A_S e & A_S \end{bmatrix}$$

is the generator of a Markov chain with state space $S \cup \{\Delta\}$ and absorbing state Δ . This Markov chain combines all the states of Q in S^* into the absorbing state Δ .

(II) Let $A_{[S]} = A_S + D(A_{SS^*})$, where $D(A_{SS^*})$ is the diagonal matrix whose s th diagonal entry is the s th entry of $A_{SS^*} e$. The matrix

$$Q_{[S]} = \begin{bmatrix} 0 & \mathbf{0} \\ -A_{[S]} e & A_{[S]} \end{bmatrix} \tag{3.8}$$

is the generator of another Markov chain with state space $S \cup \{\Delta\}$ and absorbing state Δ . This Markov chain removes all the transition rates of Q from $\mathcal{E} \setminus \{\Delta\}$ to S^* .

For any d -dimensional probability vector α and any subset $S \subseteq \mathcal{E} \setminus \{\Delta\}$, we denote by α_S the $|S|$ -dimensional subvector of α gained by removing its s th entry for all $s \notin S$. The vector $\mathbf{1}(S)$ (not to be confused with the identity matrix \mathbf{I}) denotes the column vector whose s th entry equals 1 if $s \in S$ and equals 0 otherwise. Furthermore, for any $S \subseteq \mathcal{E} \setminus \{\Delta\}$, we write $\alpha_t(S)$ for the following $|S|$ -dimensional row vector:

$$\alpha_t(S) = \frac{\alpha_S e^{tA_S}}{\alpha_S e^{tA_S} e}. \tag{3.9}$$

Note that $\alpha_t(\mathcal{E} \setminus \{\Delta\}) = \alpha_t$, where α_t is given by (2.3).

For any phase-type random vector (X_1, \dots, X_n) , Assaf *et al.* (1984) showed that the extreme values $X_{(1)} = \min\{X_1, \dots, X_n\}$ and $X_{(n)} = \max\{X_1, \dots, X_n\}$ are also of phase type. Their representations can be obtained from (3.5) and (3.6).

Lemma 3.2. *Let (X_1, \dots, X_n) be of phase type with representation $(\alpha, A, \mathcal{E}, \mathcal{E}_1, \dots, \mathcal{E}_n)$. Then*

- (i) $X_{(1)}$ is of phase type with representation $(\alpha_{\mathcal{E}_0}/\alpha_{\mathcal{E}_0} e, A_{\mathcal{E}_0}, |\mathcal{E}_0|)$, where $A_{\mathcal{E}_0}$ is defined as in (3.7), and
- (ii) $X_{(n)}$ is of phase type with representation $(\alpha, A, |\mathcal{E}| - 1)$.

Proof. By (3.5), the survival function of $X_{(1)}$ is given, for $x \geq 0$, by

$$\bar{F}_{X_{(1)}}(x) = \Pr\{X_{(1)} > x\} = \bar{F}(x, \dots, x) = \alpha e^{xA} \mathbf{g}_n \cdots \mathbf{g}_1 e.$$

Note that $\mathbf{g}_n \cdots \mathbf{g}_1 \mathbf{e} = \mathbf{I}(\mathcal{E}_0)$. Since the \mathcal{E}_i , $1 \leq i \leq n$, are all stochastically closed, we have

$$\bar{F}_{X_{(1)}}(x) = \boldsymbol{\alpha} e^{x\mathbf{A}} \mathbf{I}(\mathcal{E}_0) = \frac{\boldsymbol{\alpha}_{\mathcal{E}_0}}{\boldsymbol{\alpha}_{\mathcal{E}_0} \mathbf{e}} e^{x\mathbf{A}_{\mathcal{E}_0}} \mathbf{e},$$

which implies that $X_{(1)}$ is of phase type with representation $(\boldsymbol{\alpha}_{\mathcal{E}_0}/\boldsymbol{\alpha}_{\mathcal{E}_0} \mathbf{e}, \mathbf{A}_{\mathcal{E}_0}, |\mathcal{E}_0|)$.

Similarly, by (3.6), the distribution function of $X_{(n)}$ is given, for $x \geq 0$, by

$$F_{X_{(n)}}(x) = \Pr\{X_{(n)} \leq x\} = F(x, \dots, x) = \boldsymbol{\beta} e^{x\mathbf{Q}} \mathbf{h}_n \cdots \mathbf{h}_1 \mathbf{e} = \boldsymbol{\beta} e^{x\mathbf{Q}} \mathbf{I}(\{\Delta\}) = 1 - \boldsymbol{\alpha} e^{x\mathbf{A}} \mathbf{e}.$$

Thus, $X_{(n)}$ is of phase type with representation $(\boldsymbol{\alpha}, \mathbf{A}, |\mathcal{E}| - 1)$.

In fact, $X_{(1)}$ is the exit time of $\{X(t), t \geq 0\}$ from \mathcal{E}_0 and $X_{(n)}$ is the exit time of $\{X(t), t \geq 0\}$ from $\mathcal{E} \setminus \{\Delta\}$; thus, $X_{(1)}$ and $X_{(n)}$ are of phase type with representations given in Lemma 3.2. In general, the k th order statistic $X_{(k)}$ is the exit time of $\{X(t), t \geq 0\}$ from $\mathcal{E} \setminus \bigcup_{i_1, i_2, \dots, i_k} (\bigcap_{j=1}^k \mathcal{E}_{i_j})$ and, thus, is also of phase type. Hence, we obtain the following lemma.

Lemma 3.3. *Let $X_{(k)}$, $1 \leq k \leq n$, be the k th smallest component of (X_1, \dots, X_n) , which is of phase type with representation $(\boldsymbol{\alpha}, \mathbf{A}, \mathcal{E}, \mathcal{E}_1, \dots, \mathcal{E}_n)$. Then*

- (i) $(X_{(1)}, \dots, X_{(n)})$ is of phase type with representation $(\boldsymbol{\alpha}, \mathbf{A}, \mathcal{E}, \mathcal{O}_1, \dots, \mathcal{O}_n)$, where $\mathcal{O}_k = \bigcup_{i_1, i_2, \dots, i_k} (\bigcap_{j=1}^k \mathcal{E}_{i_j})$, $1 \leq k \leq n$.
- (ii) $X_{(k)}$ is of phase type with representation $(\boldsymbol{\alpha}_{\mathcal{E} \setminus \mathcal{O}_k}/\boldsymbol{\alpha}_{\mathcal{E} \setminus \mathcal{O}_k} \mathbf{e}, \mathbf{A}_{\mathcal{E} \setminus \mathcal{O}_k}, |\mathcal{E} \setminus \mathcal{O}_k|)$, $1 \leq k \leq n$.

Lemmas 3.2 and 3.3 and (2.4) immediately yield the CTEs of the extreme values $X_{(1)}$ and $X_{(n)}$, as follows.

Theorem 3.2. *Let (X_1, \dots, X_n) be of phase type with representation $(\boldsymbol{\alpha}, \mathbf{A}, \mathcal{E}, \mathcal{E}_1, \dots, \mathcal{E}_n)$.*

(i) *The excess loss $(X_{(1)} - t \mid X_{(1)} > t)$ of a subportfolio with the least risk is of phase type with representation $(\boldsymbol{\alpha}_t(\mathcal{E}_0), \mathbf{A}_{\mathcal{E}_0}, |\mathcal{E}_0|)$, and the CTE of $X_{(1)}$ is given by*

$$\text{CTE}_{X_{(1)}}(t) = t - \boldsymbol{\alpha}_t(\mathcal{E}_0) \mathbf{A}_{\mathcal{E}_0}^{-1} \mathbf{e}. \tag{3.10}$$

(ii) *The excess loss $(X_{(n)} - t \mid X_{(n)} > t)$ of the riskiest subportfolio is of phase type with representation $(\boldsymbol{\alpha}_t, \mathbf{A}, |\mathcal{E}| - 1)$, and the CTE of $X_{(n)}$ is given by*

$$\text{CTE}_{X_{(n)}}(t) = t - \boldsymbol{\alpha}_t \mathbf{A}^{-1} \mathbf{e}, \quad \text{where } \boldsymbol{\alpha}_t = \frac{\boldsymbol{\alpha} e^{t\mathbf{A}}}{\boldsymbol{\alpha} e^{t\mathbf{A}} \mathbf{e}}. \tag{3.11}$$

(iii) *In general, the excess loss $(X_{(k)} - t \mid X_{(k)} > t)$ of the $(n - k + 1)$ th riskiest subportfolio is of phase type with representation $(\boldsymbol{\alpha}_t(\mathcal{E} \setminus \mathcal{O}_k), \mathbf{A}_{\mathcal{E} \setminus \mathcal{O}_k}, |\mathcal{E} \setminus \mathcal{O}_k|)$, and the CTE of $X_{(k)}$ is given by*

$$\text{CTE}_{X_{(k)}}(t) = t - \boldsymbol{\alpha}_t(\mathcal{E} \setminus \mathcal{O}_k) \mathbf{A}_{\mathcal{E} \setminus \mathcal{O}_k}^{-1} \mathbf{e},$$

with $\mathcal{O}_k = \bigcup_{i_1, i_2, \dots, i_k} (\bigcap_{j=1}^k \mathcal{E}_{i_j})$, $1 \leq k \leq n$, as above.

3.3. CTEs involving different subportfolios

The Markovian method also leads to expressions for CTEs among different subportfolios. Hereafter, in a slight abuse of notation, for any d -dimensional probability vector α and any subset $S \subseteq \mathcal{E} \setminus \{\Delta\}$, we denote by $(\mathbf{0}, \alpha_S)$ the d -dimensional vector whose s th entry equals the s th entry of α if $s \in S$, and equals 0 otherwise. For example, $(\mathbf{0}, \alpha_t(S))$ denotes the d -dimensional vector whose s th entry is the s th entry of $\alpha_t(S)$ (see (3.9)) if $s \in S$, and vanishes otherwise.

Theorem 3.3. *If (X_1, \dots, X_n) is of phase type with representation $(\alpha, A, \mathcal{E}, \mathcal{E}_1, \dots, \mathcal{E}_n)$, then the random vector $((X_1 - t, \dots, X_n - t) \mid X_{(1)} > t)$ is of phase type with representation $((\mathbf{0}, \alpha_t(\mathcal{E}_0)), A, \mathcal{E}, \mathcal{E}_1, \dots, \mathcal{E}_n)$.*

Proof. Let $\{X(t), t \geq 0\}$ be the underlying Markov chain for (X_1, \dots, X_n) . Then

$$\{X_{(1)} > t\} = \{X_1 > t, \dots, X_n > t\} = \{X(t) \in \mathcal{E}_0\}.$$

It follows from the Markov property that

$$\begin{aligned} & ((X_1 - t, \dots, X_n - t) \mid X_{(1)} > t) \\ &= ((X_1 - t, \dots, X_n - t) \mid X(t) \in \mathcal{E}_0) \\ &=_{\text{st}} (\inf\{s > 0: X^*(s) \in \mathcal{E}_1\}, \dots, \inf\{s > 0: X^*(s) \in \mathcal{E}_n\}), \end{aligned}$$

where $\{X^*(s), s \geq 0\}$ is a Markov chain with the same state space and generator as $\{X(t), t \geq 0\}$, but with the initial probability vector $(\mathbf{0}, \alpha_t(\mathcal{E}_0))$ and

$$\alpha_t(\mathcal{E}_0) = \frac{1}{\Pr\{X(t) \in \mathcal{E}_0\}} (\Pr\{X(t) = i\}, i \in \mathcal{E}_0) = \frac{\alpha_{\mathcal{E}_0} e^{tA_{\mathcal{E}_0}}}{\alpha_{\mathcal{E}_0} e^{tA_{\mathcal{E}_0}} e}.$$

Thus, the random vector $((X_1 - t, \dots, X_n - t) \mid X_{(1)} > t)$ is of phase type with representation $((\mathbf{0}, \alpha_t(\mathcal{E}_0)), A, \mathcal{E}, \mathcal{E}_1, \dots, \mathcal{E}_n)$.

Hence, any marginal distribution of $((X_1 - t, \dots, X_n - t) \mid X_{(1)} > t)$ is also of phase type, and the risk contribution from the i th subportfolio, given that all the risks exceed a threshold value, can be calculated.

Corollary 3.1. *Let (X_1, \dots, X_n) be of phase type with representation $(\alpha, A, \mathcal{E}, \mathcal{E}_1, \dots, \mathcal{E}_n)$. Then the excess loss $(X_i - t \mid X_{(1)} > t)$ is of phase type with representation*

$$((\mathbf{0}, \alpha_t(\mathcal{E}_0)), A_{\mathcal{E} \setminus \mathcal{E}_i}, |\mathcal{E}| - |\mathcal{E}_i|)$$

and

$$\text{CTE}_{X_i \mid X_{(1)}}(t) = t - (\mathbf{0}, \alpha_t(\mathcal{E}_0)) A_{\mathcal{E} \setminus \mathcal{E}_i}^{-1} e. \tag{3.12}$$

Proof. By Theorem 3.3, $((X_1 - t, \dots, X_n - t) \mid X_{(1)} > t)$ has a phase-type representation $((\mathbf{0}, \alpha_t(\mathcal{E}_0)), A, \mathcal{E}, \mathcal{E}_1, \dots, \mathcal{E}_n)$. Hence, $(X_i - t \mid X_{(1)} > t)$ has a phase-type representation $((\mathbf{0}, \alpha_t(\mathcal{E}_0)), A_{\mathcal{E} \setminus \mathcal{E}_i}, |\mathcal{E}| - |\mathcal{E}_i|)$. Thus, (3.12) follows from (2.4).

In fact, from a direct calculation, we have

$$\text{CTE}_{X_i \mid X_{(1)}}(t) = E(X_i \mid X_{(1)} > t) = t + \frac{\int_t^\infty \bar{F}(t, \dots, t, x, t, \dots, t) dx}{\bar{F}(t, t, \dots, t)},$$

where, in the integrand, the i th argument of \bar{F} takes the value x . The random vector

$$(X_i, X_1, \dots, X_{i-1}, X_{i+1}, \dots, X_n)$$

has an MPH distribution with representation

$$(\alpha, A, \mathcal{E}, \mathcal{E}_i, \mathcal{E}_1, \dots, \mathcal{E}_{i-1}, \mathcal{E}_{i+1}, \dots, \mathcal{E}_n).$$

Therefore, by (3.5),

$$\begin{aligned} \bar{F}(t, \dots, t, x, t, \dots, t) &= \Pr\{X_i > x, X_1 > t, \dots, X_{i-1} > t, X_{i+1} > t, \dots, X_n > t\} \\ &= \alpha e^{tA} \prod_{k=1, k \neq i}^n g_k e^{(x-t)A} g_i e, \quad \text{for } x > t. \end{aligned}$$

Hence, for any $i, 1 \leq i \leq n$, we obtain

$$\begin{aligned} \text{CTE}_{X_i | X_{(1)}}(t) &= t + \frac{\alpha e^{tA} \prod_{k=1, k \neq i}^n g_k [\int_t^\infty e^{(x-t)A} dx] g_i e}{\alpha e^{tA} \prod_{k=1}^n g_k e} \\ &= t - \frac{\alpha e^{tA} \prod_{k=1, k \neq i}^n g_k A^{-1} g_i e}{\alpha e^{tA} \prod_{k=1}^n g_k e}. \end{aligned} \tag{3.13}$$

Expression (3.13) provides another formula for $\text{CTE}_{X_i | X_{(1)}}(t)$.

Note that (3.13) yields the same result as that of Corollary 3.1, due to the fact that

$$\begin{aligned} \frac{\alpha e^{tA} \prod_{k=1, k \neq i}^n g_k A^{-1} g_i e}{\alpha e^{tA} \prod_{k=1}^n g_k e} &= \frac{(\mathbf{0}, (\alpha_{\mathcal{E}_0} / \alpha_{\mathcal{E}_0} e) e^{tA \mathcal{E}_0}) A^{-1} g_i e}{(\alpha_{\mathcal{E}_0} / \alpha_{\mathcal{E}_0} e) e^{tA \mathcal{E}_0} e} = \frac{(\mathbf{0}, \alpha_{\mathcal{E}_0} e^{tA \mathcal{E}_0}) A_{\mathcal{E} \setminus \mathcal{E}_i}^{-1} e}{\alpha_{\mathcal{E}_0} e^{tA \mathcal{E}_0} e} \\ &= (\mathbf{0}, \alpha_t(\mathcal{E}_0)) A_{\mathcal{E} \setminus \mathcal{E}_i}^{-1} e. \end{aligned}$$

To develop a general scheme, we derive expressions for $E(X_i | X_k > t), i, k = 1, \dots, n$, where (X_1, \dots, X_n) is of phase type with representation $(\alpha, A, \mathcal{E}, \mathcal{E}_1, \dots, \mathcal{E}_n)$, and the underlying Markov chain $\{X(t), t \geq 0\}$.

Let $q_k = \Pr\{X_k > t\}$. Since \mathcal{E}_k is stochastically closed, we have

$$q_k = \Pr\{X(t) \in \mathcal{E} \setminus \mathcal{E}_k\} = \alpha e^{tA} \mathbf{I}(\mathcal{E} \setminus \mathcal{E}_k) = \frac{\alpha_{\mathcal{E} \setminus \mathcal{E}_k} e^{tA \mathcal{E} \setminus \mathcal{E}_k} e}{\alpha_{\mathcal{E} \setminus \mathcal{E}_k} e}$$

Also, for any $x \geq 0$, we have

$$\begin{aligned} \Pr\{X_i > x | X_k > t\} &= \frac{1}{q_k} \Pr\{X_i > x, X_k > t\} \\ &= \frac{1}{q_k} \Pr\{X_i > x, X(t) \in (\mathcal{E} \setminus (\mathcal{E}_i \cup \mathcal{E}_k)) \cup (\mathcal{E}_i \setminus (\mathcal{E}_i \cap \mathcal{E}_k))\} \\ &= p_{1,k} \Pr\{X_i > x | X(t) \in \mathcal{E} \setminus (\mathcal{E}_i \cup \mathcal{E}_k)\} \\ &\quad + p_{2,k} \Pr\{X_i > x | X(t) \in \mathcal{E}_i \setminus (\mathcal{E}_i \cap \mathcal{E}_k)\}, \end{aligned}$$

where

$$p_{1,k} = \frac{\Pr\{X(t) \in \mathcal{E} \setminus (\mathcal{E}_i \cup \mathcal{E}_k)\}}{q_k} \quad \text{and} \quad p_{2,k} = \frac{\Pr\{X(t) \in \mathcal{E}_i \setminus (\mathcal{E}_i \cap \mathcal{E}_k)\}}{q_k}.$$

Clearly $p_{1,k} + p_{2,k} = 1$. Thus,

$$E(X_i | X_k > t) = p_{1,k} E(X_i | X(t) \in \mathcal{E} \setminus (\mathcal{E}_i \cup \mathcal{E}_k)) + p_{2,k} E(X_i | X(t) \in \mathcal{E}_i \setminus (\mathcal{E}_i \cap \mathcal{E}_k)). \tag{3.14}$$

Since

$$\begin{aligned} \Pr\{X(t) \in \mathcal{E} \setminus (\mathcal{E}_i \cup \mathcal{E}_k)\} &= \frac{\alpha_{\mathcal{E} \setminus (\mathcal{E}_i \cup \mathcal{E}_k)} \exp(tA_{\mathcal{E} \setminus (\mathcal{E}_i \cup \mathcal{E}_k)})\mathbf{e}}{\alpha_{\mathcal{E} \setminus (\mathcal{E}_i \cup \mathcal{E}_k)}\mathbf{e}}, \\ \Pr\{X(t) \in \mathcal{E}_i \setminus (\mathcal{E}_i \cap \mathcal{E}_k)\} &= \alpha e^{tA} \mathbf{I}(\mathcal{E}_i \setminus (\mathcal{E}_i \cap \mathcal{E}_k)), \end{aligned}$$

we have

$$p_{1,k} = \left(\frac{\alpha_{\mathcal{E} \setminus \mathcal{E}_k} \mathbf{e}}{\alpha_{\mathcal{E} \setminus (\mathcal{E}_i \cup \mathcal{E}_k)} \mathbf{e}} \right) \left(\frac{\alpha_{\mathcal{E} \setminus (\mathcal{E}_i \cup \mathcal{E}_k)} \exp(tA_{\mathcal{E} \setminus (\mathcal{E}_i \cup \mathcal{E}_k)})\mathbf{e}}{\alpha_{\mathcal{E} \setminus \mathcal{E}_k} e^{tA_{\mathcal{E} \setminus \mathcal{E}_k}} \mathbf{e}} \right), \tag{3.15}$$

$$p_{2,k} = \frac{\alpha e^{tA} \mathbf{I}(\mathcal{E}_i \setminus (\mathcal{E}_i \cap \mathcal{E}_k))}{\alpha e^{tA} \mathbf{I}(\mathcal{E} \setminus \mathcal{E}_k)}. \tag{3.16}$$

To calculate the two conditional expectations in (3.14), we need the following lemma.

Lemma 3.4. *Let (X_1, \dots, X_n) have an MPH distribution with representation $(\alpha, A, \mathcal{E}, \mathcal{E}_1, \dots, \mathcal{E}_n)$.*

(i) *For any i and k , if $\Pr\{X(t) \in \mathcal{E} \setminus (\mathcal{E}_i \cup \mathcal{E}_k)\} > 0$ then*

$$E(X_i | X(t) \in \mathcal{E} \setminus (\mathcal{E}_i \cup \mathcal{E}_k)) = t - (\mathbf{0}, \alpha_t(\mathcal{E} \setminus (\mathcal{E}_i \cup \mathcal{E}_k))) A_{\mathcal{E} \setminus \mathcal{E}_i}^{-1} \mathbf{e}.$$

(ii) *For any i and k , $i \neq k$, if $\Pr\{X(t) \in \mathcal{E}_i \setminus (\mathcal{E}_i \cap \mathcal{E}_k)\} > 0$ then*

$$\begin{aligned} E(X_i | X(t) \in \mathcal{E}_i \setminus (\mathcal{E}_i \cap \mathcal{E}_k)) &= \frac{- (\mathbf{0}, \tilde{\alpha}) A_{[\mathcal{E} \setminus \mathcal{E}_k]}^{-1} \mathbf{e} - t (\mathbf{0}, \tilde{\alpha}) e^{tA_{[\mathcal{E} \setminus \mathcal{E}_k]}} \mathbf{e} + (\mathbf{0}, \tilde{\alpha}) A_{[\mathcal{E} \setminus \mathcal{E}_k]}^{-1} e^{tA_{[\mathcal{E} \setminus \mathcal{E}_k]}} \mathbf{e}}{1 - (\mathbf{0}, \tilde{\alpha}) e^{tA_{[\mathcal{E} \setminus \mathcal{E}_k]}} \mathbf{e}}, \end{aligned} \tag{3.17}$$

where $\tilde{\alpha} = \alpha_{\mathcal{E} \setminus (\mathcal{E}_i \cup \mathcal{E}_k)} / \alpha_{\mathcal{E} \setminus (\mathcal{E}_i \cup \mathcal{E}_k)} \mathbf{e}$.

Proof. We first calculate $E(X_i | X(t) \in \mathcal{E} \setminus (\mathcal{E}_i \cup \mathcal{E}_k))$. It follows from the Markov property that

$$E(X_i | X(t) \in \mathcal{E} \setminus (\mathcal{E}_i \cup \mathcal{E}_k)) = t + E(\inf\{s > 0: X^*(s) = \Delta\}),$$

where $\{X^*(t), t \geq 0\}$ is a Markov chain with state space $(\mathcal{E} \setminus \mathcal{E}_i) \cup \{\Delta\}$, subgenerator $A_{\mathcal{E} \setminus \mathcal{E}_i}$, and initial probability vector $(\mathbf{0}, \alpha_t(\mathcal{E} \setminus (\mathcal{E}_i \cup \mathcal{E}_k)))$ with

$$\alpha_t(\mathcal{E} \setminus (\mathcal{E}_i \cup \mathcal{E}_k)) = \frac{\Pr\{X(t) = j, j \in \mathcal{E} \setminus (\mathcal{E}_i \cup \mathcal{E}_k)\}}{\Pr\{X(t) \in \mathcal{E} \setminus (\mathcal{E}_i \cup \mathcal{E}_k)\}} = \frac{\alpha_{\mathcal{E} \setminus (\mathcal{E}_i \cup \mathcal{E}_k)} \exp(tA_{\mathcal{E} \setminus (\mathcal{E}_i \cup \mathcal{E}_k)})\mathbf{e}}{\alpha_{\mathcal{E} \setminus (\mathcal{E}_i \cup \mathcal{E}_k)} \exp(tA_{\mathcal{E} \setminus (\mathcal{E}_i \cup \mathcal{E}_k)})\mathbf{e}}.$$

Since $\inf\{s > 0: X^*(s) = \Delta\}$ is of phase type, we have

$$E(X_i | X(t) \in \mathcal{E} \setminus (\mathcal{E}_i \cup \mathcal{E}_k)) = t - (\mathbf{0}, \alpha_t(\mathcal{E} \setminus (\mathcal{E}_i \cup \mathcal{E}_k))) A_{\mathcal{E} \setminus \mathcal{E}_i}^{-1} \mathbf{e}.$$

To calculate $E(X_i | X(t) \in \mathcal{E}_i \setminus (\mathcal{E}_i \cap \mathcal{E}_k))$, consider

$$\{X(t) \in \mathcal{E}_i \setminus (\mathcal{E}_i \cap \mathcal{E}_k)\} = \{X_i \leq t < X_k\}.$$

We now define a new Markov chain $\{\tilde{X}(t), t \geq 0\}$ with state space $\mathcal{E} \setminus \mathcal{E}_k$, as follows. The set of absorbing states is $\mathcal{E}_i \setminus (\mathcal{E}_i \cap \mathcal{E}_k)$, the initial probability vector of $\{\tilde{X}(t), t \geq 0\}$ is $(\mathbf{0}, \tilde{\alpha})$, where $\tilde{\alpha} = \alpha_{\mathcal{E} \setminus (\mathcal{E}_i \cup \mathcal{E}_k)} / \alpha_{\mathcal{E} \setminus (\mathcal{E}_i \cup \mathcal{E}_k)} \mathbf{e}$, and the generator of $\{\tilde{X}(t), t \geq 0\}$ is given by $A_{[\mathcal{E} \setminus \mathcal{E}_k]}$ (see (3.8)). Let

$$\tilde{X}_i = \inf\{s > 0: \tilde{X}(s) \in \mathcal{E}_i \setminus (\mathcal{E}_i \cap \mathcal{E}_k)\},$$

which has a phase-type distribution. Then, from (2.5),

$$\begin{aligned} E(X_i | X(t) \in \mathcal{E}_i \setminus (\mathcal{E}_i \cap \mathcal{E}_k)) &= E(\tilde{X}_i | \tilde{X}_i \leq t) \\ &= \frac{-(\mathbf{0}, \tilde{\alpha})A_{[\mathcal{E} \setminus \mathcal{E}_k]}^{-1} \mathbf{e} - t(\mathbf{0}, \tilde{\alpha})e^{tA_{[\mathcal{E} \setminus \mathcal{E}_k]}} \mathbf{e} + (\mathbf{0}, \tilde{\alpha})A_{[\mathcal{E} \setminus \mathcal{E}_k]}^{-1} e^{tA_{[\mathcal{E} \setminus \mathcal{E}_k]}} \mathbf{e}}{1 - (\mathbf{0}, \tilde{\alpha})e^{tA_{[\mathcal{E} \setminus \mathcal{E}_k]}} \mathbf{e}}. \end{aligned}$$

Thus, (3.17) holds.

Observe that if $\mathcal{E}_i \subseteq \mathcal{E}_k$, then $X_i \geq X_k$ almost surely. It follows from (3.14) that

$$E(X_i | X_k > t) = E(X_i | X(t) \in \mathcal{E} \setminus \mathcal{E}_k) \quad \text{if } \mathcal{E}_i \subseteq \mathcal{E}_k. \tag{3.18}$$

Thus, Lemma 3.4(i) implies that $(X_i - t | X_k > t)$ is of phase type with representation $((\mathbf{0}, \alpha_t(\mathcal{E} \setminus \mathcal{E}_k)), A_{\mathcal{E} \setminus \mathcal{E}_k}, |\mathcal{E} \setminus \mathcal{E}_k|)$. This leads to the following corollaries.

Corollary 3.2. *Let (X_1, \dots, X_n) have an MPH distribution with representation $(\alpha, A, \mathcal{E}, \mathcal{E}_1, \dots, \mathcal{E}_n)$. Let $X_{(k)}, 1 \leq k \leq n$, be the k th order statistic of (X_1, \dots, X_n) .*

(i) *For any i and $k, k \leq i, (X_{(i)} - t | X_{(k)} > t)$ is of phase type with representation*

$$((\mathbf{0}, \alpha_t(\mathcal{E} \setminus \mathcal{O}_k)), A_{\mathcal{E} \setminus \mathcal{O}_k}, |\mathcal{E} \setminus \mathcal{O}_k|),$$

and

$$E(X_{(i)} | X_{(k)} > t) = t - (\mathbf{0}, \alpha_t(\mathcal{E} \setminus \mathcal{O}_k))A_{\mathcal{E} \setminus \mathcal{O}_k}^{-1} \mathbf{e}.$$

(ii) *In particular, $(X_{(n)} - t | X_{(1)} > t)$ is of phase type with representation*

$$((\mathbf{0}, \alpha_t(\mathcal{E}_0)), A, |\mathcal{E}| - 1),$$

and

$$E(X_{(n)} | X_{(1)} > t) = t - (\mathbf{0}, \alpha_t(\mathcal{E}_0))A^{-1} \mathbf{e}. \tag{3.19}$$

Proof. It follows from Lemma 3.3 that $(X_{(k)}, X_{(i)})$ is of phase type with representation $(\alpha, A, \mathcal{E}, \mathcal{O}_k, \mathcal{O}_i)$. Since $X_{(i)} \geq X_{(k)}$, the corollary follows from Lemma 3.4(i), $A_{\mathcal{E} \setminus \{\Delta\}} = A$, and $|\mathcal{E} \setminus \{\Delta\}| = |\mathcal{E}| - 1$.

Corollary 3.3. *Let (X_1, \dots, X_n) have an MPH distribution with representation $(\alpha, A, \mathcal{E}, \mathcal{E}_1, \dots, \mathcal{E}_n)$. Then $(X_{(n)} - t | X_i > t)$ is of phase type with representation $((\mathbf{0}, \alpha_t(\mathcal{E} \setminus \mathcal{E}_i)), A, |\mathcal{E}| - 1)$, and*

$$E(X_{(n)} | X_i > t) = t - (\mathbf{0}, \alpha_t(\mathcal{E} \setminus \mathcal{E}_i))A^{-1} \mathbf{e}. \tag{3.20}$$

Proof. If (X_1, \dots, X_n) is MPH distributed with representation $(\alpha, A, \mathcal{E}_1, \dots, \mathcal{E}_n)$, then $(X_i, X_{(n)})$ is of phase type with representation $(\alpha, A, \mathcal{E}, \mathcal{E}_i, \{\Delta\})$. The results of the corollary then follow from (3.18), Lemma 3.4, and the fact that $X_{(n)} \geq X_i$ almost surely for any i .

The expression for $E(X_i | X_{(n)} > t)$ is cumbersome, but can be obtained from (3.14), (3.15), (3.16), and Lemma 3.4.

4. The CTE of Marshall–Olkin distributions

In this section, we illustrate our results using the multivariate Marshall–Olkin distribution, and also show some interesting effects on the CTEs of choosing different model parameters.

Let $\{E_S, S \subseteq \{1, \dots, n\}\}$ be a sequence of independent, exponentially distributed random variables, with E_S having mean $1/\lambda_S$. Let

$$X_j = \min\{E_S : S \ni j\}, \quad j = 1, \dots, n.$$

The joint distribution of (X_1, \dots, X_n) is called the Marshall–Olkin distribution with parameters $\{\lambda_S, S \subseteq \{1, \dots, n\}\}$ (Marshall and Olkin (1967)). In the reliability context, X_1, \dots, X_n can be viewed as the lifetimes of n components operating in a random shock environment, where a fatal shock governed by a Poisson process $\{N_S(t), t \geq 0\}$ with rate λ_S destroys all the components with indices in $S \subseteq \{1, \dots, n\}$ simultaneously. Assuming that these Poisson shock arrival processes are independent, we then have

$$X_j = \inf\{t : N_S(t) \geq 1, S \ni j\}, \quad j = 1, \dots, n. \tag{4.1}$$

Let $\{\mathcal{M}_S(t), t \geq 0\}, S \subseteq \{1, \dots, n\}$, be independent Markov chains with absorbing state Δ_S , each representing the exponential distribution with parameter λ_S . It follows from (4.1) that (X_1, \dots, X_n) is of phase type, and has the underlying Markov chain, on the product space of these independent Markov chains, with absorbing classes $\mathcal{E}_j = \{(e_S) : e_S = \Delta_S \text{ for some } S \ni j\}, 1 \leq j \leq n$. It is also easy to verify that the marginal distribution of the j th component of the Marshall–Olkin-distributed random vector is exponential with mean $1/\sum_{\{S : S \ni j\}} \lambda_S$.

To calculate the CTEs, we need to simplify the underlying Markov chain for the Marshall–Olkin distribution and obtain its phase-type representation. Let $\{X(t), t \geq 0\}$ be a Markov chain with state space $\mathcal{E} = \{S : S \subseteq \{1, \dots, n\}\} = \{\Delta, e_1, \dots, e_d\}$, starting at \emptyset almost surely. The index set $\{1, \dots, n\}$ is the absorbing state Δ , $\mathcal{E}_0 = \{\emptyset\}$, and $\mathcal{E}_j = \{S : S \ni j\}, j = 1, \dots, n$.

It follows from (4.1) that its subgenerator is given by $A = (a_{i,j})$, where

$$a_{i,j} = \sum_{\{L : L \subseteq S^*, L \cup S = S^*\}} \lambda_L \quad \text{if } e_i = S, e_j = S^*, \text{ and } S \subset S^*,$$

$$a_{i,i} = \sum_{\{L : L \subseteq S\}} \lambda_L - \Lambda \quad \text{if } e_i = S \text{ and } \Lambda = \sum_S \lambda_S,$$

and the other entries of A vanish. Using the results in Sections 2–3 and these parameters, we can calculate the CTEs. To illustrate the results, we consider the bivariate case.

Example 4.1. In the Marshall–Olkin distribution, let $n = 2$, let $\mathcal{E} = \{12, 2, 1, \emptyset\}$, and let $\mathcal{E}_j = \{12, j\}, j = 1, 2$, where 12 is the absorbing state Δ . Furthermore, let the initial probability vector be $(0, \alpha)$ with $\alpha = (0, 0, 1)$. Then the subgenerator A for the two-dimensional Marshall–Olkin distribution is given by

$$A = \begin{bmatrix} -\lambda_{12} - \lambda_1 & 0 & 0 \\ 0 & -\lambda_{12} - \lambda_2 & 0 \\ \lambda_2 & \lambda_1 & -\Lambda + \lambda_\emptyset \end{bmatrix},$$

where $\Lambda = \lambda_{12} + \lambda_2 + \lambda_1 + \lambda_\emptyset$.

To study the effect of dependence on the CTEs, we calculate $CTE_S(t), CTE_{X_{(1)}}(t)$, and $CTE_{X_{(n)}}(t)$, respectively, under several different sets of model parameters. The analytic forms

TABLE 1: Effects of dependence on the CTEs of S , $X_{(1)}$, and $X_{(n)}$.

t	CTE $_S(t)$			CTE $_{X_{(1)}}(t)$			CTE $_{X_{(n)}}(t)$		
	Case 1	Case 2	Case 3	Case 1	Case 2	Case 3	Case 1	Case 2	Case 3
2	2.4667	2.5381	2.8	2.2	2.25	2.4	2.4007	2.4038	2.4
4	4.4364	4.5113	4.8	4.2	4.25	4.4	4.4000	4.4002	4.4
6	6.4250	6.5039	6.8	6.2	6.25	6.4	6.4000	6.4000	6.4
8	8.4191	8.5014	8.8	8.2	8.25	8.4	8.4000	8.4000	8.4
10	10.4154	10.5005	10.8	10.2	10.25	10.4	10.4000	10.4000	10.4

of CTE $_S(t)$, CTE $_{X_{(1)}}(t)$, and CTE $_{X_{(n)}}(t)$ in the following three cases, and the numerical values in Table 1, were easily produced from (3.4), (3.10), and (3.11) using MATHEMATICA[®]. The first column of Table 1 lists several values of t , and the next several columns list values of these CTEs in the following three cases.

Case 1: $\lambda_{12} = 0, \lambda_1 = \lambda_2 = 2.5, \lambda_{\emptyset} = 0$. In this case, X_1 and X_2 are independent, and

$$\begin{aligned} \text{CTE}_S(t) &= 0.4 + t + \frac{0.16}{0.4 + t}, \\ \text{CTE}_{X_{(1)}}(t) &= 0.2 + t, \\ \text{CTE}_{X_{(n)}}(t) &= \frac{0.8 + 2t - (0.2 + t)e^{-2.5t}}{2 - e^{-2.5t}}. \end{aligned}$$

Case 2: $\lambda_{12} = 1, \lambda_1 = \lambda_2 = 1.5, \lambda_{\emptyset} = 1$. In this case, X_1 and X_2 are positively dependent, and

$$\begin{aligned} \text{CTE}_S(t) &= \frac{1 + 2t - (0.6 + 1.5t)e^{-0.5t}}{2 - 1.5e^{-0.5t}}, \\ \text{CTE}_{X_{(1)}}(t) &= 0.25 + t, \\ \text{CTE}_{X_{(n)}}(t) &= \frac{0.8 + 2t - (0.25 + t)e^{-1.5t}}{2 - e^{-1.5t}}. \end{aligned}$$

Case 3: $\lambda_{12} = 2.5, \lambda_1 = \lambda_2 = 0, \lambda_{\emptyset} = 2.5$. This is the comonotone case in which $X_1 = X_2$ and, so, the vector (X_1, X_2) has the strongest positive dependence. In this case,

$$\text{CTE}_S(t) = 0.8 + t, \quad \text{CTE}_{X_{(1)}}(t) = \text{CTE}_{X_{(n)}}(t) = 0.4 + t.$$

In all the three cases, (X_1, X_2) has the same marginal distributions; that is, X_1 and X_2 have the exponential distributions with means $1/(\lambda_{12} + \lambda_1)$ and $1/(\lambda_{12} + \lambda_2)$, respectively. The only difference among them is in the correlation between X_1 and X_2 . It can be easily verified directly that the correlation coefficient of (X_1, X_2) in Case 1 is smaller than that in Case 2, which in turn is smaller than that in Case 3. In fact, it follows from Proposition 5.5 of Li and Xu (2000) that the random vector in Case 1 is less dependent than that in Case 2, which in turn is less dependent than that in Case 3, all in *supermodular dependence order*.

Table 1 shows that CTE $_S(t)$ becomes larger as the correlation grows. The effect of dependence on CTE $_{X_{(1)}}(t)$ is the same as it is on CTE $_S(t)$. However, the effect of dependence on CTE $_{X_{(n)}}(t)$ is different from its effect on CTE $_S(t)$ and CTE $_{X_{(1)}}(t)$. Indeed, CTE $_{X_{(n)}}(t)$ is neither increasing nor decreasing as the correlation grows.

TABLE 2: Effects of dependence and different conditions on the maximal risk.

t	$E(X_{(n)} X_{(1)} > t)$			$E(X_{(n)} X_1 > t)$			$E(X_{(n)} X_{(n)} > t)$		
	Case 1	Case 2	Case 3	Case 1	Case 2	Case 3	Case 1	Case 2	Case 3
2	2.6	2.55	2.4	2.4014	2.4075	2.4	2.4007	2.4038	2.4
4	4.6	4.55	4.4	4.4000	4.4004	4.4	4.4000	4.4002	4.4
6	6.6	6.55	6.4	6.4000	6.4000	6.4	6.4000	6.4000	6.4
8	8.6	8.55	8.4	8.4000	8.4000	8.4	8.4000	8.4000	8.4
10	10.6	10.55	10.4	10.4000	10.4000	10.4	10.4000	10.4000	10.4

Example 4.2. We now calculate $E(X_{(n)} | X_{(1)} > t)$, $E(X_{(n)} | X_1 > t)$, and $CTE_{X_{(n)}}(t) = E(X_{(n)} | X_{(n)} > t)$ under the same sets of model parameters used in Example 4.1.

The analytic forms of $E(X_{(n)} | X_{(1)} > t)$, $E(X_{(n)} | X_1 > t)$, and $E(X_{(n)} | X_{(n)} > t)$ in the following three cases, and the numerical values in Table 2, were easily produced from (3.19), (3.20), and (3.11) using MATHEMATICA. The first column of Table 2 lists several values of t , and the next several columns list values of the conditional expectations in the following three cases (corresponding to those in Example 4.1).

Case 1: $\lambda_{12} = 0, \lambda_1 = \lambda_2 = 2.5, \lambda_{\emptyset} = 0$. In this case,

$$\begin{aligned}
 E(X_{(n)} | X_{(1)} > t) &= 0.6 + t, \\
 E(X_{(n)} | X_1 > t) &= 0.4 + t + 0.2e^{-2.5t}, \\
 E(X_{(n)} | X_{(n)} > t) &= \frac{0.8 + 2t - (0.2 + t)e^{-2.5t}}{2 - e^{-2.5t}}.
 \end{aligned}$$

Case 2: $\lambda_{12} = 1, \lambda_1 = \lambda_2 = 1.5, \lambda_{\emptyset} = 1$. In this case,

$$\begin{aligned}
 E(X_{(n)} | X_{(1)} > t) &= 0.55 + t, \\
 E(X_{(n)} | X_1 > t) &= 0.4 + t + 0.15e^{-1.5t}, \\
 E(X_{(n)} | X_{(n)} > t) &= \frac{0.8 + 2t - (0.25 + t)e^{-1.5t}}{2 - e^{-1.5t}}.
 \end{aligned}$$

Case 3: $\lambda_{12} = 2.5, \lambda_1 = \lambda_2 = 0, \lambda_{\emptyset} = 2.5$. In this case,

$$E(X_{(n)} | X_{(1)} > t) = E(X_{(n)} | X_1 > t) = E(X_{(n)} | X_{(n)} > t) = 0.4 + t.$$

Table 2 shows that $E(X_{(n)} | X_{(1)} > t) \geq E(X_{(n)} | X_1 > t) \geq E(X_{(n)} | X_{(n)} > t)$ for all values of t in all three cases. However, neither $E(X_{(n)} | X_1 > t)$ nor $E(X_{(n)} | X_{(n)} > t)$ exhibits any monotonicity property as the correlation grows.

5. Concluding remarks

Using the Markovian method, we have derived explicit expressions for various conditional tail expectations for multivariate phase-type distributions in a unified fashion. These CTEs can be used, for example, to measure some right-tail risks in a financial portfolio consisting of several stochastically dependent subportfolios. We have focused on the total risk, the minimal risk, and the maximal risk of the portfolio and, as we have illustrated in the numerical examples, our CTE formulae for these risks can be easily implemented.

Some CTE function values in our numerical examples increase as the correlation among the subportfolios grows. This demonstrates that merely increasing the correlation, while fixing the marginal risk for each subportfolio, would add more risk to the entire portfolio. In fact, since the minimum statistic $X_{(1)}$ of a Marshall–Olkin-distributed random vector (X_1, \dots, X_n) has an exponential distribution, it is easy to verify directly that the CTE of $X_{(1)}$ is increasing as (X_1, \dots, X_n) becomes more dependent, in the sense of supermodular dependence order. However, whether or not a given CTE risk measure exhibits monotonicity as the correlation among the subportfolios grows remains an open and important question that needs further study.

Another problem that we have not addressed in this paper is the explicit expression for the CTE of $(X_i \mid \sum_{i=1}^n X_i > t)$, which is of interest in the risk allocation study of the total risk. Using the Markovian method, we can obtain the expression for the CTE of $(X_i \mid \sum_{i=1}^n X_i > t)$, but it is too cumbersome to have any value in practical terms. Alternatively, some kind of recursive algorithm defined on the underlying Markov structure would offer a promising computational approach.

Acknowledgements

We thank the referee for a careful reading of the paper and helpful suggestions that improved the presentation. This research was supported in part by the NSERC grant RGPIN 250031-02.

References

- ARTZNER, P., DELBAEN, F., EBER, J.-M. AND HEATH, D. (1999). Coherent measures of risks. *Math. Finance* **9**, 203–228.
- ASMUSSEN, S. (2000). *Ruin Probabilities*. World Scientific, Singapore.
- ASMUSSEN, S. (2003). *Applied Probability and Queues*, 2nd edn. Springer, New York.
- ASSAF, D., LANGBERG, N., SAVITS, T. H. AND SHAKED, M. (1984). Multivariate phase-type distributions. *Operat. Res.* **32**, 688–702.
- BOWERS, N. L. *et al.* (1997). *Actuarial Mathematics*. The Society of Actuaries, Schaumburg, IL.
- CAI, J. AND LI, H. (2005). Multivariate risk model of phase type. *Insurance Math. Econom.* **36**, 137–152.
- EMBRECHTS, P., KLÜPPELBERG, C. AND MIKOSCH, T. (1997). *Modelling Extremal Events for Insurance and Finance*. Springer, Berlin.
- KULKARNI, V. G. (1989). A new class of multivariate phase type distributions. *Operat. Res.* **37**, 151–158.
- LANDSMAN, Z. AND VALDEZ, E. (2003). Tail conditional expectations for elliptical distributions. *N. Amer. Actuarial J.* **7**, 55–71.
- LI, H. AND XU, S. H. (2000). On the dependence structure and bounds of correlated parallel queues and their applications to synchronized stochastic systems. *J. Appl. Prob.* **37**, 1020–1043.
- MARSHALL, A. W. AND OLKIN, I. (1967). A multivariate exponential distribution. *J. Amer. Statist. Assoc.* **2**, 84–98.
- NEUTS, M. F. (1981). *Matrix-Geometric Solutions in Stochastic Models. An Algorithmic Approach*. The Johns Hopkins University Press, Baltimore, MD.
- ROLSKI, T., SCHMIDL, H., SCHMIDT, V. AND TEUGELS, J. (1999). *Stochastic Processes for Finance and Insurance*. John Wiley, New York.
- SHAKED, M. AND SHANTHIKUMAR, J. G. (1994). *Stochastic Orders and Their Applications*. Academic Press, New York.