Conditional Tail Expectations for Multivariate Phase Type Distributions

Jun Cai
Department of Statistics and Actuarial Science
University of Waterloo
Waterloo, ON N2L 3G1, Canada
Telphone: 1-519-8884567, ext. 6990
Fax: 1-519-7461875
Email: jcai@uwaterloo.ca
and
Haijun Li
Department of Mathematics
Washington State University
Pullman, WA 99164, U.S.A.
Telephone: 1-509-3355279
Fax : 1-509-3351188
Email: lih@math.wsu.edu

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 for 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 formulas of 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 reveals structural insight, but also yields some new distributional properties for multivariate phase type distributions.

Keywords: Univariate and multivariate phase type distributions, value at risk (VaR), conditional tail expectation (CTE), excess loss, convolution, extreme value, Marshall-Olkin distribution.

1 Introduction

Let risk X be a non-negative random variable with cumulative distribution F, where X may refer to a claim for an insurance company or a loss on an investment portfolio. Given $0 < q < 1$, the value x_q, determined by $\bar{F}(x_q) = 1 - F(x_q) = q$ and denoted by $VaR_X(1-q)$, is called the *value at risk* (VaR) with a degree of confidence of $1-q$. The conditional expectation of X given that $X > x_q$, denoted by $CTE_X(x_q) = E(X|X > x_q)$, is called the *conditional tail expectation* (CTE) of X at VaR x_q. Observe that

$$CTE_X(x_q) = x_q + E(X - x_q|X > x_q),$$

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), respectively. Since $\frac{d}{dx}(x + E(X - x|X > x)) \geq 0$ for any continuous risk X (see page 45 of Shaked and Shanthikumar 1994), the CTE function $CTE_X(x)$ is increasing in $x \geq 0$, or equivalently, $CTE_X(x_q)$ is decreasing in $q \in (0, 1)$.

Both VaR and CTE are important measures of right-tail risks, which are frequently encountered in the insurance and financial investment. It is known that the CTE satisfies all the desirable properties of a *coherent risk measure* (Artzner et al. 1999), and that the CTE provides a more conservative measure of risk than VaR for the same level of degree of confidence (Landsman and Valdez 2003). Hence, the CTE is more preferable than the VaR in many applications, and has recently received growing attentions in the insurance and finance literature.

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

$$
\begin{aligned}
CTE_S(t) &= E(S \mid S > t), & (1.1)\\
CTE_{X_{(1)}}(t) &= E(X_{(1)} \mid X_{(1)} > t), & (1.2)\\
CTE_{X_{(n)}}(t) &= E(X_{(n)} \mid X_{(n)} > t). & (1.3)
\end{aligned}
$$

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

$$
\begin{aligned}
CTE_{X_i|S}(t) &= E(X_i \mid S > t), & (1.4)\\
CTE_{X_i|X_{(1)}}(t) &= E(X_i \mid X_{(1)} > t), & (1.5)\\
CTE_{X_i|X_{(n)}}(t) &= E(X_i \mid X_{(n)} > t), & (1.6)
\end{aligned}
$$

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

All these risk measures have physical interpretations in insurance, finance and other fields. For instance, $CTE_{X_i|S}(t)$ represents the contribution of the i-th risk X_i to the aggregate risk S since $CTE_S(t) = \sum_{i=1}^n CTE_{X_i|S}(t)$. The CTE $E(X_{(1)}|X_{(1)} > t)$ ($E(X_{(n)}|X_{(n)} > t)$) describes the expected minimal (maximal) risk in all the subportfolios given that the minimal (maximal) risk exceeds some threshold t. More interestingly, $E(X_i|X_{(1)} > t)$ represents the average contribution of the i-th risk given that all the risks exceed some value t, whereas $E(X_i|X_{(n)} > t)$ 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, these CTEs in (1.2)-(1.6) also have interpretations in life insurance. For instance, in a group life insurance, let X_i, $1 \le i \le n$, 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, $E(X_i|X_{(1)} > t)$ is the expected lifetime of member i given that all members are alive at time t, and $E(X_i|X_{(n)} > t)$ is the expected lifetime of member i given that there is at least one member who is alive at time t.

Landsman and Valdez (2003) obtained the explicit formulas of $CTE_S(t)$ and $CTE_{X_i|S}(t)$ for the multivariate elliptical distributions, which include the distributions such as multivariate normal, stable, student-t, etc. The focus of this paper is to derive the explicit formulas of various CTEs, such as $CTE_S(t)$, $CTE_{X_{(1)}}(t)$, $CTE_{X_{(n)}}(t)$, $E(X_{(n)}|X_{(1)} > t)$, $E(X_{(n)}|X_i > t)$ for $i = 1, 2, ..., n$, 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 (Neuts 1981), and in risk management and finance (Asmussen 2000 and Rolski et al. 1999). The phase type distributions have many attractive properties, and in particular, they are mathematically tractable and dense in all distributions on $[0, \infty)$. Any nonnegative distribution can be approximated by phase type distributions, and as such, the phase type distributions provide a powerful and versatile tool in probabilistic modeling. See, for example, Asmussen (2003) and Neuts (1981) for detailed discussions on these distributions and their applications.

Multivariate phase type distributions have been 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 also 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 the random variables that have a multivariate phase type distribution follows a (univariate) phase type distribution.

Due 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 multi-dimensional 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, in a unified fashion, yields the explicit expressions for most of all above-mentioned CTE functions for phase type distributions, and it also gives some new distributional properties for multivariate phase type distributions.

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

2 CTE for Univariate Phase Type Distributions

A non-negative random variable X or its distribution F is said to be of *phase type* (PH) with representation (α, 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 $\{0, 1, \ldots, d\}$, initial distribution $\beta = (0, \alpha)$, and infinitesimal generator

$$Q = \begin{bmatrix} 0 & \mathbf{0} \\ -A\boldsymbol{e} & A \end{bmatrix}, \tag{2.1}$$

where $\mathbf{0} = (0, ..., 0)$ is the d-dimensional row vector of zeros, $\boldsymbol{e} = (1, ..., 1)^T$ is the d-dimensional column vector of 1's, sub-generator A is a $d \times d$ nonsingular matrix, and $\alpha = (\alpha_1, ..., \alpha_d)$. Thus, a non-negative random variable X is of phase type with representation (α, 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 random variable X is of phase type with representation (α, A, d) if and only if

$$\bar{F}(x) = \Pr\{X(x) \in \{1, \ldots, d\}\} = \alpha\, e^{xA}\, \boldsymbol{e},\ \ x \geq 0. \tag{2.2}$$

Thus,

$$EX^k = \int_0^\infty x^k dF(x) = (-1)^k\, k!\, (\alpha A^{-k}\, \boldsymbol{e}),\ \ k = 1, 2, ... \tag{2.3}$$

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 (α, A, d), then for any $t > 0$, the excess risk $(X - t \mid X > t)$ has a PH distribution with representation

(α_t, A, d), where

$$\alpha_t = \frac{\alpha\, e^{tA}}{\alpha\, e^{tA}\, \boldsymbol{e}}. \tag{2.4}$$

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

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

It follows from (2.2) that

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

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

Remark 2.2. Proposition 2.1 can be also argued as follows. Let $\{X(t), t \geq 0\}$ be the underlying Markov chain on $\{0, 1, \ldots, d\}$ for the phase type random variable X. Since $\{X > t\} = \{X(t) \in \{1, \ldots, d\}\}$, it then follows from the Markov property that

$$(X \mid X > t) \;=\; (X \mid X(t) \in \{1, \ldots, d\}) \;=_{st}\; t + \inf\{s > 0 : X^*(s) = 0\},$$

where $\{X^*(s), s \geq 0\}$ is a Markov chain with the same state space and generator as those of $\{X(t), t \geq 0\}$, but the initial probability vector $(0, \alpha^*_t)$ and

$$\alpha^*_t = \frac{1}{\Pr\{X(t) \in \{1, \ldots, d\}\}}\,(\Pr\{X(t) = 1\}, \ldots, \Pr\{X(t) = d\}) = \frac{\alpha\, e^{tA}}{\alpha\, e^{tA}\, \boldsymbol{e}} = \alpha_t.$$

Thus, $(X - t \mid X > t) =_{st} \inf\{s > 0 : X^*(s) = 0\}$ is of phase type with representation (α_t, A, d). Such a probabilistic interpretation is the basis of our Markovian method that will be extensively used in this paper.

Corollary 2.3. If random variable X has a PH distribution with representation (α, A, d), then for any $t > 0$,

$$CTE_X(t) = t - \frac{\alpha\, A^{-1}\, e^{tA}\, \boldsymbol{e}}{\alpha\, e^{tA}\, \boldsymbol{e}}. \tag{2.5}$$

Proof. It follows from Proposition 2.1 and (2.3) that

$$CTE_X(t) = t + E(X - t \mid X > t) = t - \alpha_t\, A^{-1}\boldsymbol{e} = t - \frac{\alpha\, e^{tA}\, A^{-1}\, \boldsymbol{e}}{\alpha\, e^{tA}\, \boldsymbol{e}} = t - \frac{\alpha\, A^{-1}\, e^{tA}\, \boldsymbol{e}}{\alpha\, e^{tA}\, \boldsymbol{e}},$$

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

As one application of Corollary 2.3, we can obtain an explicit expression for conditional expectation $E(X \mid X \leq t)$ immediately as follows.

Proposition 2.4. If X has a PH distribution with representation (α, A, d), then for any $t > 0$,

$$E(X|X \le t) = \frac{-\alpha\, A^{-1}\, \boldsymbol{e} - t\, \alpha\, e^{tA}\, \boldsymbol{e} + \alpha\, A^{-1}\, e^{tA}\, \boldsymbol{e}}{1 - \alpha\, e^{tA}\, \boldsymbol{e}}, \tag{2.6}$$

$$E(t - X|X \le t) = \frac{t + \alpha\, A^{-1}\, (I - e^{tA})\, \boldsymbol{e}}{1 - \alpha\, e^{tA}\, \boldsymbol{e}}. \tag{2.7}$$

Proof. The formula (2.6) follows from

$$E(X) = E(X \mid X > t)\, P(X > t) + E(X \mid X \le t)\, P(X \le t),$$

$E(X) = -\alpha\, A^{-1}\, \boldsymbol{e}$, $P(X \le t) = 1 - \alpha\, e^{tA}\, \boldsymbol{e}$, $P(X > t) = \alpha\, e^{tA}\, \boldsymbol{e}$, and (2.5). (2.7) is obtained from $E(t - X|X \le t) = t - E(X|X \le t)$ and (2.6). □

The conditional expectation $E(X|X \le t)$ will be used later in the paper. In risk analysis, $E(t - X|X \le t)$ describes the surplus beyond the risk that has been experienced.

3 CTE for Multivariate Phase Type Distributions

Let $\{X(t), t \ge 0\}$ be a right-continuous, continuous-time Markov chain on a finite state space $\mathcal{E}$ with generator Q. Let $\mathcal{E}_i$, $i = 1, \ldots, n$, be n nonempty stochastically closed subsets of $\mathcal{E}$ such that $\cap_{i=1}^n \mathcal{E}_i$ is a proper subset of $\mathcal{E}$ (A subset of the state space is said to be stochastically closed if once the process $\{X(t), t \ge 0\}$ enters it, $\{X(t), t \ge 0\}$ never leaves). We assume that absorption into $\cap_{i=1}^n \mathcal{E}_i$ is certain. Since we are interested in the process only until it is absorbed into $\cap_{i=1}^n \mathcal{E}_i$, we may assume, without loss of generality, that $\cap_{i=1}^n \mathcal{E}_i$ consists of one state, which we shall denote by Δ. Thus, without loss of generality, we may write $\mathcal{E} = (\cup_{i=1}^n \mathcal{E}_i) \cup \mathcal{E}_0$ for some subset $\mathcal{E}_0 \subset \mathcal{E}$ with $\mathcal{E}_0 \cap \mathcal{E}_j = \emptyset$ for $1 \le j \le n$. 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

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

where $\mathbf{0} = (0, ..., 0)$ is the d-dimensional row vector of zeros, $\boldsymbol{e} = (1, ..., 1)^T$ is the d-dimensional column vector of 1's, sub-generator A is a $d \times d$ nonsingular matrix, and

$d = |\mathcal{E}| - 1$. Let β be an initial probability vector on $\mathcal{E}$ such that $\beta(\Delta) = 0$. We can write $\beta = (0, \alpha)$.

We define

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

As in Assaf et al. (1984), for simplicity, we shall assume that $P(X_1 > 0, \ldots, X_n > 0) = 1$, which means that the underlying Markov chain $\{X(t), t \geq 0\}$ starts within $\mathcal{E}_0$ almost surely. The joint distribution of $(X_1, \ldots, X_n)$ is called a *multivariate phase type* distribution (MPH) with representation $(\alpha, A, \mathcal{E}, \mathcal{E}_1, \ldots, \mathcal{E}_n)$, and $(X_1, \ldots, X_n)$ 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 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 a closed form. The set of n-dimensional MPH distributions is dense in the set of all distributions on $[0, \infty)^n$. It is 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 reliability systems.

The sum $S = X_1 + \cdots + X_n$, the extreme values $X_{(1)} = \min\{X_1, ..., X_n\}$, and $X_{(n)} = \max\{X_1, ..., X_n\}$ are all of phase type if $(X_1, ..., X_n)$ has a multivariate phase type distribution. Thus, Corollary 2.3 will yield explicit expressions of CTEs for S, $X_{(1)}$, and $X_{(n)}$ if we can provide 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 $(X_{(1)}, X_i, X_{(n)})$, $i = 1, 2, ..., n$, and then obtain the related CTEs.

3.1 CTE of Sums

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

$$\Gamma^n_{\emptyset} = \mathcal{E}_0.$$

$$\Gamma_i^{n-1} = \mathcal{E}_i - \cup_{k\neq i}(\mathcal{E}_i \cap \mathcal{E}_k),\ i = 1,\ldots,n.$$

$$\Gamma_{ij}^{n-2} = \mathcal{E}_i \cap \mathcal{E}_j - \cup_{k\neq i,\,k\neq j}(\mathcal{E}_i \cap \mathcal{E}_j \cap \mathcal{E}_k),\ i \neq j.$$

$$\ldots,\ldots.$$

For any $\mathcal{D} \subseteq \{1,\ldots,n\}$,

$$\Gamma_{\mathcal{D}}^{n-|\mathcal{D}|} = \cap_{i\in\mathcal{D}}\mathcal{E}_i - \cup_{k\notin\mathcal{D}}((\cap_{i\in\mathcal{D}}\mathcal{E}_i) \cap \mathcal{E}_k).$$

$$\ldots,\ldots,$$

$$\Gamma_{12\ldots n}^0 = \{\Delta\}.$$

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 $e \in \mathcal{E}$, define

$$k(e) = \text{number of indexes in } \{j : e \notin \mathcal{E}_j, 1 \leq j \leq n\}. \tag{3.3}$$

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

Lemma 3.1. (Cai and Li 2005) Let $(X_1,\ldots,X_n)$ be a phase type vector whose distribution has representation $(\alpha, A, \mathcal{E}, \mathcal{E}_i, i = 1,\ldots,n)$, where $A = (a_{e,e'})$. Then $\sum_{i=1}^n X_i$ has a phase type distribution with representation $(\alpha, T, |\mathcal{E}| - 1)$, where $T = (t_{e,e'})$ is given by,

$$t_{e,e'} = \frac{a_{e,e'}}{k(e)}, \tag{3.4}$$

that is, $t_{e,e'} = \frac{a_{e,e'}}{k}$ if $e \in \Gamma_{\mathcal{D}}^k$, for some $\mathcal{D} \subset \{1,...,n\}$.

Theorem 3.2. Let $(X_1,\ldots,X_n)$ be a phase type vector whose distribution has representation $(\alpha, A, \mathcal{E}, \mathcal{E}_1,...,\mathcal{E}_n)$. Then the CTE expression of $S = X_1 + \cdots + X_n$ is given by, for any $t > 0$,

$$CTE_S(t) = t - \frac{\alpha\, T^{-1}\, e^{tT}\boldsymbol{e}}{\alpha\, e^{tT}\,\boldsymbol{e}}, \tag{3.5}$$

where T is defined by (3.4).

Proof. (3.5) follows from the phase type representation $(\alpha, T, |\mathcal{E}|-1)$ of S in Lemma 3.1 and (2.5). $\square$

3.2 CTE of Order Statistics

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

$$\begin{aligned}\bar{F}(x_1, ..., x_n) &= \Pr\{X_1 > x_1, ..., X_n > x_n\} \\ &= \alpha\, e^{x_n A}\, g_n\, e^{(x_{n-1}-x_n)A}\, g_{n-1} \cdots e^{(x_1-x_2)A}\, g_1 \boldsymbol{e},\end{aligned} \tag{3.6}$$

$$\begin{aligned}F(x_1, ..., x_n) &= \Pr\{X_1 \leq x_1, ..., X_n \leq x_n\} \\ &= \beta\, e^{x_n Q}\, h_n\, e^{(x_{n-1}-x_n)Q}\, h_{n-1} \cdots e^{(x_1-x_2)Q}\, h_1 \boldsymbol{e},\end{aligned} \tag{3.7}$$

where, for $k = 1, ..., n$, g_k is defined as a diagonal $d \times d$ matrix whose e-th diagonal element for $e = 1, ..., d$ equals 1 if $e \in \mathcal{E} - \mathcal{E}_k$ and zero otherwise, and h_k is defined as a diagonal $(d+1) \times (d+1)$ matrix whose e-th diagonal element for $e = 1, ..., d+1$ equals 1 if $e \in \mathcal{E}_k$ and zero otherwise.

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

$$Q = \begin{bmatrix} 0 & \mathbf{0} & \mathbf{0} \\ -(A_S \boldsymbol{e} + A_{SS'} \boldsymbol{e}) & A_S & A_{SS'} \\ -(A_{S'S} \boldsymbol{e} + A_{S'} \boldsymbol{e}) & A_{S'S} & A_{S'} \end{bmatrix}, \tag{3.8}$$

where A_S ($A_{S'}$) is the sub-matrix of A by removing the s-th row and s-th column of A for all $s \notin S$ ($s \notin S'$), and $A_{SS'}$ ($A_{S'S}$) is the sub-matrix of A by removing the s-th row and s'-th column of A for all $s \in S'$, $s' \in S$ ($s \in S$, $s' \in S'$). In particular, we have $A_{\mathcal{E}-\{\Delta\}} = A$.

1. The matrix

$$Q_S = \begin{bmatrix} 0 & \mathbf{0} \\ -A_S\, \boldsymbol{e} & A_S \end{bmatrix}, \tag{3.9}$$

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

2. Let $A_{[S]} = A_S + D(A_{SS'})$, where $D(A_{SS'})$ is the diagonal matrix with its s-th diagonal entry being the s-th entry of $A_{SS'}\,\boldsymbol{e}$. The matrix

$$Q_{[S]} = \begin{bmatrix} 0 & \mathbf{0} \\ -A_{[S]}\,\boldsymbol{e} & A_{[S]} \end{bmatrix}, \tag{3.10}$$

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 $\mathcal{E} - \{\Delta\}$ to S'.

For any d-dimensional probability vector α and any subset $S \subseteq \mathcal{E} - \{\Delta\}$, we denote by α_S the $|S|$-dimensional sub-vector of α by removing its s-th entry for all $s \notin S$. The vector $I(S)$ denotes the column vector with the s-th entry being 1 if $s \in S$ and zero otherwise. Furthermore, for any $S \subseteq \mathcal{E} - \{\Delta\}$, we write $\alpha_t(S)$ for the following $|S|$-dimensional row vector

$$\alpha_t(S) = \frac{\alpha_S\, e^{tA_S}}{\alpha_S\, e^{tA_S}\,\boldsymbol{e}}. \tag{3.11}$$

Note that $\alpha_t(\mathcal{E} - \{\Delta\}) = \alpha_t$, where α_t is given by (2.4).

For any phase type random vector $(X_1, \ldots, X_n)$, Assaf et al. (1984) showed that the extreme values $X_{(1)} = \min\{X_1, ..., X_n\}$ and $X_{(n)} = \max\{X_1, ..., X_n\}$ are also of phase type. Their representations can be obtained from (3.6) and (3.7).

Lemma 3.3. Let $(X_1, \ldots, X_n)$ be of phase type with representation $(\alpha, A, \mathcal{E}, \mathcal{E}_1, \ldots, \mathcal{E}_n)$. Then

1. $X_{(1)}$ is of phase type with representation $\left(\frac{\alpha_{\mathcal{E}_0}}{\alpha_{\mathcal{E}_0}\boldsymbol{e}},\, A_{\mathcal{E}_0},\, |\mathcal{E}_0|\right)$, where $A_{\mathcal{E}_0}$ is defined as in (3.8).

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

Proof. By (3.6), the survival function of $X_{(1)}$ is given by, for $x \geq 0$,

$$\bar{F}_{X_{(1)}}(x) = \Pr\{X_{(1)} > x\} = \bar{F}(x, ..., x) = \alpha\, e^{xA}\, g_n \cdots g_1\, \boldsymbol{e}. \tag{3.12}$$

Note that $g_n \cdots g_1\, \boldsymbol{e} = I(\mathcal{E}_0)$, where $I(\mathcal{E}_0)$ denotes the d-dimensional vector such that the e-th component is 1 if $e \in \mathcal{E}_0$ and zero otherwise. Since $\mathcal{E}_i$, $1 \leq i \leq n$, are all stochastically closed, we have

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

which implies that $X_{(1)}$ is of phase type with representation $\left(\frac{\alpha_{\mathcal{E}_0}}{\alpha_{\mathcal{E}_0}\boldsymbol{e}},\, A_{\mathcal{E}_0},\, |\mathcal{E}_0|\right)$.

Similarly, by (3.7), the distribution function of $X_{(n)}$ is given by, for $x \geq 0$,

$$F_{X_{(n)}}(x) = \Pr\{X_{(n)} \leq x\} = F(x, ..., x) = \beta\, e^{xQ}\, h_n \cdots h_1\, \boldsymbol{e} = \beta\, e^{xQ}\, I(\{\Delta\}) = 1 - \alpha\, e^{xA}\, \boldsymbol{e}.$$

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

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 $\mathcal{E} - \{\Delta\}$, and thus $X_{(1)}$ and $X_{(n)}$ are of phase type with representations given in Lemma 3.3. In general, the k-th order statistic $X_{(k)}$ is the exit time of $\{X(t), t \geq 0\}$ from $\mathcal{E} - \cup_{\{\{i_1,i_2,...,i_k\}\}}(\cap_{j=1}^{k}\mathcal{E}_{i_j})$, and thus also of phase type. Hence, we obtain the following lemma.

Lemma 3.4. Let $X_{(k)}$, $1 \leq k \leq n$, be the k-th smallest among $X_1, \ldots, X_n$ that are of phase type with representation $(\alpha, A, \mathcal{E}, \mathcal{E}_1, \ldots, \mathcal{E}_n)$. Then we have

1. $(X_{(1)}, \ldots, X_{(n)})$ is of phase type with representation $(\alpha, A, \mathcal{E}, \mathcal{O}_1, \ldots, \mathcal{O}_n)$, where $\mathcal{O}_k = \cup_{\{\{i_1,i_2,...,i_k\}\}}(\cap_{j=1}^{k}\mathcal{E}_{i_j})$, $1 \leq k \leq n$.
2. $X_{(k)}$ is of phase type with representation $\left(\frac{\alpha_{\mathcal{E}-\mathcal{O}_k}}{\alpha_{\mathcal{E}-\mathcal{O}_k}\boldsymbol{e}},\, A_{\mathcal{E}-\mathcal{O}_k},\, |\mathcal{E} - \mathcal{O}_k|\right)$, $k = 1, \ldots, n$.

Lemmas 3.3 and 3.4, and (2.5) immediately yield the CTE expressions of the extreme values $X_{(1)}$ and $X_{(n)}$ as follows.

Theorem 3.5. Let $(X_1, \ldots, X_n)$ be of phase type with representation $(\alpha, A, \mathcal{E}, \mathcal{E}_1, ..., \mathcal{E}_n)$.

1. The excess loss $(X_{(1)} - t \mid X_{(1)} > t)$ of a subportfolio with the least risk is of phase type with representation $(\alpha_t(\mathcal{E}_0),\, A_{\mathcal{E}_0},\, |\mathcal{E}_0|)$, and the CTE of $X_{(1)}$ is given by

$$CTE_{X_{(1)}}(t) = t - \alpha_t(\mathcal{E}_0)\, A_{\mathcal{E}_0}^{-1}\, \boldsymbol{e}. \tag{3.13}$$

2. The excess loss $(X_{(n)} - t \mid X_{(n)} > t)$ of the riskiest subportfolio is of phase type with representation $(\alpha_t,\, A,\, |\mathcal{E}| - 1)$, and the CTE of $X_{(n)}$ is given by

$$CTE_{X_{(n)}}(t) = t - \alpha_t\, A^{-1}\, \boldsymbol{e}, \text{ where } \alpha_t = \frac{\alpha\, e^{tA}}{\alpha\, e^{tA}\, \boldsymbol{e}}. \tag{3.14}$$

3. 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 $(\alpha_t(\mathcal{E} - \mathcal{O}_k),\, A_{\mathcal{E}-\mathcal{O}_k},\, |\mathcal{E} - \mathcal{O}_k|)$, and the CTE of $X_{(k)}$ is given by

$$CTE_{X_{(k)}}(t) = t - \alpha_t(\mathcal{E} - \mathcal{O}_k)\, A^{-1}_{\mathcal{E}-\mathcal{O}_k}\, \boldsymbol{e}, \tag{3.15}$$

and $\mathcal{O}_k = \cup_{\{\{i_1, i_2, \ldots, i_k\}\}} (\cap_{j=1}^{k} \mathcal{E}_{i_j})$, $k = 1, \ldots, n$.

3.3 CTE Involving Different Subportfolios

The Markovian method also leads to the expressions of CTEs among different subportfolios. Hereafter, for any d-dimensional probability vector α and any subset $S \subseteq \mathcal{E} - \{\Delta\}$, we denote, with slight abuse of notations, by $(\mathbf{0}, \alpha_S)$ the d-dimensional vector whose s-th entry is the s-th entry of α if $s \in S$, and zero 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 (3.11)) if $s \in S$, and zero otherwise.

Theorem 3.6. If $(X_1, ..., X_n)$ is of phase type with representation $(\alpha,\, A,\, \mathcal{E},\, \mathcal{E}_1, ..., \mathcal{E}_n)$, then the random vector $[(X_1 - t, ..., X_n - t) \mid X_{(1)} > t]$ is of phase type with representation $((\mathbf{0},\, \alpha_t(\mathcal{E}_0)),\, A,\, \mathcal{E},\, \mathcal{E}_1, ..., \mathcal{E}_n)$.

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

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

It follows from the Markov property that

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

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

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

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

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

Corollary 3.7. Let $(X_1, ..., X_n)$ be of phase type with representation $(\alpha, A, \mathcal{E}, \mathcal{E}_1, ..., \mathcal{E}_n)$. Then, the excess loss $(X_i - t \mid X_{(1)} > t)$ is of phase type with representation

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

and

$$CTE_{X_i|X_{(1)}}(t) = t - (\mathbf{0},\ \alpha_t(\mathcal{E}_0))\, A^{-1}_{\mathcal{E}-\mathcal{E}_i}\, \boldsymbol{e}. \tag{3.16}$$

Proof. By Theorem 3.6, $[(X_1 - t, ..., X_n - t) \mid X_{(1)} > t]$ has a phase type representation $((\mathbf{0},\ \alpha_t(\mathcal{E}_0)),\ A,\ \mathcal{E},\ \mathcal{E}_1, ..., \mathcal{E}_n)$. Hence, $(X_i - t \mid X_{(1)} > t)$ has a phase type representation

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

Thus, (3.16) follows from (2.5) □

In fact, from a direct calculation, we have,

$$\begin{aligned} CTE_{X_i|X_{(1)}}(t) &= E(X_i \mid X_{(1)} > t) = \frac{\int_0^\infty \Pr\{X_i > x, X_{(1)} > t\}dx}{\Pr\{X_{(1)} > t\}} \\ &= t + \frac{\int_t^\infty \bar{F}(t, ..., t, x, t, ..., t)dx}{\bar{F}(t, t, ..., t)}. \end{aligned} \tag{3.17}$$

Since random vector $(X_i, X_1, ..., X_{i-1}, X_{i+1}, ..., X_n)$ has an MPH distribution with representation $(\alpha, A, \mathcal{E}, \mathcal{E}_i, \mathcal{E}_1, ..., \mathcal{E}_{i-1}, \mathcal{E}_{i+1}, ..., \mathcal{E}_n)$, then, by (3.6), for $x > t$,

$$\begin{aligned} \bar{F}(t, ..., t, x, t, ..., t) &= \Pr\{X_i > x, X_1 > t, ..., X_{i-1} > t, X_{i+1} > t, \dots, X_n > t\} \\ &= \alpha\, e^{tA} \left[\prod_{k=n,\, k\neq i}^{1} g_k\right] e^{(x-t)A}\, g_i\, \boldsymbol{e}, \end{aligned} \tag{3.18}$$

where, for $k = 1, ..., n$, g_k is defined as a diagonal $d \times d$ matrix whose e-th diagonal element for $e = 1, ..., d$ equals 1 if $e \in \mathcal{E} - \mathcal{E}_k$ and is 0 otherwise. Therefore, we obtain that for any $1 \le i \le n$,

$$\begin{aligned} CTE_{X_i|X_{(1)}}(t) &= t + \frac{\alpha\, e^{tA} \left[\prod_{k=n,\, k\neq i}^{1} g_k\right] \left[\int_t^\infty e^{(x-t)A} dx\right]\, g_i\, \boldsymbol{e}}{\alpha\, e^{tA} \left[\prod_{k=n}^{1} g_k\right]\, \boldsymbol{e}} \\ &= t - \frac{\alpha\, e^{tA} \left[\prod_{k=n,\, k\neq i}^{1} g_k\right]\, A^{-1}\, g_i\, \boldsymbol{e}}{\alpha\, e^{tA} \left[\prod_{k=n}^{1} g_k\right]\, \boldsymbol{e}}. \end{aligned} \tag{3.19}$$

The expression (3.19) provides another formula for $CTE_{X_i|X_{(1)}}(t)$.

Note that the expression (3.19) yields the same result as that in Corollary 3.7, due to the fact that

$$\begin{aligned}\frac{\alpha\, e^{tA}\,\left[\prod_{k=n,\,k\neq i}^{1} g_k\right]\,A^{-1}\,g_i\,\boldsymbol{e}}{\alpha\, e^{tA}\,\left[\prod_{k=n}^{1} g_k\right]\boldsymbol{e}} &= \frac{(\mathbf{0},\,\frac{\alpha_{\mathcal{E}_0}}{\alpha_{\mathcal{E}_0}\boldsymbol{e}}\,e^{tA_{\mathcal{E}_0}})\,A^{-1}\,g_i\,\boldsymbol{e}}{\frac{\alpha_{\mathcal{E}_0}}{\alpha_{\mathcal{E}_0}\boldsymbol{e}}\,e^{tA_{\mathcal{E}_0}}\boldsymbol{e}} = \frac{(\mathbf{0},\,\alpha_{\mathcal{E}_0}e^{tA_{\mathcal{E}_0}})\,A^{-1}_{\mathcal{E}-\mathcal{E}_i}\boldsymbol{e}}{\alpha_{\mathcal{E}_0}e^{tA_{\mathcal{E}_0}}\boldsymbol{e}}\\ &= (\mathbf{0},\,\alpha_t(\mathcal{E}_0))A^{-1}_{\mathcal{E}-\mathcal{E}_i}\boldsymbol{e}.\end{aligned}$$

To develop a general scheme, we drive an expression for $E(X_i \mid X_k > t)$, $i, k = 1, \ldots, n$, where $(X_1, \ldots, X_n)$ is of phase type with representation $(\alpha, A, \mathcal{E}, \mathcal{E}_1, \ldots, \mathcal{E}_n)$, 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

$$q_k = \Pr\{X(t) \in \mathcal{E} - \mathcal{E}_k\} = \alpha\, e^{tA} I(\mathcal{E} - \mathcal{E}_k) = \frac{\alpha_{\mathcal{E}-\mathcal{E}_k}}{\alpha_{\mathcal{E}-\mathcal{E}_k}\boldsymbol{e}}\, e^{tA_{\mathcal{E}-\mathcal{E}_k}}\,\boldsymbol{e}.$$

We have for any $x \geq 0$,

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

where

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

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

$$\begin{aligned} &E(X_i \mid X_k > t) = \int_0^\infty \Pr\{X_i > x \mid X_k > t\}dx\\ =\ & p_{1,k}\,E(X_i \mid X(t) \in \mathcal{E} - \mathcal{E}_i \cup \mathcal{E}_k) + p_{2,k}\,E(X_i \mid X(t) \in \mathcal{E}_i - \mathcal{E}_i \cap \mathcal{E}_k). \end{aligned} \tag{3.20}$$

Since

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

we have

$$p_{1,k} = \left(\frac{\alpha_{\mathcal{E}-\mathcal{E}_k}\boldsymbol{e}}{\alpha_{\mathcal{E}-\mathcal{E}_i\cup\mathcal{E}_k}\boldsymbol{e}}\right)\left(\frac{\alpha_{\mathcal{E}-\mathcal{E}_i\cup\mathcal{E}_k}e^{tA_{\mathcal{E}-\mathcal{E}_i\cup\mathcal{E}_k}}\boldsymbol{e}}{\alpha_{\mathcal{E}-\mathcal{E}_k}e^{tA_{\mathcal{E}-\mathcal{E}_k}}\boldsymbol{e}}\right), \tag{3.21}$$

$$p_{2,k} = \frac{\alpha e^{tA} I(\mathcal{E}_i - \mathcal{E}_i \cap \mathcal{E}_k)}{\alpha e^{tA} I(\mathcal{E} - \mathcal{E}_k)}. \tag{3.22}$$

To calculate two conditional expectations in (3.20), we need the following.

Lemma 3.8. Let $(X_1, \ldots, X_n)$ be MPH with representation $(\alpha, A, \mathcal{E}, \mathcal{E}_1, \ldots, \mathcal{E}_n)$.

1. For any i, k, if $\Pr\{X(t) \in \mathcal{E} - \mathcal{E}_i \cup \mathcal{E}_k\} > 0$,

$$E(X_i \mid X(t) \in \mathcal{E} - \mathcal{E}_i \cup \mathcal{E}_k) = t - (\mathbf{0},\, \alpha_t(\mathcal{E} - \mathcal{E}_i \cup \mathcal{E}_k))\; A^{-1}_{\mathcal{E}-\mathcal{E}_i}\, \boldsymbol{e}. \tag{3.23}$$

2. For any $i \neq k$, if $\Pr\{X(t) \in \mathcal{E}_i - \mathcal{E}_i \cap \mathcal{E}_k\} > 0$,

$$\begin{aligned} & E(X_i \mid X(t) \in \mathcal{E}_i - \mathcal{E}_i \cap \mathcal{E}_k) \qquad (3.24)\\ = \; & \frac{-(\mathbf{0},\, \alpha')\, A^{-1}_{[\mathcal{E}-\mathcal{E}_k]}\, \boldsymbol{e} - t\, (\mathbf{0}, \alpha')\, e^{tA_{[\mathcal{E}-\mathcal{E}_k]}}\, \boldsymbol{e} + (\mathbf{0},\, \alpha')\, A^{-1}_{[\mathcal{E}-\mathcal{E}_k]}\, e^{tA_{[\mathcal{E}-\mathcal{E}_k]}}\, \boldsymbol{e}}{1 - (\mathbf{0},\, \alpha')\, e^{tA_{[\mathcal{E}-\mathcal{E}_k]}}\, \boldsymbol{e}}, \end{aligned}$$

where $\alpha' = \frac{\alpha_{\mathcal{E}-\mathcal{E}_i\cup\mathcal{E}_k}}{\alpha_{\mathcal{E}-\mathcal{E}_i\cup\mathcal{E}_k}\boldsymbol{e}}$.

Proof. To calculate $E(X_i \mid X(t) \in \mathcal{E} - \mathcal{E}_i \cup \mathcal{E}_k)$, it follows from Markov property that

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

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

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

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

$$E(X_i \mid X(t) \in \mathcal{E} - \mathcal{E}_i \cup \mathcal{E}_k) = t - (\mathbf{0},\, \alpha_t(\mathcal{E} - \mathcal{E}_i \cup \mathcal{E}_k))A^{-1}_{\mathcal{E}-\mathcal{E}_i}\boldsymbol{e}. \tag{3.25}$$

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

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

Now we define a new Markov chain $\{X'(t), t \geq 0\}$ with state space $\mathcal{E} - \mathcal{E}_k$ as follows. The set of absorbing states is $\mathcal{E}_i - \mathcal{E}_i \cap \mathcal{E}_k$, the initial probability vector of $\{X'(t), t \geq 0\}$ is $(\mathbf{0}, \alpha')$, where $\alpha' = \frac{\alpha_{\mathcal{E}-\mathcal{E}_i \cup \mathcal{E}_k}}{\alpha_{\mathcal{E}-\mathcal{E}_i \cup \mathcal{E}_k} \boldsymbol{e}}$. The generator of $\{X'(t), t \geq 0\}$ is given by $A_{[\mathcal{E}-\mathcal{E}_k]}$ (see (3.10)). Let

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

which has a phase type distribution. Then, from (2.6),

$$\begin{aligned} & E(X_i \mid X(t) \in \mathcal{E}_i - \mathcal{E}_i \cap \mathcal{E}_k) = E(X'_i \mid X'_i \leq t) \\ = \quad & \frac{-(\mathbf{0}, \alpha')\, A^{-1}_{[\mathcal{E}-\mathcal{E}_k]}\, \boldsymbol{e} - t\, (\mathbf{0}, \alpha')\, e^{tA_{[\mathcal{E}-\mathcal{E}_k]}}\, \boldsymbol{e} + (\mathbf{0}, \alpha')\, A^{-1}_{[\mathcal{E}-\mathcal{E}_k]}\, e^{tA_{[\mathcal{E}-\mathcal{E}_k]}}\, \boldsymbol{e}}{1 - (\mathbf{0}, \alpha')\, e^{tA_{[\mathcal{E}-\mathcal{E}_k]}}\, \boldsymbol{e}}. \end{aligned}$$

Thus, (3.24) in Lemma 3.8 (2) holds. $\square$

Observe that if $\mathcal{E}_i \subseteq \mathcal{E}_k$, then $X_i \geq X_k$ almost surely. It follows from (3.20) that

$$E(X_i \mid X_k > t) = E(X_i \mid X(t) \in \mathcal{E} - \mathcal{E}_k), \text{ if } \mathcal{E}_i \subseteq \mathcal{E}_k. \tag{3.26}$$

Thus, Lemma 3.8 (1) implies that $(X_i - t \mid X_k > t)$ is of phase type with representation $((\mathbf{0},\ \alpha_t(\mathcal{E} - \mathcal{E}_k)),\ A_{\mathcal{E}-\mathcal{E}_i},\ |\mathcal{E} - \mathcal{E}_i|)$. This leads to the following corollaries.

Corollary 3.9. Let $(X_1, ..., X_n)$ be MPH with representation $(\alpha, A, \mathcal{E}, \mathcal{E}_1, ..., \mathcal{E}_n)$. Let $X_{(k)}$, $1 \leq k \leq n$, be the k-th order statistic of $(X_1, ..., X_n)$.

1. For any $k \leq i$, $(X_{(i)} - t \mid X_{(k)} > t)$ is of phase type with representation

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

   and

   $$E(X_{(i)} \mid X_{(k)} > t) = t - (\mathbf{0},\ \alpha_t(\mathcal{E} - \mathcal{O}_k))\ A^{-1}_{\mathcal{E}-\mathcal{O}_i}\, \boldsymbol{e}. \tag{3.27}$$

2. In particular, $(X_{(n)} - t \mid X_{(1)} > t)$ is of phase type with representation

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

   and

   $$E(X_{(n)} \mid X_{(1)} > t) = t - (\mathbf{0},\ \alpha_t(\mathcal{E}_0))\ A^{-1}\, \boldsymbol{e}. \tag{3.28}$$

Proof. It follows from Lemma 3.4 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 3.8 (1), $A_{\mathcal{E}-\{\Delta\}} = A$, and $|\mathcal{E} - \{\Delta\}| = |\mathcal{E}| - 1$. $\square$

Corollary 3.10. Let $(X_1, ..., X_n)$ be MPH with representation $(\alpha, A, \mathcal{E}, \mathcal{E}_1, ..., \mathcal{E}_n)$. Then $(X_{(n)} - t \mid X_i > t)$ is of phase type with representation

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

and

$$E(X_{(n)} \mid X_i > t) = t - (\mathbf{0},\ \alpha_t(\mathcal{E} - \mathcal{E}_i))\ A^{-1}\,\boldsymbol{e}. \tag{3.29}$$

Proof. If $(X_1, ..., X_n)$ is MPH with representation $(\alpha,\ A,\ \mathcal{E}_1, ..., \mathcal{E}_n)$, then $(X_i, X_{(n)})$ is of phase type with representation $(\alpha,\ A,\ \mathcal{E},\ \mathcal{E}_i,\ \{\Delta\})$. Then the results follow from (3.26), Lemma 3.8 and the fact that $X_{(n)} \geq X_i$ almost surely for any i. $\square$

The expression for $E(X_i \mid X_{(n)} > t)$ is cumbersome, but can be obtained from (3.20), (3.21), (3.22), and Lemma 3.8.

4 CTE of Marshall-Olkin Distributions

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

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

$$X_j = \min\{E_S :\ S \ni j\}, \quad j = 1, \ldots, n. \tag{4.1}$$

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

$$X_j = \inf\{t : N_S(t) \geq 1,\ S \ni j\}, \quad j = 1, \ldots, n. \tag{4.2}$$

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 (4.2) that $(X_1, \dots, X_n)$ is of phase type with 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 $\mathcal{E} = \{S : S \subseteq \{1, \dots, n\}\}$, and starting at $\emptyset$ almost surely. The index set $\{1, \dots, n\}$ is the absorbing state, and

$$\mathcal{E}_0 = \{\emptyset\}$$

$$\mathcal{E}_j = \{S : S \ni j\}, \quad j = 1, \dots, n.$$

It follows from (4.2) that its sub-generator is given by $A = (a_{e,e'})$, where

$$a_{e,e'} = \sum_{L:\ L \subseteq S',\, L \cup S = S'} \lambda_L, \text{ if } e = S,\ e' = S' \text{ and } S \subset S',$$

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

and zero otherwise. 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 $n = 2$, the state space $\mathcal{E} = \{12, 2, 1, 0\}$, $\mathcal{E}_j = \{12, j\}$, $j = 1, 2$, where 12 is the absorbing state. Note that the absorbing state 12 is usually denoted by 0 in Section 3, whereas 0 in this and next examples abbreviates the state $\emptyset$. Furthermore, let the initial probability vector be $(0, \alpha)$ with $\alpha = (0, 0, 1)$. Then, the sub-generator 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 $\Lambda = \lambda_{12} + \lambda_2 + \lambda_1 + \lambda_\emptyset$.

Thus, in (3.5), the matrix T is given by

$$T = \begin{bmatrix} -\lambda_1 - \lambda_{12} & 0 & 0 \\ 0 & -\lambda_2 - \lambda_{12} & 0 \\ \frac{\lambda_2}{2} & \frac{\lambda_1}{2} & -\frac{\Lambda_0}{2} \end{bmatrix},$$

where $\Lambda_0 = \lambda_1 + \lambda_2 + \lambda_{12} = \Lambda - \lambda_\emptyset$. Furthermore, in (3.13), we have $\mathcal{E}_0 = \{0\}$, $\alpha_{\mathcal{E}_0} = 1$, $A_{\mathcal{E}_0} = -\Delta + \lambda_\emptyset = -(\lambda_1 + \lambda_2 + \lambda_{12})$, and $\alpha_t(\mathcal{E}_0) = 1$.

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 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 (3.5), (3.13), and (3.14) by 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.

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

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

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

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

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

$$\begin{aligned} CTE_S(t) &= 0.8 + t, \\ CTE_{X_{(1)}}(t) &= CTE_{X_{(n)}}(t) \;=\; 0.4 + t. \end{aligned}$$

Table 1: Effects of Dependence on the CTEs of S, $X_{(1)}$, and $X_{(n)}$

	$CTE_S(t)$			$CTE_{X_{(1)}}(t)$			$CTE_{X_{(n)}}(t)$		
t	Case 1	Case 2	Case 3	Case 1	Case 2	Case 3	Case 1	Case 2	Case 3
1	1.5143	1.5835	1.8	1.2	1.25	1.4	1.4086	1.4188	1.4
2	2.4667	2.5381	2.8	2.2	2.25	2.4	2.4007	2.4038	2.4
3	3.4471	3.5201	3.8	3.2	3.25	3.4	3.4001	3.4008	3.4
4	4.4364	4.5113	4.8	4.2	4.25	4.4	4.4000	4.4002	4.4
5	5.4296	5.5066	5.8	5.2	5.25	5.4	5.4000	5.4000	5.4
6	6.4250	6.5039	6.8	6.2	6.25	6.4	6.4000	6.4000	6.4
7	7.4216	7.5023	7.8	7.2	7.25	7.4	7.4000	7.4000	7.4
8	8.4191	8.5014	8.8	8.2	8.25	8.4	8.4000	8.4000	8.4
9	9.4170	9.5008	9.8	9.2	9.25	9.4	9.4000	9.4000	9.4
10	10.4154	10.5005	10.8	10.2	10.25	10.4	10.4000	10.4000	10.4

In all the three cases, (X_1, X_2) has the same marginal distributions, namely, 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 the different 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 in 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 the $CTE_S(t)$ becomes larger as the correlation grows. The effect of dependence on $CTE_{X_{(1)}}(t)$ is the same as that on $CTE_S(t)$. However, the effect of dependence on $CTE_{X_{(n)}}(t)$ is different from those on $CTE_S(t)$ and $CTE_{X_{(1)}}(t)$. Indeed, $CTE_{X_{(n)}}(t)$ is neither increasing nor decreasing as the correlation grows.

Example 4.2. In this example, we calculate $E(X_{(n)} \mid X_{(1)} > t), E(X_{(n)} \mid X_1 > t)$ and $CTE_{X_{(n)}}(t) = E(X_{(n)} \mid X_{(n)} > t)$ under several different sets of model parameters used in Example 4.1 and discuss the effects of dependence and different conditions on the expected maximal risk $X_{(n)}$.

To apply (3.28), we have $\mathcal{E}_0 = \{0\}$, $\{\Delta\} = \{12\}$, $\mathcal{E} - \{\Delta\} = \{2, 1, 0\}$, $\alpha_{\mathcal{E}_0} = 1$, $A_{\mathcal{E}_0} = -\Delta + \lambda_\emptyset = -(\lambda_1 + \lambda_2 + \lambda_{12})$, and $\alpha_t(\mathcal{E}_0) = 1$.

To apply (3.29) for $i = 1$, we have $\mathcal{E}_1 = \{12, 1\}$ and $\mathcal{E} - \mathcal{E}_1 = \{2, 0\}$. Thus, $\alpha_{\mathcal{E}-\mathcal{E}_1} = (0, 1)$,

$$A_{\mathcal{E}-\mathcal{E}_1} = \begin{bmatrix} -\lambda_{12} - \lambda_1 & 0 \\ \lambda_2 & -\Lambda + \lambda_\emptyset \end{bmatrix},$$

and $\alpha_t(\mathcal{E} - \mathcal{E}_1) = \alpha_{\mathcal{E}-\mathcal{E}_1} e^{tA_{\mathcal{E}-\mathcal{E}_1}} / (\alpha_{\mathcal{E}-\mathcal{E}_1} e^{tA_{\mathcal{E}-\mathcal{E}_1}} \boldsymbol{e})$.

Thus, the analytic forms of $E(X_{(n)} \mid X_{(1)} > t)$, $E(X_{(n)} \mid X_1 > t)$ and $E(X_{(n)} \mid X_{(n)} > t)$ in the following three cases and the numerical values in Table 2 were easily produced from (3.28), (3.29), and (3.14) by Mathematica. The first column of Table 2 lists several values of t, and the next several columns list values of these conditional expectations in the following three cases corresponding to those in Example 4.1.

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

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

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

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

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

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

Table 2: Effects of Dependence and Different Conditions on the Maximal Risk.

	$E(X_{(n)}\|X_{(1)} > t)$			$E(X_{(n)}\|X_1 > t)$			$E(X_{(n)}\|X_{(n)} > t)$		
t	Case 1	Case 2	Case 3	Case 1	Case 2	Case 3	Case 1	Case 2	Case 3
1	1.6	1.55	1.4	1.4164	1.4335	1.4	1.4086	1.4188	1.4
2	2.6	2.55	2.4	2.4014	2.4075	2.4	2.4007	2.4038	2.4
3	3.6	3.55	3.4	3.4001	3.4017	3.4	3.4001	3.4008	3.4
4	4.6	4.55	4.4	4.4000	4.4004	4.4	4.4000	4.4002	4.4
5	5.6	5.55	5.4	5.4000	5.4001	5.4	5.4000	5.4000	5.4
6	6.6	6.55	6.4	6.4000	6.4000	6.4	6.4000	6.4000	6.4
7	7.6	7.55	7.4	7.4000	7.4000	7.4	7.4000	7.4000	7.4
8	8.6	8.55	8.4	8.4000	8.4000	8.4	8.4000	8.4000	8.4
9	9.6	9.55	9.4	9.4000	9.4000	9.4	9.4000	9.4000	9.4
10	10.6	10.55	10.4	10.400	10.4000	10.4	10.400	10.400	10.4

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

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

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

References

[1] Asmussen, S. (2000) *Ruin Probabilities.* World Scientific, Singapore.

[2] Asmussen, S. (2003) *Applied Probability and Queues*, Second Edition. Springer, New York.

[3] Artzner, P., Delbaen, F., Eber, J.M. and Heath, D. (1999) Coherent measures of risks. *Mathematical Finance* **9**, 203-228.

[4] Assaf, D., Langberg, N., Savits, T., and Shaked, M. (1984) Multivariate phase-type distributions. *Operations Research* **32**, 688–702.

[5] Bowers, N., Gerber, H., Hickman, Jones, D. and Nesbitt, C. (1997) *Actuarial Mathematics.* The Society of Actuaries, Schaumburg, Illinois.

[6] Cai, J. and Li, H. (2005) Multivariate risk model of phase type. *Insurance: Mathematics and Economics*, to appear.

[7] Embretchts, P., Klüppelberg, C. and Mikosch, T. (1997) *Modelling Extremal Events for Insurance and Finance.* Spring, Berlin.

[8] Kulkarni, V. G. (1989) A new class of multivariate phase type distributions. *Operations Research* **37**, 151–158.

[9] Landsman Z. and Valdez, E. (2003) Tail conditional expectations for elliptical distributions. *North American Actuarial Journal* **7**, 55-71.

[10] Li, H. and S. H. Xu (2000) On the dependence structure and bounds of correlated parallel queues and their applications to synchronized stochastic systems. *Journal of Applied Probability* **37**, 1020-1043.

[11] Marshall, A. W. and Olkin, I. (1967) A multivariate exponential distribution. *Journal of the American Statistical Association* **2**, 84-98.

[12] Neuts, M. F. (1981) *Matrix-Geometric Solutions in Stochastic Models: An Algorithmic Approach.* The Johns Hopkins University Press, Baltimore.

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

[14] Shaked, M. and J. G. Shanthikumar (1994) *Stochastic Orders and Their Applications*, Academic Press, New York.