INFORMS Journal on Computing
Vol. 20, No. 2, Spring 2008, pp. 179–190
ISSN 1091-9856 | EISSN 1526-5528 | 08 | 2002 | 0179

DOI 10.1287/ijoc.1070.0228

An Algorithm for Computing Minimal Coxian Representations

Qi-Ming He
Department of Industrial Engineering, Dalhousie University, Halifax, Nova Scotia B3J 2X4, Canada,
qi-ming.he@dal.ca

Hanqin Zhang
Academy of Mathematics and Systems Science, Chinese Academy of Sciences, 100080 Beijing, People's Republic of China,
hanqin@amt.ac.cn

This paper presents an algorithm for computing minimal ordered Coxian representations of phase-type distributions whose Laplace-Stieltjes transform has only real poles. We first identify a set of necessary and sufficient conditions for an ordered Coxian representation to be minimal with respect to the number of phases involved. The conditions establish a relationship between the Coxian representations of a Coxian distribution and the derivatives of its distribution function at zero. Based on the conditions, the algorithm is developed. Three numerical examples show the effectiveness of the algorithm and some geometric properties associated with ordered Coxian representations.

Key words: Coxian distribution; phase-type distribution; matrix-exponential distribution; matrix-analytic methods; nonlinear programming
History: Accepted by Edward P. C. Kao, Area Editor for Computational Probability and Analysis; received August 2005; revised May 2006; accepted April 2007. Published online in *Articles in Advance* November 26, 2007.

1. Introduction

Coxian distributions have found many applications in the study of queueing, reliability, supply chains, insurance and risk, and telecommunications (Asmussen 2000, 2003; Cox 1955a, b; Feldmann and Whitt 1998; Haddad et al. 1998; Latouche and Ramaswami 1999; Neuts 1981, 1989; Sasaki et al. 2004). For instance, Coxian distributions have been used to model service times and interarrival times in queueing models, component life times in reliability models, and interarrival times of claims in insurance and risk models. Analysis of these models usually involves complicated computational procedures using detailed information about Coxian distributions. In particular, the numbers of phases of these distributions play a significant role in computation. Consequently, reduction in the number of phases of Coxian representations can improve efficiency in computation and performance analysis. A number of studies have been carried out on Coxian and related distributions (e.g., Cox 1955a, b; Cumani 1982; Dehon and Latouche 1982; Harris et al. 1992; Heijden 1988; Mocanu and Commault 1999; O'Cinneide 1989, 1991, 1993; Osogami and Harchol-Balter 2003b).

Neuts (1975) generalized Coxian distributions into phase-type (*PH*) distributions as the distribution of the absorption time of a finite state Markov process, which made it possible to study complicated queueing models such as the *PH*/*PH*/*c* queue analytically and numerically (Takahashi 1981). It is well known that the representation of a *PH*-distribution is not unique. To reduce the time complexity of algorithms associated with *PH*-distributions, it is useful to find a *PH*-representation with the minimal number of phases for a *PH*-distribution, known as the minimal *PH*-representation problem (Aldous and Shepp 1987; Commault and Mocanu 2003; Mocanu and Commault 1999; O'Cinneide 1989, 1990, 1991, 1993; Osogami and Harchol-Balter 2003a). Chapter 2 in Neuts (1981) provided historical notes on *PH*-distributions. Commault and Mocanu (2003) and O'Cinneide (1999) reviewed studies on *PH*-distributions.

While the problem of finding a minimal *PH*-representation for a *PH*-distribution is still open, the problem of finding simpler representations for *PH*-distributions or for some subsets of *PH*-distributions has been studied extensively (Bobbio et al. 2002, 2004; Commault and Mocanu 2003; Cumani 1982; Dehon and Latouche 1982; He and Zhang 2006a, b; Mocanu and Commaut 1999; O'Cinneide 1990, 1991, 1999). Coxian representation is one of the simpler *PH*-representations being investigated. Cumani (1982) showed that any *PH*-distribution with a triangular *PH*-representation is Coxian and a Coxian representation of the same order can be found. O'Cinneide (1989, 1991, 1993) introduced a number of new concepts in the study of *PH* and Coxian distributions. Concepts such as *PH*-simplicity and triangular order shall be used in this paper. O'Cinneide (1991) proved that all *PH*-distributions whose Laplace-Stieltjes transforms have only real poles are Coxian. In O'Cinneide (1993), a necessary and sufficient condition was given for a Coxian representation to be minimal.

Approximating general probability distributions with Coxian or *PH*-distributions has been investigated extensively as well. Johnson (1993) and Johnson and Taaffe (1989; 1990a, b) introduced several methods for fitting *PH* and Coxian distributions to general distributions by matching their lower moments. Osogami and Harchol-Balter (2003b) focused on Coxian approximations of distributions. Algorithms were introduced for computing Coxian distributions matching the first three moments of the original probability distributions. Their algorithms may find Coxian representations with the minimal number of phases, though it is not guaranteed. Sasaki et al. (2004) developed an approximation method for finding Coxian distributions as approximations of general probability distributions. Other works on Coxian approximations of distributions are Asmussen et al. (1996), Feldmann and Whitt (1998), and Heijden (1988).

Coxian distributions were studied as a subset of *PH*-distributions and as approximations to general distributions. However, no algorithm for computing a minimal Coxian representation has been developed. We find a set of necessary and sufficient conditions for an ordered Coxian representation to be minimal. This set of conditions establishes an explicit relationship between the parameters of a Coxian representation of a Coxian distribution and the derivatives of its distribution function at zero. Based on these conditions, an algorithm is introduced for computing a minimal Coxian representation for any Coxian distribution or, equivalently, any *PH*-distribution whose Laplace-Stieltjes transform has only real poles. Our necessary and sufficient conditions are based on the spectral polynomial method, which was developed in He and Zhang (2006b) (also see He and Zhang 2005, 2006a) for computing bidiagonal representations of *PH*-distributions and matrix-exponential distributions. Use of spectral polynomials and the derivatives of the distribution function at zero distinguishes our method.

Section 2 gives definitions and some preliminary results for development of the algorithm. Section 3 shows how to compute the algebraic degree of a Coxian distribution, which is a lower bound on the triangular order of Coxian distribution. In §4, we present a set of necessary and sufficient conditions for an ordered Coxian representation to be minimal. In §5, we show that a minimal ordered Coxian representation can be found by solving a series of nonlinear programs for any *PH*-distribution whose Laplace-Stieltjes transform has only real poles. Three numerical examples are presented in §6 to show the effectiveness of the algorithm and some geometric properties associated with Coxian distributions. Finally, in §7, we discuss future research.

2. *PH, ME*, and Coxian Distributions

A square matrix T with negative diagonal elements, nonnegative off-diagonal elements, nonpositive row sums, and at least one negative row sum is called a *sub-generator* in the general literature of Markov process. We shall call a sub-generator T of finite size a *PH-generator*. Define an infinitesimal generator for a continuous-time Markov chain with $m+1$ states

$$\begin{pmatrix} T & -T\mathbf{e} \\ 0 & 0 \end{pmatrix},$$

where state $m+1$ is an absorption state and $\mathbf{e}$ is the column vector with all elements being one. The matrix T is an $m \times m$ *PH*-generator. We assume that states $\{1, 2, \ldots, m\}$ are transient. Let $\boldsymbol{\alpha}$ be a nonnegative vector of size m for which the sum of its elements is at most one. We call the distribution of the absorption time of the Markov chain to state $m+1$, with initial distribution $(\boldsymbol{\alpha}, 1-\boldsymbol{\alpha}\mathbf{e})$, a *phase-type distribution* (*PH*-distribution). We call the pair $(\boldsymbol{\alpha}, T)$ a *PH-representation* of that *PH*-distribution. The number m is the order of the *PH*-representation $(\boldsymbol{\alpha}, T)$. The probability distribution function of the *PH*-distribution is given as $1-\boldsymbol{\alpha}\exp\{Tt\}\mathbf{e}$ for $t \geq 0$, and the density function is given as $-\boldsymbol{\alpha}\exp\{Tt\}T\mathbf{e}$ for $t \geq 0$. If $\boldsymbol{\alpha}\mathbf{e} = 0$, the distribution has a unit mass at time 0. There is no need for a *PH*-representation for such a distribution. If $\boldsymbol{\alpha}\mathbf{e} \neq 0$, the expression $\boldsymbol{\alpha}\exp\{Tt\}\mathbf{e}$ can be written as $(\boldsymbol{\alpha}\mathbf{e})(\boldsymbol{\alpha}/(\boldsymbol{\alpha}\mathbf{e}))\exp\{Tt\}\mathbf{e}$. Thus, the study on the representation of $(\boldsymbol{\alpha}, T)$ is equivalent to that of $(\boldsymbol{\alpha}/(\boldsymbol{\alpha}\mathbf{e}), T)$. Without loss of generality, we assume that $\boldsymbol{\alpha}$ is a vector for which the sum of all its elements is one. That implies that all probability distributions we consider have a zero mass at $t=0$. See Chapter 2 in Neuts (1981) for basic properties of *PH*-distributions.

It is possible that $1-\boldsymbol{\alpha}\exp\{Tt\}\mathbf{u}$ is a probability distribution function for a row vector $\boldsymbol{\alpha}$ of size m, an $m \times m$ matrix T, and a column vector $\mathbf{u}$ of size m, where the elements of $\boldsymbol{\alpha}$, T, and $\mathbf{u}$ can be complex numbers. For this case, the 3-tuple $(\boldsymbol{\alpha}, T, \mathbf{u})$ is called a *matrix-exponential representation* (*ME*-representation) of a *matrix-exponential distribution* (*ME*-distribution). Without loss of generality, we assume that $\boldsymbol{\alpha}\mathbf{u} = 1$. Apparently, the class of *PH*-representations is a subset of the class of *ME*-representations. See Asmussen and

Bladt (1996) and Lipsky (1992) for more details about *ME*-distributions and their applications in queueing theory.

Throughout this paper, when $(\boldsymbol{\alpha}, T)$ is used, $\boldsymbol{\alpha}$ is nonnegative, T is a *PH*-generator, and $(\boldsymbol{\alpha}, T)$ is a *PH*-representation of a *PH*-distribution. When $(\boldsymbol{\alpha}, T, \mathbf{u})$ is used, $\boldsymbol{\alpha}$ may not be nonnegative and T may not be a *PH*-generator.

For $\mathbf{x} = (x_1, x_2, \ldots, x_N)$ where N is a positive integer, a bidiagonal matrix $S(\mathbf{x})$ is defined as

$$S(\mathbf{x}) = \begin{pmatrix} -x_1 & 0 & \cdots & \cdots & 0 \\ x_2 & -x_2 & \ddots & \ddots & \vdots \\ 0 & \ddots & \ddots & \ddots & \vdots \\ \vdots & \ddots & x_{N-1} & -x_{N-1} & 0 \\ 0 & \cdots & 0 & x_N & -x_N \end{pmatrix}. \quad (1)$$

If $S(\mathbf{x})$ is a *PH*-generator, it is called a *bidiagonal PH-generator*. If $\{x_1, x_2 \ldots, x_N\}$ are all real and positive and $\boldsymbol{\beta}$ is a probability vector (i.e., $\boldsymbol{\beta} \geq 0$ and $\boldsymbol{\beta}\mathbf{e} = 1$), then $(\boldsymbol{\beta}, S(\mathbf{x}))$ is called a *Coxian representation*, which represents a *Coxian distribution*. The class of Coxian distributions is a subset of the class of *PH*-distributions. Further, if $x_1 \geq x_2 \geq \cdots \geq x_N > 0$, then $(\boldsymbol{\beta}, S(\mathbf{x}))$ is called an *ordered Coxian representation*. The following theorem characterizes the class of Coxian distributions.

Theorem 1 (Theorem 4.1 in O'Cinneide 1991 and Theorem 5.2 in O'Cinneide 1993). *Every PH-distribution whose Laplace-Stieltjes transform has only real poles is a Coxian distribution and has an ordered Coxian representation.* □

Theorem 1 implies that the class of (standard) Coxian distributions defined by (1) covers many well-known probability distributions. For instance, generalized Erlang distributions, mixtures of exponential distributions, *PH*-distributions with a triangular *PH*-generator (Cumani 1982), and *PH*-distributions with a symmetric *PH*-generator (He and Zhang 2006b) are all special cases of Coxian distributions.

It is well known that the *PH*-representation of a *PH*-distribution and the ordered Coxian representation of a Coxian distribution are not unique. A *PH*-representation with the minimal number of phases is called *a minimal PH-representation*. The number of phases of a minimal *PH*-representation is called the *PH-order* of the corresponding *PH*-distribution. An ordered Coxian representation with the minimal number of phases is called *a minimal ordered Coxian representation*. The number of phases of a minimal ordered Coxian representation is called the *triangular order* of the corresponding *PH*-distribution. Theorem 6.2 in O'Cinneide (1993) gave a necessary and sufficient condition for an ordered Coxian representation to be minimal. In the following, a new set of necessary and sufficient conditions is obtained for an ordered Coxian representation to be minimal, and a series of nonlinear programs is developed for computing a minimal ordered Coxian representation.

In general, a *PH*-representation that represents a Coxian distribution has Coxian representations, but the orders of the Coxian representations may be larger than the order of the *PH*-representation (see Theorem 4.5 in He and Zhang 2006b and Example 1 in this paper).

The following *spectral polynomial algorithm* introduced in He and Zhang 2006b is useful for computing bidiagonal representations of *ME*-distributions and *PH*-distributions. Let $(\boldsymbol{\alpha}, T, \mathbf{u})$ be an *ME*-representation of order m. Denote by $\mathbf{x} = (x_1, x_2, \ldots, x_N)$ a vector whose elements are complex numbers, where N is a positive integer. If all elements of $\mathbf{x}$ are nonzero, define

$$\begin{cases} \mathbf{p}_1 = -T\mathbf{u}/x_1; \\ \mathbf{p}_n = (x_{n-1}I + T)\mathbf{p}_{n-1}/x_n, \quad 2 \leq n \leq N; \\ \mathbf{p}_{N+1} = (x_N I + T)\mathbf{p}_N. \end{cases} \quad (2)$$

Let $P = (\mathbf{p}_1, \mathbf{p}_2, \ldots, \mathbf{p}_N)$, which is an $m \times N$ matrix. Similarly to the proof of Proposition 3.1 in He and Zhang (2006b), the following results can be proved.

Theorem 2. *If* $\mathbf{p}_{N+1} = 0$, *we have* $TP = PS(\mathbf{x})$ *and* $P\mathbf{e} = \mathbf{u}$. *Further, the matrix-exponential representation* $(\boldsymbol{\alpha}, T, \mathbf{u})$ *has an equivalent bidiagonal representation* $(\boldsymbol{\beta}, S(\mathbf{x}), \mathbf{e})$ *of order* N *with* $\boldsymbol{\beta}\mathbf{e} = \boldsymbol{\alpha}\mathbf{u} = 1$ *and* $\boldsymbol{\beta} = \boldsymbol{\alpha}P$. □

The above spectral polynomial algorithm is called the *Post-T spectral polynomial algorithm* in He and Zhang (2006b). We remove "Post-*T*" because it is nonessential.

For any two vectors $\mathbf{x}$ and $\mathbf{y}$, we say that $\mathbf{x}$ is *part of* $\mathbf{y}$ if any element that appears k (≥ 0) times in $\mathbf{x}$ appears at least k times in $\mathbf{y}$. A particular choice of $\mathbf{x}$ was given in He and Zhang (2006b). Denote by $\{-\lambda_1, -\lambda_2, \ldots, -\lambda_m\}$ the spectrum of T (counting multiplicities, i.e., with repeated elements). If $(\lambda_1, \lambda_2, \ldots, \lambda_m)$ is part of $\mathbf{x}$, by the Cayley-Hamilton theorem (Lancaster and Tismenetsky 1985), we must have $\mathbf{p}_{N+1} = 0$. If $x_1 \geq x_2 \geq \cdots \geq x_N > 0$ and $\boldsymbol{\alpha}P$ is nonnegative, $(\boldsymbol{\alpha}P, S(\mathbf{x}))$ is an ordered Coxian representation. For this case, an ordered Coxian representation is obtained for $(\boldsymbol{\alpha}, T, \mathbf{u})$. For more about the spectral polynomial algorithm, see He and Zhang (2005; 2006a, b).

In O'Cinneide (1989), a *PH*-generator T is called *PH-simple* if each *PH*-distribution $(\boldsymbol{\alpha}, T)$ has a unique representation of the form $(\boldsymbol{\alpha}, T)$, i.e., if $\boldsymbol{\alpha}_1$ and $\boldsymbol{\alpha}_2$ are two probability vectors and $\boldsymbol{\alpha}_1 \neq \boldsymbol{\alpha}_2$, then $(\boldsymbol{\alpha}_1, T)$ and $(\boldsymbol{\alpha}_2, T)$ represent two different *PH*-distributions. By

Theorem 1 in O'Cinneide (1989), a *PH*-generator T is *PH*-simple if and only if $\{T\mathbf{e}, T^2\mathbf{e}, \ldots, T^m\mathbf{e}\}$ are independent vectors. By Theorem 2 in O'Cinneide (1989), a *PH*-generator T is *PH*-simple if and only if it has no left eigenvector orthogonal to $\mathbf{e}$.

The above definition of *PH*-simplicity can be generalized to *ME*-representations of the form $(\boldsymbol{\alpha}, T, \mathbf{e})$, where T is a *PH*-generator, and $\boldsymbol{\alpha}\mathbf{e} = 1$ (but $\boldsymbol{\alpha}$ may not be nonnegative). That is, a *PH*-generator T is *PH*-simple if each *ME*-distribution $(\boldsymbol{\alpha}, T, \mathbf{e})$ has a unique representation of the form $(\boldsymbol{\alpha}, T, \mathbf{e})$. This generalization of simplicity is exactly the same as the property of simplicity for the case of a *PH*-generator. It is easy to check that both Theorems 1 and 2 in O'Cinneide (1989) are still true under the generalized definition of *PH*-simplicity. Next, we establish a relationship between *PH*-simplicity and the spectral polynomial algorithm.

Theorem 3. *Assume that* (1) *T is PH-simple*; (2) $\mathbf{u} = \mathbf{e}$; (3) *all elements of* $\mathbf{x}$ *are nonzero*; *and* (4) $N = m$. *Then the matrix P obtained from* (2) *is invertible.*

Proof. From (2), we obtain

$$\mathbf{p}_n = c_{n,1}T\mathbf{e} + c_{n,2}T^2\mathbf{e} + \cdots + c_{n,n}T^n\mathbf{e}, \quad 1 \le n \le m,$$

where $\{c_{n,j}, 1 \le j \le n \le m\}$ are some constants and $c_{n,n} \neq 0$, for $1 \le n \le m$. Since T is *PH*-simple, by Theorem 1 in O'Cinneide (1989), vectors $T\mathbf{e}, T^2\mathbf{e}, \ldots, T^m\mathbf{e}$ are independent. Together with $c_{n,n} \neq 0$, for $1 \le n \le m$, it is easy to see that the vectors $\mathbf{p}_1, \mathbf{p}_2, \ldots, \mathbf{p}_m$ are independent, and consequently, the matrix P is invertible. □

3. A Minimal *ME*-Representation

We assume that the Laplace-Stieltjes transform of a *PH*-distribution $(\boldsymbol{\alpha}, T)$ of order m has only real poles. According to Theorem 1, $(\boldsymbol{\alpha}, T)$ represents a Coxian distribution. We will develop a method for finding the triangular order of $(\boldsymbol{\alpha}, T)$ and for computing a minimal ordered Coxian representation of $(\boldsymbol{\alpha}, T)$. We begin with computation of a minimal *ME*-representation.

We assume that the *PH*-generator T has in total K distinct eigenvalues $\{-\mu_1, -\mu_2, \ldots, -\mu_K\}$. Assume that the algebraic multiplicity of the eigenvalue $-\mu_k$ is m_k, i.e., $-\mu_k$ appears a total of m_k times in the vector $(-\lambda_1, -\lambda_2, \ldots, -\lambda_m)$. Let $\boldsymbol{\mu} = (\mu_1, \mu_2, \ldots, \mu_K)$ and $\mathbf{m} = (m_1, m_2, \ldots, m_K)$. Define

$$S(\boldsymbol{\mu}, \mathbf{m}) = \begin{pmatrix} E(\mu_1, m_1) & & \\ & \ddots & \\ & & E(\mu_K, m_K) \end{pmatrix},$$

where the matrix $E(\mu_k, m_k) = S((\mu_k, \mu_k, \ldots, \mu_k))$ is of size m_k, $1 \le k \le K$. Note that if μ_k is positive real, $E(\mu_k, m_k)$ is the *PH*-generator of an Erlang distribution of order m_k with parameter μ_k.

Lemma 1. *If $\mu_1, \mu_2, \ldots, \mu_K$ are distinct, then no left eigenvector of $S(\boldsymbol{\mu}, \mathbf{m})$ is orthogonal to the vector* $\mathbf{e}$. *Furthermore, if $\mu_1, \mu_2, \ldots, \mu_K$ are distinct positive real numbers, then the PH-generator $S(\boldsymbol{\mu}, \mathbf{m})$ is PH-simple.*

Proof. Suppose that vector $\mathbf{v}$ is a left eigenvector of $S(\boldsymbol{\mu}, \mathbf{m})$ corresponding to eigenvalue μ_k. Since $\mu_1, \mu_2, \ldots, \mu_K$ are distinct, $\mathbf{v}$ has structure $(0, \ldots, 0, \mathbf{v}_k, 0, \ldots 0)$, where $\mathbf{v}_k$ is an eigenvector of the matrix $E(\mu_k, m_k)$. Post-multiplying $\mathbf{e}$ on both sides of $-\mu_k\mathbf{v}_k = \mathbf{v}_k E(\mu_k, m_k)$, we obtain

$$-\mu_k\mathbf{v}_k\mathbf{e} = \mathbf{v}_k E(\mu_k, m_k)\mathbf{e} = -\mu_k\mathbf{v}_k(1, 0, \ldots, 0)', \qquad (3)$$

where $(1, 0, \ldots, 0)'$ is the transpose of $(1, 0, \ldots, 0)$. If $\mathbf{v}\mathbf{e} = 0$, then $\mathbf{v}_k\mathbf{e} = 0$. Equation (3) implies that the first element of $\mathbf{v}_k$ is zero. From $-\mu_k\mathbf{v}_k = \mathbf{v}_k E(\mu_k, m_k)$, it can be shown that the vector $\mathbf{v}_k$ is zero. Thus, $\mathbf{v}$ is zero, which is a contradiction. Therefore, no left eigenvector of $S(\boldsymbol{\mu}, \mathbf{m})$ is orthogonal to $\mathbf{e}$.

If $\mu_1, \mu_2, \ldots, \mu_K$ are distinct positive real numbers, then $S(\boldsymbol{\mu}, \mathbf{m})$ is a *PH*-generator. Because no left eigenvector of $S(\boldsymbol{\mu}, \mathbf{m})$ is orthogonal to $\mathbf{e}$, by Theorem 2 in O'Cinneide (1989), $S(\boldsymbol{\mu}, \mathbf{m})$ is *PH*-simple. □

Next, we use the spectral polynomial algorithm to find an expression for the distribution function of $(\boldsymbol{\alpha}, T)$ in a few steps.

1. By using the spectral polynomial algorithm, we obtain $TP(\boldsymbol{\lambda}) = P(\boldsymbol{\lambda})S(\boldsymbol{\lambda})$, where $P(\boldsymbol{\lambda})$ is an $m \times m$ matrix with unit row sums and $\boldsymbol{\lambda} = (\lambda_1, \lambda_2, \ldots, \lambda_m)$. (See Theorem 2 and the ensuing discussion.)
2. By using the spectral polynomial algorithm again, we obtain $S(\boldsymbol{\mu}, \mathbf{m})P_1 = P_1 S(\boldsymbol{\lambda})$, where P_1 is a matrix with unit row sums. By Lemma 1, similarly to the proof of Theorem 3, it can be shown that P_1 is invertible. Then the equation $S(\boldsymbol{\mu}, \mathbf{m})P_1 = P_1 S(\boldsymbol{\lambda})$ can be written as $P_1^{-1}S(\boldsymbol{\mu}, \mathbf{m}) = S(\boldsymbol{\lambda})P_1^{-1}$.
3. Combining the above results, we obtain $TP(\boldsymbol{\lambda}) \cdot P_1^{-1} = P(\boldsymbol{\lambda})P_1^{-1}S(\boldsymbol{\mu}, \mathbf{m})$ and $P(\boldsymbol{\lambda})P_1^{-1}\mathbf{e} = \mathbf{e}$. It is easy to check that $(\boldsymbol{\beta}, S(\boldsymbol{\mu}, \mathbf{m}), \mathbf{e})$ is an *ME*-representation of $(\boldsymbol{\alpha}, T)$, where $\boldsymbol{\beta} = \boldsymbol{\alpha}P(\boldsymbol{\lambda})P_1^{-1}$ (see the proof of Proposition 2.1 in He and Zhang 2006b).

Lemma 2. *The distribution function $F(t)$ of $(\boldsymbol{\alpha}, T)$ can be obtained as* (*note* $m_0 = 0$)

$$F(t) = 1 - \sum_{k=1}^{K}\left(\sum_{i=0}^{m_k-1}\frac{t^i\mu_k^i}{i!}\left(\sum_{j=m_1+\cdots+m_{k-1}+1+i}^{m_1+\cdots+m_{k-1}+m_k}\beta_j\right)\right)\exp\{-\mu_k t\}, \quad t \ge 0.$$

Proof. By definitions and the structure of $S(\boldsymbol{\mu}, \mathbf{m})$, the distribution function $F(t)$ of $(\boldsymbol{\alpha}, T)$ can be obtained as

$$\begin{aligned} F(t) &= 1 - \boldsymbol{\alpha}\exp\{Tt\}\mathbf{e} \\ &= 1 - \boldsymbol{\beta}\exp\{S(\boldsymbol{\mu}, \mathbf{m})t\}\mathbf{e} \\ &= 1 - \sum_{k=1}^{K}(\beta_{m_1+\cdots+m_{k-1}+1}, \beta_{m_1+\cdots+m_{k-1}+2}, \ldots, \\ &\qquad \beta_{m_1+\cdots+m_{k-1}+m_k})\exp\{E(\mu_k, m_k)t\}\mathbf{e}, \end{aligned} \tag{4}$$

which leads to (4) by routine calculations. □

The *PH*-generator T may have non-real eigenvalues when the Laplace-Stieltjes transform of $(\boldsymbol{\alpha}, T)$ has only real poles. However, because the Laplace-Stieltjes transform of the *PH*-distribution $(\boldsymbol{\alpha}, T)$ has only real poles, the coefficient of $\exp\{-\mu_k t\}$ in (4) must be zero if $-\mu_k$ is not a real number. Without loss of generality and to simplify notation, we assume that all K eigenvalues $\{-\mu_1, -\mu_2, \ldots, -\mu_K\}$ in (4) are real. Since T is a *PH*-generator, all its eigenvalues must have negative real parts (Minc 1988, Seneta 1973). Thus, $-\mu_1, -\mu_2, \ldots, -\mu_K$ are negative real numbers and, consequently, $S(\boldsymbol{\mu}, \mathbf{m})$ is a *PH*-generator. Further, if $\beta_{m_1+\cdots+m_{k-1}+i} = 0$, $1 \le i \le m_k$, we remove all items associated with $-\mu_k$. Thus, we assume that at least one of $\{\beta_{m_1+\cdots+m_{k-1}+i}, i = 1, 2, \ldots, m_k\}$ is not zero, for $k = 1, 2, \ldots, K$. If $\beta_{m_1+\cdots+m_k} = 0$, then we choose the largest i such that $\beta_{m_1+\cdots+m_{k-1}+i} \ne 0$, $1 \le i \le m_k$, and redefine m_k to i to ensure $\beta_{m_1+\cdots+m_k} \ne 0$. Let

$$N_m = m_1 + m_2 + \cdots + m_K, \tag{5}$$

which is the *algebraic degree* of the corresponding *PH*-distribution (O'Cinneide 1993).

It is readily seen that $N_m \le m$. (4) indicates that $F(t)$ is associated with K Jordan blocks corresponding to distinct real eigenvalues $\{-\mu_1, -\mu_2, \ldots, -\mu_K\}$ of T. Next, we use the expression of $F(t)$ given in (4) to prove that the algebraic degree N_m is a lower bound on the triangular order of the *PH*-distribution $(\boldsymbol{\alpha}, T)$. It is well known that the triangular order and *PH*-order of a *PH*-distribution are as large as the algebraic degree of the *PH*-distribution (O'Cinneide 1993). We provide the following proof for completeness.

Lemma 3. *Consider a PH-distribution with a PH-representation $(\boldsymbol{\alpha}, T)$ of order m. We assume $\mu_1 > \mu_2 > \cdots > \mu_K$ and $\beta_{m_1+\cdots+m_k} \ne 0$ for $k = 1, 2, \ldots, K$. Then any ME-representation of the PH-distribution $(\boldsymbol{\alpha}, T)$ must have at least N_m phases. Consequently, the triangular order and the PH-order of $(\boldsymbol{\alpha}, T)$ are larger than or equal to N_m.*

Proof. By (4), since $\min\{\mu_1, \ldots, \mu_{K-1}\} > \mu_K$, we have

$$\lim_{t\to\infty} \frac{F(t)}{\frac{t^{m_K-1}\mu_K^{m_K-1}}{(m_K-1)!}\exp\{-\mu_K t\}} = \beta_{m_1+m_2+\cdots+m_K}. \tag{6}$$

Similarly to (6), for $1 \le k \le K-1$, we have

$$\lim_{t\to\infty} \frac{F(t) - \sum_{u=k+1}^{K}\sum_{i=0}^{m_u-1}\frac{t^i\mu_u^i}{i!}\left(\sum_{j=m_1+\cdots+m_{u-1}+1+i}^{m_1+\cdots+m_{u-1}+m_u}\beta_j\right)\exp\{-\mu_u t\}}{\frac{t^{m_k-1}\mu_k^{m_k-1}}{(m_k-1)!}\exp\{-\mu_k t\}} = \beta_{m_1+\cdots+m_k}.$$

Since $\beta_{m_1+\cdots+m_k} \ne 0$ for $1 \le k \le K$, it is easy to see that any expression of the distribution function of $(\boldsymbol{\alpha}, T)$ must be equivalent to the one given in (4). Thus, the Laplace-Stieltjes transform of any representation of $F(t)$ must have the poles $-\mu_1, -\mu_2, \ldots$, and $-\mu_K$ with multiplicities $m_1, m_2, \ldots$, and m_K, respectively. Therefore, any *ME*-representation of $(\boldsymbol{\alpha}, T)$ must have at least N_m phases. Consequently, the triangular order and the *PH*-order of $(\boldsymbol{\alpha}, T)$ are larger than or equal to N_m. □

According to Lemma 3, $(\boldsymbol{\beta}, S(\boldsymbol{\mu}, \mathbf{m}), \mathbf{e})$ is a minimal *ME*-representation of $(\boldsymbol{\alpha}, T)$ and the 3-tuple $\{\boldsymbol{\beta}, \boldsymbol{\mu}, \mathbf{m}\}$ is invariant to the probability distribution $(\boldsymbol{\alpha}, T)$.

4. A Set of Necessary and Sufficient Conditions

In this section, we use $\{\boldsymbol{\beta}, \boldsymbol{\mu}, \mathbf{m}\}$ defined in §3 to find a set of necessary and sufficient conditions for a Coxian representation of $(\boldsymbol{\alpha}, T)$ to be minimal.

We begin with the construction of an *ME*-representation with an ordered Coxian generator from $\{\boldsymbol{\beta}, \boldsymbol{\mu}, \mathbf{m}\}$. Let $\boldsymbol{\eta} = (\eta_1, \eta_2, \ldots, \eta_{N_m})$ with

$$\eta_j = \mu_k \quad \text{for } m_1+m_2+\cdots+m_{k-1}+1 \le j \le m_1+m_2+\cdots+m_k \text{ and } 1 \le k \le K.$$

Since $\mu_1 > \mu_2 > \cdots > \mu_K$, the elements of $\boldsymbol{\eta}$ are in nonincreasing order. By using the spectral polynomial algorithm, we can find $P(\boldsymbol{\eta})$ satisfying $S(\boldsymbol{\mu}, \mathbf{m})P(\boldsymbol{\eta}) = P(\boldsymbol{\eta})S(\boldsymbol{\eta})$ and $P(\boldsymbol{\eta})\mathbf{e} = \mathbf{e}$. Since $S(\boldsymbol{\mu}, \mathbf{m})$ is lower triangular, by Corollary 2 in O'Cinneide (1989) (also see Theorem 4.3 in He and Zhang 2006b), $P(\boldsymbol{\eta})$ is nonnegative so it is a stochastic matrix. Because all elements of $\boldsymbol{\eta}$ are positive, the determinant of

$$(S(\boldsymbol{\eta})\mathbf{e}, (S(\boldsymbol{\eta}))^2\mathbf{e}, \ldots, (S(\boldsymbol{\eta}))^{N_m}\mathbf{e})$$
$$= -\begin{pmatrix} \eta_1 & -\eta_1^2 & \eta_1^3 & \cdots & (-1)^{N_m-1}\eta_1^{N_m} \\ 0 & \eta_1\eta_2 & -(\eta_1^2\eta_2+\eta_1\eta_2^2) & \cdots & \vdots \\ 0 & 0 & \eta_1\eta_2\eta_3 & \cdots & \vdots \\ \vdots & \ddots & \ddots & \ddots & \vdots \\ 0 & \cdots & 0 & 0 & \eta_1\eta_2\eta_3\cdots\eta_{N_m} \end{pmatrix}$$

is nonzero. Thus, the vectors $\{S(\boldsymbol{\eta})\mathbf{e}, S(\boldsymbol{\eta})^2\mathbf{e}, \ldots, S(\boldsymbol{\eta})^{N_m}\mathbf{e}\}$ are independent. Therefore, by Theorem 1 in O'Cinneide (1989), the *PH*-generator $S(\boldsymbol{\eta})$ is *PH*-simple.

According to Theorem 2, the *PH*-representation $(\boldsymbol{\alpha}, T)$ has an equivalent *ME*-representation $(\boldsymbol{\beta}P(\boldsymbol{\eta}), S(\boldsymbol{\eta}), \mathbf{e})$, where $S(\boldsymbol{\eta})$ is an ordered Coxian generator. If $\boldsymbol{\beta}P(\boldsymbol{\eta})$ is nonnegative, $(\boldsymbol{\beta}P(\boldsymbol{\eta}), S(\boldsymbol{\eta}))$ is an ordered Coxian representation. By Lemma 3, $(\boldsymbol{\beta}P(\boldsymbol{\eta}), S(\boldsymbol{\eta}))$ is a minimal ordered Coxian representation of $(\boldsymbol{\alpha}, T)$ and the triangular order of $(\boldsymbol{\alpha}, T)$ is N_m. If $\boldsymbol{\beta}P(\boldsymbol{\eta})$ is not nonnegative, however, the triangular order of $(\boldsymbol{\alpha}, T)$ must be larger than N_m. The reason is that by Lemma 3, any ordered Coxian representation of $(\boldsymbol{\alpha}, T)$ of order N_m must have a representation of the form $(\mathbf{y}, S(\boldsymbol{\eta}), \mathbf{e})$. Since $S(\boldsymbol{\eta})$ is *PH*-simple, we must have $\mathbf{y} = \boldsymbol{\beta}P(\boldsymbol{\eta})$, which is not nonnegative. Therefore, the triangular order of $(\boldsymbol{\alpha}, T)$ is larger than N_m.

Suppose that $(\boldsymbol{\gamma}, S(\mathbf{x}), \mathbf{e})$ of order n $(\geq N_m)$ is an *ME*-representation of $(\boldsymbol{\alpha}, T)$, where $\mathbf{x}$ equals $(x_1, x_2, \ldots, x_n)$ with positive elements. By Lemma 3, it is apparent that $(\eta_1, \eta_2, \ldots, \eta_{N_m})$ must be part of $(x_1, x_2, \ldots, x_n)$. By using the spectral polynomial algorithm, we obtain $S(\boldsymbol{\mu}, \mathbf{m})P(\mathbf{x}) = P(\mathbf{x})S(\mathbf{x})$, $\boldsymbol{\gamma} = \boldsymbol{\beta}P(\mathbf{x})$, and $P(\mathbf{x}) = (\mathbf{p}_1(\mathbf{x}), \mathbf{p}_2(\mathbf{x}), \ldots, \mathbf{p}_n(\mathbf{x}))$, where

$$\begin{aligned}
\mathbf{p}_1(\mathbf{x}) &= -S(\boldsymbol{\mu}, \mathbf{m})\mathbf{e}/x_1;\\
\mathbf{p}_j(\mathbf{x}) &= \frac{(-1)}{x_j \cdots x_1}(x_{j-1}I + S(\boldsymbol{\mu}, \mathbf{m})) \cdots (x_1 I + S(\boldsymbol{\mu}, \mathbf{m}))\\
&\quad \cdot S(\boldsymbol{\mu}, \mathbf{m})\mathbf{e}, \quad 2 \leq j \leq n; \qquad (7)\\
0 &= \frac{(-1)}{x_n \cdots x_1}(x_n I + S(\boldsymbol{\mu}, \mathbf{m})) \cdots (x_1 I + S(\boldsymbol{\mu}, \mathbf{m}))\\
&\quad \cdot S(\boldsymbol{\mu}, \mathbf{m})\mathbf{e}.
\end{aligned}$$

Our goal is to find the smallest n such that $(\boldsymbol{\gamma}, S(\mathbf{x}), \mathbf{e})$ is an ordered Coxian representation. To that end, we first show that $\boldsymbol{\gamma}$ can be obtained from $\mathbf{x}$ and the original representation $(\boldsymbol{\alpha}, T)$ (more specifically, the derivatives of the *PH*-distribution at zero). Let $\Omega(k, j)$ be the set of all the subsets of $\{1, 2, \ldots, j\}$ with exactly k elements, for $k \leq j$. Define

$$\begin{aligned}
h_{i,1} &= F^{(i)}(0), \quad 1 \leq i \leq n;\\
h_{i,j} &= \sum_{k=1}^{j-1} F^{(k+i-1)}(0)\left(\sum_{\{i_1, \ldots, i_{j-k}\} \subseteq \Omega(j-k, j-1)} x_{i_1} \cdots x_{i_{j-k}}\right) \qquad (8)\\
&\quad + F^{(j+i-1)}(0), \quad 3 \leq i+j \leq n+1,
\end{aligned}$$

where $F^{(j)}(0)$ is the j-th derivative of $F(t)$ at $t = 0$ for $j \geq 0$. Note that

$$F(t) = 1 - \boldsymbol{\alpha}\exp\{Tt\}\mathbf{e} = 1 - \boldsymbol{\beta}\exp\{S(\boldsymbol{\mu}, \mathbf{m})t\}\mathbf{e}$$

and

$$F^{(j)}(0) = -\boldsymbol{\alpha}T^j\mathbf{e} = -\boldsymbol{\beta}(S(\boldsymbol{\mu}, \mathbf{m}))^j\mathbf{e}, \quad j \geq 1.$$

The relationship between $\boldsymbol{\gamma}$ and $\{h_{i,j}, i, j \geq 1\}$ is given as follows.

Lemma 4. *For ME-representation $(\boldsymbol{\gamma}, S(\mathbf{x}), \mathbf{e})$ defined above, we have $\gamma_j = h_{1,j}/(x_1 x_2 \cdots x_j)$, $1 \leq j \leq n$.*

Proof. Since $\boldsymbol{\gamma} = \boldsymbol{\beta}P(\mathbf{x})$, we have $\gamma_j = \boldsymbol{\beta}\mathbf{p}_j(\mathbf{x})$, $1 \leq j \leq n$. The results are obtained by (7) and (8) and routine calculations. □

Lemma 5. *The ME-representation $(\boldsymbol{\gamma}, S(\mathbf{x}), \mathbf{e})$ defined above is an ordered Coxian representation of $(\boldsymbol{\alpha}, T)$ if and only if*

(1) $x_1 \geq x_2 \geq \cdots \geq x_n > 0$;
(2) $\boldsymbol{\eta}$ *is part of* $\mathbf{x}$; *and*
(3) $h_{1,j} \geq 0$, $1 \leq j \leq n$.

Proof. If $(\boldsymbol{\gamma}, S(\mathbf{x}), \mathbf{e})$ is an ordered Coxian representation of $(\boldsymbol{\alpha}, T)$, we must have $x_1 \geq x_2 \geq \cdots \geq x_n > 0$. Thus, condition (1) holds. By Lemma 2, $\boldsymbol{\eta}$ must be part of $\mathbf{x}$, which implies condition (2). Since $\boldsymbol{\gamma} \geq 0$, by Lemma 4, condition (3) holds. This proves the necessity of the conditions.

On the other hand, suppose that conditions (1), (2), and (3) hold. Condition (1) implies that $S(\mathbf{x})$ is a *PH*-generator. Condition (2) implies $S(\boldsymbol{\mu}, \mathbf{m})P(\mathbf{x}) = P(\mathbf{x})S(\mathbf{x})$ and $P(\mathbf{x})\mathbf{e} = \mathbf{e}$. Condition (3) implies that $\boldsymbol{\gamma} = \boldsymbol{\beta}P(\mathbf{x})$ is nonnegative. In conclusion, $(\boldsymbol{\gamma}, S(\mathbf{x}), \mathbf{e})$ is an equivalent ordered Coxian representation of $(\boldsymbol{\alpha}, T)$. This proves the sufficiency of the conditions. □

Theorem 6.2 in O'Cinneide (1993) provides a necessary and sufficient condition for a *PH*-representation with a triangular *PH*-generator to be minimal. Compared to that condition, the conditions given in Lemma 5 are easier to check numerically. In fact, the conditions in Lemma 5 lead to an algorithm for computing a minimal Coxian representation for any Coxian distribution. Lemma 5 establishes a relationship between the Coxian representation of a Coxian distribution and the derivatives of its distribution function at zero. Now, we are ready to state and prove the main result of this section.

Theorem 4. *Let n^* be the minimal n such that conditions* (1), (2), *and* (3) *in Lemma* 5 *are satisfied. Then n^* is the triangular order of $(\boldsymbol{\alpha}, T)$. The number n^* is finite and is at least N_m (defined in* (5)*). The corresponding representation $(\boldsymbol{\gamma}, S(\mathbf{x}), \mathbf{e})$ is a minimal ordered Coxian representation of $(\boldsymbol{\alpha}, T)$.*

Proof. The first conclusion is obvious from Lemma 5. By Lemma 3, we must have $n^* \geq N_m$. According to Theorem 1, any *PH*-distribution with only real poles has a triangular representation or, equivalently, an ordered Coxian representation. Therefore, n^* is finite. □

It is clear from the proofs of the above lemmas that Theorem 4 is valid for any Coxian distribution with an *ME*-representation $(\boldsymbol{\alpha}, T, \mathbf{e})$, where $\boldsymbol{\alpha}$ may not be nonnegative. Consequently, the algorithm developed in §5 can be used for computing minimal ordered Coxian representations for Coxian distributions with *ME*-representations $(\boldsymbol{\alpha}, T, \mathbf{e})$ (see Example 1, §6).

5. A Nonlinear-Programming Approach

Based on Lemma 5 and Theorem 4, a nonlinear system can be developed for computing a minimal ordered Coxian representation for $(\boldsymbol{\alpha}, T)$. First, we prove a property for minimal ordered Coxian representations.

LEMMA 6. *For a minimal ordered Coxian representation* $(\boldsymbol{\gamma}, S(\mathbf{x}))$ *of order* n^*, *we must have* $x_j \geq \eta_{N_m}$, $1 \leq j \leq n^*$.

PROOF. Note that $x_1 \geq x_2 \geq \cdots \geq x_{n^*}$. If there exists x_j such that $x_j < \eta_{N_m}$, then x_{n^*} is the smallest one with that property. Suppose that $x_{n^*-1} > x_{n^*}$. By routine calculations, we can obtain

$$\boldsymbol{\gamma}\exp\{S(x)t\}\mathbf{e} = \sum_{j=1}^{n^*-1} c_j(t)\exp\{-x_j t\} + \gamma_{n^*}\exp\{-x_{n^*}t\},$$

where $\{c_j(t), 1 \leq j \leq n^*\}$ are some polynomials of a finite order. By the proof of Lemma 3, we must have $\gamma_{n^*} = 0$. Thus, the ordered Coxian representations $((\gamma_1, \gamma_2, \ldots, \gamma_{n^*-1}), S(x_1, x_2, \ldots, x_{n^*-1}))$ and $(\boldsymbol{\gamma}, S(\mathbf{x}))$ represent the same probability distribution. This further implies that $(\boldsymbol{\gamma}, S(\mathbf{x}))$ is not a minimal Coxian representation, which gives a contradiction. Therefore, for a minimal Coxian representation $(\boldsymbol{\gamma}, S(\mathbf{x}))$, we must have $x_j \geq \eta_{N_m}$ for $1 \leq j \leq n^*$. The case $x_{n^*-1} = x_{n^*}$ can be proved similarly. □

Based on Lemma 5, Theorem 4, and Lemma 6, a Coxian representation can be constructed from any feasible solution to the following nonlinear system (NLS):

$$\begin{aligned}
&(1)\quad x_1 \geq x_2 \geq \cdots \geq x_n \geq \eta_{N_m}, \quad x_1 \geq \eta_1;\\
&(2)\quad (\eta_1, \eta_2, \ldots, \eta_{N_m}) \text{ is part of } (x_1, x_2, \ldots, x_n);\\
&(3)\quad h_{i,j} \geq 0, \quad 1 \leq j \leq n;\\
&(4)\quad (8) \text{ holds.}
\end{aligned} \tag{9}$$

Next, we discuss how to find a feasible solution to NLS (9). First, we can replace constraint (1) in (9) by $x_1 \geq \eta_1$, and $x_i \geq \eta_{N_m}$, $2 \leq i \leq n$. The solution to the modified nonlinear system (if exists) can be used to construct a Coxian representation, which may not be an ordered Coxian representation. By Property 5.1 in He and Zhang (2006b), we can use the spectral polynomial algorithm to find an ordered Coxian representation of the same order from that solution. Second, we rewrite (8) as the following recursive equations:

$$\begin{aligned}
h_{i,1} &= F^{(i)}(0), \quad 1 \leq i \leq n;\\
h_{i,j} &= x_{j-1}h_{i,j-1} + h_{i+1,j-1},\\
&\qquad 3 \leq j+i \leq n+1,\ j \geq 2,\ i \geq 1.
\end{aligned} \tag{10}$$

Equation (10) can be obtained in a straightforward manner from (8). Equalities in (10) are polynomials of degree 2 in $\{x_1, x_2, \ldots, x_n\}$ and $\{h_{i,j}, h_2, \ldots, h_{n+1}\}$, which is significantly smaller than that of (8). This may bring efficiency in solving the nonlinear system.

Let $H = \{h_{i,j}, 2 \leq i+j \leq n+1\}$. Combining Lemmas 5 and 6, Theorem 4, the above observations, and (10), we introduce the following nonlinear program (NLP) to find an ordered Coxian representation $(\boldsymbol{\gamma}, S(\mathbf{x}))$ for $(\boldsymbol{\alpha}, T)$ and a given n ($\geq N_m$):

$$\begin{aligned}
&\min_{(x,H)} \sum_{i=1}^{n} x_i\\
&\text{s.t.}\ (1)\quad x_1 \geq \eta_1, x_i \geq \eta_{N_m}, \quad \text{for } 2 \leq i \leq n;\\
&\qquad (2)\quad (\eta_1, \eta_2, \ldots, \eta_{N_m}) \text{ is part of } (x_1, x_2, \ldots, x_n);\\
&\qquad (3)\quad h_{1,j} \geq 0, \quad 1 \leq j \leq n;\\
&\qquad (4)\quad (10) \text{ holds.}
\end{aligned} \tag{11}$$

If NLP (11) has a solution, that solution corresponds to a Coxian representation of order n for $(\boldsymbol{\alpha}, T)$ (Lemma 5). By using the spectral polynomial algorithm, we can find an ordered Coxian representation of order n from that solution. Otherwise, the triangular order of $(\boldsymbol{\alpha}, T)$ is larger than n.

The derivative $F^{(j)}(0)$ can be extremely large or small for large or even moderate j, which can cause problems in computation. One suggestion to solve the problem is to rescale the time, i.e., to scale T to δT where δ is a positive number. For instance, we can choose $\delta^{-1} = \max_{1\leq i\leq m}\{|\lambda_i|\}$. Numerical examples demonstrate that such a change of scale makes the algorithm more stable.

From numerical experiments, we learned that choosing an appropriate initial search point is important for solving NLP (11) efficiently. An initial search point we use is $\mathbf{x} = (\lambda, \ldots, \lambda, \eta_1, \eta_2, \ldots, \eta_{N_m})$, where $\lambda > \eta_1$.

It is easy to see that the (minimal) ordered Coxian representation of $(\boldsymbol{\alpha}, T)$ may not be unique. In fact, any feasible solution to (11) corresponds to an ordered Coxian representation of $(\boldsymbol{\alpha}, T)$. Also, the optimal solution to (11) can be different if a different objective function is used. In (11), the objective function is a simple linear function. Such a linear function is chosen to make it easier to solve (11). The following property supports the nonuniqueness of minimal ordered Coxian representation.

LEMMA 7. *Suppose that*

$$(\boldsymbol{\gamma}(0), S(\mathbf{x}(0))) \quad and \quad (\boldsymbol{\gamma}(1), S(\mathbf{x}(1)))$$

are two ordered Coxian representations of order n *for* $(\boldsymbol{\alpha}, T)$, where $\mathbf{x}(0)$ *and* $\mathbf{x}(1)$ *are identical except for one element* (*i.e.*, $x_i(0) = x_i(1)$, $1 \leq i \leq n$ *and* $i \neq i_0$). *Then for any convex combination* $\mathbf{x}(\lambda) = \lambda\mathbf{x}(1) + (1-\lambda)\mathbf{x}(0)$ *with* $0 \leq \lambda \leq 1$, *there is an ordered Coxian representation* $(\boldsymbol{\gamma}(\lambda), S(\mathbf{x}(\lambda)))$ *of order* n *for* $(\boldsymbol{\alpha}, T)$.

Proof. Because all elements of $\mathbf{x}(0)$ and $\mathbf{x}(1)$ are positive, all elements of $\mathbf{x}(\lambda)$ are positive. Since both $\mathbf{x}(0)$ and $\mathbf{x}(1)$ are ordered sequences, $\mathbf{x}(\lambda)$ is an ordered sequence as well. By (8), $\{h_{i,j},\ 2 \le i+j \le n+1\}$, as functions of the elements of $\mathbf{x}$, are linear functions in each element of $\mathbf{x}$. Define $\{h_{i,j}(0),\ 2 \le i+j \le n+1\}$ by (10) for $\mathbf{x}(0)$, $\{h_{i,j}(1), 2 \le i+j \le n+1\}$ for $\mathbf{x}(1)$, and $\{h_{i,j}(\lambda),\ 2 \le i+j \le n+1\}$ for $\mathbf{x}(\lambda)$. Since $x_i(\lambda) = x_i(0) = x_i(1)$, $1 \le i \le n$ and $i \ne i_0$, and $x_{i_0}(\lambda) = \lambda x_{i_0}(1) + (1-\lambda)x_{i_0}(0)$, we have $h_{i,j}(\lambda) = \lambda h_{i,j}(1) + (1-\lambda) \cdot h_{i,j}(0)$, $2 \le i+j \le n+1$. Therefore, $\{h_{1,j}(\lambda),\ 1 \le j \le n\}$ are nonnegative. Let $\boldsymbol{\gamma}(\lambda) = (\gamma_1(\lambda), \gamma_2(\lambda), \ldots, \gamma_n(\lambda))$, where $\gamma_j(\lambda) = h_{1,j}(\lambda)/(x_1(\lambda)x_2(\lambda)\cdots x_j(\lambda))$, $1 \le j \le n$. Because $\boldsymbol{\eta}$ is part of $\mathbf{x}(0)$ and part of $\mathbf{x}(1)$, it is easy to see that $\boldsymbol{\eta}$ is part of $\mathbf{x}(\lambda)$. Thus, $\{\boldsymbol{\gamma}(\lambda), \mathbf{x}(\lambda)\}$ satisfies the conditions in Lemma 5. Therefore, $(\boldsymbol{\gamma}(\lambda), S(\mathbf{x}(\lambda)))$ is an ordered Coxian representation of $(\boldsymbol{\alpha}, T)$. □

According to Aldous and Shepp (1987), for any PH-distribution $(\boldsymbol{\alpha}, T)$ of order m, we must have $\mathrm{cv}(\boldsymbol{\alpha}, T) \ge 1/m$, where $\mathrm{cv}(\boldsymbol{\alpha}, T)$ is the coefficient of variation of $(\boldsymbol{\alpha}, T)$. That implies that $m \ge 1/\mathrm{cv}(\boldsymbol{\alpha}, T) = (\boldsymbol{\alpha}T^{-1}\mathbf{e})^2/[\boldsymbol{\alpha}T^{-2}\mathbf{e} - (\boldsymbol{\alpha}T^{-1}\mathbf{e})^2]$. Let $N^* = \max\{N_m, \lceil 1/\mathrm{cv}(\boldsymbol{\alpha}, T)\rceil\}$. Apparently, N^* is a lower bound of the triangular order of $(\boldsymbol{\alpha}, T)$.

Finally, we introduce an algorithm for computing a minimal ordered Coxian representation of $(\boldsymbol{\alpha}, T)$. For this purpose, one needs to solve (11) for $n = N^*, N^*+1, \ldots,$ until a solution is found. Theorem 4 ensures that the smallest n for which (11) has a solution is the triangular order of $(\boldsymbol{\alpha}, T)$. We summarize the computational steps as follows.

Minimal Coxian Representation Algorithm. Assume that the Laplace-Stieltjes transform of PH-distribution $(\boldsymbol{\alpha}, T)$ has only real poles.

Step 1. Find the spectrum of T. Use the spectral polynomial algorithm to compute $\{\boldsymbol{\beta}, S(\boldsymbol{\mu}, \mathbf{m})\}$ and N_m (defined in (5)). Remove all eigenvalues whose corresponding elements in $\boldsymbol{\beta}$ are all zero.

Step 2. Use the spectral polynomial algorithm to compute $P(\boldsymbol{\eta})$ and $\{\boldsymbol{\beta}P(\boldsymbol{\eta}), S(\boldsymbol{\eta})\}$ from $\{\boldsymbol{\beta}, S(\boldsymbol{\mu}, \mathbf{m})\}$. If $\boldsymbol{\beta}P(\boldsymbol{\eta})$ is nonnegative, $(\boldsymbol{\beta}P(\boldsymbol{\eta}), S(\boldsymbol{\eta}))$ is a minimal ordered Coxian representation of order N_m for $(\boldsymbol{\alpha}, T)$. Otherwise, compute N^* and let $n = N^*$.

Step 3. Calculate $F^{(j)}(0) = -\boldsymbol{\alpha}T^j\mathbf{e}$, $1 \le j \le n+1$. Solve (11).

Step 4. If there is no solution to (11), set $n =: n+1$ and go to Step 3. Otherwise, use Lemma 4 to compute $\boldsymbol{\gamma}$ from the optimal solution. Then $(\boldsymbol{\gamma}, S(\mathbf{x}))$ is a minimal Coxian representation of order n for $(\boldsymbol{\alpha}, T)$. If $(\boldsymbol{\gamma}, S(\mathbf{x}))$ is not an ordered Coxian representation (i.e., the elements in $\mathbf{x}$ are not in nonincreasing order), use the spectral polynomial algorithm to compute a minimal ordered Coxian representation from $(\boldsymbol{\gamma}, S(\mathbf{x}))$.

According to Cumani (1982) and He and Zhang (2006b), if T is triangular or symmetric, the search for the triangular order is between N^* and m. In general, by Theorem 4.1 in O'Cinneide (1991), the search process will be terminated after a finite number of iterations. However, there is no known upper bound for the search process. Thus, finding an upper bound of the triangular order is useful for the above algorithm and is an interesting issue for future research.

6. Three Numerical Examples

In this section, we demonstrate the effectiveness of the algorithm developed in §§3, 4, and 5, show some geometric properties associated with ordered Coxian representations, and explain why finding a minimal ordered Coxian representation is not straightforward. Note that all distributions under consideration have a zero mass at $t = 0$.

Example 1. We show that the triangular order of a PH-representation whose Laplace-Stieltjes transform has three poles can be arbitrarily large, as pointed out by O'Cinneide (1991). We consider a PH-generator T of order three given by

$$T = \begin{pmatrix} -7 & 0 & 0.3 \\ 3 & -4 & 0 \\ 0 & 3.5 & -4 \end{pmatrix}. \tag{12}$$

The eigenvalues of T are $\{-\lambda_1 = -6.4933, -\lambda_2 = -5.4056, -\lambda_3 = -3.1012\}$. Let $\{\boldsymbol{\alpha}_1, \boldsymbol{\alpha}_2, \boldsymbol{\alpha}_3\}$ be the corresponding eigenvectors normalized by $\boldsymbol{\alpha}_1\mathbf{e} = \boldsymbol{\alpha}_2\mathbf{e} = \boldsymbol{\alpha}_3\mathbf{e} = 1$. Then $\{(\boldsymbol{\alpha}_i, T, \mathbf{e}),\ 1 \le i \le 3\}$ represent three exponential distributions with parameters $\{\lambda_1, \lambda_2, \lambda_3\}$, respectively. Similarly to Example 3.2 in He and Zhang (2006a) (also see examples in Dehon and Latouche 1982), we construct a PH-invariant polytope $\mathrm{conv}\{\mathbf{q}_1, \mathbf{q}_{12}, \mathbf{q}_{123}\}$ from $\{\boldsymbol{\alpha}_1, \boldsymbol{\alpha}_2, \boldsymbol{\alpha}_3\}$ as follows:

$$\mathbf{q}_1 = \boldsymbol{\alpha}_1;$$
$$\mathbf{q}_{12} = \frac{\lambda_2}{\lambda_2 - \lambda_1}\boldsymbol{\alpha}_1 + \frac{\lambda_1}{\lambda_1 - \lambda_2}\boldsymbol{\alpha}_2;$$
$$\mathbf{q}_{123} = \frac{\lambda_3\lambda_2}{(\lambda_3 - \lambda_1)(\lambda_2 - \lambda_1)}\boldsymbol{\alpha}_1 + \frac{\lambda_1\lambda_3}{(\lambda_1 - \lambda_2)(\lambda_3 - \lambda_2)}\boldsymbol{\alpha}_2 + \frac{\lambda_1\lambda_2}{(\lambda_1 - \lambda_3)(\lambda_2 - \lambda_3)}\boldsymbol{\alpha}_3.$$

Let Q be a 3×3 matrix with $\mathbf{q}_1$ as its first row, $\mathbf{q}_{12}$ as its second row, and $\mathbf{q}_{123}$ as its third row. It can be shown that $QT = S(\lambda)Q$. Since $S(\lambda)$ is a PH-generator, we say that the polytope $\mathrm{conv}\{\mathbf{q}_1, \mathbf{q}_{12}, \mathbf{q}_{123}\}$ is PH-invariant under T.

A polytope is a convex set with a finite number of extreme points. See Rockafellar (1970) for more about convex sets and polytopes. The concept of PH-invariant polytopes was introduced in O'Cinneide (1991), who discusses more about PH-invariant polytopes.

According to Dehon and Latouche (1982), conv $\{\mathbf{q}_1, \mathbf{q}_{12}, \mathbf{q}_{123}\}$ contains all $\boldsymbol{\alpha}$ corresponding to probability distributions $(\boldsymbol{\alpha}, T, \mathbf{e})$ with a bidiagonal *PH*-representation of order three or smaller. Denote by Ω_3 the set of all vectors $\boldsymbol{\alpha}$ with a unit sum corresponding to probability functions that are affine combinations of exponential distributions $\{(\boldsymbol{\alpha}_i, T, \mathbf{e}), 1 \le i \le 3\}$. Denote by $\mathbf{e}_i$ the unit row vector with all elements being zero except the i-th element, which is one, $1 \le i \le 3$. By definition, all probability vectors are in the polytope conv$\{\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3\}$. We have conv$\{\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3\} \subseteq \Omega_3$ and conv$\{\mathbf{q}_1, \mathbf{q}_{12}, \mathbf{q}_{123}\} \subseteq \Omega_3$. In Figure 1, the probability vector polytope conv$\{\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3\}$, the polytope conv$\{\mathbf{q}_1, \mathbf{q}_{12}, \mathbf{q}_{123}\}$, and the convex set Ω_3 are plotted.

For any *PH*-representation $(\boldsymbol{\alpha}, T)$ with T given in (12), if $\boldsymbol{\alpha}$ is in conv$\{\mathbf{q}_1, \mathbf{q}_{12}, \mathbf{q}_{123}\}$, $(\boldsymbol{\alpha}, T)$ has an ordered Coxian representation of order three or smaller. Otherwise, $(\boldsymbol{\alpha}, T, \mathbf{e})$ does not have an ordered Coxian representation of order three. He and Zhang (2006b) show that *PH*-representation $(\boldsymbol{\alpha}, T)$ of order three always has an ordered Coxian representation of order four or smaller. For instance, the triangular order of *PH*-distribution $(\mathbf{e}_1, T)$ is four, three for $(\mathbf{e}_2, T)$, and three for $(\mathbf{e}_3, T)$. Note that the triangular order of $(\mathbf{e}_1, T)$ is larger than its *PH*-order (which is three). If $\boldsymbol{\alpha}$ is not in conv$\{\mathbf{q}_1, \mathbf{q}_{12}, \mathbf{q}_{123}\}$ nor in conv$\{\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3\}$, then the triangular order of $(\boldsymbol{\alpha}, T, \mathbf{e})$ is at least four and can be arbitrarily large.

Next, we consider the following vectors $\boldsymbol{\alpha}$ in Ω_3 and use the nonlinear program developed in §§3, 4, and 5 to find the triangular order and a minimal ordered Coxian representation of $(\boldsymbol{\alpha}, T, \mathbf{e})$. If $\boldsymbol{\alpha} = \boldsymbol{\alpha}_i$, $i = 1, 2$, or 3, we have $n^* = 1$. If $\boldsymbol{\alpha} = y\boldsymbol{\alpha}_i + (1 - y)\boldsymbol{\alpha}_3$, for $0 < y < 1$ and $i = 1$ or 2, we have $n^* = 2$. If $\boldsymbol{\alpha} = y\boldsymbol{\alpha}_1 + z\boldsymbol{\alpha}_3 + (1 - y - z)\boldsymbol{\alpha}_3$ for $0 < y, z, y + z < 1$, we have $n^* = 3$. Now, we choose $\boldsymbol{\alpha}$ outside of conv$\{\mathbf{q}_1, \mathbf{q}_{12}, \mathbf{q}_{123}\}$ as follows: $\boldsymbol{\alpha} = 0.5(1, 0, 0) + 0.5\mathbf{q}_{123} + \delta(1, -1, 0)$, where δ is to be determined. For the following $\boldsymbol{\alpha}$, we use NLP (11) to find the triangular order n^* of $(\boldsymbol{\alpha}, T, \mathbf{e})$:

$$
\begin{aligned}
&\delta = 0: && \boldsymbol{\alpha}_4 = (0.4714, -0.1458, 0.6744), && n^* = 4;\\
&\delta = 0.7: && \boldsymbol{\alpha}_5 = (1.1714, -0.8458, 0.6744), && n^* = 5;\\
&\delta = 0.8: && \boldsymbol{\alpha}_6 = (1.2714, -0.9458, 0.6744), && n^* = 6;\\
&\delta = 0.9: && \boldsymbol{\alpha}_7 = (1.3714, -1.0458, 0.6744), && n^* = 8;\\
&\delta = 1.0: && \boldsymbol{\alpha}_8 = (1.4714, -1.1458, 0.6744), && n^* = 12;\\
&\delta = 1.1: && \boldsymbol{\alpha}_9 = (1.5714, -1.2458, 0.6744), && n^* = 31;\\
&\delta = 1.2: && \boldsymbol{\alpha}_{10} = (1.6714, -1.3458, 0.6744), && \\
& && && n^* = \text{nonexistent}.
\end{aligned}
$$

The vectors $\{\boldsymbol{\alpha}_i, 4 \le i \le 10\}$ are plotted in Figure 1 as well. It is clear that when $\boldsymbol{\alpha}$ approaches the boundary of Ω_3, the triangular order of $(\boldsymbol{\alpha}, T, \mathbf{e})$ can be very large. When $\delta = 1.2$, the corresponding $\boldsymbol{\alpha}_{10}$ is outside of Ω_3. Thus $(\boldsymbol{\alpha}_{10}, T, \mathbf{e})$ does not represent a probability distribution. Numerical experimentations demonstrate that the minimal Coxian representation algorithm is efficient if $\boldsymbol{\alpha}$ is not close to the boundary of Ω_3.

For Example 1, the original *PH*-generator T is not an ordered Coxian generator. In the next two examples, we consider cases for which T is an ordered Coxian generator. As will be shown, finding a minimal Coxian representation can be equally complicated even if the original representation is an ordered Coxian representation.

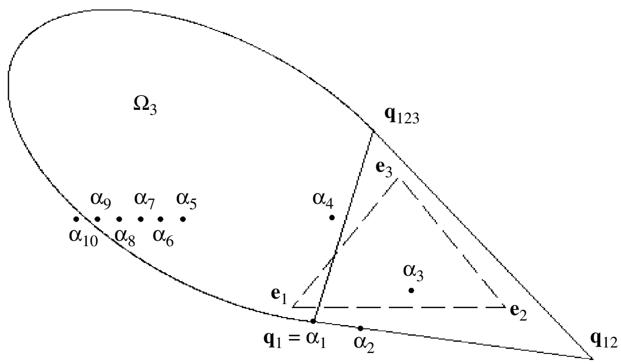

Figure 1 Ω_3 **and *PH*-Invariant Polytopes for Example 1**

Example 2. We consider a *PH*-generator $S(\boldsymbol{\lambda})$ of order four given by

$$
S(\boldsymbol{\lambda}) = \begin{pmatrix} -10 & 0 & 0 & 0 \\ 5 & -5 & 0 & 0 \\ 0 & 1.5 & -1.5 & 0 \\ 0 & 0 & 1 & -1 \end{pmatrix}.
$$

Eigenvalues of $S(\boldsymbol{\lambda})$ are $-\lambda_1 = -10$, $-\lambda_2 = -5$, $-\lambda_3 = -1.5$, $-\lambda_4 = -1$ and the corresponding eigenvectors are $\boldsymbol{\alpha}_1 = (1, 0, 0, 0)$, $\boldsymbol{\alpha}_2 = (0.5, 0.5, 0, 0)$, $\boldsymbol{\alpha}_3 = (0.15, 0.255, 0.595, 0)$, $\boldsymbol{\alpha}_4 = (0.1, 0.18, 0.48, 0.24)$, respectively. All eigenvectors are normalized to have a unit sum. According to He and Zhang (2006b), conv$\{\boldsymbol{\alpha}_1, \boldsymbol{\alpha}_2, \boldsymbol{\alpha}_3, \boldsymbol{\alpha}_4\}$ is a *PH*-invariant polytope under $S(\boldsymbol{\lambda})$. Again, by using the method developed in Dehon and Latouche (1982), the *PH*-invariant polytope conv$\{\boldsymbol{\alpha}_1, \boldsymbol{\alpha}_2, \boldsymbol{\alpha}_3, \boldsymbol{\alpha}_4\}$ can be expanded to a *PH*-invariant polytope conv$\{\mathbf{q}_1, \mathbf{q}_{12}, \mathbf{q}_{123}, \mathbf{q}_{1234}\}$ as follows:

$$
\mathbf{q}_1 = \boldsymbol{\alpha}_1; \qquad \mathbf{q}_{12} = \frac{\lambda_2}{\lambda_2 - \lambda_1}\boldsymbol{\alpha}_1 + \frac{\lambda_1}{\lambda_1 - \lambda_2}\boldsymbol{\alpha}_2;
$$

$$
\mathbf{q}_{123} = \frac{\lambda_3\lambda_2}{(\lambda_3 - \lambda_1)(\lambda_2 - \lambda_1)}\boldsymbol{\alpha}_1 + \frac{\lambda_1\lambda_3}{(\lambda_1 - \lambda_2)(\lambda_3 - \lambda_2)}\boldsymbol{\alpha}_2 + \frac{\lambda_1\lambda_2}{(\lambda_1 - \lambda_3)(\lambda_2 - \lambda_3)}\boldsymbol{\alpha}_3;
$$

$$\begin{aligned}\mathbf{q}_{1234} = {} & \frac{\lambda_4\lambda_3\lambda_2}{(\lambda_4-\lambda_1)(\lambda_3-\lambda_1)(\lambda_2-\lambda_1)}\boldsymbol{\alpha}_1 \\ & + \frac{\lambda_4\lambda_3\lambda_1}{(\lambda_4-\lambda_2)(\lambda_3-\lambda_2)(\lambda_1-\lambda_2)}\boldsymbol{\alpha}_2 \\ & + \frac{\lambda_4\lambda_2\lambda_1}{(\lambda_4-\lambda_3)(\lambda_2-\lambda_3)(\lambda_1-\lambda_3)}\boldsymbol{\alpha}_3 \\ & + \frac{\lambda_3\lambda_2\lambda_1}{(\lambda_3-\lambda_4)(\lambda_2-\lambda_4)(\lambda_1-\lambda_4)}\boldsymbol{\alpha}_4.\end{aligned}$$

For this example, because $S(\boldsymbol{\lambda})$ is itself an ordered Coxian generator, it can be verified that $\mathbf{q}_1 = \mathbf{e}_1, \mathbf{q}_{12} = \mathbf{e}_2, \mathbf{q}_{123} = \mathbf{e}_3$, and $\mathbf{q}_{1234} = \mathbf{e}_4$. According to Dehon and Latouche (1982), the simplex $\Gamma_{1,2,3,4} = \text{conv}\{\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3, \mathbf{e}_4\}$ consists of all probability vectors $\boldsymbol{\alpha}$ such that $(\boldsymbol{\alpha}, S(\boldsymbol{\lambda}), \mathbf{e})$ is a representation of an ordered Coxian distribution of order four or smaller. Let $\Omega_{1,2,3,4}$ be the set of all vectors $\boldsymbol{\alpha}$ with a unit sum such that $(\boldsymbol{\alpha}, S(\boldsymbol{\lambda}), \mathbf{e})$ is an *ME*-representation of a distribution.

We consider four sub-affine sets generated by $\{\boldsymbol{\alpha}_1, \boldsymbol{\alpha}_2, \boldsymbol{\alpha}_3\}$, $\{\boldsymbol{\alpha}_1, \boldsymbol{\alpha}_2, \boldsymbol{\alpha}_4\}$, $\{\boldsymbol{\alpha}_1, \boldsymbol{\alpha}_3, \boldsymbol{\alpha}_4\}$, and $\{\boldsymbol{\alpha}_2, \boldsymbol{\alpha}_3, \boldsymbol{\alpha}_4\}$, respectively. Denote the sub-affine sets as aff$\{\boldsymbol{\alpha}_1, \boldsymbol{\alpha}_2, \boldsymbol{\alpha}_3\}$, aff$\{\boldsymbol{\alpha}_1, \boldsymbol{\alpha}_2, \boldsymbol{\alpha}_4\}$, aff$\{\boldsymbol{\alpha}_1, \boldsymbol{\alpha}_3, \boldsymbol{\alpha}_4\}$, and aff$\{\boldsymbol{\alpha}_2, \boldsymbol{\alpha}_3, \boldsymbol{\alpha}_4\}$. By Dehon and Latouche (1982), the set of all probability distributions in each sub-affine set can be identified, which are denoted by $\Omega_{1,2,3}$, $\Omega_{1,3,4}$, $\Omega_{1,2,4}$, and $\Omega_{2,3,4}$, respectively. These subsets of probability distributions and polytopes are plotted in Figure 2. Similarly to the expansion from conv$\{\boldsymbol{\alpha}_1, \boldsymbol{\alpha}_2, \boldsymbol{\alpha}_3, \boldsymbol{\alpha}_4\}$ to conv$\{\mathbf{q}_1, \mathbf{q}_{12}, \mathbf{q}_{123}, \mathbf{q}_{1234}\}$, we can expand the *PH*-invariant polytopes conv$\{\boldsymbol{\alpha}_1, \boldsymbol{\alpha}_2, \boldsymbol{\alpha}_3\}$, conv$\{\boldsymbol{\alpha}_1, \boldsymbol{\alpha}_2, \boldsymbol{\alpha}_4\}$, conv$\{\boldsymbol{\alpha}_1, \boldsymbol{\alpha}_3, \boldsymbol{\alpha}_4\}$, and conv$\{\boldsymbol{\alpha}_2, \boldsymbol{\alpha}_3, \boldsymbol{\alpha}_4\}$ to $\Gamma_{1,2,3} = \text{conv}\{\mathbf{q}_1, \mathbf{q}_{12}, \mathbf{q}_{123}\}$, $\Gamma_{1,2,4} = \text{conv}\{\mathbf{q}_1, \mathbf{q}_{12}, \mathbf{q}_{124}\}$, $\Gamma_{1,3,4} = \text{conv}\{\mathbf{q}_1, \mathbf{q}_{13}, \mathbf{q}_{134}\}$, and $\Gamma_{2,3,4} = \text{conv}\{\mathbf{q}_2, \mathbf{q}_{23}, \mathbf{q}_{234}\}$, respectively. All the ordered Coxian representations of order three with eigenvalues chosen from $\{-\lambda_1, -\lambda_2, -\lambda_3, -\lambda_4\}$ are in these *PH*-invariant polytopes.

Intuitively, for a probability vector $\boldsymbol{\alpha}$ in some subspace, it is possible that $(\boldsymbol{\alpha}, S(\boldsymbol{\lambda}), \mathbf{e})$ has an ordered Coxian representation of lower order. In fact, if a probability vector $\boldsymbol{\alpha}$ is in aff$\{\boldsymbol{\alpha}_1, \boldsymbol{\alpha}_2, \boldsymbol{\alpha}_3\}$, aff$\{\boldsymbol{\alpha}_1, \boldsymbol{\alpha}_2, \boldsymbol{\alpha}_4\}$, or aff$\{\boldsymbol{\alpha}_1, \boldsymbol{\alpha}_3, \boldsymbol{\alpha}_4\}$, then $(\boldsymbol{\alpha}, S(\boldsymbol{\lambda}), \mathbf{e})$ has an ordered Coxian representation of order three or smaller. The reason is that $\Omega_{1,2,3} \cap \Gamma_{1,2,3,4} = \Gamma_{1,2,3}$, $\Omega_{1,2,4} \cap \Gamma_{1,2,3,4} = \Gamma_{1,2,4}$, and $\Omega_{1,3,4} \cap \Gamma_{1,2,3,4} = \Gamma_{1,3,4}$, as shown in Figure 2.

However, if a probability vector $\boldsymbol{\alpha}$ is in aff$\{\boldsymbol{\alpha}_2, \boldsymbol{\alpha}_3, \boldsymbol{\alpha}_4\}$, it is possible that $F(t)$ of $(\boldsymbol{\alpha}, S(\boldsymbol{\lambda}), \mathbf{e})$ has an ordered Coxian representation of order four whose Laplace-Stieltjes transform has poles $\{-\lambda_2, -\lambda_3, -\lambda_4\}$ and the triangular order of $F(t)$ is four, i.e., $F(t) \in \Gamma_{1,2,3,4}$ but $F(t) \notin \Gamma_{2,3,4}$. The reason is that $\Gamma_{2,3,4} \subset \Omega_{2,3,4} \cap \Gamma_{1,2,3,4}$ and $\Gamma_{2,3,4} \neq \Omega_{2,3,4} \cap \Gamma_{1,2,3,4}$, as shown in Figure 2. In fact, $\Omega_{2,3,4} \cap \Gamma_{1,2,3,4} = \text{conv}\{\mathbf{q}_2, \mathbf{q}_{23}, \mathbf{q}_{234}, \mathbf{q}_0\}$, where $\mathbf{q}_0$ is in the set conv$\{\mathbf{q}_1, \mathbf{q}_{1234}\}$ (the point "$*$" in Figure 2).

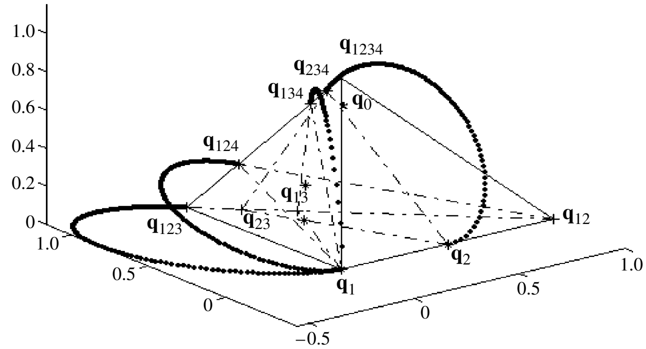

Figure 2 $\Omega_{1,2,3}, \Omega_{1,2,4}, \Omega_{1,3,4}, \Omega_{2,3,4}$, **and *PH*-Invariant Polytopes for Example 2**

Any ordered Coxian distribution in the interior of the set conv$\{\mathbf{q}_2, \mathbf{q}_{234}, \mathbf{q}_0\}$ has a triangular order four.

In summary, for any Coxian distribution in $\Gamma_{1,2,3}$, $\Gamma_{1,3,4}$, and $\Gamma_{1,2,4}$, the number of the poles of its Laplace-Stieltjes transform equals its triangular order. For a Coxian distribution in $\Gamma_{2,3,4}$, it is possible that the number of the poles of its Laplace-Stieltjes transform is smaller than its triangular order.

Example 3. We consider a Coxian distribution $(\boldsymbol{\alpha}, S(\boldsymbol{\lambda}))$ with $m = 5$, $\boldsymbol{\alpha} = (0.0787, 0.0778, 0.0184, 0.1174, 0.7076)$, and

$$S(\boldsymbol{\lambda}) = \begin{pmatrix} -20 & 0 & 0 & 0 & 0 \\ 10 & -10 & 0 & 0 & 0 \\ 0 & 7.5 & -7.5 & 0 & 0 \\ 0 & 0 & 5 & -5 & 0 \\ 0 & 0 & 0 & 1.5 & -1.5 \end{pmatrix}, \tag{13}$$

where $\boldsymbol{\lambda} = (20, 10, 7.5, 5, 1.5)$. It can be shown that the *PH*-distribution $(\boldsymbol{\alpha}, S(\boldsymbol{\lambda}))$ has an *ME*-representation $(\boldsymbol{\beta}, S((7.5, 5, 1.5)), \mathbf{e})$ of order three with poles $\{-7.5, -5, -1.5\}$ and $\boldsymbol{\beta} = (0.21, -0.11, 0.9)$. By Lemma 2, the triangular order of $(\boldsymbol{\alpha}, S(\boldsymbol{\lambda}))$ must be at least four.

To see if $(\boldsymbol{\alpha}, S(\boldsymbol{\lambda}))$ has an ordered Coxian representation of order four, we consider ordered Coxian generators with eigenvalues $\{-10, -7.5, -5, -1.5\}$ and $\{-20, -7.5, -5, -1.5\}$, where -10 and -20 are the eigenvalues of the original *PH*-generator given in (13). Using the spectral polynomial algorithm, it can be shown that $(\boldsymbol{\alpha}, S(\boldsymbol{\lambda}))$ is equivalent to *ME*-representations $(\boldsymbol{\beta}_1, S((10, 7.5, 5, 1.5)), \mathbf{e})$ and $(\boldsymbol{\beta}_2, S((20, 7.5, 5, 1.5)), \mathbf{e})$, where $\boldsymbol{\beta}_1 = (0.1575, -0.0025, 0.0800, 0.7650)$ and $\boldsymbol{\beta}_2 = (0.0787, 0.1038, -0.0150, 0.8235)$. Unfortunately, none of them is a *PH*-representation. Therefore, we need to try a value other than -10 and -20. By considering -15, i.e., $\boldsymbol{\lambda} = (15, 7.5, 5, 1.5)$, we find $\boldsymbol{\beta}_3 = (0.1050,$

0.0683, 0.0167, 0.8100), which implies that $(\boldsymbol{\beta}_3, S((15, 7.5, 5, 1.5)))$ is an alternative *PH*-representation of $(\boldsymbol{\alpha}, S(\boldsymbol{\lambda}))$. Since $(\boldsymbol{\beta}_3, S((15, 7.5, 5, 1.5)))$ is an ordered Coxian representation, the triangular order of $(\boldsymbol{\alpha}, S(\boldsymbol{\lambda}))$ is four. Furthermore, using Lemma 7, we find that we can replace -15 by any number between -10.88 and -18.00 to generate an alternative ordered Coxian representation of $(\boldsymbol{\alpha}, S(\boldsymbol{\lambda}))$.

An interesting observation is that the eigenvalue -15 is not an eigenvalue of the original generator $S(\boldsymbol{\lambda})$. Thus, to find a minimal ordered Coxian representation for $(\boldsymbol{\alpha}, S(\boldsymbol{\lambda}))$, we must find the eigenvalue -15 (or a number between -10.88 and -18.00), which is not straightforward.

7. Discussion on Future Research

An interesting issue for future research is to explore the structure of NLP (11) to develop more efficient ways to solve it. Also, we would like to find an upper bound on the triangular order, which makes it possible to combine all the nonlinear programs into a single nonlinear program to find a minimal Coxian representation.

Another area of future research is to extend the spectral polynomial method to parameter estimation and fitting of *PH*-distributions, which are of great practical significance. Possible results in this area may complement to those in Asmussen et al. (1996), Bobbio et al. (2002, 2004), Feldmann and Whitt (1998), and Johnson and Taaffe (1989; 1990a, b).

Acknowledgments

The authors thank an anonymous referee for detailed and constructive comments and suggestions on an earlier version of this paper. This research project was financially supported by the National Sciences and Engineering Research Council (NSERC) of Canada and the Chinese Academy of Sciences (People's Republic of China).

References

Aldous, D., L. Shepp. 1987. The least variable phase type distribution is Erlang. *Stochastic Models* **3** 467–473.

Asmussen, S. 2000. *Ruin Probabilities*. World Scientific Publishing, Singapore.

Asmussen, S. 2003. *Applied Probability and Queues*, 2nd ed. Springer-Verlag, New York.

Asmussen, S., M. Bladt. 1996. Renewal theory and queueing algorithms for matrix-exponential distributions. A. S. Alfa, S. Chakravarthy, eds. *Proc. First Internat. Conf. Matrix Analytic Methods in Stochastic Models*, Marcel Dekker, New York, 313–341.

Asmussen, S., O. Nerman, M. Olsson. 1996. Fitting phase-type distributions via the EM algorithm. *Scand. J. Statist.* **23** 419–441.

Bobbio, A., A. Horváth, M. Telek. 2004. The scale factor: A new degree of freedom in phase type approximation. *Performance Eval.* **56** 121–144.

Bobbio, A., A. Horváth, M. Scarpa, M. Telek. 2002. A cyclic discrete phase type distribution: Properties and a parameter estimation algorithm. *Performance Eval.* **54** 1–32.

Commault, C., S. Mocanu. 2003. Phase-type distributions and representations: Some results and open problems for system theory. *Internat. J. Control* **76** 566–580.

Cox, D. R. 1955a. On the use of complex probabilities in the theory of stochastic processes. *Proc. Camb. Phil. Soc.* **51** 313–319.

Cox, D. R. 1955b. The analysis of non-Markovian stochastic processes by the inclusion of supplementary variables. *Proc. Camb. Phil. Soc.* **51** 433–441.

Cumani, A. 1982. On the canonical representation of Markov processes modeling failure time distributions. *Microelectronics Reliability* **22** 583–602.

Dehon, M., G. Latouche. 1982. A geometric interpretation of the relations between the exponential and the generalized Erlang distributions. *Adv. Appl. Probab.* **14** 885–897.

Feldmann, A., W. Whitt. 1998. Fitting mixtures of exponentials to long-tail distributions to analyze network performance models. *Performance Eval.* **32** 245–279.

Haddad, S., P. Moreaux, G. Chiola. 1998. Cox and phase-type distributions in stochastic Petri nets—Efficient derivation of solutions. *RAIRO Recherche Operationnelle-Oper. Res.* **32** 289–323.

Harris, C. M., W. G. Marchal, R. F. Botta. 1992. A note on generalized hyperexponential distributions. *Stochastic Models* **8** 179–191.

He, Q. M., H. Q. Zhang. 2005. A note on unicyclic representations of phase type distributions. *Stochastic Models* **21** 465–483.

He, Q. M., H. Q. Zhang. 2006a. *PH*-invariant polytopes and Coxian representations of phase type distributions. *Stochastic Models* **22** 383–409.

He, Q. M., H. Q. Zhang. 2006b. Spectral polynomial algorithms for computing bi-diagonal representations for phase-type distributions and matrix-exponential distributions. *Stochastic Models* **22** 289–317.

Heijden, M. C. 1988. On the three moment approximation of a general distribution by a Coxian distribution. *Probab. Engrg. Inform. Sci.* **2** 257–261.

Johnson, M. A. 1993. Selecting parameters of phase type distributions: Combining nonlinear programming, heuristics and Erlang distributions. *ORSA J. Comput.* **5** 69–83.

Johnson, M. A., M. R. Taaffe. 1989. Matching moments to phase distributions: Mixtures of Erlang distributions of common order. *Stochastic Models* **5** 711–743.

Johnson, M. A., M. R. Taaffe. 1990a. Matching moments to phase distributions: Density function shapes. *Stochastic Models* **6** 283–306.

Johnson, M. A., M. R. Taaffe. 1990b. Matching moments to phase distributions: Nonlinear programming approaches. *Stochastic Models* **6** 259–281.

Lancaster, P., M. Tismenetsky. 1985. *The Theory of Matrices*. Academic Press, New York.

Latouche, G., V. Ramaswami. 1999. *Introduction to Matrix Analytic Methods in Stochastic Modelling*. ASA & SIAM, Philadelphia.

Lipsky, L. 1992. *Queueing Theory: A Linear Algebraic Approach*. McMillan and Company, New York.

Minc, H. 1988. *Nonnegative Matrix*. John Wiley & Sons, New York.

Mocanu, S., C. Commault. 1999. Sparse representations of phase type distributions. *Stochastic Models* **15** 759–778.

Neuts, M. F. 1975. Probability distributions of phase type. *Liber Amicorum Prof. Emeritus H. Florin*, Department of Mathematics, University of Louvain, Leuven, Belgium, 173–206.

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

Neuts, M. F. 1989. *Structured Stochastic Matrices of M/G/ 1 Type and their Applications*. Marcel Dekker, New York.

O'Cinneide, C. A. 1989. On non-uniqueness of representations of phase-type distributions. *Stochastic Models* **5** 247–259.

O'Cinneide, C. A. 1990. Characterization of phase-type distributions. *Stochastic Models* **6** 1–57.

O'Cinneide, C. A. 1991. Phase-type distributions and invariant polytopes. *Adv. Appl. Probab.* **23** 515–535.

O'Cinneide, C. A. 1993. Triangular order of triangular phase-type distributions. *Stochastic Models* **9** 507–529.

O'Cinneide, C. A. 1999. Phase-type distributions: Open problems and a few properties. *Stochastic Models* **15** 731–757.

Osogami, T., M. Harchol-Balter. 2003a. A closed-form solution for mapping general distributions to minimal *PH* distributions. *13th Internat. Conf. Model. Techniques and Tools for Comput. Performance Eval.* (*Tools 03*), 200–217, http://www.cs.cmu.edu/~osogami/publication.html.

Osogami, T., M. Harchol-Balter. 2003b. Necessary and sufficient conditions for representing general distributions by Coxians. *13th Internat. Conf. Model. Techniques and Tools for Comput. Performance Evaluation* (*Tools 03*), 182–199, http://www.cs.cmu.edu/~osogami/publication.html.

Rockafellar, R. T. 1970. *Convex Analysis*. Princeton University Press, Princeton, NJ.

Sasaki, Y., H. Imai, M. Tsunoyama, I. Ishii. 2004. Approximation of probability distribution functions by Coxian distribution to evaluate multimedia systems. *Systems Comput. Japan* **35** 16–24.

Seneta, E. 1973. *Non-Negative Matrices: An Introduction to Theory and Applications*. John Wiley & Sons, New York.

Takahashi, Y. 1981. Asymptotic exponentiality of the tail of the waiting time distribution in a *PH/PH/c* queue. *Adv. Appl. Probab.* **13** 619–630.