Stochastic Models, 22:383–409, 2006

ISSN: 1532-6349 print/1532-4214 online
DOI: 10.1080/15326340600820349

PH-INVARIANT POLYTOPES AND COXIAN REPRESENTATIONS OF PHASE TYPE DISTRIBUTIONS

Qi-Ming He □ *Department of Industrial Engineering, Dalhousie University, Halifax, N.S., Canada*

Hanqin Zhang □ *Academy of Mathematics and Systems Science, Chinese Academy of Sciences, Beijing, P.R. China*

□ *In this paper, we study the relationship between phase type (PH) distributions, PH-invariant polytopes, and Coxian representations of PH-distributions. Explicit links are established between vectors in PH-invariant polytopes, PH-representations, and probability measures. A method is developed for the construction of PH-invariant polytopes associated with ordered Coxian representations. The relationship between PH-invariant polytopes and spectral polynomial algorithms is explored as well. Furthermore, a generic algorithm for computing bi-diagonal PH-representations of PH-distributions is developed. Numerical examples are provided to gain insight into the problems of interest.*

Keywords Aggregated Markov chain; Coxian distribution; Jordan canonical form; Matrix analytic methods; *PH*-distribution; Polytope.

Mathematics Subject Classification Primary 60J27; Secondary 15A21.

1. INTRODUCTION

The *PH*-distribution was introduced in Neuts[25] nearly three decades ago. Since then, *PH*-distributions have been used widely in stochastic modeling of manufacturing, service, and telecommunication systems (Alfa and Chakravarthy[1], Asmussen[2], Asmussen et al.[4], Chakravarthy and Alfa[7], Latouche and Ramaswami[21], Latouche and Taylor[22,23], and Neuts[26]). One of the main advantages of *PH*-distributions is the use of a phase variable to keep track of the status of an underlying process. Consequently, the stochastic system of interest becomes analytically and numerically tractable. However, this approach often has the dimensionality

Received August 2004; Accepted October 2005
Address correspondence to Qi-Ming He, Department of Industrial Engineering, Dalhousie University, Halifax, B3J 2X4 N.S., Canada; E-mail: Qi-Ming.he@dal.ca

problem, i.e., the space and time complexity of algorithms developed for the related stochastic models is high. Thus, the reduction of the orders of *PH*-distributions, which is related to the minimal representation problem of *PH*-distributions, is an important issue.

It is well known that the *PH*-representation of a *PH*-distribution is not unique (Neuts[25,26] and O'Cinneide[27]). Thus, finding simpler and smaller *PH*-representations for *PH*-distributions is an interesting and important issue. Cumani[14] proved that any *PH*-representation with a triangular *PH*-generator has an ordered Coxian representation. That result showed for the first time that *PH*-representations can be drastically simplified. Dehon and Latouche[15] showed how to construct ordered Coxian distributions from exponential distributions. Their work indicated a close relationship between *PH*-invariant polytopes and bi-diagonal *PH*-representations. In O'Cinneide[28–30], *PH*-invariant polytopes were introduced and were used in the study of a number of problems related to *PH*-distributions. This work showed that *PH*-invariant polytopes can play an important role in the study of the minimal and sparse representation problem of *PH*-distributions. In O'Cinneide[30], it was shown that any *PH*-distribution with only real poles has a bi-diagonal *PH*-representation. In a series of papers by Commault et al.[8–12,24], several types of sparse *PH*-representations for *PH*-distributions were introduced. This work made significant progress on the simpler and smaller *PH*-representation problem, but no easy-to-use algorithm was developed for computing simpler and smaller *PH*-representations. In Commault and Mocanu[12], He and Zhang[17], and O'Cinneide[31], more detailed reviews on the study of the minimal representation, sparse representation, and other related issues of *PH*-representations were provided.

In He and Zhang[17], spectral polynomial algorithms were introduced for computing bi-diagonal representations of matrix-exponential distributions (Asmussen and Bladt[3]). The algorithms are also useful for computing Coxian representations for *PH*-representations. In this paper, based on *PH*-invariant polytopes, we take a different approach and develop new algorithms for computing bi-diagonal representations for *PH*-distributions. The results obtained in this paper provide a geometric interpretation to the algorithms developed in Ref.[17] and this paper.

The main contribution of this paper is the establishment of a relationship between *PH*-representations and ordered Coxian representations. The media for that objective are *PH*-invariant polytopes. For a *PH*-generator T, we show explicitly how to construct *PH*-invariant polytopes in which every vector corresponds to an ordered Coxian distribution, which makes it easier to give (counter) examples for the relationship between *PH*-distributions and other subsets of probability distributions. We generalize the results obtained in Dehon and Latouche[15] from the case with distinct real eigenvalues to the case with real Jordan blocks. We generalize the result in Cumani[14] in the sense that a *PH*-representation with a general *PH*-generator may have a bi-diagonal *PH*-representation as well. Compared to previous work, we focus more on *PH*-generators, rather than on specific *PH*-distributions.

This paper is also related to the identifiability problem of aggregated Markov processes (Blackwell and Koopmans[5]; Ito et al.[19]; and Ryden[33]). Intuitively, the identifiability problem is to find out when the function processes of two Markov processes are equivalent. It was shown in Ryden[33] that a *PH*-distribution can be considered as a function of a Markov process. In Ito et al.[19] and Ryden[33], necessary and sufficient conditions were obtained for two *PH*-representations to be equivalent. Since we only consider the equivalence of *PH*-representations and Coxian representations, the conditions given in this paper (as well as in He and Zhang[17]) are simpler. Furthermore, we develop algorithms for computing some of the conditions given in Ito et al.[19] and Ryden[33] and provide geometric interpretations for them.

We note that all results obtained in this paper are valid for discrete time *PH*-distributions.

The rest of the paper is organized as follows. In Section 2, some notation used in this paper is introduced. Section 3 presents three simple examples of *PH*-invariant polytopes and simple *PH*-representations associated with them. In Section 4, a method is developed for constructing an expanded *PH*-invariant polytope from some existing *PH*-invariant polytopes. Based on the results in Section 4, an order reduction algorithm is proposed in Section 5. In Section 6, the relationship between *PH*-invariant polytopes and spectral polynomial algorithms are investigated. Section 7 summarizes the results obtained in this paper.

2. PRELIMINARIES

This section introduces notation and definitions to be used throughout this paper. We include this section for easy reading and completeness. An $m \times m$ matrix T with negative diagonal elements, nonnegative off-diagonal elements, and nonpositive row sums (at least one negative row sum) is called a *subgenerator* in the general literature of Markov process. We shall call a subgenerator T a *PH-generator*. Define a continuous time Markov chain with $m+1$ states and an infinitesimal generator:

$$\begin{pmatrix} T & -T\mathbf{e} \\ 0 & 0 \end{pmatrix}, \tag{2.1}$$

where the $(m+1)$st state is an absorption state and $\mathbf{e}$ is the column vector with all elements being one. We assume that states $\{1, 2, \ldots, m\}$ are transient. Assume that $\boldsymbol{\alpha}$ is a nonnegative vector of size m with the sum of its elements less than or equal to 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 2-tuple $(\boldsymbol{\alpha}, T)$ a *PH-representation* of that *PH*-distribution. The integer m is the order of the *PH*-representation $(\boldsymbol{\alpha}, T)$. The *PH-order* of a *PH*-distribution is defined as the order of its *PH*-representation(s) with the minimal number of states. We refer to Chapter 2 in Neuts[26] for basic properties of *PH*-distributions. The probability distribution function of the *PH*-distribution is given as $1-\boldsymbol{\alpha}\exp\{Tt\}\mathbf{e}$ for $t \geq 0$, and its density function is given as $-\boldsymbol{\alpha}\exp\{Tt\}T\mathbf{e}$ for $t \geq 0$. Without loss of generality, we shall assume that $\boldsymbol{\alpha}$ is a vector with a unit sum throughout this paper. It is possible that $1-\boldsymbol{\alpha}\exp\{Tt\}\mathbf{e}$ is a probability distribution function when the vector $\boldsymbol{\alpha}$ is not nonnegative. For that case, $1-\boldsymbol{\alpha}\exp\{Tt\}\mathbf{e}$ is called a *matrix-exponential distribution.* The 3-tuple $(\boldsymbol{\alpha}, T, \mathbf{e})$ is a *matrix-exponential representation* of that distribution. We refer to Asmussen and Bladt[3] for more details about matrix-exponential distributions.

For a given $m \times m$ matrix T, denote by $\{-\lambda_i, 1 \leq i \leq m\}$ the spectrum of T (i.e., all the roots of the characteristic polynomial of T). It is well known that $\{-\lambda_i, 1 \leq i \leq m\}$ includes all eigenvalues of T. We refer to Lancaster and Tismenetsky[20] for the theory of matrices.

Throughout this paper, we shall work with the usual vector space $\Re^m$ of real m-tuples, where $\Re$ denotes the set of all real numbers. Given a set of N vectors $\{\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_N\}$ of size m, the *affine set* $\text{aff}\{\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_N\}$ is defined as $\{\mathbf{x} = \lambda_1\mathbf{x}_1 + \lambda_2\mathbf{x}_2 + \cdots + \lambda_N\mathbf{x}_N : \lambda_1 + \lambda_2 + \cdots + \lambda_N = 1 \text{ and } \lambda_i \in \Re, 1 \leq i \leq N\}$. We call $\mathbf{x}$ an affine combination of $\{\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_N\}$ if $\mathbf{x} \in \mathit{aff}(\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_N\}$. Denote by $\mathbf{e}_k$ the row vector of size m with the kth element being one and all others zero, $1 \leq k \leq m$. The affine set $\text{aff}\{\mathbf{e}_k, 1 \leq k \leq m\}$ is of particular importance to us, since it consists of all vectors with elements summing to one and it contains all probability vectors. We shall work with $\text{aff}(\mathbf{e}_1, \mathbf{e}_2, \ldots, \mathbf{e}_m)$ throughout this paper. The *convex set* $\text{conv}\{\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_N\}$ is defined as $\{\mathbf{x} = \lambda_1\mathbf{x}_1 + \lambda_2\mathbf{x}_2 + \cdots + \lambda_N\mathbf{x}_N : \lambda_1 + \lambda_2 + \cdots + \lambda_N = 1 \text{ and } \lambda_i \geq 0, 1 \leq i \leq N\}$. We call $\mathbf{x}$ a convex combination of $\{\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_N\}$ if $\mathbf{x} \in \text{conv}\{\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_N\}$. The convex set $\text{conv}\{\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_N\}$ is also called a *polytope* (if N is finite). The polytope $\text{conv}\{\mathbf{e}_k, 1 \leq k \leq m\}$ will be referred to as the *probability vector polytope.* We refer to Ewald[16] and Rockafellar[32] for more details about vectors, affine sets, convex sets, and polytopes.

The *PH*-generator T will be considered as a linear mapping. A polytope $\text{conv}\{\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_N\}$ is *PH-invariant under T* if

$$\mathbf{x}_i T = \sum_{j=1}^{N} s_{i,j}\mathbf{x}_j, \quad 1 \leq i \leq N, \tag{2.2}$$

and the $N \times N$ matrix $S = (s_{i,j})$ is a *PH*-generator. We call S the *PH*-generator corresponding to the *PH*-invariant polytope conv$\{\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_N\}$. The term *PH*-invariant is called *R*-invariant in O'Cinneide[28]. Obviously, the polytope conv$\{\mathbf{e}_i, 1 \le i \le m\}$ is *PH*-invariant under T. Denote by Q an $N \times m$ matrix with rows $\{\mathbf{x}_i, 1 \le i \le N\}$, i.e., with $\mathbf{x}_1$ as its first row, $\mathbf{x}_2$ the second row, ..., and $\mathbf{x}_N$ the Nth row. Then Equation (2.2) becomes $QT = SQ$. If the matrix Q has unit row sums (i.e., $Q\mathbf{e} = \mathbf{e}$) and if there exists a probability vector $\boldsymbol{\beta}$ such that $\boldsymbol{\alpha} = \boldsymbol{\beta} Q$, then the *PH*-distribution $(\boldsymbol{\alpha}, T)$ has another *PH*-representation $(\boldsymbol{\beta}, S)$ (O'Cinneide[27]). In this paper, we are interested in *PH*-invariant polytopes conv$\{\mathbf{x}_i, 1 \le i \le N\}$ under T such that the corresponding *PH*-generator S has a bi-diagonal form. A *Coxian representation* $(\boldsymbol{\beta}, S(\boldsymbol{\lambda}))$ is a *PH*-representation, where the *PH*-generator $S(\boldsymbol{\lambda})$ is a *Coxian generator* given as

$$S(\boldsymbol{\lambda}) = \begin{pmatrix} -\lambda_1 & 0 & \cdots & \cdots & 0 \\ \lambda_2 & -\lambda_2 & \ddots & \ddots & \vdots \\ 0 & \ddots & \ddots & \ddots & \vdots \\ \vdots & \ddots & \lambda_{N-1} & -\lambda_{N-1} & 0 \\ 0 & \cdots & 0 & \lambda_N & -\lambda_N \end{pmatrix}, \tag{2.3}$$

and $\boldsymbol{\lambda} = \{\lambda_1, \lambda_2, \ldots, \lambda_N\}$ is a set of positive real numbers. An *ordered Coxian representation* $(\boldsymbol{\beta}, S(\boldsymbol{\lambda}))$ is a Coxian representation with $\lambda_1 \ge \lambda_2 \ge \ldots \ge \lambda_N > 0$ (Botta et al.[6]; Commault and Mocanu[12]; and Cox[13]). The objective of this paper is to find Coxian representations of distributions with *PH*-representation $(\boldsymbol{\alpha}, T)$. The idea is to find *PH*-invariant polytopes with corresponding Coxian generators for T. If $\boldsymbol{\alpha}$ is in such a *PH*-invariant polytope, then an equivalent Coxian representation is found for $(\boldsymbol{\alpha}, T)$. In order to cover more probability vectors $\boldsymbol{\alpha}$, we look for the largest possible *PH*-invariant polytope with corresponding Coxian generator. Based on these results, an order reduction method is then introduced for computing smaller *PH*-representations.

A triangular *PH*-representation is a *PH*-representation $(\boldsymbol{\alpha}, T)$ for which T is either upper triangular or lower triangular. The *triangular order* of a *PH*-distribution is the order of its triangular *PH*-representation(s) (if it exists) with the minimal number of states. It is easy to see that a Coxian representation is a triangular *PH*-representation. It has been shown in Cumani[14] (also see Dehon and Latouche[15] and O'Cinneide[27]) that any triangular *PH*-representation has an ordered Coxian representation of the same or a smaller order. Thus, the triangular order of a *PH*-distribution is also the minimal order of all its Coxian representations (if they exist).

For a given *PH*-generator T, we denote by $PH(T)$ the set of all *PH*-distributions with a *PH*-generator T. A *PH*-generator T is *PH-simple* if each distribution in $PH(T)$ has a unique PH-representation in the form $(\boldsymbol{\alpha}, T)$. It was shown in O'Cinneide[27] that T is PH-simple if and only if the vectors $\{T^i\mathbf{e}, 0 \leq i \leq m-1\}$ are independent. For two PH-generators $\{T, S\}$, S is said to *PH-majorize* T if $PH(T) \subseteq PH(S)$. It was shown in Ref.[27] that S PH-majorizes T if and only if there exists a nonnegative matrix P with unit row sums for which $TP = PS$. For later use, we present a simple result for PH-simple PH-generators.

Proposition 2.1. *Assume that T is PH-simple. Then all Jordan blocks of T have different eigenvalues. If $TP = PS(\boldsymbol{\lambda})$, then P is invertible, i.e., T and $S(\boldsymbol{\lambda})$ are similar to each other.*

Proof. First, we show that all Jordan blocks of T have different eigenvalues. If the conclusion does not hold, then there must exist two eigenvectors $\mathbf{u}_1$ and $\mathbf{u}_2$ corresponding to a common eigenvalue of T. It is easy to verify that $\mathbf{u}_1 - \mathbf{u}_2$ is also an eigenvector of that eigenvalue. Since T is PH-simple, by Theorem 2 in O'Cinneide[27], $\mathbf{u}_1\mathbf{e}$ and $\mathbf{u}_2\mathbf{e}$ are not zero. Then we can normalize $\mathbf{u}_1$ and $\mathbf{u}_2$ to $\mathbf{u}_1/\mathbf{u}_1\mathbf{e}$ and $\mathbf{u}_2/\mathbf{u}_2\mathbf{e}$. Then $(\mathbf{u}_1 - \mathbf{u}_2)\mathbf{e} = 0$, which contradicts to Theorem 2 in O'Cinneide[27]. Therefore, all Jordan blocks of T have different eigenvalues.

From the first part, it is clear that T and $S(\boldsymbol{\lambda})$ have the same Jordan canonical form (Lancaster and Tismenetsky[20]). That implies that T and $S(\boldsymbol{\lambda})$ are similar, which leads to the second conclusion. This completes the proof of Proposition 2.1.

3. THREE EXAMPLES OF *PH*-INVARIANT POLYTOPES

Finding simpler or smaller PH-representations for PH-distributions is a problem related to PH-invariant polytopes under PH-generator T. In this section, we have a look at three simple PH-invariant polytopes.

Example 3.1 (Diagonal PH-Representations for PH-Distributions). A *fundamental observation* is that every eigenvector $\boldsymbol{\alpha}$ of T (corresponding to a real eigenvalue $-\lambda$) with a nonzero sum is associated with an exponential distribution. Suppose that $\boldsymbol{\alpha}T = -\lambda\boldsymbol{\alpha}$ and $\boldsymbol{\alpha}\mathbf{e} \neq 0$. We normalize $\boldsymbol{\alpha}$ to have a unit sum, i.e., $\boldsymbol{\alpha} =: \boldsymbol{\alpha}/(\boldsymbol{\alpha}\mathbf{e})$. It is easy to see that the one member set $\{\boldsymbol{\alpha}\}$ is a PH-invariant polytope under T. The corresponding PH-generator $S = (-\lambda)$ is associated with an exponential distribution with parameter λ. In fact, it is easy to show that $\boldsymbol{\alpha}\exp\{Tt\}\mathbf{e} = \exp\{-\lambda t\}$, $t \geq 0$. We shall say that $\boldsymbol{\alpha}$ represents an exponential distribution even though $\boldsymbol{\alpha}$ may not be a probability vector. This point of view on vectors in the affine set $\text{aff}\{\mathbf{e}_i, 1 \leq i \leq m\}$ builds a bridge between the *geometric interpretation and probabilistic significance* of vectors in that affine set.

Furthermore, suppose that $\{\boldsymbol{\alpha}_1, \boldsymbol{\alpha}_2, \ldots, \boldsymbol{\alpha}_n\}$ are eigenvectors of T corresponding to real eigenvalues $\{-\lambda_1, -\lambda_2, \ldots, -\lambda_n\}$, respectively, and $\boldsymbol{\alpha}_i\mathbf{e} \neq 0$, $1 \leq i \leq n$. We normalize $\boldsymbol{\alpha}_i$ to $\boldsymbol{\alpha}_i/(\boldsymbol{\alpha}_i\mathbf{e})$, $1 \leq i \leq n$. It is easy to see that the polytope conv$\{\boldsymbol{\alpha}_i, 1 \leq i \leq n\}$ is *PH*-invariant under T, which corresponds to a diagonal *PH*-generator $D(\boldsymbol{\lambda})$, where $D(\boldsymbol{\lambda})$ is an $n \times n$ matrix with diagonal elements $\{-\lambda_1, -\lambda_2, \ldots, -\lambda_n\}$ and all others zero. It is easy to verify that $QT = D(\boldsymbol{\lambda})Q$, where Q is a matrix with rows $\{\boldsymbol{\alpha}_i, 1 \leq i \leq n\}$. For any probability vector $\boldsymbol{\alpha}$ in conv$\{\boldsymbol{\alpha}_i, 1 \leq i \leq n\}$, i.e., $\boldsymbol{\alpha} = \beta_1\boldsymbol{\alpha}_1 + \cdots + \boldsymbol{\beta}_n\boldsymbol{\alpha}_n$, where $\boldsymbol{\beta} = (\beta_1, \ldots, \beta_n)$ is a probability vector, $(\boldsymbol{\alpha}, T)$ has a diagonal *PH*-representation $(\boldsymbol{\beta}, D(\boldsymbol{\lambda}))$ (O'Cinneide[27]). Since $D(\boldsymbol{\lambda})$ is diagonal, the simplification from $(\boldsymbol{\alpha}, T)$ to $(\boldsymbol{\beta}, D(\boldsymbol{\lambda}))$ is significant.

In Figure 1, an example of the polytope conv$\{\boldsymbol{\alpha}_i, 1 \leq i \leq n\}$ is plotted for $n = 3$ and $\lambda_1 > \lambda_2 > \lambda_3$ (Note that, since $\boldsymbol{\alpha}_i\mathbf{e} = 1$, $1 \leq i \leq 3$, the polytope conv$\{\boldsymbol{\alpha}_i, 1 \leq i \leq 3\}$ is in the two-dimensional affine space aff$\{\mathbf{e}_k, 1 \leq k \leq 3\}$.) In Figure 1, the polytope conv$\{\mathbf{e}_i, 1 \leq i \leq 3\}$ of all probability vectors is plotted. A *PH*-invariant polytope conv$\{\mathbf{q}_{1\ldots i}, 1 \leq i \leq 3\}$ that corresponds to an ordered Coxian generator is plotted as well (see Example 3.2 for the construction of this polytope). Since the intersection of conv$\{\boldsymbol{\alpha}_i, 1 \leq i \leq 3\}$ and conv$\{\mathbf{e}_i, 1 \leq i \leq 3\}$ is not empty, there are some probability vectors $\boldsymbol{\alpha}$ for which $(\boldsymbol{\alpha}, T)$ has an equivalent diagonal *PH*-representation. The reason is that the eigenvector $\boldsymbol{\alpha}_3$ of T corresponding to the eigenvalue with the largest real part can be chosen nonnegative.

However, as shown in Figure 1, the polytope conv$\{\boldsymbol{\alpha}_i, 1 \leq i \leq n\}$ covers only a small portion of the probability vector polytope conv$\{\mathbf{e}_k, 1 \leq k \leq m\}$. Thus, we need to expand the polytope conv$\{\boldsymbol{\alpha}_i, 1 \leq i \leq n\}$ as much as

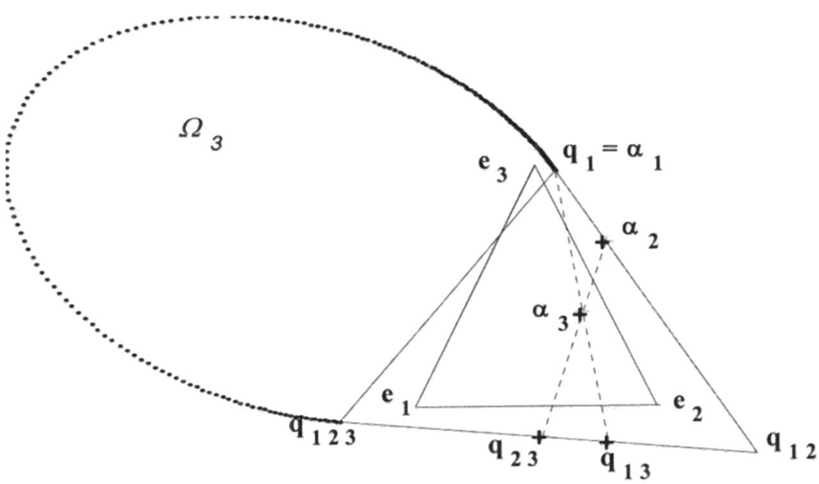

FIGURE 1 *PH*-invariant polytopes for Example 3.2.

possible to cover more probability vectors that are associated with simple *PH*-representations such as the bi-diagonal *PH*-representation. Next, we use two simple examples to demonstrate how such expansions can be done.

Example 3.2 (Dehon-Latouche's *PH*-Invariant Polytope for $m = 3$). Assume that T is a *PH*-generator of order 3. Suppose that $\{\boldsymbol{\alpha}_1, \boldsymbol{\alpha}_2, \boldsymbol{\alpha}_3\}$ are eigenvectors of T corresponding to three distinct real eigenvalues $-\boldsymbol{\lambda} = (-\lambda_1, -\lambda_2, -\lambda_3)$, respectively. We assume that $\lambda_1 > \lambda_2 > \lambda_3$. Suppose that $\{\boldsymbol{\alpha}_1, \boldsymbol{\alpha}_2, \boldsymbol{\alpha}_3\}$ have unit sums. Based on a method introduced in Dehon and Latouche[15], we expand the *PH*-invariant polytope $\text{conv}\{\boldsymbol{\alpha}_1, \boldsymbol{\alpha}_2, \boldsymbol{\alpha}_3\}$ to $\text{conv}\{\mathbf{q}_{1\ldots i}, 1 \le i \le 3\}$ as follows:

$$\begin{aligned}
\mathbf{q}_1 &= \boldsymbol{\alpha}_1; \\
\mathbf{q}_{12} &= \frac{\lambda_2}{\lambda_2 - \lambda_1}\boldsymbol{\alpha}_1 + \frac{\lambda_1}{\lambda_1 - \lambda_2}\boldsymbol{\alpha}_2; \\
\mathbf{q}_{123} &= \frac{\lambda_3\lambda_2}{(\lambda_3 - \lambda_1)(\lambda_2 - \lambda_1)}\boldsymbol{\alpha}_1 + \frac{\lambda_1\lambda_3}{(\lambda_1 - \lambda_2)(\lambda_3 - \lambda_2)}\boldsymbol{\alpha}_2 \\
&\quad + \frac{\lambda_1\lambda_2}{(\lambda_1 - \lambda_3)(\lambda_2 - \lambda_3)}\boldsymbol{\alpha}_3.
\end{aligned} \tag{3.1}$$

It is easy to verify that $\mathbf{q}_{1\ldots i}\mathbf{e} = 1$, $1 \le i \le 3$. By Theorem 4.1 below, $\text{conv}\{\boldsymbol{\alpha}_1, \boldsymbol{\alpha}_2, \boldsymbol{\alpha}_3\}$ is a subset of $\text{conv}\{\mathbf{q}_{1\ldots i}, 1 \le i \le 3\}$. Denote by $Q_{1,2,3}$ a 3×3 matrix with rows $\{\boldsymbol{\alpha}_1, \boldsymbol{\alpha}_2, \boldsymbol{\alpha}_3\}$ and $Q_{1,12,123}$ a 3×3 matrix with rows $\{\mathbf{q}_{1\ldots i}, 1 \le i \le 3\}$. Immediately, we have $Q_{1,2,3}T = D(\boldsymbol{\lambda})Q_{1,2,3}$. Equation (3.1) leads to

$$Q_{1,12,123}T = S(\boldsymbol{\lambda})Q_{1,12,123}. \tag{3.2}$$

Thus, for all $\boldsymbol{\alpha}$ in $\text{conv}\{\boldsymbol{\alpha}_1, \boldsymbol{\alpha}_2, \boldsymbol{\alpha}_3\}$, $(\boldsymbol{\alpha}, T)$ has a diagonal *PH*-representation $(\boldsymbol{\beta}, D(\boldsymbol{\lambda}))$ with $\boldsymbol{\alpha} = \boldsymbol{\beta}Q_{1,2,3}$. For all $\boldsymbol{\alpha}$ in $\text{conv}\{\mathbf{q}_{1\ldots i}, 1 \le i \le 3\}$, $(\boldsymbol{\alpha}, T)$ has an ordered Coxian representation $(\boldsymbol{\beta}, S(\boldsymbol{\lambda}))$, where $\boldsymbol{\beta}$ satisfies $\boldsymbol{\alpha} = \boldsymbol{\beta}Q_{1,12,123}$. By Dehon and Latouche[15], $\text{conv}\{\mathbf{q}_{1\ldots i}, 1 \le i \le 3\}$ contains all probability distributions with a bi-diagonal *PH*-representation of order 3 or a smaller order with eigenvalues $-\lambda^1, -\lambda^2$, and $-\lambda^3$.

For the *PH*-generator T given in Equation (3.3), its three eigenvalues are $-\boldsymbol{\lambda} = (-6.7151, -5.0333, -3.2515)$.

$$T = \begin{pmatrix} -4 & 3.5 & 0 \\ 0 & -4 & 3 \\ 0.2 & 0 & -7 \end{pmatrix} \quad \text{and} \quad Q_{1,12,123} = \begin{pmatrix} -0.0721 & 0.0930 & 0.9792 \\ -0.3127 & 1.5121 & -0.1994 \\ 1.3434 & -0.2868 & -0.0566 \end{pmatrix}. \tag{3.3}$$

Denote by Ω_3 the set of all probability functions that are affine combinations of exponential distributions $\{(\boldsymbol{\alpha}_i, T), 1 \leq i \leq 3\}$. In Figure 1, three polytopes conv$\{\boldsymbol{\alpha}_1, \boldsymbol{\alpha}_2, \boldsymbol{\alpha}_3\}$, conv$\{\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3\}$, and conv$\{\mathbf{q}_{1\ldots i}, 1 \leq i \leq 3\}$, and the convex set Ω_3 are plotted (see Dehon and Latouche[15] for the construction of Ω_3). For any *PH*-representation $(\boldsymbol{\alpha}, T)$, if $\boldsymbol{\alpha}$ is in conv$\{\mathbf{q}_{1\ldots i}, 1 \leq i \leq 3\}$, $(\boldsymbol{\alpha}, T)$ has an ordered Coxian representation of order 3. Otherwise, $(\boldsymbol{\alpha}, T)$ does not have an ordered Coxian representation of order 3. It was shown in He and Zhang[17] that a *PH*-representation $(\boldsymbol{\alpha}, T)$ of order 3 always has an ordered Coxian representation of order 4 or a smaller order. It was shown in O'Cinneide[29] that the *PH*-order of $(\boldsymbol{\alpha}, T)$ can be very large (e.g., if $\boldsymbol{\alpha}$ is close to the extreme point $\mathbf{q}_{123}$ but is outside of conv$\{\mathbf{q}_{1\ldots i}, 1 \leq i \leq 3\}$). On the other hand, for probability vector $\boldsymbol{\alpha}$ near $\mathbf{e}_3$, $(\boldsymbol{\alpha}, T)$ has *PH*-order 3, but the triangular order of $(\boldsymbol{\alpha}, T)$ is 4 (Theorem 4.5 in He and Zhang[17]).

In addition, we can construct other (smaller) *PH*-invariant polytopes under T as follows:

$$\begin{aligned} \mathbf{q}_{13} &= \frac{\lambda_3}{\lambda_3 - \lambda_1}\boldsymbol{\alpha}_1 + \frac{\lambda_1}{\lambda_1 - \lambda_3}\boldsymbol{\alpha}_3; \\ \mathbf{q}_{23} &= \frac{\lambda_3}{\lambda_3 - \lambda_2}\boldsymbol{\alpha}_2 + \frac{\lambda_2}{\lambda_2 - \lambda_3}\boldsymbol{\alpha}_3. \end{aligned} \tag{3.4}$$

It can be verified that polytopes conv$\{\mathbf{q}_1, \mathbf{q}_{13}\}$ and conv$\{\boldsymbol{\alpha}_2, \mathbf{q}_{23}\}$ are *PH*-invariant under T, which are associated with ordered Coxian generators of order 2. Therefore, for any $\boldsymbol{\alpha}$ in these subsets, $(\boldsymbol{\alpha}, T)$ has an ordered Coxian representation of order 2 or 1. For example, if $\boldsymbol{\alpha} = \beta_1\mathbf{q}_1 + \beta_3\mathbf{q}_{13}$ with $\beta_1 + \beta_3 = 1$, $\beta_1 \geq 0$, and $\beta_3 \geq 0$, then $((\beta_1, \beta_3), S(\lambda_1, \lambda_3))$ represents the same distribution as $(\boldsymbol{\alpha}, T)$.

Even though the *PH*-generator T is not triangular, $(\boldsymbol{\alpha}, T)$ may have an ordered Coxian representation. This example shows the relationship between different subsets of probability distributions (Botta et al.[6]).

Example 3.3 (*PH*-Invariant Polytope Associated with a Jordan Block)**.** Suppose that the Jordan chain of T corresponding to a Jordan block (Lancaster and Tismenetsky[20]) of real eigenvalue $-\lambda$ is given as $\{\boldsymbol{\alpha}_1, \boldsymbol{\alpha}_2, \ldots, \boldsymbol{\alpha}_n\}$ satisfying:

$$\begin{cases} \boldsymbol{\alpha}_1 T = -\lambda\boldsymbol{\alpha}_1; \\ \boldsymbol{\alpha}_k T = -\lambda\boldsymbol{\alpha}_k + \boldsymbol{\alpha}_{k-1}, \quad 2 \leq k \leq n. \end{cases} \tag{3.5}$$

Assume that $\boldsymbol{\alpha}_1\mathbf{e} \neq 0$. We normalize $\boldsymbol{\alpha}_1$ to $\boldsymbol{\alpha}_1/(\boldsymbol{\alpha}_1\mathbf{e})$ so that $\boldsymbol{\alpha}_1\mathbf{e} = 1$. We can then make $\boldsymbol{\alpha}_k\mathbf{e} \neq 0$ by replacing $\boldsymbol{\alpha}_k$ with $\boldsymbol{\alpha}_k + \boldsymbol{\alpha}_1$ if $\boldsymbol{\alpha}_k\mathbf{e} = 0$, $2 \leq k \leq n$.

After normalizing all vectors to have a unit sum, Equation (3.5) becomes

$$\begin{cases} \boldsymbol{\alpha}_1 T = -\lambda \boldsymbol{\alpha}_1; \\ \boldsymbol{\alpha}_2 T = -\lambda \boldsymbol{\alpha}_2 + \lambda_{2,1} \boldsymbol{\alpha}_1; \\ \boldsymbol{\alpha}_k T = -\lambda \boldsymbol{\alpha}_k + \lambda_{k,k-1} \boldsymbol{\alpha}_{k-1} + \lambda_{k,1} \boldsymbol{\alpha}_1, \quad 3 \le k \le n, \end{cases} \tag{3.6}$$

with $\boldsymbol{\alpha}_k \mathbf{e} = 1,\ 1 \le k \le n$. Note that $\lambda_{k,k-1} \neq 0,\ 2 \le k \le n$. Otherwise, the Jordan chain is broken. The polytope conv$\{\boldsymbol{\alpha}_k, 1 \le k \le n\}$ is invariant under T, but it may not be *PH*-invariant under T. Thus, the vectors $\{\boldsymbol{\alpha}_k, 2 \le k \le n\}$ may not correspond to probability distributions. Next, we construct a *PH*-invariant polytope conv$\{\mathbf{q}_k, 1 \le k \le n\}$ that corresponds to a Coxian generator. Let $x_{1,1} = 1$, and, for $2 \le i \le n$,

$$\begin{cases} x_{i,i} = \dfrac{\lambda}{\lambda_{i,i-1}} x_{i-1,i-1}, \quad x_{i,j} = x_{i-1,j-1}, \ 3 \le j \le i-1; \\ x_{i,2} = x_{i-1,1} - \dfrac{\lambda_{i,1}}{\lambda_{i,i-1}} x_{i-1,i-1}; \quad x_{i,1} = 1 - \displaystyle\sum_{j=2}^{i} x_{i,j}. \end{cases} \tag{3.7}$$

We introduce the following vectors:

$$\mathbf{q}_i = x_{i,i} \boldsymbol{\alpha}_i + \sum_{j=1}^{i-1} x_{i,j} \mathbf{q}_j, \quad 1 \le i \le n. \tag{3.8}$$

Let Q be an $n \times m$ matrix with rows $\{\mathbf{q}_i, 1 \le i \le n\}$.

Lemma 3.1 (Jordan Block Polytope Lemma). *Assume that the eigenvector* $\boldsymbol{\alpha}_1$ *corresponding to the Jordan block of interest is not orthogonal to* $\mathbf{e}$. *For vectors* $\{\mathbf{q}_i, 1 \le i \le n\}$ *given in Equation* (3.8), *we have* $\mathbf{q}_i \mathbf{e} = 1,\ 1 \le i \le n$. *The polytope* conv$\{\mathbf{q}_i, 1 \le i \le n\}$ *is PH-invariant under* T *and the corresponding PH-generator is* $S(\lambda, n)$, *which is defined as*

$$S(\lambda, n) = \begin{pmatrix} -\lambda & 0 & \cdots & \cdots & 0 \\ \lambda & -\lambda & \ddots & \vdots & \vdots \\ 0 & \ddots & \ddots & \ddots & \vdots \\ \vdots & \ddots & \ddots & \ddots & 0 \\ 0 & \cdots & 0 & \lambda & -\lambda \end{pmatrix}_{n \times n}. \tag{3.9}$$

For T and $S(\lambda, n)$, we have $QT = S(\lambda, n) Q$. For any vector $\boldsymbol{\alpha}$ in conv$\{\mathbf{q}_k, 1 \le k \le n\}$, the matrix-exponential distribution $(\boldsymbol{\alpha}, T, \mathbf{e})$ is a mixed Erlang distribution with a *PH*-representation $(\boldsymbol{\beta}, S(\lambda, n))$, where $\boldsymbol{\beta}$ is a probability vector satisfying $\boldsymbol{\alpha} = \boldsymbol{\beta} Q$. (Note: Mixed Erlang distributions are special Coxian distributions.)

Proof. By definition (see Equation (3.7)), we have $x_{i,i} \neq 0$ and $x_{i,1} + x_{i,2} + \cdots + x_{i,i} = 1, 1 \leq i \leq n$. It is then easy to verify that $\mathbf{q}_i \mathbf{e} = 1, 1 \leq i \leq n$. Furthermore, $QT = S(\lambda, n)Q$ can be shown by routine calculations as follows. By Equations (3.7) and (3.8), we have $\mathbf{q}_1 T = -\lambda \mathbf{q}_1$ and, for $2 \leq i \leq n$,

$$\begin{aligned}
\mathbf{q}_i T &= x_{i,i} \boldsymbol{\alpha}_i T + \sum_{j=1}^{i-1} x_{i,j} \mathbf{q}_j T \\
&= x_{i,i}(-\lambda \boldsymbol{\alpha}_i + \lambda_{i,i-1} \boldsymbol{\alpha}_{i-1} + \lambda_{i,1} \boldsymbol{\alpha}_1) - x_{i,1} \lambda \mathbf{q}_1 + \sum_{j=2}^{i-1} x_{i,j}(-\lambda \mathbf{q}_j + \lambda \mathbf{q}_{j-1}) \\
&= -\lambda \mathbf{q}_i + \frac{x_{i,i} \lambda_{i,i-1}}{x_{i-1,i-1}} \left[\mathbf{q}_{i-1} - \sum_{j=1}^{i-2} x_{i-1,j} \mathbf{q}_j \right] + x_{i,i} \lambda_{i,1} \mathbf{q}_1 + \lambda \sum_{j=1}^{i-2} x_{i,j+1} \mathbf{q}_j \\
&= -\lambda \mathbf{q}_i + \lambda \mathbf{q}_{i-1} + (x_{i,i} \lambda_{i,1} + \lambda x_{i,2} - \lambda x_{i-1,1}) \mathbf{q}_1 \\
&= -\lambda \mathbf{q}_i + \lambda \mathbf{q}_{i-1}.
\end{aligned} \tag{3.10}$$

This completes the proof of Lemma 3.1.

Note 3.1. Let $\mathcal{A}$ be an $n \times m$ matrix with rows $\{\boldsymbol{\alpha}_i, 1 \leq i \leq n\}$. Let Λ be an $n \times n$ matrix with (i,i)th element $-\lambda$, $(i, i-1)$st element $\lambda_{i,i-1}$, $(i, 1)$st element $\lambda_{i,1}$, and all other elements zero. Equation (3.6) becomes $\mathcal{A} T = \Lambda \mathcal{A}$. Using matrix notation, finding the *PH*-invariant polytope conv$\{\mathbf{q}_i, 1 \leq i \leq n\}$ is equivalent to solving a linear system $X\Lambda = S(\lambda, n)X$ and $X\mathbf{e} = \mathbf{e}$. Then $\{\mathbf{q}_i, 1 \leq i \leq n\}$ can be obtained by $Q = X\mathcal{A}$.

Lemma 3.1 indicates that the extreme point $\mathbf{q}_i$ of the constructed polytope conv$\{\mathbf{q}_i, 1 \leq i \leq n\}$ represents an Erlang distribution of order i with parameter λ, i.e., $(\mathbf{q}_i, T)$ is an Erlang distribution, $1 \leq i \leq n$. Thus, for any $\boldsymbol{\alpha}$ in the polytope conv$\{\mathbf{q}_i, 1 \leq i \leq n\}$, $(\boldsymbol{\alpha}, T, \mathbf{e})$ is a mixture of Erlang distributions. However, the polytope conv$\{\mathbf{q}_i, 1 \leq i \leq n\}$ may not include every probability vector. Consider the following *PH*-generator:

$$T = \begin{pmatrix} -2 & 0 & 0 & 0.1 & 0.5 \\ 0.7 & -2 & 0 & 0.2 & 0 \\ 1 & 0.5 & -2 & 0 & 0.1 \\ 0 & 0 & 0 & -3 & 0.5 \\ 0 & 0 & 0 & 0 & -3 \end{pmatrix}. \tag{3.11}$$

For T, there are two Jordan blocks of orders 2 and 3 corresponding to eigenvalues -3 and -2, respectively. The two Jordan chains corresponding to the two Jordan blocks are $\{\boldsymbol{\alpha}_i, 1 \le i \le 2\}$ and $\{\boldsymbol{\alpha}_i, 3 \le i \le 5\}$, where

$$\boldsymbol{\alpha}_1 = (0.0000,\ 0.0000,\ 0.0000,\ 0.0000,\ 1.0000);$$
$$\boldsymbol{\alpha}_2 = (0.0000,\ 0.0000,\ 0.0000,\ -1.3489,\ -2.3489);$$
$$\boldsymbol{\alpha}_3 = (0.6061,\ 0.0000,\ 0.0000,\ 0.0606,\ 0.3333);$$
$$\boldsymbol{\alpha}_4 = (0.4487,\ 0.3205,\ 0.0000,\ 0.0865,\ 0.1442);$$
$$\boldsymbol{\alpha}_5 = (0.5052,\ -0.6186,\ 1.4433,\ -0.2680,\ -0.0619).$$

By Lemma 3.1, we construct two *PH*-invariant polytopes conv$\{\mathbf{q}_i, 1 \le i \le 2\}$ and conv$\{\mathbf{q}_i, 3 \le i \le 5\}$ from $\{\boldsymbol{\alpha}_i, 1 \le i \le 2\}$ and $\{\boldsymbol{\alpha}_i, 3 \le i \le 5\}$, respectively, where

$$\mathbf{q}_1 = (0.0000,\ 0.0000,\ 0.0000,\ 0.0000,\ 1.0000);$$
$$\mathbf{q}_2 = (0.0000,\ 0.0000,\ 0.0000,\ 6.0000,\ -5.0000);$$
$$\mathbf{q}_3 = (0.6061,\ 0.0000,\ 0.0000,\ 0.0606,\ 0.3333);$$
$$\mathbf{q}_4 = (-0.2440,\ 1.7316,\ 0.0000,\ 0.2007,\ -0.6883);$$
$$\mathbf{q}_5 = (3.8643,\ -10.5920,\ 6.9264,\ -2.1334,\ 2.9347).$$

It is easy to see that neither conv$\{\mathbf{q}_i, 1 \le i \le 2\}$ nor conv$\{\mathbf{q}_i, 3 \le i \le 5\}$ covers all probability vectors. The *PH*-invariant polytope conv$\{\mathbf{q}_i, 1 \le i \le 5\}$ does not cover all probability vectors either. In fact, it can be verified that $\{\mathbf{e}_i, 1 \le i \le 3\}$ are not in conv$\{\mathbf{q}_i, 1 \le i \le 5\}$ while $\{\mathbf{e}_i, 4 \le i \le 5\}$ are in conv$\{\mathbf{q}_i, 1 \le i \le 5\}$. In the next section, we show how to further expand these *PH*-invariant polytopes to cover more (possibly all) probability vectors that are associated with *PH*-distributions with bi-diagonal *PH*-representations.

4. EXPANSION OF *PH*-INVARIANT POLYTOPES

In Dehon and Latouche[15], an expanded *PH*-invariant polytope was constructed in the way shown in Example 3.2 (i.e., from $\{\boldsymbol{\alpha}_1, \boldsymbol{\alpha}_2, \boldsymbol{\alpha}_3\}$ to $\{\mathbf{q}_1, \mathbf{q}_{12}, \mathbf{q}_{123}\}$). That approach works if all the Jordan blocks of T have order 1. In this section, we extend that approach to cases with Jordan blocks having orders larger than 1.

First, we formulate the expansion problem formally. We consider K Jordan blocks corresponding to real eigenvalues $\{\mu_1, \mu_2, \ldots, \mu_K\}$ of T. We assume that $\mu_1 > \mu_2 > \cdots > \mu_K$, where μ_k is the eigenvalue corresponding to a Jordan block of order n_k, $1 \le k \le K$. We also assume that the *PH*-invariant polytope introduced in Example 3.3 has been constructed for each of the K Jordan blocks: $\{\mathbf{q}_j, n_1+n_2+\cdots+n_{k-1} < j \le n_1+n_2+\cdots+n_k\}$ for the kth Jordan block, $1 \le k \le K$. (Note: $n_0 = 0$.) Let $N = n_1 + n_2 + \cdots + n_K$. Let $\boldsymbol{\lambda} = (\lambda_1, \ldots, \lambda_N)$, where $\lambda_j = \mu_k$ if $n_1 + n_2 + \cdots + n_{k-1} < j \le n_1 + n_2 + \cdots + n_k$ for $1 \le k \le K$, and assume that

$$\begin{cases} \mathbf{q}_1 T = -\lambda_1 \mathbf{q}_1; \\ \mathbf{q}_i T = -\lambda_i \mathbf{q}_i + \lambda_{i,i-1}\mathbf{q}_{i-1}, \quad 2 \le i \le N, \end{cases} \tag{4.1}$$

where $\lambda_{i,i-1} = \lambda_i$, if $\mathbf{q}_i$ and $\mathbf{q}_{i-1}$ correspond to the same Jordan block; 0, otherwise. Let $Q_{1,2,\ldots,N}$ be an $N \times m$ matrix with rows $\{\mathbf{q}_i, 1 \le i \le N\}$. Then Equation (4.1) can be written into matrix form:

$$Q_{1,2,\ldots,N} T = \begin{pmatrix} S(\mu_1, n_1) & & \\ & \ddots & \\ & & S(\mu_K, n_K) \end{pmatrix} Q_{1,2,\ldots,N} \equiv S(\boldsymbol{\mu}, \mathbf{n}) Q_{1,2,\ldots,N}, \tag{4.2}$$

where $\boldsymbol{\mu} = (\mu_1, \mu_2, \ldots, \mu_K)$ and $\mathbf{n} = (n_1, n_2, \ldots, n_K)$. The matrix $S(\boldsymbol{\mu}, \mathbf{n})$ is a *PH*-generator and $\mathrm{conv}\{\mathbf{q}_i, 1 \le i \le N\}$ is a *PH*-invariant polytope under T. If $\boldsymbol{\alpha}$ is in the *PH*-invariant polytope $\mathrm{conv}\{\mathbf{q}_i, 1 \le i \le N\}$, $(\boldsymbol{\beta}, S(\boldsymbol{\mu}, \mathbf{n}))$ is a bi-diagonal *PH*-representation of $(\boldsymbol{\alpha}, T, \mathbf{e})$, where $\boldsymbol{\alpha} = \boldsymbol{\beta} Q_{1,2,\ldots,N}$.

Next, we expand the polytope $\mathrm{conv}\{\mathbf{q}_i, 1 \le i \le N\}$. Define $\{y_{i,j}, 1 \le j \le i \le N\}$ as: $y_{1,1} = 1$, and for $2 \le i \le N$, if $\lambda_{i,i-1} = 0$,

$$\begin{cases} y_{i,i-1} = -\dfrac{\lambda_i}{\lambda_{i-1} - \lambda_i}, \quad y_{i,j} = \dfrac{\lambda_{j+1}}{\lambda_j - \lambda_i} y_{i,j+1}, \;\; 1 \le j \le i-2; \\ y_{i,i} = 1 - \displaystyle\sum_{j=1}^{i-1} y_{i,j}; \end{cases} \tag{4.3}$$

and if $\lambda_{i,i-1} = \lambda_i$, $y_{i,i} = y_{i-1,i-1}$, and

$$\begin{pmatrix} y_{i,1} \\ y_{i,2} \\ \vdots \\ \vdots \\ y_{i,i-1} \end{pmatrix} = \begin{pmatrix} 1 & 1 & \cdots & \cdots & 1 \\ \frac{\lambda_i - \lambda_1}{\lambda_2} & 1 & 0 & \cdots & 0 \\ 0 & \ddots & \ddots & \ddots & \vdots \\ \vdots & \ddots & \frac{\lambda_i - \lambda_{i-3}}{\lambda_{i-2}} & 1 & 0 \\ 0 & \cdots & 0 & \frac{\lambda_i - \lambda_{i-2}}{\lambda_{i-1}} & 1 \end{pmatrix}^{-1} \begin{pmatrix} 1 - y_{i,i} \\ \frac{\lambda_i}{\lambda_2} y_{i-1,1} \\ \vdots \\ \vdots \\ \frac{\lambda_i}{\lambda_{i-1}} y_{i-1,i-2} \end{pmatrix}. \tag{4.4}$$

Lemma 4.1. *Assume that* $\mu_1 > \mu_2 > \cdots > \mu_K$. *For* $\{y_{i,j}, 1 \le j \le i, 1 \le i \le N\}$ *defined by Equations* (4.3) *and* (4.4), *we have* $y_{i,i} \ge 1$, $1 \le i \le N$, $y_{i,j} \le 0$, $1 \le j \le i-1$, *and* $y_{i,1} + y_{i,2} + \cdots + y_{i,i} = 1$, $1 \le i \le N$.

Proof. If $\lambda_{i,i-1} = 0$, we have $\lambda_i < \lambda_j$, $1 \le j \le i-1$. Therefore, $\{y_{i,j}, 1 \le j \le i\}$ in Equation (4.3) are well-defined and the claim is true. By routine calculations, we obtain the determination of the matrix on the right hand side of Equation (4.4) as

$$1 + \frac{(\lambda_1 - \lambda_i)}{\lambda_2}\left(1 + \cdots \left(1 + \frac{(\lambda_{i-3} - \lambda_i)}{\lambda_{i-2}}\left(1 + \frac{\lambda_{i-2} - \lambda_i}{\lambda_{i-1}}\right)\right)\cdots\right), \tag{4.5}$$

which is no less than one, since elements in $\boldsymbol{\lambda}$ are ordered in non-increasing order. Therefore, the inverse matrix exists and $\{y_{i,j}, 1 \le j \le i-1\}$ in Equation (4.4) are well-defined. It is readily seen that Equation (4.4) is equivalent to

$$\begin{cases} y_{i,i} = y_{i-1,i-1}, \\ 1 = \sum_{j=1}^{i} y_{i,j}, \\ y_{i,j} = \dfrac{\lambda_i}{\lambda_j} y_{i-1,j-1} + \dfrac{(\lambda_{j-1} - \lambda_i)}{\lambda_j} y_{i,j-1}, \quad 2 \le j \le i-1. \end{cases} \tag{4.6}$$

Since $y_{1,1} = 1$, (by induction), $y_{i-1,i-1} \ge 1$. Thus, $y_{i,i} \ge 1$. By the second equality in Equation (4.6) and routine calculations, we obtain, for $2 \le j \le i-1$,

$$y_{i,j} = \frac{\lambda_i}{\lambda_j} y_{i-1,j-1} + \sum_{k=1}^{j-2} \frac{\lambda_i}{\lambda_{k+1}}\left(\prod_{s=k+1}^{j-1}\left(\frac{\lambda_s - \lambda_i}{\lambda_{s+1}}\right)\right) y_{i-1,k} + \prod_{s=1}^{j-1}\left(\frac{\lambda_s - \lambda_i}{\lambda_{s+1}}\right) y_{i,1}. \tag{4.7}$$

Taking summation on both sides of Equation (4.7), we obtain

$$\begin{aligned} \sum_{j=1}^{i-1} y_{i,j} = \sum_{k=1}^{i-2} \frac{\lambda_i}{\lambda_{k+1}}\left(1 + \sum_{j=k+2}^{i-1} \prod_{s=k+1}^{j-1}\left(\frac{\lambda_s - \lambda_i}{\lambda_{s+1}}\right)\right) y_{i-1,k} \\ + y_{i,1}\left(1 + \sum_{j=2}^{i-1} \prod_{s=1}^{j-1}\left(\frac{\lambda_s - \lambda_i}{\lambda_{s+1}}\right)\right). \end{aligned} \tag{4.8}$$

Using $y_{i,i} = y_{i-1,i-1}$ and equation (4.8), we have

$$\begin{aligned} & y_{i,1}\left(1 + \sum_{j=2}^{i-1} \prod_{s=1}^{j-1}\left(\frac{\lambda_s - \lambda_i}{\lambda_{s+1}}\right)\right) \\ & = \sum_{k=1}^{i-2} y_{i-1,k}\left(1 - \frac{\lambda_i}{\lambda_{k+1}}\left(1 + \sum_{j=k+2}^{i-1} \prod_{s=k+1}^{j-1}\left(\frac{\lambda_s - \lambda_i}{\lambda_{s+1}}\right)\right)\right). \end{aligned} \tag{4.9}$$

It is easy to verify that $\frac{\lambda_i}{\lambda_{i-1}} \leq 1$ and $\frac{\lambda_i}{\lambda_{i-2}}\left(1 + \frac{\lambda_{i-2}-\lambda_i}{\lambda_{i-1}}\right) = \frac{\lambda_i}{\lambda_{i-2}} + \left(\frac{\lambda_{i-2}-\lambda_i}{\lambda_{i-2}}\right)\frac{\lambda_i}{\lambda_{i-1}} \leq 1$. By induction, we have, for $1 \leq k \leq i-3$,

$$
\begin{aligned}
&\frac{\lambda_i}{\lambda_{k+1}}\left(1 + \sum_{j=k+2}^{i-2} \prod_{s=k+1}^{j-1} \left(\frac{\lambda_s - \lambda_i}{\lambda_{s+1}}\right)\right) \\
&= \frac{\lambda_i}{\lambda_{k+1}}\left(1 + \frac{(\lambda_{k+1} - \lambda_i)}{\lambda_{k+2}}\left(1 + \sum_{j=k+3}^{i-1} \prod_{s=k+2}^{j-1} \left(\frac{\lambda_s - \lambda_i}{\lambda_{s+1}}\right)\right)\right) \\
&= \frac{\lambda_i}{\lambda_{k+1}} + \left(\frac{\lambda_{k+1} - \lambda_i}{\lambda_{k+1}}\right)\left(\frac{\lambda_i}{\lambda_{k+2}}\right)\left(1 + \sum_{j=k+3}^{i-1} \prod_{s=k+2}^{j-1} \left(\frac{\lambda_s - \lambda_i}{\lambda_{s+1}}\right)\right) \\
&\leq \frac{\lambda_i}{\lambda_{k+1}} + \frac{\lambda_{k+1} - \lambda_i}{\lambda_{k+1}} = 1. \qquad (4.10)
\end{aligned}
$$

By induction and Equations (4.9) and (4.10), we have $y_{i,1} \leq 0$. By Equation (4.7) and induction, we have $y_{i,j} \leq 0$, $1 \leq j \leq i-1$. This completes the proof of Lemma 4.1.

Now, we expand the polytope $\text{conv}\{\mathbf{q}_i, 1 \leq i \leq N\}$ to $\text{conv}\{\mathbf{q}_{1\ldots i}, 1 \leq i \leq N\}$ as follows:

$$
\mathbf{q}_{1\ldots i} = y_{i,i}\mathbf{q}_i + \sum_{j=1}^{i-1} y_{i,j}\mathbf{q}_{1\ldots j}, \quad 1 \leq i \leq N. \qquad (4.11)
$$

Let $Q_{1,12,\ldots,1\ldots N}$ be a matrix with rows $\{\mathbf{q}_{1\ldots i}, 1 \leq i \leq N\}$. Note that Lemma 4.1 implies that the points $\{\mathbf{q}_i, 1 \leq i \leq N\}$ are all located in the new polytope $\text{conv}\{\mathbf{q}_{1\ldots i}, 1 \leq i \leq N\}$. That conclusion is similar to that of Theorem 3 in Dehon and Latouche[15].

Theorem 4.1 (Expanded *PH*-Invariant Polytope). *Let T be a PH-generator and $\{-\mu_1, -\mu_2, \ldots, -\mu_K\}$ be K real eigenvalues of T, which are ordered as $\mu_1 > \mu_2 > \cdots > \mu_K$. Assume that the PH-invariant polytopes associated with Jordan blocks for eigenvalues $\{-\mu_1, -\mu_2, \ldots, -\mu_K\}$ and extreme points $\{\mathbf{q}_i, 1 \leq i \leq N\}$ are constructed. Then the polytope* $\text{conv}\{\mathbf{q}_{1\ldots i}, 1 \leq i \leq N\}$ *is PH-invariant under T with an ordered Coxian generator $S(\boldsymbol{\lambda})$, i.e., $Q_{1,12,\ldots,1\ldots N}T = S(\boldsymbol{\lambda})Q_{1,12,\ldots,1\ldots N}$. In addition, we have* $\text{conv}\{\mathbf{q}_i, 1 \leq i \leq N\} \subseteq$ *conv*$\{\mathbf{q}_{1\ldots i}, 1 \leq i \leq N\}$. *For any PH-distribution $(\boldsymbol{\alpha}, T)$, if $\boldsymbol{\alpha}$ is in* $\text{conv}\{\mathbf{q}_{1\ldots i}, 1 \leq i \leq N\}$, *then there exists a nonnegative vector $\boldsymbol{\beta}$ such that $\boldsymbol{\alpha} = \boldsymbol{\beta}Q_{1,12,\ldots,1\ldots N}$ and $(\boldsymbol{\beta}, S(\boldsymbol{\lambda}))$ is an ordered Coxian representation of $(\boldsymbol{\alpha}, T)$.*

Note 4.1. Implicitly, we assume that none of the eigenvectors associated with the Jordan blocks is orthogonal to $\mathbf{e}$.

Proof. First, it is easy to see that $\mathbf{q}_{1\ldots i}\mathbf{e} = 1$, $1 \leq i \leq N$. Then we check that the polytope conv$\{\mathbf{q}_{1\ldots i},\ 1 \leq i \leq N\}$ is *PH*-invariant under T. It is readily seen that $\mathbf{q}_1 T = -\lambda_1 \mathbf{q}_1$ and, for $2 \leq i \leq N$, by Equations (4.1) and (4.11),

$$
\begin{aligned}
\mathbf{q}_{1\ldots i} T &= y_{i,i}(-\lambda_i \mathbf{q}_i + \lambda_{i,i-1}\mathbf{q}_{i-1}) + \sum_{j=1}^{i-1} y_{i,j}(-\lambda_j \mathbf{q}_{1\ldots j} + \lambda_j \mathbf{q}_{1\ldots(j-1)}) \\
&= -\lambda_i \mathbf{q}_{1\ldots i} + y_{i,i}\lambda_{i,i-1}\mathbf{q}_{i-1} + \lambda_i \sum_{j=1}^{i-1} y_{i,j}\mathbf{q}_{1\ldots j} + \sum_{j=1}^{i-1} y_{i,j}(-\lambda_j \mathbf{q}_{1\ldots j} + \lambda_j \mathbf{q}_{1\ldots(j-1)}) \\
&= -\lambda_i \mathbf{q}_{1\ldots i} + \frac{y_{i,i}\lambda_{i,i-1}}{y_{i-1,i-1}}\left(\mathbf{q}_{1\ldots(i-1)} - \sum_{j=1}^{i-2} y_{i-1,j}\mathbf{q}_{1\ldots j}\right) \\
&\quad + \sum_{j=1}^{i-1} y_{i,j}(\lambda_i \mathbf{q}_{1\ldots j} - \lambda_j \mathbf{q}_{1\ldots j} + \lambda_j \mathbf{q}_{1\ldots(j-1)}) \\
&= -\lambda_i \mathbf{q}_{1\ldots i} + \left(\frac{y_{i,i}\lambda_{i,i-1}}{y_{i-1,i-1}} + (\lambda_i - \lambda_{i-1})y_{i,i-1}\right)\mathbf{q}_{1\ldots(i-1)} \\
&\quad + \sum_{j=1}^{i-2}\left(-\frac{y_{i,i}\lambda_{i,i-1}}{y_{i-1,i-1}} y_{i-1,j} + (\lambda_i - \lambda_j)y_{i,j} + \lambda_{j+1}y_{i,j+1}\right)\mathbf{q}_{1\ldots j} \\
&= -\lambda_i \mathbf{q}_{1\ldots i} + \lambda_i \mathbf{q}_{1\ldots(i-1)}.
\end{aligned}
\tag{4.12}
$$

Therefore, the polytope conv$\{\mathbf{q}_{1\ldots i}, 1 \leq i \leq N\}$ is *PH*-invariant under T, and corresponds to an ordered Coxian generator $S(\boldsymbol{\lambda})$, i.e., $Q_{1,12,\ldots,1\ldots N}T = S(\boldsymbol{\lambda})Q_{1,12,\ldots,1\ldots N}$. Finally, By Lemma 4.1, $y_{i,i} \geq 1$, and $y_{i,j} \leq 0$, $1 \leq j < i$, $1 \leq i \leq m$. Thus, Equation (4.11) shows clearly that $\{\mathbf{q}_i, 1 \leq i \leq N\} \subseteq$ conv$\{\mathbf{q}_{1\ldots i}, 1 \leq i \leq N\}$, which implies conv$\{\mathbf{q}_i, 1 \leq i \leq N\} \subseteq$ conv$\{\mathbf{q}_{1\ldots i}, 1 \leq i \leq N\}$. This completes the proof Theorem 4.1.

Note 4.2. In matrix form, finding the *PH*-invariant polytope conv$\{\mathbf{q}_{1\ldots i}, 1 \leq i \leq N\}$ is equivalent to solving the linear system $YS(\boldsymbol{\mu}, \mathbf{n}) = S(\boldsymbol{\lambda})Y$ and $Y\mathbf{e} = \mathbf{e}$. Then we have $Q_{1,12,\ldots,1\ldots N} = YQ_{1,2,\ldots,N}$.

For two polytopes, we say one polytope is *larger* if the other is a subset of it. In Lemma 4.1 and Theorem 4.1, $\mu_1 > \mu_2 > \cdots > \mu_K$ is required. Without that requirement, a *PH*-invariant polytope can be constructed by using Lemma 4.1 and Theorem 4.1, as long as the eigenvalues are different. However, without the condition $\mu_1 > \mu_2 > \cdots > \mu_K$, the constructed *PH*-invariant polytope is "smaller" and may not include some of the (original) *PH*-invariant polytopes associated with Jordan blocks (See Example 6.1 for more discussion on this issue.) In the rest of this section, we compare *PH*-invariant polytopes constructed by using Lemma 4.1 and Theorem 4.1.

Probabilistically, every vector $\mathbf{q}_{1\ldots i}$ corresponds to a Coxian distribution, for $1 \le i \le N$. It is interesting to see that every polytope conv$\{\mathbf{q}_{1\ldots i}, 1 \le i \le n\}$, for $1 \le n \le N$, is *PH*-invariant under T. If $n_1 = n_2 = \cdots = n_K = 1$, the *PH*-invariant polytope conv$\{\mathbf{q}_{1\ldots i}, 1 \le i \le K\}$ was introduced and investigated in Dehon and Latouche[15]. For this case, Equation (4.11) can be reduced to

$$\mathbf{q}_{1\ldots i} = \sum_{j=1}^{i} \left(\prod_{\substack{k=1 \\ k\neq j}}^{i} \frac{\lambda_k}{\lambda_k - \lambda_j} \right) \mathbf{q}_j, \quad 1 \le i \le K \tag{4.13}$$

which is consistent with Ref.[15]. Dehon and Latouche's theorems in Ref.[15] imply that the polytope conv$\{\mathbf{q}_{1\ldots i}, 1 \le i \le K\}$ is the *largest PH-invariant polytope* under T that is associated with a triangular representation that can be constructed from the individual Coxian distributions represented by $\{(\mathbf{q}_i, T), 1 \le i \le K\}$. By using a perturbation argument, for Jordan block cases, we can conclude that conv$\{\mathbf{q}_{1\ldots i}, 1 \le i \le N\}$ with $\mu_1 > \mu_2 > \cdots > \mu_K$ is the *largest PH-invariant polytope* under T that corresponds to a triangular generator. Thus, similar to Example 3.2, for any $\boldsymbol{\alpha}$ located outside of the polytope conv$\{\mathbf{q}_{1\ldots i}, 1 \le i \le N\}$ associated with all Jordan blocks (note that $\boldsymbol{\alpha}$ is in aff$\{\mathbf{e}_i, 1 \le i \le m\}$), $(\boldsymbol{\alpha}, T, \mathbf{e})$ is either not a *PH*-distribution, or a *PH*-distribution with an triangular order as large as m+1.

Although the polytope conv$\{\mathbf{q}_{1\ldots i}, 1 \le i \le N\}$ is larger than conv$\{\mathbf{q}_i, 1 \le i \le N\}$, there is no guarantee that it covers the entire probability vector polytope (see Example 3.2). Nonetheless, numerical results show that, if we include all Jordan blocks with distinct eigenvalues to construct an expanded *PH*-invariant polytope, it does cover all probability vectors for many cases (if all eigenvalues of T are real) (see Example 6.1).

Our next result shows that the set of all expanded *PH*-invariant polytopes built from Jordan blocks are partially ordered.

Proposition 4.1. *Let T be a PH-generator. Consider any two expanded PH-invariant polytopes constructed with a certain number of the Jordan blocks with real eigenvalues, by using the methods developed in Lemma* 3.1 *and Theorem* 4.1. *Assume that none of the eigenvalues associated with the second expanded PH-invariant polytope has an eigenvector orthogonal to the vector* $\mathbf{e}$. *Assume that the eigenvalues corresponding to the first expanded PH-invariant polytope are ordered increasingly. If the set of Jordan blocks of the second polytope is a subset of that of the first one, then the second expanded PH-invariant polytope is a subset of the first one and, consequently, the set of all ordered Coxian distributions corresponding to the second polytope is a subset of that of the first polytope.*

Proof. Denote by $\{\boldsymbol{\mu}, \mathbf{n}\}$ the set of eigenvalues corresponding to the first expanded *PH*-invariant polytope and orders of their corresponding Jordan blocks. By Theorem 4.1, the expanded *PH*-invariant polytope associated with $\{\boldsymbol{\mu}, \mathbf{n}\}$ satisfies $Q_a T = S(\boldsymbol{\lambda})Q_a$, $Q_a\mathbf{e} = \mathbf{e}$, where $\boldsymbol{\lambda}$ is defined from $\{\boldsymbol{\mu}, \mathbf{n}\}$ and the rows of the matrix Q_a are the extreme points of the expanded *PH*-invariant polytope. The set of eigenvalues corresponding to the second expanded *PH*-invariant polytope and orders are $\{\gamma, \mathbf{k}\}$. By the assumption, γ is a subset of $\boldsymbol{\mu}$ and $\mathbf{k}$ is a subset of $\mathbf{n}$. By Theorem 4.1, the expanded *PH*-invariant polytope associated with $\{\gamma, \mathbf{k}\}$ satisfies $Q_b T = S(\boldsymbol{\eta})Q_b$, and $Q_b\mathbf{e} = \mathbf{e}$, where $\boldsymbol{\eta}$ is defined from $\{\gamma, \mathbf{k}\}$, similar to the way that $\boldsymbol{\lambda}$ was defined from $\{\boldsymbol{\mu}, \mathbf{n}\}$.

By Proposition 5.1 in He and Zhang[17], there exists a nonnegative matrix P with unit row sums such that $S(\boldsymbol{\eta})P = PS(\boldsymbol{\lambda})$ (Note that $\mu_1 > \mu_2 > \cdots > \mu_K$). Then we have $PQ_aT = S(\boldsymbol{\eta})PQ_a$. Combining with $Q_bT = S(\boldsymbol{\eta})Q_b$, we obtain $(PQ_a - Q_b)T = S(\boldsymbol{\eta})(PQ_a - Q_b)$. Denote the first row of $PQ_a - Q_b$ by $\mathbf{u}_1$. Then we have $\mathbf{u}_1 T = -\eta_1\mathbf{u}_1$ and $\mathbf{u}_1\mathbf{e} = 0$. By the assumption, we must have $\mathbf{u}_1 = 0$. Denote the second row of $PQ_a - Q_b$ by $\mathbf{u}_2$. Then we have $\mathbf{u}_2 T = -\eta_2\mathbf{u}_2 + \eta_2\mathbf{u}_1 = -\eta_2\mathbf{u}_2$ and $\mathbf{u}_2\mathbf{e} = 0$. Again, by the assumption, we must have $\mathbf{u}_2 = 0$. By induction, it can be shown that $PQ_a - Q_b = 0$, i.e., $PQ_a = Q_b$. Since P is nonnegative with unit row sums, it is readily seen that the second polytope is a subset of the first one. This completes the proof of Proposition 4.1.

Proposition 4.1 indicates that, in order to include more probability vectors, we should choose as many Jordan blocks as possible in the construction of expanded *PH*-invariant polytopes. The largest expanded *PH*-invariant polytope is the one constructed with all Jordan blocks with distinct real eigenvalues in nondecreasing order (see Example 6.1).

Example 4.1 (Example 3.3 Continued). In Example 3.3, two *PH*-invariant polytopes $\{\mathbf{q}_i, 1 \le i \le 2\}$ and $\{\mathbf{q}_i, 3 \le i \le 5\}$ were constructed from two Jordan Blocks. By using Lemma 4.1 and Theorem 4.1, an expanded *PH*-invariant polytope can be constructed as follows:

$$\mathbf{q}_1 = (0.0000,\ 0.0000,\ 0.0000,\ 0.0000,\ 1.0000);$$

$$\mathbf{q}_{12} = (0.0000,\ 0.0000,\ 0.0000,\ 6.0000,\ -5.0000);$$

$$\mathbf{q}_{123} = (5.4545,\ 0.0000,\ 0.0000,\ -11.4545,\ 7.0000);$$

$$\mathbf{q}_{1234} = (-24.0142,\ 15.5844,\ 0.0000,\ 23.6246,\ -14.1948);$$

$$\mathbf{q}_{12345} = (109.0170,\ -157.6657,\ 62.3377,\ -67.8806,\ 55.1916).$$

It can be verified that $\{\mathbf{e}_i, 1 \le i \le 5\}$ are convex combinations of $\{\mathbf{q}_1, \mathbf{q}_{12}, \mathbf{q}_{123}, \mathbf{q}_{1234}, \mathbf{q}_{12345}\}$ as follows: (i.e., $\mathbf{e}_i = \boldsymbol{\beta}_i Q_{1,12,\ldots,12345}, 1 \le i \le 5$)

$$\mathbf{e}_1 : \boldsymbol{\beta}_1 = (0.4667,\ 0.3500,\ 0.1833,\ 0.0000,\ 0.0000);$$

$$\mathbf{e}_2 : \boldsymbol{\beta}_2 = (0.3667,\ 0.2867,\ 0.2827,\ 0.0642,\ 0.0000);$$

$$\mathbf{e}_3 : \boldsymbol{\beta}_3 = (0.1333,\ 0.2944,\ 0.3939,\ 0.1623,\ 0.0160);$$

$$\mathbf{e}_4 : \boldsymbol{\beta}_4 = (0.8333,\ 0.1667,\ 0.0000,\ 0.0000,\ 0.0000);$$

$$\mathbf{e}_5 : \boldsymbol{\beta}_5 = (1.0000,\ 0.0000,\ 0.0000,\ 0.0000,\ 0.0000).$$

Since $\{\boldsymbol{\beta}_i, 1 \le i \le 5\}$ are all nonnegative, we know that the probability vector polytope conv$\{\mathbf{e}_i, 1 \le i \le 5\}$ is a subset of the expanded *PH*-invariant polytope conv$\{\mathbf{q}_1, \mathbf{q}_{12}, \mathbf{q}_{123}, \mathbf{q}_{1234}, \mathbf{q}_{12345}\}$. Thus, all *PH*-distributions $(\boldsymbol{\alpha}, T)$ have an ordered Coxian representation of order 5 or a smaller order.

Finally, we like to point out that adding more Jordan blocks increases the order of the corresponding matrix representation. For a particular initial probability vector $\boldsymbol{\alpha}$, we aim to choose the minimal set of Jordan blocks to construct a *PH*-invariant polytope that includes $\boldsymbol{\alpha}$. One method to select such a subset of Jordan blocks is introduced in the next section.

5. AN ORDER REDUCTION ALGORITHM

In this section, we develop an algorithm that can be used to find a (possibly) smaller *PH*-representation for a given *PH*-representation. The algorithm is designed to reduce the order of a *PH*-representation by considering the locations of the vector $\mathbf{e}$ and the initial probability vector $\boldsymbol{\alpha}$, two parameter vectors of a *PH*-representation $(\boldsymbol{\alpha}, T, \mathbf{e})$. The algorithm is based on two general observations (Propositions 5.1 and 5.2).

Given N vectors $\{\mathbf{q}_1, \mathbf{q}_2, \ldots, \mathbf{q}_N\}$ in $\Re^m$, we assume that the polytope conv$\{\mathbf{q}_1, \mathbf{q}_2, \ldots, \mathbf{q}_N\}$ is invariant under T. We also assume that, for some n, $1 \le n < N$, the set conv$\{\mathbf{q}_1, \ldots, \mathbf{q}_n\}$ is invariant under T as well. The assumptions can be expressed in the following matrix form:

$$QT = \begin{pmatrix} S_1 & 0 \\ S_{21} & S_2 \end{pmatrix} Q, \tag{5.1}$$

where S_1 is an $n \times n$ matrix, S_{21} an $(N-n) \times n$ matrix, S_2 is an $(N-n) \times (N-n)$ matrix, Q is an $N \times n$ matrix with rows $\{\mathbf{q}_1, \ldots, \mathbf{q}_N\}$.

Proposition 5.1 ($\mathbf{e}$-Reduction Method). *For a PH-generator T, we assume that $\mathbf{q}_i\mathbf{e} = 0$, for $1 \le i \le n$, and $\mathbf{q}_i\mathbf{e} = 1$, for $n+1 \le i \le N$. We also assume that S_2 is a PH-generator. For a given initial probability vector $\boldsymbol{\alpha}$, there* *exists* $\mathbf{x} \in \Re^N$ *such that* $\boldsymbol{\alpha} = \mathbf{x}Q$. *Then the PH-distribution* $(\boldsymbol{\alpha}, T)$ *has a matrix-exponential representation* $(\boldsymbol{\beta}_{[n+1,N]}, S_2, \mathbf{e})$, *where* $\boldsymbol{\beta}_{[n+1,N]} = (x_{n+1}, \ldots, x_N)$. *If* $\boldsymbol{\beta}_{[n+1,N]}$ *is nonnegative, then* $(\boldsymbol{\alpha}, T)$ *has a new PH-representation* $(\boldsymbol{\beta}_{[n+1,N]}, S_2)$ *of order* $N - n$.

Proof. By Equation (5.1), we have $QT^k = \begin{pmatrix} S_1 & 0 \\ S_{21} & S_2 \end{pmatrix}^k Q$, for $k \geq 0$. That leads to, for $t \geq 0$,

$$\begin{aligned} \mathbf{x}Q\exp\{Tt\}\mathbf{e} &= \mathbf{x}\exp\left\{\begin{pmatrix} S_1 & 0 \\ S_{21} & S_2 \end{pmatrix} t\right\} Q\mathbf{e} \\ &= \mathbf{x}\begin{pmatrix} \exp\{S_1 t\} & 0 \\ S_{21}(t) & \exp\{S_2 t\} \end{pmatrix}\begin{pmatrix} 0 \\ \mathbf{e} \end{pmatrix} \\ &= \boldsymbol{\beta}_{[n+1,N]}\exp\{S_2 t\}\mathbf{e}. \end{aligned} \tag{5.2}$$

This completes the proof of Proposition 5.1.

Denote by $\boldsymbol{\alpha} = x_1\mathbf{q}_1 + x_2\mathbf{q}_2 + \cdots + x_N\mathbf{q}_N \equiv \boldsymbol{\alpha}_1 + \boldsymbol{\alpha}_2$, where $\boldsymbol{\alpha}_1 \in \text{span}\{\mathbf{q}_1, \ldots, \mathbf{q}_n\}$ (the subspace generated by $\{\mathbf{q}_1, \ldots, \mathbf{q}_n\}$) and $\boldsymbol{\alpha}_2 \in \text{conv}\{\mathbf{q}_{n+1}, \ldots, \mathbf{q}_N\}$. By Proposition 5.1, $(\boldsymbol{\alpha}, T)$ has the same probability distribution as $(\boldsymbol{\alpha}_2, T)$ for any $\boldsymbol{\alpha}_1 \in \text{span}\{\mathbf{q}_1, \ldots, \mathbf{q}_n\}$. Thus, an alternative *PH*-representation for $(\boldsymbol{\alpha}_2, T)$ is an alternative *PH*-representation for $(\boldsymbol{\alpha}_1 + \boldsymbol{\alpha}_2, T)$ for any $\boldsymbol{\alpha}_1$ in $\text{span}\{\mathbf{q}_1, \ldots, \mathbf{q}_n\}$. The implication is that when we construct *PH*-invariant polytopes for a simpler *PH*-representation for $(\boldsymbol{\alpha}, T)$, we can concentrate on the subset $\text{conv}\{\mathbf{q}_{n+1}, \ldots, \mathbf{q}_N\}$.

We can use Proposition 5.1 to exclude vectors in the Jordan chains from the construction of expanded *PH*-invariant polytopes. According to Section 4, a *PH*-invariant polytope corresponding to an Erlang representation can be constructed for each Jordan block of T if the corresponding eigenvectors have nonzero sums. Suppose that we have two such *PH*-invariant polytopes ($\text{conv}\{\boldsymbol{\alpha}_k, 1 \leq k \leq n\}$ and $\text{conv}\{\gamma_k, 1 \leq k \leq n+s\}$) constructed for two Jordan blocks with the same eigenvalue $-\lambda$, satisfying:

$$\begin{aligned} \boldsymbol{\alpha}_1 T = -\lambda\boldsymbol{\alpha}_1; \quad \boldsymbol{\alpha}_k T = -\lambda\boldsymbol{\alpha}_k + \lambda\boldsymbol{\alpha}_{k-1}, \quad 2 \leq k \leq n; \\ \gamma_1 T = -\lambda\gamma_1; \quad \gamma_k T = -\lambda\gamma_k + \lambda\gamma_{k-1}, \quad 2 \leq k \leq n+s, \end{aligned} \tag{5.3}$$

with $\boldsymbol{\alpha}_k\mathbf{e} = 1$, $1 \leq k \leq n$, and $\gamma_k\mathbf{e} = 1$, $1 \leq k \leq n+s$. Then we can obtain, for $2 \leq k \leq n$,

$$\begin{aligned} &(\boldsymbol{\alpha}_1 - \gamma_1)T = -\lambda(\boldsymbol{\alpha}_1 - \gamma_1); \\ &(\boldsymbol{\alpha}_k - \gamma_k)T = -\lambda(\boldsymbol{\alpha}_k - \gamma_k) + \lambda(\boldsymbol{\alpha}_{k-1} - \gamma_{k-1}). \end{aligned} \tag{5.4}$$

Thus, we can use $\{\boldsymbol{\alpha}_k - \gamma_k, 1 \le k \le n\}$ to replace $\{\boldsymbol{\alpha}_k, 1 \le k \le n\}$. Since $(\boldsymbol{\alpha}_k - \gamma_k)\mathbf{e} = 0, 1 \le k \le n$, and the polytope $\text{conv}\{\boldsymbol{\alpha}_k - \gamma_k, 1 \le k \le n\}$ is invariant under T, by Proposition 5.1, it can be excluded from the construction of expanded *PH*-invariant polytopes. The implication is that, for each real eigenvalue of T, we only need to consider its Jordan block with the largest order.

Next, we look at possible order reduction if $\boldsymbol{\alpha}$ is located in a subspace in $\Re^m$. Suppose that there is a *PH*-invariant polytope $\text{conv}\{\mathbf{q}_1, \mathbf{q}_2, \ldots, \mathbf{q}_n\}$ under T with a *PH*-generator S and $n < m$. Suppose that $\boldsymbol{\alpha} = \sum_{i=1}^{n} \beta_i \mathbf{q}_i$ and denote by $\boldsymbol{\beta} = (\beta_1, \ldots, \beta_n)$.

Proposition 5.2 ($\boldsymbol{\alpha}$-Reduction Method). *Assume that* $\mathbf{q}_i\mathbf{e} = 1$, *for* $1 \le i \le n$, *and the polytope* $\text{conv}\{\mathbf{q}_1, \mathbf{q}_2, \ldots, \mathbf{q}_n\}$ *is PH-invariant under* T. *Then the PH-distribution* $(\boldsymbol{\alpha}, T)$ *has a matrix-exponential representation* $(\boldsymbol{\beta}, S, \mathbf{e})$. *If* $\boldsymbol{\beta}$ *is nonnegative, then* $(\boldsymbol{\alpha}, T)$ *has a new PH-representation* $(\boldsymbol{\beta}, S)$ *of order* n.

Note 5.1. The basic idea of the $\mathbf{e}$-reduction method and the $\boldsymbol{\alpha}$-reduction method is, if $\mathbf{e}$ or $\boldsymbol{\alpha}$ is located in an invariant subspace under T, then $(\boldsymbol{\alpha}, T)$ may have a *PH*-representation of a smaller order. Using the $\mathbf{e}$-reduction and $\boldsymbol{\alpha}$-reduction methods, it seems possible to develop an iterative algorithm for computing a minimal *PH*-representation, which is an interesting topic for more research.

Now, we combine all the methods to propose an algorithm for computing an ordered Coxian representation for $(\boldsymbol{\alpha}, T)$.

A generic order reduction algorithm (we assume that T is a *PH*-generator)

Step 1. Compute the Jordan canonical form of T. Find the cyclic chains.

Step 2. ($\mathbf{e}$-reduction step) Use the $\mathbf{e}$-reduction method to exclude Jordan chains with zero sums and duplicate Jordan chains.

Step 3. ($\boldsymbol{\alpha}$-reduction step) Use the $\boldsymbol{\alpha}$-reduction method to exclude Jordan chains independent of $\boldsymbol{\alpha}$.

Step 4. For each remaining Jordan chain, find its corresponding *PH*-invariant polytope by using the method developed in Lemma 3.1.

Step 5. Use the method developed in Section 4 to construct an expanded *PH*-invariant polytope by using all the *PH*-invariant polytopes found in Step 4.

Step 6. Find the corresponding matrix-exponential representation for $(\boldsymbol{\alpha}, T)$ by Theorem 4.1.

Example 5.1. We consider *PH*-generator T given in Equation (5.5). T has four eigenvalues: -15, -2, -4.3059, and -36.6941, where the eigenvalue -2 corresponds to the Jordan block of order 2.

$$T = \begin{pmatrix} -11 & 9 & 0 & 0 & 1 \\ 4 & -6 & 0 & 0 & 0 \\ 0 & 0 & -15 & 11 & 2 \\ 0 & 0 & 23 & -25 & 0 \\ 0 & 0 & 1 & 0 & -3 \end{pmatrix}, \quad S(15,\ 2,\ 2) = \begin{pmatrix} -15 & 0 & 0 \\ 2 & -2 & 0 \\ 0 & 2 & -2 \end{pmatrix}. \tag{5.5}$$

Jordan chains of T are given as (after normalization whenever appropriate):

$$\begin{aligned} \text{for } -15 : \boldsymbol{\alpha}_1 &= (-13.0000,\ 13.0000,\ -0.0431,\ -0.0474,\ 1.0905); \\ \text{for } -2 : \boldsymbol{\alpha}_2 &= (0.0000,\ 0.0000,\ 0.2875,\ 0.1375,\ 0.5750); \\ \boldsymbol{\alpha}_3 &= (0.3413,\ 0.7679,\ -0.0725,\ -0.0367,\ 0.0000); \\ \text{for } -4.3059 : \boldsymbol{\alpha}_4 &= (0.0000,\ 0.0000\ -0.6529,\ -0.3471,\ 1.0000); \\ \text{for } -36.6941 : \boldsymbol{\alpha}_5 &= (0.0000,\ 0.0000,\ -16.8471,\ 15.8471,\ 1.0000). \end{aligned}$$

Since $\boldsymbol{\alpha}_4\mathbf{e} = \boldsymbol{\alpha}_5\mathbf{e} = 0$, we shall only use $\{\boldsymbol{\alpha}_1, \boldsymbol{\alpha}_2, \boldsymbol{\alpha}_3\}$ to construct an expanded *PH*-invariant polytope. The *PH*-invariant polytope associated with the Jordan block of eigenvalue -2 can be obtained as: $\{\mathbf{q}_2 = \boldsymbol{\alpha}_2 = (0, 0, 0.2875, 0.1375, 0.5750), \mathbf{q}_3 = (2.00, 4.50, -1.8222, -0.8834, -2.7944)\}$. The expanded *PH*-invariant polytope of $\{\boldsymbol{\alpha}_1, \mathbf{q}_2, \mathbf{q}_3\}$ is obtained as:

$$\begin{aligned} \mathbf{q}_1 &= (-13.0000, 13.0000, -0.0431, -0.0474, 1.0905); \\ \mathbf{q}_{12} &= (2.0000, -2.0000, 0.3384, 0.1659, 0.4957); \\ \mathbf{q}_{123} &= (2.0000, 5.5000, -2.1546, -1.0449, -3.3005). \end{aligned}$$

Let Q be a matrix with rows $\{\mathbf{q}_1, \mathbf{q}_{12}, \mathbf{q}_{123}, \boldsymbol{\alpha}_4, \boldsymbol{\alpha}_5\}$. For any probability vector $\boldsymbol{\alpha}$, denote by $\mathbf{x} = \boldsymbol{\alpha}Q^{-1}(= (x_1, x_2, x_3, x_4, x_5))$. Let $\boldsymbol{\beta} = (x_1, x_2, x_3)$. Since $\boldsymbol{\alpha}\mathbf{e} = 1, \boldsymbol{\alpha}_1\mathbf{e} = \boldsymbol{\alpha}_2\mathbf{e} = \boldsymbol{\alpha}_3\mathbf{e} = 1$, and $\boldsymbol{\alpha}_4\mathbf{e} = \boldsymbol{\alpha}_5\mathbf{e} = 0$, we must have $\boldsymbol{\beta}\mathbf{e} = 1$. Every *PH*-distribution $(\boldsymbol{\alpha}, T)$ has a representation $(\boldsymbol{\beta}, S(15, 2, 2), \mathbf{e})$ of order 3, where $S(15, 2, 2)$ is given in equation (5.5). For $\{\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3, \mathbf{e}_4, \mathbf{e}_5\}$, their corresponding $\boldsymbol{\beta}$ are given as:

$$\begin{aligned} \mathbf{e}_1 : \boldsymbol{\beta} &= (0.0667,\ 0.8000,\ 0.1333); \\ \mathbf{e}_2 : \boldsymbol{\beta} &= (0.1333,\ 0.7333,\ 0.1333); \\ \mathbf{e}_3, \mathbf{e}_4, \mathbf{e}_5 : \boldsymbol{\beta} &= (0.1333,\ 0.8667,\ 0.0000). \end{aligned}$$

Since all the above $\boldsymbol{\beta}$ are nonnegative, *every* *PH*-distribution $(\boldsymbol{\alpha}, T)$ has an ordered Coxian representation of the form $(\boldsymbol{\beta}, S(15, 2, 2))$. Furthermore, if $\boldsymbol{\alpha}$ is independent of $\boldsymbol{\alpha}_3$, then it can be shown that $(\boldsymbol{\alpha}, T)$ has an ordered Coxian representation $(\boldsymbol{\beta}, S(15, 2))$ of order 2. If $\boldsymbol{\alpha}$ is independent of $\boldsymbol{\alpha}_1$, then it can be shown that $(\boldsymbol{\alpha}, T)$ has an Erlang representation $(\boldsymbol{\beta}, S(2, 2))$ of order 2.

6. EXPANDED *PH*-INVARIANT POLYTOPES AND SPECTRAL POLYNOMIAL ALGORITHM

In this section, we establish a connection between *PH*-invariant polytopes and the spectral polynomial algorithms introduced in He and Zhang[17]. The results obtained in this paper give a geometric and probabilistic interpretation to the spectral polynomial method. To begin with, we expand the equation $Q_{1,12,\ldots,1\ldots m}T = S(\boldsymbol{\lambda})Q_{1,12,\ldots,1\ldots m}$ in Theorem 4.1 to obtain

$$\mathbf{q}_{1\ldots n} = \frac{1}{\lambda_m \cdots \lambda_{n+1}} \mathbf{q}_{1\ldots m}(\lambda_m I + T)\ldots(\lambda_{n+1} I + T), \quad 1 \le n \le m-1, \tag{6.1}$$

$\mathbf{q}_1 T = -\lambda_1 \mathbf{q}_1$, and $\mathbf{q}_{1\ldots m}$ is so determined that $\mathbf{q}_{1\ldots n}\mathbf{e} = 1$ for $1 \le n \le m$. The spectrum $\{-\lambda_1, -\lambda_2, \ldots, -\lambda_m\}$ is *not required* to be all real nor to be ordered for the above expansion. Let $W_m = I$, and $W_n = (\lambda_m I + T)\ldots(\lambda_{n+1} I + T)/(\lambda_m \ldots \lambda_{n+1})$, $1 \le n \le m-1$. We have $\mathbf{q}_{1\ldots n} = \mathbf{q}_{1\ldots m} W_n$, $1 \le n \le m$. Then $\mathbf{q}_{1\ldots m}$ must satisfy the equation $\mathbf{q}_{1\ldots m}(W_m\mathbf{e}, W_{m-1}\mathbf{e}, \ldots, W_1\mathbf{e}) = (1, 1, \ldots, 1)$. This shows that the construction of the *PH*-invariant polytope may be accomplished by using the spectral polynomial approach shown in equation (6.1). In fact, the next proposition shows that the two approaches are indeed equivalent if T is *PH*-simple.

Proposition 6.1. *If T is PH-simple (or, equivalently,* $\det(\lambda I - T)$ *is the minimal polynomial of T), then the matrix $(W_m\mathbf{e}, W_{m-1}\mathbf{e}, \ldots, W_1\mathbf{e})$ is invertible and $\mathbf{q}_{1\ldots m}$ can be chosen as $\mathbf{q}_{1\ldots m} = (1, \ldots, 1)(W_m\mathbf{e}, W_{m-1}\mathbf{e}, \ldots, W_1\mathbf{e})^{-1}$. The matrix $Q_{1,12,\ldots,1\ldots m}$ is invertible and is the unique solution to equation $XT = S(\boldsymbol{\lambda})X$ and $X\mathbf{e} = \mathbf{e}$, where X is an $m \times m$ matrix.*

Proof. The equivalence of the *PH*-simplicity of T and the minimality of $\det(\lambda I - T)$ is from O'Cinneide[27]. If $Q_{1,12,\ldots,1\ldots m}$ is not unique, then we have a nonzero solution to $XT = S(\boldsymbol{\lambda})X$ and $X\mathbf{e} = 0$. This is impossible since T is *PH*-simple. By Proposition 2.1, $Q_{1,12,\ldots,1\ldots m}$ is invertible. This completes the proof of Proposition 6.1.

Denote by $P = (Q_{1,12,\ldots,1\ldots m})^{-1}$. Then we have $TP = PS(\boldsymbol{\lambda})$. Note that P can be obtained from the Post-T spectral polynomial developed in

He and Zhang[17]. Geometrically, P is nonnegative (i.e., $S(\boldsymbol{\lambda})$ *PH*-majorizes T) if and only if the probability vector polytope is a subset of the polytope generated by the rows of $Q_{1,12,\ldots,1\ldots m} = P^{-1}$. Based on the above relationships, we propose the following algorithm to construct a Coxian representation for a *PH*-simple T.

A Pre-T spectral polynomial algorithm (for a given *PH*-representation $(\boldsymbol{\alpha}, T)$, we assume that T is *PH*-simple):

Step 1. Find the spectrum $\{-\lambda_1, -\lambda_2, \ldots, -\lambda_m\}$ of T.

Step 2. Compute $\mathbf{q}_{1\ldots m} = (1, \ldots, 1)(W_m\mathbf{e}, W_{m-1}\mathbf{e}, \ldots, W_1\mathbf{e})^{-1}$ and $\{\mathbf{q}_{1\ldots n}, 1 \le n \le m-1\}$ by equation (6.1).

Step 3. Construct a Coxian generator $S(\boldsymbol{\lambda})$ of size m with $\boldsymbol{\lambda} = (\lambda_1, \lambda_2, \ldots, \lambda_m)$. Define a matrix Q with row vectors $\{\mathbf{q}_{1\ldots n}, 1 \le n \le m\}$. Compute $\boldsymbol{\beta} = \boldsymbol{\alpha} Q^{-1}$.

The representation $(\boldsymbol{\beta}, S(\boldsymbol{\lambda}), \mathbf{e})$ found by the above algorithm is a matrix-exponential representation of $(\boldsymbol{\alpha}, T)$. If $\boldsymbol{\beta}$ is nonnegative and $\{-\lambda_1, -\lambda_2, \ldots, -\lambda_m\}$ are all real, then $(\boldsymbol{\beta}, S(\boldsymbol{\lambda}))$ is a Coxian representation. *The main advantage of the spectral polynomial approach is that only information about the spectrum of T is required for the construction of the expanded PH-invariant polytope.*

Note 6.1. To use the above Pre-T spectral polynomial algorithm, T must be *PH*-simple. Therefore, that algorithm cannot replace the algorithms developed in He and Zhang[17].

There is another way to expand the *PH*-invariant polytope defined at the beginning of Section 4. For $S(\boldsymbol{\lambda})$ and $S(\boldsymbol{\mu}, \mathbf{n})$ defined at the beginning of Section 4, by Theorem 3.3 in He and Zhang[17], there exists a nonnegative matrix P with unit row sums such that $S(\boldsymbol{\mu}, \mathbf{n})P = PS(\boldsymbol{\lambda})$.

Proposition 6.2. *Assume that $(\boldsymbol{\beta}, S(\boldsymbol{\mu}, \mathbf{n}))$ is a PH-representation of $(\boldsymbol{\alpha}, T)$. Let $\gamma = \boldsymbol{\beta} P$. Then $(\gamma, S(\boldsymbol{\lambda}))$ is an ordered Coxian representation of $(\boldsymbol{\alpha}, T)$.*

Proof. Since $(\boldsymbol{\beta}, (\boldsymbol{\mu}, \mathbf{n}))$ is a *PH*-representation, $\boldsymbol{\beta}$ is nonnegative. By Theorem 3.3 in He and Zhang[17], P is nonnegative. Thus, $\gamma = \boldsymbol{\beta} P$ is nonnegative. Therefore, $(\gamma, S(\boldsymbol{\lambda}))$ is a *PH*-representation. $(\gamma, S(\boldsymbol{\lambda}))$ and $(\boldsymbol{\alpha}, T)$ have the same distribution since they both have the same distribution as $(\boldsymbol{\beta}, S(\boldsymbol{\mu}, \mathbf{n}))$. This completes the proof of Proposition 6.2.

Note that γ can be nonnegative even if $\boldsymbol{\beta}$ is not, since P is nonnegative. That implies that the above method does expand the set of *PH*-distributions with an ordered Coxian representation.

To end this section, we show graphically that if the eigenvalues in $\boldsymbol{\lambda}$ are ordered differently, the corresponding *PH*-invariant polytopes can be dramatically different.

Example 6.1. We consider the *PH*-generator

$$T = \begin{pmatrix} -5 & 1 & 0.5 \\ 3 & -4 & 0 \\ 1 & 0.1 & -2 \end{pmatrix}. \tag{6.2}$$

The eigenvalues of T are $\{-6.3673, -2.9192, -1.7135\}$ (i.e., $\lambda_1 = 6.3673$, $\lambda_2 = 2.9192$, $\lambda_3 = 1.7135$ with $\lambda_1 > \lambda_2 > \lambda_3$). By using the Pre-$T$ spectral polynomial algorithm, we found the *PH*-invariant polytopes $\text{conv}\{\mathbf{q}_1, \mathbf{q}_{12}, \mathbf{q}_{123}\}$ for $\boldsymbol{\lambda} = (\lambda_1, \lambda_2, \lambda_3)$, which is the triangle with dashed lines in Figure 2. In Figure 2, the triangle with solid lines is the probability vector polytope $\text{conv}\{\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3\}$.

By using the Pre-T spectral polynomial algorithm, we can find the *PH*-invariant polytopes when $\{\lambda_1, \lambda_2, \lambda_3\}$ are arranged in a total of six different orders:

$$(\lambda_1, \lambda_2, \lambda_3) : \text{conv}\{\boldsymbol{\alpha}_1, \mathbf{q}_{12}, \mathbf{q}_{123}\};$$
$$(\lambda_1, \lambda_3, \lambda_2) : \text{conv}\{\boldsymbol{\alpha}_1, \boldsymbol{\alpha}_{13}, \mathbf{q}_{123}\};$$
$$(\lambda_2, \lambda_1, \lambda_3) : \text{conv}\{\boldsymbol{\alpha}_2, \mathbf{q}_{12}, \mathbf{q}_{123}\};$$
$$(\lambda_2, \lambda_3, \lambda_1) : \text{conv}\{\boldsymbol{\alpha}_2, \boldsymbol{\alpha}_{23}, \mathbf{q}_{123}\};$$

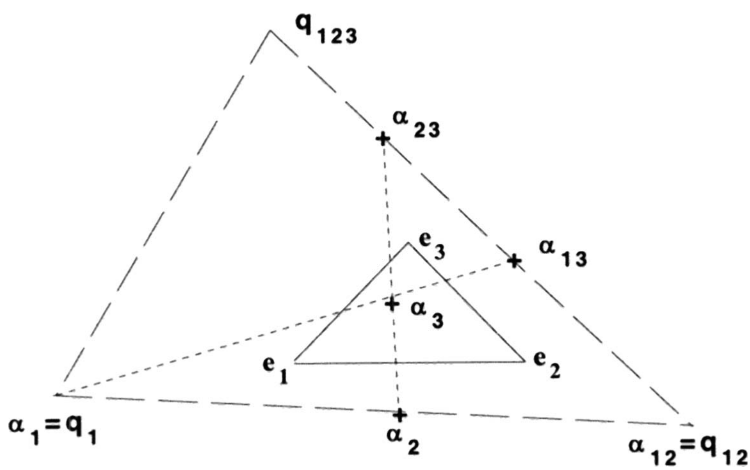

FIGURE 2 *PH*-invariant polytopes for Example 6.1.

$$(\lambda_3, \lambda_1, \lambda_2) : \text{conv}\{\boldsymbol{\alpha}_3, \boldsymbol{\alpha}_{13}, \mathbf{q}_{123}\};$$
$$(\lambda_3, \lambda_2, \lambda_1) : \text{conv}\{\boldsymbol{\alpha}_3, \boldsymbol{\alpha}_{23}, \mathbf{q}_{123}\}.$$

Each of the above six *PH*-invariant polytopes corresponds to a Coxian generator of order 3. The polytope $\text{conv}\{\mathbf{q}_1, \mathbf{q}_{12}, \mathbf{q}_{123}\}$ corresponding to $(\lambda_1, \lambda_2, \lambda_3)$ covers all other *PH*-invariant polytopes. Thus, to find simpler *PH*-representations, we would like to focus on $\text{conv}\{\mathbf{q}_1, \mathbf{q}_{12}, \mathbf{q}_{123}\}$, which explains why we emphasized the case $\{\lambda_1, \lambda_2, \lambda_3\}$ with $\lambda_1 > \lambda_2 > \lambda_3$ in He and Zhang[17] as well as in this paper. Note that $\text{conv}\{\mathbf{q}_1, \mathbf{q}_{12}, \mathbf{q}_{123}\}$ does not always cover the probability vector polytope (Example 3.2), but it does always cover the other *PH*-invariant polytopes with a Coxian generator (Theorem 4.1 and Proposition 4.1 of this paper, and Proposition 5.1 in He and Zhang[17]).

7. SUMMARY

In this paper, we studied the relationship between *PH*-generators, *PH*-invariant polytopes, and Coxian representations. Geometric and probabilistic interpretations to the role played by expanded *PH*-invariant polytopes are provided. A method is developed for the construction of the expanded *PH*-invariant polytope associated with Jordan blocks with real eigenvalues. A generic algorithm for computing a smaller bi-diagonal matrix-exponential representation for any *PH*-representation with real eigenvlues is developed. With a better understanding of the relationship between *PH*-invariant polytopes and *PH*-generators, the next step is to find the minimal bi-diagonal *PH*-representations for *PH*-representations. Results will be reported in He and Zhang[18].

ACKNOWLEDGMENTS

The authors would like to thank an anonymous referee for valuable comments and suggestions on an early version of the paper. This research is supported by a NSERC research grant.

REFERENCES

1. Alfa, A.S.; Chakravarthy, S.R., Eds. *Advances in Matrix Analytic Methods for Stochastic Models*; Notable Publications: New Jersey, 1998.
2. Asmussen, S. *Applied Probability and Queues*; Springer: New York, 2003.
3. Asmussen, S.; Bladt, M. Renewal theory and queueing algorithms for matrix-exponential distributions. In *Proceedings of the First International Conference on Matrix Analytic Methods in Stochastic Models*; Alfa, A.S., Chakravarthy, S., Eds.; Marcel Dekker: New York, 1996; pp. 313–341.
4. Asmussen, S.; Nerman, O.; Olsson, M. Fitting phase-type distributions via the EM algorithm. Scand. J. Statist. **1996**, *23*, 419–441.
5. Blackwell, D.; Koopmans, L. On he identifiability problem for functions of finite Markov chains. Ann. Math. Statist. **1957**, *28*, 1011–1015.

6. Botta, R.F.; Harris, C.M.; Marchal, W.G. Characterizations of generalized hyperexponential distribution functions. Stochastic Models **1987**, *3*, 115–148.
7. Chakravarthy, S.R.; Alfa, A.S., Eds. *Matrix-Analytic Methods in Stochastic Models*; Marcel Dekker: New York, 1996.
8. Commault, C. Linear positive systems and phase-type representations. Positive Systems, Proceedings **2003**, *294*, 281–288.
9. Commault, C.; Chelma, J.P. On dual and minimal phase-type representations. Stochastic Models **1993**, *9*, 421–434.
10. Commault, C.; Chemla, J.P. An invariant of representations of phase-type distributions and some applications. Journal of Applied Probability **1996**, *33*, 368–381.
11. Commault, C.; Mocanu, S. A generic property of phase-type representations. Journal of Applied Probability **2002**, *39*, 775–785.
12. Commault, C.; Mocanu, S. Phase-type distributions and representations: some results and open problems for system theory. International Journal of Control **2003**, *76*, 566–580.
13. Cox, D.R. The analysis of non-Markovian stochastic processes by the inclusion of supplementary variables. Proc. Camb. Phil. Soc. **1955**, *51*, 433–441.
14. Cumani, A. On the canonical representation of Markov processes modeling failure time distributions. Microelectronics and Reliability **1982**, *22*, 583–602.
15. Dehon, M.; Latouche, G. A geometric interpretation of the relations between the exponential and the generalized Erlang distributions. Advances in Applied Probability **1982**, *14*, 885–897.
16. Ewald, G. *Combinatorial Convexity and Algebraic Goemetry*; Springer-Verlag: New York, 1996.
17. He, Q.-M.; Zhang, H. Spectral polynomial algorithms for computing bi-diagonal representations for phase type distributions and matrix-exponential distributions. Stochastic Models *22*, 289–317.
18. He, Q.-M.; Zhang, H. An algorithm for computing minimal Coxian representations (submitted for publication).
19. Ito, H.; Amari, S.I.; Kobayashi, K. Identifiability of hidden Markov information sources and their minimum degrees of freedom. IEEE Trans. Inf. Theory **1992**, *38*, 324–333.
20. Lancaster, P.; Tismenetsky, M. *The Theory of Matrices*; Academic Press: New York, 1985.
21. Latouche, G.; Ramaswami, V. *Introduction to Matrix Analytic Methods in Stochastic Modelling*; ASA & SIAM: Philadelphia, USA, 1999.
22. Latouche, G.; Taylor, P. *Advances in Algorithmic Methods for Stochastic Models*; Notable Publications: New Jersey, 2000.
23. Latouche, G.; Taylor, P. *Matrix-Analytic Methods: Theory and Applications*; World Scientific: New Jersey, 2002.
24. Mocanu, S.; Commault, C. Sparse representations of phase type distributions. Stochastic Models **1999**, *15*, 759–778.
25. Neuts, M.F. Probability distributions of phase type. In *Liber Amicorum Prof. Emeritus H. Florin*; University of Louvain, 1975; pp. 173–206.
26. Neuts, M.F. *Matrix-Geometric Solutions in Stochastic Models – An Algorithmic Approach*; The Johns Hopkins University Press: Baltimore, 1981.
27. O'Cinneide, C.A. On non-uniqueness of representations of phase-type distributions. Stochastic Models **1989**, *5*, 247–259.
28. O'Cinneide, C.A. Characterization of phase-type distributions. Stochastic Models **1990**, *6*, 1–57.
29. O'Cinneide, C.A. Phase-type distributions and invariant polytope. Advances in Applied Probability **1991**, *23*, 515–535.
30. O'Cinneide, C.A. Triangular order of triangular phase-type distributions. Stochastic Models **1993**, *9*, 507–529.
31. O'Cinneide, C.A. Phase-type distributions: open problems and a few properties. Stochastic Models **1999**, *15*, 731–757.
32. Rockafellar, R.T. *Convex Analysis*; Princeton University Press: New Jersey, 1970.
33. Ryden, T. On identifiability and order of continuous-time aggregated markov chains, Markov-modulated Poisson processes, and phase-type distributions. Journal of Applied Probability **1996**, *33*, 640–653.