THE ALGEBRAIC DEGREE OF PHASE-TYPE DISTRIBUTIONS

MARK FACKRELL,* *University of Melbourne*

QI-MING HE,** *Dalhousie University*

PETER TAYLOR,* *University of Melbourne*

HANQIN ZHANG,*** *Chinese Academy of Sciences and NUS Business School*

Abstract

This paper is concerned with properties of the algebraic degree of the Laplace–Stieltjes transform of phase-type (PH) distributions. The main problem of interest is: given a PH generator, how do we find the maximum and the minimum algebraic degrees of all irreducible PH representations with that PH generator? Based on the matrix exponential (ME) order of ME distributions and the spectral polynomial algorithm, a method for computing the algebraic degree of a PH distribution is developed. The maximum algebraic degree is identified explicitly. Using Perron–Frobenius theory of nonnegative matrices, a lower bound and an upper bound on the minimum algebraic degree are found, subject to some conditions. Explicit results are obtained for special cases.

Keywords: Phase-type distribution; rational Laplace–Stieltjes transform; algebraic degree; matrix-analytic method

2010 Mathematics Subject Classification: Primary 60E05; 15A09

1. Introduction

The distribution of the time to absorption in a finite-state continuous-time Markov chain with a single absorbing state is known as a phase-type (PH) distribution. Any PH distribution is completely defined by an initial state probability vector $\boldsymbol{\alpha}$ of order m and an $m \times m$ PH generator $\boldsymbol{T}$. The pair $(\boldsymbol{\alpha}, \boldsymbol{T})$ is said to be an order-m representation for the PH distribution.

Since their introduction by Neuts in 1975, PH distributions have been used in a wide range of stochastic modeling applications in areas as diverse as telecommunications, teletraffic modeling, biostatistics, queueing theory, reliability theory, and health and social care modeling. PH distributions have enjoyed such popularity because they constitute a very versatile class of distributions defined on the nonnegative real numbers that lead to models that are algorithmically tractable. Stochastic models that incorporate PH distributions retain a Markovian structure, but have considerably greater flexibility in modeling individual lifetimes.

A number of performance measures of models involving PH distributions can be written in terms of the Laplace–Stieltjes transform (LST) of the distributions (see Botta *et al.* (1987) and Neuts (1984)). Thus, it is convenient and useful to consider the LST of a PH distribution. The LST of any PH distribution is a rational function, and its algebraic degree is defined to

be the degree of the denominator polynomial when the LST is expressed as the ratio of two coprime polynomials (see O'Cinneide (1990)). In this paper we are interested in finding the algebraic degree for a given PH representation $(\boldsymbol{\alpha}, \boldsymbol{T})$ and specifically in the following question: given a PH generator $\boldsymbol{T}$, how can we choose an $\boldsymbol{\alpha}$ so that the algebraic degree is maximized or minimized? The answer to this question is not simple. To find the minimum algebraic degree, an obvious first suggestion is to choose $\boldsymbol{\alpha}$ to be the (stochastic) left eigenvector that corresponds to the eigenvalue of maximal real part, in which case the algebraic degree will be 1. However, the resultant PH representation may not be irreducible and so some of the phases in the representation $(\boldsymbol{\alpha}, \boldsymbol{T})$ may be superfluous. If we stipulate that the representation that gives the minimal algebraic degree must be irreducible, then the problem becomes much more interesting and challenging.

The algebraic degree has also been studied for matrix exponential (ME) distributions (see, for example, O'Cinneide (1990), van de Liefvoort (1990), Asmussen and Bladt (1997), Commault and Chemla (1996), and He and Zhang (2007)). In He and Zhang (2006), a spectral polynomial method was introduced for studying PH and ME representations. He and Zhang (2007) called the algebraic degree the ME order. They developed an approach using Hankel matrices for computing the ME order of ME distributions. Perron–Frobenius theory has been used extensively in the study of all kinds of stochastic models (see Gantmacher (1959), (1960), Berman and Plemmons (1979), and Seneta (2006)) and can be applied here. We combine these three approaches to obtain results for the algebraic degree, and the maximum and minimum algebraic degrees.

The remainder of the paper is organized as follows. In Section 2, PH distributions and PH representations are introduced, and the algebraic degree is defined. Section 3 gives a number of properties of the algebraic degree and develops an algorithm for computing it. In Section 4, the maximum algebraic degree is identified. It is shown that the poles of the LSTs of all PH distributions with maximum algebraic degree are the same. Section 5 gives a lower bound and an upper bound for the minimum algebraic degree, subject to some conditions. Some exact results are also obtained for the algebraic degree and the corresponding poles.

2. PH distributions, PH representations, and the algebraic degree

We refer the reader to Neuts (1981) and Latouche and Ramaswami (1999) for the material in this section concerning PH distributions and to Gantmacher (1960) for the material on the Jordan canonical form.

A (continuous-time) PH generator $\boldsymbol{T}$ of order m is an invertible $m \times m$ matrix with negative diagonal elements, nonnegative off-diagonal elements, and nonpositive row sums. A *substochastic vector* $\boldsymbol{\alpha}$ of order m is such that $\boldsymbol{\alpha} \geq 0$ and $\boldsymbol{\alpha}\boldsymbol{e} \leq 1$, where $\boldsymbol{e}$ is a column vector of 1s. A vector $\boldsymbol{\alpha}$ is called *stochastic* if it is substochastic and $\boldsymbol{\alpha}\boldsymbol{e} = 1$. The pair $(\boldsymbol{\alpha}, \boldsymbol{T})$ is called a *PH representation* of a *PH distribution*. The distribution function corresponding to $(\boldsymbol{\alpha}, \boldsymbol{T})$ is given by $1 - \boldsymbol{\alpha} \exp(\boldsymbol{T}t)\boldsymbol{e}$, $t \geq 0$. Let $\boldsymbol{T}^0 = -\boldsymbol{T}\boldsymbol{e}$, a nonnegative column vector. A PH representation $(\boldsymbol{\alpha}, \boldsymbol{T})$ is *irreducible* if the matrix $\boldsymbol{T} + \boldsymbol{T}^0\boldsymbol{\alpha}/(\boldsymbol{\alpha}\boldsymbol{e})$ is an irreducible infinitesimal generator of a continuous-time Markov chain with m states.

Let $f^*(s)$ be the LST of the PH distribution with PH representation $(\boldsymbol{\alpha}, \boldsymbol{T})$. For convenience, we also call $f^*(s)$ the LST of $(\boldsymbol{\alpha}, \boldsymbol{T})$. It is well known that $f^*(s)$ is a rational function that can be written in the form

$$f^*(s) = 1 - \boldsymbol{\alpha}\boldsymbol{e} + \boldsymbol{\alpha}(s\boldsymbol{I} - \boldsymbol{T})^{-1}\boldsymbol{T}^0 = \frac{p(s)}{q(s)}, \qquad s \geq 0, \tag{2.1}$$

where $\boldsymbol{I}$ is the identity matrix, and $p(s)$ and $q(s)$ are polynomial functions. If $p(s)$ and $q(s)$ are coprime polynomial functions, then the degree of $q(s)$ is called the *algebraic degree* of the PH distribution, which is denoted by $\kappa(\boldsymbol{\alpha}, \boldsymbol{T})$ in this paper. For convenience, we also call $\kappa(\boldsymbol{\alpha}, \boldsymbol{T})$ the algebraic degree of the PH representation $(\boldsymbol{\alpha}, \boldsymbol{T})$. By (2.1), it is easy to see that $\kappa(\boldsymbol{\alpha}, \boldsymbol{T}) = \kappa(\boldsymbol{\alpha}/(\boldsymbol{\alpha}\boldsymbol{e}), \boldsymbol{T})$ if $\boldsymbol{\alpha}\boldsymbol{e} > 0$. Thus, without loss of generality, we assume that $\boldsymbol{\alpha}$ is stochastic.

Some results on the algebraic degree can be obtained immediately from the Jordan canonical form of the matrix $\boldsymbol{T}$. Denote by $\{-\lambda_1, -\lambda_2, \ldots, -\lambda_m\}$ the eigenvalues of $\boldsymbol{T}$ (counting multiplicity) and by $\{-\mu_1, \ldots, -\mu_N\}$ the distinct eigenvalues of $\boldsymbol{T}$, with $N \leq m$. Throughout the paper, we will order the eigenvalues according to their increasing moduli. Thus, $-\lambda_1 = -\mu_1$ is the maximal (Perron–Frobenius) eigenvalue of $\boldsymbol{T}$, which will be denoted by $\rho(\boldsymbol{T})$. Let N_i be the *algebraic number* of $-\mu_i$ (that is, the multiplicity of $-\mu_i$ as a zero of $\det(s\boldsymbol{I} - \boldsymbol{T})$), let m_i be the *geometric number* of eigenvalue $-\mu_i$ (that is, the number of Jordan blocks corresponding to $-\mu_i$ in the Jordan canonical form of $\boldsymbol{T}$), let $n_{i,j}$ be the size of the jth Jordan block corresponding to $-\mu_i$, and let $\nu_i = \max_{\{1 \leq j \leq m_i\}}\{n_{i,j}\}$ be the *degree* of $-\mu_i$ (that is, the size of the largest Jordan block associated with $-\mu_i$). Then $N_i = \sum_{j=1}^{m_i} n_{i,j}$ and $\sum_{i=1}^{N} \nu_i$ is the degree of the minimal polynomial of $\boldsymbol{T}$, which is the monic polynomial $g(x)$ of the smallest degree satisfying $g(\boldsymbol{T}) = 0$.

Equation (2.1) leads to the following results.

(R1) $1 \leq \kappa(\boldsymbol{\alpha}, \boldsymbol{T}) \leq \sum_{i=1}^{N} \nu_i \leq m$.

(R2) The set of poles of the LST of $(\boldsymbol{\alpha}, \boldsymbol{T})$ is a subset of $\{-\mu_1, \ldots, -\mu_N\}$. The multiplicity of each pole is less than or equal to its degree as an eigenvalue of $\boldsymbol{T}$.

In addition to the above, the following results can be obtained easily.

(R3) For any PH generator $\boldsymbol{T}$ satisfying $\boldsymbol{T}\boldsymbol{e} = -\lambda\boldsymbol{e}$, $\kappa(\boldsymbol{\alpha}, \boldsymbol{T}) = 1$ for any stochastic vector $\boldsymbol{\alpha}$.

(R4) For any PH generator $\boldsymbol{T}$, if the stochastic vector $\boldsymbol{\alpha}$ is an eigenvector corresponding to $\rho(\boldsymbol{T})$, then $\kappa(\boldsymbol{\alpha}, \boldsymbol{T}) = 1$.

In general, however, finding the algebraic degree for a given PH representation is not straightforward. Consider the following examples.

Example 2.1. Consider the PH generator

$$
\boldsymbol{T} = \begin{pmatrix} -\tau_1 & & & & \\ \tau_2 & -\tau_2 & & & \\ & \tau_3 & -\tau_3 & & \\ & & \ddots & \ddots & \\ & & & \tau_m & -\tau_m \end{pmatrix}, \tag{2.2}
$$

where $\tau_i > 0, 1 \leq i \leq m$. Although the matrix $\boldsymbol{T}$ has a simple structure, the algebraic degree associated with $(\boldsymbol{\alpha}, \boldsymbol{T})$ exhibits interesting behavior. If $\boldsymbol{\alpha} = (0, \ldots, 0, 1, 0, \ldots, 0)$, with the ith component equal to 1, it is easy to see that $\kappa(\boldsymbol{\alpha}, \boldsymbol{T}) = i$. However, such PH representations are not irreducible except when $i = m$. If $(\boldsymbol{\alpha}, \boldsymbol{T})$ is required to be irreducible then the algebraic degree of $(\boldsymbol{\alpha}, \boldsymbol{T})$ depends on the actual values of the diagonal elements in $\boldsymbol{T}$. For example, if $\tau_1 > \tau_2 > \cdots > \tau_m$ then $-\tau_m$ is the Perron–Frobenius eigenvalue of $\boldsymbol{T}$ and the Perron–Frobenius eigenvector $\boldsymbol{\alpha}$ can be chosen to be *positive*. Then $(\boldsymbol{\alpha}, \boldsymbol{T})$ is an irreducible PH representation with $\kappa(\boldsymbol{\alpha}, \boldsymbol{T}) = 1$. On the other hand, if $\tau_1 < \tau_2 < \cdots < \tau_m$, by Example 3.1

or Example 5.1, below, $\kappa(\boldsymbol{\alpha}, \boldsymbol{T}) = m$ for any stochastic $\boldsymbol{\alpha}$ for which $(\boldsymbol{\alpha}, \boldsymbol{T})$ is an irreducible PH representation.

Example 2.2. Consider the PH generator

$$
\boldsymbol{T} = \begin{pmatrix}
-3 & 1.5 & 0 & 0 & 0 & 0 \\
2 & -5 & 0 & 0 & 0 & 0 \\
1 & 1 & -3 & 0.5 & 0 & 0 \\
1 & 0 & 2 & -3 & 0 & 0 \\
1 & 0 & 1 & 0 & -10 & 0 \\
1 & 1 & 0 & 1 & 0 & -10
\end{pmatrix}. \tag{2.3}
$$

The eigenvalues of $\boldsymbol{T}$ are $\{-2, -2, -4, -6, -10, -10\}$. The degree ν_1 of the eigenvalue -2 is equal to 2, and the degree ν_4 of the eigenvalue -10 is equal to 1. Thus, $\sum_{i=1}^{N} \nu_i = 5$ and, by (R1), the algebraic degree of any PH representation $(\boldsymbol{\alpha}, \boldsymbol{T})$ must lie between 1 and 5. If irreducibility is required, however, we will see in Example 3.2 that the algebraic degree of $(\boldsymbol{\alpha}, \boldsymbol{T})$ must lie in the range 3 to 5.

3. Properties of the algebraic degree

A triplet $(\boldsymbol{\beta}, \boldsymbol{S}, \boldsymbol{u})$ of order m is called an *ME representation* of an ME if $1 - \boldsymbol{\beta} \exp(\boldsymbol{S}t)\boldsymbol{u}$, $t \geq 0$, is a probability distribution (see Lipsky (1992)). The *ME order* of an ME distribution is defined as the smallest order of all the ME representations of that *ME* distribution. It is easy to see that PH distributions and PH representations are special cases of ME distributions and ME representations, respectively. Thus, every PH distribution has an ME order.

For a PH representation $(\boldsymbol{\alpha}, \boldsymbol{T})$ of order m, (2.1) can be written explicitly as

$$
f^*(s) = \frac{\boldsymbol{\alpha} \operatorname{adj}(s\boldsymbol{I} - \boldsymbol{T})\boldsymbol{T}^0}{\det(s\boldsymbol{I} - \boldsymbol{T})} = \frac{\sum_{i=0}^{\kappa(\boldsymbol{\alpha}, \boldsymbol{T})-1} p_i s^i}{s^{\kappa(\boldsymbol{\alpha}, \boldsymbol{T})} + \sum_{i=0}^{\kappa(\boldsymbol{\alpha}, \boldsymbol{T})-1} q_i s^i} \equiv \frac{p(s)}{q(s)}, \qquad s \geq 0, \tag{3.1}
$$

where $\operatorname{adj}(s\boldsymbol{I} - \boldsymbol{T})$ is the adjoint matrix of $s\boldsymbol{I} - \boldsymbol{T}$, $\det(s\boldsymbol{I} - \boldsymbol{T})$ is the determinant of $s\boldsymbol{I} - \boldsymbol{T}$, $q_0 = \det(-\boldsymbol{T}) \neq 0$, and $p(s)$ and $q(s)$ are coprime polynomials. The third expression in (3.1) leads to an equivalent ME representation $(\boldsymbol{\beta}, \boldsymbol{S}, \boldsymbol{u})$ of the PH representation $(\boldsymbol{\alpha}, \boldsymbol{T})$ (see Asmussen and Bladt (1997)), where

$$
\boldsymbol{\beta} = (p_0, p_1, \ldots, p_{(\boldsymbol{\alpha}, \boldsymbol{T})-1}),
$$

$$
\boldsymbol{S} = \begin{pmatrix}
0 & 1 & & & \\
& 0 & 1 & & \\
& & \ddots & \ddots & \\
& & & 0 & 1 \\
-q_0 & -q_1 & \cdots & \cdots & -q_{\kappa(\boldsymbol{\alpha}, \boldsymbol{T})-1}
\end{pmatrix}, \qquad
\boldsymbol{u} = \begin{pmatrix}
1/q_0 \\
0 \\
\vdots \\
0 \\
0
\end{pmatrix}.
$$

From Theorem 3.1 of He and Zhang (2007), the ME order of an ME distribution can be obtained from any of its ME representations via the expression, for any integer k,

$$
O_{\mathrm{ME}}(\boldsymbol{\alpha}, \boldsymbol{T}) = \max\{n \mid \det(\boldsymbol{\Lambda}_n^{[k]}) \neq 0, \ n \geq 1\},
$$

where $\boldsymbol{\Lambda}_n^{[k]}$ is an $n \times n$ Hankel matrix defined by

$$
\boldsymbol{\Lambda}_n^{[k]} = (a_{i,j}^{[k]}), \quad \text{where} \quad a_{i,j}^{[k]} = \boldsymbol{\alpha} \boldsymbol{T}^{i+j-2+k} \boldsymbol{e} = \boldsymbol{\beta} \boldsymbol{S}^{i+j-2+k} \boldsymbol{u}, \quad 1 \leq i, j \leq n.
$$

Note that the elements of the matrix $\mathbf{\Lambda}_n^{[k]}$ are independent of the ME representation and the matrix $\mathbf{\Lambda}_n^{[k]}$ can be written as

$$\mathbf{\Lambda}_n^{[k]} = \begin{pmatrix} \boldsymbol{\alpha} \\ \boldsymbol{\alpha T} \\ \vdots \\ \boldsymbol{\alpha T}^{n-1} \end{pmatrix} \boldsymbol{T}^k(\boldsymbol{e}, \boldsymbol{Te}, \dots, \boldsymbol{T}^{n-1}\boldsymbol{e}) = \begin{pmatrix} \boldsymbol{\beta} \\ \boldsymbol{\beta S} \\ \vdots \\ \boldsymbol{\beta S}^{n-1} \end{pmatrix} \boldsymbol{S}^k(\boldsymbol{u}, \boldsymbol{Su}, \dots, \boldsymbol{S}^{n-1}\boldsymbol{u}). \quad (3.2)$$

For convenience, we state a known relationship between $\kappa(\boldsymbol{\alpha}, \boldsymbol{T})$ and $O_{\mathrm{ME}}(\boldsymbol{\alpha}, \boldsymbol{T})$ in the following proposition.

Proposition 3.1. *For a PH representation* $(\boldsymbol{\alpha}, \boldsymbol{T}), \kappa(\boldsymbol{\alpha}, \boldsymbol{T}) = O_{\mathrm{ME}}(\boldsymbol{\alpha}, \boldsymbol{T})$. *In addition, we have* $\kappa(\boldsymbol{\alpha}, \boldsymbol{T}) \leq \dim \mathrm{Span}\{\boldsymbol{e}, \boldsymbol{Te}, \boldsymbol{T}^2\boldsymbol{e}, \dots\}$ *and* $\kappa(\boldsymbol{\alpha}, \boldsymbol{T}) \leq \dim \mathrm{Span}\{\boldsymbol{\alpha}, \boldsymbol{\alpha T}, \boldsymbol{\alpha T}^2, \dots\}$.

For a PH generator $\boldsymbol{T}$ and a stochastic vector $\boldsymbol{\alpha}$, define a (nonlinear) mapping $\mathcal{M}_{\boldsymbol{T}} \colon \boldsymbol{\alpha} \to \boldsymbol{\alpha T}^{-1}/(\boldsymbol{\alpha T}^{-1}\boldsymbol{e})$. Note that, since the matrix $-\boldsymbol{T}$ is an M-matrix, all elements of $-\boldsymbol{T}^{-1}$ are nonnegative. It is well known that $-[\boldsymbol{T}^{-1}]_{i,j}$ is the expected sojourn time in state j before absorption, given that the initial state was i, and so $-\boldsymbol{\alpha T}^{-1}\boldsymbol{e}$ is the mean of the PH distribution and is positive. Thus, $\mathcal{M}_{\boldsymbol{T}}(\boldsymbol{\alpha}) = \boldsymbol{\alpha T}^{-1}/(\boldsymbol{\alpha T}^{-1}\boldsymbol{e})$ is a well-defined stochastic vector. Some properties related to the mapping $\mathcal{M}_{\boldsymbol{T}}$ are collected in the following proposition.

Proposition 3.2. (i) *We have* $\kappa(\boldsymbol{\alpha}, \boldsymbol{T}) = \kappa(\mathcal{M}_{\boldsymbol{T}}(\boldsymbol{\alpha}), \boldsymbol{T})$, *that is, the poles of the LST of* $(\mathcal{M}_{\boldsymbol{T}}(\boldsymbol{\alpha}), \boldsymbol{T})$ *(counting multiplicities) are the same as those of* $(\boldsymbol{\alpha}, \boldsymbol{T})$.

(ii) *The PH representation* $(\boldsymbol{\alpha}, \boldsymbol{T})$ *is irreducible if and only if the vector* $\mathcal{M}_{\boldsymbol{T}}(\boldsymbol{\alpha})$ *is positive.*

Proof. (i) Equation (3.2) implies that the ME orders of $(\boldsymbol{\alpha}, \boldsymbol{T})$ and $(\mathcal{M}_{\boldsymbol{T}}(\boldsymbol{\alpha}), \boldsymbol{T})$ are equal. By Proposition 3.1 we must have $\kappa(\boldsymbol{\alpha}, \boldsymbol{T}) = \kappa(\mathcal{M}_{\boldsymbol{T}}(\boldsymbol{\alpha}), \boldsymbol{T})$.

By Section 2, the poles of the LSTs are contained in the set of eigenvalues of $\boldsymbol{T}$ (counting multiplicities). For each pole, we need only show that the multiplicity of that pole as a zero of the function $q(s)$ is the same for $(\boldsymbol{\alpha}, \boldsymbol{T})$ and $(\mathcal{M}_{\boldsymbol{T}}(\boldsymbol{\alpha}), \boldsymbol{T})$. We use the identity

$$(s\boldsymbol{I} - \boldsymbol{T}) \operatorname{adj}(s\boldsymbol{I} - \boldsymbol{T}) = \det(s\boldsymbol{I} - \boldsymbol{T})\boldsymbol{I}$$

to prove the result. Define $\boldsymbol{h}^*(s) = \operatorname{adj}(s\boldsymbol{I} - \boldsymbol{T})\boldsymbol{T}^0$. Then we obtain $(s\boldsymbol{I} - \boldsymbol{T})\boldsymbol{h}^*(s) = \det(s\boldsymbol{I} - \boldsymbol{T})\boldsymbol{T}^0$. Differentiating both sides of the equation with respect to s, we obtain, for any eigenvalue $-\mu$ of $\boldsymbol{T}$,

$$\begin{aligned} (s\boldsymbol{I} - \boldsymbol{T})\boldsymbol{h}^{*(0)}(s)|_{s=-\mu} &= 0, \\ (s\boldsymbol{I} - \boldsymbol{T})\boldsymbol{h}^{*(n)}(s)|_{s=-\mu} &= -n\boldsymbol{h}^{*(n-1)(s)}|_{s=-\mu} \quad \text{if } 1 < n \leq N_{-\mu} - 1, \end{aligned} \quad (3.3)$$

where $\boldsymbol{h}^{*(n)}(s)$ is the nth derivative of $\boldsymbol{h}^*(s)$ and $N_{-\mu}$ is the algebraic number of the eigenvalue $-\mu$. Multiplying on the left by $\boldsymbol{T}^{-1}$ in (3.3) yields

$$\begin{aligned} s\boldsymbol{T}^{-1}\boldsymbol{h}^{*(0)}(s)|_{s=-\mu} &= \boldsymbol{h}^{*(0)}(s)|_{s=-\mu}, \\ s\boldsymbol{T}^{-1}\boldsymbol{h}^{*(n)}(s)|_{s=-\mu} &= \boldsymbol{h}^{*(n)}(s)|_{s=-\mu} - n\boldsymbol{T}^{-1}\boldsymbol{h}^{*(n-1)}(s)|_{s=-\mu} \quad \text{if } 0 < n \leq N_{-\mu} - 1. \end{aligned} \quad (3.4)$$

By (3.4), for $\mu \neq 0$ and $0 \leq k \leq N_{-\mu} - 1$, it is easy to show that $\boldsymbol{\alpha T}^{-1}\boldsymbol{h}^{*(n)}(s)|_{s=-\mu} = 0$ for $0 \leq n \leq k$ if and only if $\boldsymbol{\alpha h}^{*(n)}(s)|_{s=-\mu} = 0$ for $0 \leq n \leq k$. This implies that the multiplicities of $-\mu$ as zeros of $\boldsymbol{\alpha h}^*(s)$ and $\boldsymbol{\alpha T}^{-1}\boldsymbol{h}^*(s)$ are equal. By (3.1), the multiplicities

of $-\mu$ as poles of the LSTs of $(\boldsymbol{\alpha}, \boldsymbol{T})$ and $(\mathcal{M}_{\boldsymbol{T}}(\boldsymbol{\alpha}), \boldsymbol{T})$ are equal. Consequently, the LSTs of $(\boldsymbol{\alpha}, \boldsymbol{T})$ and $(\mathcal{M}_{\boldsymbol{T}}(\boldsymbol{\alpha}), \boldsymbol{T})$ have the same set of poles, each with the same multiplicity. This completes the proof of part (i).

(ii) First note that $\mathcal{M}_{\boldsymbol{T}}(\boldsymbol{\alpha})(\boldsymbol{T} + \boldsymbol{T}^0\boldsymbol{\alpha}) = 0$, that is, $\mathcal{M}_{\boldsymbol{T}}(\boldsymbol{\alpha})$ is the Perron–Frobenius eigenvector of the infinitesimal generator $\boldsymbol{T} + \boldsymbol{T}^0\boldsymbol{\alpha}$. If $(\boldsymbol{\alpha}, \boldsymbol{T})$ is an irreducible representation then $\boldsymbol{T} + \boldsymbol{T}^0\boldsymbol{\alpha}$ is an irreducible matrix, and, by Perron–Frobenius theory, $\mathcal{M}_{\boldsymbol{T}}(\boldsymbol{\alpha})$ is positive.

Now, we assume that $\mathcal{M}_{\boldsymbol{T}}(\boldsymbol{\alpha})$ is positive. If $(\boldsymbol{\alpha}, \boldsymbol{T})$ is reducible then $\boldsymbol{T} + \boldsymbol{T}^0\boldsymbol{\alpha}$ must have the structure

$$\begin{pmatrix} \boldsymbol{A}_{1,1} & \boldsymbol{A}_{1,} \\ 0 & \boldsymbol{A}_{2,2} \end{pmatrix},$$

where $\boldsymbol{A}_{1,1}$ is a $j \times j$ matrix with $1 \le j < m$. Since $\boldsymbol{T}^0\boldsymbol{\alpha}$ is nonnegative and all off-diagonal elements of $\boldsymbol{T}$ are nonnegative, it is clear that $\boldsymbol{T}$ must also have the same structure, so that $\boldsymbol{T}$ can be written in the form

$$\begin{pmatrix} \boldsymbol{T}_{1,1} & \boldsymbol{T}_{1,2} \\ 0 & \boldsymbol{T}_{2,2} \end{pmatrix}.$$

From the structure of the matrix $\boldsymbol{T}^0\boldsymbol{\alpha}$, it is clear that we must have either $\alpha_1 = \alpha_2 = \cdots = \alpha_j = 0$ or $t^0_{j+1} = t^0_{j+2} = \cdots = t^0_m = 0$, where α_n is the nth element of $\boldsymbol{\alpha}$ and t^0_n is the nth element of $\boldsymbol{T}^0$. If $t^0_{j+1} = t^0_{j+2} = \cdots = t^0_m = 0$, we must have $\boldsymbol{T}_{2,2} = \boldsymbol{A}_{2,2}$. Since $\boldsymbol{T} + \boldsymbol{T}^0\boldsymbol{\alpha}$ is an infinitesimal generator, $\boldsymbol{T}_{2,2}$ must be an infinitesimal generator, which is a singular matrix. This implies that the matrix $\boldsymbol{T}$ is not invertible, which contradicts the fact that $\boldsymbol{T}$ is a PH generator. If $\alpha_1 = \alpha_2 = \cdots = \alpha_j = 0$ then the relationship $\mathcal{M}_{\boldsymbol{T}}(\boldsymbol{\alpha})\boldsymbol{T} = (\mathcal{M}_{\boldsymbol{T}}(\boldsymbol{\alpha})\boldsymbol{T}\boldsymbol{e})\boldsymbol{\alpha}$ leads to

$$((\mathcal{M}_{\boldsymbol{T}}(\boldsymbol{\alpha}))_1, (\mathcal{M}_{\boldsymbol{T}}(\boldsymbol{\alpha}))_2, \ldots, (\mathcal{M}_{\boldsymbol{T}}(\boldsymbol{\alpha}))_j)\boldsymbol{T}_{1,1} = (\alpha_1, \alpha_2, \ldots, \alpha_j) = 0.$$

Since $\boldsymbol{T}_{1,1}$ is invertible, this leads to $((\mathcal{M}_{\boldsymbol{T}}(\boldsymbol{\alpha}))_1, (\mathcal{M}_{\boldsymbol{T}}(\boldsymbol{\alpha}))_2, \ldots, (\mathcal{M}_{\boldsymbol{T}}(\boldsymbol{\alpha}))_j) = 0$, which contradicts the assumption that $\mathcal{M}_{\boldsymbol{T}}(\boldsymbol{\alpha})$ is positive. This completes the proof of Proposition 3.2.

There are some interesting implications of Proposition 3.2. Let the stochastic vector $\boldsymbol{\theta}$ be an eigenvector of $\boldsymbol{T}$ corresponding to $\rho(\boldsymbol{T})$. Then $(\boldsymbol{\theta}, \boldsymbol{T})$ represents an exponential distribution and $\kappa(\boldsymbol{\theta}, \boldsymbol{T}) = 1$. For any stochastic vector $\boldsymbol{\alpha}[0]$, define $\boldsymbol{\alpha}[n] = \mathcal{M}_{\boldsymbol{T}}(\boldsymbol{\alpha}[n-1])$, $n > 0$. By Proposition 3.2, $\kappa(\boldsymbol{\alpha}[n], \boldsymbol{T}) = \kappa(\boldsymbol{\alpha}[0], \boldsymbol{T})$ and the LST of $(\boldsymbol{\alpha}[n], \boldsymbol{T})$ has the same poles for all $n \ge 0$. However, it is well known that, if $\boldsymbol{T}$ is irreducible, $\{\boldsymbol{\alpha}[n], \ n \ge 0\}$ converges pointwise to $\boldsymbol{\theta}$. The implication is that there are PH distributions of any possible algebraic degree that are arbitrarily close to the exponential distribution associated with $\rho(\boldsymbol{T})$. From a topological point of view, in the m-dimensional space parameterized by $\{\alpha_1, \ldots, \alpha_m\}$, distributions of lower degree fall on a finite number of hyperplanes. Thus, the exponential distribution is arbitrarily close to distributions of maximal algebraic degree. Our result extends this observation to distributions of any possible algebraic degree. In addition, the poles of the LSTs of these PH distributions can be quite different from that of the exponential distribution.

Combining parts (i) and (ii) of Proposition 3.2, we obtain the following useful result.

Corollary 3.1. *For any n between 1 and m, if there exists a stochastic vector $\boldsymbol{\alpha}$ such that $\kappa(\boldsymbol{\alpha}, \boldsymbol{T}) = n$ and $(\boldsymbol{\alpha}, \boldsymbol{T})$ is an irreducible PH representation, then there exists a positive $\boldsymbol{\theta}$ (e.g. $\boldsymbol{\theta} = \mathcal{M}_{\boldsymbol{T}}(\boldsymbol{\alpha})$) such that $\kappa(\boldsymbol{\theta}, \boldsymbol{T}) = n$.*

Next, we adopt a method introduced in He and Zhang (2008) for finding $\kappa(\boldsymbol{\alpha}, \boldsymbol{T})$. This is performed in three steps. To begin with, we establish a relationship between the algebraic

degree and a linear space generated by the matrix T and e, which leads to an algorithm for computing the algebraic degree. Define

$$\chi(T) = \dim \text{Span}\{e, Te, T^2e, \dots\}. \tag{3.5}$$

By Proposition 3.1 we know that $\kappa(\alpha, T) \leq \chi(T)$. Definition (3.5) implies that $\chi(T) \leq \nu_1 + \cdots + \nu_N$, the degree of the minimal polynomial function of T. Since T is invertible, it is easy to show that the set of vectors $\{Te, T^2e, T^3e, \dots, T^{\chi(T)}e\}$ is linearly independent, while the set $\{e, Te, T^2e, T^3e, \dots, T^{\chi(T)}e\}$ is linearly dependent. This implies that there exists $\{a_0, a_1, \dots, a_{\chi(T)}\}$ such that $a_{\chi(T)} = 1$, $a_0 \neq 0$, and $\sum_{n=0}^{\chi(T)} a_n T^n e = 0$. Define

$$g_T(x) = \sum_{n=0}^{\chi(T)} a_n x^n.$$

Let $-\gamma = \{-\gamma_1, \dots, -\gamma_{\chi(T)}\}$ contain the $\chi(T)$ roots of the polynomial $g_T(x)$ (counting multiplicities), and let the distinct roots of $g_T(x)$ be denoted by $\{-\delta_1, \dots, -\delta_M\}$, with δ_j repeating itself $m_{g,j}$ times in $\{-\gamma_1, \dots, -\gamma_{\chi(T)}\}$. Then $m_{g,1} + \cdots + m_{g,M} = \chi(T)$ and

$$g_T(T)e = \left(\sum_{n=0}^{\chi(T)} a_n T^n\right)e = \left(\prod_{n=1}^{\chi(T)} (\gamma_n I + T)\right)e = 0. \tag{3.6}$$

Using the Jordan canonical form of T and (3.6), it can be shown that $\{-\gamma_1, \dots, -\gamma_{\chi(T)}\}$ is a subset of the eigenvalues of T.

Second, we use the spectral polynomial algorithm (SPA) (see He and Zhang (2006)) to find an equivalent bidiagonal ME representation of the order $\chi(T)$ for the PH representation (α, T). For $\{-\gamma_1, \dots, -\gamma_{\chi(T)}\}$, define

$$p_1 = -\frac{Te}{\gamma_1}, \qquad p_n = \frac{(\gamma_{n-1}I + T)p_{n-1}}{\gamma_n}, \quad 2 \leq n \leq \chi(T), \tag{3.7}$$
$$p_{\chi(T)+1} = (\gamma_{\chi(T)}I + T)p_{\chi(T)}.$$

Let $P = (p_1, p_2, \dots, p_{\chi(T)})$, which is an $m \times \chi(T)$ matrix. Equation (3.6) implies that

$$\left(\prod_{n=1}^{\chi(T)} \gamma_n\right) p_{\chi(T)+1} = -T g_T(T)e = 0.$$

Then (3.7) can be rewritten as $TP = PS(\gamma, \chi(T))$, where

$$S(\gamma, \chi(T)) = \begin{pmatrix} -\gamma_1 & 0 & \cdots & & \cdots & 0 \\ \gamma_2 & -\gamma_2 & \ddots & & \ddots & \vdots \\ 0 & \ddots & \ddots & & \ddots & \vdots \\ \vdots & \ddots & & \gamma_{\chi(T)-1} & -\gamma_{\chi(T)-1} & 0 \\ 0 & \cdots & & 0 & \gamma_{\chi(T)} & -\gamma_{\chi(T)} \end{pmatrix}. \tag{3.8}$$

Since $p_{\chi(T)+1} = 0$ and T is invertible, it can be shown that $Pe = e$ (see Proposition 2.1 and Propositions 3.1 of He and Zhang (2006)). Then $(\alpha P, S(\gamma, \chi(T)), e)$ represents the same

PH distribution as $(\boldsymbol{\alpha}, \boldsymbol{T})$, that is, $1 - \boldsymbol{\alpha} \boldsymbol{P} \exp(S(\boldsymbol{\gamma}, \chi(\boldsymbol{T}))t)\boldsymbol{e} = 1 - \boldsymbol{\alpha} \exp(\boldsymbol{T}t)\boldsymbol{e}$, $t \geq 0$. Thus, $(\boldsymbol{\alpha} \boldsymbol{P}, S(\boldsymbol{\gamma}, \chi(\boldsymbol{T})), \boldsymbol{e})$ is an ME representation of order $\chi(\boldsymbol{T})$. Since $a_0 \neq 0$, none of the roots $\{-\gamma_1, \ldots, -\gamma_{\chi(\boldsymbol{T})}\}$ is zero. Since the vectors $\{\boldsymbol{T}\boldsymbol{e}, \boldsymbol{T}^2\boldsymbol{e}, \boldsymbol{T}^3\boldsymbol{e}, \ldots, \boldsymbol{T}^{\chi(\boldsymbol{T})}\boldsymbol{e}\}$ are linearly independent, the vectors $\{\boldsymbol{p}_1, \boldsymbol{p}_2, \ldots, \boldsymbol{p}_{\chi(\boldsymbol{T})}\}$ are also linearly independent. Therefore, the rank of $\boldsymbol{P}$ must be $\chi(\boldsymbol{T})$.

Define

$$\boldsymbol{J}(\boldsymbol{\gamma}) = \begin{pmatrix} \boldsymbol{J}(\delta_1) & & \\ & \ddots & \\ & & \boldsymbol{J}(\delta_M) \end{pmatrix}, \qquad \boldsymbol{J}(\delta_n) = \begin{pmatrix} -\delta_n & & & \\ \delta_n & \ddots & & \\ & \ddots & \ddots & \\ & & \delta_n & -\delta_n \end{pmatrix}_{m_{g,n} \times m_{g,n}}$$

for $1 \leq n \leq M$. Applying the spectral polynomial algorithm to $\boldsymbol{J}(\boldsymbol{\gamma})$, we obtain $\boldsymbol{J}(\boldsymbol{\gamma})\boldsymbol{Q} = \boldsymbol{Q}S(\boldsymbol{\gamma}, \chi(\boldsymbol{T}))$ and $\boldsymbol{Q}\boldsymbol{e} = \boldsymbol{e}$, where the columns $\{\boldsymbol{q}_1, \boldsymbol{q}_2, \ldots, \boldsymbol{q}_{\chi(\boldsymbol{T})}\}$ of $\boldsymbol{Q}$ are defined by

$$\boldsymbol{q}_1 = -\frac{\boldsymbol{J}(\boldsymbol{\gamma})\boldsymbol{e}}{\gamma_1}, \qquad \boldsymbol{q}_n = \frac{(\gamma_{n-1}\boldsymbol{I} + \boldsymbol{J}(\boldsymbol{\gamma}))\boldsymbol{q}_{n-1}}{\gamma_n}, \quad 2 \leq n \leq \chi(\boldsymbol{T}),$$

and all matrices are of order $\chi(\boldsymbol{T})$. By direct calculations, it can be shown that the matrix $\boldsymbol{Q}$ is invertible. Then we have $\boldsymbol{Q}^{-1}\boldsymbol{J}(\boldsymbol{\gamma}) = S(\boldsymbol{\gamma}, \chi(\boldsymbol{T}))\boldsymbol{Q}^{-1}$ and $\boldsymbol{Q}^{-1}\boldsymbol{e} = \boldsymbol{e}$. Together with $\boldsymbol{T}\boldsymbol{P} = \boldsymbol{P}S(\boldsymbol{\gamma}, \chi(\boldsymbol{T}))$ and $\boldsymbol{P}\boldsymbol{e} = \boldsymbol{e}$, we obtain

$$\boldsymbol{T}\boldsymbol{P}\boldsymbol{Q}^{-1} = \boldsymbol{P}\boldsymbol{Q}^{-1}\boldsymbol{J}(\boldsymbol{\gamma}), \qquad \boldsymbol{P}\boldsymbol{Q}^{-1}\boldsymbol{e} = \boldsymbol{e}.$$

Thus, for any stochastic vector $\boldsymbol{\alpha}$, the PH representation $(\boldsymbol{\alpha}, \boldsymbol{T})$ is equivalent to

$$(\boldsymbol{\alpha} \boldsymbol{P}\boldsymbol{Q}^{-1}, \boldsymbol{J}(\boldsymbol{\gamma}), \boldsymbol{e}).$$

It is easy to see that $m_{g,j}$ is less than or equal to the degree of $-\delta_j$ as an eigenvalue of $\boldsymbol{T}$.

Let

$$\boldsymbol{\beta} = \boldsymbol{\alpha} \boldsymbol{P}\boldsymbol{Q}^{-1} = (\boldsymbol{\beta}(1), \ldots, \boldsymbol{\beta}(M)),$$

where $\boldsymbol{\beta}(j) = (\boldsymbol{\beta}(j, 1), \ldots, \boldsymbol{\beta}(j, m_{g,j}))$. By routine calculations, the LST of the PH distribution $(\boldsymbol{\alpha}, \boldsymbol{T})$ can be expressed in the form

$$f^*(s) = \boldsymbol{\beta}(s\boldsymbol{I} - \boldsymbol{J}(\boldsymbol{\gamma}))^{-1}(-\boldsymbol{J}(\boldsymbol{\gamma}))\boldsymbol{e} = \sum_{j=1}^{M} \sum_{n=1}^{m_{g,j}} \beta(j, n) \left(\frac{\delta_j}{s + \delta_j}\right)^n. \tag{3.9}$$

Finally, we find $\kappa(\boldsymbol{\alpha}, \boldsymbol{T})$. Define, for $1 \leq j \leq M$,

$$\xi_j(\boldsymbol{\alpha}, \boldsymbol{T}) = \begin{cases} 0 & \text{if the set}\{i : \beta(j, i) \neq 0, \ 1 \leq i \leq m_{g,j}\} \\ & \text{is empty,} \\ \max\{i : \beta(j, i) \neq 0, \ 1 \leq i \leq m_{g,j}\} & \text{otherwise.} \end{cases}$$

The above analysis leads to the following results.

Proposition 3.3. *For a PH representation $(\boldsymbol{\alpha}, \boldsymbol{T})$, the following statements hold.*

(i) $\kappa(\boldsymbol{\alpha}, \boldsymbol{T}) = \xi_1(\boldsymbol{\alpha}, \boldsymbol{T}) + \xi_2(\boldsymbol{\alpha}, \boldsymbol{T}) + \cdots + \xi_M(\boldsymbol{\alpha}, \boldsymbol{T}).$

(ii) *The set of poles of the LST of $(\boldsymbol{\alpha}, \boldsymbol{T})$ is a subset of $\{-\gamma_1, \ldots, -\gamma_{\chi(\boldsymbol{T})}\}$, which is a subset of the set of eigenvalues of $\boldsymbol{T}$.*

(iii) *The multiplicity of the pole $-\delta_j$ of the LST of $(\boldsymbol{\alpha}, \boldsymbol{T})$ is less than or equal to $m_{g,j}$, $1 \leq j \leq M$.*

Proof. Part (i) follows from an argument similar to the proof of Lemma 3 of He and Zhang (2008), noting that $\{-\delta_1, \ldots, -\delta_M\}$ are distinct, and using (3.9). Parts (ii) and (iii) can be obtained from the structure of the matrix $\boldsymbol{J}(\boldsymbol{\gamma})$. This completes the proof.

Example 3.1. (*Example 2.1 continued.*) For the generator $\boldsymbol{T}$ defined in (2.2), it is easy to verify that $\chi(\boldsymbol{T}) = m$. Proposition 3.3 shows that, without the irreducibility requirement, the algebraic degree of $(\boldsymbol{\alpha}, \boldsymbol{T})$ can be any integer from 1 to m. If irreducibility is required, the following two cases can be solved.

If $\tau_1 > \tau_2 > \cdots > \tau_m$, we set $\gamma_i = \tau_i$, $1 \leq i \leq m$. By (3.7), it can be verified that $\boldsymbol{P} = \boldsymbol{I}$ and $\boldsymbol{Q}$ is a lower triangular matrix with all the elements on the diagonal and in the lower triangular part positive. Since $\boldsymbol{\beta} = \boldsymbol{\alpha} \boldsymbol{Q}^{-1}$, we obtain $\boldsymbol{\alpha} = \boldsymbol{\beta} \boldsymbol{Q}$. For any $1 \leq i \leq m$, choose a stochastic vector $\boldsymbol{\beta}$ with exactly i positive elements. Then $(\boldsymbol{\beta}, \boldsymbol{J}(\boldsymbol{\gamma}), \boldsymbol{e})$ represents a mixture of i exponential distributions. If $\boldsymbol{\beta}_m > 0$ then $\alpha_m > 0$, which implies that $(\boldsymbol{\alpha}, \boldsymbol{T})$ is irreducible and $\kappa(\boldsymbol{\alpha}, \boldsymbol{T}) = i$. Thus, with the irreducibility requirement, the algebraic degree of $(\boldsymbol{\alpha}, \boldsymbol{T})$ can be any integer from 1 to m.

If $\tau_1 < \tau_2 < \cdots < \tau_m$, we set $\gamma_i = \tau_{m-i+1}$, $1 \leq i \leq m$. By routine calculations, it can be shown that $\boldsymbol{P} \boldsymbol{Q}^{-1}$ is a lower triangular matrix with elements given by

$$(\boldsymbol{P} \boldsymbol{Q}^{-1})_{i,j} = \begin{cases} 0 & \text{if } i < j, \\ 1 & \text{if } i = j = 1, \\ \displaystyle\prod_{\{n=1: \, n \neq j\}}^{i} \frac{\tau_n}{\tau_n - \tau_j} & \text{if } i \geq 2, i \geq j. \end{cases}$$

Thus, the nonzero elements in each column of $\boldsymbol{P} \boldsymbol{Q}^{-1}$ have the same sign and all elements in the last row of $\boldsymbol{P} \boldsymbol{Q}^{-1}$ are nonzero. Consequently, all elements of $\boldsymbol{\beta} = \boldsymbol{\alpha} \boldsymbol{P} \boldsymbol{Q}^{-1}$ are nonzero if and only if $\alpha_m > 0$. Note that $(\boldsymbol{\alpha}, \boldsymbol{T})$ is irreducible if and only if $\alpha_m > 0$. Therefore, $\kappa(\boldsymbol{\alpha}, \boldsymbol{T}) = m$ if and only if $(\boldsymbol{\alpha}, \boldsymbol{T})$ is irreducible.

Example 3.2. (*Example 2.2 continued.*) For the PH generator $\boldsymbol{T}$ defined in (2.3), by applying Proposition 3.3, we conclude that $\chi(\boldsymbol{T}) = 5$, $(-\gamma_1, \ldots, -\gamma_{\chi(\boldsymbol{T})}) = (-2, -2, -4, -6, -10)$, $M = 4$, $(-\delta_1, \ldots, -\delta_M) = (-2, -4, -6, -10)$, and

$$\boldsymbol{P} \boldsymbol{Q}^{-1} = \begin{pmatrix} 1.1250 & 0 & 0 & -0.1250 & 0 \\ 0.7500 & 0 & 0 & 0.2500 & 0 \\ 0.4766 & 0.6094 & -0.0312 & -0.0547 & 0 \\ -0.3594 & 1.2187 & 0.0625 & 0.0781 & 0 \\ 0.1812 & 0.0762 & -0.0052 & -0.0449 & 0.7928 \\ 0.1514 & 0.1523 & 0.0104 & 0.0508 & 0.6351 \end{pmatrix}.$$

The ordering of the rows is the same as the ordering of the states in the original matrix $\boldsymbol{T}$, while the ordering of the columns is determined by the order in which the roots of $g_{\boldsymbol{T}}(x)$ were inserted into $\boldsymbol{J}(\boldsymbol{\gamma})$. The structure of $\boldsymbol{T}$ implies that $(\boldsymbol{\alpha}, \boldsymbol{T})$ is irreducible if and only if α_5 and α_6 are both positive. Since the second column (corresponding to the second occurrence of -2

as a root of $g_T(x)$) and the last column (corresponding to the root -10) of the matrix PQ^{-1} are nonnegative with positive fifth and sixth components, it is clear that $\boldsymbol{\beta}(1, 2)$ and $\boldsymbol{\beta}(4, 1)$ are nonzero for any irreducible $(\boldsymbol{\alpha}, T)$. Therefore, by Proposition 3.3, the algebraic degree of any irreducible $(\boldsymbol{\alpha}, T)$ is greater than or equal to 3, and there must be at least three poles $\{-2, -2, -10\}$. In fact, algebraic degrees of three, four, and five are possible, as the following examples illustrate.

- If $\boldsymbol{\alpha} = (0.1704, 0.1185, 0.237, 0.1185, 0.237, 0.1185)$ then $(\boldsymbol{\alpha}, T)$ is irreducible,

$$\boldsymbol{\beta} = (0.4118, 0.325, 0, 0, 0.2632), \qquad \kappa(\boldsymbol{\alpha}, T) = 3,$$

  and the poles are $\{-2, -2, -10\}$.

- If $\boldsymbol{\alpha} = (0, 0, 0, 0, \frac{2}{3}, \frac{1}{3})$ then $(\boldsymbol{\alpha}, T)$ is irreducible,

$$\boldsymbol{\beta} = (0.1713, 0.1016, 0, -0.013, 0.7402), \qquad \kappa(\boldsymbol{\alpha}, T) = 4,$$

  and the poles are $\{-2, -2, -6, -10\}$.

- If $\boldsymbol{\alpha} = (0, 0, 0, 0, 0.5, 0.5)$ then $(\boldsymbol{\alpha}, T)$ is irreducible,

$$\boldsymbol{\beta} = (0.1663, 0.1142, 0.0026, 0.0029, 0.7139), \qquad \kappa(\boldsymbol{\alpha}, T) = 5,$$

  and the poles are $\{-2, -2, -4, -6, -10\}$.

4. Maximum algebraic degree

For a given PH generator T, define

$$\kappa_{\max}(T) = \max\{\kappa(\boldsymbol{\alpha}, T) \colon (\boldsymbol{\alpha}, T) \text{ is an irreducible PH representation, } \boldsymbol{\alpha}e > 0\}.$$

The following theorem characterizes $\kappa_{\max}(T)$, and the PH distributions at which $\kappa_{\max}(T)$ is attained.

Theorem 4.1. *For any PH generator T, we have $\kappa_{\max}(T) = \chi(T)$. In addition, the poles of the LST of any PH distribution $(\boldsymbol{\alpha}, T)$ with maximum algebraic degree are $\{-\gamma_1, \ldots, -\gamma_{\chi(T)}\}$.*

Proof. By Proposition 3.1 we have $\kappa_{\max}(T) \leq \chi(T)$. Thus, we need only prove that $\kappa_{\max}(T = \chi(T)$. First, we mention two observations: (i) if every element of $\boldsymbol{\alpha}$ is positive then $(\boldsymbol{\alpha}, T)$ is an irreducible PH representation; (ii) if $\boldsymbol{\beta}(j, m_{g,j}) \neq 0$ for all $1 \leq j \leq M$ then the algebraic degree of $(\boldsymbol{\beta}, J(\gamma), e)$ is $\chi(T)$. Define, for $1 \leq j \leq M$,

$$\Omega_j = \{x \colon x \in \mathbb{R}^m, (xPQ^{-1})_{m_{g,1}+\cdots+m_{g,j}} = 0\},$$

where $\mathbb{R}^m$ is the Euclidean space of dimension m. Since the rank of $P(\chi(T))$ is $\chi(T)$ and Q is invertible, the rank of PQ^{-1} is $\chi(T)$. Since PQ^{-1} is an $m \times \chi(T)$ matrix, then PQ^{-1} has no zero column. Thus, the dimension of Ω_j is $m - 1$, $1 \leq j \leq M$. Define $\Omega = \bigcup_{j=1}^M \Omega_j$. The set $\mathbb{R}^m - \Omega$ is dense in $\mathbb{R}^m$. Thus, there exists $\boldsymbol{\alpha}$ with all elements being positive such that $\boldsymbol{\alpha}PQ^{-1} \notin \Omega$, i.e. it has the property that $\boldsymbol{\beta}(j, m_{g,j}) \neq 0$ for $1 \leq j \leq M$. Therefore, the algebraic degree $\kappa(\boldsymbol{\alpha}, T)$ is equal to $\chi(T)$. This completes the proof.

O'Cinneide (1989) defined a PH generator T to be *PH simple* if, for any two different stochastic vectors $\boldsymbol{\alpha}$ and $\boldsymbol{\beta}$, the PH distributions represented by $(\boldsymbol{\alpha}, T)$ and $(\boldsymbol{\beta}, T)$ are different.

By Theorem 1 of O'Cinneide (1989), T is PH simple if and only if $\chi(T) = m$. Consequently, if the PH generator T is PH simple then we have $\kappa_{\max}(T) = \chi(T) = m$.

For Example 2.1, by applying Theorem 4.1, we find that $\kappa_{\max}(T) = \chi(T) = m$ for the PH generator T given in (2.2), and that the corresponding poles are $\{-\tau_1, -\tau_2, \ldots, -\tau_m\}$. For Example 2.2, we find that $\kappa_{\max}(T) = \chi(T) = 5$ for the PH generator T given in (2.3), and that the corresponding poles are $\{-2, -2, -4, -6, -10\}$.

5. Minimum algebraic degree

For a given PH generator T, define

$$\kappa_{\min}(T) = \min\{\kappa(\boldsymbol{\alpha}, T) \colon (\boldsymbol{\alpha}, T) \text{ is an irreducible PH representation, } \boldsymbol{\alpha}e > 0\}.$$

Unlike $\kappa_{\max}(T)$, $\kappa_{\min}(T)$ has much to do with the special structure of T. For instance, there are explicit results for the following special cases.

- There exists λ such that $Te = -\lambda e$. Then $\kappa_{\min}(T) = 1$.

- There is no λ such that $Te = -\lambda e$, but T is irreducible. Then $\kappa_{\min}(T) = 1$.

If there is no λ such that $Te = -\lambda e$, and T is reducible, the problem is complicated. In the rest of this section, under some conditions, we give a lower bound and an upper bound on $\kappa_{\min}(T)$.

First, we need to introduce a number of concepts. We use the notation from Section 2.3 of Berman and Plemmons (1979). Assume that A is a nonnegative matrix of order m and write $A^n = (a_{i,j}^{(n)})$. We say that state j is *accessible* from state i if $a_{i,j}^{(n)} > 0$ for some n. A *class* of states consists of all states for which any two states are accessible from each other. Class φ is accessible from class ϕ if any state in ϕ has access to any state in φ. A class is *initial* if no other class has access to it. Class φ is *basic* if $\rho(A[\varphi]) = \rho(A)$, where $A[\varphi]$ is the submatrix of A based on the states in φ, and *nonbasic* if $\rho(A[\varphi]) < \rho(A)$.

A collection of basic classes $\{\varphi_1, \varphi_2, \ldots, \varphi_n\}$ is a *chain* from φ_1 to φ_n if φ_k has access to φ_{k+1}, $k = 1, 2, \ldots, n - 1$. The *length* of a chain is the number of basic classes it contains. A class φ has access to a class ϕ in j *steps* if the length of the longest chain from φ to ϕ is j. The *height* of a class φ is the length of the longest chain of classes that begins in φ.

The *degree* of eigenvalue $\rho(A)$, denoted as $\nu(A)$, is defined as the size of the largest Jordan block corresponding to $\rho(A)$. The null space, $\mathcal{N}((\rho(A)I - A)^{\nu(A)})$, is called the *algebraic eigenspace* of A and its elements are called *generalized eigenvectors*. From Rothblum (1975) (or Exercise (3.28) of Berman and Plemmons (1979)), $\nu(A)$ equals the length of the longest chain. Assume that the geometric number of $\rho(A)$ is L. Then there are L Jordan blocks associated with $\rho(A)$. Denote by $\{m_{\rho(A),i}, 1 \le i \le L\}$ the sizes of the L Jordan blocks.

By Rothblum (1975), a chain of length $\nu(A)$ has $\nu(A)$ nonzero and nonnegative generalized eigenvectors, which correspond to a Jordan block of the size $\nu(A)$. Thus, Jordan blocks of size $\nu(A)$ correspond to disjoint chains of length $\nu(A)$. Denote by $\{x^{(i,1)}, x^{(i,2)}, \ldots, x^{(i,m_{\rho(A),i})}\}$ the row generalized eigenvectors and by $\{y^{(i,1)}, y^{(i,2)}, \ldots, y^{(i,m_{\rho(A),i})}\}$ the column generalized eigenvectors corresponding to the ith Jordan block. Then the vectors satisfy

$$
\begin{aligned}
x^{(i,1)}A &= \rho(A)x^{(i,1)}, \\
x^{(i,j)}A &= \rho(A)x^{(i,j)} + x^{(i,j-1)}, & 2 \le j \le m_{\rho(A),i}, \\
Ay^{(i,m_{\rho(A),i})} &= \rho(A)y^{(i,m_{\rho(A),i})}, \\
Ay^{(i,j)} &= \rho(A)y^{(i,j)} + y^{(i,j+1)}, & 1 \le j \le m_{\rho(A),i} - 1.
\end{aligned}
$$

For any Jordan block with $m_{\rho(A),i} = \nu(A)$, assume that the basic classes on the corresponding chain are $\{\varphi_{i,1}, \varphi_{i,2}, \ldots, \varphi_{i,m_{\rho(A),i}}\}$, where $\varphi_{i,1}$ is accessible from $\varphi_{i,2}$, $\varphi_{i,2}$ is accessible from $\varphi_{i,3}$, etc. Then the class $\varphi_{i,1}$ has no access to any other basic class; otherwise, the length of the chain would be greater than $\nu(A)$ and the degree of the eigenvalue $\rho(A)$ would be greater than $\nu(A)$. Similarly, the class $\varphi_{i,m_{\rho(A),i}}$ is not accessible from any basic class. Consequently, if $m_{\rho(A),i} = \nu(A)$, the eigenvectors $\boldsymbol{x}^{(i,1)}$ and $\boldsymbol{y}^{(i,m_{\rho(A),i})}$ can be chosen to be nonnegative. Furthermore, because all classes that can be reached from class $\varphi_{i,1}$ are not basic, $(\boldsymbol{x}^{(i,1)})_k > 0$ if and only if the state k is accessible from the class $\varphi_{i,1}$ and, because all classes that have access to class $\varphi_{i,m_{\rho(A),i}}$ are not basic, $(\boldsymbol{y}^{(i,m_{\rho(A),i})})_k > 0$ if and only if the state k has access to the class $\varphi_{i,m_{\rho(A),i}}$. The generalized eigenvectors associated with Jordan blocks of order smaller than $\nu(A)$ may not have the above property.

For a PH generator $\boldsymbol{T}$, choose λ sufficiently large such that $\boldsymbol{A} = \lambda \boldsymbol{I} + \boldsymbol{T}$ is a nonnegative matrix. Then all of the above definitions can be applied to $\boldsymbol{T}$ by using the matrix $\boldsymbol{A}$. It is easy to see that the definitions are independent of λ as long as $\boldsymbol{A}$ is a nonnegative matrix. All the results hold for the matrix $\boldsymbol{T}$ as well.

Now we are ready to state and prove a lower bound on $\kappa_{\min}(\boldsymbol{T})$.

Theorem 5.1. *For any PH generator $\boldsymbol{T}$, $\kappa_{\min}(\boldsymbol{T}) \geq \nu(\boldsymbol{T})$. For any irreducible PH representation $(\boldsymbol{\alpha}, \boldsymbol{T})$ with algebraic degree $\kappa_{\min}(\boldsymbol{T})$, $\rho(\boldsymbol{T})$ is a pole of its LST with multiplicity $\nu(\boldsymbol{T})$. If all initial classes of $\boldsymbol{T}$ are also basic then we have $\kappa_{\min}(\boldsymbol{T}) = \nu(\boldsymbol{T})$ and $\rho(\boldsymbol{T})$ is the only pole of the LST with multiplicity $\nu(\boldsymbol{T})$ for any PH representation $(\boldsymbol{\alpha}, \boldsymbol{T})$ of minimal algebraic degree.*

Proof. Assume that the order of $\boldsymbol{T}$ is m. Recall that $-\lambda = \{-\lambda_1, \ldots, -\lambda_m\}$ are the eigenvalues of $\boldsymbol{T}$, which are ordered by their increasing moduli. Then we have $\rho(\boldsymbol{T}) = -\lambda_1$. Let $N_1 = m_{\rho(\boldsymbol{T}),1} + m_{\rho(\boldsymbol{T}),2} + \cdots + m_{\rho(\boldsymbol{T}),L}$, where L is the geometric number of the eigenvalue $-\lambda_1$ and N_1 is the algebraic number of $-\lambda_1$,

Suppose that $\kappa_{\min}(\boldsymbol{T})$ is attained by the irreducible PH representation $(\boldsymbol{\alpha}, \boldsymbol{T})$. Since $(\boldsymbol{\alpha}, \boldsymbol{T})$ is irreducible, by Proposition 3.2, the vector $\mathcal{M}_{\boldsymbol{T}}(\boldsymbol{\alpha}) = \boldsymbol{\alpha} \boldsymbol{T}^{-1}/(\boldsymbol{\alpha} \boldsymbol{T}^{-1} \boldsymbol{e})$ is positive and $\kappa(\boldsymbol{\alpha}, \boldsymbol{T}) = \kappa(\mathcal{M}_{\boldsymbol{T}}(\boldsymbol{\alpha}), \boldsymbol{T})$. Let $\boldsymbol{\theta} = \mathcal{M}_{\boldsymbol{T}}(\boldsymbol{\alpha})$, which is positive by Proposition 3.2. To prove the first part of Theorem 5.1, we need to show that $\kappa(\boldsymbol{\theta}, \boldsymbol{T}) \geq \nu(\boldsymbol{T})$.

Denote the Jordan canonical form of $\boldsymbol{T}$ by $\boldsymbol{J}$. Based on the above discussion on the algebraic eigenspace, the matrix $\boldsymbol{T}$ can be written as

$$\boldsymbol{T} = \boldsymbol{Y} \boldsymbol{J} \boldsymbol{X} = \boldsymbol{Y} \begin{pmatrix} \boldsymbol{J}_1 & & & \\ & \ddots & & \\ & & \boldsymbol{J}_L & \\ & & & \boldsymbol{J}^* \end{pmatrix} \boldsymbol{X},$$

where

$$\boldsymbol{J}_i = \begin{pmatrix} -\lambda_1 & & & \\ 1 & -\lambda_1 & & \\ & \ddots & \ddots & \\ & & 1 & -\lambda_1 \end{pmatrix}_{m_i \times m_i}, \quad 1 \leq i \leq L, \qquad \boldsymbol{X} = \begin{pmatrix} \boldsymbol{x}^{(1,1)} \\ \boldsymbol{x}^{(1,2)} \\ \vdots \\ \boldsymbol{x}^{(L,m_{\rho(\boldsymbol{T}),L})} \\ \boldsymbol{X}^* \end{pmatrix};$$

$$\boldsymbol{Y} = (\boldsymbol{y}^{(1,1)}, \boldsymbol{y}^{(1,2)}, \boldsymbol{y}^{(1,m_{\rho(\boldsymbol{T}),1})}, \boldsymbol{y}^{(2,1)}, \ldots, \boldsymbol{y}^{(L,m_{\rho(\boldsymbol{T}),L})}, \boldsymbol{Y}^*),$$

and J^*, X^*, and Y^* are the remaining Jordan blocks in J, and the remaining parts of the matrices X and Y, respectively. Note that all the eigenvalues associated with Jordan blocks in J^* are different from $-\lambda_1$. By the definition of Jordan canonical form we have $XY = I$. By routine calculations we obtain

$$
\begin{aligned}
\text{adj}(sI - T) &= (sI - T)^{-1} \det(sI - T) \\
&= Y(sI - J)^{-1} X \det(sI - T) \\
&= Y \begin{pmatrix} (sI - J_1)^{-1} & & & \\ & \ddots & & \\ & & (sI - J_L)^{-1} & \\ & & & (sI - J^*)^{-1} \end{pmatrix} X \det(sI - T).
\end{aligned}
$$
(5.1)

Furthermore, we have, for $1 \le i \le L$,

$$
\begin{aligned}
&(sI - J_i)^{-1} \\
&= \begin{pmatrix} (s+\lambda_1)^{N_1-1} & & & \\ (s+\lambda_1)^{N_1-2} & (s+\lambda_1)^{N_1-1} & & \\ \vdots & \ddots & \ddots & \\ (s+\lambda_1)^{N_1-m_{\rho(T),i}} & \cdots & (s+\lambda_1)^{N_1-2} & (s+\lambda_1)^{N_1-1} \end{pmatrix} \frac{1}{(s+\lambda_1)^{N_1}}.
\end{aligned}
$$
(5.2)

Note that $\det(sI - T)(s + \lambda_1)^{-N_1}$ is a polynomial function for which $s = -\lambda_1$ is not a root. Equation (5.2) indicates that all elements of the matrix $\text{adj}(sI - T)$ are polynomial functions for which $s = -\lambda_1$ is a root of multiplicity greater than or equal to $N_1 - \nu(T)$. Taking derivatives on both sides of (5.2) and letting $s = -\lambda_1$, we obtain, for $1 \le i \le L$,

$$
\begin{aligned}
((sI - J_i)^{-1})^{(n)}_{s=-\lambda_1} &= \begin{cases} 0 & \text{if } 0 \le n \le N_1 - \nu(T) - 1, \\ 0 & \text{if } n = N_1 - \nu(T), m_{\rho(T),i} < \nu(T), \end{cases} \\
((sI - J^*)^{-1})^{(n)}_{s=-\lambda_1} &= 0 \quad \text{if } 0 \le n \le N_1 - 1,
\end{aligned}
$$
(5.3)

and

$$
\begin{aligned}
&((sI - J_i)^{-1})^{(N_1-\nu(T))}_{s=-\lambda_1} \\
&= \left(\frac{1}{(s+\lambda_1)^{N_1}} \right)_{s=-\lambda_1} \begin{pmatrix} 0 & & & \\ \vdots & & \ddots & \\ 0 & & \ddots & \ddots \\ (N_1-\nu(T))! & 0 & \cdots & 0 \end{pmatrix} \quad \text{if } m_{\rho(T),i} = \nu(T).
\end{aligned}
$$
(5.4)

Recall that $h^*(s) = \text{adj}(sI - T)T^0$. By (5.1)–(5.4), we obtain

$$
h^{*(n)}(s)|_{s=-\lambda_1} = 0 \quad \text{if } 1 \le n \le N_1 - \nu(T) - 1,
$$
(5.5)

and

$$h^{*(N_1 - \nu(T))}(s)|_{s=-\lambda_1}$$

$$= \left(\frac{\det(sI - T)}{(s + \lambda_1)^{N_1}}\right)_{s=-\lambda_1} (N_1 - \nu(T))! \sum_{\{i=1 : m_{\rho(T),i}=\nu(T)\}}^{L} y^{(i,m_{\rho(T),i})}(x^{(i,1)}T^0). \quad (5.6)$$

Equation (5.6) leads to

$$\theta h^{*(N_1 - \nu(T))}(s)|_{s=-\lambda_1}$$

$$= \left(\frac{\det(sI - T)}{(s + \lambda_1)^{N_1}}\right)_{s=-\lambda_1} (N_1 - \nu(T))! \sum_{\{i=1 : m_{\rho(T),i}=\nu(T)\}}^{L} (\theta y^{(i,m_{\rho(T),i})})(x^{(i,1)}T^0).$$

$$(5.7)$$

Since the underlying Markov chain associated with the PH representation is absorbing with probability 1, we must have $x^{(i,1)}T^0 > 0$ if $m_{\rho(T),i} = \nu(T)$. Since θ is positive, and $y^{(i,m_{\rho(T),i})}$ is nonzero and nonnegative, if $m_{\rho(T),i} = \nu(T)$, we must have $\theta y^{(i,m_{\rho(T),i})} > 0$. Therefore, (5.5) and (5.7) lead to

$$\theta h^{*(n)}(s)|_{s=-\lambda_1} \begin{cases} = 0 & \text{if } 1 \leq n \leq N_1 - \nu(T) - 1, \\ \neq 0 & \text{if } n = N_1 - \nu(T). \end{cases} \quad (5.8)$$

Equation (5.8) implies that $s = -\lambda_1$ is a root with multiplicity $N_1 - \nu(T)$ of the polynomial function $\theta h^*(s)$. Therefore, by (3.1), the degree of the polynomial function $q(s)$ is at least $N_1 - (N_1 - \nu(T)) = \nu(T)$, which implies that $\kappa(\theta, T) \geq \nu(T)$ and the multiplicity of $-\lambda_1$ as a pole of the LST of (θ, T) is $\nu(T)$. This completes the first part of Theorem 5.1.

To prove the second part of Theorem 5.1, we first use the SPA introduced in Section 3 (see (3.7)) to find an equivalent bidiagonal ME representation for (α, T). Using the SPA, we find an $m \times m$ matrix P such that $TP = PS(\lambda, m)$ and $Pe = e$ (see (3.8) for the definition of $S(\lambda, m)$).

Consider any initial class φ_h. By assumption, φ_h is also a basic class. We assume that the height of φ_h is h. Then the longest chain beginning at φ_h has exactly h basic classes $\{\varphi_1, \varphi_2, \ldots, \varphi_h\}$, where φ_i has access to φ_j if $i > j$. If $h = \nu(T)$, by Theorem 3.20 of Berman and Plemmons (1979) (also see Rothblum (1975)), there exist generalized eigenvectors $\{x^{(1)}, x^{(2)}, \ldots, x^{(h)}\}$ of $-\lambda_1$, such that

$$x^{(k)}(-\lambda_1 I - T)^k = 0, \qquad 1 \leq k \leq h. \quad (5.9)$$

More specifically (see Exercise (3.29) of Berman and Plemmons (1979)), we have $x^{(1)}T = \rho(T)x^{(1)}$ and $x^{(k)}T = -\lambda_1 x^{(k)} + x^{(k-1)}$, $2 \leq k \leq h$, so that

$$x^{(k)} = (\lambda_1 I + T)^{h-k}x^{(h)}, \qquad 1 \leq k \leq h - 1. \quad (5.10)$$

The vectors $\{x^{(1)}, x^{(2)}, \ldots, x^{(h)}\}$ are nonnegative. For any state i that can be reached from the class φ_h, $(x^{(h)})_i > 0$. If $h < \nu(T)$, we can introduce fictitious states to extend the chain $\{\varphi_1, \varphi_2, \ldots, \varphi_h\}$ to be the longest chain and φ_h is still an initial class. Then Theorem 3.20 of Berman and Plemmons (1979) still applies. Therefore, for any initial basic class φ_h, there

exists a nonnegative generalized eigenvector $x^{(h)}$ such that $(x^{(h)})_i > 0$ for all states accessible from the class φ_h.

Suppose that $x^{(h)}$ is normalized to a stochastic vector. Consider the PH representation $(x^{(h)}, T)$. Equations (3.7) and (5.10) lead to an equivalent ME representation

$$(x^{(h)} P, S(\lambda, m), e),$$

where

$$
\begin{aligned}
x^{(h)} P &= \left(\frac{x^{(h)} T^0}{\lambda_1}, \frac{x^{(h)}(\lambda_1 I + T) T^0}{\lambda_1^2}, \ldots, \frac{x^{(h)}(\lambda_1 I + T)^{h-1} T^0}{\lambda_1^h}, 0, \ldots, 0 \right) \\
&= \left(\frac{x^{(h)} T^0}{\lambda_1}, \frac{x^{(h-1)} T^0}{\lambda_1^2}, \ldots, \frac{x^{(1)} T^0}{\lambda_1^h}, 0, \ldots, 0 \right).
\end{aligned}
$$

Note that $\lambda_j = \lambda_1$ for $1 \leq j \leq N_1$. For any state i in a class that can be reached from the class φ_1, we have $(x^{(1)})_i > 0$. Since the underlying Markov chain associated with the PH representation is absorbing, we must have $x^{(1)} T^0 > 0$. Therefore, $(x^{(h)}, T)$ is a generalized Erlang distribution of order h with parameter λ_1.

Define

$$\alpha = \left(\left(\sum_h x^{(h)} \right) e \right)^{-1} \left(\sum_h x^{(h)} \right),$$

where the summation is over all h corresponding to basic classes φ_h. Since all initial classes are basic, any class must be reached from at least one initial basic class. Consequently, the vector α is positive and stochastic. Then (α, T) is an irreducible PH representation and represents a convex combination of generalized Erlang distributions of order $\nu(T)$ or less with the same parameter λ_1. Since $\nu(T)$ is the length of at least one chain, (α, T) represents a generalized Erlang distribution of the order $\nu(T)$ with parameter λ_1. Hence, we must have $\kappa(\alpha, T) = \nu(T)$ and $\kappa_{\min}(T) = \nu(T)$, which leads to the second part of the lemma. This completes the proof of Theorem 5.1.

Theorem 5.1 can be used to find the minimum algebraic degree for many PH generators and it leads to the second main result of this section. It is easy to show (see Berman and Plemmons (1979)) that any PH generator T can be written in the form

$$
T = \begin{pmatrix}
T_{1,1} & & & \\
T_{2,1} & T_{2,2} & & \\
\vdots & \ddots & \ddots & \\
T_{K,1} & \cdots & T_{K,K-1} & T_{K,K}
\end{pmatrix},
\tag{5.11}
$$

where $T_{1,1}, T_{2,2}, \ldots, T_{K,K}$ satisfy the following properties.

- $\rho(T_{1,1}) > \rho(T_{2,2}) > \cdots > \rho(T_{K,K})$.

- Taken as classes of the matrix $T_{k,k}$, all initial classes are basic.

This decomposition expression for T is obtained by rearranging states of T according to the basic classes. First, all classes that are accessible from a basic class of T are included in $T_{1,1}$. From the remaining classes, find all the classes with the largest Perron–Frobenius eigenvalue, which are called basic classes in $T_{2,2}$. All classes that are accessible from this set of basic

classes are included in $T_{2,2}$. Expression (5.11) is obtained by repeating this process until all classes are exhausted. Note that $T_{k,k}$ may not be irreducible. Let $\nu(T_{k,k})$ denote the degree of the Perron–Frobenius eigenvalue $\rho(T_{k,k})$ of the matrix $T_{k,k}$.

Theorem 5.2. *Consider the PH generator T given in (5.11). Assume that $\rho(T_{k,k})$ is not an eigenvalue of any diagonal block except $T_{k,k}$ for $1 \le k \le K$. Then we have*

$$\kappa_{\min}(T) \le \sum_{k=1}^{K} \nu(T_{k,k}). \tag{5.12}$$

Proof. The ideas of our proof are similar to those of the proof of Theorem 5.1. Define, for $1 \le k \le K$,

$$B_k = \begin{pmatrix} T_{k,k} & & & \\ T_{k+1,k} & T_{k+1,k+1} & & \\ \vdots & \ddots & \ddots & \\ T_{K,k} & \cdots & T_{K,K-1} & T_{K,K} \end{pmatrix}.$$

Denote by N_k the algebraic number of the eigenvalue $\rho(T_{k,k})$ of the matrix $T_{k,k}$, $1 \le k \le K$. By the assumption that $\rho(T_{k,k})$ is not an eigenvalue of any block except $T_{k,k}$, N_k is also the algebraic number of the eigenvalue $\rho(T_{k,k})$ of the matrix T and the algebraic number of the eigenvalue $\rho(T_{k,k})$ of the matrix B_k, $1 \le k \le K$. Again, by the assumption, $\nu(T_{k,k})$ is also the degree of $\rho(T_{k,k})$ as an eigenvalue of T and the degree of the eigenvalue $\rho(T_{k,k})$ of the matrix B_k, $1 \le k \le K$.

To prove that $\kappa_{\min}(T) \le \sum_{k=1}^{K} \nu(T_{k,k})$, we construct a PH representation with algebraic degree less than or equal to $\sum_{k=1}^{K} \nu(T_{k,k})$. For $1 \le k \le K$, by applying Theorem 5.1, there exists a positive stochastic vector $x^{(k,k)}$ such that $(x^{(k,k)}, T_{k,k})$ is an irreducible PH representation of algebraic degree $\nu(T_{k,k})$. In fact, the proof of Lemma 5.1 shows that $(x^{(k,k)}, T_{k,k})$ represents a generalized Erlang distribution of order $\nu(T_{k,k})$ with parameter $\rho(T_{k,k})$. Rewrite $(-\rho(T_{k,k})I + T)^n$ so that

$$(-\rho(T_{k,k})I + T)^n = \begin{pmatrix} T_{1,1}^{(k,n)} & & & \\ T_{2,1}^{(k,n)} & T_{2,2}^{(k,n)} & & \\ \vdots & \ddots & \ddots & \\ T_{K,1}^{(k,n)} & \cdots & T_{K,K-1}^{(k,n)} & T_{K,K}^{(k,n)} \end{pmatrix}, \tag{5.13}$$

where $T_{i,i}^{(k,n)} = (-\rho(T_{k,k})I + T_{i,i})^n$, $1 \le i \le K$. Define, for $1 \le k \le K$ and $1 \le j \le k\text{-}1$,

$$x^{(k,j)} = -\left(\sum_{i=j+1}^{k} x^{(k,i)} T_{i,j}^{(k,\nu(T_{k,k}))} \right)(-\rho(T_{k,k})I + T_{j,j})^{-\nu(T_{k,k})}$$

for $1 \le k \le K$, $1 \le j \le k - 1$. The inverse matrices in (5.13) are well defined since $\rho(T_{k,k})$ is not an eigenvalue of $T_{j,j}$ for $j < k$. Define

$$\alpha^{(k)} = (x^{(k,1)}, x^{(k,2)}, \ldots, x^{(k,k)}, 0, \ldots, 0) \quad \text{for } 1 \le k \le K.$$

Since $x^{(k,k)}(-\rho(T_{k,k})I + T_{k,k})^{\nu(T_{k,k})} = 0$, it is easy to verify that

$$\alpha^{(k)}(-\rho(T_{k,k})I + T)^{\nu(T_{k,k})} = 0.$$

The SPA method (see (3.7)) implies that $(\boldsymbol{\alpha}^{(k)}, \boldsymbol{T}, \boldsymbol{e})$ is the representation of a function similar to the generalized Erlang distribution with algebraic degree less than or equal to $\nu(\boldsymbol{T}_{k,k})$ and parameter $-\rho(\boldsymbol{T}_{k,k})$. This function may not be a probability distribution function since $\boldsymbol{\alpha}^{(k)}$ may not be nonnegative. Nevertheless, the LST of the function is a rational function and the denominator of that transform has a degree less than or equal to $\nu(\boldsymbol{T}_{k,k})$ if the denominator and the numerator are coprime polynomials. For $\varepsilon > 0$, define

$$\boldsymbol{\alpha} = \left(\sum_{k=1}^{K} \varepsilon^{k-1} \boldsymbol{\alpha}^{(k)} \right) \Big/ \left(\sum_{k=1}^{K} \varepsilon^{k-1} \boldsymbol{\alpha}^{(k)} \boldsymbol{e} \right).$$

Since $\{\boldsymbol{x}^{(k,k)}, 1 \le k \le K\}$ are positive vectors, it is easy to see that the vector $\boldsymbol{\alpha}$ becomes positive if ε is sufficiently small. Then $(\boldsymbol{\alpha}, \boldsymbol{T})$ is a PH representation and its algebraic degree is less than or equal to $\sum_{k=1}^{K} \nu(\boldsymbol{T}_{k,k})$. This completes the proof of Theorem 5.2.

Applying Theorem 5.1 or 5.2 to Example 2.1, we find that if $\tau_1 > \tau_2 > \cdots > \tau_m$ then $\kappa_{\min}(\boldsymbol{T}) = 1$ and the corresponding pole is $-\tau_m$. Applying Theorem 5.2 to Example 2.2, we obtain $\kappa_{\min}(\boldsymbol{T}) \le 3$ and the corresponding poles are in $\{-2, -2, -10\}$.

Since any PH generator can be rewritten in the form given in (5.12), Theorem 5.2 can be used to find a bound on $\kappa_{\min}(\boldsymbol{T})$ for $\boldsymbol{T}$ that do not satisfy the conditions of Theorem 5.1. On the other hand, there are examples that do not satisfy the conditions in Theorem 5.2, and more work is needed to even bound $\kappa_{\min}(\boldsymbol{T})$ for these cases.

To finish the paper, we will make a few further remarks on the relationship between $\kappa_{\min}(\boldsymbol{T})$ and $\sum_{k=1}^{K} \nu(\boldsymbol{T}_{k,k})$.

Example 5.1. (*Example 2.1 continued.*) Assume that $\tau_1 < \tau_2 < \cdots < \tau_m$. Applying the method used in the first part of the proof of Proposition 3.2, we show that $\kappa_{\min}(\boldsymbol{T}) = m$. By routine calculations we obtain

$$\boldsymbol{\alpha} h^*(-\tau_i) = \sum_{k=i}^{m} \alpha_k \left(\prod_{j=1}^{k} \tau_j \right) \left(\prod_{j=k+1}^{m} (\tau_j - \tau_k) \right) > 0, \qquad i = 1, 2, \ldots, m,$$

for any positive stochastic vector $\boldsymbol{\alpha}$. Therefore, $\{-\tau_1, -\tau_2, \ldots, -\tau_m\}$ are poles of the LST of any irreducible PH representation $(\boldsymbol{\alpha}, \boldsymbol{T})$. Consequently, we have $\kappa_{\min}(\boldsymbol{T}) = m$. Note that $\kappa_{\min}(\boldsymbol{T}) = m$ has been shown in Example 3.1 using a different method.

Example 5.2. Consider the PH generator

$$\boldsymbol{T} = \begin{array}{c} 1 \\ 2 \\ 3 \\ 4 \\ 5 \\ 6 \end{array} \begin{pmatrix} -3 & 1.5 & 0 & 0 & 0 & 0 \\ 2 & -5 & 0 & 0 & 0 & 0 \\ 1 & 1 & -3 & 0.5 & 0 & 0 \\ 1 & 0 & 2 & -3 & 0 & 0 \\ 1 & 0 & 1 & 0 & -4 & 0 \\ 1 & 1 & 0 & 1 & 0 & -10 \end{pmatrix}.$$

The states form four irreducible classes, $\{1, 2\}$, $\{3, 4\}$, $\{5\}$, and $\{6\}$. Since the Perron–Frobenius eigenvalues associated with classes $\{1, 2\}$ and $\{3, 4\}$ are both -2, in decomposition (5.12) we see that $K = 3$, with the states corresponding to $\boldsymbol{T}_{1,1}$ given by $\{1, 2, 3, 4\}$, the state corresponding to $\boldsymbol{T}_{2,2}$ equal to $\{5\}$, and the state corresponding to $\boldsymbol{T}_{3,3}$ equal to $\{6\}$. The eigenvalues of $\boldsymbol{T}$ are $\{-2, -2, -4, -4, -6, -10\}$. Since -4 is the Perron–Frobenius eigenvalue of the block $\boldsymbol{T}_{2,2}$, but also an eigenvalue of the block $\boldsymbol{T}_{1,1}$, the conditions of Theorem 5.2 are not satisfied.

After some calculation, we see that $\chi(\boldsymbol{T}) = 6$, $(-\gamma_1, \ldots, -\gamma_{\chi(\boldsymbol{T})}) = (-2, -2, -4, -4, -6, -10)$, $M = 4$, $(-\delta_1, \ldots, -\delta_M) = (-2, -4, -6, -10)$, and

$$
\boldsymbol{P}\boldsymbol{Q}^{-1} = \begin{pmatrix} 1.125 & 0 & 0 & 0 & -0.125 & 0 \\ 0.75 & 0 & 0 & 0 & 0.25 & 0 \\ 0.4766 & 0.6094 & -0.0313 & 0 & -0.0547 & 0 \\ -0.3594 & 1.2188 & 0.0625 & 0 & 0.0781 & 0 \\ 0.4961 & 0.3047 & 0.1172 & -0.0078 & 0.0898 & 0 \\ 0.1514 & 0.1523 & 0.0104 & 0 & 0.0508 & 0.6351 \end{pmatrix}.
$$

By reasoning similar to that used in our consideration of Example 3.2, for an irreducible representation $(\boldsymbol{\alpha}, \boldsymbol{T})$, α_5 and α_6 must be positive. Since columns 2 and 6 are nonnegative and column 4 is nonpositive, it follows that $\boldsymbol{\beta}(1, 2)$, $\boldsymbol{\beta}(2, 2)$, and $\boldsymbol{\beta}(4, 1)$ are nonzero for any irreducible $(\boldsymbol{\alpha}, \boldsymbol{T})$. By Proposition 3.3, any irreducible $(\boldsymbol{\alpha}, \boldsymbol{T})$ must have poles $\{-2, -2, -4, -4, -10\}$, which implies that $\kappa_{\min}(\boldsymbol{T}) \geq 5 > \sum_{k=1}^{3} \nu(\boldsymbol{T}_{k,k}) = 2 + 1 + 1 = 4$. In fact, we find $\boldsymbol{\alpha} = (0.36, 0, 0, 0, 0.32, 0.32)$ such that the fifth element in $\boldsymbol{\alpha}\boldsymbol{P}\boldsymbol{Q}^{-1}$ is 0, which implies that $\kappa_{\min}(\boldsymbol{T}) = 5$.

Example 5.1 indicates that $\kappa_{\min}(\boldsymbol{T}) = \sum_{k=1}^{K} \nu(\boldsymbol{T}_{k,k})$ is possible. Example 5.2 demonstrates that $\kappa_{\min}(\boldsymbol{T}) > \sum_{k=1}^{K} \nu(\boldsymbol{T}_{k,k})$ is possible. Thus, if the matrix $\boldsymbol{T}$ is reducible, the relationship between $\kappa_{\min}(\boldsymbol{T})$ and $\sum_{k=1}^{K} \nu(\boldsymbol{T}_{k,k})$ is complicated. We leave the problem of identifying $\kappa_{\min}(\boldsymbol{T})$ for the general case as an open problem for future research.

Acknowledgement

References

ASMUSSEN, S. AND BLADT, M. (1997). Renewal theory and queueing algorithms for matrix-exponential distributions. In *Matrix-Analytic Methods in Stochastic Models* (Lecture Notes Pure Appl. Math. **183**), eds S. R. Chakravarthy and A. S. Alfa, Marcel Dekker, New York, pp. 313–341.

BERMAN, A. AND PLEMMONS, R. J. (1979). *Nonnegative Matrices in the Mathematical Sciences*. Academic Press, New York.

BOTTA, R. F., HARRIS, C. M. AND MARCHAL, W. G. (1987). Characterizations of generalized hyper-exponential distribution functions. *Commun. Statist. Stoch. Models* **3**, 115–148.

COMMAULT, C. AND CHEMLA, J. P. (1996). An invariant of representations of phase-type distributions and some applications. *J. Appl. Prob.* **33**, 368–381.

GANTMACHER, F. R. (1959). *Applications of the Theory of Matrices*. Interscience, New York.

GANTMACHER, F. R. (1960). *Matrix Theory*, Vol. I. Chelsea, New York.

HE, Q.-M. AND ZHANG, H. (2006). Spectral polynomial algorithms for computing bi-diagonal representations for phase type distributions and matrix-exponential distributions. *Stoch. Models* **22**, 289–317.

HE, Q.-M. AND ZHANG, H. (2007). On matrix exponential distributions. *Adv. Appl. Prob.* **39**, 271–292.

HE, Q.-M. AND ZHANG, H. (2008). An algorithm for computing the minimal Coxian representation. *INFORMS J. Computing* **20**, 179–190.

LATOUCHE, G. AND RAMASWAMI, V. (1999). *Introduction to Matrix Analytic Methods in Stochastic Modeling*. Society for Industrial and Applied Mathematics, Philadelphia, PA.

LIPSKY, L. (1992). *Queueing Theory: A Linear Algebraic Approach*. Macmillan, New York.

NEUTS, M. F. (1975). Probability distributions of phase type. In *Liber Amicorum Prof. Emeritus H. Florin*, University of Louvain, pp. 173–206.

NEUTS, M. F. (1981). *Matrix-Geometric Solutions in Stochastic Models*. Johns Hopkins University Press, Baltimore, MD.

NEUTS, M. F. (1984). The abscissa of convergence of the Laplace–Stieltjes transform of a PH-distribution. *Commun. Statist. B Simul. Computing* **13**, 367–373.

O'CINNEIDE, C. A. (1989). On nonuniqueness of representations of phase-type distributions. *Commun. Statist. Stoch. Models* **5,** 247–259.

O'CINNEIDE, C. A. (1990). Characterization of phase-type distributions. *Commun. Statist. Stoch. Models* **6,** 1–57.

ROTHBLUM, U. G. (1975). Algebraic eigenspaces of nonnegative matrices. *Linear Algebra Appl.* **12,** 281–292.

SENETA, E. (2006). *Non-Negative Matrices and Markov Chains*. Springer, New York.

VAN DE LIEFVOORT, A. (1990). The moment problem for continuous distributions. Tech. Rep. CM-1990-02, University of Missouri.