J. Appl. Prob. **58,** 880–889 (2021)
*doi:*10.1017/jpr.2021.18

BOUNDS ON THE MEAN AND SQUARED COEFFICIENT OF VARIATION OF PHASE-TYPE DISTRIBUTIONS

QI-MING HE,[*] *University of Waterloo*

Abstract

We consider a class of phase-type distributions (PH-distributions), to be called the MMPP class of PH-distributions, and find bounds of their mean and squared coefficient of variation (SCV). As an application, we have shown that the SCV of the event-stationary inter-event time for Markov modulated Poisson processes (MMPPs) is greater than or equal to unity, which answers an open problem for MMPPs. The results are useful for selecting proper PH-distributions and counting processes in stochastic modeling.

Keywords: Phase-type distribution; mean; squared coefficient of variation; Markov modulated Poisson process; doubly stochastic matrix; M-matrix

2010 Mathematics Subject Classification: Primary 60J27
Secondary 60E15

1. Introduction

Recently, the study of the variability of Markov modulated Poisson processes (MMPPs) (see [6] and [7]) and Markov switched Poisson processes (MSPPs) (see [5]) has attracted the attention of many researchers (e.g. [2] and references therein). One of the main issues of interest is to show that the squared coefficient of variations (SCV) of the event-stationary inter-arrival time is greater than or equal to unity. Bounds of this kind are useful for selecting proper PH-distributions and counting processes in stochastic modeling for applications in telecommunications, transportation, manufacturing, finance/insurance, and healthcare. While the issue for MSPPs has been resolved largely in [2] and in the existing literature on hyper-exponential distributions (e.g. [3]), the one for MMPP remained as an open problem, which is answered in this paper. Since the event-stationary inter-event time, as well as the time-stationary inter-arrival time, is a special phase-type distribution (PH-distribution), this paper investigates the class of phase-type distributions related to MMPPs, and finds bounds for them, which lead to the desired bounds for MMPPs.

PH-distributions were introduced by Marcel Neuts [14] in 1975. Since then, PH-distributions have been investigated extensively (e.g. [17] and [18]) and have found applications in queueing theory and many other areas, including telecommunications, risk analysis, and biology (e.g. [19] and [21]). PH-distributions play a key role in the area of matrix-analytic methods [8, 10, 16]. Various bounds for PH-distributions have been obtained [1, 9, 18]. For instance, Aldous and Shepp [1] found that the SCV of all PH-distributions is greater than or equal to $1/m$, where m is the order of the PH-representations. That bound depends only on the order of the PH-distributions. O'Cinneide [18] obtained a number of bounds, notably the

Received 20 August 2020; revision received 14 January 2021.

[*] Postal address: Department of Management Sciences, University of Waterloo, 200 University Avenue West, Waterloo, Canada N2L 3G1. q7he@uwaterloo.ca

steepest increase property of the density function of PH-distributions. He *et al.* [9] obtained bounds for the moments, variance, and SCV of PH-distributions with infinite or finite support. In this paper, for the aforementioned applications to MMPPs, we find order-free bounds for the mean and parameter-free bounds for the SCV of the MMPP class of PH-distributions. Those results are useful for the selection of PH-distributions in practice.

The rest of the paper is organized as follows. In Section 2 we introduce the MMPP class of PH-distributions and show a few properties related to the PH-representations of such PH-distributions. In Section 3 we state and prove the main results of the paper: bounds on the mean and SCV of PH-distributions in the MMPP class. Section 4 concludes the paper by (i) discussing applications to MMPPs, (ii) briefly talking about extensions to the discrete-time PH-distributions, and (iii) listing problems for future research.

2. The MMPP class of PH-distributions and properties

A phase-type distribution, to be called a PH-distribution, is defined as the distribution of the absorption time of a special state in a finite-state discrete/continuous-time Markov chain. Since there is a one-to-one correspondence between the results obtained in this paper for discrete- and continuous-time cases, we shall focus on continuous-time PH-distributions. Consider a continuous-time Markov chain (CTMC) $\{Y(t), t \geq 0\}$ with $m+1$ states and infinitesimal generator

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

where T is an $m \times m$ matrix with negative diagonal elements and non-negative off-diagonal element, $\mathbf{T}^0$ is a non-negative and non-zero column vector of order m, $T\mathbf{e} + \mathbf{T}^0 = 0$, and $\mathbf{e}$ is the column vector with all elements being one. Define $X = \min\{t\colon\ Y(t) = m+1, t \geq 0\}$. Assume that $Y(0)$ is distributed according to $(\boldsymbol{\alpha}, \alpha_{m+1})$. It is well known that

$$\begin{aligned} \mathbb{P}\{X < t\} &= 1 - \boldsymbol{\alpha} \exp\{Tt\}\mathbf{e} \quad \text{for } t \geq 0, \\ \mathbb{E}[X] &= -\boldsymbol{\alpha} T^{-1}\mathbf{e}, \\ \mathbb{E}[X^2] &= 2\boldsymbol{\alpha} T^{-2}\mathbf{e}, \\ \mathrm{SCV}\,(X) &= \frac{2\boldsymbol{\alpha} T^{-2}\mathbf{e}}{(\boldsymbol{\alpha} T^{-1}\mathbf{e})^2} - 1. \end{aligned} \tag{2.1}$$

Random variable X has a PH-distribution and the pair $(\boldsymbol{\alpha}, T)$ is called a PH-representation of X. We refer to Neuts [16], He [8], and Buchholz, Kriege, and Felko [4] for the general theory on phase-type distributions.

For the MMPP class of PH-distributions, we choose $T = Q - D$, where Q is an infinitesimal generator for an irreducible CTMC and D is a non-zero diagonal matrix with non-negative diagonal elements $\{d_1, d_2, \ldots, d_m\}$. In Section 4 we will use MMPPs to explain why T has that structure. Let $\boldsymbol{\pi} = (\pi_1, \pi_2, \ldots, \pi_m)$ be the stationary distribution of Q, i.e. $\boldsymbol{\pi} Q = 0$ and $\boldsymbol{\pi}\mathbf{e} = 1$. Let $\lambda = \boldsymbol{\pi} D\mathbf{e}$ and $\boldsymbol{\pi}_D = \boldsymbol{\pi} D/\lambda$. We define the set of PH-random variables X with PH-representation $(\boldsymbol{\pi}, Q - D)$ and X_D with PH-representation $(\boldsymbol{\pi}_D, Q - D)$ as the *MMPP class of PH-distributions*. In the rest of this section we show several properties related to the PH-representations $(\boldsymbol{\pi}, Q - D)$ and $(\boldsymbol{\pi}_D, Q - D)$.

Let $P_Q = I + \mathrm{diag}\,(\boldsymbol{\pi})Q/c_Q$, where $\mathrm{diag}\,(\boldsymbol{\pi})$ is a diagonal matrix with $\boldsymbol{\pi}$ on its diagonal, and c_Q is a constant satisfying $c_Q \geq \max\{-q_{i,i}, i = 1, 2, \ldots, m\}$, and $Q = (q_{i,j})$.

Lemma 2.1. *The matrix P_Q is a doubly stochastic matrix.*

Proof. Since Q is an infinitesimal generator, it is easy to see that P_Q is non-negative, $P_Q\mathbf{e} = \mathbf{e}$, and $\mathbf{e}'P_Q = \mathbf{e}' + \boldsymbol{\pi} Q/c_Q = \mathbf{e}'$, where $\mathbf{e}'$ is the transpose of $\mathbf{e}$. Thus P_Q is a doubly stochastic matrix. □

Lemma 2.2. *The matrix $D - Q - \lambda\mathbf{e}\boldsymbol{\pi}$ is non-invertible if and only if $D = \lambda I$, where I is the identity matrix.*

Proof. If $D = \lambda I$, we have $\boldsymbol{\pi}(D - Q - \lambda\mathbf{e}\boldsymbol{\pi}) = \boldsymbol{\pi}(\lambda I - \lambda\mathbf{e}\boldsymbol{\pi}) = 0$. Thus $D - Q - \lambda\mathbf{e}\boldsymbol{\pi}$ has a zero eigenvalue and so is non-invertible. This proves the sufficiency of the condition $D = \lambda I$.

To show the necessity of the condition, let $\mathbf{u} = (u_1, u_2, \ldots, u_m)$ be a non-zero left eigenvector of $D - Q - \lambda\mathbf{e}\boldsymbol{\pi}$ corresponding to the eigenvalue zero. Then we obtain

$$0 = \mathbf{u}(D - Q - \lambda\mathbf{e}\boldsymbol{\pi}) = \mathbf{u}(D - Q) - \lambda(\mathbf{ue})\boldsymbol{\pi}. \tag{2.2}$$

If $\mathbf{ue} = 0$, we have $\mathbf{u}(D - Q) = 0$. Since $D - Q$ is an M-matrix, $D - Q$ is invertible and $(D - Q)^{-1}$ is positive elementwise (see [13]). That leads to $\mathbf{u} = 0$, which is a contradiction. Thus $\mathbf{ue}$ cannot be zero.

If $\mathbf{ue}$ is non-zero, without loss of generality, we assume that $\mathbf{ue} = 1$. By (2.2), we obtain $\mathbf{u} = \lambda\boldsymbol{\pi}(D - Q)^{-1}$. Thus $\mathbf{u}$ is positive elementwise and, in fact, is a stochastic vector. Multiplying both sides of (2.2) by $\mathbf{e}$, we obtain $\sum_{i=1}^{m} u_i d_i = \mathbf{u}D\mathbf{e} = \lambda$. By (2.2), we also obtain

$$\mathbf{u}(D - \lambda\mathbf{e}\boldsymbol{\pi}) = \mathbf{u}Q = \mathbf{u}(\operatorname{diag}(\boldsymbol{\pi}))^{-1}\operatorname{diag}(\boldsymbol{\pi})Q = c_Q(\mathbf{u}(\operatorname{diag}(\boldsymbol{\pi}))^{-1})(P_Q - I). \tag{2.3}$$

Multiplying both sides of the above equation by $(\operatorname{diag}(\boldsymbol{\pi}))^{-1}\mathbf{u}'$, we obtain

$$\mathbf{u}(D - \lambda\mathbf{e}\boldsymbol{\pi})(\operatorname{diag}(\boldsymbol{\pi}))^{-1}\mathbf{u}' = c_Q(\mathbf{u}(\operatorname{diag}(\boldsymbol{\pi}))^{-1}P_Q(\operatorname{diag}(\boldsymbol{\pi}))^{-1}\mathbf{u}' - \mathbf{u}(\operatorname{diag}(\boldsymbol{\pi}))^{-2}\mathbf{u}'). \tag{2.4}$$

By Lemma 2.1, P_Q is a doubly stochastic matrix. By the well-known Birkhoff theorem for doubly stochastic matrices (see [13], or [12, A.2 Theorem]), we have

$$P_Q = \sum_{\sigma\in\Omega} c_\sigma P_\sigma,$$

where Ω is the set of all permutations on $\{1, 2, \ldots, m\}$, $\sigma \in \Omega$, P_σ is the permutation matrix associated with σ, and $\{c_\sigma, \sigma \in \Omega\}$ are non-negative numbers with unit sum. For any real row vector $\mathbf{x} = (x_1, x_2, \ldots, x_m)$ and permutation σ, it is well known that $\mathbf{x}P_\sigma\mathbf{x}' \le \mathbf{xx}'$, which can be shown easily as follows:

$$\mathbf{x}P_\sigma\mathbf{x}' = \sum_{i=1}^{m} x_i x_{\sigma(i)} = \frac{1}{2}\left(\sum_{i=1}^{m} 2x_i x_{\sigma(i)}\right) \le \frac{1}{2}\left(\sum_{i=1}^{m}\left(x_i^2 + x_{\sigma(i)}^2\right)\right) = \mathbf{xx}'. \tag{2.5}$$

Consequently, the right-hand side of (2.4) is less than or equal to zero, which leads to $\mathbf{u}(D - \lambda\mathbf{e}\boldsymbol{\pi})(\operatorname{diag}(\boldsymbol{\pi}))^{-1}\mathbf{u}' \le 0$. This last expression can be rewritten explicitly as follows:

$$\sum_{i=1}^{m} u_i \frac{d_i}{\pi_i} u_i - \lambda(\mathbf{ue})^2 = \sum_{i=1}^{m}\left(\frac{1}{d_i}\right)\frac{d_i^2 u_i^2}{\pi_i^2}\pi_i - \lambda \le 0. \tag{2.6}$$

Next we consider two cases: (i) all $\{d_1, d_2, \ldots, d_m\}$ are positive, and (ii) some $\{d_1, d_2, \ldots, d_m\}$ are zero.

(i) If all $\{d_1, d_2, \ldots, d_m\}$ are positive, we choose $\boldsymbol{\pi}$ as a measure defined on $\{1, 2, \ldots, m\}$, $p=-1$, $q=1/2$ (so that $1/p+1/q=1$), $f(i)=1/d_i$, and $g(i)=(d_iu_i/\pi_i)^2$. By Hölder's inequality (see (iii) in [12, D.1.C]),

$$\int fg\,\mathrm{d}\boldsymbol{\pi} \geq \left(\int f^p\,\mathrm{d}\boldsymbol{\pi}\right)^{1/p}\left(\int g^q\,\mathrm{d}\boldsymbol{\pi}\right)^{1/q},$$

we obtain

$$\begin{aligned}\sum_{i=1}^m \left(\frac{1}{d_i}\right)\frac{d_i^2u_i^2}{\pi_i^2}\pi_i &\geq \left(\sum_{i=1}^m \left(\frac{1}{d_i}\right)^{-1}\pi_i\right)^{-1}\left(\sum_{i=1}^m \left(\frac{d_i^2u_i^2}{\pi_i^2}\right)^{0.5}\pi_i\right)^2\\ &= \left(\sum_{i=1}^m d_i\pi_i\right)^{-1}\left(\sum_{i=1}^m d_iu_i\right)^2\\ &= \lambda, \end{aligned} \tag{2.7}$$

in which we have used the facts $\boldsymbol{\pi}D\mathbf{e}=\mathbf{u}D\mathbf{e}=\lambda$. Combining (2.6) and (2.7), we obtain $\lambda \leq \sum_{i=1}^m u_i^2 d_i/\pi_i \leq \lambda$, which implies that $\sum_{i=1}^m u_i^2 d_i/\pi_i=\lambda$. On the other hand, the equality in (2.7) holds if and only if f^p and g^q are proportional, i.e. f^p/g^q is constant, which implies that $d_i/(d_iu_i/\pi_i)=\pi_i/u_i$ are the same for $i=1, 2, \ldots, m$. Since $\mathbf{ue}=\boldsymbol{\pi}\mathbf{e}=1$, we must have $\boldsymbol{\pi}=\mathbf{u}$, which leads to $\boldsymbol{\pi}D=\lambda\boldsymbol{\pi}$ by (2.2). Consequently, we must have $d_i=\lambda$ for all $i=1, 2, \ldots, m$, i.e. $D=\lambda I$.

(ii) Next we consider the case for which some $\{d_1, \ldots, d_m\}$ are zero. We show that this is impossible if $D-Q-\lambda\mathbf{e}\boldsymbol{\pi}$ is non-invertible. Without loss of generality we suppose that $d_i=0$ for $i=1, 2, \ldots, k$, and $d_i>0$ for $i=k+1, \ldots, m$. Accordingly, we partition vectors $\mathbf{u}=(\mathbf{u}_1, \mathbf{u}_2)$, $\boldsymbol{\pi}=(\boldsymbol{\pi}_1, \boldsymbol{\pi}_2)$, and matrices

$$Q=\begin{pmatrix} Q_{1,1} & Q_{1,2}\\ Q_{2,1} & Q_{2,2}\end{pmatrix}, \quad D=\begin{pmatrix} 0 & 0\\ 0 & D_2\end{pmatrix}.$$

In the above proof, since $d_i=0$ for $i=1, \ldots, k$, we can restrict $\boldsymbol{\pi}$ to $\boldsymbol{\pi}_2$, $\mathbf{u}$ to $\mathbf{u}_2$, and D to D_2 in (2.7) to obtain $\mathbf{u}_2=\gamma\boldsymbol{\pi}_2$ for some positive constant γ. We claim that $\gamma=1$. By $\mathbf{u}(D-Q-\lambda\mathbf{e}\boldsymbol{\pi})=0$, we obtain

$$\begin{aligned}\mathbf{u}_1(Q_{1,1}+\lambda\mathbf{e}\boldsymbol{\pi}_1)+\mathbf{u}_2(Q_{2,1}+\lambda\mathbf{e}\boldsymbol{\pi}_1)&=0,\\ \mathbf{u}_1(Q_{1,2}+\lambda\mathbf{e}\boldsymbol{\pi}_2)+\mathbf{u}_2(-D_2+Q_{2,2}+\lambda\mathbf{e}\boldsymbol{\pi}_2)&=0.\end{aligned}$$

Using $\mathbf{u}_2=\gamma\boldsymbol{\pi}_2$ and $\mathbf{ue}=\mathbf{u}_1\mathbf{e}+\mathbf{u}_2\mathbf{e}=1$, the above equations become

$$\begin{aligned}\mathbf{u}_1Q_{1,1}+\gamma\boldsymbol{\pi}_2Q_{2,1}+\lambda\boldsymbol{\pi}_1&=0,\\ \mathbf{u}_1Q_{1,2}+\gamma\boldsymbol{\pi}_2(Q_{2,2}-D_2)+\lambda\boldsymbol{\pi}_2&=0.\end{aligned}$$

Using $\boldsymbol{\pi}_1=\boldsymbol{\pi}_2Q_{2,1}(-Q_{1,1})^{-1}$ and the above two equations, we obtain

$$\boldsymbol{\pi}_2(\gamma(Q_{2,2}+Q_{2,1}(-Q_{1,1})^{-1}Q_{1,2}-D_2)+\lambda(I+Q_{2,1}(-Q_{1,1})^{-2}Q_{1,2}))=0.$$

Using $\boldsymbol{\pi}_2(Q_{2,2}+Q_{2,1}(-Q_{1,1})^{-1}Q_{1,2})=0$, the above equation can be reduced to

$$\boldsymbol{\pi}_2(\gamma D_2-\lambda(I+Q_{2,1}(-Q_{1,1})^{-2}Q_{1,2}))=0.$$

Note that $\lambda=\boldsymbol{\pi}D\mathbf{e}=\boldsymbol{\pi}_2D_2\mathbf{e}>0$. Multiplying both sides of the above equation by $\mathbf{e}$ yields, after canceling λ,

$$\begin{aligned}\gamma&=\boldsymbol{\pi}_2(I+Q_{2,1}(-Q_{1,1})^{-2}Q_{1,2})\mathbf{e}\\&=\boldsymbol{\pi}_2\mathbf{e}+\boldsymbol{\pi}_2Q_{2,1}(-Q_{1,1})^{-1}(-Q_{1,1})^{-1}Q_{1,2}\mathbf{e}=\boldsymbol{\pi}_2\mathbf{e}+\boldsymbol{\pi}_1\mathbf{e}=1,\end{aligned}$$

which proves the claim that $\gamma=1$. On the other hand, we rewrite (2.3) as

$$\frac{1}{c_Q}\mathbf{u}(D-\lambda\mathbf{e}\boldsymbol{\pi})+\mathbf{u}(\mathrm{diag}\,(\boldsymbol{\pi}))^{-1}=(\mathbf{u}(\mathrm{diag}\,(\boldsymbol{\pi}))^{-1})P_Q.$$

For two vectors $\mathbf{x}=(x_1,x_2,\ldots,x_m)$ and $\mathbf{y}=(y_1,y_2,\ldots,y_m)$, we reorganize their elements in ascending order to obtain $(x_{(1)},x_{(2)},\ldots,x_{(m)})$ and $(y_{(1)},y_{(2)},\ldots,y_{(m)})$. Vector $\mathbf{x}$ majorizes $\mathbf{y}$ if $\sum_{i=1}^{j}x_{(i)}\le\sum_{i=1}^{j}y_{(i)}$, for $j=1,2,\ldots,m-1$, and $\mathbf{xe}=\mathbf{ye}$ (see [12, A.1]). Since P_Q is a doubly stochastic matrix, vector $\mathbf{u}(\mathrm{diag}\,(\boldsymbol{\pi}))^{-1}$ majorizes $\mathbf{u}(D-\lambda\mathbf{e}\boldsymbol{\pi})/c_Q+\mathbf{u}(\mathrm{diag}\,(\boldsymbol{\pi}))^{-1}$ (see [12, B.2 Theorem (Hardy, Littlewood, and Pólya)]). Define

$$\begin{aligned}i_{(1)}&=\arg\min\left\{\frac{u_i}{\pi_i},\ i=1,2,\ldots,m\right\},\\ j_{(1)}&=\arg\min\left\{\frac{u_j}{\pi_j}+\frac{1}{c_Q}(u_jd_j-\lambda\pi_j),\ j=1,2,\ldots,m\right\}.\end{aligned}$$

If $u_{i_{(1)}}/\pi_{i_{(1)}}\ge1$, we have $u_i/\pi_i\ge1$ for all $i=1,2,\ldots,m$. By $\mathbf{ue}=\boldsymbol{\pi}\mathbf{e}=1$, we must have $\mathbf{u}=\boldsymbol{\pi}$, which leads to $\boldsymbol{\pi}D=\lambda\boldsymbol{\pi}$ and $D=\lambda I$ by (2.2) and contradicts to $d_i=0$ for $i=1,2,\ldots,k$. Now we must have $u_{i_{(1)}}/\pi_{i_{(1)}}<1$. Then we must have $i_{(1)}\le k$ and $d_{i_{(1)}}=0$, since $u_i/\pi_i=\gamma=1$ for $i=k+1,\ldots,m$. By the above definition of vector majorization, we must have

$$\frac{u_{i_{(1)}}}{\pi_{i_{(1)}}}+\frac{1}{c_Q}(u_{i_{(1)}}d_{i_{(1)}}-\lambda\pi_{i_{(1)}})\ge\frac{u_{j_{(1)}}}{\pi_{j_{(1)}}}+\frac{1}{c_Q}(u_{j_{(1)}}d_{j_{(1)}}-\lambda\pi_{j_{(1)}})\ge\frac{u_{i_{(1)}}}{\pi_{i_{(1)}}},$$

which leads to $-\lambda\pi_{i_{(1)}}\ge0$, which is a contradiction. Consequently, if $D-Q-\lambda\mathbf{e}\boldsymbol{\pi}$ is non-invertible, then $\{d_1,\ldots,d_m\}$ must all be positive and equal to λ, i.e. $D=\lambda I$. This completes the proof. $\square$

Lemma 2.3. *We have* $\lim_{x\to\infty}(D-xQ)^{-1}=\mathbf{e}\boldsymbol{\pi}/\lambda$.

Proof. For convenience, let $\Phi(x)=(D-xQ)^{-1}$, for $x>0$. Since $D-xQ$ is an M-matrix for $x>0$, $\Phi(x)$ exists and is positive elementwise. Then we have $(D-xQ)\Phi(x)=I$. Multiplying both sides of the expression by $\boldsymbol{\pi}$, we obtain $\boldsymbol{\pi}D\Phi(x)=\boldsymbol{\pi}$. Since $\boldsymbol{\pi}$ is positive elementwise, $\lim_{x\to\infty}D\Phi(x)$ exists. Since $xQ\Phi(x)=I-D\Phi(x)$, $\lim_{x\to\infty}Q\Phi(x)=\lim_{x\to\infty}(I-D\Phi(x))/x=0$. Since Q is an irreducible infinitesimal generator, its left and right eigenvectors of eigenvalue zero (a.k.a. Perron–Frobenius eigenvectors [13]) are unique up to a constant. Thus we must have $\lim_{x\to\infty}\Phi(x)=\mathbf{ey}$, where $\mathbf{y}$ is a positive vector. In a similar way, we can show that $\lim_{x\to\infty}\Phi(x)Q=0$. Thus we must have $\lim_{x\to\infty}\Phi(x)=\rho\mathbf{e}\boldsymbol{\pi}$, where ρ is a positive constant. Letting x go to infinity in $\boldsymbol{\pi}D\Phi(x)=\boldsymbol{\pi}$, we obtain $\boldsymbol{\pi}D\rho\mathbf{e}\boldsymbol{\pi}=\boldsymbol{\pi}$, which leads to $\rho=1/(\boldsymbol{\pi}D\mathbf{e})=1/\lambda$. $\square$

3. Main results for the MMPP class of PH-distributions

First, we show that $1/\lambda$ is a lower bound of the mean of X. We state the result in a more general manner, which is also needed in the proof of the result. Let X_x be the PH-random variable with PH-representation $(\boldsymbol{\pi}, xQ - D)$ for all $x > 0$. We recall that D is diagonal, non-zero, and non-negative, and Q is irreducible throughout this paper.

Theorem 3.1. *For* $\{D, Q, \boldsymbol{\pi}\}$ *defined in Section 2, we have*

$$\mathbb{E}[X_x] = \boldsymbol{\pi}(D - xQ)^{-1}\mathbf{e} \geq \frac{1}{\lambda} \quad \textit{for } x > 0.$$

The equality holds in the equation if and only if all diagonal elements of D are equal to each other (i.e. X_x has an exponential distribution with parameter λ). In addition, we have $\lim_{x\to 0} \mathbb{E}[X_x] = \boldsymbol{\pi} D^{-1}\mathbf{e}$ *and* $\lim_{x\to\infty} \mathbb{E}[X_x] = 1/\lambda$.

Proof. If $D = \lambda I$, it is easy to verify that $\mathbb{E}[X_x] = 1/\lambda$ for all $x > 0$. Thus all results hold.

Now we suppose that $\{d_1, \ldots, d_m\}$ are not all equal to each other. First, we show that $\lim_{x\to 0} \mathbb{E}[X_x] > 1/\lambda$. If one of the diagonal elements of D is zero, $\lim_{x\to 0} \mathbb{E}[X_x] = \infty > 1/\lambda$. If all diagonal elements of D are positive, $\lim_{x\to 0} \mathbb{E}[X_x] = \mathbb{E}[X_0] = \boldsymbol{\pi} D^{-1}\mathbf{e} = \sum_{i=1}^{m} \pi_i/d_i$. Then $\lim_{x\to 0} \mathbb{E}[X_x] > 1/\lambda$ is equivalent to

$$\left(\sum_{i=1}^{m} \pi_i/d_i\right)\left(\sum_{i=1}^{m} \pi_i d_i\right) > 1,$$

which can be verified easily as follows:

$$\begin{aligned}
\left(\sum_{i=1}^{m} \pi_i/d_i\right)\left(\sum_{i=1}^{m} \pi_i d_i\right) &= \sum_{i=1}^{m} \pi_i^2 + \sum_{i<j} \pi_i\pi_j\left(\frac{d_i}{d_j} + \frac{d_j}{d_i}\right) \\
&> \sum_{i=1}^{m} \pi_i^2 + 2\sum_{i<j} \pi_i\pi_j \\
&= \left(\sum_{i=1}^{m} \pi_i\right)^2 \\
&= 1.
\end{aligned}$$

Thus we must have $\mathbb{E}[X_x] > 1/\lambda$, if x is close to zero (recall that $\{d_1, \ldots, d_m\}$ are not all equal to each other). If $\mathbb{E}[X_x] > 1/\lambda$ holds for all $x > 0$, the desired result is obtained. Otherwise, there must be $x > 0$ such that $\mathbb{E}[X_x] = 1/\lambda$, i.e. $\boldsymbol{\pi}(D - xQ)^{-1}\mathbf{e} = 1/\lambda$, since $\mathbb{E}[X_x]$ is continuous in x for $x > 0$.

Next we show that $\boldsymbol{\pi}(D - xQ)^{-1}\mathbf{e} = 1/\lambda$ if and only if matrix $D - xQ - \lambda\mathbf{e}\boldsymbol{\pi}$ is non-invertible. If the matrix is non-invertible, there exists a non-zero row vector $\mathbf{u}$ such that $\mathbf{u}(D - xQ - \lambda\mathbf{e}\boldsymbol{\pi}) = 0$. By the proof of Lemma 2.2, $\mathbf{ue}$ cannot be zero, and without loss of generality we suppose that $\mathbf{ue} = 1$. By (2.2) we obtain $\mathbf{u} = \lambda\boldsymbol{\pi}(D - xQ)^{-1}$. Multiplying both sides of the expression by $\mathbf{e}$, we obtain $1 = \mathbf{ue} = \lambda\boldsymbol{\pi}(D - xQ)^{-1}\mathbf{e}$. So we have shown the sufficiency of the condition. On the other hand, if $1 = \lambda\boldsymbol{\pi}(D - xQ)^{-1}\mathbf{e}$, we can verify that $\lambda\boldsymbol{\pi}(D - xQ)^{-1}(D - xQ - \lambda\mathbf{e}\boldsymbol{\pi}) = 0$, which implies that $D - xQ - \lambda\mathbf{e}\boldsymbol{\pi}$ has an eigenvalue zero. Thus $D - xQ - \lambda\mathbf{e}\boldsymbol{\pi}$ is non-invertible. This proves the necessity of the condition.

Since we suppose that $\{d_1, \ldots, d_m\}$ are not all equal to each other, by Lemma 2.2, matrix $D - xQ - \lambda \mathbf{e}\boldsymbol{\pi}$ is invertible. Thus $\lambda\boldsymbol{\pi}(D - xQ)^{-1}\mathbf{e} = 1$ cannot be true, which is a contradiction. Consequently, $\mathbb{E}[X_x]$ has to be above $1/\lambda$ for all $x > 0$ for this case.

Combining results for the two cases (i.e. $\{d_1, \ldots, d_m\}$ are all equal or not), we have shown $\boldsymbol{\pi}(D - xQ)^{-1}\mathbf{e} \geq 1/\lambda$ for $x > 0$.

Finally, the result $\lim_{x\to\infty} \mathbb{E}[X_x] = 1/\lambda$ can be obtained directly from Lemma 2.3. □

Let $X_{D,x}$ be the PH-random variable with PH-representation $(\boldsymbol{\pi}_D, xQ - D)$ for all $x > 0$. The squared coefficient of variation of $X_{D,x}$ is defined as

$$\mathrm{SCV}\,(X_{D,x}) = \mathbb{E}[(X_{D,x} - \mathbb{E}[X_{D,x}])^2]/(\mathbb{E}[X_{D,x}])^2.$$

Theorem 3.2. *For $\{D, Q, \boldsymbol{\pi}\}$ defined in Section 2, we have $\mathbb{E}[X_{D,x}] = 1/\lambda$ and*

$$\mathrm{SCV}\,(X_{D,x}) = 2\lambda\mathbb{E}[X_x] - 1 \geq 1 \quad \text{for } x > 0. \tag{3.1}$$

The equality holds if and only if $D = \lambda I$.

Proof. Since $\boldsymbol{\pi} D(D - xQ)^{-1} = \boldsymbol{\pi}$, we have

$$\mathbb{E}[X_{D,x}] = \boldsymbol{\pi}_D(D - xQ)^{-1}\mathbf{e} = \frac{\boldsymbol{\pi} D}{\lambda}(D - xQ)^{-1}\mathbf{e} = \frac{\boldsymbol{\pi}\mathbf{e}}{\lambda} = \frac{1}{\lambda}.$$

By the definition of SCV, we have

$$\mathrm{SCV}\,(X_{D,x}) = \mathbb{E}[X_{D,x}^2]/(\mathbb{E}[X_{D,x}])^2 - 1.$$

Further, by Theorem 3.1, we obtain

$$\begin{aligned}\mathrm{SCV}\,(X_{D,x}) &= \frac{\mathbb{E}[X_{D,x}^2]}{(\mathbb{E}[X_{D,x}])^2} - 1 \\ &= \frac{2\boldsymbol{\pi}_D(D - xQ)^{-2}\mathbf{e}}{(\mathbb{E}[X_{D,x}])^2} - 1 \\ &= 2\lambda\boldsymbol{\pi}(D - xQ)^{-1}\mathbf{e} - 1 \\ &\geq 2 - 1 \\ &= 1.\end{aligned} \tag{3.2}$$

By Theorem 3.1, the equality holds in (3.2) if and only if $D = \lambda I$, i.e. when X_x and $X_{D,x}$ are exponential random variables with parameter λ. □

Define $\mathbb{E}[X_0] = \boldsymbol{\pi} D^{-1}\mathbf{e}$ if $d_i > 0$ for $i = 1, 2, \ldots, m$, and ∞ otherwise. We note that, with the definition of $\mathbb{E}[X_0]$, Theorems 3.1 and 3.2 hold for $x \geq 0$.

Theorem 3.3. *We have $\mathbb{E}[X_x] \leq \mathbb{E}[X_0]$ (i.e. $\boldsymbol{\pi}(D - xQ)^{-1}\mathbf{e} \leq \boldsymbol{\pi} D^{-1}\mathbf{e}$), for $x > 0$.*

Proof. If at least one of $\{d_1, d_2, \ldots, d_m\}$ is zero, then $\mathbb{E}[X_0] = \infty$ and the desired result is obtained.

Now suppose that all $\{d_1, d_2, \ldots, d_m\}$ are positive. Then $\mathbb{E}[X_x] \leq \mathbb{E}[X_0]$ is equivalent to

$$\begin{aligned}\boldsymbol{\pi}(D - xQ)^{-1}\mathbf{e} - \boldsymbol{\pi} D^{-1}\mathbf{e} &= \boldsymbol{\pi}(D - xQ)^{-1}(D - (D - xQ))D^{-1}\mathbf{e} \\ &= x\boldsymbol{\pi}(D - xQ)^{-1}QD^{-1}\mathbf{e} \\ &\leq 0.\end{aligned} \tag{3.3}$$

Let $Q_1 = D\,\mathrm{diag}\,(\boldsymbol{\pi})(D - xQ)^{-1}Q$. It can be verified that $Q_1\mathbf{e} = 0$ and $\mathbf{e}'Q_1 = 0$, noting that $D\,\mathrm{diag}\,(\boldsymbol{\pi}) = \mathrm{diag}\,(\boldsymbol{\pi})D$. Then $P_1 = I + Q_1/c_1$ is a doubly stochastic matrix for $c_1 \geq \max\{-(Q_1)_{i,i},\ i = 1, 2, \ldots, m\}$. Then (3.3) can be rewritten as

$$\begin{aligned} x\mathbf{e}'D^{-1}D\,\mathrm{diag}\,(\boldsymbol{\pi})(D - xQ)^{-1}QD^{-1}\mathbf{e} &= xc_1(D^{-1}\mathbf{e})'(P_1 - I)D^{-1}\mathbf{e} \\ &\leq xc_1(D^{-1}\mathbf{e})'(I - I)D^{-1}\mathbf{e} \\ &= 0, \end{aligned}$$

where the inequality can be proved by using (2.5). The proof is complete. □

Combining Theorems 3.1 and 3.3, we obtain $1/\lambda \leq \mathbb{E}[X_x] \leq \mathbb{E}[X_0]$, for $x \geq 0$.

For some special cases, we can further show that $\mathbb{E}[X_x]$ and SCV $(X_{D,x})$ are decreasing in x, and SCV $(X_x) \geq 1$. A CTMC is called time-reversible if its infinitesimal generator Q satisfies $Q = (\mathrm{diag}\,(\boldsymbol{\pi}))^{-1}Q'\,\mathrm{diag}\,(\boldsymbol{\pi})$ (see [20, Section 6.6]).

Theorem 3.4. *If the CTMC with infinitesimal generator Q is time-reversible, then* (i) $\mathbb{E}[X_x]$ *and* SCV $(X_{D,x})$ *are non-increasing in x, for $x > 0$, and* (ii) SCV $(X_x) \geq 1$, *for $x > 0$.*

Proof. To prove part (i), we show that the derivative of $\mathbb{E}[X_x]$ is non-positive for $x > 0$. By routine calculations, we can obtain

$$\frac{\mathrm{d}\mathbb{E}[X_x]}{\mathrm{d}x} = \boldsymbol{\pi}(D - xQ)^{-1}Q(D - xQ)^{-1}\mathbf{e} \quad \text{for } x > 0.$$

We rewrite the above expression as follows:

$$\frac{\mathrm{d}\mathbb{E}[X_x]}{\mathrm{d}x} = \mathbf{e}'(D - x\ \mathrm{diag}\,(\boldsymbol{\pi})Q(\mathrm{diag}\,(\boldsymbol{\pi}))^{-1})^{-1}(\mathrm{diag}\,(\boldsymbol{\pi})Q)(D - xQ)^{-1}\mathbf{e}.$$

Note that $\boldsymbol{\pi} = \mathbf{e}'\,\mathrm{diag}\,(\boldsymbol{\pi})$ and $\mathrm{diag}\,(\boldsymbol{\pi})D(\mathrm{diag}\,(\boldsymbol{\pi}))^{-1} = D$. Let

$$\mathbf{u} = \boldsymbol{\pi}(D - xQ)^{-1}(\mathrm{diag}\,(\boldsymbol{\pi}))^{-1} = \mathbf{e}'(D - x\,\mathrm{diag}\,(\boldsymbol{\pi})Q(\mathrm{diag}\,(\boldsymbol{\pi}))^{-1})^{-1}$$

and $\mathbf{v} = (D - xQ)^{-1}\mathbf{e}$. Since Q is time-reversible, we obtain $\mathbf{v} = \mathbf{u}'$. Using Lemma 2.1, similar to the proof of Lemma 2.2, we obtain

$$\frac{\mathrm{d}\mathbb{E}[X_x]}{\mathrm{d}x} = c_Q\mathbf{u}(P_Q - I)\mathbf{u}' \leq 0 \quad \text{for } x > 0.$$

Therefore $\mathbb{E}[X_x]$ is non-increasing in x, and so is SCV $(X_{D,x})$ by (3.1).

To prove part (ii), by (2.1), it is sufficient to show $\boldsymbol{\pi}(D - xQ)^{-2}\mathbf{e} \geq (\boldsymbol{\pi}(D - xQ)^{-1}\mathbf{e})^2$. Using vectors $\mathbf{u}$ and $\mathbf{v}$ defined above, the inequality can be rewritten as

$$\boldsymbol{\pi}(D - xQ)^{-2}\mathbf{e} = \mathbf{u}\,\mathrm{diag}\,(\boldsymbol{\pi})\mathbf{v} \geq \mathbf{u}\boldsymbol{\pi}'\boldsymbol{\pi}\mathbf{v} = (\boldsymbol{\pi}(D - xQ)^{-1}\mathbf{e})^2.$$

If Q is time-reversible, then $\mathbf{v} = \mathbf{u}'$, and the above inequality is equivalent to $\sum_{i=1}^m \pi_i u_i^2 \geq (\sum_{i=1}^m u_i\pi_i)^2$. This last inequality can be proved as follows:

$$\left(\sum_{i=1}^m u_i\pi_i\right)^2 = \sum_{i=1}^m u_i^2\pi_i^2 + \sum_{i<j}\pi_i\pi_j(u_iu_j + u_ju_i) \leq \sum_{i=1}^m u_i^2\pi_i^2 + \sum_{i<j}\pi_i\pi_j(u_i^2 + u_j^2) = \sum_{i=1}^m \pi_i u_i^2.$$

This proves the theorem. □

Corollary 3.1. *Results in Theorem 3.4 hold for X_x and $X_{D,x}$ with the following types of Q matrix:* (i) $m = 2$, (ii) *the CTMC with Q is a (truncated) birth and death process, and* (iii) *Q is symmetric.*

4. Applications, extensions, and future research

4.1. Applications to the MMPPs

Markovian arrival processes (MAP) were introduced as a counting processes by Marcel Neuts [15] in 1979. An MAP has a matrix representation (D_0, D_1), where D_0 is a matrix for the transition rates without an arrival/event, and D_1 is a matrix for the transition rates with an arrival. The MAP is associated with a underlying Markov chain with infinitesimal generator $D_0 + D_1$. We refer to Lucantoni [11] for more on the definition and notation of MAPs.

An MMPP is a special Markovian arrival process with matrix representation $(D_0 = Q - D, D_1 = D)$, where Q is an irreducible infinitesimal generator of order m and for transitions without arrivals, and D is a non-negative diagonal matrix for the arrival rates. Within each state of the underlying Markov chain Q, the arrivals form a Poisson process. For instance, within state i, arrivals form a Poisson process with parameter d_i. Thus the arrival rate is modulated by the underlying Markov chain, which is why the counting process is called a Markov modulated Poisson process. Since Q is irreducible and m is finite, the stationary distribution of the underlying Markov chain Q exists and is denoted by $\boldsymbol{\pi}$. Then the time-stationary inter-arrival times have a PH-distribution with PH-representation $(\boldsymbol{\pi}, Q - D)$, which is the random variable X defined in Section 2. The event-stationary inter-arrival times have a PH-distribution with PH-representation $(\boldsymbol{\pi}_D, Q - D)$, which is the random variable X_D defined in Section 2. Consequently, Theorems 3.1 and 3.2 apply to the two types of inter-arrival times for MMPPs. Those results imply that (i) on average, the time-stationary inter-arrival time is longer than the event-stationary inter-event time, and (ii) the SCV of the event-stationary inter-event time is greater than or equal to unity. The only MMPP whose SCV is unity is the Poisson process. The insight is that MMPPs are not suitable for modeling input processes whose SCV is less than one, which is useful for engineers and researchers to select proper point processes in stochastic modeling. Since MMPPs have been used in stochastic modeling in science and engineering for more than half a century, the results are theoretically and practically interesting.

4.2. Extensions to the discrete-time PH-distributions

All results obtained in this paper for the continuous-time case can be translated to the discrete-time case. We consider a discrete-time PH-distribution with PH-representation $(\boldsymbol{\pi}, P)$, where P is an irreducible substochastic matrix and $\boldsymbol{\pi}$ is the stationary distribution of a discrete-time Markov chain with transition probability matrix $P + \mathrm{diag}\,(\mathbf{p}_0)$, where $\mathbf{p}_0 = (I - P)\mathbf{e}$. Define $Q = P + \mathrm{diag}\,(\mathbf{p}_0) - I$ and $D = \mathrm{diag}\,(\mathbf{p}_0)$. It is easy to verify that $\boldsymbol{\pi} Q = 0$. We define two discrete-time PH-distributions X_d and $X_{D,d}$ with PH-representations $(\boldsymbol{\pi}, P)$ and $(\boldsymbol{\pi}_D, P)$, respectively. Recall that X and X_D are defined in Section 2 with PH-representations $(\boldsymbol{\pi}, Q - D)$ and $(\boldsymbol{\pi}_D, Q - D)$, respectively. It is easy to verify that $\mathbb{E}[X] = \mathbb{E}[X_d]$, $\mathbb{E}[X^2] = \mathbb{E}[X_d^2]$, $\mathbb{E}[X_D] = \mathbb{E}[X_{D,d}]$, and $\mathbb{E}[X_D^2] = \mathbb{E}[X_{D,d}^2]$. Consequently, means and SCVs of the discrete case have the same bounds as the continuous case. For instance, the SCV of $X_{D,d}$ with $(\boldsymbol{\pi}_D, P)$ is greater than or equal to unity.

4.3. Future research

Numerically, all continuous-time MMPP type PH-distributions have the following properties:

- X_x is stochastically larger than $X_{D,x}$ and X_x is stochastically larger than the exponential distribution with parameter $\lambda = \boldsymbol{\pi} D\mathbf{e}$,

- SCV $(X_x) \geq 1$, for $x > 0$,
- $\mathbb{E}[X_x]$ and SCV $(X_{D,x})$ are convex and non-increasing in x, for $x \geq 0$,
- $\mathbb{E}[X_x]$ is convex in $(d_1, d_2, \ldots, d_m)$.

These are interesting questions for future studies.

Acknowledgements

I would like to thank two reviewers for their constructive comments and suggestions. I also want to thank Krishna Sabareesh Rajangom and Haoran Wu for proofreading the manuscript. The research was supported by an NSERC Discovery Grant.

References

[1] Aldous, D. and Shepp, L. (1987). The least variable phase type distribution is Erlang. *Stoch. Models* **3**, 467–473.
[2] Asanjarani, A. and Nazarathy, Y. (2019). Stationary Markovian arrival processes, results and open problems. Available at arXiv:1905.01736v1.
[3] Barlow, R. E. and Proschan, F. (1996). *Mathematical Theory of Reliability*. SIAM, Philadelphia.
[4] Buchholz, P., Kriege, J. and Felko, I. (2014). *Input Modeling with Phase-Type Distributions and Markov Models: Theory and Applications*. Springer.
[5] Dan, L. and Neuts, M. F. (1991). Counter-examples involving Markovian arrival processes. *Stoch. Models* **7**, 499–509.
[6] Fischer, W. and Meier-Hellstern, K. (1993). The Markov-modulated Poisson process (MMPP) cookbook. *Performance Evaluation* **18**, 149–171.
[7] Grandell, J. (2006). *Doubly Stochastic Poisson Processes* (Lecture Notes Math. **529**). Springer.
[8] He, Q.-M. (2014). *Fundamentals of Matrix-Analytic Methods*. Springer.
[9] He, Q.-M., Horváth, G., Horváth, I. and Telek, M. (2019). Moment bounds of PH distributions with infinite or finite support based on the steepest increase property. *Adv. Appl. Prob.* **51**, 168–183.
[10] Latouche, G. and Ramaswami, V. (2011). *Introduction to Matrix Analytic Methods in Stochastic Modeling* (ASA-SIAM Series on Statistics and Applied Probability), 2nd edn. SIAM, Philadelphia.
[11] Lucantoni, D. M. (1991). New results on the single server queue with a batch Markovian arrival process. *Stoch. Models* **7**, 1–46.
[12] Marshall, A. W., Olkin, I. and Arnold, B. C. (2011). *Inequalities: Theory of Majorization and its Applications* (Mathematics in Science and Engineering **143**), 2nd edn. Springer.
[13] Minc, H. (1988). *Non-Negative Matrices*. John Wiley, New York.
[14] Neuts, M. F. (1975). Probability distributions of phase type. In *Liber Amicorum Prof. Emeritus H. Florin*, pp. 173–206. University of Louvain, Belgium.
[15] Neuts, M. F. (1979). A versatile Markovian point process. *J. Appl. Prob.* **16**, 764–779.
[16] Neuts, M. F. (1981). *Matrix-Geometric Solution in Stochastic Model: An Algorithmic Application*. The Johns Hopkins University Press, Baltimore.
[17] O'Cinneide, C. A. (1990). Characterization of phase-type distributions. *Stoch. Models* **6**, 1–57.
[18] O'Cinneide, C. A. (1999). Phase-type distributions: open problems and a few properties. *Stoch. Models* **15**, 731–757.
[19] Olsson, F. and Hössjer, O. (2015). Equilibrium distributions and simulation methods for age structured populations. *Math. Biosci.* **268**, 45–51.
[20] Ross, S. M. (2014). *Introduction to Probability Models*. Academic Press.
[21] Xing, Y., Li, L., Bi, Z., Wilamowska-Korsak, M. and Zhang, L. (2013). Operations research (OR) in service industries: a comprehensive review. *Systems Res. Behavioral Sci.* **30**, 300–353.