Adv. Appl. Prob. **51,** 168–183 (2019)
*doi:*10.1017/apr.2019.7

MOMENT BOUNDS OF PH DISTRIBUTIONS WITH INFINITE OR FINITE SUPPORT BASED ON THE STEEPEST INCREASE PROPERTY

QI-MING HE,* *University of Waterloo*
GÁBOR HORVÁTH,** *** *Budapest University of Technology and Economics*
ILLÉS HORVÁTH,**** *MTA-BME Information Systems Research Group*
MIKLÓS TELEK,** ***** *Budapest University of Technology and Economics*

Abstract

The steepest increase property of phase-type (PH) distributions was first proposed in O'Cinneide (1999) and proved in O'Cinneide (1999) and Yao (2002), but since then has received little attention in the research community. In this work we demonstrate that the steepest increase property can be applied for proving previously unknown moment bounds of PH distributions with infinite or finite support. Of special interest are moment bounds free of specific PH representations except the size of the representation. For PH distributions with infinite support, it is shown that such a PH distribution is stochastically smaller than or equal to an Erlang distribution of the same size. For PH distributions with finite support, a class of distributions which was introduced and investigated in Ramaswami and Viswanath (2014), it is shown that the squared coefficient of variation of a PH distribution with finite support is greater than or equal to $1/(m(m+2))$, where m is the size of its PH representation.

Keywords: Phase-type distribution; infinite support; finite support; moment bound

2010 Mathematics Subject Classification: Primary 60E15;
Secondary 62E15

1. Introduction

Phase-type (PH) distributions were introduced in 1975 by Neuts [3] for the study of queueing systems. Since then, PH distributions have been a subject of research, and have found applications in many areas of applied probability. For example, a succession of papers by O'Cinneide ([5]–[8]) revealed some fundamental properties of PH distributions. In [1] it was shown that the squared coefficient of variation (SCV) of a PH distribution with a PH representation of size m is greater than or equal to $1/m$. In [11] the minimal SCV of discrete PH distributions was found. In [2] stochastic comparison was utilized in the study of PH distributions, which also led to moment bounds of PH distributions. These results not only

Received 19 June 2018; revision received 5 December 2018.
* Postal address: Department of Management Sciences, University of Waterloo, Waterloo, Ontario N2L 3G1, Canada. Email address: q7he@uwaterloo.ca
** Postal address: Department of Networked Systems and Services, Budapest University of Technology and Economics, PO Box 91, 1521 Budapest, Hungary.
*** Email address: ghorvath@hit.bme.hu
**** Postal address: MTA-BME Information Systems Research Group, PO Box 91, 1521 Budapest, Hungary. Email address: horvath.illes.antal@gmail.com
***** Email address: telek@hit.bme.hu

deepen our understanding of PH distributions and PH representations, but also facilitate the applications of PH distributions significantly.

The steepest increase property of PH distributions was first proposed in O'Cinneide [8] and proved by O'Cinneide [8] and Yao [12]. We think that it is a very interesting property of PH distributions, which received little attention in the literature. In this paper we apply the property to find moment bounds of PH distributions, which demonstrates the usefulness of the property.

PH distributions with finite support were introduced and investigated in [9]. This generalization extends the applications of PH distributions significantly.

In this paper we find a number of stochastic and moment bounds for PH distributions with finite and infinite support. Many of the moment bounds depend only on the size of PH representations and the eigenvalue with the largest real part of PH generators. A highlight of this paper is that the SCV of a bounded PH distribution with a PH representation of size m is greater than or equal to $1/(m(m+2))$. The moment bounds of PH distributions reveal fundamental properties of such probability distributions, e.g. they indicate which type of general distributions can be closely approximated. The results are useful in, but not limited to,

- finding moment bounds of performance measures for stochastic models, and
- selecting PH representations in the parameter estimation of PH distributions.

The rest of the paper is organized as follows. In Section 2 we introduce PH distributions with infinite support together with the steepest increase property and with some of their moment bounds. In Section 3 we introduce PH distributions with finite support and study their moment bounds. Section 4 concludes the paper.

2. PH distributions with infinite support

In this section we first define PH distributions and their corresponding PH representations. Then we review the so-called 'steepest increase lemma' (see [12]) on the density function of (ordinary) PH distributions, since this turned out to be useful to derive the moment bounds. We further extend the 'steepest increase lemma' and derive new moment bounds for ordinary PH distributions.

A nonnegative random variable $\mathcal{Y}$ has a PH distribution if it is the absorption time in a finite-state, continuous-time Markov chain (see [4]). We assume that $\mathcal{Y}$ has a PH representation $(\boldsymbol{\alpha}, \boldsymbol{A})$ of size m, where $\boldsymbol{\alpha}$ is the initial distribution of the underlying continuous-time Markov chain and $\boldsymbol{A}$ contains the transition rates among the transient states of the underlying continuous-time Markov chain (referred to as a PH generator or sub-intensity matrix). That is, $\boldsymbol{\alpha}$ is a nonnegative probability vector and $\boldsymbol{A}$ has nonnegative off-diagonal and negative diagonal elements such that the diagonal element dominates each row, $\boldsymbol{A}\mathbf{1} \le 0$ (elementwise), where $\mathbf{1}$ is a column vector of 1s with the appropriate size. Let us denote the density function of $\mathcal{Y}$ by $f(t) = \boldsymbol{\alpha} \mathrm{e}^{\boldsymbol{A}t}(-\boldsymbol{A}\mathbf{1})$ and the cumulative distribution function (CDF) by $F(t) = 1 - \boldsymbol{\alpha}\mathrm{e}^{\boldsymbol{A}t}\mathbf{1}$ (both for $t \ge 0$). To avoid the trivial case $\mathcal{Y} \equiv 0$, we assume that $0 < \boldsymbol{\alpha}\mathbf{1} \le 1$. We also note that, by [6], $f(t) > 0$ for all $t > 0$.

The steepest increase conjecture was first published and partially proven by O'Cinneide [8]. Its complete proof was given by Yao [12]. The steepest increase lemma states that $f'(t)/f(t) \le (m-1)/t$ for $t > 0$, or, equivalently, $f(t)/t^{m-1}$ is decreasing in t for $t > 0$. Next, we present and show a 'sharp' form of the steepest increase lemma.

Lemma 1. *For a PH distribution with representation* $(\boldsymbol{\alpha}, \boldsymbol{A})$ *of size m and with density function* $f(t)$*, we have*

$$\frac{f'(t)}{f(t)} \leq \frac{m-1}{t} - \lambda \quad \textit{for } t > 0, \tag{1}$$

where λ *is the absolute value of the eigenvalue of* $\boldsymbol{A}$ *with the largest real part (which is real and negative). Here* $-\lambda$ *is also referred to as the dominant eigenvalue of matrix* $\boldsymbol{A}$*. In (1) the equality holds when* $\mathcal{Y}$ *is Erlang*(m, λ)*-distributed (i.e.* $\mathcal{Y}$ *is the sum of m independent exponential random variables with the same parameter* λ*).*

Proof. The inequality just before Equation (12) of [8], combined with the proof in [12], leads to $((m-1-\lambda)\boldsymbol{I} - \boldsymbol{A})\mathrm{e}^{\boldsymbol{A}} \geq 0$ for PH generator $\boldsymbol{A}$, where $\boldsymbol{I}$ is the identity matrix. For PH generator $\boldsymbol{A}$, $\boldsymbol{A}t$ is also a PH generator for $t > 0$. The eigenvalue with the largest real part of $\boldsymbol{A}t$ is $-\lambda t$ for $t > 0$. Setting $\boldsymbol{A} =: \boldsymbol{A}t$ in the inequality, we obtain, for $t > 0$,

$$((m-1-\lambda t)\boldsymbol{I} - \boldsymbol{A}t)\mathrm{e}^{\boldsymbol{A}t} \geq 0,$$

which leads to

$$\mathrm{e}^{\boldsymbol{A}t}\boldsymbol{A}t \leq (m-1-\lambda t)\mathrm{e}^{\boldsymbol{A}t}. \tag{2}$$

Pre-multiplying and post-multiplying both sides of (2) by $\boldsymbol{\alpha}$ and $-\boldsymbol{A}\mathbf{1}$, respectively, we obtain $f'(t)t \leq (m-1-\lambda t)f(t)$, which proves the lemma. □

Similar to Lemma 1, $-\lambda$ stands for the dominant eigenvalue of $\boldsymbol{A}$ in the sequel.

Equation (1) can be written in several alternative forms. For $t \geq 0$,

$$\begin{aligned} tf'(t) &\leq (m-1)f(t) - \lambda t f(t), \\ \frac{\mathrm{d}}{\mathrm{d}t}(tf(t)) &\leq (m-\lambda t)f(t), \end{aligned} \tag{3}$$

and utilizing the fact that $\lambda > 0$, we also have, for $t > 0$,

$$\frac{\mathrm{d}}{\mathrm{d}t}(tf(t)) < mf(t). \tag{4}$$

Lemma 2. *For* $\mathcal{Y}$ *with PH generator* $\boldsymbol{A}$ *of size m and dominant eigenvalue* $-\lambda$*, we have, for* $0 < t_1 < t_2$,

$$\frac{f(t_2)}{f(t_1)} \leq \frac{f_{(m,\lambda)}(t_2)}{f_{(m,\lambda)}(t_1)},$$

where $f_{(m,\lambda)}(t) = \lambda^m t^{m-1}\mathrm{e}^{-\lambda t}/(m-1)!$ *is the density function of the Erlang random variable with parameters* (m, λ)*.*

Proof. For $t > 0$, (1) can be written as,

$$(m-1)(\ln(t))' - (\lambda t)' \geq (\ln(f(t)))'. \tag{5}$$

Integrating both sides of (5) from $t_1 > 0$ to $t_2 > t_1$ yields

$$\ln\left(\frac{(t_2/t_1)^{m-1}f(t_1)}{f(t_2)}\right) \geq \lambda(t_2 - t_1), \tag{6}$$

where $f(t) > 0$ for all $t > 0$ ensures that the integration can be done properly. Taking the exponent of both sides of (6), we obtain

$$f(t_2) \leq \frac{f(t_1)}{t_1^{m-1}} \mathrm{e}^{\lambda t_1} t_2^{m-1} \mathrm{e}^{-\lambda t_2} = f(t_1) \frac{f_{(m,\lambda)}(t_2)}{f_{(m,\lambda)}(t_1)} \quad \text{for } t_2 > t_1 > 0,$$

which leads to the desired result. □

Lemma 2 implies that $f(t)/f_{(m,\lambda)}(t)$ is decreasing in t, which is a generalization of the monotonicity of the function $f(t)/t^{m-1}$. Furthermore, Lemma 2 leads to the following stochastic comparison result between $\mathcal{Y}$ and the Erlang random variable $\mathcal{X}_{(m,\lambda)}$ with parameters (m, λ). A random variable Y is stochastically smaller than or equal to a random variable X, denoted as $Y \leq_d X$, if $F_Y(t) \geq F_X(t)$ holds for all real t (see [10]).

Corollary 2.1. *The PH-distributed random variable $\mathcal{Y}$ (of size m and with dominant eigenvalue $-\lambda$) is stochastically smaller than or equal to $\mathcal{X}_{(m,\lambda)}$. Consequently, we have, for $n \geq 1$,*

$$\mathbb{E}[\mathcal{Y}^n] \leq \mathbb{E}[\mathcal{X}^n_{(m,\lambda)}] = \frac{(m+n-1)!}{(m-1)!\,\lambda^n}.$$

Proof. Since both $f(t)$ and $f_{(m,\lambda)}(t)$ are density functions on $[0, \infty)$, there must be at least one intersection in $(0, \infty)$. If t^* is an intersection (i.e. $f(t^*) = f_{(m,\lambda)}(t^*)$), by Lemma 2, we must have $f(t) \leq f_{(m,\lambda)}(t)$ for $t > t^*$ and $f(t) \geq f_{(m,\lambda)}(t)$ for $t < t^*$. Thus, there are only three possible cases:

- $f(t)$ and $f_{(m,\lambda)}(t)$ are identical;
- $f(t)$ and $f_{(m,\lambda)}(t)$ have exactly two intersections, $t = 0$ and $t = t^*$;
- $f(t)$ and $f_{(m,\lambda)}(t)$ have exactly one intersection, $t = t^*$.

Then we must have $f(t) \geq f_{(m,\lambda)}(t)$ for $0 < t \leq t^*$ and $f(t) \leq f_{(m,\lambda)}(t)$ for $t \geq t^*$ (> 0), which leads to $F(t) \geq F_{\mathcal{X}_{(m,\lambda)}}(t)$, where $F_{\mathcal{X}_{(m,\lambda)}}(t)$ is the CDF of $\mathcal{X}_{(m,\lambda)}$ for $0 < t \leq t^*$, and $1 - F(t) \leq 1 - F_{\mathcal{X}_{(m,\lambda)}}(t)$ for $t > t^*$. Consequently, we obtain $F(t) \geq F_{\mathcal{X}_{(m,\lambda)}}(t)$ for $t > 0$, which leads to the first result. All the moment bounds can be obtained from $\mathcal{Y} \leq_d \mathcal{X}_{(m,\lambda)}$ directly. □

A random variable X is smaller in the mean residual life than a random variable Y, denoted as $X \leq_c Y$, if $\mathbb{E}\,[\max\{0, X - t\}] \leq \mathbb{E}\,[\max\{0, Y - t\}]$ holds for all real t (see [10]). By [7], it is known that the Erlang distribution with parameters $(m, m/\mathbb{E}\,[\mathcal{Y}]\,)$ is smaller in the mean residual life than $\mathcal{Y}$, which has the same mean. Then the moments of $\mathcal{Y}$ are bounded from above and below as follows:

$$\frac{(m+n-1)!}{(m-1)!} \left(\frac{\mathbb{E}\,[\mathcal{Y}]}{m} \right)^n \leq \mathbb{E}[\mathcal{Y}^n] \leq \frac{(m+n-1)!}{(m-1)!} \frac{1}{\lambda^n}.$$

A by-product of the above inequalities is that $\mathbb{E}[\mathcal{Y}] = m/\lambda$ if and only if $\mathcal{Y}$ has an Erlang distribution, which can also be obtained from inequality (3) directly.

Based on Lemma 1, the next lemma refines the upper bound for the $(n+1)$th moment based on the nth moment, which gives Corollary 2.1 an alternative proof.

Lemma 3. *For $n = 0, 1, \ldots,$ the $(n+1)$th moment of $\mathcal{Y}$ (of size m and with dominant eigenvalue $-\lambda$) is bounded by*

$$\mathbb{E}[\mathcal{Y}^{n+1}] \leq \frac{m+n}{\lambda} \mathbb{E}[\mathcal{Y}^n], \tag{7}$$

and the equality holds when $\mathcal{Y} = \mathcal{X}_{(m,\lambda)}$.

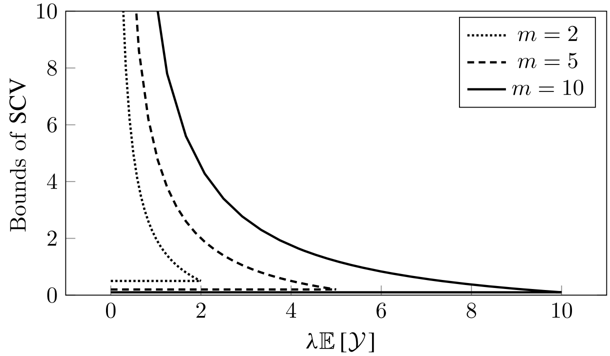

FIGURE 1: Bounds of the SCV for ordinary PH distributions. The lower bound is the well known $1/m$ bound provided in [1] and [7], while the upper bound is provided based on the steepest increase property.

Proof. Multiplying both sides of (3) by t^n and integrating from 0 to ∞ gives the following identities for the left-hand side (LHS) and the right-hand side (RHS):

$$\text{LHS} = \int_{t=0}^{\infty} t^n \,\mathrm{d}(tf(t)) = [t^{n+1}f(t)]_0^{\infty} - \int_{t=0}^{\infty} tf(t)\,\mathrm{d}t^n = -n\int_{t=0}^{\infty} tf(t)t^{n-1}\,\mathrm{d}t = -n\mathbb{E}[\mathcal{Y}^n];$$

$$\begin{aligned}\text{RHS} &= \int_{t=0}^{\infty} t^n(m-\lambda t)f(t)\,\mathrm{d}t \\ &= m\int_{t=0}^{\infty} t^n f(t)\,\mathrm{d}t - \lambda\int_{t=0}^{\infty} t^{n+1}f(t)\,\mathrm{d}t \\ &= m\mathbb{E}[\mathcal{Y}^n] - \lambda\mathbb{E}[\mathcal{Y}^{n+1}].\end{aligned}$$

Thus

$$-n\mathbb{E}[\mathcal{Y}^n] \le m\mathbb{E}[\mathcal{Y}^n] - \lambda\mathbb{E}[\mathcal{Y}^{n+1}].$$

When $\mathcal{Y}$ is Erlang(λ, m) distributed, the equality in (7) comes from the fact that Lemma 1 gives equality for an Erlang distribution for all $t > 0$. □

Applying Lemma 3 for $n = 0$ and $n = 1$ enables us to derive the following upper bounds on the mean $\mathbb{E}[\mathcal{Y}]$ and the squared coefficient of variation $\text{SCV}_{\mathcal{Y}} = \mathbb{E}[\mathcal{Y}^2]/\mathbb{E}[\mathcal{Y}]^2 - 1$:

$$\mathbb{E}[\mathcal{Y}] \le \frac{m}{\lambda}, \tag{8}$$

$$\text{SCV}_{\mathcal{Y}} \le \frac{m+1-\lambda\mathbb{E}[\mathcal{Y}]}{\lambda\mathbb{E}[\mathcal{Y}]}. \tag{9}$$

Interestingly, (9) gives an upper bound for $\text{SCV}_{\mathcal{Y}}$, while the lower bound for $SCV_{\mathcal{Y}}$ is much more widely known as $\text{SCV}_{\mathcal{Y}} \ge 1/m$. Hence, we have (see Figure 1)

$$\frac{1}{m} \le \text{SCV}_{\mathcal{Y}} \le \frac{m+1}{\lambda\mathbb{E}[\mathcal{Y}]} - 1.$$

3. PH distributions with finite support

PH distributions with finite support were introduced in [9], where three classes of finite support distributions were considered (matrix exponential densities from the lower bound to

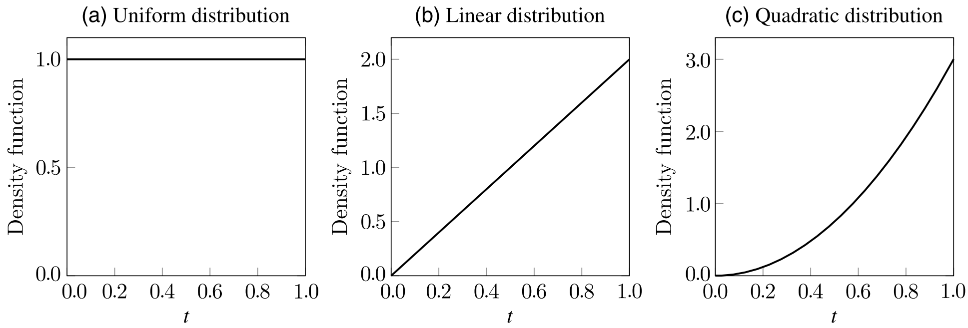

FIGURE 2: Some interesting members of the FTPH_1 class.

the upper bound, and from the upper bound to the lower bound, and a convex combination of the two). Instead, we define $\mathcal{Z}$ to have distribution $b + (\mathcal{Y} \mid \mathcal{Y} < T)$. The support of $\mathcal{Z}$ is thus $[b, B)$ with $b < B$ and $B = b + T$. Recall that $\mathcal{Y}$ is the ordinary (or infinite support) PH distribution with density function $f(t) = \boldsymbol{\alpha} \mathrm{e}^{\boldsymbol{A}t}(-\boldsymbol{A})\mathbf{1}$. Then the density function of $\mathcal{Z}$ is given by

$$f_{\mathcal{Z}}(t) = \frac{\boldsymbol{\alpha} \mathrm{e}^{\boldsymbol{A}(t-b)}(-\boldsymbol{A})\mathbf{1}}{1 - \boldsymbol{\alpha} \mathrm{e}^{\boldsymbol{A}T}\mathbf{1}}$$

for $t \in [b, B)$, and $f_{\mathcal{Z}}(t) = 0$ for $t \notin [b, B)$.

We denote this class of finite PH distributions by FTPH_1 (with some further similar classes FTPH_2 and FTPH_3 in mind: FTPH_2 is defined as $\mathcal{Z}_2 = B - (\mathcal{Y} \mid \mathcal{Y} < T)$ and FTPH_3 as the convex combination of $\mathcal{Z}$ and $\mathcal{Z}_2$; these are subject to future investigations). In this paper we primarily focus on the moments of FTPH_1 distributions.

Although FTPH_1 is obtained by just a simple truncation of an ordinary PH distribution, it has some very interesting members. A truncated exponential distribution with a very small intensity parameter leads to a uniform distribution (Figure 2(a)), since

$$\lim_{\lambda \to 0} f_{\mathcal{Z}}^{\mathrm{Exp}}(t) = \lim_{\lambda \to 0} \frac{\lambda \mathrm{e}^{\lambda t}}{1 - \mathrm{e}^{-\lambda}} = 1.$$

Similarly, a truncated Erlang-N distribution with very small intensity gives

$$\lim_{\lambda \to 0} f_{\mathcal{Z}}^{\mathrm{Erl}-N}(t) = N\, t^{N-1},$$

yielding linear, quadratic, and cubic distributions (see Figure 2(b) and 2(c)).

The importance of these special FTPH_1 members is that density functions of such shapes are notoriously difficult to capture by ordinary PH distributions. Ordinary PH distributions with a limited number of phases purely approximate them. In practical computations the $\lambda \to 0$ limit also causes difficulties, but easily computable small positive λ values give reasonably good approximations. These approximation issues and parameter estimations of PH distributions with finite support will be addressed in a separate paper.

A general formula for computing the moments for PH distributions with finite support can be obtained as follows.

Theorem 1. *The nth moment of an* FTPH_1 *random variable,* $\mathcal{Z}$ *with parameters* $(\boldsymbol{\alpha}, \boldsymbol{A})$ *over interval* $[b, B)$*, is expressed by*

$$\mathbb{E}[\mathcal{Z}^n] = \frac{b^n(1-\boldsymbol{\alpha}\mathbf{1}) + \sum_{d=0}^{n}\binom{n}{d}d!\,\boldsymbol{\alpha}(-\boldsymbol{A})^{-d}(b^{n-d}\boldsymbol{I} - T^{n-d}\mathrm{e}^{\boldsymbol{A}T})\mathbf{1}}{1-\boldsymbol{\alpha}\mathrm{e}^{\boldsymbol{A}T}\mathbf{1}},$$

where $T = B - b$.

Proof. The theorem is proved by routine calculations. □

3.1. Moment bounds for the case with $b = 0$

In this subsection we assume that $b = 0$, which is extended to $b > 0$ in the next subsection. A random variable in FTPH_1 is denoted as $\mathcal{W} = \mathcal{Y} \mid \mathcal{Y} < T$. We derive and prove the lower and upper bounds for the moments of $\mathcal{W}$. For $i = 0, 1, \ldots,$ we introduce the notation $E_i(T) = \mathbb{E}[\mathbb{I}_{\{\mathcal{Y}<T\}}\mathcal{Y}^i] = \int_{t=0}^{T} t^i f(t)\,\mathrm{d}t$, where $\mathbb{I}_{\{a\}}$ denotes the indicator of a. Then $\mathbb{E}[\mathcal{W}^i]$ can be written as, for $i = 0, 1, \ldots,$

$$\mathbb{E}[\mathcal{W}^i] = \frac{E_i(T)}{E_0(T)}.$$

The next lemma is similar to Lemma 3 (for ordinary PH distributions), but it does not depend on the dominant eigenvalue $-\lambda$.

Lemma 4. *The moments of an* FTPH_1 *random variable,* $\mathcal{W}$*, with support on* $(0, T)$*, satisfies*

$$\mathbb{E}[\mathcal{W}^n] \le \frac{(m+n-1)T}{m+n}\mathbb{E}[\mathcal{W}^{n-1}]. \tag{10}$$

Proof. Multiplying both sides of (4) by $t^{n-1}(T-t)$ and integrating from 0 to T gives the following identities for the left-hand side:

$$\begin{aligned}
\text{LHS} &= \int_{t=0}^{T} t^{n-1}(T-t)\frac{\mathrm{d}}{\mathrm{d}t}(tf(t))\,\mathrm{d}t \\
&= -\int_{t=0}^{T} tf(t)\,\mathrm{d}(t^{n-1}(T-t)) \\
&= \int_{t=0}^{T} (nt^n - (n-1)Tt^{n-1})f(t)\,\mathrm{d}t \\
&= nE_n(T) - (n-1)TE_{n-1}(T).
\end{aligned}$$

We use Stieltjes integration by parts in the first step,

$$\int_a^b f(x)\,\mathrm{d}g(x) = f(b)g(b) - f(a)g(a) - \int_a^b g(x)\,\mathrm{d}f(x),$$

and

$$\int_a^b g(x)\,\mathrm{d}f(x) = \int_a^b g(x)f'(x)\,\mathrm{d}x$$

in the second step. For the right-hand side, we have

$$
\begin{aligned}
\mathrm{RHS} &= \int_{t=0}^{T} t^{n-1}(T-t)mf(t)\,\mathrm{d}t \\
&= \int_{t=0}^{T} mTt^{n-1}f(t)\,\mathrm{d}t - m\int_{t=0}^{T} t^{n}f(t)\,\mathrm{d}t \\
&= mTE_{n-1}(T) - mE_n(T),
\end{aligned}
$$

from which

$$
\begin{aligned}
nE_n(T) - (n-1)TE_{n-1}(T) &\le mTE_{n-1}(T) - mE_n(T), \\
(m+n)E_n(T) &\le (m+n-1)TE_{n-1}(T), \\
\frac{E_n(T)}{E_0(T)} &\le \frac{(m+n-1)T}{m+n}\frac{E_{n-1}(T)}{E_0(T)},
\end{aligned} \tag{11}
$$

which leads to the desired result. □

In the following corollary the upper bound for the nth moment is provided, independent from the lower-order moments.

Corollary 1. *The nth moment of an* FTPH_1 *random variable,* $\mathcal{W}$*, with support on* $(0, T)$*, is bounded by*

$$\mathbb{E}[\mathcal{W}^n] \le \frac{mT^n}{m+n}, \tag{12}$$

and the upper bound is strict except when $\mathcal{Y}$ *is Erlang(*λ, m*) distributed and* λ *tends to* 0*. In particular, we have* $\mathbb{E}[\mathcal{W}] \le mT/(m+1)$*, which indicates that no PH distribution with finite support can have a mean close to the upper bound* $T = B$*.*

Proof. Recursively applying (10) for moments $1, \ldots, n$ gives the upper bound. The statement on the equality comes from the fact that the right-hand side of (12) gives equality only when (11) gives equality for $1, \ldots, n$, and it occurs only when $\mathcal{Y}$ is Erlang(λ, m) distributed and λ tends to 0, because (4) gives equality only for an Erlang(λ, m) distribution when λ tends to 0. □

Having derived these moment bounds, we now consider how to reach the extreme values, and the FTPH_1 structure which realizes the moment bounds. According to the next lemma, Erlang distributions play an important role in this respect.

Lemma 5. *The upper bound in Lemma 4 is strict when* $\mathcal{W}$ *equals* 0 *with probability* $p = 1 - (m+n-1)\mu_{n-1}/mT^{n-1}$ *and is truncated Erlang(*λ, m*) distributed with probability* $1-p$ *such that* λ *tends to* 0*, where* $\mu_{n-1} = \mathbb{E}[\mathcal{W}^{n-1}]$*.*

Proof. To prove the statement, we show that compared to the extreme distribution of Lemma 5 any valid change in the probability of the mass at 0, p, decreases the nth moment. First we note that increasing p is not a valid change, because to maintain $\mathbb{E}[\mathcal{W}^{n-1}]$ the $(n-1)$th moment of a strictly positive FTPH_1 needs to be increased above $mT^{n-1}/(m+n-1)$ (which is not possible according to Corollary 1).

Let us now try to decrease rather than increase p. Consider a distribution whose mass at 0 has probability $\hat{p} = p - \Delta\mu_{n-1}$, where Δ is a small positive number. In this case, the $(n-1)$th moment of the strictly positive part, μ^+_{n-1}, is

$$\begin{aligned}\mu^+_{n-1} &= \frac{\mu_{n-1}}{1-\hat{p}} \\ &= \frac{\mu_{n-1}}{1-p+\Delta\mu_{n-1}} \\ &= \frac{\mu_{n-1}}{(m+n-1)\mu_{n-1}/mT^{n-1}+\Delta\mu_{n-1}} \\ &= \frac{mT^{n-1}}{(m+n-1)+\Delta mT^{n-1}} \\ &< \frac{mT^{n-1}}{m+n-1}.\end{aligned}$$

When the $(n-1)$th moment of the strictly positive part, $\mathcal{W}^+$, is μ^+_{n-1}, its nth moment is bounded by (10), and using that, we can write

$$\begin{aligned}&(1-\hat{p})\mathbb{E}[\mathcal{W}^{+n}] \\ &\quad< \underbrace{\mu_{n-1}\left(\frac{(m+n-1)+\Delta mT^{n-1}}{mT^{n-1}}\right)}_{1-\hat{p}} \frac{(m+n-1)T}{m+n} \underbrace{\left(\frac{mT^{n-1}}{(m+n-1)+\Delta mT^{n-1}}\right)}_{\mu^+_{n-1}} \\ &\quad= \frac{(m+n-1)T\mu_{n-1}}{m+n},\end{aligned}$$

where the inequality is strict, because the distribution of $\mathcal{W}^+$ is different from an Erlang(λ, m) distribution such that $\lambda\to 0$, since its $(n-1)$th moment is less than $mT^{n-1}/(m+n-1)$. □

While the moment bounds derived in Lemma 4 are independent of the dominant eigenvalue, the following results provide moment bounds as a function of λ.

Lemma 6. *For $n = 1, 2, \ldots$, the $(n+1)$th moment of $\mathcal{W}$ is bounded by*

$$\begin{aligned}\frac{m+n+\lambda T}{\lambda}\mathbb{E}[\mathcal{W}^n] - \frac{(m+n-1)T}{\lambda}\mathbb{E}[\mathcal{W}^{n-1}] &\le \mathbb{E}[\mathcal{W}^{n+1}] \\ &\le \frac{m+n}{\lambda}\mathbb{E}[\mathcal{W}^n] - \frac{T^{n+1}f(T)}{\lambda E_0(T)}.\end{aligned}$$

Proof. Multiplying both sides of (3) by $(T-t)t^{n-1}$ and integrating from 0 to T gives

$$\begin{aligned}\text{LHS} &= \int_{t=0}^{T} (T-t)t^{n-1}\frac{\mathrm{d}}{\mathrm{d}t}(tf(t))\,\mathrm{d}t \\ &= -\int_{t=0}^{T} tf(t)\,\mathrm{d}(Tt^{n-1}-t^n) \\ &= -\int_{t=0}^{T} tf(t)(T(n-1)t^{n-2}-nt^{n-1})\,\mathrm{d}t \\ &= -T(n-1)E_{n-1}(T)+nE_n(T), \\ \text{RHS} &= \int_{t=0}^{T} (Tt^{n-1}-t^n)(m-\lambda t)f(t)\,\mathrm{d}t \\ &= mTE_{n-1}(T)-mE_n(T)-\lambda TE_n(T)+\lambda E_{n+1}(T),\end{aligned}$$

from which it follows that

$$\begin{aligned} -T(n-1)E_{n-1}(T)+nE_n(T) &\le mTE_{n-1}(T)-(m+\lambda T)E_n(T)+\lambda E_{n+1}(T),\\ (m+n+\lambda T)E_n(T)-(m+n-1)TE_{n-1}(T) &\le \lambda E_{n+1}(T),\\ \frac{m+n+\lambda T}{\lambda}\frac{E_n(T)}{E_0(T)}-\frac{(m+n-1)T}{\lambda}\frac{E_{n-1}(T)}{E_0(T)} &\le \frac{E_{n+1}(T)}{E_0(T)}. \end{aligned}$$

On the other hand, multiplying both sides of (3) by t^n and integrating from 0 to T gives

$$\begin{aligned} \text{LHS} &= \int_{t=0}^{T} t^n \frac{\mathrm{d}}{\mathrm{d}t}(tf(t))\,\mathrm{d}t\\ &= T^{n+1}f(T)-\int_{t=0}^{T} tf(t)\,\mathrm{d}t^n\\ &= T^{n+1}f(T)-\int_{t=0}^{T} tf(t)nt^{n-1}\,\mathrm{d}t\\ &= T^{n+1}f(T)-nE_n(T),\\ \text{RHS} &= \int_{t=0}^{T} t^n(m-\lambda t)f(t)\,\mathrm{d}t\\ &= m\int_{t=0}^{T} t^n f(t)\,\mathrm{d}t-\lambda\int_{t=0}^{T} t^{n+1}f(t)\,\mathrm{d}t\\ &= mE_n(T)-\lambda E_{n+1}(T), \end{aligned}$$

from which we obtain

$$\begin{aligned} T^{n+1}f(T)-nE_n(T) &\le mE_n(T)-\lambda E_{n+1}(T),\\ \lambda E_{n+1}(T) &\le (m+n)E_n(T)-T^{n+1}f(T),\\ \frac{E_{n+1}(T)}{E_0(T)} &\le \frac{m+n}{\lambda}\frac{E_n(T)}{E_0(T)}-\frac{T^{n+1}f(T)}{\lambda E_0(T)}, \end{aligned}$$

which leads to the desired results. □

Lemma 6 gives a tight moment bound, for which the upper and the lower limits are identical if $\mathcal{Y}$ is Erlang distributed. To get rid of the density function in the boundary, a loose version of Lemma 6 is

$$\begin{aligned} \frac{m+n+\lambda T}{\lambda}\mathbb{E}[\mathcal{W}^n]-\frac{(m+n-1)T}{\lambda}\mathbb{E}[\mathcal{W}^{n-1}] &\le \mathbb{E}[\mathcal{W}^{n+1}]\\ &\le \frac{m+n}{\lambda}\mathbb{E}[\mathcal{W}^n]-\frac{T^{n+1}f(T)}{\lambda E_0(T)}\\ &< \frac{m+n}{\lambda}\mathbb{E}[\mathcal{W}^n], \end{aligned}$$

where the strict inequality indicates the loose boundary.

Now we look at the lower-order moments by focusing on moment bounds for the mean and SCV for FTPH_1 distributions with $b=0$.

Corollary 2. *The SCV of* $\mathcal{W}$, $\mathrm{SCV}_{\mathcal{W}} = \mathbb{E}[\mathcal{W}^2]/\mathbb{E}[\mathcal{W}]^2 - 1$, *is bounded by*

$$\frac{m+1+\lambda T}{\lambda \mathbb{E}\,[\mathcal{W}]} - \frac{mT}{\lambda(\mathbb{E}\,[\mathcal{W}]\,)^2} - 1 \leq \mathrm{SCV}_{\mathcal{W}} < \frac{m+1}{\lambda \mathbb{E}\,[\mathcal{W}]} - 1.$$

Proof. For $n = 1$, Lemma 6 gives

$$\frac{m+1+\lambda T}{\lambda}\mathbb{E}\,[\mathcal{W}] - \frac{mT}{\lambda} \leq \mathbb{E}[\mathcal{W}^2] \leq \frac{m+1}{\lambda}\mathbb{E}\,[\mathcal{W}] - \frac{T^2 f(T)}{\lambda E_0(T)},$$

whose right-hand side can be upper bounded by $(m+1)\mathbb{E}[\mathcal{W}]/\lambda$, from which the corollary follows by dividing with $(\mathbb{E}\,[\mathcal{W}]\,)^2$ and subtracting 1. □

We note that the difference between the upper and the lower limits of $\mathrm{SCV}_{\mathcal{W}}$ in Corollary 2 is

$$\frac{T}{(\mathbb{E}\,[\mathcal{W}]\,)^2}\Big(\frac{m}{\lambda} - \mathbb{E}\,[\mathcal{W}]\Big),$$

for which, according to (8) and the definition of $\mathcal{W}$, we have

$$\mathbb{E}\,[\mathcal{W}] = \mathbb{E}\,[\mathcal{Y} \mid \mathcal{Y} < T] < \mathbb{E}\,[\mathcal{Y}] \leq \frac{m}{\lambda}.$$

That is, when λ tends to 0, the upper bound converges to ∞. In this case, the λ independent upper limit $mT/(m+1)$ from Corollary 1 can be applied. Combining the results, we obtain $\mathbb{E}\,[\mathcal{W}] \leq \min\{m/\lambda, mT/(m+1)\}$.

Our final result of this subsection gives a lower bound of $\mathrm{SCV}_{\mathcal{W}}$ in terms of m only, which generalizes the result of Aldous and Shepp [1] mentioned in the introduction.

Theorem 2. *The SCV of the* FTPH_1 *random variable* $\mathcal{W}$ *with support on* $(0, B)$ *is bounded by* $\mathrm{SCV}_{\mathcal{W}} \geq 1/(m(m+2))$.

Proof. Recall that $T = B - b = B$. Define

$$g(t) = \frac{(m-1)f(t) - tf'(t)}{mF(T) - Tf(T)} \quad \text{for } 0 \leq t \leq T.$$

By Lemma 1, $(m-1-\lambda t)f(t) - tf'(t) \geq 0$ for $t > 0$, which leads to $(m-1)f(t) - tf'(t) > 0$ for $t > 0$. By integrating from 0 to T, we obtain

$$mF(T) - Tf(T) = \int_0^T ((m-1)f(t) - tf'(t))\,\mathrm{d}t > 0. \tag{13}$$

Consequently, $g(t)$ is a density function of a random variable, to be called Y_T, with support $[0, T)$. Note that $\int_0^T t^n \,\mathrm{d}F(t) = \mathbb{E}[\mathcal{Y}^n \mathbb{I}_{\{\mathcal{Y}<T\}}]$ for $n = 0, 1, 2, \ldots$. By routine calculations, we obtain

$$\mathbb{E}[Y_T] = \frac{(m+1)\mathbb{E}[\mathcal{Y}\mathbb{I}_{\{\mathcal{Y}<T\}}] - T^2 f(T)}{mF(T) - Tf(T)}, \qquad \mathbb{E}[Y_T^2] = \frac{(m+2)\mathbb{E}[\mathcal{Y}^2\mathbb{I}_{\{\mathcal{Y}<T\}}] - T^3 f(T)}{mF(T) - Tf(T)}.$$

It is well known that $\mathbb{E}\left[Y_T^2\right]/(\mathbb{E}\,[Y_T]\,)^2 \geq 1$. Using the above expressions, we obtain

$$(m+2)\mathbb{E}[\mathcal{Y}^2\mathbb{I}_{\{\mathcal{Y}<T\}}] \geq T^3 f(T) + \frac{((m+1)\mathbb{E}[\mathcal{Y}\mathbb{I}_{\{\mathcal{Y}<T\}}] - T^2 f(T))^2}{mF(T) - Tf(T)}.$$

Recall that $\mathcal{W} = \mathcal{Y} \mid \mathcal{Y} < T$. We also note that $\mathbb{E}[\mathcal{W}^2] = \mathbb{E}[\mathcal{Y}^2 \mathbb{I}_{\{\mathcal{Y}<T\}}]/F(T)$ and $\mathbb{E}[\mathcal{W}] = \mathbb{E}[\mathcal{Y}\mathbb{I}_{\{\mathcal{Y}<T\}}]/F(T)$. The above equation leads to

$$\begin{aligned}
&\frac{\mathbb{E}[\mathcal{W}^2]}{(\mathbb{E}[\mathcal{W}])^2} \\
&\quad = \frac{\mathbb{E}[\mathcal{Y}^2\mathbb{I}_{\{\mathcal{Y}<T\}}]F(T)}{(\mathbb{E}[\mathcal{Y}\mathbb{I}_{\{\mathcal{Y}<T\}}])^2} \\
&\quad \geq \frac{F(T)(T^3 f(T) + ((m+1)\mathbb{E}[\mathcal{Y}\mathbb{I}_{\{\mathcal{Y}<T\}}] - T^2 f(T))^2/(mF(T) - Tf(T)))}{(m+2)(\mathbb{E}[\mathcal{Y}\mathbb{I}_{\{\mathcal{Y}<T\}}])^2} \\
&\quad = \frac{F(T)(((m+1)\mathbb{E}[\mathcal{Y}\mathbb{I}_{\{\mathcal{Y}<T\}}])^2 - 2(m+1)T^2 f(T)\mathbb{E}[\mathcal{Y}\mathbb{I}_{\{\mathcal{Y}<T\}}] + mT^3 f(T)F(T))}{(m+2)(mF(T) - Tf(T))(\mathbb{E}[\mathcal{Y}\mathbb{I}_{\{\mathcal{Y}<T\}}])^2} \\
&\quad = \frac{(m+1)^2}{m(m+2)} \\
&\qquad \times \left(\frac{1 - 2T^2 f(T)/((m+1)\mathbb{E}[\mathcal{Y}\mathbb{I}_{\{\mathcal{Y}<T\}}]) + mT^3 f(T)F(T)/((m+1)\mathbb{E}[\mathcal{Y}\mathbb{I}_{\{\mathcal{Y}<T\}}])^2}{1 - Tf(T)/(mF(T))}\right) \\
&\quad = \frac{(m+1)^2}{m(m+2)}\Theta(T).
\end{aligned}$$

We want to show that $\Theta(T) \geq 1$ for all $T > 0$. Since $mF(T) > Tf(T)$ according to (13), $\Theta(T) \geq 1$ is equivalent to

$$\frac{Tf(T)}{mF(T)} + \frac{mT^3 f(T)F(T)}{((m+1)\mathbb{E}[\mathcal{Y}\mathbb{I}_{\{\mathcal{Y}<T\}}])^2} \geq 2\frac{T^2 f(T)}{(m+1)\mathbb{E}[\mathcal{Y}\mathbb{I}_{\{\mathcal{Y}<T\}}]},$$

which is equivalent to

$$\frac{1}{mF(T)} + \frac{T^2 mF(T)}{((m+1)\mathbb{E}[\mathcal{Y}\mathbb{I}_{\{\mathcal{Y}<T\}}])^2} \geq 2\frac{T}{(m+1)\mathbb{E}[\mathcal{Y}\mathbb{I}_{\{\mathcal{Y}<T\}}]}.$$

The last equation holds by applying the well-known inequality $a^2 + b^2 \geq 2ab$ for any real numbers a and b. Thus, we have shown that $\Theta(T) \geq 1$ for all $T > 0$. Consequently, we have shown that

$$\frac{\mathbb{E}[\mathcal{W}^2]}{(\mathbb{E}\,[\mathcal{W}]\,)^2} \geq \frac{(m+1)^2}{m(m+2)},$$

which is equivalent to $\mathrm{SCV}_{\mathcal{W}} \geq 1/(m(m+2))$. □

By Corollary 1, the lower bound of $\mathrm{SCV}_{\mathcal{W}}$ is strict for all PH distributions with finite support. The following lemma and corollary show how the lower bound of $\mathrm{SCV}_{\mathcal{W}}$ can be attained approximately by bounded Erlang distributions. For $t \geq 0$, denote by

$$F_m(t) = 1 - \mathrm{e}^{-\lambda t} - \lambda t\mathrm{e}^{-\lambda t} - \cdots - \frac{(\lambda t)^{m-1}}{(m-1)!}\mathrm{e}^{-\lambda t},$$
$$f_m(t) = f_{(m,\lambda)}(t) = \frac{\lambda^m t^{m-1}}{(m-1)!}\mathrm{e}^{-\lambda t},$$

the distribution function and density function of an Erlang random variable $\mathcal{X}_{(m,\lambda)}$, respectively. By routine calculations, we obtain

$$\mathbb{E}[\mathcal{X}_{(m,\lambda)} \mid \mathcal{X}_{(m,\lambda)} < T] = \frac{\int_0^T tf_m(t)\,\mathrm{d}t}{F_m(T)} = \frac{m}{\lambda}\frac{F_{m+1}(T)}{F_m(T)},$$

$$\mathbb{E}[\mathcal{X}^2_{(m,\lambda)} \mid \mathcal{X}_{(m,\lambda)} < T] = \frac{\int_0^T t^2 f_m(t)\,\mathrm{d}t}{F_m(T)} = \frac{m(m+1)}{\lambda^2}\frac{F_{m+2}(T)}{F_m(T)},$$

$$\mathrm{SCV}_{\{\mathcal{X}_{(m,\lambda)} \mid \mathcal{X}_{(m,\lambda)}<T\}} = \frac{\mathbb{E}[\mathcal{X}^2_{(m,\lambda)} \mid \mathcal{X}_{(m,\lambda)} < T]}{(\mathbb{E}[\mathcal{X}_{(m,\lambda)} \mid \mathcal{X}_{(m,\lambda)} < T])^2} - 1 = \left(\frac{m+1}{m}\right)\frac{F_m(T)F_{m+2}(T)}{(F_{m+1}(T))^2} - 1.$$

Lemma 7. *For all $T > 0$, the following bounds apply for the distribution functions of Erlang random variables:*

$$\frac{m+1}{m+2} \le \frac{F_m(T)F_{m+2}(T)}{(F_{m+1}(T))^2} \le 1.$$

In addition, we have $\lim_{T\to 0} F_m(T)F_{m+2}(T)/(F_{m+1}(T))^2 = (m+1)/(m+2)$ *and* $\lim_{T\to\infty} F_m(T)F_{m+2}(T)/(F_{m+1}(T))^2 = 1$.

Proof. For convenience, we denote λT as t in this proof. First, we prove the upper bound. Since $\mathrm{e}^{\lambda t} = \sum_{k=0}^{\infty} (\lambda t)^k/k!$, we need to show that

$$\left(\sum_{k=m+2}^{\infty}\frac{t^k}{k!}\right)\left(\frac{t^m}{m!} + \frac{t^{m+1}}{(m+1)!} + \sum_{k=m+2}^{\infty}\frac{t^k}{k!}\right) \le \left(\frac{t^{m+1}}{(m+1)!} + \sum_{k=m+2}^{\infty}\frac{t^k}{k!}\right)^2,$$

which can be reduced to

$$\left(\sum_{k=m+2}^{\infty}\frac{t^k}{k!}\right)\frac{t^m}{m!} \le \left(\sum_{k=m+2}^{\infty}\frac{t^k}{k!}\right)\frac{t^{m+1}}{(m+1)!} + \left(\frac{t^{m+1}}{(m+1)!}\right)^2.$$

Next, we compare the coefficients of t^k on both sides. For $k = 2m+2$, the left-hand side is $1/((m!(m+2)!)$ and the right-hand side is $1/((m+1)!)^2$. From

$$\frac{1}{m!(m+2)!} = \frac{m+1}{m+2}\left(\frac{1}{(m+1)!}\right)^2,$$

it follows that the left-hand side is smaller than the right-hand side. For $k \ge 2m+3$, we have, for $k = j+m$,

$$\frac{1}{j!\,m!} \le \frac{1}{(j-1)!\,(m+1)!},$$

which is equivalent to $m+1 \le j$, and holds since $j = k-m \ge m+3$. Consequently, we have shown the upper bound.

The proof of the lower bound is similar but tedious. The lower-bound expression can be rewritten as $(m+1)F^2_{m+1}(t/\lambda) \le (m+2)F_m(t/\lambda)F_{m+2}(t/\lambda)$, which can be rewritten explicitly as

$$(m+1)\left(\sum_{k=m}^{\infty}\frac{t^k}{k!} - \frac{t^m}{m!}\right)^2 \le (m+2)\left(\sum_{k=m}^{\infty}\frac{t^k}{k!}\right)\left(\sum_{k=m}^{\infty}\frac{t^k}{k!} - \frac{t^m}{m!} - \frac{t^{m+1}}{(m+1)!}\right),$$

which leads to

$$(m+2)\left(\sum_{k=m}^{\infty}\frac{t^k}{k!}\right)\frac{t^{m+1}}{(m+1)!}+(m+1)\left(\frac{t^m}{m!}\right)^2\le\left(\sum_{k=m}^{\infty}\frac{t^k}{k!}\right)^2+m\left(\sum_{k=m}^{\infty}\frac{t^k}{k!}\right)\frac{t^m}{m!}.$$

To prove the above inequality, we compare the coefficients of t^n on both sides. For $n=2m$, we have

$$\frac{m+1}{m!\,m!}\le\frac{1}{m!\,m!}+\frac{m}{m!\,m!},$$

which holds. For $n\ge 2m+1$, we need to prove that

$$\frac{m+2}{(n-m-1)!\,(m+1)!}\le\frac{m}{(n-m)!\,m!}+\sum_{i=m}^{n-m}\frac{1}{i!(n-i)!}.$$

Separating the first and last terms of the summation and applying $k!=k(k-1)!$, we obtain

$$\frac{(m+2)(n-m)}{(n-m)!\,(m+1)!}\le\frac{(m+2)(m+1)}{(n-m)!\,(m+1)!}+\sum_{i=m+1}^{n-m-1}\frac{1}{i!\,(n-i)!},$$

which leads to

$$\frac{(m+2)(n-2m-1)}{(n-m)!\,(m+1)!}\le\sum_{i=m+1}^{n-m-1}\frac{1}{i!\,(n-i)!}. \tag{14}$$

For any $i\in\{m+1,\ldots,n-m-1\}$, we have $i\le n-m$ and $m+1\le n-i$, from which we can write

$$\left(\frac{i}{n-m}\frac{i-1}{n-m-1}\cdots\frac{m+2}{n-i+2}\right)\frac{m+2}{n-i+1}\le 1,$$

$$(i(i-1)\cdots(m+2))(m+2)\le(n-m)(n-m-1)\cdots(n-i+1),$$

$$\frac{i!}{(m+1)!}(m+2)\le\frac{(n-m)!}{(n-i)!},\qquad\frac{m+2}{(n-m)!(m+1)!}\le\frac{1}{i!\,(n-i)!},$$

Considering that $n-m-1-(m+1)+1=n-2m-1$ terms are summed on the right-hand side of (14), each of which is greater than or equal to $(m+2)/((n-m)!\,(m+1)!)$, inequality (14) as well as the lower bound of Lemma 7 are proved. This completes the proof of the lemma. □

Immediate consequences of Lemma 7 are a lower bound and an upper bound of the SCV for bounded $\mathcal{X}_{(m,\lambda)}$.

Corollary 3. *Assume that $\mathcal{X}$ has an Erlang distribution with parameters (m,λ). For all $T>0$, we have*

$$\frac{1}{m(m+2)}\le \mathrm{SCV}_{\{\mathcal{X}_{(m,\lambda)}\mid\mathcal{X}_{(m,\lambda)}<T\}}\le\frac{1}{m}.$$

3.2. The $b>0$ case

Let $\mathcal{Z}=b+\mathcal{W}=b+(\mathcal{Y}mid\mathcal{Y}<T)$. Then, for $\mathbb{E}\left[\mathcal{Z}^n\right]$, we have

$$\mathbb{E}[\mathcal{Z}^n]=\sum_{i=0}^{n}\binom{n}{i}b^{n-i}\frac{E_i(T)}{E_0(T)}=\sum_{i=0}^{n}\binom{n}{i}b^{n-i}\mathbb{E}[\mathcal{W}^i],$$

where $E_i(T)$ is defined as before. That is, for $n = 1, 2,$ we have

$$\mathbb{E}[\mathcal{Z}] = b + \mathbb{E}[\mathcal{W}] \quad \text{and} \quad \mathbb{E}[\mathcal{Z}^2] = b^2 + 2b\mathbb{E}[\mathcal{W}] + \mathbb{E}[\mathcal{W}^2].$$

Corollary 4. *The nth moment of an* FTPH_1 *random variable,* $\mathcal{Z}$*, with support on* (b, B) *is bounded by*

$$b^n \le \mathbb{E}[\mathcal{Z}^n] \le \sum_{i=0}^{n} \binom{n}{i} b^{n-i} \frac{mT^i}{m+i}. \tag{15}$$

Proof. Equation (15) directly follows from Corollary 1. □

For lower-order moments, according to (15), the mean of $\mathcal{Z}$ is bounded by

$$b \le \mathbb{E}[\mathcal{Z}] \le b + \frac{mT}{m+1} = \frac{b+mB}{m+1} < B,$$

where both moment bounds are tight. The lower boundary is reached when $\mathbb{E}[\mathcal{Y}]$ tends to 0 and the upper boundary is reached when $\mathcal{Y}$ is Erlang(λ, m) distributed and λ tends to 0.

For the SCV, we have the following corollary.

Corollary 5. *The* $\mathrm{SCV}_{\mathcal{Z}}$ *is bounded by the following* λ *independent and dependent moment bounds:*

$$\mathrm{SCV}_{\mathcal{Z}} = \frac{\mathbb{E}[\mathcal{W}^2] - \mathbb{E}[\mathcal{W}]^2}{(b + \mathbb{E}[\mathcal{W}])^2} \le \frac{(m+1)\mathbb{E}[\mathcal{W}]\,T/(m+2) - \mathbb{E}[\mathcal{W}]^2}{(b + \mathbb{E}[\mathcal{W}])^2},$$

$$\frac{-mT + (m+1+\lambda T)\mathbb{E}[\mathcal{W}] - \lambda\mathbb{E}[\mathcal{W}]^2}{\lambda(b + \mathbb{E}[\mathcal{W}])^2} \le \mathrm{SCV}_{\mathcal{Z}} < \frac{(m+1)\mathbb{E}[\mathcal{W}] - \lambda\mathbb{E}[\mathcal{W}]^2}{\lambda(b + \mathbb{E}[\mathcal{W}])^2}.$$

Proof. From Lemmas 4 and 6 we have $\mathbb{E}[\mathcal{W}^2] \le (m+1)T\mathbb{E}[\mathcal{W}]/(m+2)$ and

$$\frac{m+1+\lambda T}{\lambda}\mathbb{E}[\mathcal{W}] - \frac{mT}{\lambda} \le \mathbb{E}[\mathcal{W}^2] < \frac{m+1}{\lambda}\mathbb{E}[\mathcal{W}], \tag{16}$$

respectively. Subtracting $\mathbb{E}[\mathcal{W}]^2$ and then dividing by $(b + \mathbb{E}[\mathcal{W}])^2$ in (16) gives the corollary. □

Different from the $b = 0$ case, the $\mathrm{SCV}_{\mathcal{Z}}$ can reach 0 in the $b > 0$ case.

4. Discussion and conclusion

In this paper we presented new moment bounds on PH distributions with infinite and finite supports by using the steepest increase property. For PH distributions with infinite support and a PH representation $(\boldsymbol{\alpha}, \boldsymbol{A})$ of size m, denoted as $\mathcal{Y}$, we have

- shown that any PH distribution is stochastically smaller than or equal to an Erlang distribution $\mathcal{X}_{(m,\lambda)}$ with λ the absolute value of the dominant eigenvalue of $\boldsymbol{A}$; and
- obtained upper bounds of moments in terms of m and λ (e.g. $\mathbb{E}[\mathcal{Y}] \le m/\lambda$).

For PH distributions with finite support (for the set FTPH_1), denoted as $\mathcal{W} = \mathcal{Y} \mid \mathcal{Y} < T$, we have

- obtained upper bounds of moments in terms of m and T;

- obtained lower and upper bounds of moments depending on λ;
- shown that $\mathbb{E}[\mathcal{W}] \le \min\{mT/(m+1), m/\lambda\}$; and
- shown that $\mathrm{SCV}_{\mathcal{W}} \ge 1/(m(m+2))$.

For the finite support case, we focused on the distribution set FTPH_1. Results for the set FTPH_2 can be obtained similarly. The set FTPH_3 is a convex mixture of FTPH_1 and FTPH_2. Moment bounds can also be obtained as a convex mixture of the moment bounds obtained for FTPH_1 and FTPH_2, but this is outside the scope of the current work.

Acknowledgements

This work was supported by the Hungarian research project OTKA K123914, a NSERC discovery grant (Canada), and by the János Bolyai Research Scholarship of the Hungarian Academy of Sciences. The authors would like to thank two anonymous referees and an Associate Editor for their exceptionally careful reviews of the paper.

References

[1] Aldous, D. and Shepp, L. (1987). The least variable phase type distribution is Erlang. *Commun. Statist. Stoch. Models* **3,** 467–473.
[2] He, Q.-M., Zhang, H., and Vera, J. (2013). Majorization and extremal PH-distributions. *In Matrix-Analytic Methods in Stochastic Models* (Springer Proc. Math. Statist. **27**), eds G. Latouche *et al.*, Springer, New York, pp. 107-121.
[3] Neuts, M. F. (1975). Probability distributions of phase type. *In Liber Amicorum Prof. Emeritus H. Florin,* University of Louvain, Belgium, pp. 173-206.
[4] Neuts, M. F. (1981). *Matrix-Geometric Solutions in Stochastic Models: An Algorithmic Approach.* Johns Hopkins University Press, Baltimore.
[5] O'Cinneide, C. A. (1989). On non-uniqueness of representations of phase-type distributions. *Commun. Statist. Stoch. Models* **5**, 247–259.
[6] O'Cinneide, C. A. (1990). Characterization of phase-type distributions. *Commun. Statist. Stoch. Models* **6**, 1–57.
[7] O'Cinneide, C. A. (1991). Phase-type distributions and majorization. *Ann. Appl. Prob.* **1**, 219-227.
[8] O'Cinneide, C. A. (1999). Phase-type distributions: open problems and a few properties. *Commun. Statist. Stoch. Models*, **15**, 731-757,
[9] Ramaswami, V. and Viswanath, N. C. (2014). Phase type distributions with finite support. *Stoch. Models*, **30**, 576-597.
[10] Stoyan, D. (1983). *Comparison Methods for Queues and Other Stochastic Models.* John Wiley, Chichester.
[11] Telek, M. (2000). The minimal coefficient of variation of discrete phase type distributions. *Proc. 3rd Internat. Conf. on Matrix Analytic Methods in Stochastic Models*, eds G. Latouche and P. G. Taylor, Notable Publications, New Jersey.
[12] Yao, R. (2002). A proof of the steepest increase conjecture of a phase-type density. *Stoch. Models*, **18**, 1-6.