On the Canonical Representation of Phase Type Distributions

Gábor Horváth and Miklós Telek

Department of Telecommunications, Budapest University of Technology and Economics,
H-1521 Budapest, Hungary

Abstract

The characterization and the canonical representation of order-n phase type distributions (PH(n)) is an open research problem.

This problem is solved for $n = 2$, since the equivalence of the acyclic and the general PH distributions has been proven for a long time. However, no canonical representations have been introduced for the general PH distribution class so far for $n > 2$. In this paper we summarize the related results for $n = 3$. Starting from these results we provide a canonical representation of the PH(3) class (that is a minimal representation, too) and present a symbolical transformation procedure to obtain the canonical representation based on any (not only Markovian) vector-matrix representation of the distribution. We show that – using the same approach – no symbolical results can be derived for the order-4 PH distributions, thus probably the PH(3) class is the highest order PH class for which a symbolical canonical transformation exists.

Using the transformation method to canonical form for PH(3) we numerically evaluate the moment bounds of the PH(3) distribution set, compare it to the order-3 acyclic PH distribution (APH(3)) class, and present other possible applications of the canonical form.

Email address: {`ghorvath,telek`}`@hit.bme.hu` (Gábor Horváth and Miklós Telek).

[1] This work is partially supported by the Italian-Hungarian R&D project 9/2003 and by the OTKA K61709 grant.

[2] This paper is an extended version of [1].

1 Introduction

Markovian structures are efficiently applied in various fields of stochastic modeling because of their computability and numerical stability. Phase type distributions are non-negative distributions with Markovian structure [2,3]. They are widely used in distribution approximation due to their computational advantages and easy integration in complex stochastic models.

The most common representation of a phase type distribution is the definition of its initial probability vector $\underline{\alpha}$ and generator matrix $\boldsymbol{A}$. This representation is known to be non-unique and non-minimal, thus there might be a vector $\underline{\alpha}'$ and a matrix $\boldsymbol{A}'$, which define the same distribution. Furthermore, the number of parameters (non-determined elements) of this representation is $n^2 + n - 1$ when the cardinality of vector $\underline{\alpha}$ and square matrix $\boldsymbol{A}$ is n (since $\boldsymbol{A}$ has n^2 elements and $\underline{\alpha}$ has $n - 1$ assuming no probability mass at zero), while the Laplace transform of PH(n) distributions – that uniquely determines the distribution – has $2n - 1$ roots and zeros.

To overcome these drawbacks a unique, minimal representation is required which is commonly referred to as canonical representation. A canonical representation is available for any order acyclic phase type distributions by Cumani [4], and it is also known that any PH(2) distribution can be transformed to an acyclic form [5] and this way the same canonical form is applicable of PH(2).

The canonical representation of PH(n) distributions is not known for $n \geq 4$ and we present a proposal for the canonical representation of the PH(3) class in this paper. The proposed representation has a special $\underline{\alpha}$ vector and $\boldsymbol{A}$ matrix such that it has exactly $2n - 1 = 5$ parameters and it is proved to exist for all PH(3) distributions. We also provide a procedure for transforming any (not only Markovian) vector-matrix representation of the distribution to the canonical form. The transformation procedure is composed of explicit computational steps, whose most complex element is the evaluation of the eigenvalues of the generator matrix (finding the roots of an order-3 polynomial, for which symbolic solution is available).

Our results are very much based on the results of [6], where the unicyclic representation of PH(3) distributions is proved. Indeed, the presented canonical representation is unicyclic, but it extends the results of [6] with the careful analysis of the initial probability vector of the canonical representation, which is not taken into consideration in [6].

By means of this transformation procedure, which fails only when the input vector-matrix pair cannot be transformed into a valid PH(3) representation, we investigate also the moments bounds of the PH(3) class. Some results on the bounds of the first three moments of PH(3) distributions are provided in

[7], but the behaviour of the fourth and fifth moments are unknown to the best of our knowledge.

The rest of the paper is organized as follows. Section 2 gives the definition and the basic properties of PH(3) distributions. The unicyclic transformation of PH(3) distributions is summarized in Section 3 and the proposed canonical representation is presented in Section 4. The possible canonical forms of PH(4) distributions are investigated in Section 5. Section 6 lists some applications of the canonical form and the associated transformation method including a numerical study of the moment bounds. The paper is concluded in Section 7.

2 PH distributions

Let $\mathcal{X}$ be a continuous non-negative random variable with cumulative distribution function

$$F(t) = Pr(\mathcal{X} < t) = 1 - \underline{v}e^{\boldsymbol{H}t}\mathbb{1} \ ,$$

where the row vector $\underline{v}$ is referred to as the initial vector, square matrix $\boldsymbol{H}$ as the generator and $\mathbb{1}$ as the closing vector. Without loss of generality [8], we assume that the closing vector $\mathbb{1}$ is a column vector of ones, i.e., $\mathbb{1} = [1, 1, \ldots, 1]^T$. Since $\mathcal{X}$ is a continuous random variable, it has no probability mass at zero, i.e., $\underline{v}\mathbb{1} = 1$. The density, the Laplace transform and the moments of $\mathcal{X}$ are

$$f(t) = \underline{v}e^{\boldsymbol{H}t}(-\boldsymbol{H})\mathbb{1} \ , \tag{1}$$

$$f^*(s) = E(e^{-s\mathcal{X}}) = \underline{v}(s\boldsymbol{I} - \boldsymbol{H})^{-1}(-\boldsymbol{H})\mathbb{1} \ , \tag{2}$$

$$\mu_n = E(\mathcal{X}^n) = n!\underline{v}(-\boldsymbol{H})^{-n}\mathbb{1} \ . \tag{3}$$

When the cardinality of vector $\underline{v}$ and of square matrix $\boldsymbol{H}$ is n, we have the following cases [9]:

- If $f(t) \geq 0$ and $\int_0^\infty f(t)dt = 1$, then $\mathcal{X}$ has an order-n matrix exponential (ME(n)) distribution. The elements of $\underline{v}$ and $\boldsymbol{H}$ may be arbitrary real numbers.
- If $\underline{v}$ is a probability vector and $\boldsymbol{H}$ is a transient Markovian generator matrix (i.e., the generator matrix of a transient continuous-time Markov chain (CTMC)), then $\mathcal{X}$ has a PH(n) distribution. (The set of PH(n) distributions form a true subset of the ME(n) set for $n > 2$.)

Vector $\underline{v}$ is a probability vector when $v_i \geq 0$, $\underline{v}\mathbb{1} = 1$ and matrix $\boldsymbol{H}$ is a transient Markovian generator when $\boldsymbol{H}$ is non-singular, $\boldsymbol{H}_{ii} < 0$, $\boldsymbol{H}_{ij} \geq 0$ for $i \neq j$, $\boldsymbol{H}\mathbb{1} \leq \boldsymbol{0}$, $\boldsymbol{H}\mathbb{1} \neq \boldsymbol{0}$. Scalars like $\boldsymbol{H}_{ij}$ denote the ijth element of matrix $\boldsymbol{H}$.

Definition 1 *The* $(\underline{v}, \boldsymbol{H})$ *representation is a* Markovian representation, *if* $\underline{v}$ *is a probability vector and* $\boldsymbol{H}$ *is a transient Markovian generator matrix.*

In general it is not easy to check whether an $f(t)$ in (1) corresponding to a $(\underline{v}, \boldsymbol{H})$ pair is a density function. We have the following necessary conditions (those that we use in the sequel, [10]):

- the eigenvalues of $\boldsymbol{H}$ have negative real part,
- the largest eigenvalue of $\boldsymbol{H}$ is real, and
- the initial value of the density function is non-negative:

$$f(0) = -\underline{v}\boldsymbol{H}\mathbb{1} \geq 0 \ . \tag{4}$$

Definition 2 *Assuming* $\boldsymbol{B}$ *is a non-singular matrix such that* $\boldsymbol{B}\mathbb{1} = \mathbb{1}$ *then the vector-matrix pair* $\underline{v}\boldsymbol{B}$, $\boldsymbol{B}^{-1}\boldsymbol{H}\boldsymbol{B}$ *define a* similarity transform *of the vector-matrix pair* $\underline{v}$, $\boldsymbol{H}$.

Note that the vector-matrix pairs $\underline{v}$, $\boldsymbol{H}$ and $\underline{v}\boldsymbol{B}$, $\boldsymbol{B}^{-1}\boldsymbol{H}\boldsymbol{B}$ represent the same distribution, since

$$\hat{F}(t) = 1 - \underline{v}\boldsymbol{B}e^{\boldsymbol{B}^{-1}\boldsymbol{H}\boldsymbol{B}t}\mathbb{1} = 1 - \underline{v}\boldsymbol{B}\boldsymbol{B}^{-1}e^{\boldsymbol{H}t}\boldsymbol{B}\mathbb{1} = 1 - \underline{v}e^{\boldsymbol{H}t}\mathbb{1} = F(t) \ .$$

Example 1

$$\underline{v} = \begin{bmatrix} 0.1 & 0.5 & 0.4 \end{bmatrix}, \quad \boldsymbol{H} = \begin{bmatrix} -5 & 2 & 1 \\ 1 & -2 & 1 \\ 1 & 0 & -4 \end{bmatrix}$$

and

$$\underline{z} = \begin{bmatrix} -1.1 & 2.5 & -0.4 \end{bmatrix}, \quad \boldsymbol{G} = \begin{bmatrix} -11 & 10 & -1 \\ -6.6 & 6 & -1 \\ -15 & 20 & -6 \end{bmatrix}$$

represent the same distribution, since $\underline{z} = \underline{v}\boldsymbol{B}$ *and* $\boldsymbol{G} = \boldsymbol{B}^{-1}\boldsymbol{H}\boldsymbol{B}$ *with* $\boldsymbol{B} = \begin{bmatrix} 1 & 0 & 0 \\ -4 & 5 & 0 \\ 2 & 0 & -1 \end{bmatrix}$. $(\underline{z}, \boldsymbol{G})$ *is a non-Markovian representation of this PH(3) distribution.*

Now, we can refine the above definition of PH(n) distributions by means of similarity transforms.

Definition 3 *The random variable* $\mathcal{X}$ *with density function (1) is* PH distributed *if there is a non-singular matrix* $\boldsymbol{B}$, *such that* $\boldsymbol{B}\mathbb{1} = \mathbb{1}$, *and* $(\underline{v}\boldsymbol{B}, \boldsymbol{B}^{-1}\boldsymbol{H}\boldsymbol{B})$ *is a Markovian representation.*

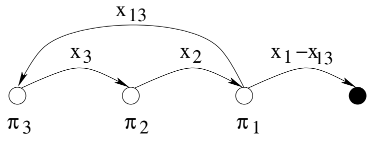

Fig. 1. The structure of the considered unicyclic PH(3) distribution

Note that this definition implies that $f(t) \geq 0$.

One of the main goals of this paper is to decide if such similarity transform exists for a given non-Markovian vector-matrix pair, since the definition is obvious when the vector-matrix pair is Markovian.

3 Unicyclic representation of PH(3) distributions

The results of this paper are based on the unicyclic transformation of PH(3) distributions presented in [6]. We summarize the related results, in a slightly modified way, for completeness.

Theorem 1 *[6] If $(\underline{v}, \boldsymbol{H})$ is a Markovian representation of a PH(3) distribution then it can be similarity transformed to the following unicyclic Markovian representation*

$$\underline{\pi} = \begin{bmatrix} \pi_1 & \pi_2 & \pi_3 \end{bmatrix}, \quad \boldsymbol{A} = \begin{bmatrix} -x_1 & 0 & x_{13} \\ x_2 & -x_2 & 0 \\ 0 & x_3 & -x_3 \end{bmatrix}, \tag{5}$$

where $x_1 \geq x_2 \geq x_3 > 0,\ 0 \leq x_{13} < x_1,\ 0 \leq \pi_1, \pi_2, \pi_3,\ \pi_1 + \pi_2 + \pi_3 = 1$ and the procedure in Figure 2 generates this unicyclic representation.

The structure of the resulting unicyclic PH distribution is depicted in Figure 1.

The main difference between Theorem 1 ([6]) and the goal of this paper is that Theorem 1 assumes that $(\underline{v}, \boldsymbol{H})$ is Markovian, while we look for a transformation which is applicable for any non-Markovian $(\underline{v}, \boldsymbol{H})$ representation. For example the procedure of Figure 2 gives a proper unicyclic representation when it is called with the $(\underline{v}, \boldsymbol{H})$ pair of Example 1, but it gives complex results when it is called with the $(\underline{z}, \boldsymbol{G})$ representation of the same PH(3) distribution.

Let $\lambda_1, \lambda_2, \lambda_3$ denote the eigenvalues of $-\boldsymbol{H}$ which are ordered such that $Re(\lambda_1) \geq Re(\lambda_2) \geq Re(\lambda_3)$ and a_0, a_1, a_2 the coefficients of the character-

function PH(3)–to–unicyclic PH(3)

input: $\underline{v}, \boldsymbol{H}$ (Markovian)

output: $\underline{\pi}, \boldsymbol{A}$ (unicyclic)

begin

$\lambda_1, \lambda_2, \lambda_3 \;=\;$ decreasingly ordered eigenvalues of $-\boldsymbol{H}$,

$a_0 = \lambda_1\,\lambda_2\,\lambda_3, \quad a_1 = \lambda_1\,\lambda_2 + \lambda_1\,\lambda_3 + \lambda_2\,\lambda_3, \quad a_2 = \lambda_1 + \lambda_2 + \lambda_3,$

$\gamma_u = \frac{1}{3}\,(a_2 + 2\sqrt{a_2^2 - 3\,a_1}), \quad \gamma_0 = \frac{1}{3}\,(a_2 + \sqrt{a_2^2 - 3\,a_1}),$

$\gamma_\ell = \begin{cases} \lambda_1 \ \textbf{if}\ \lambda_1 \in \text{real}, \\ \gamma_0 \ \textbf{if}\ \lambda_1 \in \text{complex}, \end{cases}$

$\phi = \max\{-\boldsymbol{H}_{1,1},\ -\boldsymbol{H}_{2,2},\ -\boldsymbol{H}_{3,3}\},$

$x_1 = \max\{\phi,\ \gamma_\ell\},$

$x_{13} = x_1 - a_0 \;/\; (x_1^2 - a_2\,x_1 + a_1),$

$x_2 = \frac{1}{2}\left(a_2 - x_1 + \sqrt{(a_2 - x_1)^2 \,-\, 4\,(x_1^2 - a_2\,x_1 + a_1)}\right),$

$x_3 = \frac{1}{2}\left(a_2 - x_1 - \sqrt{(a_2 - x_1)^2 \,-\, 4\,(x_1^2 - a_2\,x_1 + a_1)}\right),$

$\pi_1 = \underline{v}\,\boldsymbol{H}\,\mathbb{1} \;/\; (x_{13} - x_1),$

$\pi_2 = \underline{v}\,(x_1\,\boldsymbol{I} + \boldsymbol{H})\,\boldsymbol{H}\,\mathbb{1} \;/\; \left((x_{13} - x_1)\,x_2\right),$

$\pi_3 = \underline{v}\,(x_2\,\boldsymbol{I} + \boldsymbol{H})\,(x_1\,\boldsymbol{I} + \boldsymbol{H})\,\boldsymbol{H}\,\mathbb{1} \;/\; \left((x_{13} - x_1)\,x_2\,x_3\right),$

return $\underline{\pi} = \begin{bmatrix} \pi_1 & \pi_2 & \pi_3 \end{bmatrix}, \ \boldsymbol{A} = \begin{bmatrix} -x_1 & 0 & x_{13} \\ x_2 & -x_2 & 0 \\ 0 & x_3 & -x_3 \end{bmatrix},$

end

Fig. 2. Unicyclic transformation of PH(3) distributions

istic polynomial of $-\boldsymbol{H}$, i.e., $x^3 + a_2x^2 + a_1x + a_0 = 0$, where

$$a_0 = \lambda_1\lambda_2\lambda_3,\ a_1 = \lambda_1\lambda_2 + \lambda_1\lambda_3 + \lambda_2\lambda_3,\ a_2 = \lambda_1 + \lambda_2 + \lambda_3. \tag{6}$$

A simple interpretation of Theorem 1 is that the similarity transform with matrix $\boldsymbol{B}$ makes the transformed matrix to be unicyclic if $\boldsymbol{B}$ is composed of the column vectors $\{\underline{b_1}, \underline{b_2}, \underline{b_3}\}$ where

$$\begin{aligned} \underline{b_1} &= \frac{1}{x_{13} - x_1}\ \boldsymbol{H}\mathbb{1}, \\ \underline{b_2} &= \frac{1}{(x_{13} - x_1)x_2}\ (x_1\boldsymbol{I} + \boldsymbol{H})\boldsymbol{H}\mathbb{1}, \\ \underline{b_3} &= \frac{1}{(x_{13} - x_1)x_2x_3}\ (x_2\boldsymbol{I} + \boldsymbol{H})(x_1\boldsymbol{I} + \boldsymbol{H})\boldsymbol{H}\mathbb{1}, \end{aligned} \tag{7}$$

and

$$
\begin{aligned}
x_{13} &= x_1 - \frac{a_0}{x_1^2 - a_2 x_1 + a_1}, \\
x_2 &= \frac{a_2 - x_1 + \sqrt{(a_2 - x_1)^2 - 4(x_1^2 - a_2 x_1 + a_1)}}{2}, \\
x_3 &= \frac{a_2 - x_1 - \sqrt{(a_2 - x_1)^2 - 4(x_1^2 - a_2 x_1 + a_1)}}{2}.
\end{aligned} \tag{8}
$$

These expressions are obtained from the fact that the resulting generator $\boldsymbol{A}$ has the same characteristic polynomial as the original $\boldsymbol{H}$, i.e., the parameters are obtained from the solution of the equations

$$
a_0 = (x_1 - x_{13}) x_2 x_3, \; a_1 = x_1 x_2 + x_2 x_3 + x_3 x_1, \; a_2 = x_1 + x_2 + x_3. \tag{9}
$$

The transformation matrix $\boldsymbol{B}$ and the transformed unicyclic representation $\boldsymbol{A}$ depend on the choice of x_1. [6] showed the following properties of PH(3) distributions and this similarity transform.

P1) When $\boldsymbol{H}$ is a Markovian generator then

$$
\gamma_u = \frac{a_2 + 2\sqrt{a_2^2 - 3a_1}}{3}, \tag{10}
$$

$$
\gamma_0 = \frac{a_2 + \sqrt{a_2^2 - 3a_1}}{3}, \tag{11}
$$

$$
\gamma_\ell = \begin{cases} \lambda_1, & \text{if } \lambda_1 \text{ is real,} \\ \gamma_0, & \text{if } \lambda_1 \text{ is complex} \end{cases} \tag{12}
$$

are real and positive such that $\gamma_0 \le \gamma_\ell \le \gamma_u$.

P2) When $\gamma_\ell \le x_1 \le \gamma_u$ then the transformed generator matrix, $\boldsymbol{A} = \boldsymbol{B}^{-1}\boldsymbol{H}\boldsymbol{B}$ is Markovian such that $x_1 \ge x_2 \ge x_3 > 0$.

Indeed, property **P2** holds also for any non-Markovian matrix $\boldsymbol{H}$ if its eigenvalues satisfy the requirements of PH(3) distributions:

- λ_3 is real and positive,
- $a_2^2 - 3a_1 \ge 0$.

Due to the fact that the similarity transform leaves the eigenvalues unchanged, this generalization of property **P2** is a consequence of property **P1** and Theorem 1.

We can summarize the results of [6] as follows. It defines a similarity transformation of PH(3) distributions to a unicyclic representation. This transformation depends on a parameter, x_1. [6] also defines the range of parameter x_1,

(γ_ℓ, γ_u), where the transformed generator matrix is Markovian. The problem which remains open is how to set parameter x_1 such that the initial vector is Markovian, i.e., is a proper probability vector.

In the procedure in Figure 2 parameter ϕ is used to ensure the positivity of the initial vector. Unfortunately that approach is not sufficient when we have a non-Markovian $(\underline{v}, \boldsymbol{H})$ representation, as it is the case with the non-Markovian representation of Example 1. The next section investigates the range of x_1 where the initial vector is Markovian.

4 Canonical representation of PH(3) distributions

Using the similarity matrix defined in (7) the elements of the initial vector $\underline{\pi} = \underline{v}\boldsymbol{B}$ are:

$$\pi_1 = \frac{-\underline{v}\boldsymbol{H}\mathbb{1}}{x_1 - x_{13}} = \frac{d_1}{x_1 - x_{13}}, \tag{13}$$

$$\pi_2 = \frac{-\underline{v}(x_1\boldsymbol{I} + \boldsymbol{H})\boldsymbol{H}\mathbb{1}}{(x_1 - x_{13})x_2} = \frac{x_1 d_1 + d_2}{(x_1 - x_{13})x_2}, \tag{14}$$

$$\pi_3 = \frac{-\underline{v}(x_2\boldsymbol{I} + \boldsymbol{H})(x_1\boldsymbol{I} + \boldsymbol{H})\boldsymbol{H}\mathbb{1}}{(x_1 - x_{13})x_2 x_3} = \frac{x_1 x_2 d_1 + (x_1 + x_2)d_2 + d_3}{(x_1 - x_{13})x_2 x_3}, \tag{15}$$

where $d_i = -\underline{v}\boldsymbol{H}^i\mathbb{1}, i = 1, 2, 3$. The derivatives of the density function at zero are closely related to these parameters since $f^{(i)}(0) = d_{i+1} = -\underline{v}\boldsymbol{H}^{i+1}\mathbb{1}$. Consequently, for a Markovian $(\underline{v}, \boldsymbol{H})$ pair

P3) $d_1 > 0$, or $d_1 = 0$ and $d_2 \geq 0$,

must hold for having a non-negative density around zero.

The canonical form we propose in this paper is based on the following theorem.

Theorem 2 *If $(\underline{v}, \boldsymbol{H})$ has a Markovian representation, then the similarity transform with matrix $\boldsymbol{B}$, defined in (7), with parameter*

$$x_1 = \begin{cases} \max\{\gamma_2, \gamma_\ell\}, & \textbf{if } \underline{v}\,\boldsymbol{H}\,\mathbb{1} < 0, \\ \gamma_\ell, & \textbf{if } \underline{v}\,\boldsymbol{H}\,\mathbb{1} = 0, \end{cases} \tag{16}$$

$$\gamma_2 = -\frac{\underline{v}\boldsymbol{H}^2\mathbb{1}}{\underline{v}\boldsymbol{H}\mathbb{1}}, \tag{17}$$

provides a Markovian representation.

Proof Due to Theorem 1 and $\boldsymbol{B}\mathbb{1} = \mathbb{1}$, if $(\underline{v}, \boldsymbol{H})$ has a Markovian representation, then $\boldsymbol{B}^{-1}\boldsymbol{H}\boldsymbol{B}$ is Markovian, and $x_1 - x_{13}$, x_2, x_3 are positive, when x_1 is in the $[\gamma_\ell, \gamma_u]$ interval. Thus, it is enough to prove that vector $\{\pi_1, \pi_2, \pi_3\}$, defined in (13)-(15), is non-negative when x_1 takes is value according to (16).

$\pi_1 \geq 0$ follows immediately from (4), since if $(\underline{v}, \boldsymbol{H})$ has a Markovian representation, then its density is non-negative at zero.

When $\underline{v}\boldsymbol{H}\mathbb{1} = 0$, π_2 must be non-negative according to property **P3**.

When $\underline{v}\boldsymbol{H}\mathbb{1} < 0$, we can re-write (14) as:

$$\pi_2 = \frac{-\underline{v}\boldsymbol{H}\mathbb{1}}{(x_1 - x_{13})x_2}(x_1 - \gamma_2). \tag{18}$$

The first term of (18) is positive and the second term is non-negative when $x_1 = \max\{\gamma_2, \gamma_\ell\}$ according to (16).

For the analysis of π_3 we re-write (15) as

$$\pi_3 = \frac{1}{(x_1 - x_{13})x_2x_3} \underbrace{(x_1x_2d_1 + (x_1 + x_2)d_2 + d_3)}_{g(x_1)} \tag{19}$$

The first term is positive again, thus it remains to prove that $g(x_1) \geq 0$ if x_1 is according to (16). The first derivative of $g(x_1)$ has at most two roots:

$$\frac{d}{dx_1}\, g(x_1) = 0 \quad \Leftrightarrow \quad x_1 = \frac{a_2 \pm \sqrt{a_2^2 - 3a_1}}{3}. \tag{20}$$

If $\sqrt{a_2^2 - 3a_1} = 0$ then $\gamma_u = \gamma_\ell = \gamma_0$ and $x_1 = \gamma_\ell$ is the only valid value according to Theorem 1.

If $\sqrt{a_2^2 - 3a_1} > 0$ then the larger root of (20) equals to γ_0, hence $g(x_1)$ is a monotone function when $x_1 > \gamma_0$. In the $x_1 > \gamma_0$ region the increasing/decreasing behaviour of $g(x_1)$ is determined by the sign of the second derivative at $x_1 = \gamma_0$:

$$\frac{d^2}{dx_1^2}\, g(x_1)|_{x_1=\gamma_0} = \frac{-2(a_2d_1 + 4d_1\sqrt{a_2^2 - 3a_1} + 3d_2)}{3\sqrt{a_2^2 - 3a_1}} \tag{21}$$

When $d_1 = -\underline{v}\boldsymbol{H}\mathbb{1} = 0$, then (21) is non-positive because the numerator is non-positive due to property **P3** and the denominator is positive. In this case we have 2 subcases. If $d_2 = 0$, then $g(x_1)$ is constant and x_1 does not effect the sign of π_3, when $\gamma_\ell \le x_1 \le \gamma_u$. If $d_2 > 0$, then $g(x_1)$ is monotone decreasing and the minimal x_1 value of the valid range ($\gamma_\ell \le x_1 \le \gamma_u$ and $\gamma_2 \le x_1$) ensures the non-negativity of π_3 (assuming that a Markovian representation exists).

When $d_1 = -\underline{v}\boldsymbol{H}\mathbb{1} > 0$ we have

$$\begin{aligned} \frac{d^2}{dx_1^2}\, g(x_1)|_{x_1=\gamma_0} &= \frac{-2\, d_1\, (a_2 + 4\sqrt{a_2^2 - 3a_1} - 3\gamma_2)}{3\sqrt{a_2^2 - 3a_1}} \\ &= -\underbrace{\frac{2\, d_1}{3\sqrt{a_2^2 - 3a_1}}}_{>0} \left[3\underbrace{(\gamma_u - \gamma_2)}_{\ge 0} + \underbrace{(3\gamma_u - a_2)}_{>0} \right] \le 0, \end{aligned} \tag{22}$$

where the positivity of the under-braced terms follows from $\sqrt{a_2^2 - 3a_1} > 0$, and the non-negativity of the second term must hold since $(\underline{v}, \boldsymbol{H})$ has a Markovian representation (according to the condition of the theorem) and according to Theorem 1 it must have a unicyclic representation ($x_1 \le \gamma_u$) with a non-negative π_2 ($x_1 \ge \gamma_2$).

If the second derivative in (22) is negative then $g(x_1)$ is monotone decreasing at $x_1 > \gamma_0$ and the minimal x_1 value of the valid range ($\gamma_\ell \le x_1 \le \gamma_u$ and $\gamma_2 \le x_1$) ensures the non-negativity of π_3 (assuming that a Markovian representation exists).

If the second derivative in (22) equals to zero (i.e., $\gamma_u = \gamma_2$) it means that there is only a single x_1 value, $x_1 = \gamma_u = \gamma_2$, which results in a Markovian representation, because for $x_1 > \gamma_u$ matrix $\boldsymbol{A}$ is non-Markovian and for $x_1 < \gamma_2$ vector $\underline{\pi}$ is not a probability vector.

When $\sqrt{a_2^2 - 3a_1} > 0$, the possible behaviors of $g(x_1)$ and the associated choices of x_1 are summarized in the following table.

Cases	$g(x_1)$ at $x_1 > \gamma_0$	constraint of x_1	choice of x_1
$d_1 = 0, d_2 > 0$	mon. decreasing		minimal value
$d_1 = 0, d_2 = 0$	constant		minimal value
$d_1 > 0, \gamma_u > \gamma_2$	mon. decreasing		minimal value
$d_1 > 0, \gamma_u = \gamma_2$		$x_1 = \gamma_u = \gamma_2$	constraint

That is, (16) sets x_1 such that the obtained representation is Markovian when a Markovian representation exists. □

function Canonical–PH(3)–transformation
  input: $\underline{v}, \boldsymbol{H}$ (any matrix representation)
  output: $\underline{\pi}, \boldsymbol{A}$ (Canonical representation if $v, \boldsymbol{H}$ is a PH(3))
begin
  if $v_1 + v_2 + v_3 \neq 1$
    error "Probability mass at 0",
  $\lambda_1, \lambda_2, \lambda_3 \;=\;$ decreasingly ordered eigenvalues of $-\boldsymbol{H}$,
  if $\lambda_3 < 0$ **or** $\lambda_3 \in \mathbb{C}$ **or** $\underline{v}\,\boldsymbol{H}\,\mathbb{1} < 0$
    error "Invalid eigenvalues",
  $a_0 = \lambda_1\,\lambda_2\,\lambda_3, \quad a_1 = \lambda_1\,\lambda_2 + \lambda_1\,\lambda_3 + \lambda_2\,\lambda_3, \quad a_2 = \lambda_1 + \lambda_2 + \lambda_3$
  if $a_2^2 - 3\,a_1 < 0$
    error "Invalid characteristic polynomial",
  $\gamma_u = \frac{1}{3}\,(a_2 + 2\sqrt{a_2^2 - 3\,a_1}), \quad \gamma_0 = \frac{1}{3}\,(a_2 + \sqrt{a_2^2 - 3\,a_1}),$
  $\gamma_\ell = \begin{cases} \lambda_1 & \textbf{if } \lambda_1 \in \text{real}, \\ \gamma_0 & \textbf{if } \lambda_1 \in \text{complex}, \end{cases}$
  if $\underline{v}\,\boldsymbol{H}\,\mathbb{1} > 0$ **or** $(\underline{v}\,\boldsymbol{H}\,\mathbb{1} == 0$ **and** $\underline{v}\,\boldsymbol{H}^2\,\mathbb{1} > 0)$
    error "Negative density around 0",
  $\gamma_2 = \begin{cases} -\underline{v}\,\boldsymbol{H}^2\,\mathbb{1}\;/\;\underline{v}\,\boldsymbol{H}\,\mathbb{1} & \textbf{if } \underline{v}\,\boldsymbol{H}\,\mathbb{1} < 0, \\ 0 & \textbf{if } \underline{v}\,\boldsymbol{H}\,\mathbb{1} == 0, \end{cases}$
  if $\gamma_2 > \gamma_u$
    error "π_2 is negative",
  $x_1 = \max\{\gamma_2, \gamma_\ell\},$
  $x_{13} = x_1 - a_0 \;/\; (x_1^2 - a_2\,x_1 + a_1),$
  $x_2 = \frac{1}{2}\left(a_2 - x_1 + \sqrt{(a_2 - x_1)^2 \,-\, 4\,(x_1^2 - a_2\,x_1 + a_1)}\right),$
  $x_3 = \frac{1}{2}\left(a_2 - x_1 - \sqrt{(a_2 - x_1)^2 \,-\, 4\,(x_1^2 - a_2\,x_1 + a_1)}\right),$
  $\pi_1 = \underline{v}\,\boldsymbol{H}\,\mathbb{1} \;/\; (x_{13} - x_1),$
  $\pi_2 = \underline{v}\,(x_1\,\boldsymbol{I} + \boldsymbol{H})\,\boldsymbol{H}\,\mathbb{1} \;/\; ((x_{13} - x_1)\,x_2),$
  $\pi_3 = \underline{v}\,(x_2\,\boldsymbol{I} + \boldsymbol{H})\,(x_1\,\boldsymbol{I} + \boldsymbol{H})\,\boldsymbol{H}\,\mathbb{1} \;/\; ((x_{13} - x_1)\,x_2\,x_3),$
  if $\pi_3 < 0$
    error "π_3 is negative",
  return $\underline{\pi} = \begin{bmatrix} \pi_1 & \pi_2 & \pi_3 \end{bmatrix}, \; \boldsymbol{A} = \begin{bmatrix} -x_1 & 0 & x_{13} \\ x_2 & -x_2 & 0 \\ 0 & x_3 & -x_3 \end{bmatrix},$
end

Fig. 3. Canonical transformation of PH(3) distributions

4.1 The canonical transformation procedure

The transformation procedure is presented in Figure 3. If the procedure exits with one of the error messages then the input does not represent a PH(3) distribution. If the procedure completes, it gives back the canonical representation of the given PH(3) distribution, which is Markovian, minimal and unique as it is discussed in the next subsection.

If $\underline{v}$ is an arbitrary vector and $\boldsymbol{H}$ is an arbitrary matrix of cardinality three such that $(\underline{v}, \boldsymbol{H})$ represents an order-3 phase type distribution, then $(\underline{\pi}, \boldsymbol{A})$ is a Markovian representation of this PH(3) distribution.

$(\underline{\pi}, \boldsymbol{A})$ is unique, in the sense that for any $(\underline{v}, \boldsymbol{H})$ representation of a PH(3) distribution the procedure provides the same $(\underline{\pi}, \boldsymbol{A})$ pair.

The PH(3) distributions are known to be determined by five parameters. E.g., the first five moments or the five coefficients of the Laplace rational transform uniquely determines a PH(3) distribution. Although not obvious at first sight, the presented canonical form is also determined by exactly five independent parameters. In the unicyclic form [6] there are six parameters $(x_1, x_2, x_3, x_{13}, \pi_1, \pi_2)$ and in the transformation procedure presented in this paper one of these parameters is additionally set to a special value. The following constraint decreases the number of parameters to five:

f1) λ_1 real, $\gamma_2 < \gamma_\ell \quad \rightarrow \quad x_{13} = 0$,
f2) λ_1 complex, $\gamma_2 < \gamma_\ell \quad \rightarrow \quad x_1 = x_2$,
f3) $\gamma_\ell < \gamma_2 \quad \rightarrow \quad \pi_2 = 0$.

Indeed, these cases represent three different forms of the canonical representation.

It is an additional nice feature of the proposed canonical form that it is compatible with the widely used canonical representation of acyclic phase type distributions [4], since when $(\underline{v}, \boldsymbol{H})$ represents an order-3 acyclic phase type distribution, then form f1 gives Cumani's canonical representation of that distribution.

5 Unicyclic representation of PH(4) distributions

5.1 Transformation to unicyclic representation

Based on the structure of the canonical representation of PH(3) distributions we study the following unicyclic PH(4) structure.

Let $(\underline{v}, \boldsymbol{H})$ be a general matrix representation of a PH(4) distribution and $\{\lambda_1, \lambda_2, \lambda_3, \lambda_4\}$ its eigenvalues. The characteristic polynomial of $\boldsymbol{H}$ is $x^4 + a_3 x^3 + a_2 x^2 + a_1 x + a_0$ where

$$a_0 = \lambda_1\lambda_2\lambda_3\lambda_4, \tag{23}$$
$$a_1 = \lambda_1\lambda_2\lambda_3 + \lambda_1\lambda_2\lambda_4 + \lambda_1\lambda_3\lambda_4 + \lambda_2\lambda_3\lambda_4, \tag{24}$$
$$a_2 = \lambda_1\lambda_2 + \lambda_1\lambda_3 + \lambda_2\lambda_3 + \lambda_1\lambda_4 + \lambda_2\lambda_4 + \lambda_3\lambda_4, \tag{25}$$
$$a_3 = \lambda_1 + \lambda_2 + \lambda_3 + \lambda_4 \tag{26}$$

Theorem 3 *The* $(\underline{v}, \boldsymbol{H})$ *representation can be transformed to the* $(\underline{\pi}, \boldsymbol{A})$ *unicyclic form where* $\underline{\pi} = \underline{v}\boldsymbol{B}$*,* $\boldsymbol{A} = \boldsymbol{B}^{-1}\boldsymbol{H}\boldsymbol{B}$*,* $\boldsymbol{B}\mathbb{1} = \mathbb{1}$*, and matrix* $\boldsymbol{A}$ *has the form*

$$\boldsymbol{A} = \begin{bmatrix} -x_1 & 0 & x_{13} & x_{14} \\ x_2 & -x_2 & 0 & 0 \\ 0 & x_3 & -x_3 & 0 \\ 0 & 0 & x_4 & -x_4 \end{bmatrix}.$$

The similarity matrix of this transformation, $\boldsymbol{B} = [\underline{b_1}, \underline{b_2}, \underline{b_3}, \underline{b_4}]$*, is composed of the following column vectors*

$$\underline{b_1} = \frac{-\boldsymbol{H}\mathbb{1}}{x_1 - x_{13} - x_{14}}, \quad \underline{b_2} = \frac{(x_1\boldsymbol{I} + \boldsymbol{H})\underline{b_1}}{x_2},$$

$$\underline{b_3} = \frac{(x_2\boldsymbol{I} + \boldsymbol{H})\underline{b_2}}{x_3}, \quad \underline{b_4} = \frac{(x_3\boldsymbol{I} + \boldsymbol{H})\underline{b_3}}{x_4} - \frac{x_{13}\underline{b_1}}{x_4},$$

where x_1 *and* x_{13} *are arbitrary parameters and* x_{14}, x_2, x_3, x_4 *are the solution of the following set of equations*

$$a_0 = (x_1 - x_{13} - x_{14})x_2x_3x_4, \tag{27}$$
$$a_1 = (x_1 - x_{13})x_2x_3 + x_1x_2x_4 + x_1x_3x_4 + x_2x_3x_4, \tag{28}$$
$$a_2 = x_1x_2 + x_1x_3 + x_2x_3 + x_1x_4 + x_2x_4 + x_3x_4, \tag{29}$$
$$a_3 = x_1 + x_2 + x_3 + x_4 \ . \tag{30}$$

Proof The coefficients of the characteristic polynomial of $\boldsymbol{A}$ are given at the right hand side of (27)-(30). $\boldsymbol{H}$ and $\boldsymbol{A}$ are similar since their characteristic polynomials are identical due to (27)-(30). The columns of the similarity matrix $\boldsymbol{B}$ can be obtained from the columns of the matrix equation $\boldsymbol{H}\boldsymbol{B} = \boldsymbol{B}\boldsymbol{A}$, which are

$$\boldsymbol{H}\underline{b_1} = -x_1\underline{b_1} + x_2\underline{b_2} \ , \tag{31}$$
$$\boldsymbol{H}\underline{b_2} = -x_2\underline{b_2} + x_3\underline{b_3} \ , \tag{32}$$
$$\boldsymbol{H}\underline{b_3} = -x_3\underline{b_3} + x_4\underline{b_4} + x_{13}\underline{b_1} \ , \tag{33}$$
$$\boldsymbol{H}\underline{b_4} = -x_4\underline{b_4} + x_{14}\underline{b_1} \ . \tag{34}$$

Summing up (31)-(34) and using $\boldsymbol{B}\mathbb{1} = \mathbb{1}$ we have

$$\boldsymbol{H}\mathbb{1} = -x_1\underline{b_1} + x_{13}\underline{b_1} + x_{14}\underline{b_1} \ , \tag{35}$$

from which $\underline{b_1} = \dfrac{-\boldsymbol{H}\mathbb{1}}{x_1 - x_{13} - x_{14}}$. Consecutively substituting the result into (31)-(33) we obtain $\underline{b_2}, \underline{b_3}, \underline{b_4}$, respectively. □

Corollary 1 *Starting from (27) - (30) and having x_1 and x_{13} fixed, x_{14}, x_2, x_3, x_4 are obtained as the solution of an order-6 equation.*

Consequently, there is no symbolic transformation method to the $(\underline{\pi}, \boldsymbol{A})$ unicyclic form.

Corollary 1 remains valid also when $x_{13} = 0$.

5.2 Experimentation with the $(\underline{\pi}, \boldsymbol{A})$ representation

We have implemented the transformation method defined in Theorem 3 and additionally we implemented transformation methods to the following simple order-4 generators

$$\boldsymbol{A_{14}} = \boldsymbol{A} \text{ with } x_{13} = 0,$$

$$\boldsymbol{A_{13}} = \begin{bmatrix} -x_1 & 0 & x_{13} & 0 \\ x_2 & -x_2 & 0 & 0 \\ 0 & x_3 & -x_3 & 0 \\ 0 & 0 & x_4 & -x_4 \end{bmatrix}, \quad \boldsymbol{A_{24}} = \begin{bmatrix} -x_1 & 0 & 0 & 0 \\ x_2 & -x_2 & 0 & x_{24} \\ 0 & x_3 & -x_3 & 0 \\ 0 & 0 & x_4 & -x_4 \end{bmatrix}.$$

Having these transformation methods we checked if general PH(4) distributions can be transformed to the given specific forms. We found that none of the $\boldsymbol{A}_{13}$, $\boldsymbol{A}_{14}$ and $\boldsymbol{A}_{24}$ forms are sufficiently general to transform all PH(4) distributions into that form. Indeed, we found that it is usually impossible to transform between these forms. I.e., having a PH(4) distribution whose generator has the form of $\boldsymbol{A}_{24}$, it is commonly not possible to transform it to the form of $\boldsymbol{A}_{13}$ and $\boldsymbol{A}_{14}$, and so on.

In contrast, we found that the $(\underline{\pi}, \boldsymbol{A})$ representation, with properly chosen x_1 and x_{13} parameters, is general enough to cover all PH(4) examples we tried with.

The $(\underline{\pi}, \boldsymbol{A})$ representation is defined by nine parameters, x_1, x_2, x_3, x_4, x_{13}, x_{14}, π_1, π_2, π_3.

Assuming that the $(\underline{\pi}, \boldsymbol{A})$ representation is a candidate for the canonical representation of PH(4) distributions and that the canonical representation of PH(4) distributions contains the minimal number of parameters (which is seven), two additional constraints should apply. Some of the possible constraints are $x_{13} = 0$, $x_{14} = 0$, $\pi_2 = 0$, $\pi_3 = 0$, $x_1 = x_2$, $x_2 = x_3$, $x_3 = x_4$. Considering only these constraints we have a wide variety of different constraint-pairs. Some of them might be too restrictive, but e.g., $x_{13} = x_{14} = 0$ results in the acyclic subclass of PH(4) distributions.

6 Practical application of the PH(3) canonical form

6.1 Phase type fitting

There is a large number of PH distribution fitting methods available in the literature (for a survey see [11]). Some of them operate on the full PH class while others look for the solution in a subclass of the PH distributions. The most commonly used subclasses for fitting purposes are the APH class, the hyper-exponential distributions and the hyper-Erlang structure. At first sight these structural restrictions seem to decrease the efficiency of the fitting methods, since they look for the best fit in a smaller class of distributions. However, based on practical experiments, the opposite seems to be true: fitting a distribution with a restricted PH sub-class often provides better results, both in terms of distance and speed. The reason is that methods optimizing the full PH generator matrix and initial probability vector are often circling around different representations of the same distribution. Methods operating on the restricted PH sub-classes have an easier job, since they optimize fewer parameters.

The canonical form of PH(3) distributions can be utilized to develop more efficient PH fitting methods. These canonical forms are minimal representations, thus the optimization methods find the solution more easily.

Since there are three different canonical forms, the optimization has to be performed with all three structures and the best fit should be selected as a final result.

To show the benefits of canonical forms in distribution fitting we present some numerical examples. We developed a simple fitting method in Matlab. This method uses the built-in optimization procedure of Matlab (based on the line search algorithm) with the subject function set to cross entropy. Cross entropy is a popular quantity to measure the goodness of fit, because for discrete sample it equals to the log-likelihood. It is defined by:

$$-\hat{H} = \int_0^\infty \log \hat{f}(t)\, dF(t),$$

where $\hat{f}(t)$ is the density function of the fitting PH distribution and $F(t)$ denotes the cdf of the distribution to fit. The initial point was the best selected from hundred random PHs. During the numerical experiments the target distributions were W1, U1 and ME distributions defined in [12]:

$$f_{W1}(t) = \frac{\beta}{\eta}\left(\frac{t}{\eta}\right)^{\beta-1} e^{-\left(\frac{t}{\eta}\right)^{\beta}} \quad \text{with } \eta = 1, \beta = 1.5$$
$$f_{U1}(t) = 1, \qquad 0 \le t \le 1$$
$$f_{ME}(t) = \left(1 + \frac{1}{(2\pi)^2}\right)(1 - \cos(2\pi t))e^{-t}$$

	W1		U1		ME	
	Distance	Time	Distance	Time	Distance	Time
Full PH3	0.0018	56.5 s	21.4404	75 s	18.2410	38.6586 s
Form f1)	0.0018	19.4 s	16.67	9.8 s	18.0000	9.1 s
Form f2)	0.0018	11 s	16.67	14.3 s	17.9174	14.64 s
Form f3)	0.0018	24.17 s	16.67	11.2 s	17.9174	13.97 s

Table 1
Summary of the PH fitting results

The results are summarized in Table 1. In the first case (W1) all the considered PH structures resulted in equally good fits, however the canonical forms found the optimal fit faster. In the second test (U1) the fit with the full PH distribtion was slightly worse compared to the canonical forms. The table reflects the significant difference in the optimization speed in this case, too. In the third case (ME), the worst results are obtained by the full PH(3) structure, followed by the acyclic form f1), and the best results are achieved by f2) and f3). Again, optimization is faster with the usage of the canonical forms.

6.2 Moment matching with PH(3)

The presented transformation procedure is also applicable for moment matching with PH(3) distributions. For a given set of $\{\mu_1, \dots \mu_5\}$ moments we can generate a PH(3) distribution, whose first five moments are $\{\mu_1, \dots \mu_5\}$. This moments fitting procedure is composed of the following two steps.

- The first step is to compute a vector and matrix pair, $(\underline{v}, \boldsymbol{H})$, for which $i!\underline{v}(-\boldsymbol{H})^{-i}\mathbb{1} = \mu_i, i = 1, \dots, 5$. The procedure of Appie van de Liefvoort in [13] produces such $(\underline{v}, \boldsymbol{H})$ pair with a proper transformation of the closing

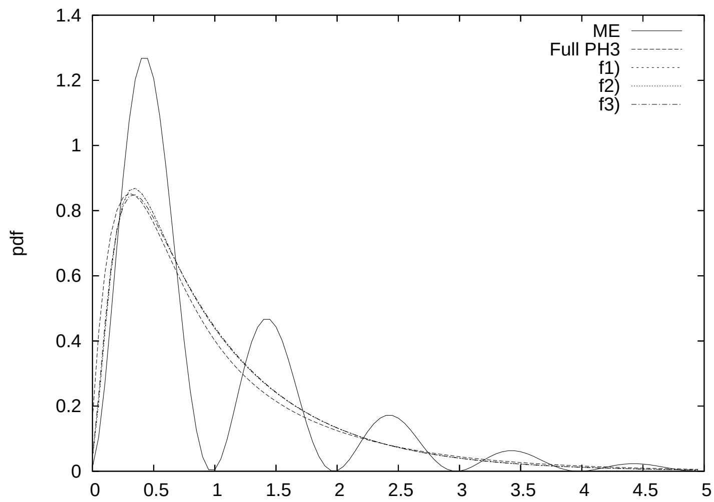

Fig. 4. PH fitting results for the ME distribution

vector [3].

- Starting from $(\underline{v}, \boldsymbol{H})$ the canonical PH(3) transformation procedure generates the Markovian representation of the PH(3) distribution, whose first five moments are $\{\mu_1, \dots \mu_5\}$.

Example 2 *When the first five moments are* $\{1.85111, 5.45136, 22.2838, 118.094, 774.513\}$ *the procedure of [13] gives*

$$\underline{v} = \begin{bmatrix} 1/3 & 1/3 & 1/3 \end{bmatrix}, \quad \boldsymbol{H} = \begin{bmatrix} -2.92628 & 44.7789 & -40.8522 \\ -0.398989 & -3.56926 & 3.0189 \\ -0.267678 & 2.9026 & -3.68557 \end{bmatrix},$$

and the canonical transformation procedure gives

$$\underline{\pi} = \begin{bmatrix} 0.0865519 & 0.124609 & 0.788839 \end{bmatrix}, \quad \boldsymbol{A} = \begin{bmatrix} -4.20997 & 0 & 0.360255 \\ 4.20997 & -4.20997 & 0 \\ 0 & 1.76118 & -1.76118 \end{bmatrix}.$$

6.3 Moments bounds of the PH(3) class

The presented transformation procedure is also applicable for evaluating the borders of the PH(3) distribution class. Indeed the above moment fitting pro-

[3] In [13] the initial and the closing vector are $\{1, 0, 0, \dots, 0\}$. In our case the closing vector is $\{1, 1, \dots, 1\}$, hence a similarity transformation is required as described in [9].

cedure terminates properly only when $\{\mu_1, \ldots \mu_5\}$ are the moments of a PH(3) distribution and the moment matching method aborts with some error if there is no PH(3) distribution whose moments are $\{\mu_1, \ldots \mu_5\}$.

To demonstrate the moment bounds of the PH(3) distribution set we first introduce the normalized moments $n_i = \frac{\mu_i}{\mu_1 \mu_{i-1}}$, $n_1 = 1$. The normalized moments are time unit independent "normalized" quantities, which carry the structural information of the moments apart from a time unit dependent scaling factor. n_2 is closely associated with the squared coefficient of variation (c_v^2) as $n_2 = c_v^2 + 1$.

Closed form symbolical bounds for the second and third normalized moments of APH(n) distributions are published in [7]. The moment formulas of PH(3) distributions are much more complex than the ones of the APH(3). We were not able to derive symbolic bounds for the normalized moments of PH(3) distributions. Instead, we investigated the moment bounds by an exhaustive search in the space of first five moments. In case of any given set of first five moments we applied the method detailed in Section 6.2 to check the PH(3) feasibility of the moments.

6.3.1 Bounds of the third normalized moment

With our numerical procedure we found that **the bounds of the third normalized moments of the APH(3) and of the general PH(3) distributions are the same**. Thus, in terms of the first three moments the general PH(3) distributions do not add extra flexibility over the APH(3) class. This statement is confirmed later by observing that the feasible range of n_4 reduces to zero at the upper and lower bounds of the third normalized moments of the APH(3) class (see Figure 8).

The bounds for n_3 are provided by [7] for the APH(3) class, depicted in Figure 5. Two sets can be distinguished according to the figure, n_3 is lower and upper bounded in the first one (corresponding to $4/3 \leq n_2 < 3/2$, called SET-1 in the sequel), while it is only lower bounded in the second (where $n_2 > 3/2$, called SET-2 in the sequel).

6.3.2 Bounds of the fourth normalized moment

We also investigated the fourth normalized moment as a function of the second and third normalized moments. We found that n_4 is lower bounded over the feasible n_2, n_3 range, the lower bound is depicted in Figure 6. We repeated the same experiments with the APH(3) class, and found that the lower bounds are different: somewhat tighter compared to the general PH(3) class.

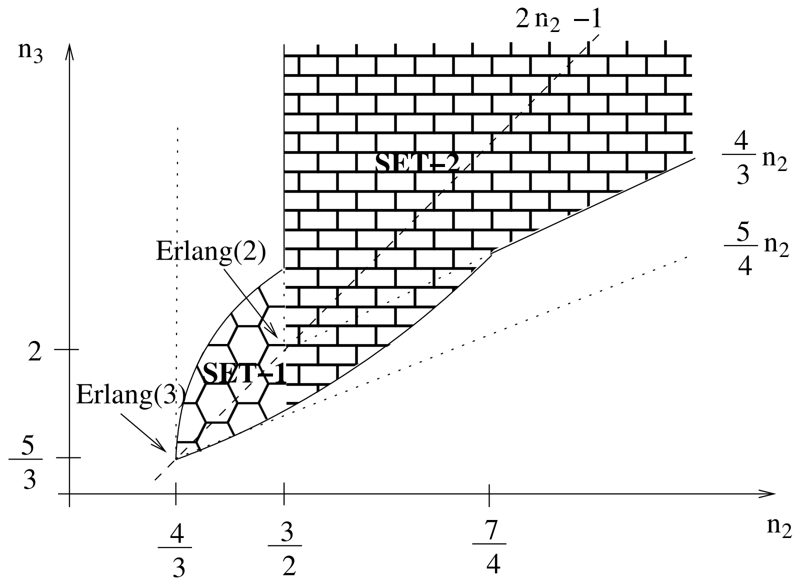

Fig. 5. Bounds of the third normalized moment of APH(3) distributions

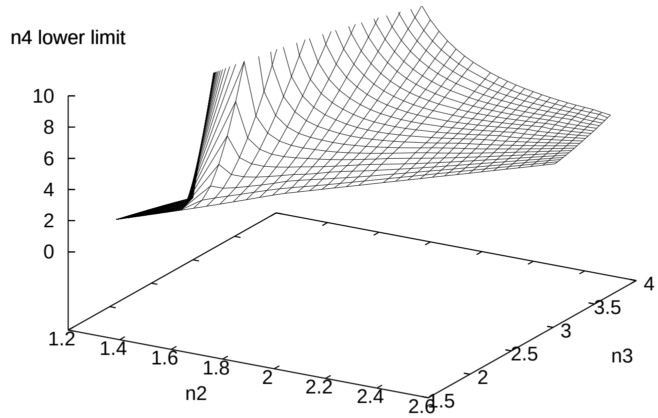

Fig. 6. Bounds of the fourth normalized moment as a function of (n_2, n_3)

When investigating the upper bound of n_4 we found that it is unbounded in SET-2 both in case of APH(3) and general PH(3) distributions. However in SET-1 it is upper bounded, with different upper bounds for the APH(3) and PH(3) case. Figure 7 shows the difference between the maximum feasible n_4 as a function of n_2 and n_3. This figure clearly shows that the difference between the bounds is small, the largest difference is 0.004.

The range of n_4 feasible by the APH(3) and PH(3) classes (thus, the upper bound minus the lower bound) is depicted in Figure 8 in SET-1. (The jagged lines are due to numerical inaccuracies). At the borders of the feasible regions the upper and lower bounds of n_4 are equal.

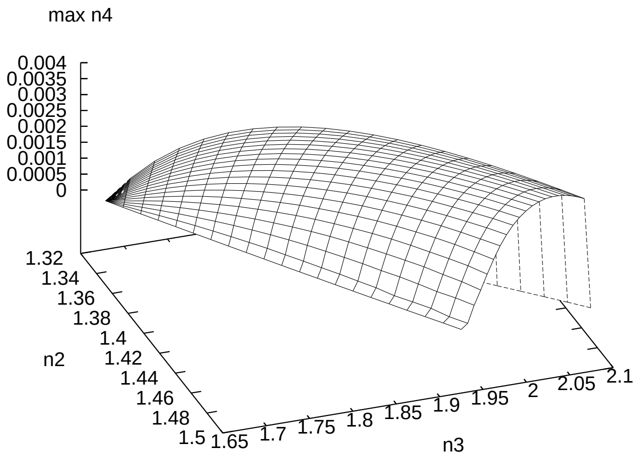

Fig. 7. The difference between the maximal feasible n_4 of PH(3) and APH(3) classes

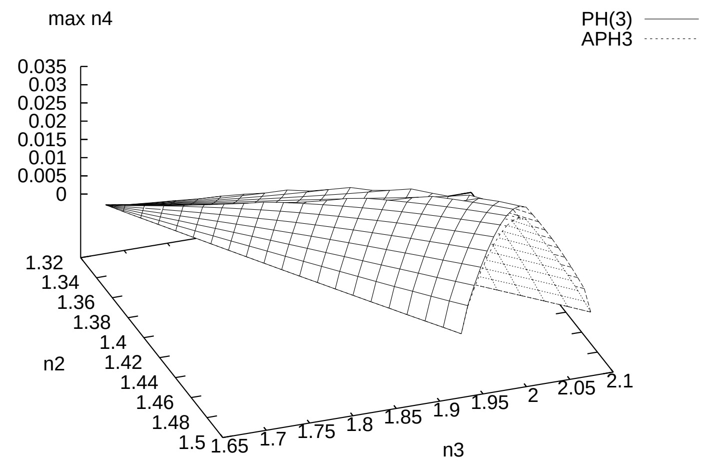

Fig. 8. The feasible range of n_4

6.3.3 The bounds of the fourth and fifth normalized moments

On the next figures we compare the feasible (n_4, n_5) regions of the general PH(3) and of the APH(3) distributions with different n_2, n_3 settings ($n_2 = 1.91696$ and $n_3 = 2.8699$ in Figure 9 and $n_2 = 1.49$ and $n_3 = 1.99$ in Figure 10).

The figures have been generated as follows. We applied the canonical transformation procedure for each pixel (representing an n_2, n_3, n_4, n_5 tuple) on the figure. The white pixels mean that the corresponding moments are not feasible with a PH(3) distribution. Light grey pixels are used where the solution is of form f1 (thus, an APH(3)), dark grey pixels where it is of form f2, and black pixels where it is of form f3. The results conform to our findings in the previous section, namely that the majority of the moments feasible by a PH(3) distribution is actually feasible for an APH(3) distribution, when $n_2 > 1.5$. In this case the feasible n_4, n_5 values seems to follow a similar structure as the one reported in [14] for feasible n_2, n_3 values (a lower bounded triangle with a bounded extension). We found that the difference between the PH(3) and APH(3) class gets to be significant when $n_2 < 1.5$. It seems that in this case the feasible n_4, n_5 values are bounded according to Figure 10 (and the lower bounded triangle part is missing).

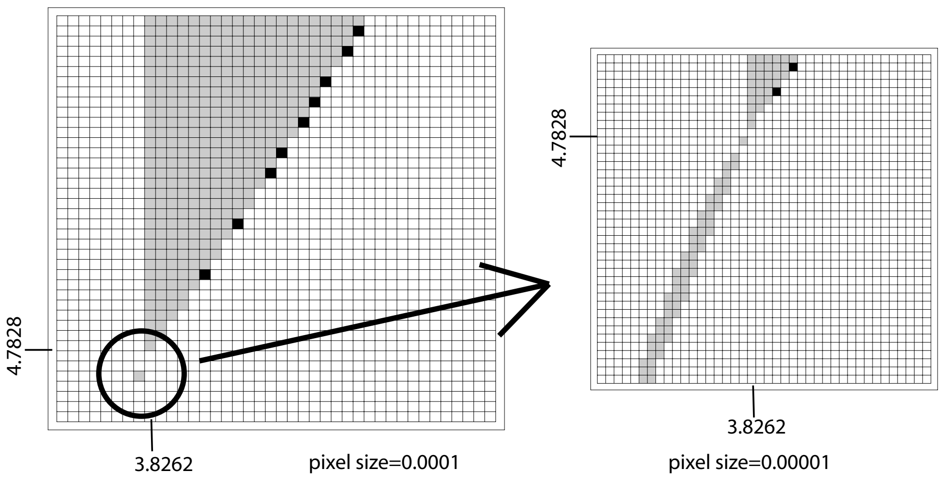

Fig. 9. Feasible n_4, n_5 normalized moments of PH(3) and APH(3) distributions when $n_2 = 1.91696$ and $n_3 = 2.8699$

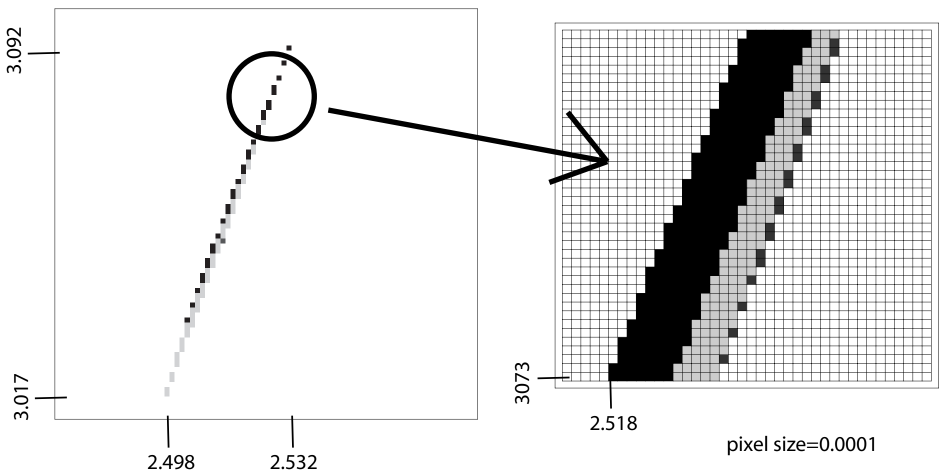

Fig. 10. Feasible n_4, n_5 normalized moments of PH(3) and APH(3) distributions when $n_2 = 1.49$ and $n_3 = 1.99$

Acknowledgement

The authors thank the help of Laura Fábián whose numerical investigations led to the canonical representation of PH(3) distributions. We also thank the detailed comments of the reviewers, which improved a lot the presentation of the paper.

7 Conclusion

In a number of practical applications it is very efficient to use the canonical representation of PH distributions that have as few parameters as possible. The problem of canonical representation of high order PH distributions is still open, but in this paper we presented a canonical representation for order-3 PH distributions. This canonical representation uses the unicyclic structure of He and Zhang and additionally ensures that the initial vector is non-negative.

We demonstrated potential applications of the canonical form and the associated transformation method through the analysis of the moments bounds of the PH(3) class. Furthermore, we presented and evaluated candidates for the canonical form of the PH(4) distribution class.

References

[1] G. Horváth, M. Telek, A canonical representation of order 3 phase type distributions, in: K. Wolter (Ed.), Formal methods and stochastic models for Performance Evaluation, LNCS 4748, Springer, 2007, pp. 48–62, ePEW.

[2] M. Neuts, Matrix-Geometric Solutions in Stochastic Models, John Hopkins University Press, Baltimore, MD, USA, 1981.

[3] G. Latouche, V. Ramaswami, Introduction to Matrix-Analytic Methods in Stochastic Modeling, Series on statistics and applied probability, ASA-SIAM, 1999.

[4] A. Cumani, On the canonical representation of homogeneous Markov processes modelling failure-time distributions, Microelectronics and Reliability 22 (1982) 583–602.

[5] D. R. Cox, A use of complex probabilities in the theory of stochastic processes, Proc. Cambridge Phil. Soc. 51 (1955) 313–319.

[6] Q.-M. He, H. Zhang, A note on unicyclic representation of PH-distributions, Stochastic Models 21 (2005) 465–483.

[7] A. Bobbio, A. Horváth, M. Telek, Matching three moments with minimal acyclic phase-type distributions, Stochastic Models 21 (2-3) (2005) 303–323.

[8] L. Lipsky, Queueing Theory: A linear algebraic approach, MacMillan, New York, 1992.

[9] M. Telek, G. Horváth, A minimal representation of Markov arrival processes and a moments matching method, Performance Evaluation 64 (9-12) (2007) 1153–1168.

[10] S. Mocanu, C. Commault, Sparse representations of phase-type distributions, Commun. Stat., Stochastic Models 15 (4) (1999) 759 – 778.

[11] A. Lang, J. L. Arthur, Parameter approximations for phase-type distributions, in: Proc. 1st Int. Conf. on Matrix-Analytic Methods in Stochastic Models, Notable Publications, 1996, pp. 151–206.

[12] A. Bobbio, M. Telek, A benchmark for PH estimation algorithms: results for Acyclic-PH, Stochastic Models 10 (1994) 661–677.

[13] A. van de Liefvoort, The moment problem for continuous distributions, Tech. rep., University of Missouri, WP-CM-1990-02, Kansas City (1990).

[14] M. Telek, A. Heindl, Matching moments for acyclic discrete and continuous phase-type distributions of second order, International Journal of Simulation Systems, Science & Technology 3(3-4) (2002) 47–57, special Issue on: Analytical & Stochastic Modelling Techniques.