A Canonical Representation of Order 3 Phase Type Distributions*

Gábor Horváth and Miklós Telek

Department of Telecommunications
Budapest University of Technology and Economics
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 recommend a canonical representation of the PH(3) class and present a transformation procedure to obtain the canonical representation based on any (not only Markovian) vector-matrix representation of the distribution.
Using this canonical transformation method we evaluate the moment bounds of the PH(3) distribution set and present the results of our numerical investigations.

Keywords: Phase Type Distribution, Canonical Form, Moment Bounds.

1 Introduction

The 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 [10, 7]. 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 α, and generator matrix $\boldsymbol{A}$. This representation is known to be non-unique and non-minimal, thus there might be a vector α' 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 α' and square matrix $\boldsymbol{A}'$ is n (since $\boldsymbol{A}$ has n^2 elements and α 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.

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

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 [3] 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 α 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 [5], where the unicyclic representation of PH(3) distributions is proved. Indeed, the presented canonical representation is unicyclic, but it extends the results of [5] with the careful analysis of the initial probability vector of the canonical representation, which is not taken into consideration in [5], because it aims to solve a different problem.

With the help 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 3 moments of PH(3) distributions are provided in [2], but the behaviour of the 4th and 5th 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. Section 5 lists some applications of the canonical form and the associated transformation method and Section 6 demonstrates the behaviour of the parameters used in the transformation procedure. The paper is concluded in Section 7.

2 PH(3) distributions

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

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

where the row vector 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., $v\mathbb{1} = 1$. The density, the Laplace transform and the moments of $\mathcal{X}$ are

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

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

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

When the cardinality of vector $\boldsymbol{v}$ and of square matrix $\boldsymbol{H}$ is 3, we have the following cases:

- If $f(t) \geq 0$ and $\int_0^\infty f(t)dt = 1$, then X has an order 3 matrix exponential (ME(3)) distribution. The elements of $\boldsymbol{v}$ and $\boldsymbol{H}$ may be arbitrary real numbers.
- If $\boldsymbol{v}$ 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 X has a PH(3) distribution. (The set of PH(3) distributions form a true subset of the ME(3) set.)

Vector $\boldsymbol{v}$ is a probability vector when $\boldsymbol{v}_i \geq 0$, $\boldsymbol{v}\mathbb{1} = 1$ and matrix $\boldsymbol{H}$ is a transient Markovian generator when $\boldsymbol{H}_{ii} < 0$, $\boldsymbol{H}_{ij} \geq 0$ 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* $(v, \boldsymbol{H})$ *representation is a* Markovian representation*, if* 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 $(v, \boldsymbol{H})$ pair is a density function. We have the following necessary conditions (those that we use in the sequel, [9]):

- 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:

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

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

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

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

Example 1.

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

and

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

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

Now, we can refine the above definition of PH(3) distributions with the help of similarity transform.

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

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 [5]. We summarize the related results, in a bit modified way, for completeness.

Theorem 1. *[5] If $(v, \boldsymbol{H})$ is a Markovian representation of a PH(3) distribution then it can be similarity transformed to the following unicyclic Markovian representation*

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

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

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

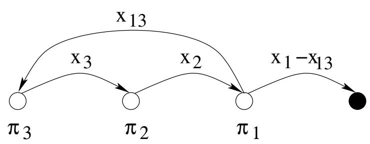

Fig. 1. The structure of the considered unicyclic PH(3) distribution

The main difference between Theorem 1 ([5]) and the goal of this paper is that Theorem 1 assumes that $(v, \boldsymbol{H})$ is Markovian, while we look for a transformation

function PH(3)–to–unicyclic PH(3)
  input: $v, \boldsymbol{H}$ (Markovian)
  output: $\pi, \boldsymbol{A}$ (unicyclic)
begin
  $\lambda_1, \lambda_2, \lambda_3$ = decreasingly ordered eigenvalues of $-\boldsymbol{H}$,
  $a_0 = \lambda_1\, \lambda_2\, \lambda_3, \quad a_1 = \lambda_1\, \lambda_2 + \lambda_1\, \lambda_3 + \lambda_2\, \lambda_3, \quad a_2 = \lambda_1 + \lambda_2 + \lambda_3,$
  $\gamma_u = \frac{1}{3}\,(a_2 + 2\sqrt{a_2^2 - 3\,a_1}), \quad \gamma_0 = \frac{1}{3}\,(a_2 + \sqrt{a_2^2 - 3\,a_1}),$
  $\gamma_\ell = \begin{array}{l} \lambda_1 \textbf{ if } \lambda_1 \in \text{real}, \\ \gamma_0 \textbf{ if } \lambda_1 \in \text{complex}, \end{array}$
  $\phi = \max\{-\boldsymbol{H}_{1,1},\ -\boldsymbol{H}_{2,2},\ -\boldsymbol{H}_{3,3}\},$
  $x_1 = \max\{\phi,\ \gamma_\ell\},$
  $x_{13} = x_1 - a_0\ /\ (x_1^2 - a_2\, x_1 + a_1),$
  $x_2 = \frac{1}{2}\ \ a_2 - x_1 + \sqrt{(a_2 - x_1)^2\ -\ 4\,(x_1^2 - a_2\, x_1 + a_1)}\ \ ,$
  $x_3 = \frac{1}{2}\ \ a_2 - x_1 - \sqrt{(a_2 - x_1)^2\ -\ 4\,(x_1^2 - a_2\, x_1 + a_1)}\ \ ,$
  $\pi_1 = v\, \boldsymbol{H}\, 1\!\!1\ /\ (x_{13} - x_1),$
  $\pi_2 = v\,(x_1\, \boldsymbol{I} + \boldsymbol{H})\, \boldsymbol{H}\, 1\!\!1\ /\ (x_{13} - x_1)\, x_2,$
  $\pi_3 = v\,(x_2\, \boldsymbol{I} + \boldsymbol{H})\,(x_1\, \boldsymbol{I} + \boldsymbol{H})\, \boldsymbol{H}\, 1\!\!1\ /\ (x_{13} - x_1)\, x_2\, x_3,$
  return $\pi = \ \ \pi_1\ \pi_2\ \pi_3\ \ , \ \boldsymbol{A} = \begin{bmatrix} -x_1 & 0 & x_{13} \\ x_2 & -x_2 & 0 \\ 0 & x_3 & -x_3 \end{bmatrix},$
end

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

which is applicable for any non-Markovian $(v, \boldsymbol{H})$ representation. For example the procedure of Figure 2 gives a proper unicyclic representation when it is called with the $(v, \boldsymbol{H})$ pair of Example 1, but it gives complex results when it is called with the $(z, \boldsymbol{G})$ 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 characteristic polynomial of $-\boldsymbol{H}$, i.e.,

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

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 by the column vectors $\{b_1, b_2, b_3\}$ where

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

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. [5] 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 $\gamma_\ell \leq \gamma_u$.

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

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

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

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 [5] as follows. It defines a similarity transformation of PH(3) distributions to a unicyclic representation. This transformation depends on a parameter, x_1. [5] 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 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 $(v, \boldsymbol{H})$ 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 $\pi = v\boldsymbol{B}$ are:

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

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

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

where $d_i = -v\boldsymbol{H}^i\mathbb{1}, i = 1, 2, 3$. The derivatives of the density function at 0 are closely related with these parameters since $f^{(i)}(0) = d_{i+1} = -v\boldsymbol{H}^{i+1}\mathbb{1}$. Consequently, for a Markovian $(v, \boldsymbol{H})$ 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 $(v, \boldsymbol{H})$ has a Markovian representation, then the similarity transform with matrix $\boldsymbol{B}$, defined in (7), with parameter*

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

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

provides a Markovian representation.

Proof Due to Theorem 1 and $\boldsymbol{B}\mathbb{1} = \mathbb{1}$ it is enough to prove that $\pi_1, \pi_2, \pi_3 \geq 0$ in (13), (14), (15), for some x_1 in the $[\gamma_\ell, \gamma_u]$ interval, where $x_1 - x_{13}$, x_2, x_3 are positive and $[\gamma_\ell, \gamma_u]$ is not empty.

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

When $v\boldsymbol{H}\mathbb{1} = 0$, π_2 must be non-negative according to property P3. When $v\boldsymbol{H}\mathbb{1} < 0$, we can re-write (14) as:

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

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

$$\pi_3 = \frac{1}{(x_1 - x_{13})x_2 x_3} \underbrace{(x_1 x_2 d_1 + (x_1 + x_2)d_2 + d_3)}_{g(x_1)} \tag{19}$$

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

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

The larger root 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 $x_1 = \gamma_0$:

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

When $d_1 = -v\boldsymbol{H}\mathbb{1} = 0$, then the second derivative is non-positive due to property P1 and P3 and when $d_1 = -v\boldsymbol{H}\mathbb{1} > 0$ we have

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

where the non-negativity of the first under-braced term follows from property P1, the non-negativity of the second term must hold since $(v, \boldsymbol{H})$ is Markovian and according to Theorem 1 it must have a unicyclic representation ($x_1 \leq \gamma_u$) with a non-negative π_2 ($x_1 \geq \gamma_2$). The non-negativity of the third under-braced term follows from $Re(\lambda_1) \leq \gamma_u$ and the fact that $Re(\lambda_1) \geq Re(\lambda_2) \geq Re(\lambda_3)$.

If the second derivative in (22) equals to 0 it means that there is only a single x_1 value, $x_1 = \gamma_u$, which results in a Markovian representation.

If the second derivative in (22) is negative then $g(x_1)$ has a local maximum at $x_1 = \gamma_0$, and it is monotone decreasing function at $x_1 > \gamma_0$. To obtain a valid generator $x_1 > \gamma_\ell$ must hold as well, and since $\gamma_\ell \geq \gamma_0$, the largest feasible π_3 value is obtained at $x_1 = \gamma_\ell$. Since $(v, \boldsymbol{H})$ has a unicyclic representation according to Theorem 1, π_3 is non-negative in this point. □

We demonstrate the numerical behaviour of π_2 and π_3 as a function of x_1 in Section 6.

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.

function Canonical–PH(3)–transformation
    input: $v, \boldsymbol{H}$ (any matrix representation)
    output: $\pi, \boldsymbol{A}$ (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** $v\, \boldsymbol{H}\, \mathbb{1} < 0$
        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",
    $\gamma_u = \frac{1}{3}\,(a_2 + 2\sqrt{a_2^2 - 3\, a_1}), \quad \gamma_0 = \frac{1}{3}\,(a_2 + \sqrt{a_2^2 - 3\, a_1}),$
    $\gamma_\ell = \begin{cases} \lambda_1 \text{ if } \lambda_1 \in \text{real}, \\ \gamma_0 \text{ if } \lambda_1 \in \text{complex}, \end{cases}$
    if $v\, \boldsymbol{H}\, \mathbb{1} > 0$ **or** $(v\, \boldsymbol{H}\, \mathbb{1} == 0$ **and** $v\, \boldsymbol{H}^2\, \mathbb{1} > 0)$
        error "Negative density around 0",
    $\gamma_2 = \begin{cases} -v\, \boldsymbol{H}^2\, \mathbb{1} \,/\, v\, \boldsymbol{H}\, \mathbb{1} & \text{if } v\, \boldsymbol{H}\, \mathbb{1} < 0, \\ 0 & \text{if } v\, \boldsymbol{H}\, \mathbb{1} == 0, \end{cases}$
    if $\gamma_2 > \gamma_u$
        error "π_2 is negative",
    $x_1 = \max\{\gamma_2, \gamma_\ell\},$
    $x_{13} = x_1 - a_0 \,/\, (x_1^2 - a_2\, x_1 + a_1),$
    $x_2 = \frac{1}{2}\ \ a_2 - x_1 + \sqrt{(a_2 - x_1)^2 \ - \ 4\,(x_1^2 - a_2\, x_1 + a_1)} \ ,$
    $x_3 = \frac{1}{2}\ \ a_2 - x_1 - \sqrt{(a_2 - x_1)^2 \ - \ 4\,(x_1^2 - a_2\, x_1 + a_1)} \ ,$
    $\pi_1 = v\, \boldsymbol{H}\, \mathbb{1} \,/\, (x_{13} - x_1),$
    $\pi_2 = v\,(x_1\, \boldsymbol{I} + \boldsymbol{H})\, \boldsymbol{H}\, \mathbb{1} \,/\, (x_{13} - x_1)\, x_2,$
    $\pi_3 = v\,(x_2\, \boldsymbol{I} + \boldsymbol{H})\,(x_1\, \boldsymbol{I} + \boldsymbol{H})\, \boldsymbol{H}\, \mathbb{1} \,/\, (x_{13} - x_1)\, x_2\, x_3,$
    if $\pi_3 < 0$
        error "π_3 is negative",
    return $\pi = \ \ \pi_1\ \pi_2\ \pi_3 \ \ , \ \boldsymbol{A} = \begin{bmatrix} -x_1 & 0 & x_{13} \\ x_2 & -x_2 & 0 \\ 0 & x_3 & -x_3 \end{bmatrix},$
end

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

4.2 Properties of the proposed canonical form

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

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

The PH(3) distributions are known to be determined by 5 parameters. E.g., the first 5 moments, or the 5 coefficients of the Laplace rational transform uniquely determines a PH(3) distribution. Although not obvious from the first sight, the presented canonical form is also determined by exactly 5 independent parameters. In the unicyclic form [5] there are 6 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 5:

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 $(v, \boldsymbol{H})$ represents an order 3 acyclic phase type distribution, then form f1 gives the Cumani's canonical representation of that distribution.

5 Practical application of the canonical form and the transformation procedure

5.1 Phase type fitting

The currently available PH(3) fitting methods are either restricted to the acyclic subclass of PH(3) distributions (e.g., [6]) or they are not restricted, but their performance is limited by the fact that they optimize too many parameters (e.g., [1]). The canonical representation allows to eliminate the weakness of the second type of fitting methods. Using the 3 potential forms of the canonical representation one can compose 3 fitting methods (for form f1, f2 and f3) with minimal number of parameters and the best of the 3 gives the best fit over the whole PH(3) class.

5.2 Moment matching with PH(3)

The presented transformation procedure is also applicable for moment matching with PH(3) distributions. For a given set of $\{\mu_1, \ldots \mu_5\}$ moments we can generate a PH(3) distribution, whose first five moments are the same. This moments fitting procedure is composed by the following 2 steps.

– The first step is to compute a vector and matrix pair, $v, \boldsymbol{H}$, for which $i!v(-\boldsymbol{H})^{-i}\mathbb{1} = \mu_i, i = 1, \ldots, 5$. The procedure of Appie van de Liefvoort in [12] produces such $v, \boldsymbol{H}$ pair with a proper transformation of the closing vector[1].
– Starting from $v, \boldsymbol{H}$ the canonical PH(3) transformation procedure generates the Markovian representation of the PH(3) distribution, whose first 5 moments are $\{\mu_1, \ldots \mu_5\}$.

Example 2. For example, when the first 5 moments are $\{1.85111, 5.45136, 22.2838, 118.094, 774.513\}$ the procedure of [12] gives

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

and the canonical transformation procedure gives

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

5.3 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 described moment fitting procedure 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 $n_i = \frac{\mu_i}{\mu_1 \mu_{i-1}}$. The normalized moments are time unit independent “normalized” quantities, which carry the structural information of the moments apart of a time unit dependent scaling factor. n_2 is closely associated with the squared coefficient of variation, c_v^2. $n_2 = c_v^2 + 1$. The second and third normalized moments of APH(n) distributions are studied in [2, 11].

Example 3. We study the fourth and fifth normalized moments of PH(3) distributions with two pairs of second and third normalized moments.

The first point, $n_2 = 1.6$ and $n_3 = 2.3$, is taken in the $n_2 < 2$ range, where the coefficient of variation is less than 1, while the second point, $n_2 = 2.018$ and $n_3 = 3.036$, is taken in the $n_2 > 2$ range. The feasible range of normalized moment n_4 and n_5 are depicted in Figure 4 and 5, respectively. It is interesting to see that the fifth normalized moment, n_5, is both upper and lower bounded as well in the first case, while it is only lower bounded in the second case.

The presented canonical transformation procedure gives a tool for the numerical investigation of the moments bounds, but the detailed qualitative investigation of these moments bounds is out of the scope of this paper.

[1] In [12] the initial and the closing vector are $\{1, 0, 0, \ldots, 0\}$. In our case the closing vector is $\{1, 1, \ldots, 1\}$, hence a similarity transformation is required.

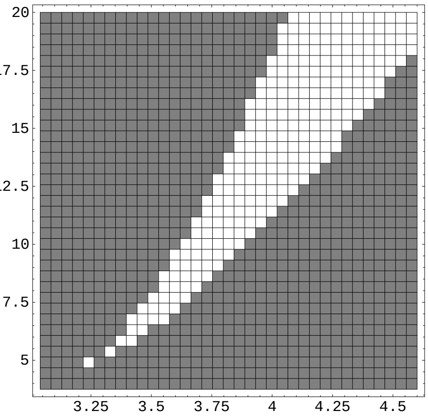

Fig. 4. Legal n_4, n_5 normalized moments of PH(3) distributions when $n_2 = 1.6$ and $n_3 = 2.3$

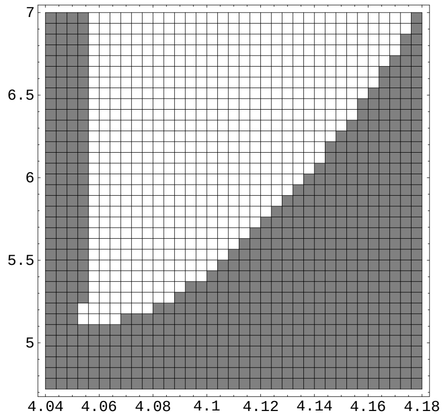

Fig. 5. Legal n_4, n_5 normalized moments of PH(3) distributions when $n_2 = 2.018$ and $n_3 = 3.036$

6 Numerical examples

6.1 Dependence of bounding quantities on the matrix elements

We demonstrate the dependence of the bounding quantities of the canonical representation, $\gamma_0, \gamma_\ell, \gamma_2, \gamma_u$, on the elements of the PH representation through some numerical examples.

We study the dependence of the bounding quantities on the initial distribution using the following representation, $v = \begin{bmatrix} x \; 0.8 - x \; 0.2 \end{bmatrix}$ and $\boldsymbol{H} = \begin{bmatrix} -3 & 0 & 2.5 \\ 2 & -2 & 0 \\ 0 & 1 & -1 \end{bmatrix}$. The result is presented in Figure 6. In this case all quantities which are associated with the Markovian representation of the generator matrix (the coefficients of the characteristic polynomial, a_0, a_1, a_2, the eigenvalues, $\lambda_1, \lambda_2, \lambda_3$ and the associated bounding quantities, $\gamma_0 = 2.57735, \gamma_\ell = 2.57735, \gamma_u = 3.1547$) remain constant and only γ_2 changes which is associated with the Markovian representation of the initial vector. The x_1 value of the canonical representation is determined by γ_ℓ if $x < 0.660434$ and it is determined by γ_2 for larger x values.

The dependence on the feedback element, x_{13}, is investigated using $v = \begin{bmatrix} 0.62 \; 0.246 \; 0.134 \end{bmatrix}$ and $\boldsymbol{H} = \begin{bmatrix} -3 & 0 & x \\ 2 & -2 & 0 \\ 0 & 1 & -1 \end{bmatrix}$. The curves in Figure 7 indicates another behaviour. $\gamma_2 = 2.20645$ is independent of the feedback element, but in this case some other, generator matrix related quantities, are constant as well. The a_1 and the a_2 coefficients of the characteristic polynomial are constant. As a consequence $\gamma_0 = 2.57735$ and $\gamma_u = 3.1547$ are independent on x_1. Only the a_0 coefficient of the characteristic polynomial changes with x, which makes the eigenvalues depend on x as well. In the $x \in \{0, 0.2\}$ range the λ_1 eigenvalue is real and it determines the x_1 value of the canonical representation. When x is greater γ_0 determines the x_1 value.

The most complicated behaviour has been obtained when the intensity of a transition is changing. For $v = \begin{bmatrix}0.62\ 0.246\ 0.134\end{bmatrix}$ and $\boldsymbol{H} = \begin{bmatrix} -3 & 0 & 1 \\ x & -x & 0 \\ 0 & 1 & -1 \end{bmatrix}$ the bounding quantities are depicted in Figure 8. In this case γ_2 has a linearly decreasing behaviour starting from 3, the γ_u function has a minimum at $x = 1$, the λ_1 eigenvalue is real and equals to γ_ℓ while $x < 1.1$ and it is complex and $\gamma_\ell = \gamma_0$ when $x > 1.1$. γ_0 is an increasing function of x starting from 2.21525. The x_1 value equals to γ_2 when $x < 2.01$ and it equals to γ_0 for larger x.

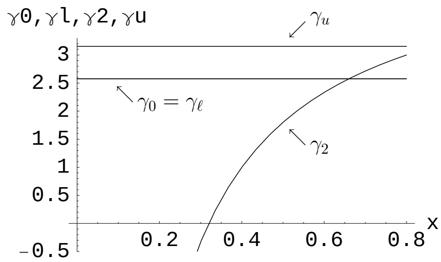

Fig. 6. Dependence of bounding quantities on the initial vector

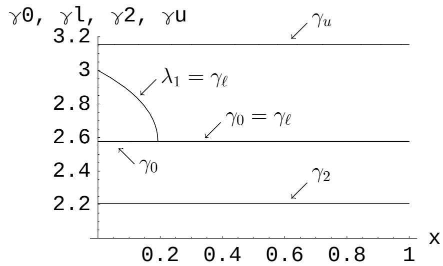

Fig. 7. Dependence of bounding quantities on the feedback element

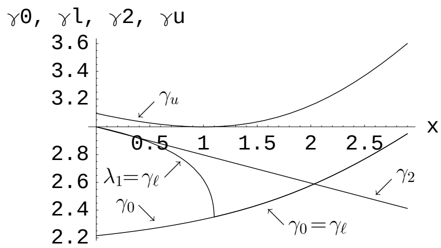

Fig. 8. Dependence of bounding quantities on a transition rate

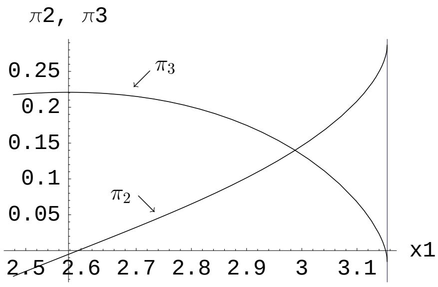

Fig. 9. The function of π_2 π_3 as a function of x_1

6.2 Dependence of the unicyclic representation on x_1

In the majority of the cases γ_ℓ and/or γ_u allows a Markovian representation. The case when $v = \begin{bmatrix}0.72\ 0.146\ 0.134\end{bmatrix}$ and $\boldsymbol{H} = \begin{bmatrix} -3 & 0 & 2.025 \\ 2 & -2 & 0 \\ 0 & 1 & -1 \end{bmatrix}$, is different, since in this case $\gamma_2 = 2.59444 > \gamma_\ell = \gamma_0 = 2.57735$ and $\gamma_u = 3.1547 > \gamma_z = 3.15186$, i.e., none of γ_ℓ and γ_u results in a Markovian representation. The behaviours of π_2 and π_3 are depicted in Figure 9. The y axis is set to $\gamma_\ell = \gamma_0$ and the grid line to γ_u. It is also visible that π_3 has a maximum at γ_0.

Acknowledgement

The authors thank the effort of Laura Fábián whose numerical investigations led to the basic idea of this paper.

7 Conclusion

In a number of practical applications it is very efficient using 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 positive.

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.

References

1. S. Asmussen and O. Nerman. Fitting Phase-type distributions via the EM algorithm. In *Proceedings: "Symposium i Advent Statistik"*, pages 335–346, Copenhagen, 1991.
2. A. Bobbio, A. Horváth, and M. Telek. Matching three moments with minimal acyclic phase-type distributions. *Stochastic Models*, 21(2-3):303–323, 2005.
3. D. R. Cox. A use of complex probabilities in the theory of stochastic processes. *Proc. Cambridge Phil. Soc.*, 51:313–319, 1955.
4. A. Cumani. On the canonical representation of homogeneous Markov processes modelling failure-time distributions. *Microelectronics and Reliability*, 22:583–602, 1982.
5. Qi-Ming He and Hanqin Zhang. A note on unicyclic representation of ph-distributions. *Stochastic Models*, 21:465–483, 2005.
6. A. Horváth and M. Telek. PhFit: A general purpose phase type fitting tool. In *Tools 2002*, pages 82–91, London, England, April 2002. Springer, LNCS 2324.
7. G. Latouche and V. Ramaswami. *Introduction to Matrix-Analytic Methods in Stochastic Modeling*. Series on statistics and applied probability. ASA-SIAM, 1999.

8. L. Lipsky. *Queueing Theory: A linear algebraic approach.* MacMillan, New York, 1992.
9. Stefanita Mocanu and Christian Commault. Sparse representations of phase-type distributions. *Commun. Stat., Stochastic Models*, 15(4):759 – 778, 1999.
10. M. Neuts. *Matrix-Geometric Solutions in Stochastic Models.* John Hopkins University Press, Baltimore, MD, USA, 1981.
11. T. Osogami and M. Harchol-Balter. A closed form solution for mapping general distributions to minimal ph distributions. In *International Conference on Performance Tools – TOOLS 2003*, pages 200–217, Urbana, IL, USA, Sept 2003. Springer, LNCS 2794.
12. A. van de Liefvoort. The moment problem for continuous distributions. Technical report, University of Missouri, WP-CM-1990-02, Kansas City, 1990.