Canonical representation of discrete phase type distributions of order 2 and 3

János Papp[1], Miklós Telek[1,2]

[1] Technical University of Budapest, Hungary

[2] MTA-BME Information systems research group

e-mail: papp.janos.90@gmail.com, telek@hit.bme.hu

August 8, 2013

Abstract

In spite of the fact that discrete phase type (DPH) distributions are used almost as often as their continuous phase type (CPH) distributions canonical representations are not available for general (cyclic) DPH distributions.

In this paper we investigate the canonical representation of DPH distributions of order 2 and 3. During the course of this investigation we find DPH properties which essentially differ from the ones of CPH, a subset of DPH distributions which can be mapped with CPH ones with respect to their canonical representations and an other subset of DPH distributions which requires a completely different treatment and exhibits different canonical forms.

1 Introduction

Stochastic performance models were restricted to "memory less" distributions (exponential in case of continuous time models and geometrical in case of discrete time models) for a long time in order to utilize the nice computational properties of discrete state Markov models. Phase-Type distributions [1, 2] have been introduced for relaxing this modeling limitation on the considered distributions, while maintaining the nice Markovian behavior.

For a period of time continuous time stochastic models with continuous phase type (PH) distributions were more often applied in performance modeling of computer and communication systems, but also in this period the analysis of the continuous time models were often based on the method of embedded Markov chains, which transforms the analysis problem into discrete time. Later on, with the rise of slotted time telecommunication protocols (e.g., ATM) discrete time models become primary modeling tools (for recent surveys see [3, 4]). As a consequence, approximation of experimental data set with CPH gained more attention for a period of time. Especially, the acyclic subset of CPH distributions gained popularity due to the simple canonical forms available for their representation [5]. The use of acyclic PH distributions has a further important consequence. A lot of properties of the acyclic CPH and the acyclic DPH distributions are identical. For example the same canonical representations apply for acyclic DPH distributions as for acyclic CPH ones [6]. Due to this similarity the problem of fitting DPH distributions was considered to be similar to the one of fitting CPH distributions, but this similarity is limited to the acyclic PH distributions only, as it is indicated through a counterexample in [7]. In this paper we investigate how far the similarities of cyclic CPH and cyclic DPH distributions extend and study the cases when the canonical representation of CPH and DPH distributions differ.

The canonical representation of order 3 CPH distributions is provided in [8]. Basically we investigate the similar results for order 2 and 3 DPH distributions. The rest of the paper is organized as follows. The next section provides a short introduction of DPH distributions. Section 3 and 4 present the canonical forms for order 2 and 3 DPH distributions respectively. The paper is concluded in Section 6.

2 Introduction

2.1 Discrete phase type and matrix geometric distributions

We define DPH [1] and matrix geometric (MG) distributions and their continuous counterparts CPH [2] and matrix exponential (ME) distributions [9] first.

Definition 1. *Let $\mathcal{X}$ be a discrete positive random variable with probability mass function (pmf)*

$$p_i = Pr(\mathcal{X} = i) = \alpha \boldsymbol{A}^{i-1} a, \quad i = 1, 2, \ldots, \tag{1}$$

where α is an initial row vector of size n, $\boldsymbol{A}$ is a square matrix of size $n \times n$, $a = (\mathbb{1} - \boldsymbol{A}\mathbb{1})$, $\mathbb{1}$ is the column vector of ones of size n and $\alpha\mathbb{1} = 1$ (there is no probability mass at $t = 0$). In this case, we say that $\mathcal{X}$ is matrix geometrically distributed with representation $\alpha, \boldsymbol{A}$, or shortly, MG($\alpha, \boldsymbol{A}$) distributed.

Definition 2. *If $\mathcal{X}$ is an MG($\alpha, \boldsymbol{A}$) distributed random variable, where α and $\boldsymbol{A}$ have the following properties:*

- $\alpha_i \geq 0$,
- $A_{ij} \geq 0$, $\boldsymbol{A}\mathbb{1} \leq 1$,
- $\boldsymbol{I} - \boldsymbol{A}$ *is non-singular, where $\boldsymbol{I}$ is the unity matrix,*

then we say that $\mathcal{X}$ is discrete phase type distributed with representation $\alpha, \boldsymbol{A}$, or shortly, DPH($\alpha, \boldsymbol{A}$) distributed.

The vector-matrix representations satisfying the conditions of Definition 2 are called Markovian.

Definition 3. *If $\mathcal{X}$ is an DPH($\alpha, \boldsymbol{A}$) distributed random variable and $\boldsymbol{A}$ is an upper triangular matrix then we say that $\mathcal{X}$ is acyclic discrete phase type distributed with representation $\alpha, \boldsymbol{A}$, or shortly, ADPH($\alpha, \boldsymbol{A}$) distributed.*

The sets of ADPH, DPH, and MG distributions that can be described with size n representations are referred to as order n ADPH, DPH, and MG distributions, respectively. From Definition 1 – 3 it follows that order n ADPH $\subset$ order n DPH $\subset$ order n MG.

2.2 Continuous phase type and matrix exponential distributions

The continuous counterparts of these distributions are the CPH and the matrix exponential distributions.

Definition 4. *Let $\mathcal{X}$ be a continuous positive random variable with cumulative distribution function (cdf)*

$$F_X(x) = Pr(\mathcal{X} < x) = 1 - \alpha e^{\boldsymbol{A}x}\mathbb{1},$$

where α is an initial row vector of size n, $\boldsymbol{A}$ is a square matrix of size $n \times n$, $\mathbb{1}$ is the column vector of ones of size n and $\alpha\mathbb{1} = 1$ (there is no probability mass at $t = 0$). In this case, we say that $\mathcal{X}$ is matrix exponentially distributed with representation $\alpha, \boldsymbol{A}$, or shortly, ME($\alpha, \boldsymbol{A}$) distributed.

Definition 5. *If $\mathcal{X}$ is an ME($\alpha, \boldsymbol{A}$) distributed random variable, where α and $\boldsymbol{A}$ have the following properties: $\alpha_i \geq 0$, $A_{ii} < 0$, $A_{ij} \geq 0$ for $i \neq j$, $\boldsymbol{A}\mathbb{1} \leq 0$, $\boldsymbol{A}$ is non-singular, then we say that $\mathcal{X}$ is continuous phase type distributed with representation $\alpha, \boldsymbol{A}$, or shortly, CPH($\alpha, \boldsymbol{A}$) distributed.*

Definition 6. *If $\mathcal{X}$ is a CPH($\alpha, \boldsymbol{A}$) distributed random variable, where $\boldsymbol{A}$ is an upper triangular matrix then we say that $\mathcal{X}$ is acyclic continuous phase type distributed with representation $\alpha, \boldsymbol{A}$, or shortly, ACPH($\alpha, \boldsymbol{A}$) distributed.*

Definition 7. *Any order n ACPH($\alpha, \boldsymbol{A}$) can be represented with the following vector matrix pair*

$$[\gamma_1, \gamma_2, \ldots, \gamma_n], \begin{bmatrix} -\lambda_1 & \lambda_1 & & \\ & \ddots & \ddots & \\ & & -\lambda_{n-1} & \lambda_{n-1} \\ & & & -\lambda_n \end{bmatrix}$$

where $0 \leq \gamma_i \leq 1$ and λ_i are the eigenvalues of $-\boldsymbol{A}$ such that $\lambda_i \geq \lambda_{i-1}$. This representation is referred to as Cumani's canonical form.

The vector-matrix representations satisfying the conditions of Definition 5 are called Markovian. By these definitions we have the following relations: order n ACPH $\subset$ order n CPH $\subset$ order n ME. Further more for order 2 we have: order 2 ACPH $\equiv$ order 2 CPH $\equiv$ order 2 ME and for order 3 the ACPH set is a valid subset of the CPH set, which is a valid subset of the ME set.

In the sequel we focus on discrete distributions, the continuous ones are introduced for indicating the relations of DPH and CPH distributions.

2.3 Similarity transformation

A given DPH($\alpha, \boldsymbol{A}$) distribution can be represented with more than one vector matrix pair.

Theorem 1. *Let $\boldsymbol{B}$ a square matrix of size n such that $\boldsymbol{B}$ is invertible and $\boldsymbol{B}\mathbb{1} = \mathbb{1}$. Then the vector matrix pair $\gamma = \alpha\boldsymbol{B}, \boldsymbol{G} = \boldsymbol{B}^{-1}\boldsymbol{A}\boldsymbol{B}$ is another representation of DPH($\alpha, \boldsymbol{A}$).*

Proof.

$$\begin{aligned} \bar{p}_i &= Pr(\bar{\mathcal{X}} = i) = \gamma \boldsymbol{G}^{i-1}(\mathbb{1} - \boldsymbol{G}\mathbb{1}) \\ &= \alpha\boldsymbol{B}(\boldsymbol{B}^{-1}\boldsymbol{A}\boldsymbol{B})^{i-1}(\mathbb{1} - \boldsymbol{B}^{-1}\boldsymbol{A}\boldsymbol{B}\mathbb{1}) \\ &= \alpha\boldsymbol{A}^{i-1}(\mathbb{1} - \boldsymbol{A}\mathbb{1}) = p_i. \end{aligned} \tag{2}$$

□

There are important consequences of Theorem 1. The $\boldsymbol{B}^{-1}\boldsymbol{A}\boldsymbol{B}$ transformation of matrix $\boldsymbol{A}$, referred to as similarity transformation, maintains the eigenvalues of matrix $\boldsymbol{A}$ and only modifies the associated eigenvectors. This way the eigenvalues of the matrix of any representation are strongly related with the distribution and can be used to characterize different distribution subclasses.

Further more, an infinite set of vector matrix pairs represent a given ADPH, DPH, or MG distribution and ADPH and DPH distributions can be described with non-Markovian vector matrix pairs.

Definition 8. *A canonical representation is a convenient vector matrix pair chosen from the infinite set of vector matrix pairs defining the same distribution.*

For the convenient canonical representation of DPH distributions we follow the same principles as in [8]. That is the canonical representation is Markovian, takes Cumani's acyclic canonical form [5] if possible and contains the maximal number of zero elements. In some cases these principles completely define the canonical representation, while additional selection criteria are applied in other cases.

3 Canonical form of order 2 DPH distributions

We start with characterizing the properties all possible order 2 MG distributions for which (1) gives non-negative probabilities.

Theorem 2. *An order 2 MG distribution has one of the following two forms*

- *different eigenvalues:*
$$p_i = a_1 s_1^{i-1} + a_2 s_2^{i-1}, \tag{3}$$
where s_1, s_2 are real, $0 < s_1 < 1$, $s_1 > |s_2|$, and a_1, a_2 are such that $0 < a_1 \leq \frac{(1-s_1)(1-s_2)}{s_1 - s_2}$ and $a_2 = (1 - s_2)\left(1 - \frac{a_1}{1-s_1}\right)$;

- *identical eigenvalues:*
$$p_i = (a_1(i-1) + a_2)s^{i-1}, \tag{4}$$
where s is real $0 < s < 1$, and a_1, a_2 are such that $0 < a_1 \leq \frac{(1-s)^2}{s}$ and $a_2 = \frac{(1-s)^2 - a_1 s}{1-s}$.

A vector matrix representation of the first form is

$$\alpha = \left[\frac{a_1}{1-s_1}, \frac{a_2}{1-s_2}\right], \boldsymbol{A} = \begin{bmatrix} s_1 & 0 \\ 0 & s_2 \end{bmatrix}, a = \begin{bmatrix} 1 - s_1 \\ 1 - s_2 \end{bmatrix}, \tag{5}$$

and the second form is

$$\alpha = \left[\tfrac{a_1}{1-s}, \tfrac{a_2(1-s)-a_1(1-2s)}{(1-s)^2}\right],$$
$$\boldsymbol{A} = \begin{bmatrix} s & s \\ 0 & s \end{bmatrix}, \; a = \begin{bmatrix} 1-2s \\ 1-s \end{bmatrix}. \tag{6}$$

Proof. The first form covers the cases when the eigenvalues of $\boldsymbol{A}$ (s_1, s_2) are different and the second one when the eigenvalues ($s_1 = s_2 = s$) are identical. We discuss the cases separately.

- different eigenvalues:

  First we shows that the eigenvalues are real. Assume, that $\boldsymbol{A}$ has a complex eigenvalue. In this case the other eigenvalue has to be its complex conjugate and a_1 and a_2 must be conjugates, too, to obtain real $p_i = a_1 s_1^{i-1} + a_2 s_2^{i-1}$ values. Let φ be the argument of a_1 ($a_1 = |a_1| e^{i\varphi}$), and ψ the argument of s_1. Moreover assume that $\psi \in (0, \pi)$. From $i = 1$ we get that $\varphi \in \left[-\frac{\pi}{2}, \frac{\pi}{2}\right]$. Now consider the case $i = \lceil \frac{\pi}{\psi} \rceil + 1$. The argument of $a_1 s_1^{i-1}$ is $\varphi + (i-1)\psi$, and it is in $\left[\frac{\pi}{2}, \frac{3\pi}{2}\right]$. This means that p_i is negative since $a_1 s_1^{i-1}$ and $a_2 s_2^{i-1}$ are conjugates. So we get, that eigenvalues are real.

  The two real eigenvalues have to be such that the one with the larger absolute value (s_1) is positive, because it becomes dominant for large i values and p_i would become negative for large i values with negative dominant eigenvalue. Additionally the dominant eigenvalue has to be less than one to ensure that the p_i series have finite sum.

  The relation of the a_1, a_2 coefficients is obtained from $\sum_i p_i = 1$. The $a_1 > 0$ bound of a_1 comes from the fact that $p_i \sim a_1 s_1^{i-1}$ for large i, where s_1 is positive. A negative a_1 would result in negative p_i for large i. The upper bound of a_1 comes from $p_1 \geq 0$, since

$$\begin{aligned} 0 &\leq p_1 = a_1 + a_2 \\ 0 &\leq a_1 + (1-s_2)\left(1 - \frac{a_1}{1-s_1}\right) \\ 0 &\leq a_1 \frac{s_2 - s_1}{1-s_1} + (1-s_2) \\ a_1 &\leq \frac{(1-s_1)(1-s_2)}{s_1 - s_2} \end{aligned}$$

- identical eigenvalues:

  First we shows that the eigenvalue is real and non-negative. If s is complex or negative in (4) then $p_i \sim a_1 (i-1) s^{i-1}$ for large i, which becomes complex or negative, respectively, for any a_1 in case of two consecutive large i values.

  $s < 1$ comes from the fact that the p_i series have finite sum.

  Similar to the previous case, the relation of the a_1, a_2 coefficients is obtained from $\sum_i p_i = 1$ and $a_1 > 0$ bound of a_1 comes from the fact that $p_i \sim a_1(i-1)s^{i-1}$ for large i, where s is positive. A negative a_1 would result in negative p_i for large i. The upper bound of a_1 comes from $p_1 \geq 0$, since

$$\begin{aligned} 0 &\leq p_1 = a_2 \\ 0 &\leq \frac{(1-s)^2 - a_1 s}{1-s} \\ a_1 &\leq \frac{(1-s)^2}{s} \end{aligned}$$

□

Theorem 3. *If $\mathcal{X}$ is $MG(2)$ distributed with two distinct positive eigenvalues ($0 < s_2 < s_1 < 1$) then it can be represented as ADPH($\alpha, \boldsymbol{A}$), where*

$$\alpha = \left[\frac{a_1(s_1 - s_2)}{(s_1-1)(s_2-1)}, \frac{a_1 + a_2}{1 - s_2}\right], \qquad \boldsymbol{A} = \begin{bmatrix} s_1 & 1-s_1 \\ 0 & s_2 \end{bmatrix}.$$

Proof. $(\alpha, \boldsymbol{A})$ are such that $p_i = \alpha \boldsymbol{A}^{i-1} a = a_1 s_1^{i-1} + a_2 s_2^{i-1}$. Matrix $\boldsymbol{A}$ and vector $\mathbb{1} - \boldsymbol{A}\mathbb{1}$ obviously satisfies the conditions of Definition 3 when $0 < s_2 < s_1 < 1$. It remained to show the nonnegativity of α when $0 < s_2 < s_1 < 1$, $0 < a_1$ and $p_1 \geq 0$. In the first element of α we have $a_1 > 0$, $s_1 - s_2 > 0$, $s_1 - 1 < 0$, $s_2 - 1 < 0$, from which it is positive. In the second element we have $a1 + a2 = p_1 \geq 0$ and $1 - s_2 > 0$. Note that $\alpha \mathbb{1} = 1$ when $a_2 = (1 - s_2)\left(1 - \frac{a_1}{1-s_1}\right)$. □

Theorem 4. *If $\mathcal{X}$ is $MG(2)$ distributed with a dominant positive and a negative eigenvalues ($s_2 < 0 < s_1 < 1$ and $s_1 + s_2 > 0$) then it can be represented as DPH($\alpha, \boldsymbol{A}$), where*

$$\alpha = \left[\frac{a_1 s_1 + a_2 s_2}{(s_1-1)(s_2-1)}, \frac{(a_1+a_2)(1-s_1-s_2)}{(s_1-1)(s_2-1)}\right],$$
$$\boldsymbol{A} = \begin{bmatrix} 1-\beta_1 & \beta_1 \\ \beta_2 & 0 \end{bmatrix},$$

$\beta_1 = 1 - s_1 - s_2$ *and* $\beta_2 = \frac{s_1 s_2}{s_1+s_2-1}$.

Proof. The eigenvalues of $\boldsymbol{A}$ are s_1, s_2 and $(\alpha, \boldsymbol{A})$ are such that $p_i = \alpha \boldsymbol{A}^{i-1} a = a_1 s_1^{i-1} + a_2 s_2^{i-1}$.

β_1 and β_2 are positive and less than 1. from which matrix $\boldsymbol{A}$ and vector $\mathbb{1} - \boldsymbol{A}\mathbb{1}$ satisfies the conditions of Definition 3.

It remained to show the nonnegativity of α when $s_2 < 0 < s_1 < 1$, $1 > s_1 > s_1 + s_2 > 0$ and $p_1, p_2 \geq 0$. For the first element of α we have $a_1 s_1 + a_2 s_2 = p_2 \geq 0$ $s_1 - 1 < 0$, $s_2 - 1 < 0$ and for the numerator of the second element we have $a_1 + a_2 = p_1 \geq 0$ and $1 - s_1 - s_2 > 0$. The denominator of the second element is the same as the first one. □

Theorem 5. *If $\mathcal{X}$ is $MG(2)$ distributed with two identical eigenvalues ($0 < s = s_2 = s_1 < 1$) then it can be represented as ADPH($\alpha, \boldsymbol{A}$), where*

$$\alpha = \left[\frac{a_1 s}{(1-s)^2}, \frac{a_2}{1-s}\right], \qquad \boldsymbol{A} = \begin{bmatrix} s & 1-s \\ 0 & s \end{bmatrix}.$$

Proof. $(\alpha, \boldsymbol{A})$ are such that $p_i = \alpha \boldsymbol{A}^{i-1} a = (a_1(i-1) + a_2)s^{i-1}$ and matrix $\boldsymbol{A}$ and vector $\mathbb{1} - \boldsymbol{A}\mathbb{1}$ satisfies the conditions of Definition 3 when $0 < s < 1$.

It remained to show the nonnegativity of α when $0 < s < 1$, $0 < a_1$ and $p_1 \geq 0$. All terms of the elements of α are non-negative since $a2 = p_1 \geq 0$. □

Theorem 3 – 5 have the following consequences.

Corollary 1. *The vector matrix representations in Theorem 3 – 5 can be used for canonical representation of order 2 DPH and MG distributions.*

Corollary 2.

order 2 *DPH* $\equiv$ *order* 2 *MG*

order 2 *ADPH* $\equiv$ *order* 2 *MG with positive eigenvalues*

Corollary 3. *If the eigenvalues of the order 2 MG($\gamma, \boldsymbol{G}$) are positive and its canonical representation is ADPH($\alpha, \boldsymbol{A}$) then ME($\gamma, \boldsymbol{G} - \boldsymbol{I}$) is a matrix exponential distribution, whose canonical ACPH representation (Cumani's canonical form) is ACPH($\alpha, \boldsymbol{A} - \boldsymbol{I}$).*

Proof. The matrix of the canonical representation ADPH($\alpha, \boldsymbol{A}$) has the form $\begin{bmatrix} s_1 & 1-s_1 \\ 0 & s_2 \end{bmatrix}$, where $1 > s_1 \geq s_2 > 0$. Consequently $\boldsymbol{A} - \boldsymbol{I}$ is a matrix of an ACPH distribution in Cumani's canonical form with eigenvalues $0 > s_1 - 1 \geq s_2 - 1 > -1$.

Further more, due to the fact that ME($\gamma, \boldsymbol{G} - \boldsymbol{I}$) and ACPH($\alpha, \boldsymbol{A} - \boldsymbol{I}$) represent the same distribution ME($\gamma, \boldsymbol{G} - \boldsymbol{I}$) is a valid ME distribution. □

3.1 Canonical transformations

The introduced canonical representations can be obtained from a general vector matrix representation with the following similarity transformation.

Corollary 4. *If the eigenvalues of the order 2 MG($\gamma, \boldsymbol{G}$) are $0 < s_2 < s_1 < 1$ then its canonical representation is ADPH($\alpha = \gamma \boldsymbol{B}, \boldsymbol{A} = \boldsymbol{B}^{-1}\boldsymbol{G}\boldsymbol{B}$), where matrix $\boldsymbol{B}$ is composed by column vectors $b_1 = \mathbb{1} - b_2$ and $b_2 = \frac{1}{1-s_2}(\mathbb{1} - \boldsymbol{G}\mathbb{1})$.*

Proof. Matrix $\boldsymbol{B}$ is obtained as the solution of $\boldsymbol{B}\mathbb{1} = b_1 + b_2 = \mathbb{1}$ and $\boldsymbol{G}\boldsymbol{B} = \boldsymbol{B}\begin{bmatrix} s_1 & 1-s_1 \\ 0 & s_2 \end{bmatrix}$, whose column vector form is $\boldsymbol{G}b_1 = s_1 b_1$ and $\boldsymbol{G}b_2 = (1-s_1)b_1 + s_2 b_2$. Consequently, $\boldsymbol{A} = \begin{bmatrix} s_1 & 1-s_1 \\ 0 & s_2 \end{bmatrix}$. □

The subsequent proofs of this section follow the same pattern and are omitted.

Corollary 5. *If the eigenvalues of the order 2 MG($\gamma, \boldsymbol{G}$) are $s_2 < 0 < s_1 < 1$ then its canonical representation is ADPH($\gamma\boldsymbol{B}, \begin{bmatrix} s_1 + s_2 & 1-s_1-s_2 \\ \frac{s_1 s_2}{s_1+s_2-1} & 0 \end{bmatrix}$), where matrix $\boldsymbol{B}$ is composed by column vectors $b_1 = \mathbb{1} - b_2$ and $b_2 = \frac{1-s_1-s_2}{(s_1-1)(s_2-1)}(\mathbb{1} - \boldsymbol{G}\mathbb{1})$.*

Corollary 6. *If the eigenvalues of the order 2 MG($\gamma, \boldsymbol{G}$) are $s = s_1 = s_2 < 1$ then its canonical representation is ADPH($\gamma\boldsymbol{B}, \begin{bmatrix} s & 1-s \\ 0 & s \end{bmatrix}$), where matrix $\boldsymbol{B}$ is composed by column vectors $b_1 = \mathbb{1} - b_2$ and $b_2 = \frac{1}{1-s}(\mathbb{1} - \boldsymbol{G}\mathbb{1})$.*

3.2 Procedure

The presented similarity transformations can be used as transformation methods to compute the canonical representation from a general (Markovian or non-Markovian) vector matrix representation. For example a simple implementation of Corollary 4 is presented in Figure 1.

procedure `CanonicalDPH-PP`($\gamma, \boldsymbol{G}$)
    $[s_1, s_2] = \text{eig}(\boldsymbol{G})$;
    $e = [1; 1]$;
    $b_2 = \frac{1}{1-s_2} * (e - \boldsymbol{G} * e)$;
    $b_1 = e - b_2$;
    return $(\gamma * [b_1, b2], \begin{bmatrix} s_1 & 1-s_1 \\ 0 & s_2 \end{bmatrix})$
end procedure

Figure 1: Canonical order 2 DPH representation based on Corollary 4

4 Canonical form of order 3 DPH distributions

In the previous section we intended to prove that the whole order 2 MG class can be represented with Markovian vector matrix pairs. That is why we started with the characterization of the order 2 MG class. In this section we intend to show that all order 3 DPH can be represented with one of the canonical forms defined below. Due to this difference we follow a different approach here and show only that the transformation with a given similarity matrix results in a Markovian canonical form.

Similar to the order 2 case the canonical representations of order 3 DPH distributions are classified according to the eigenvalue structure of the distribution. We encode the eigenvalues in decreasing absolute value and denote the ones with negative real part by N and the ones with non-negative real part by P. For example PNP means that $1 \geq |s_1| \geq |s_2| \geq |s_3|$ and $\text{Re}(s_1) \geq \text{Re}(s_3) \geq 0 > \text{Re}(s_2)$, where s_i, $i = 1, 2, 3$ denote the eigenvalues. Due to the fact that the eigenvalue with the largest absolute value (dominant) has to be real and positive (to ensure positive probabilities in (2) for large i) we have the following cases PPP, PPN, PNP, PNN. Complex (conjugate) eigenvalues can occur only in case of PPP and PNN.

4.1 Case PPP

Following the pattern of Corollary 3 we define the canonical form in the PPP case based on the canonical representation of order 3 CPH distribution.

Theorem 6. *If the eigenvalues of the order 3 DPH($\gamma, \boldsymbol{G}$) are all non-negative we define the canonical form as follows. The vector matrix pair $(\gamma, \boldsymbol{G}-\boldsymbol{I})$ define an order 3 CPH. Let $(\alpha, \boldsymbol{A})$ be the canonical representation of CPH($\gamma, \boldsymbol{G}-\boldsymbol{I}$) as defined in [8]. The canonical representation of DPH($\gamma, \boldsymbol{G}$) is $(\alpha, \boldsymbol{A}+\boldsymbol{I})$.*

Proof. The complete proof of the theorem requires the introduction of the procedure defined in [8]. Here we only demonstrate the result for the case when the canonical representation of CPH($\gamma, \boldsymbol{G}-\boldsymbol{I}$) is acyclic. When the eigenvalues of $\boldsymbol{G}$ are $1 > s_1 \geq s_2 \geq s_3 > 0$ the eigenvalues of $\boldsymbol{G}-\boldsymbol{I}$ are $0 > s_1-1 \geq s_2-1 \geq s_3-1 > -1$. In this case the matrix of the acyclic canonical form of CPH($\gamma, \boldsymbol{G}-\boldsymbol{I}$) is $\boldsymbol{A} = \begin{bmatrix} s_3-1 & 0 & s^*=0 \\ 1-s_2 & s_2-1 & 0 \\ 0 & 1-s_1 & s_1-1 \end{bmatrix}$ and the associated vector α is non-negative. Finally, $\boldsymbol{A}+\boldsymbol{I} = \begin{bmatrix} s_3 & 0 & s^*=0 \\ 1-s_2 & s_2 & 0 \\ 0 & 1-s_1 & s_1 \end{bmatrix}$ is non-negative and the associated exit probability vector, $\mathbb{1} - \boldsymbol{A}\mathbb{1} = [1-s_3, 0, 0]^T$, is non-negative as well.

In the general case s^* might be positive and $s_i - 1$, $i = 1,2,3$ are not the eigenvalues of $\boldsymbol{A}$, but also in that case it holds that the elements of $\boldsymbol{A}+\boldsymbol{I}$ and $\mathbb{1} - \boldsymbol{A}\mathbb{1}$ are non-negative. □

The rest of the cases require the introduction of new canonical structures.

4.2 Case PPN

Theorem 7. *If the eigenvalues of the order 3 DPH($\gamma, \boldsymbol{G}$) are $1 > |s_1| \geq |s_2| \geq |s_3|$ and $\mathrm{Re}(s_1) \geq \mathrm{Re}(s_2) > 0 > \mathrm{Re}(s_3)$ then its canonical representation is DPH($\gamma\boldsymbol{B}, \boldsymbol{A}$), where*

$$\boldsymbol{A} = \begin{bmatrix} x_1 & 1-x_1 & 0 \\ 0 & x_2 & 1-x_2 \\ 0 & x_3 & 0 \end{bmatrix},$$

$x_1 = s_1$, $x_2 = s_2 + s_3$, $x_3 = \frac{-s_2 s_3}{1-s_2-s_3}$ and matrix $\boldsymbol{B}$ is composed by column vectors $b_1 = \mathbb{1} - b_2 - b_3$, $b_2 = \frac{1}{(1-x_2)(1-x_3)}\boldsymbol{G}(\mathbb{1} - \boldsymbol{G}\mathbb{1})$, $b_3 = \frac{1}{1-x_3}(\mathbb{1} - \boldsymbol{G}\mathbb{1})$.

Proof. The eigenvalues of the canonical matrix are s_1, s_2, s_3. We need to prove that $0 \leq x_i < 1$ and $\gamma b_i \geq 0$ for $i = 1,2,3$. Based on the eigenvalue conditions of the PPN case the validity of x_1 and x_2 readable. For x_3 it is readable that $x_3 > 0$ and for the other limit we have

$$\begin{aligned} \frac{-s_2 s_3}{1-s_2-s_3} &< 1 \\ -s_2 s_3 &< 1 - s_2 - s_3 \\ 0 &< 1 - s_2 - s_3 + s_2 s_3 \\ 0 &< \underbrace{(1-s_2)}_{>0}\underbrace{(1-s_3)}_{>0} \end{aligned}$$

b_2 and b_3 are non-negative vectors, because $(\mathbb{1} - \boldsymbol{G}\mathbb{1})$ and $\boldsymbol{G}(\mathbb{1} - \boldsymbol{G}\mathbb{1})$ are the one and two steps exit probability vector of DPH($\gamma, \boldsymbol{G}$) and $0 \leq x_2, x_3 < 1$.

Finally, from the first column of the matrix equation $\boldsymbol{G}\boldsymbol{B} = \boldsymbol{B}\boldsymbol{A}$ we have another expression for b_1, $x_1 b_1 = \boldsymbol{G} b_1$. That is, $x_1 = s_1$ is the largest eigenvalue of $\boldsymbol{G}$ and b_1 is the associated eigenvector which is positive according to the Perron-Frobenius theorem. □

4.3 Case PNP

Theorem 8. *If the eigenvalues of the order 3 DPH($\gamma, \boldsymbol{G}$) are $1 > |s_1| \geq |s_2| \geq |s_3|$, $\mathrm{Re}(s_1) \geq \mathrm{Re}(s_3) \geq 0 > \mathrm{Re}(s_2)$ and $\gamma b_1 \geq 0$ then its canonical representation is DPH($\gamma\boldsymbol{B}, \boldsymbol{A}$), where*

$$\boldsymbol{A} = \begin{bmatrix} x_1 & 1-x_1 & 0 \\ x_2 & 0 & 1-x_2 \\ 0 & x_3 & 0 \end{bmatrix},$$

$x_1 = -a_2$, $x_2 = \frac{a_0 - a_1 a_2}{a_2(1+a_2)}$, $x_3 = \frac{a_0(1+a_2)}{a_0 - a_2 - a_1 a_2 - a_2^2}$ *the matrix elements are defined based on the coefficients of the characteristic polynomial of* $\boldsymbol{G}$, $a_0 = -s_1 s_2 s_3$, $a_1 = s_1 s_2 + s_1 s_3 + s_2 s_3$, $a_2 = -s_1 - s_2 - s_3$. *and matrix* $\boldsymbol{B}$ *is composed by column vectors* $b_1 = \mathbb{1} - b_2 - b_3$, $b_2 = \frac{1}{(1-x_2)(1-x_3)} \boldsymbol{G}(\mathbb{1} - \boldsymbol{G}\mathbb{1})$, $b_3 = \frac{1}{1-x_3}(\mathbb{1} - \boldsymbol{G}\mathbb{1})$.

Proof. The eigenvalues of the canonical matrix are s_1, s_2, s_3. We need to prove that $0 \leq x_i < 1$ and $\gamma b_i \geq 0$ for $i = 1, 2, 3$.

Let $\lambda_i = -s_i$ for $i = 1, 2, 3$. In this case λ_2 is strictly positive and so λ_1 is also strictly negative. λ_3 is non-positive. So $a_0 = \lambda_1 \lambda_2 \lambda_3 \geq 0$. The positivity of $x_1 = -a_2$ follows from the fact that the sum of the eigenvalues of $\boldsymbol{G}$ is positive.

$$\begin{aligned} 1 + a_2 = \underbrace{1 + \lambda_1}_{>0} + \underbrace{\lambda_2 + \lambda_3}_{\geq 0} &> 0 \\ 1 &> -a_2 \\ 1 &> x_1 \end{aligned}$$

The first inequality follows from $-1 < \lambda_1$ and $|\lambda_3| \leq |\lambda_2|$. The next inequality also follows from $-1 < \lambda_1, \lambda_3$ and $0 < \lambda_2$.

$$1 + a_0 + a_1 + a_2 = (1 + \lambda_1)(1 + \lambda_2)(1 + \lambda_3) > 0$$

In the following we use that $-a_2 < 1$. From that we get $a_0 \geq -a_2 a_0$.

$$a_0 - a_2 - a_1 a_2 - a_2^2 \geq -\underbrace{a_2}_{<0} \underbrace{(1 + a_1 + a_2 + a_0)}_{>0} > 0$$

The above expression is the denominator of x_3. In its nominator a_0 is non-negative and $1 + a_2$ is positive, so x_3 is non-negative too. We need to show that $x_3 < 1$:

$$\begin{aligned} x_3 &< 1 \\ a_0 + a_0 a_2 &< a_0 - a_2 - a_1 a_2 - a_2^2 \\ 0 &< -a_2(1 + a_0 + a_1 + a_2). \end{aligned}$$

We saw that above. At the end of this case consider x_2:

$$\begin{aligned} x_2 &< 1 \\ a_0 - a_1 a_2 &> a_2(1 + a_2) \\ a_0 - a_2 - a_1 a_2 - a_2^2 &> 0. \end{aligned}$$

We use here that the eigenvalues of λ_i are decreasing and only λ_2 is positive:

$$x_2 = \frac{-\overbrace{(\lambda_1 + \lambda_2)}^{\leq 0}\overbrace{(\lambda_1 + \lambda_3)}^{\leq 0}\overbrace{(\lambda_2 + \lambda_3)}^{\geq 0}}{-\underbrace{x_1}_{>0}\underbrace{(1 - x_1)}_{>0}} \geq 0$$

b_2 and b_3 are non-negative vectors, because $(\mathbb{1} - \boldsymbol{G}\mathbb{1})$ and $\boldsymbol{G}(\mathbb{1} - \boldsymbol{G}\mathbb{1})$ are the one and the two steps exit probability vectors of DPH$(\gamma, \boldsymbol{G})$ and $0 \leq x_2, x_3 < 1$. □

Theorem 9. *If the eigenvalues of the order 3 DPH(*$\gamma, \boldsymbol{G}$*) are* $1 > |s_1| \geq |s_2| \geq |s_3|$, $\mathrm{Re}(s_1) \geq \mathrm{Re}(s_3) \geq 0 > \mathrm{Re}(s_2)$ *such that* $p_i = a_1 s_1^{i-1} + a_2 s_2^{i-1} + a_3 s_3^{i-1}$ *and the* $\gamma b_1 < 0$ *condition of Theorem 8 does not hold then* *its canonical representation is DPH($\alpha, \boldsymbol{A}$), where*

$$\alpha = \left[\frac{a_3}{1-s_3}, \frac{a_1 s_1 + a_2 s_2}{(s_1-1)(s_2-1)}, \frac{(a_1+a_2)(1-s_1-s_2)}{(s_1-1)(s_2-1)}\right],$$

$$\boldsymbol{A} = \begin{bmatrix} x_1 & 0 & 0 \\ 0 & x_2 & 1-x_2 \\ 0 & x_3 & 0 \end{bmatrix},$$

$x_1 = s_3$, $x_2 = s_1 + s_2$, $x_3 = \frac{-s_1 s_2}{1-s_1-s_2}$.

Proof. It is easy to see, that matrix $\boldsymbol{A}$ is Markovian, because the eigenvalue conditions ensure that $0 < x_1, x_2, x_3 < 1$. It remains to prove that $\alpha \geq 0$ when the $\gamma b_1 < 0$ condition of Theorem 8 does not hold. This part of the proof is still missing. □

4.4 Case PNN

Theorem 10. *If the eigenvalues of the order 3 DPH($\gamma, \boldsymbol{G}$) are* $1 > |s_1| \geq |s_2| \geq |s_3|$, $\mathrm{Re}(s_1) > 0 > \mathrm{Re}(s_3) \geq \mathrm{Re}(s_2)$ *and* $|s_2|^2 \leq 2s_1(-\mathrm{Re}(s_2))$ *then its canonical representation is DPH($\gamma\boldsymbol{B}, \boldsymbol{A}$), where*

$$\boldsymbol{A} = \begin{bmatrix} x_1 & 1-x_1 & 0 \\ x_2 & 0 & 1-x_2 \\ x_3 & 0 & 0 \end{bmatrix},$$

$x_1 = -a_2$, $x_2 = \frac{-a_1}{1+a_2}$, $x_3 = \frac{-a_0}{1+a_1+a_2}$, *the matrix elements are defined based on the coefficients of the characteristic polynomial of* $\boldsymbol{G}$, $a_0 = -s_1 s_2 s_3$, $a_1 = s_1 s_2 + s_1 s_3 + s_2 s_3$, $a_2 = -s_1 - s_2 - s_3$. *and matrix* $\boldsymbol{B}$ *is composed by column vectors* $b_1 = \mathbb{1} - b_2 - b_3$, $b_2 = \frac{1}{(1-x_2)(1-x_3)}\boldsymbol{G}(\mathbb{1} - \boldsymbol{G}\mathbb{1})$, $b_3 = \frac{1}{1-x_3}(\mathbb{1} - \boldsymbol{G}\mathbb{1})$.

Proof. The eigenvalues of the canonical matrix are s_1, s_2, s_3. We need to prove that $0 \leq x_i < 1$ and $\gamma b_i \geq 0$ for $i = 1, 2, 3$.

Let $\lambda_i = -s_i$ for $i = 1, 2, 3$. The statements about a_2 in the PNP case are also valid for this case. The trace of matrix $\boldsymbol{G}$ (the sum of its diagonal elements) equals to the sum of its eigenvalues and so the sum of the eigenvalues as well as $-a_2$ are non-negative. Consequently, $0 \leq x_1 < 1$. Now we consider x_2. $(1 + a_2)$ is positive, so we need to show that a_1 is non-positive.

If the eigenvalues are all real, then we can write

$$a_1 = \underbrace{s_1 s_2}_{<0} + \underbrace{s_3}_{<0} \underbrace{(s_1 + s_2)}_{\geq 0},$$

that is the sum of a negative and a non-positive numbers, so the result will be also negative.

If s_2 and s_3 are complex conjugates, we can write them as $s_2 = -u + \mathrm{i}v$ and $s_3 = -u - \mathrm{i}v$ where u, v are positive reals. With these notations:

$$a_1 = s_1(-u + \mathrm{i}v) + s_1(-u - \mathrm{i}v) + (u^2 + v^2) = u^2 + v^2 - 2s_1 u \leq 0$$

where the last inequality comes from $|s_2|^2 \leq 2s_1(-\mathrm{Re}(s_2))$.

Now we show that x_2 is less than 1:

$$x_2 < 1$$
$$-a_1 < 1 + a_2$$
$$0 < 1 + a_1 + a_2$$

We can see the last inequality if we write $1 + a_1 + a_2$ in the following way:

$$1 + a_1 + a_2 = \underbrace{(1 + \lambda_1)(1 + \lambda_2)(1 + \lambda_3)}_{>0} - \underbrace{\lambda_1 \lambda_2 \lambda_3}_{<0} > 0$$

$\lambda_1 \lambda_2 \lambda_3$ is a_0 so we also get, that x_3 is positive:

$$x_3 = \frac{-\overbrace{a_0}^{<0}}{\underbrace{1+a_1+a_2}_{>0}} > 0$$

And the upper bound of x_3 also follows:

$$\begin{aligned} x_3 &< 1 \\ -a_0 &< 1 + a_1 + a_2 \\ 0 &< 1 + a_0 + a_1 + a_2 \\ 0 &< (1+\lambda_1)(1+\lambda_2)(1+\lambda_3) \end{aligned}$$

b_2 and b_3 are non-negative vectors, because $(\mathbb{1} - \boldsymbol{G}\mathbb{1})$ and $\boldsymbol{G}(\mathbb{1} - \boldsymbol{G}\mathbb{1})$ are the one and two steps exit probability vector of DPH$(\gamma, \boldsymbol{G})$ and $0 \leq x_2, x_3 < 1$.

Finally, from the matrix equation $\boldsymbol{GB} = \boldsymbol{BA}$ we have an explicit expression for b_1, $b_1 = \frac{1}{(1-x_1)(1-x_2)(1-x_3)}\boldsymbol{G}^2(\mathbb{1} - \boldsymbol{G}\mathbb{1})$. That is b_1 is the three steps exit probability vector multiplied with a positive constant. □

Theorem 10 does not cover the case when $|s_2|^2 > 2s_1(-\mathrm{Re}(s_2))$. This case can occur only when s_2 and s_3 are complex conjugate eigenvalues. The following theorem applies in this case.

Theorem 11. *If the eigenvalues of the order 3 DPH$(\gamma, \boldsymbol{G})$ are $1 \geq |s_1| \geq |s_2| \geq |s_3|$, $\mathrm{Re}(s_1) > 0 > \mathrm{Re}(s_3) \geq \mathrm{Re}(s_2)$ and $|s_2|^2 > 2s_1(-\mathrm{Re}(s_2))$ then we use the same canonical form as in case of PPP in Theorem 6.*

Proof. Similar to the proof of Theorem 6 we need to introduce the procedure of [8] in order to prove the theorem, which we avoid here. □

5 Examples

In this section we only demonstrate the special cases of the PNP and the PNN classes.

PNN example The eigenvalues of DPH$(\gamma, \boldsymbol{G})$ with

$$\gamma = \{0.23, 0.46, 0.31\}, \quad \boldsymbol{G} = \begin{bmatrix} 0.47 & 0.34 & 0.03 \\ 0.17 & 0.2 & 0.1 \\ 0.99 & 0 & 0 \end{bmatrix}$$

are $0.715116, -0.0225579 \pm 0.195587i$ for which $|s_2|^2 - 2s_1(-\mathrm{Re}(s_2)) = 0.0065 > 0$. The transformation of DPH$(\gamma, \boldsymbol{G})$ according to Theorem 10 gives

$$\{0.509526, 0.18130549, 0.3091685\}, \\ \begin{bmatrix} 0.67 & 0.33 & 0 \\ -0.01969 & 0 & 1.01969 \\ 0.0823774 & 0 & 0 \end{bmatrix}$$

and the canonical transformation of DPH$(\gamma, \boldsymbol{G})$ according to Theorem 11 gives

$$\{0.31067766, 0.18516351, 0.5041588\}, \\ \begin{bmatrix} 0.0049046 & 0 & 0.0819303 \\ 0.9950954 & 0.0049 & 0 \\ 0 & 0.3398092 & 0.66019 \end{bmatrix}$$

PNP example The eigenvalues of DPH$(\gamma, \boldsymbol{G})$ with

$$\gamma = \{0.01, 0.48, 0.51\}, \quad \boldsymbol{G} = \begin{bmatrix} 0.19 & 0.8 & 0 \\ 0.88 & 0.02 & 0 \\ 0 & 0.12 & 0.5 \end{bmatrix}$$

are $0.948342, -0.738342, 0.5$. This DPH violates the γb_1 condition of Theorem 8 which can be seen from the result of applying the transformation of Theorem 8

$$\{-0.716466, 0.704157, 1.0123\},\\ \begin{bmatrix} 0.71 & 0.29 & 0 \\ 0.352 & 0 & 0.648 \\ 0 & 0.761 & 0 \end{bmatrix}$$

The canonical transformation of DPH($\gamma, \boldsymbol{G}$) according to Theorem 9 gives

$$\{0.378826, 0.159427, 0.461747\},\\ \begin{bmatrix} 0.5 & 0 & 0 \\ 0 & 0.21 & 0.79 \\ 0 & 0.886329 & 0 \end{bmatrix}$$

Acknowledgement

M. Telek was partially supported by the TÁMOP-4.2.2C-11/1/KONV-2012-0001 and the OTKA K101150 projects.

6 Conclusions

We have classified the set of order 2 and 3 DPH distributions based on their eigenvalue structures. This classification contains a class of DPH distributions which behave very similar to the CPH class of the same order. It is the class of DPH distributions whose eigenvalues have positive real part. For the other classes of DPH distributions (i.e., the ones which have at least one eigenvalue with negative real part) we propose canonical forms with different structural properties.

For the full characterization of the order 3 DPH class the proof of Theorem 9 should be completed.

References

[1] M. Neuts, “Probability distributions of phase type,” in *Liber Amicorum Prof. Emeritus H. Florin*. University of Louvain, 1975, pp. 173–206.

[2] ——, *Matrix Geometric Solutions in Stochastic Models*. Baltimore: Johns Hopkins University Press, 1981.

[3] A. Alfa, “Discrete time queues and matrix-analytic methods,” *Top*, vol. 10, pp. 147–185, 2002. [Online]. Available: http://dx.doi.org/10.1007/BF02579008

[4] L. Lakatos, L. Szeidl, and M. Telek, *Introduction to Queueing Systems with Telecommunication Applications*. Springer, 2013. [Online]. Available: http://www.springer.com/mathematics/probability/book/978-1-4614-5316-1

[5] A. Cumani, “On the canonical representation of homogeneous Markov processes modelling failure-time distributions,” *Microelectronics and Reliability*, vol. 22, pp. 583–602, 1982.

[6] A. Bobbio, A. Horváth, M. Scarpa, and M. Telek, “Acyclic discrete phase type distributions: Properties and a parameter estimation algorithm,” *Performance Evaluation*, vol. 54, no. 1, pp. 1–32, 2003.

[7] M. Telek and A. Heindl, “Matching moments for acyclic discrete and continuous phase-type distributions of second order,” *International Journal of Simulation Systems, Science & Technology*, vol. 3, no. 3-4, pp. 47–57, dec. 2002, special Issue on: Analytical & Stochastic Modelling Techniques.

[8] G. Horváth and M. Telek, “On the canonical representation of phase type distributions,” *Performance Evaluation*, vol. 66, no. 8, pp. 396 – 409, 2009.

[9] M. Bladt and M. F. Neuts, “Matrix-exponential distributions: Calculus and interpretations via flows,” *Stochastic Models*, vol. 19, no. 1, pp. 113–124, 2003.