

Does a given vector-matrix pair correspond to a PH distribution?

Philipp Reinecke¹ and Miklós Telek^{2,3,4}

¹ HP Labs, Bristol

² Department of Telecommunications, Technical University of Budapest

³ MTA-BME Information systems research group

⁴ Inter-University Center for Telecommunications and Informatics Debrecen

e-mail: philipp.reinecke@hp.com, telek@hit.bme.hu

July 22, 2014

Abstract

The analysis of practical queueing problems benefits if realistic distributions can be used as parameters. Phase type (PH) distributions can approximate many distributions arising in practice, but their practical applicability has always been limited when they are described by a non-Markovian vector-matrix pair. In this case it is hard to check whether the non-Markovian vector-matrix pair defines a non-negative matrix-exponential function or not. In this paper we propose a numerical procedure for checking if the matrix-exponential function defined by a non-Markovian vector-matrix pair can be represented by a Markovian vector-matrix pair with potentially larger size. If so, then the matrix-exponential function is non-negative.

The proposed procedure is based on O'Connide's characterization result, which says that a non-Markovian vector-matrix pair with strictly positive density on $(0, \infty)$ and with a real dominant eigenvalue has a Markovian representation. Our method checks the existence of a potential Markovian representation in a computationally efficient way utilizing the structural properties of the applied representation transformation procedure.

Keywords: Phase-type distribution, vector-matrix representation, randomization, uniformization.

1 Introduction

Phase-type (PH) distributions are defined as the time to absorption in a discrete-state Markov chain with all but one transient states and one absorbing state [14]. In this paper we focus on continuous PH distributions whose underlying stochastic process is a continuous-time Markov chain. This background Markov-chain-based stochastic interpretation of PH distributions made their use quite efficient in applied stochastic modeling. In the Markovian framework the PH distribution is defined by the initial probability distribution of the Markov chain (given as a row vector of initial state probabilities – non-negative numbers whose sum is one) and its generator matrix (whose off-diagonal elements hold non-negative transition rates and diagonal elements are not greater than the sum of all transition rates out of the associated state times minus one). The vector-matrix pair satisfying these structural restrictions is referred to as Markovian and allows the application of the background Markov-chain-based stochastic interpretation of PH distributions. Unfortunately, the vector-matrix based definition of PH distributions is not unique. Infinitely many vector-matrix pairs can define a given PH distribution and infinitely many of them are non-Markovian [10].

The use of a non-Markovian vector-matrix pair for describing a PH distribution is rather limited in applied stochastic modeling mainly due to the following two reasons:

1. it does not allow the application of the background Markov-chain-based stochastic interpretation of PH distributions,
2. it is rather complex to check if the associated matrix-exponential function is non-negative.

Recent results indicate that the lack of a stochastic interpretation is not that crucial. Following the general results in [10] it has been shown [1, 4] that the numerical analysis of a stochastic model with non-Markovian vector-matrix representation of PH distributions can be performed with the same procedures as the ones used with Markovian representations.

The focus of this paper is on the second important disadvantage listed above. We present a numerical procedure to decide if a matrix-exponential function defined by a non-Markovian vector-matrix pair defines a PH distribution or not. The proposed procedure utilizes results about the properties of PH distributions, Markov chains and randomization. In particular, we make use of the fact that any matrix-exponential function with positive density on $(0, \infty)$ and with a unique dominant eigenvalue (whose multiplicity can be greater than one) has a PH representation [16]. Furthermore, our approach is built on the monocyclic representation of PH distributions [15, 6]. The monocyclic representation of PH distribution bridges several Markov-chain-related results with the analysis of matrix-exponential functions.

A motivating application for this method is the problem of moment-matching with PH distributions. There are numerical procedures [19, 18] to generate non-Markovian vector-matrix pairs whose associated moments match a set of predefined values (for example, experimental moments). The applicability of this non-Markovian vector-matrix pair in stochastic models depends on the fact whether it defines a valid PH distribution or not. Therefore, for the use of this moments matching approach it is crucial to check if the obtained non-Markovian vector-matrix pair defines a PH distribution. In this paper we propose an efficient numerical method to decide this question which might eliminate the remaining technical difficulties from the practical application of non-Markovian vector-matrix pair definition of PH distributions.

The rest of the paper is organized as follows: We first summarize the notation and important properties of phase type distributions. We then describe the proposed method. In Section 3 we give an algorithm for computing a positive representation, if one exists. We then discuss a refinement of the algorithm and conclude the paper by giving some numerical examples.

2 Preliminaries

2.1 Basic definitions

Definition 1. A pair $(\boldsymbol{\alpha}, \mathbf{A})$, where $\boldsymbol{\alpha}$ is a vector of size $1 \times n$ and \mathbf{A} is a matrix of size $n \times n$ is said to be a vector-matrix pair of size n .

Definition 2. The vector-matrix pair of size n , $(\boldsymbol{\alpha}, \mathbf{A})$, is said to be Markovian if $\boldsymbol{\alpha}$ and \mathbf{A} have the following properties:

- $\alpha_i \geq 0$,
- $A_{ii} < 0$, $A_{ij} \geq 0$ for $i \neq j$, $\mathbf{A}\mathbf{1} \leq 0$,
- \mathbf{A} is non-singular,

PH distributions can be defined as follows.

Definition 3. Let X be a random variable with cumulative distribution function (CDF) $F_X(x) = Pr(X \leq x)$. X is PH distributed if there is a finite size Markovian vector-matrix pair, $(\boldsymbol{\alpha}, \mathbf{A})$, for which

$$F_X(x) = Pr(X \leq x) = 1 - \boldsymbol{\alpha}e^{\mathbf{A}x}\mathbf{1}, \quad (1)$$

In this case we say that X is phase-type distributed with representation $(\boldsymbol{\alpha}, \mathbf{A})$, $PH(\boldsymbol{\alpha}, \mathbf{A})$ distributed, for short.

In Definition 3, vector $\boldsymbol{\alpha}$ is often referred to as initial row vector and matrix \mathbf{A} as transient generator matrix. In this paper we assume that $\boldsymbol{\alpha}\mathbf{1} = 1$, which means that there is no probability mass at $t = 0$ and $F_X(0) = 0$. The probability density function (PDF), the Laplace transform and the moments of X are

$$f_X(x) = \boldsymbol{\alpha}e^{\mathbf{A}x}(-\mathbf{A})\mathbf{1}, \quad (2)$$

$$f_X^*(s) = E(e^{-sX}) = \boldsymbol{\alpha}(s\mathbf{I} - \mathbf{A})^{-1}(-\mathbf{A})\mathbf{1}, \quad (3)$$

$$\mu_n = E(X^n) = n!\boldsymbol{\alpha}(-\mathbf{A})^{-n}\mathbf{1}. \quad (4)$$

(2) defines the matrix-exponential function associated with the vector-matrix pair $(\boldsymbol{\alpha}, \mathbf{A})$, both for Markovian and non-Markovian vector-matrix pairs.

2.2 Different Representations of PH Distributions

The vector-matrix representation of $f_X(x)$ is not unique. There are different vector-matrix pairs, both with identical size and different sizes, resulting in the same matrix-exponential function, as it is summarized in the following theorems. The proofs are available in the original papers.

Theorem 1. [18] Let $(\boldsymbol{\alpha}, \mathbf{A})$ and $(\boldsymbol{\gamma}, \mathbf{G})$ be two vector-matrix pairs of size n with associated matrix-exponential function $f_X(x)$ and $f_Y(x)$, respectively. Then $f_X(x)$ and $f_Y(x)$ are identical iff there exists a non-singular matrix \mathbf{B} of size $n \times n$, such that $\boldsymbol{\gamma} = \boldsymbol{\alpha}\mathbf{B}$, $\mathbf{G} = \mathbf{B}^{-1}\mathbf{A}\mathbf{B}$ and $\mathbf{B}\mathbf{1} = \mathbf{1}$.

Theorem 2. [5] Let $(\boldsymbol{\alpha}, \mathbf{A})$ be a vector-matrix pair of size n and $(\boldsymbol{\gamma}, \mathbf{G})$ be a vector-matrix pair of size m ($m > n$) with associated matrix-exponential functions $f_X(x)$ and $f_Y(x)$, respectively. If there exists a matrix \mathbf{W} of cardinality $n \times m$, such that $\boldsymbol{\alpha}\mathbf{W} = \boldsymbol{\gamma}$, $\mathbf{A}\mathbf{W} = \mathbf{W}\mathbf{G}$, $\mathbf{W}\mathbf{1}_m = \mathbf{1}_n$ then $f_X(x) \equiv f_Y(x)$.

Definition 4. X is PH distributed with density function $f_X(x)$. The $(\boldsymbol{\alpha}, \mathbf{A})$ vector-matrix pair of size n is said to be a minimal representation of X if there is no vector-matrix pair of smaller size whose associated matrix-exponential function is $f_X(x)$.

If the $(\boldsymbol{\alpha}, \mathbf{A})$ representation of $f_X(x)$ is not minimal then, for example, the STAIRCASE method from [5] (implemented in [3]) can be applied to obtain a minimal representation. The minimal representation of PH distributions has some easy-to-check necessary conditions [16]:

- C1 The eigenvalues of \mathbf{A} have negative real part (to avoid divergence of $f_X(x)$).
- C2 There is a real eigenvalue of \mathbf{A} with maximal real part (to avoid asymptotic oscillations of $f_X(x)$).
- C3 $\boldsymbol{\alpha}\mathbf{1} = 1$ (normalizing condition without probability mass at zero).
- C4 $f_X(0)$ is either positive or zero. If it is zero then the first non-zero derivative of $f_X(x)$ at $x = 0$ is positive (to avoid negative $f_X(x)$ values close to $x = 0$).

If these necessary conditions are violated then the vector-matrix pair $(\boldsymbol{\alpha}, \mathbf{A})$ does not define a valid PH distribution. Further necessary conditions are available in [17, 8] and the references therein. Apart from these necessary conditions there are some necessary and sufficient conditions for PH distributions whose minimal vector-matrix representation is of size 3 [9, 2], but there are no easy-to-check necessary and sufficient conditions for the general case.

The representation-transformation methods in Theorem 1 and 2 have important properties with respect to the eigenvalues of the different representations. If $(\boldsymbol{\alpha}, \mathbf{A})$ and $(\boldsymbol{\gamma}, \mathbf{G})$ are both minimal and the conditions of Theorem 1 hold, then the eigenvalues of \mathbf{A} and \mathbf{G} are identical with the same multiplicities.

If $(\boldsymbol{\alpha}, \mathbf{A})$ of size n is minimal, the size of $(\boldsymbol{\gamma}, \mathbf{G})$ is m ($m > n$) and the conditions of Theorem 2 hold, then every eigenvalue of \mathbf{A} is an eigenvalue of \mathbf{G} at most with the same multiplicity.

2.3 The main elements of the proposed method

It is an important property of Definition 3 that the vector-matrix representation of a PH distribution is not unique, that is, that different vector-matrix pairs can define the same distribution. In particular, it might happen that a *non-Markovian* vector-matrix pair and another, *Markovian*, vector-matrix pair define the same PH distribution.

Using these basic definitions it is easy to present the main idea of the proposed numerical procedure: given a non-Markovian vector-matrix pair the procedure looks for a Markovian vector-matrix pair (of potentially much larger, but finite, size) which defines the same matrix-exponential function. If there exists such a Markovian vector-matrix pair then the non-Markovian vector-matrix pair defines a valid PH distribution.

The theoretical basis of our procedure is the seminal result of O’Cinneide [16], which states that any matrix-exponential function with a unique dominant eigenvalue (of potentially higher multiplicity) and a strictly positive density in $(0, \infty)$ has a Markovian vector-matrix representation. Based on this result Mocanu and Commault [11] recommended a particular unique representation for PH distributions. Additionally, Mocanu developed a tool [12], called MoMI tool, which implements the generation of such unique PH representations based on the poles and the zeros of the rational Laplace transform. Indeed the MoMI tool can also be used for deciding the validity of a non-Markovian vector-matrix pair, but our proposed procedure

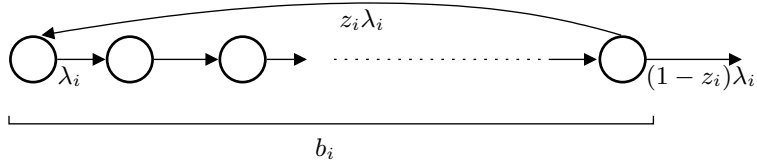


Figure 1: FE-diagonal block.

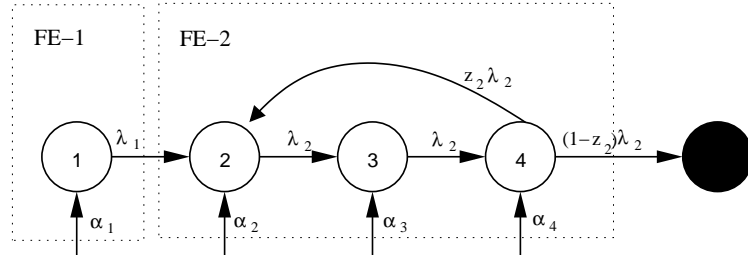


Figure 2: FE-diagonal representation of a generator with a real eigenvalue (λ_1) and a pair of complex ones.

makes an efficient use of additional properties (for example, Theorem 4) which are not considered in the MoMI tool.

Employing the approach from [11] the following section presents a way for constructing a representation with a Markovian matrix (and a potentially non-Markovian initial vector) for a non-Markovian vector-matrix pair. The core of the approach is that every eigenvalue is represented separately with a small Markovian matrix block.

2.4 Representation of PH distributions with Markovian matrix

The Laplace transform of a PH distributions may have complex poles which are identical with the non-vanishing eigenvalues of the matrix of the associated vector-matrix pairs. [11] proposed the use of Feedback-Erlang (FE) blocks to represent pairs of complex eigenvalues with proper Markovian generator blocks:

Definition 5. [11] A Feedback-Erlang (FE) block with parameters (b, λ, z) is a chain of b states with transition rate λ and one transition from the b th state to the first state, with rate $z\lambda$ (c.f. Figure 1). The probability $z \in [0, 1)$ is called the feedback probability.

Feedback-Erlang blocks with length $b = 1$ or feedback probability $z = 0$ are called *degenerate* FE blocks. Note that an FE block (b, λ, z) with length $b = 1$ corresponds to an exponential distribution with rate λ , while $z = 0$ gives the Erlang- b distribution with rate λ (the sum of b independent exponentially distributed random variables with parameter λ). In both cases, the eigenvalue of the matrix block is $-\lambda$.

A non-degenerate FE block where b is odd has a real eigenvalue and $(b-1)/2$ complex conjugate eigenvalue pairs. A non-degenerate FE block where b is even has 2 real eigenvalues and $(b-2)/2$ complex conjugate eigenvalue pairs. In both cases the eigenvalues are located on a circle in the complex plane. The dominant eigenvalue (the one with the largest real part) of the FE block with parameters (b, λ, z) is always real and given by $r = -\lambda \left(1 - z^{\frac{1}{b}}\right)$ [11]. Given the eigenvalues $\sigma_1, \dots, \sigma_n$ of a non-Markovian matrix whose dominant eigenvalue (which is negative and real according to condition C1 and C2) is σ_1 , it is possible to compose FE blocks for representing these eigenvalues as follows

- if σ_j is real, the corresponding FE block is a degenerate block; thus the parameters are:

$$\lambda_j = -\sigma_j, \quad b_j = 1, \quad z_j = 0, \quad (5)$$

- if $\sigma_j = -a_j \pm ic_j$ is a complex conjugate pair, the parameters are:

$$b_i = \left\lceil \frac{2\pi}{\pi - 2 \arctan\left(\frac{c_i}{a_i + \sigma_1}\right)} \right\rceil, \quad (6)$$

$$\lambda_i = \frac{1}{2} \left(2a_i - c_i \tan \frac{\pi}{b_i} + c_i \cot \frac{\pi}{b_i} \right), \quad (7)$$

$$z_i = \left(1 - \frac{a_i - c_i \tan \frac{\pi}{b_i}}{\lambda_i} \right)^{b_i}, \quad (8)$$

where $\lceil x \rceil$ denotes the smallest integer strictly greater than x .

This construction of the FE blocks ensures that σ_1 remains the dominant eigenvalue, that is, the real parts of all eigenvalues of the FE blocks are not greater than σ_1 . Based on these FE blocks it is possible to construct a proper Markovian generator which possesses all eigenvalues of the non-Markovian representation and some additional eigenvalues which are the additional eigenvalues of the FE blocks:

Definition 6. A monocyclic representation of a non-Markovian matrix with a unique negative real dominant eigenvalue is a Markovian matrix which consists of h (the number of complex conjugate eigenvalue pairs plus the number of real eigenvalues) Feedback-Erlang blocks $(b_i, \lambda_i, z_i), i = 1, \dots, h$, such that it exhibits all eigenvalues of the non-Markovian matrix with at least the same multiplicity. The only non-zero entries of the Markovian matrix are in the FE blocks along the diagonal and the transition rates from the last state of a FE block to the first state of the next one $((1 - z_i)\lambda_i)$. The size of the monocyclic representation is $m = \sum_{i=1}^h b_i$ and the dominant eigenvalues of the FE blocks are ordered by increasing absolute value, $|r_i| \leq |r_j|$, for $1 \leq i \leq j \leq h$.

Figure 2 depicts an example of a Markovian generator which is the monocyclic representation of a generator with a real eigenvalue (λ_1) and a pair of complex conjugate ones in FE-diagonal form. In this representation there are two FE blocks, one of length $b_1 = 1$ with rate $q_1 = \lambda_1$, and one of length $b_2 = 3$ with rate λ_2 and feedback probability z_2 . The associated generator matrix is

$$\mathbf{G} = \left(\begin{array}{c|ccc} -\lambda_1 & \lambda_1 & 0 & 0 \\ \hline 0 & -\lambda_2 & \lambda_2 & 0 \\ 0 & 0 & -\lambda_2 & \lambda_2 \\ 0 & z\lambda_2 & 0 & -\lambda_2 \end{array} \right).$$

Let (α, \mathbf{A}) be a non-Markovian vector-matrix pair with a dominant real eigenvalue. Then we can compute the monocyclic representation of the matrix \mathbf{A} , denoted as matrix \mathbf{G} , by the eigenvalues of \mathbf{A} and the associated FE blocks. If we look for an equivalent representation of (α, \mathbf{A}) with monocyclic matrix \mathbf{G} , then we need to compute a vector γ , for which the matrix-exponential function associated with (α, \mathbf{A}) is identical with the matrix-exponential function associated with (γ, \mathbf{G}) . There are different methods for this computation, for example in [11], but here we present an approach based on the similarity transformations of Theorems 1 - 2 [5]. Let n and m ($n \leq m$) be the size of \mathbf{A} and \mathbf{G} , respectively. Compute matrix $\widehat{\mathbf{W}}$ of size $n \times m$ as the unique solution to

$$\mathbf{A}\widehat{\mathbf{W}} = \widehat{\mathbf{W}}\mathbf{G}, \quad \mathbf{W}\mathbb{1} = \mathbb{1}, \quad (9)$$

and based on $\widehat{\mathbf{W}}$ the initial vector is

$$\underline{\gamma} = \underline{\alpha} \cdot \widehat{\mathbf{W}}. \quad (10)$$

Since \mathbf{G} is Markovian, the obtained (γ, \mathbf{G}) representation is Markovian if γ is non-negative, but this is not necessarily the case. The case when γ has negative elements is considered in the following subsection.

2.5 Markovian representation of PH distributions

If the $(\boldsymbol{\gamma}, \mathbf{G})$ representation of the non-Markovian vector-matrix pair $(\boldsymbol{\alpha}, \mathbf{A})$ with strictly positive density on $(0, \infty)$ and with monocyclic matrix \mathbf{G} is not Markovian (that is, $\boldsymbol{\gamma}$ has at least one negative element), then the $(\boldsymbol{\gamma}, \mathbf{G})$ representation can be extended with an appropriate number of degenerate FE blocks with appropriate parameter λ in the following way

$$\mathbf{B} = \begin{pmatrix} \mathbf{G} & -\mathbf{G}\mathbb{1} & & & \\ & -\lambda & \lambda & & \\ & & \ddots & \ddots & \\ & & & \ddots & -\lambda \end{pmatrix},$$

such that the representation $(\boldsymbol{\beta}, \mathbf{B})$, where $\boldsymbol{\beta} = \boldsymbol{\gamma}\mathbf{W}$ and \mathbf{W} is the unique solution of $\mathbf{G}\mathbf{W} = \mathbf{W}\mathbf{B}$, $\mathbf{W}\mathbb{1} = \mathbb{1}$, is Markovian. The column vector $-\mathbf{G}\mathbb{1}$ is such that its only non-zero element is the last element which is the exit rate from the last FE block, $(1 - z_h)\lambda_h$. This result is from [11, 13, 12], which ensures the wide applicability of the monocyclic representation according to the following theorem:

Theorem 3. *Every non-Markovian vector-matrix pair $(\boldsymbol{\alpha}, \mathbf{A})$ with a dominant real eigenvalue and with a strictly positive associated matrix-exponential function on $(0, \infty)$ has a monocyclic representation in the form of matrix \mathbf{B} with a Markovian initial vector $\boldsymbol{\beta}$.*

We note that a Markovian monocyclic representation of a non-Markovian vector-matrix pair $(\boldsymbol{\alpha}, \mathbf{A})$ can also be computed using the MoMI tool of S. Mocanu. Unfortunately, we have experienced various computational and operational problems with that tool, and its source code is not publicly available. The individual steps of the representation-transformation procedure proposed above are available as part of the Butools library [3].

2.6 Available steps of the procedure

We want to decide if a matrix-exponential function defined by a non-Markovian vector-matrix pair $(\boldsymbol{\alpha}, \mathbf{A})$ defines a proper PH distribution or not. To this end we propose to perform the following steps.

S1 Check if $(\boldsymbol{\alpha}, \mathbf{A})$ satisfies conditions C1 - C4.

S2 Transform $(\boldsymbol{\alpha}, \mathbf{A})$ to $(\boldsymbol{\gamma}, \mathbf{G})$ according to Section 2.4 such that each eigenvalue and complex eigenvalue pair of \mathbf{A} is represented by an FE block in \mathbf{G} . This transformation ensures that matrix \mathbf{G} is Markovian. If $\boldsymbol{\gamma}$, the initial vector associated with matrix \mathbf{G} , is non-negative then $(\boldsymbol{\alpha}, \mathbf{A})$ defines a PH distribution with strictly positive density. Otherwise step S3 needs to be applied.

S3 Transform $(\boldsymbol{\gamma}, \mathbf{G})$ to $(\boldsymbol{\beta}, \mathbf{B})$, with non-negative $\boldsymbol{\beta}$ according to Section 2.5 if possible.

Unfortunately, Theorem 3 is not supported with an efficient procedure to find the number of additional degenerate FE blocks and parameter λ for composing matrix \mathbf{B} . In the rest of the paper we essentially focus on step S3 and present numerical methods for finding the number of additional blocks and λ if possible.

3 Positive realization

Let $(\boldsymbol{\gamma}, \mathbf{G})$ be a vector-matrix pair where $\boldsymbol{\gamma}$ is a row vector with at least one negative element and \mathbf{G} is a monocyclic generator composed of FE blocks (that is, matrix \mathbf{G} is Markovian). The size of vector $\boldsymbol{\gamma}$ and matrix \mathbf{G} is u and the matrix-exponential function associated with $(\boldsymbol{\gamma}, \mathbf{G})$ is $f(t) = -\boldsymbol{\gamma}e^{\mathbf{G}t}\mathbf{G}\mathbb{1}$. We transform the $(\boldsymbol{\gamma}, \mathbf{G})$ representation to another equivalent representation of larger size, by extending the $(\boldsymbol{\gamma}, \mathbf{G})$ representation with additional n phases in the following way:

$$\mathbf{B}(n, \lambda) = \begin{pmatrix} \mathbf{G} & -\mathbf{G}\mathbb{1} & & & \\ & -\lambda & \lambda & & \\ & & \ddots & \ddots & \\ & & & \ddots & -\lambda \end{pmatrix},$$

where $\mathbf{B}(n, \lambda)$ is of size $u + n$ (the size of the upper left block of $\mathbf{B}(n, \lambda)$ is u , the remaining n blocks are of size one). $-\mathbf{G}\mathbb{1}$ is a non-negative column vector of size u due to the properties of \mathbf{G} . Let vector $\beta(n, \lambda)$ of size $u + n$ be $\beta(n, \lambda) = \gamma\mathbf{W}$, where matrix \mathbf{W} (of size $u \times (u + n)$) is such that $\mathbf{G}\mathbf{W} = \mathbf{W}\mathbf{B}(n, \lambda)$, $\mathbf{W}\mathbb{1}_{u+n} = \mathbb{1}_n$. $(\beta(n, \lambda), \mathbf{B}(n, \lambda))$ is a different vector-matrix representation of the same matrix-exponential function, because

$$\begin{aligned} -\beta(n, \lambda)e^{\mathbf{B}(n, \lambda)t}\mathbf{B}(n, \lambda)\mathbb{1} &= -\gamma\mathbf{W}e^{\mathbf{B}(n, \lambda)t}\mathbf{B}(n, \lambda)\mathbb{1} \\ &= -\gamma e^{\mathbf{G}t}\mathbf{W}\mathbf{B}(n, \lambda)\mathbb{1} \\ &= -\gamma e^{\mathbf{G}t}\mathbf{G}\mathbf{W}\mathbb{1} \\ &= -\gamma e^{\mathbf{G}t}\mathbf{G}\mathbb{1} \\ &= f(t). \end{aligned}$$

Our main goal is to find $n \in \mathbb{N}^+$ and $\lambda \in \mathbb{R}^+$ such that $\beta(n, \lambda)$ is non-negative. The transformation between different representations of a matrix-exponential function is rather complex in general. Fortunately, due to the special structure of matrix $\mathbf{B}(n, \lambda)$ the transformation between representations (γ, \mathbf{G}) and $(\beta(n, \lambda), \mathbf{B}(n, \lambda))$ is rather regular, as summarized in the following theorem.

Theorem 4. \mathbf{W} has the form

$$\mathbf{W} = ((\mathbf{I} + \mathbf{G}/\lambda)^n \mid (\mathbf{I} + \mathbf{G}/\lambda)^{n-1}g(\lambda) \mid (\mathbf{I} + \mathbf{G}/\lambda)^{n-2}g(\lambda) \mid \dots \mid g(\lambda)),$$

where the size of the first block is $u \times u$, the size of the remaining blocks is $u \times 1$ and $g(\lambda) = -\mathbf{G}\mathbb{1}/\lambda$ is a column vector of size u .

Proof. First we show that $\mathbf{G}\mathbf{W} = \mathbf{W}\mathbf{B}(n, \lambda)$ and then we show that $\mathbf{W}\mathbb{1} = \mathbb{1}$. On the one hand, using that \mathbf{G} and $(\mathbf{I} + \mathbf{G}/\lambda)$ commute we have

$$\mathbf{G}\mathbf{W} = ((\mathbf{I} + \mathbf{G}/\lambda)^n\mathbf{G} \mid (\mathbf{I} + \mathbf{G}/\lambda)^{n-1}\mathbf{G}g(\lambda) \mid (\mathbf{I} + \mathbf{G}/\lambda)^{n-2}\mathbf{G}g(\lambda) \mid \dots \mid \mathbf{G}g(\lambda)),$$

on the other hand,

$$\begin{aligned} \mathbf{W}\mathbf{B}(n, \lambda) &= \\ &= ((\mathbf{I} + \mathbf{G}/\lambda)^n \mid (\mathbf{I} + \mathbf{G}/\lambda)^{n-1}g(\lambda) \mid (\mathbf{I} + \mathbf{G}/\lambda)^{n-2}g(\lambda) \mid \dots \mid g(\lambda)) \begin{pmatrix} \mathbf{G} & -\mathbf{G}\mathbb{1} & & & \\ & -\lambda & \lambda & & \\ & & & \ddots & \\ & & & & \ddots & \\ & & & & & -\lambda \end{pmatrix} \\ &= ((\mathbf{I} + \mathbf{G}/\lambda)^n\mathbf{G} \mid (\mathbf{I} + \mathbf{G}/\lambda)^{n-1}\mathbf{G}g(\lambda) \mid (\mathbf{I} + \mathbf{G}/\lambda)^{n-2}\mathbf{G}g(\lambda) \mid \dots \mid \mathbf{G}g(\lambda)). \end{aligned}$$

Furthermore

$$\begin{aligned} \mathbf{W}\mathbb{1} &= ((\mathbf{I} + \mathbf{G}/\lambda)^n \mid (\mathbf{I} + \mathbf{G}/\lambda)^{n-1}g(\lambda) \mid (\mathbf{I} + \mathbf{G}/\lambda)^{n-2}g(\lambda) \mid \dots \mid g(\lambda)) \mathbb{1} \\ &= (\mathbf{I} + \mathbf{G}/\lambda)^n\mathbb{1} + \sum_{i=0}^{n-1} (\mathbf{I} + \mathbf{G}/\lambda)^{n-1}g(\lambda) \\ &= (\mathbf{I} + \mathbf{G}/\lambda)^n\mathbb{1} + (\mathbf{I} - (\mathbf{I} + \mathbf{G}/\lambda)^n)(\mathbf{I} - (\mathbf{I} + \mathbf{G}/\lambda))^{-1}g(\lambda) \\ &= (\mathbf{I} + \mathbf{G}/\lambda)^n\mathbb{1} + (\mathbf{I} - (\mathbf{I} + \mathbf{G}/\lambda)^n)(-\mathbf{G}/\lambda)^{-1}(-\mathbf{G}/\lambda)\mathbb{1} \\ &= (\mathbf{I} + \mathbf{G}/\lambda)^n\mathbb{1} + (\mathbf{I} - (\mathbf{I} + \mathbf{G}/\lambda)^n)\mathbb{1} \\ &= \mathbb{1}. \end{aligned}$$

□

This regular structure of the transformation matrix allows us to design efficient numerical computation methods using the following stochastic interpretation.

3.1 Stochastic interpretation

\mathbf{G} is a Markovian monocyclic generator. Let λ be greater than $\|\mathbf{G}\|$ (the absolute value of the element of \mathbf{G} with maximal absolute value). In this case the transformation from (γ, \mathbf{G}) to $(\beta(n, \lambda), \mathbf{B}(n, \lambda))$ gains a nice stochastic interpretation through uniformization (randomization) [7]. Apart from efficient numerical computation of transient probabilities of a CTMC, uniformization is a technique to interpret the behavior

of a CTMC through a discrete time Markov chain and a Poisson process. We utilize this approach for matrix-exponential functions below.

The matrix \mathbf{G} is a transient generator matrix and consequently the matrix $(\mathbf{I} + \mathbf{G}/\lambda)$ is the transient transition probability matrix of the uniformized version of the same process, where the uniformization rate is λ . The non-Markovian element of the $(\boldsymbol{\gamma}, \mathbf{G})$ representation is vector $\boldsymbol{\gamma}$, which contains negative elements. Our main question is if vector $\boldsymbol{\beta}(n, \lambda) = \boldsymbol{\gamma}\mathbf{W}$ is non-negative. We have

$$\boldsymbol{\gamma}\mathbf{W} = (\boldsymbol{\gamma}(\mathbf{I} + \mathbf{G}/\lambda)^n \mid \boldsymbol{\gamma}(\mathbf{I} + \mathbf{G}/\lambda)^{n-1}g(\lambda) \mid \boldsymbol{\gamma}(\mathbf{I} + \mathbf{G}/\lambda)^{n-2}g(\lambda) \mid \dots \mid \boldsymbol{\gamma}g(\lambda)). \quad (11)$$

The first block of this vector is of size u and the remaining n blocks are of size 1. The stochastic interpretation of vector $\boldsymbol{\gamma}\mathbf{W}$ is as follows: The last n elements of the vector are the first n probabilities of a discrete-time matrix-geometric distribution with initial vector $\boldsymbol{\gamma}$ and matrix $\mathbf{I} + \mathbf{G}/\lambda$. The first u elements of the vector are the transient value of the initial vector of the same matrix-geometric distribution after n steps.

The matrix-geometric series defined by the vector-matrix pair $(\boldsymbol{\gamma}, \mathbf{I} + \mathbf{G}/\lambda)$, $-\boldsymbol{\gamma}(\mathbf{I} + \mathbf{G}/\lambda)^{i-1}\mathbf{G}\mathbf{1}/\lambda$ is denoted by $\text{MG}(\boldsymbol{\gamma}, \mathbf{I} + \mathbf{G}/\lambda)$. To better approach the stochastic interpretation of $\boldsymbol{\gamma}\mathbf{W} > 0$ we first we show that $\boldsymbol{\gamma}(\mathbf{I} + \mathbf{G}/\lambda)^i g(\lambda) > 0$ for $i = 0, 1, \dots, n-1$ and $\boldsymbol{\gamma}(\mathbf{I} + \mathbf{G}/\lambda)^n > 0$ implies that $\text{MG}(\boldsymbol{\gamma}, \mathbf{I} + \mathbf{G}/\lambda)$ is positive for all $i \geq 1$.

The first condition ensures that the first n elements of the matrix-geometric series are positive. Since the matrix \mathbf{G} is a proper Markovian generator, the matrix $(\mathbf{I} + \mathbf{G}/\lambda)$ is also a proper transient transition probability matrix (with non-negative elements and all row sums less than or equal to 1). Due to the fact that $\boldsymbol{\gamma}(\mathbf{I} + \mathbf{G}/\lambda)^n$ is positive and $(\mathbf{I} + \mathbf{G}/\lambda)$ is a transient transition probability matrix we have that the rest of the elements of the probability mass function are also positive, since for $i > n$ we have

$$\boldsymbol{\gamma}(\mathbf{I} + \mathbf{G}/\lambda)^i g(\lambda) = \underbrace{\boldsymbol{\gamma}(\mathbf{I} + \mathbf{G}/\lambda)^n}_{>0} \underbrace{(\mathbf{I} + \mathbf{G}/\lambda)^{i-n}g(\lambda)}_{\geq 0 \text{ and } \neq 0} > 0.$$

In the second step we show that $f(t) > 0$ on $(0, \infty)$ if $\text{MG}(\boldsymbol{\gamma}, \mathbf{I} + \mathbf{G}/\lambda)$ is non-negative for all $i \geq 1$. According to the uniformization argument, for $f(t)$ we can write

$$\begin{aligned} f(t) &= -\boldsymbol{\gamma}e^{\mathbf{G}t}\mathbf{G}\mathbf{1} \\ &= -\boldsymbol{\gamma} \sum_{i=1}^{\infty} \frac{t^{i-1}}{(i-1)!} \mathbf{G}^{i-1}\mathbf{G}\mathbf{1} \\ &= -\boldsymbol{\gamma} \sum_{i=1}^{\infty} \frac{(\lambda t)^{i-1}}{(i-1)!} (\mathbf{G}/\lambda)^{i-1}\mathbf{G}\mathbf{1} \\ &= -\boldsymbol{\gamma} \sum_{i=1}^{\infty} \frac{(\lambda t)^{i-1}}{(i-1)!} e^{-\lambda t} (\mathbf{I} + \mathbf{G}/\lambda)^{i-1}\mathbf{G}\mathbf{1} \\ &= \sum_{i=1}^{\infty} \underbrace{\frac{\lambda(\lambda t)^{i-1}}{(i-1)!} e^{-\lambda t}}_{\text{Erlang}(i, \lambda) \text{ PDF}} \underbrace{\boldsymbol{\gamma}(\mathbf{I} + \mathbf{G}/\lambda)^{i-1}g(\lambda)}_{i\text{th element of } \text{MG}(\boldsymbol{\gamma}, \mathbf{I} + \mathbf{G}/\lambda)}. \end{aligned} \quad (12)$$

Which means that $f(t)$ is a convex combination of $\text{Erlang}(i, \lambda)$ probability density functions which are strictly positive in $(0, \infty)$ and the weights are the positive elements of the $\text{MG}(\boldsymbol{\gamma}, \mathbf{I} + \mathbf{G}/\lambda)$ series.

3.2 Basic algorithm

According to (11), in order to check if a non-Markovian vector-matrix pair $(\boldsymbol{\gamma}, \mathbf{G})$ with Markovian generator \mathbf{G} defines a PH distribution we need to find λ for which $\text{MG}(\boldsymbol{\gamma}, \mathbf{I} + \mathbf{G}/\lambda)$ series is positive for all $i \geq 1$. Furthermore, to check if the elements of the $\text{MG}(\boldsymbol{\gamma}, \mathbf{I} + \mathbf{G}/\lambda)$ series are positive for all $i \geq 1$ it is not necessary to check all elements of the series, but it is enough to check if the first elements of the series are positive up to the point when the transient vector becomes strictly positive.

To implement a procedure based on this approach we still need to describe a method to find an appropriate λ value. An intuitive interpretation of (12) might help in understanding the behavior of $\text{MG}(\boldsymbol{\gamma}, \mathbf{I} + \mathbf{G}/\lambda)$


```

1. POSITIVEMATRIXEXP( $\alpha$ ,  $\mathbf{A}$ )
2.  $\{\gamma, \mathbf{G}\} = \text{MONOCYCLICREPRESENTATION}(\alpha, \mathbf{A});$ 
3.  $\lambda = \text{MAX}(\text{REALPART}(\text{EIGENVALUES}(-\mathbf{A})));$ 
4. For( $i = 1, i \leq 20, i ++$ , % doubles  $\lambda$  20 times
5.    $vec = \gamma;$ 
6.   While( $\text{Min}(vec) < 0$ ,
7.      $vec = vec * (\mathbf{I} + \mathbf{G}/\lambda);$ 
8.      $out = -vec * \mathbf{G}/\lambda * \mathbf{1};$ 
9.     If( $out < 0$ , BreakWhile);
10.  ) % ENDWHILE
11.  If( $\text{MIN}(vec) \geq 0$ , Return(TRUE)); # otherwise  $out < 0$ 
12.   $\lambda = 2 * \lambda;$ 
13. ) % ENDFOR
14. Return(UNDECIDED);

```

Figure 3: Pseudo-code of the basic algorithm

as a function of λ . If λ is large then $\mathbf{I} + \mathbf{G}/\lambda$ is very close to the identity matrix and this way the product of γ and $\mathbf{I} + \mathbf{G}/\lambda$ remains close to γ and a lot of multiplication is needed to obtain a strictly positive vector $\gamma(\mathbf{I} + \mathbf{G}/\lambda)^n$. On the other hand if λ is small then $\mathbf{I} + \mathbf{G}/\lambda$ is far from the identity matrix and the elements of the $\gamma(\mathbf{I} + \mathbf{G}/\lambda)^n g(\lambda)$ series vary more. Consequently, the occurrence of a negative element is more likely.

Based on this behavior our proposed algorithm tries to check the positivity of the $\text{MG}(\gamma, \mathbf{I} + \mathbf{G}/\lambda)$ series with small λ values first, which is computationally cheaper, and in case of negative result it doubles λ . Here the negative result means the occurrence of a negative $\gamma(\mathbf{I} + \mathbf{G}/\lambda)^i g(\lambda)$ value before the vector $\gamma(\mathbf{I} + \mathbf{G}/\lambda)^i$ becomes positive. The pseudo-code of the algorithm which implements these steps is provided in Figure 3. Note that the inner loop (lines 6–10) is guaranteed to terminate, as $\gamma(\mathbf{I} + \mathbf{G}/\lambda)^i$ converges to the dominant eigenvector of $(\mathbf{I} + \mathbf{G}/\lambda)$, which is non-negative.

To avoid an infinite search for an appropriate λ value, the basic algorithm doubles λ at most 20 times. Therefore, the result of the algorithm has the following meaning: TRUE means that non-Markovian vector-matrix pair (α, \mathbf{A}) defines a PH distribution with strictly positive density in $(0, \infty)$. UNDECIDED means that $\gamma(\mathbf{I} + \mathbf{G}/\lambda)^i g(\lambda)$ for $i = 0, 1, \dots$ is not a strictly positive matrix-geometric series for the investigated λ values, whose maximum is 2^{20} times the absolute value of the dominant eigenvalue. In this case, (α, \mathbf{A}) might still define a valid PH distribution, but the algorithm cannot determine if this is the case.

3.2.1 Approximate PH distribution based on the basic method.

Theorem 4 allows to compute a $(\beta(n, \lambda), \mathbf{B}(n, \lambda))$ representation of (γ, \mathbf{G}) for a given (n, λ) pair. The size of the $(\beta(n, \lambda), \mathbf{B}(n, \lambda))$ representation ($u+n$) is increasing with n , and the potential nonzero elements of $\beta(n, \lambda)$ depend on both λ and n . The $(\beta(n, \lambda), \mathbf{B}(n, \lambda))$ representation obtained with any (n, λ) pair can be used for a PH approximation of the matrix-exponential function associated with (γ, \mathbf{G}) . If $\beta(n, \lambda)$ is obtained from $\beta(n, \lambda)$ by setting the negative elements to zero and re-normalizing, then the obtained Markovian vector-matrix representation, $(\tilde{\beta}(n, \lambda), \mathbf{B}(n, \lambda))$, is going to approximate the original density closely everywhere except around local minima.

The basic procedure searches for (n, λ) pairs which result in a non-negative vector $\beta(n, \lambda)$. If we eliminate line 9 of the POSITIVEMATRIXEXP procedure and store the out and vec values into vector $\beta(n, \lambda)$ and finally approximate $\beta(n, \lambda)$ by $\tilde{\beta}(n, \lambda)$ then $(\tilde{\beta}(n, \lambda), \mathbf{B}(n, \lambda))$ is computed by a (n, λ) pair which is optimal according to the POSITIVEMATRIXEXP procedure with the given computation limit.

```

1. POSITIVEMATRIXEXPWITHSHIFTING( $\alpha$ ,  $\mathbf{A}$ )
2.  $f[x] = -\alpha e^{\mathbf{A}x} \mathbf{A} \mathbf{1}$ ;
3.  $\lambda = \text{MAX}(\text{REALPART}(\text{EIGENVALUES}(-\mathbf{A})))$ ;
4.  $\{\gamma, \mathbf{G}\} = \text{MONOCYCLICGENERATOR}(\alpha, \mathbf{A})$ ;
5.  $iter = 0$ ;  $shift = 0$ ;  $ivec = \gamma$ ;
6. For( $i = 1, i \leq 20, i ++$ , % doubles  $\lambda$  20 times
7.    $vec = ivec$ ;
8.   While( $\text{Min}(vec) < 0$ ,
9.      $iter ++$ ;
10.     $vec = vec * (\mathbf{I} + \mathbf{G}/\lambda)$ ;
11.     $out = -vec * \mathbf{G}/\lambda * \mathbf{1}$ ;
12.    If( $out < 0, \text{BreakWhile}$ );
13.  ) % ENDWHILE
14.  If( $\text{MIN}(vec) \geq 0, \text{Return}(\text{TRUE})$ ); # otherwise  $out < 0$ 
15.  If( $f[shift + iter/\lambda] = < 0, \text{Return}(\text{FALSE})$ ); # otherwise  $f[shift + iter/\lambda] > 0$ 
16.   $shift += iter/\lambda$ ;
17.   $ivec = ivec e^{\mathbf{G} iter/\lambda}$ ;
18.   $\lambda = 2 * \lambda$ ;
19. ) % ENDFOR
20. Return( $\text{UNDECIDED}$ );

```

Figure 4: Pseudo-code of the algorithm with shifting

4 Shifting

The computational cost of checking the non-negativity of $\text{MG}(\gamma, \mathbf{I} + \mathbf{G}/\lambda)$ is increasing with increasing λ . It is a weak point of the basic algorithm that the expensive checking steps for large λ values need to be completed when $f(t) = \alpha e^{\mathbf{A}t}(-\mathbf{A})\mathbf{1}$ is negative for some t . Practically it means that the basic method is efficient when the result is TRUE, but it is inefficient otherwise. To enhance the efficiency of the procedure by efficiently obtaining FALSE result we further need to investigate the meaning of (12).

The Erlang(i, λ) probability density is concentrated around i/λ with variance i/λ^2 . It means that the contribution of the $\gamma(\mathbf{I} + \mathbf{G}/\lambda)^i g(\lambda)$ coefficient in $f(t)$ is most dominant at $t = 1/\lambda$. This way if $\gamma(\mathbf{I} + \mathbf{G}/\lambda)^i g(\lambda)$ is negative then $f(t)$ might also be negative at $t = 1/\lambda$, which can be used for a cost-efficient check for FALSE result.

Another way to improve the efficiency of the method is to exclude the part of the analysis of $f(t)$ which was found to be non-negative in the previous cycle of the procedure (with the previous value of λ). A positive $\gamma(\mathbf{I} + \mathbf{G}/\lambda)^i g(\lambda)$ series for $i = 0, 1, \dots, n-1$ indicates that the Erlang(i, λ) ($i = 0, 1, \dots, n-1$) density functions which are dominant over the $(0, n/\lambda)$ interval are weighted with positive coefficients. This way if $f(t)$ is positive at $t = n/\lambda$ then it is positive over the whole $(0, n/\lambda)$ interval. We utilize this fact by shifting the $f(t)$ function by n/λ to the left. A consequence of the shifting is that the potentially negative point of $f(t)$ gets closer to the origin and the required number of iterations of our algorithm to identify this negative value through the $\gamma(\mathbf{I} + \mathbf{G}/\lambda)^i g(\lambda)$ coefficients reduces accordingly. Another advantage of the explicit check of negative $f(t)$ values is that the outcome of the procedure can be better specified. The shifting method returns TRUE or FALSE if a positive or a negative answer is obtained within the predefined limit of precision (doubling λ 20 times in Figure 4) and returns UNDECIDED if the function cannot be classified with the given precision limit. The pseudo-code of the algorithm with these enhancements is provided in Figure 4.

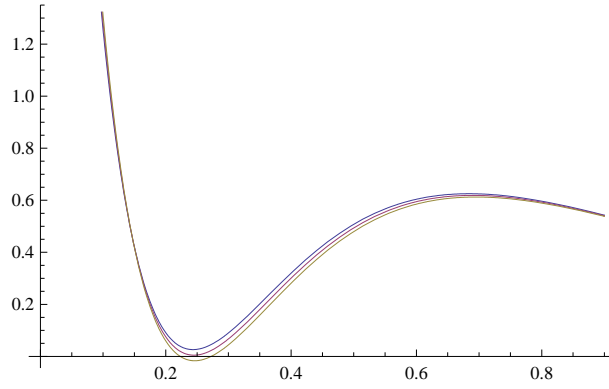


Figure 5: Matrix-exponential functions with a minimum close to zero.

5 Numerical examples

This section presents a number of numerical examples to illustrate the proposed method.

5.1 Densities close to zero

First we investigate the behavior of the algorithm with a series of vector-matrix pairs whose matrix-exponential functions have a minimum in $(0, \infty)$ which is close to zero. The evaluated vector-matrix pairs are $(\alpha_1, \mathbf{A}), \dots, (\alpha_4, \mathbf{A})$, where

$$\alpha_i = [1, -x_i, x_i], \quad i = 1, 2, 3, 4, \quad \mathbf{A} = \begin{pmatrix} -2 & 2 & 0 \\ 0 & -5 & 5 \\ 0 & 0 & -9 \end{pmatrix},$$

and

$$x_1 = 0.7, \quad x_2 = 0.72, \quad x_3 = 0.74, \quad x_4 = 0.7244301.$$

The matrix-exponential functions are plotted in Figure 5 and the following table summarizes the main parameters of the procedure.

	Result	Min. around 0.24	Iterations		Final shift
			outer loop	inner loop	
α_1, \mathbf{A}	TRUE	0.0261394	5	22	0.208333
α_2, \mathbf{A}	TRUE	0.00475063	6	29	0.229167
α_3, \mathbf{A}	FALSE	-0.0167315	4	17	0.138889
α_4, \mathbf{A}	UNDECIDED	0	20	80	0.245182

Except for α_4, \mathbf{A} the procedure requires 3–6 iterations for λ values with these examples. The first two examples indicate that the closer the minimum is to zero, the higher the computational cost of the procedure. The third example indicates that the introduced error-checking method efficiently recognizes negative density values. In spite of the lower absolute value of the minimum of (α_3, \mathbf{A}) with respect to (α_1, \mathbf{A}) (0.0167315 versus 0.0261394), the computational cost of obtaining the FALSE answer for (α_3, \mathbf{A}) is less than the one for obtaining the TRUE answer for (α_1, \mathbf{A}) . The last example indicate that the procedure does not decide the positivity of the density if its local minimum is zero. In this case the procedure runs until the predefined limit of the computations defined in Line 6 of Figure 4.

5.2 Multi-modal density

The proposed shifting method intends adopting the peculiarities of the matrix-exponential functions which occur due to the coefficients with different signs associated with the elementary functions of the eigenvalues.

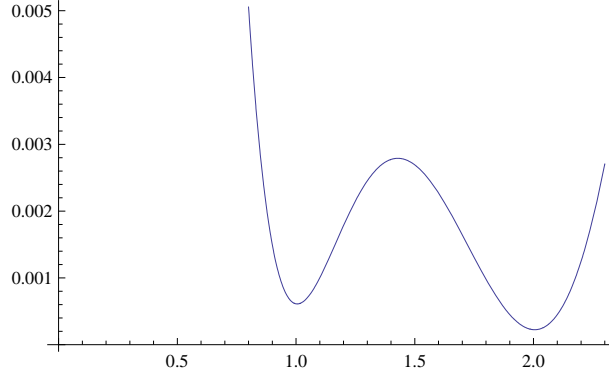


Figure 6: Matrix-exponential function with two local minima.

For example, (α_5, \mathbf{A}_5) with

$$\alpha_5 = [3.99334, -5.99002, 4.32612, -1.99667, 0.667221], \mathbf{A}_5 = \begin{pmatrix} -1 & 1 & 0 & 0 & 0 \\ 0 & -1 & 1 & 0 & 0 \\ 0 & 0 & -1 & 1 & 0 \\ 0 & 0 & 0 & -1 & 1 \\ 0 & 0 & 0 & 0 & -1 \end{pmatrix},$$

exhibits two local minima, as depicted in Figure 6. The main parameters in the table below indicate that the procedure increases the λ value three times in order to determine that the minimum at 1 is not negative, and after that it efficiently shifts the representation close to the next minimum (1.875).

	Result	Minimum	Iterations		Shift 1	Shift 2	Shift 3	Shift 4
			outer loop	inner loop				
α_5, \mathbf{A}_5	TRUE	0.000224626	6	66	0.25	0.5	0.75	1.875

5.3 Complex eigenvalues

In case of complex eigenvalues the matrix-exponential functions can have several minima. According to our intuitive understanding the procedure needs to investigate the minima caused by the coefficients with different signs associated with the elementary functions of the eigenvalues, but the infinitely alternating terms associated with complex eigenvalues do not cause any extra problem. For example, the matrix-exponential function defined by

$$\alpha_6 = [-3.07692, 2.30769, 1.76923], \mathbf{A}_6 = \begin{pmatrix} -2 & 0.56 & 0 \\ -0.56 & -2 & 1 \\ 0 & 0 & -1 \end{pmatrix},$$

and depicted in Figure 7 is identified with the following parameters.

	Result	Minimum	Iterations		Shift 1	Shift 2	Shift 3
			outer loop	inner loop			
α_6, \mathbf{A}_6	TRUE	0.00443014	5	25	0.25	0.625	0.9375

5.4 Qualitative properties

The proposed method is applicable up to a given numerical precision. The example in Section 5.1 demonstrates this property. If the local minimum of the density decreases to zero the number of required iterations increases to infinity. If the proposed computation is stopped at any finite limit, then there are positive densities with very small positive value at the local minimum which cannot be decided by the algorithm within that finite limit. In this case the shifting method returns with UNDECIDED.

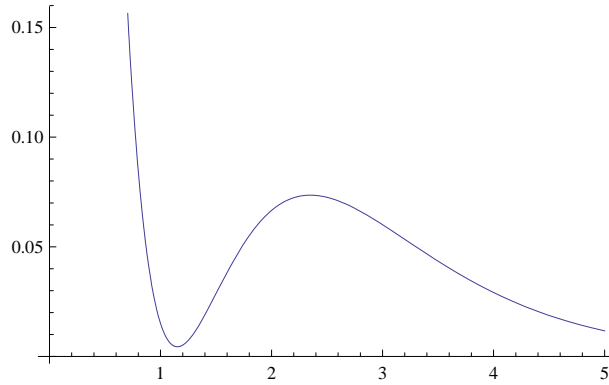


Figure 7: Matrix-exponential function with complex eigenvalues.

The two main conditions of Theorem 3 are that the dominant eigenvalue is real and that the density is positive on $(0, \infty)$. Both of these conditions can be approached arbitrarily closely by a phase-type distribution. If the limit of the first condition is approached, i.e., there is a complex eigenvalue pair, whose real part is smaller than, but very close to the dominant eigenvalue, then the size of the monocyclic representation (matrix \mathbf{G}) gets very large, because the size of the FE block associated with this eigenvalue (defined by (6)) gets very large. If the limit of the second condition is approached (i.e. the density is very close to zero at some point), then the number of required iterations (steps to increase λ) tends to infinity.

5.5 Performance comparison

Based on the available literature and documentation we assume that non of the previously proposed procedures applies the proposed efficient computation of a feasible λ, n pair and the explicit transformation presented in Theorem 4 in step S3. The only tool of similar purpose, the authors are aware of, is the MoMI tool [12] which was developed for demonstrating the computability of the monocyclic representation and not for efficient analysis of matrix exponential functions. The MoMI tool is partially documented in S. Mocanu's Ph.D. [13], but unfortunately, the implementation of step S3 is not detailed there. We used the IDA disassembler (<https://www.hex-rays.com/products/ida/>) to gain information on the related procedures of MoMI. The obtained low level program code is rather hard to interpret. It seems that the MoMI tool checks the time (t^*) at which $\gamma e^{\mathbf{G}t}$ is non-negative at fixed (0.1) time units and searches for feasible λ, n pair with $\lambda = n/t^*$ and n incremented from 1 till 16. We evaluated the previous examples also with the direct method of Figure 8. The POSITIVEMATRIXEXPDIRECT procedure returned UNDECIDED for all examples of this section. In case of Example (α_1, \mathbf{A}) the largest x value for which the procedure returned TRUE is $x = 0.49$ (i.e., $(\{1, -0.49, 0.49\}, \mathbf{A})$) and the facts that procedure POSITIVEMATRIXEXPDIRECT obtains β by a solution of a set of linear equation and generates the large $\{\beta, \mathbf{B}\}$ representation makes this direct method slow and numerically sensitive.

6 Conclusions

Despite the characterization result of O'Kinneide [16] from 1990, the practical application of non-Markovian vector-matrix pairs has always been limited by the lack of efficient methods for checking if a non-Markovian vector-matrix pair defines a valid PH distribution. Our proposed checking method is also based on [16], but the characterization result is combined with several tricks and model transformation steps to improve its efficiency. The applied problem transformation steps include transformations from the general non-Markovian generator matrix to the Markovian one proposed in [11] and from a continuous-time description to a kind of uniformized one.

The Mathematica implementation of the presented methods is publicly available as part of the BuTools package [3].

1. `POSITIVEMATRIXEXPDIRECT(α , \mathbf{A})`
2. `{ γ , \mathbf{G} } = MONOCYCLICREPRESENTATION(α , \mathbf{A});`
3. `If(MIN(γ) \geq 0, Return(TRUE));`
4. `t = 0;`
5. `While(Min($\gamma e^{0.1t\mathbf{G}}$) $<$ 0, t ++);`
6. `For(n = 1, n \leq 16, n ++, % size of Erlang tail`
7. `$\lambda = n/(0.1t)$;`
8. `{ β , \mathbf{B} } = EXTENDEDREPRESENTATION(γ , \mathbf{G} , n, λ);`
9. `If(MIN(β) \geq 0, Return(TRUE));`
10. `) % ENDFOR`
11. `Return(UNDECIDED);`

Figure 8: Pseudo-code of the direct algorithm

Acknowledgement

This work is partially supported by the Hungarian Government through the TÁMOP-4.2.2C-11/1/KONV-2012-0001 and the OTKA K101150 projects.

We would like to thank the anonymous reviewers for their invaluable input, which has greatly helped us to improve the paper and Tamás Izsó and Gábor Horváth for their help in disassembling and interpreting the MoMI tool.

References

- [1] N. G. Bean and B. F. Nielsen. Quasi-birth-and-death processes with rational arrival process components. *Stochastic Models*, 26(3):309–334, 2010.
- [2] Nigel G. Bean, Mark Fackrell, and Peter Taylor. Characterization of matrix-exponential distributions. *Stochastic Models*, 24(3):339–363, 2008.
- [3] Levente Bodrog, Peter Buchholz, Armin Heindl, András Horváth, Gábor Horváth, István Kolossváry, Zoltán Németh, Philipp Reinecke, Miklós Telek, and Miklós Vécsei. Butools: Program packages for computations with PH, ME distributions and MAP, RAP processes. <http://webspn.hit.bme.hu/~butools>, October 2011.
- [4] P. Buchholz and M. Telek. Stochastic Petri nets with matrix exponentially distributed firing times. *Performance Evaluation*, 67(12):1373–1385, 2010.
- [5] Peter Buchholz and Miklós Telek. On minimal representation of rational arrival processes. *Annals of Operations Research*, 2011. DOI 10.1007/s10479-011-1001-5.
- [6] C. Commault and S. Mocanu. Phase-type distributions and representations: some open problems for system theory. *Int. J. Control*, 76(6):566–580, 2003.
- [7] Winfried Grassmann. Transient solutions in markovian queues: An algorithm for finding them and determining their waiting-time distributions. *European Journal of Operational Research*, 1(6):396 – 402, 1977.
- [8] C. M. Harris, W. G. Marchal, and R. F. Botta. A note on generalized hyperexponential distributions. *Communications in Statistics. Stochastic Models*, 8(1):179–191, 1992.
- [9] I. Kolossváry and M. Telek. Explicit evaluation of me(3) membership. In *8th International Conference on Quantitative Evaluation of SysTems (QEST), fast abstract*, pages 11–13, Aachen, Germany, sept. 2011. Centre for Telematics and Information Technology (CTIT), University of Twente.

- [10] L. Lipsky. *Queueing Theory: A Linear Algebraic Approach*. Springer, 2008.
- [11] S. Mocanu and C. Commault. Sparse representations of phase-type distributions. *Commun. Stat., Stochastic Models*, 15(4):759 – 778, 1999.
- [12] Stefan Mocanu. Monocyclic markovian identifier (MoMI) home page. Website: <http://membres.multimania.fr/mocanu/>, 1998. last checked: 12.06.2013, unavailable in 2014.
- [13] Stefan Mocanu. *Construction et propriétés des représentations monocycliques des lois de type phase*. PhD thesis, Institut Polytechnique de Grenoble, 1999.
- [14] M.F. Neuts. *Matrix Geometric Solutions in Stochastic Models*. Johns Hopkins University Press, Baltimore, 1981.
- [15] C. Art O’Cinneide. Phase-type distributions: Open problems and a few properties. *Stochastic Models*, 15(4):731–757, 1999.
- [16] Colm Art O’Cinneide. Characterization of phase-type distributions. *Communications in Statistics. Stochastic Models*, 6(1):1–57, 1990.
- [17] Ushio Sumita and Yasushi Masuda. Classes of probability density functions having laplace transforms with negative zeros and poles. *Advances in Applied Probability*, 19(3):632–651, 1987.
- [18] M. Telek and G. Horváth. A minimal representation of Markov arrival processes and a moments matching method. *Performance Evaluation*, 64(9-12):1153–1168, 2007.
- [19] A. van de Liefvoort. The moment problem for continuous distributions. Technical report, University of Missouri, WP-CM-1990-02, Kansas City, 1990.