J. Appl. Probab. **60,** 1–13 (2023)
*doi:*10.1017/jpr.2022.25

A MARKOV JUMP PROCESS ASSOCIATED WITH THE MATRIX-EXPONENTIAL DISTRIBUTION

OSCAR PERALTA ,* ** *University of Lausanne and the University of Adelaide*

Abstract

Let f be the density function associated to a matrix-exponential distribution of parameters $(\boldsymbol{\alpha}, T, \boldsymbol{s})$. By exponentially tilting f, we find a probabilistic interpretation which generalizes the one associated to phase-type distributions. More specifically, we show that for any sufficiently large $\lambda \geq 0$, the function $x \mapsto \left(\int_0^\infty e^{-\lambda s} f(s) \mathrm{d}s\right)^{-1} e^{-\lambda x} f(x)$ can be described in terms of a finite-state Markov jump process whose generator is tied to T. Finally, we show how to revert the exponential tilting in order to assign a probabilistic interpretation to f itself.

Keywords: Phase-type distribution; matrix-analytic methods; exponential tilting; finite-state Markov jump process

2020 Mathematics Subject Classification: Primary 60J27
Secondary 60E05; 60K25

1. Introduction

A phase-type distribution corresponds to the law of $Y := \inf\{t \geq 0 : J_t = \star\}$ where $\{J_t\}_{t\geq 0}$ is a Markov jump process with state space $\{1, \ldots, p\} \cup \{\star\}$, with $\{1, \ldots, p\}$ assumed to be transient states and $\{\star\}$ absorbing. If $\{J_t\}_{t\geq 0}$ has a block-partitioned initial distribution $(\boldsymbol{\pi}, 0)$ and intensity matrix given by

$$\begin{bmatrix} A & \boldsymbol{b} \\ \boldsymbol{0} & 0 \end{bmatrix} \quad \text{with} \quad \boldsymbol{b} = -A\boldsymbol{1}, \tag{1}$$

where $\boldsymbol{0}$ represents a p-dimensional row vector of 0s and $\boldsymbol{1}$ a p-dimensional column vector of 1s, then we say that the phase-type distribution is of parameters $(\boldsymbol{\pi}, A)$. Via simple probabilistic arguments, it can be shown that the density function of a phase-type distribution of parameters $(\boldsymbol{\pi}, A)$ is of the form

$$g(x) = \boldsymbol{\pi} e^{Ax} \boldsymbol{b}, \quad x \geq 0. \tag{2}$$

Indeed, the vector $\boldsymbol{\pi} e^{Ax}$ yields the probabilities of $\{J_t\}_{t\geq 0}$ being in some state $\{1, \ldots, p\}$ at time x, and $\boldsymbol{b}$ corresponds to the intensity vector of an absorption happening immediately after. Phase-type distributions were first introduced in [14] with the aim of constructing a robust and tractable class of distributions on $\mathbb{R}_+$ to be used in econometric problems. A more

Received 3 March 2021; revision received 30 January 2022.
* Postal address: University of Lausanne, Faculty of Business and Economics, Quartier de Chambronne, 1015 Lausanne, Switzerland. Email address: oscar.peraltagutierrez@unil.ch
** Postal address: The University of Adelaide, School of Mathematical Sciences, SA 5005, Australia. Email address: oscar.peraltagutierrez@adelaide.edu.au

comprehensive study of phase-type distributions was carried on by Neuts [15, 16], whose work popularized their use in more general stochastic models.

On the other hand, a *matrix-exponential distribution* of dimension $p \geq 1$ is an absolutely continuous distribution on $(0, \infty)$ whose density function can be written as

$$f(x) = \boldsymbol{\alpha} e^{Tx} \boldsymbol{s}, \quad x \geq 0, \tag{3}$$

where $\boldsymbol{\alpha} = (\alpha_1, \ldots, \alpha_p)$ is a p-dimensional row vector, $T = \{t_{ij}\}_{i,j \in \{1,\ldots,p\}}$ is a $(p \times p)$-dimensional square matrix, and $\boldsymbol{s} = (s_1, \ldots, s_p)^\intercal$ is a p-dimensional column vector, all with complex entries. If the dimension need not be specified, we refer to such a distribution simply as *matrix-exponential*. It follows from (2) and (3) that the class of phase-type distributions is a subset of those that are matrix-exponential, with the inclusion being strict (see [17] for details on the latter).

Matrix-exponential distributions were first studied in [8, 9] through the concept of complex-valued transition probabilities. More precisely, these papers showed that certain systems with complex-valued elements can be formally studied by analytical means without assigning a specific physical interpretation to their components. While their method provided mathematical rigour to systems 'driven' by complex-valued intensity matrices, it failed to provide a physical meaning to each individual component, as opposed to the case of Markov jump processes with genuine intensity matrices. Later on, it was proved in [5, 17] that matrix-exponential distributions have an interpretation in terms of a Markov process with continuous state space, as opposed to the finite-state-space one that phase-type distributions enjoy. Even after the discovery of these physical interpretations of matrix-exponential distributions, however, properties of this class of distributions are still not as well understood as they are for its phase-type counterpart. One of the main reasons for this is that processes with continuous state space are more difficult to handle, so that studying matrix-exponential distributions by physical means requires a more sophisticated framework. For example, this is the case in [2, 4, 3], where the theory of piecewise deterministic Markov processes is used to study models with matrix-exponential components. Thus, having a finite-state system interpretation for matrix-exponential distributions available may potentially lead to the discovery of new properties, as has traditionally been the case for phase-type distributions.

Functions of the form (3) also play an important role in control theory, more specifically, in the topic of single-input–single-output (SISO) linear systems. Such systems are described by a column-vector function $\boldsymbol{x} : \mathbb{R}_+ \to \mathbb{R}^p$ and $y : \mathbb{R}_+ \to \mathbb{R}$ which satisfy the ordinary differential equations

$$\frac{\mathrm{d}\boldsymbol{x}(t)}{\mathrm{d}t} = T_0 \boldsymbol{x}(t) + \boldsymbol{b}_0 u(t),$$
$$y(t) = \boldsymbol{\alpha}_0 \boldsymbol{x}(t);$$

here u is called the input function, $\boldsymbol{x}$ the state function, and y the output function. SISO linear systems which produce a nonnegative output from a nonnegative input are said to be externally positive. It can be shown [12, Theorem 1] that if $\boldsymbol{x}(0) = \boldsymbol{0}$, then the output function takes the form

$$y(t) = \int_0^t h_0(t - z) u(s) \mathrm{d}z,$$

where $h_0(z) = \boldsymbol{\alpha}_0 e^{T_0 z} \boldsymbol{s}_0$. From this, one can deduce that the system is externally positive if and only if h_0 is a nonnegative function. If, additionally, h_0 is bounded, then h_0 is essentially a

https://doi.org/10.1017/jpr.2022.25 Published online by Cambridge University Press

scaled matrix-exponential density function. This nicely links the theory of externally positive SISO linear systems with that of matrix-exponential distributions, both of which share some fundamental research questions, such as the positive realization and minimality problems. See [7] for a detailed account of the duality of phase-type and matrix-exponential distributions in control theory.

In this paper we give a physical interpretation to *each element* of the parameters $(\boldsymbol{\alpha}, T, \boldsymbol{s})$ associated to the matrix-exponential density function (3) satisfying the following conditions:

A1. The elements of $\boldsymbol{\alpha}$, T and $\boldsymbol{s}$ are real.

A2. The dominant eigenvalue of T, denoted by σ_0, is real and strictly negative.

Since it can be shown that for a given matrix-exponential density of the form (3) the parameters $(\boldsymbol{\alpha}, T, \boldsymbol{s})$ can be chosen is such a way that **A1** and **A2** hold (see [1]), the interpretation that we develop essentially completes the picture laid out in [8, 9]. Our method, inspired by the recent work in [18], provides a transparent interpretation of $(\boldsymbol{\alpha}, T, \boldsymbol{s})$ in terms of a finite-state Markov jump process. To do so, we employ the technique known as *exponential tilting*, which means that we focus on the density proportional to $e^{-\lambda \cdot} f(\cdot)$ for large enough $\lambda > 0$. After we perform this transformation, we construct a Markov jump process on a finite state space formed by two groups: the *original states* and the *anti-states*, the latter being a copy of the former. Heuristically, jumps within the set of original states or within the set of anti-states occur according to the off-diagonal nonnegative 'jump intensities' of T, while jumps between the original and the anti-states occur according to the negative 'jump intensities' of T. Eventual absorption or termination happens, and each realization 'carries' a positive or negative sign depending only on its initial and final state. Our main contribution is to show that this mechanism yields the exponentially tilted matrix-exponential distribution, and, by reverting the exponential tilting, to provide some probabilistic insight into the original matrix-exponential distribution as well.

The structure of the paper is as follows. In Section 2 we provide a brief exposition on exponential tilting and how it affects the representation of a matrix-exponential distribution. In Section 3 we present our main results, Theorem 3.2 and Corollary 3.1, where we give a precise interpretation of an exponentially tilted matrix-exponential density in terms of a Markov jump process. Finally, in Section 4 we provide methods to recover formulae and probabilistic interpretations for matrix-exponential distributions for which the assumptions **A1** and **A2** hold, based on the results of their exponentially tilted versions.

2. Preliminaries

Exponential tilting, also known as the *Escher transform*, is a technique which transforms any probability density function f with support on $[0, \infty)$ into a new probability density function f_λ defined by

$$f_\lambda(x) = \frac{e^{-\lambda x} f(x)}{\int_0^\infty e^{-\lambda r} f(r) \mathrm{d}r}, \qquad x \geq 0,$$

where $\lambda \geq 0$ is the *tilting rate*. The use of exponential tilting goes back at least to [11], where it was used to build upon Cramér's classical actuarial models [10]. Later on, the exponential tilting method played a prominent role in the theory of option pricing [13].

The exponentially tilted version of a matrix-exponential distribution has a simple form which happens to be matrix-exponential itself. To see this, notice that if f is of the

https://doi.org/10.1017/jpr.2022.25 Published online by Cambridge University Press

form (3), then

$$\int_0^\infty e^{-\lambda r} f(r)\mathrm{d}r = \int_0^\infty e^{-\lambda r}(\boldsymbol{\alpha} e^{Tr}\boldsymbol{s})\mathrm{d}r = \boldsymbol{\alpha}(\lambda I - T)^{-1}\boldsymbol{s},$$

where we used the fact that $T-\lambda I$ has eigenvalues with strictly negative real parts and thus $e^{(T-\lambda I)r}$ vanishes as $r\to\infty$. Thus,

$$f_\lambda(x) = \frac{e^{-\lambda x}(\boldsymbol{\alpha} e^{Tx}\boldsymbol{s})}{\boldsymbol{\alpha}(\lambda I - T)^{-1}\boldsymbol{s}} = \left(\frac{\boldsymbol{\alpha}}{\boldsymbol{\alpha}(\lambda I - T)^{-1}\boldsymbol{s}}\right) e^{(T-\lambda I)x}\boldsymbol{s}, \quad x \geq 0, \tag{4}$$

implying that f_λ corresponds to the density function of a matrix-exponential distribution of parameters $\left(\frac{\boldsymbol{\alpha}}{\boldsymbol{\alpha}(\lambda I - T)^{-1}\boldsymbol{s}}, T-\lambda I, \boldsymbol{s}\right)$.

Recall that the parameters $(\boldsymbol{\alpha}, T, \boldsymbol{s})$ need not have a probabilistic meaning in terms of a finite-state-space Markov chain, as opposed to the parameters associated to phase-type distributions. For instance, the parameters

$$\boldsymbol{\alpha} = (1, 0, 0), \quad T = \begin{bmatrix} -1 & -1 & 2/3 \\ 1 & -1 & -2/3 \\ 0 & 0 & -1 \end{bmatrix}, \quad \boldsymbol{s} = \begin{bmatrix} 4/3 \\ 2/3 \\ 1 \end{bmatrix} \tag{5}$$

yield a valid matrix-exponential distribution whose density function is given by $f(x) = \frac{2}{3}e^{-x}(1+\cos(x))$, and where the dominant eigenvalue of T is -1 (see [6, Example 4.5.21] for details). In the following section we show how to assign a probabilistic meaning to the exponentially tilted version of (5), and more generally to those having the properties **A1** and **A2**, in terms of a finite-state Markov jump process.

3. Main results

Let $(\boldsymbol{\alpha}, T, \boldsymbol{s})$ be parameters associated to a p-dimensional matrix-exponential distribution which have the properties **A1** and **A2**. For $1 \leq i, j \leq p$ denote by t_{ij} the (i, j) entry of T, and denote by s_i the ith entry of $\boldsymbol{s}$. For $\ell \in \{+, -\}$, define the $(p \times p)$-dimensional matrix $T^\ell = \left\{t^\ell_{ij}\right\}_{1\leq i,j\leq p}$ and the p-dimensional column vector $\boldsymbol{s}^\ell = \left(s^\ell_1, \ldots, s^\ell_p\right)^\intercal$ where

$$t^\pm_{ij} = \max\{0, \pm t_{ij}\} \quad \forall\, i \neq j,$$

$$t^\pm_{ii} = \pm\min\{0, \pm t_{ii}\} \quad \forall\, i, \text{ and}$$

$$s^\pm_i = \max\{0, \pm s_i\} \quad \forall\, i.$$

It follows that T^+ has nonnegative off-diagonal elements and nonpositive diagonal elements, T^- is a nonnegative matrix, $\boldsymbol{s}^+$ and $\boldsymbol{s}^-$ are nonnegative column vectors, $T = T^+ - T^-$, and $\boldsymbol{s} = \boldsymbol{s}^+ - \boldsymbol{s}^-$. Now, let

$$\lambda_0 = \min\left\{ r \geq 0 : s^+_i + s^-_i + \sum_{j=1}^p (t^+_{ij} + t^-_{ij}) \leq r \text{ for all } 1 \leq i \leq p \right\}.$$

https://doi.org/10.1017/jpr.2022.25 Published online by Cambridge University Press

For some fixed $\lambda \geq \lambda_0$, consider a (possibly) terminating Markov jump process $\{\varphi_t^\lambda\}_{t\geq 0}$ driven by the block-partitioned subintensity matrix

$$G = \begin{bmatrix} T^+ - \lambda I & T^- & s^+ & s^- \\ T^- & T^+ - \lambda I & s^- & s^+ \\ \mathbf{0} & \mathbf{0} & 0 & 0 \\ \mathbf{0} & \mathbf{0} & 0 & 0 \end{bmatrix} \tag{6}$$

evolving on the state space $\mathcal{E} = \mathcal{E}^o \cup \mathcal{E}^a \cup \{\Delta^o\} \cup \{\Delta^a\}$ where $\mathcal{E}^o := \{1^o, 2^o, \ldots, p^o\}$ and $\mathcal{E}^a = \{1^a, 2^a, \ldots, p^a\}$. The state space $\mathcal{E}$ may be thought as the union of two sets: a collection of *original states* $\mathcal{E}^o \cup \{\Delta^o\}$ and a collection of *anti-states* $\mathcal{E}^a \cup \{\Delta^a\}$, where both Δ^o and Δ^a are absorbing. In the case $\lambda > \lambda_0$, the process $\{\varphi_t^\lambda\}_{t\geq 0}$ alternates between sojourn times in $\mathcal{E}^o$ and $\mathcal{E}^a$ up until one of the following happens: (a) get absorbed into Δ^o, (b) get absorbed into Δ^a, or (c) undergo termination due to the defect of (6). If $\lambda = \lambda_0$, the states $\mathcal{E}^o \cup \mathcal{E}^a$ may or may not be transient, their status depending on the values of T.

In Theorem 3.1 we establish a link between the absorption probabilities of $\{\varphi_t^\lambda\}_{t\geq 0}$ and the vector $e^{(T-\lambda I)x}\boldsymbol{s}$ appearing in the exponentially tilted matrix-exponential density (4). More specifically, we express each element in $e^{(T-\lambda I)x}\boldsymbol{s}$ as the sum of some positive density function and some negative density function, where the positive density is associated to an absorption of $\{\varphi_t^\lambda\}_{t\geq 0}$ to Δ^o, while the negative density function corresponds to an absorption of $\{\varphi_t^\lambda\}_{t\geq 0}$ to Δ^a. To shorten notation, from now on we denote by $\mathbb{P}_j$ ($\mathbb{E}_j$), $j \in \mathcal{E}$, the probability measure (expectation) associated to $\{\varphi_t^\lambda\}_{t\geq 0}$ conditional on the event $\{\varphi_0^\lambda = j\}$.

Theorem 3.1. *Let $\lambda \geq \lambda_0$ be such that the states $\mathcal{E}^o \cup \mathcal{E}^a$ are transient. Define*

$$\tau = \inf\{x \geq 0 : \varphi_x^\lambda \notin \mathcal{E}^o \cup \mathcal{E}^a\}. \tag{7}$$

Then, for $i \in \{1, \ldots, p\}$ and $x \geq 0$,

$$\left(\boldsymbol{e}_i^\mathsf{T} e^{(T-\lambda I)x}\boldsymbol{s}\right)\mathrm{d}x = \mathbb{E}_{i^o}\left[\mathbb{1}\{\tau \in [x, x+\mathrm{d}x]\}\beta(\varphi_\tau^\lambda)\right] \tag{8}$$

$$= (\boldsymbol{e}_i^\mathsf{T}, \mathbf{0}) \exp\left(\begin{bmatrix} T^+ - \lambda I & T^- \\ T^- & T^+ - \lambda I \end{bmatrix} x\right) \begin{bmatrix} \boldsymbol{s} \\ -\boldsymbol{s} \end{bmatrix} \mathrm{d}x, \tag{9}$$

where $\boldsymbol{e}_i$ denotes the column vector with 1 as its ith entry and 0 elsewhere, and $\beta(j) := \mathbb{1}\{j = \Delta^o\} - \mathbb{1}\{j = \Delta^a\}$. Moreover,

$$\left(-\boldsymbol{e}_i^\mathsf{T} e^{(T-\lambda I)x}\boldsymbol{s}\right)\mathrm{d}x = \mathbb{E}_{i^a}\left[\mathbb{1}\{\tau \in [x, x+\mathrm{d}x]\}\beta(\varphi_\tau^\lambda)\right] \tag{10}$$

$$= (\mathbf{0}, \boldsymbol{e}_i^\mathsf{T}) \exp\left(\begin{bmatrix} T^+ - \lambda I & T^- \\ T^- & T^+ - \lambda I \end{bmatrix} x\right) \begin{bmatrix} \boldsymbol{s} \\ -\boldsymbol{s} \end{bmatrix} \mathrm{d}x. \tag{11}$$

Proof. The block structure of (6) implies that

$$\mathbb{P}_{i^o}(\tau \in [x, x+\mathrm{d}x], \varphi_\tau^\lambda = \Delta^o) = (\boldsymbol{e}_i^\mathsf{T}, \mathbf{0}) \exp\left(\begin{bmatrix} T^+ - \lambda I & T^- \\ T^- & T^+ - \lambda I \end{bmatrix} x\right) \begin{bmatrix} \boldsymbol{s}^+ \\ \boldsymbol{s}^- \end{bmatrix} \mathrm{d}x,$$

$$\mathbb{P}_{i^o}(\tau \in [x, x+\mathrm{d}x], \varphi_\tau^\lambda = \Delta^a) = (\boldsymbol{e}_i^\mathsf{T}, \mathbf{0}) \exp\left(\begin{bmatrix} T^+ - \lambda I & T^- \\ T^- & T^+ - \lambda I \end{bmatrix} x\right) \begin{bmatrix} \boldsymbol{s}^- \\ \boldsymbol{s}^+ \end{bmatrix} \mathrm{d}x;$$

https://doi.org/10.1017/jpr.2022.25 Published online by Cambridge University Press

therefore, the right-hand side of (8) is equal to (9). Next, we prove that (8) holds.

Define the collection of $(p \times p)$-dimensional matrices $\{\Phi_{oa}(x)\}_{x\geq 0}$, $\{\Phi_{ao}(x)\}_{x\geq 0}$, $\{\Phi_{oo}(x)\}_{x\geq 0}$, and $\{\Phi_{aa}(x)\}_{x\geq 0}$ by

$$(\Phi_{oa}(x))_{ij} = \mathbb{P}_{i^o}(\tau > x, \varphi^\lambda_x = j^a), \qquad (\Phi_{ao}(x))_{ij} = \mathbb{P}_{i^a}(\tau > x, \varphi^\lambda_x = j^o),$$
$$(\Phi_{oo}(x))_{ij} = \mathbb{P}_{i^o}(\tau > x, \varphi^\lambda_x = j^o), \qquad (\Phi_{aa}(x))_{ij} = \mathbb{P}_{i^a}(\tau > x, \varphi^\lambda_x = j^a),$$

for all $i, j \in \{1, \ldots, p\}$. By the symmetry of the subintensity matrix G it is clear that for all $x \geq 0$, $\Phi_{oa}(x) = \Phi_{ao}(x)$ and $\Phi_{oo}(x) = \Phi_{aa}(x)$, even if their probabilistic interpretations differ. For all $x \geq 0$, let $\Phi_o(x) := \Phi_{oo}(x) - \Phi_{oa}(x)$ and $\Phi_a(x) := \Phi_{aa}(x) - \Phi_{ao}(x)$. Define

$$\gamma = \inf\{r \geq 0 : \varphi^\lambda_r \notin \mathcal{E}^o\}.$$

Then, for $i, j \in \{1, \ldots, p\}$,

$$\begin{aligned}
\boldsymbol{e}_i^\intercal \Phi_o(x)\boldsymbol{e}_j &= \mathbb{P}_{i^o}(\tau > x, \varphi^\lambda_x = j^o) - \mathbb{P}_{i^o}(\tau > x, \varphi^\lambda_x = j^a)\\
&= \left\{\mathbb{P}_{i^o}(\gamma > x, \tau > x, \varphi^\lambda_x = j^o) + \mathbb{P}_{i^o}(\gamma \leq x, \tau > x, \varphi^\lambda_x = j^o)\right\}\\
&\quad - \left\{\mathbb{P}_{i^o}(\gamma > x, \tau > x, \varphi^\lambda_x = j^a) + \mathbb{P}_{i^o}(\gamma \leq x, \tau > x, \varphi^\lambda_x = j^a)\right\}\\
&= \left\{\mathbb{P}_{i^o}(\gamma > x, \varphi^\lambda_x = j^o) + \int_0^x \mathbb{P}_{i^o}(\gamma \in [r, r + \mathrm{d}r], \tau > x, \varphi^\lambda_x = j^o)\right\}\\
&\quad - \left\{\mathbb{P}_{i^o}(\gamma > x, \varphi^\lambda_x = j^a) + \int_0^x \mathbb{P}_{i^o}(\gamma \in [r, r + \mathrm{d}r], \tau > x, \varphi^\lambda_x = j^a)\right\}\\
&= \mathbb{P}_{i^o}(\gamma > x, \varphi^\lambda_x = j^o)\\
&\quad + \int_0^x \sum_{k=1}^p \mathbb{P}_{i^o}(\gamma \in [r, r + \mathrm{d}r], \varphi^\lambda_\gamma = k^a)\mathbb{P}_{k^a}(\tau > x - r, \varphi^\lambda_{x-r} = j^o)\\
&\quad - \int_0^x \sum_{k=1}^p \mathbb{P}_{i^o}(\gamma \in [r, r + \mathrm{d}r], \varphi^\lambda_\gamma = k^a)\mathbb{P}_{k^a}(\tau > x - r, \varphi^\lambda_{x-r} = j^a), \qquad (12)
\end{aligned}$$

where in the last equality we used that $\{\gamma > x, \varphi^\lambda_x = j^a\} = \varnothing$ and the Markov property of $\{\varphi^\lambda_x\}_{x\geq 0}$. Note that all the elements in (12) correspond to transition probabilities or intensities that can be expressed in matricial form as follows:

$$\mathbb{P}_{i^o}(\gamma > x, \varphi^\lambda_x = j^o) = \boldsymbol{e}_i^\intercal e^{(T^+ - \lambda I)x}\boldsymbol{e}_j,$$
$$\mathbb{P}_{i^o}(\gamma \in [r, r + \mathrm{d}r], \varphi^\lambda_\gamma = k^a) = \boldsymbol{e}_i^\intercal e^{(T^+ - \lambda I)r}T^-\boldsymbol{e}_k\mathrm{d}r,$$
$$\mathbb{P}_{k^a}(\tau > x - r, \varphi^\lambda_{x-r} = j^o) = \boldsymbol{e}_k^\intercal \Phi_{ao}(x - r)\boldsymbol{e}_j,$$
$$\mathbb{P}_{k^a}(\tau > x - r, \varphi^\lambda_{x-r} = j^a) = \boldsymbol{e}_k^\intercal \Phi_{aa}(x - r)\boldsymbol{e}_j.$$

Substituting these expressions into (12) and using the identity $I=\sum_{k=1}^{p}\boldsymbol{e}_k\boldsymbol{e}_k^{\mathsf{T}}$ gives

$$\begin{aligned}\boldsymbol{e}_i^{\mathsf{T}}\Phi_o(x)\boldsymbol{e}_j &= \boldsymbol{e}_i^{\mathsf{T}}e^{(T^+-\lambda I)x}\boldsymbol{e}_j + \int_0^x \sum_{k=1}^{p}\left(\boldsymbol{e}_i^{\mathsf{T}}e^{(T^+-\lambda I)r}T^-\boldsymbol{e}_k\right)\left(\boldsymbol{e}_k^{\mathsf{T}}\Phi_{ao}(x-r)\boldsymbol{e}_j\right)\mathrm{d}r\\ &\quad -\int_0^x \sum_{k=1}^{p}\left(\boldsymbol{e}_i^{\mathsf{T}}e^{(T^+-\lambda I)r}T^-\boldsymbol{e}_k\right)\left(\boldsymbol{e}_k^{\mathsf{T}}\Phi_{aa}(x-r)\boldsymbol{e}_j\right)\mathrm{d}r\\ &= \boldsymbol{e}_i^{\mathsf{T}}\left(e^{(T^+-\lambda I)x}+\int_0^x e^{(T^+-\lambda I)r}T^-\left[\Phi_{ao}(x-r)-\Phi_{aa}(x-r)\right]\mathrm{d}r\right)\boldsymbol{e}_j\\ &= \boldsymbol{e}_i^{\mathsf{T}}\left(e^{(T^+-\lambda I)x}+\int_0^x e^{(T^+-\lambda I)r}T^-\left[\Phi_{oa}(x-r)-\Phi_{oo}(x-r)\right]\mathrm{d}r\right)\boldsymbol{e}_j\\ &= \boldsymbol{e}_i^{\mathsf{T}}\left(e^{(T^+-\lambda I)x}+\int_0^x e^{(T^+-\lambda I)r}(-T^-)\Phi_o(x-r)\mathrm{d}r\right)\boldsymbol{e}_j,\end{aligned}$$

so that $\{\Phi_o(x)\}_{x\geq 0}$ is the bounded solution to the matrix-integral equation

$$\Phi_o(x)=e^{(T^+-\lambda I)x}+\int_0^x e^{(T^+-\lambda I)r}(-T^-)\Phi_o(x-r)\mathrm{d}r.$$

By [3, Theorem 3.10],

$$\Phi_o(x)=e^{[(T^+-\lambda I)+(-T^-)]x}=e^{(T-\lambda I)x}.$$

The Markov property implies that

$$\begin{aligned}&\mathbb{P}_{i^o}(\tau\in[x,x+\mathrm{d}x],\,\varphi_\tau^\lambda=\Delta^o)\\ &=\sum_{k=1}^{p}\mathbb{P}_{i^o}(\tau>x,\,\varphi_x^\lambda=k^o)\mathbb{P}_{k^o}(\tau\in[x,x+\mathrm{d}x],\,\varphi_\tau^\lambda=\Delta^o)\\ &\quad+\sum_{k=1}^{p}\mathbb{P}_{i^o}(\tau>x,\,\varphi_x^\lambda=k^a)\mathbb{P}_{k^a}(\tau\in[x,x+\mathrm{d}x],\,\varphi_\tau^\lambda=\Delta^o)\\ &=\sum_{k=1}^{p}(\boldsymbol{e}_i^{\mathsf{T}}\Phi_{oo}(x)\boldsymbol{e}_k)(\boldsymbol{e}_k^{\mathsf{T}}\boldsymbol{s}^+)\mathrm{d}x+\sum_{k=1}^{p}(\boldsymbol{e}_i^{\mathsf{T}}\Phi_{oa}(x)\boldsymbol{e}_k)(\boldsymbol{e}_k^{\mathsf{T}}\boldsymbol{s}^-\mathrm{d}x)\\ &=\boldsymbol{e}_i^{\mathsf{T}}(\Phi_{oo}(x)\boldsymbol{s}^++\Phi_{oa}(x)\boldsymbol{s}^-)\mathrm{d}x.\end{aligned}$$

Similarly,

$$\mathbb{P}_{i^o}(\tau\in[x,x+\mathrm{d}x],\,\varphi_\tau^\lambda=\Delta^a)=\boldsymbol{e}_i^{\mathsf{T}}(\Phi_{oa}(x)\boldsymbol{s}^++\Phi_{oo}(x)\boldsymbol{s}^-)\mathrm{d}x.$$

Thus,

$$\begin{aligned}&\mathbb{P}_{i^o}(\tau\in[x,x+\mathrm{d}x],\,\varphi_\tau^\lambda=\Delta^o)-\mathbb{P}_{i^o}(\tau\in[x,x+\mathrm{d}x],\,\varphi_\tau^\lambda=\Delta^a)\\ &=\boldsymbol{e}_i^{\mathsf{T}}([\Phi_{oo}(x)\boldsymbol{s}^++\Phi_{oa}(x)\boldsymbol{s}^-]-[\Phi_{oa}(x)\boldsymbol{s}^++\Phi_{oo}(x)\boldsymbol{s}^-])\mathrm{d}x\\ &=\boldsymbol{e}_i^{\mathsf{T}}([\Phi_{oo}(x)-\Phi_{oa}(x)]\boldsymbol{s}^+-[\Phi_{oo}(x)-\Phi_{oa}(x)]\boldsymbol{s}^-)\mathrm{d}x\\ &=\boldsymbol{e}_i^{\mathsf{T}}\Phi_o(x)\boldsymbol{s}\mathrm{d}x=\boldsymbol{e}_i^{\mathsf{T}}e^{(T-\lambda I)x}\boldsymbol{s}\mathrm{d}x,\end{aligned}$$

https://doi.org/10.1017/jpr.2022.25 Published online by Cambridge University Press

so that (8) holds. Analogous arguments follow for (10) and (11), which completes the proof. □

Heuristically, Equations (8) and (10) imply that initiating $\{\varphi_t^\lambda\}_{t\geq 0}$ in the anti-state i^a has the opposite effect, in terms of sign, to initiating in the original state i^o. In the following we exploit this fact to provide a probabilistic interpretation not only for the elements of $e^{(T-\lambda I)x}\boldsymbol{s}$, but also for the exponentially tilted matrix-exponential density $\boldsymbol{\alpha} e^{(T-\lambda I)x}\boldsymbol{s}$.

Define w^+ and w^- by

$$w^\pm = \sum_{i=1}^{p} \max\{0, \pm\alpha_i\},$$

and define $\boldsymbol{\alpha}^+ = (\alpha_1^+, \ldots, \alpha_p^+)$ and $\boldsymbol{\alpha}^- = (\alpha_1^-, \ldots, \alpha_p^-)$ by

$$\alpha_i^\pm = \begin{cases} \frac{1}{w^\pm}\max\{0, \pm\alpha_i\} & \text{if} \quad w^\pm > 0, \\ 0 & \text{if} \quad w^\pm = 0. \end{cases}$$

If $w^\pm > 0$, then $\boldsymbol{\alpha}^\pm$ is a probability vector, and in general,

$$\boldsymbol{\alpha} = w^+\boldsymbol{\alpha}^+ - w^-\boldsymbol{\alpha}^-. \tag{13}$$

In some sense, $(w^+ + w^-)^{-1}\boldsymbol{\alpha}$ can be thought as a mixture of the probability vectors $\boldsymbol{\alpha}^+$ and $\boldsymbol{\alpha}^-$, with the latter contributing 'negative mass'. Fortunately, this 'negative mass' in the context of $\boldsymbol{\alpha} e^{(T-\lambda I)x}\boldsymbol{s}$ can be given a precise probabilistic interpretation by means of anti-states as follows.

Theorem 3.2. *Let* $f_\lambda(x) = (\boldsymbol{\alpha}(\lambda I - T)^{-1}\boldsymbol{s})^{-1}\boldsymbol{\alpha} e^{(T-\lambda I)x}\boldsymbol{s}$, $x \geq 0$, *be the density of the exponentially tilted matrix-exponential distribution of parameters* $(\boldsymbol{\alpha}, T, \boldsymbol{s})$. *Define the vectors*

$$\widehat{\boldsymbol{\alpha}}^+ := \tfrac{w^+}{w^+ + w^-}\boldsymbol{\alpha}^+ \quad \text{and} \quad \widehat{\boldsymbol{\alpha}}^- := \tfrac{w^-}{w^+ + w^-}\boldsymbol{\alpha}^-,$$

and suppose $\varphi_0^\lambda \sim (\widehat{\boldsymbol{\alpha}}^+, \widehat{\boldsymbol{\alpha}}^-)$. *Then*

$$f_\lambda(x)\mathrm{d}x = \frac{(w^+ + w^-)}{\boldsymbol{\alpha}(\lambda I - T)^{-1}\boldsymbol{s}}\mathbb{E}\left[\mathbb{1}\{\tau \in [x, x+\mathrm{d}x]\}\beta(\varphi_\tau^\lambda)\right] \tag{14}$$

$$= \frac{(w^+ + w^-)}{\boldsymbol{\alpha}(\lambda I - T)^{-1}\boldsymbol{s}}\left((\widehat{\boldsymbol{\alpha}}^+, \widehat{\boldsymbol{\alpha}}^-)\exp\left(\begin{bmatrix} T^+ - \lambda I & T^- \\ T^- & T^+ - \lambda I \end{bmatrix} x\right)\begin{bmatrix} \boldsymbol{s} \\ -\boldsymbol{s} \end{bmatrix}\right)\mathrm{d}x, \tag{15}$$

where τ *and* $\beta(\cdot)$ *are defined as in Theorem 3.1.*

Proof. Equation (13) implies that

$$f_\lambda(x) = \frac{1}{\boldsymbol{\alpha}(\lambda I - T)^{-1}\boldsymbol{s}}\left(\sum_{i=1}^{p} w^+\alpha_i^+\left(\boldsymbol{e}_i^\intercal e^{(T-\lambda I)x}\boldsymbol{s}\right) + \sum_{i=1}^{p} w^-\alpha_i^-\left(-\boldsymbol{e}_i^\intercal e^{(T-\lambda I)x}\boldsymbol{s}\right)\right)$$

$$= \frac{(w^+ + w^-)}{\boldsymbol{\alpha}(\lambda I - T)^{-1}\boldsymbol{s}}\left(\sum_{i=1}^{p} \tfrac{w^+}{w^+ + w^-}\alpha_i^+\left(\boldsymbol{e}_i^\intercal e^{(T-\lambda I)x}\boldsymbol{s}\right) + \sum_{i=1}^{p} \tfrac{w^-}{w^+ + w^-}\alpha_i^-\left(-\boldsymbol{e}_i^\intercal e^{(T-\lambda I)x}\boldsymbol{s}\right)\right). \tag{16}$$

https://doi.org/10.1017/jpr.2022.25 Published online by Cambridge University Press

Equality (14) follows from (16), (8), and (10). Equality (15) follows from (16), (9), and (11). □

Example 3.1. Let $(\boldsymbol{\alpha}, T, \boldsymbol{s})$ be the matrix-exponential parameters corresponding to (5). As noted previously, these parameters by themselves lack a probabilistic interpretation, so we apply Theorem 3.1 to construct one. For such parameters we take the tilting parameter $\lambda := \lambda_0 = 2$, leading to the block-partitioned matrices

$$\begin{bmatrix} T^+ - \lambda I & T^- \\ T^- & T^+ - \lambda I \end{bmatrix} = \left[\begin{array}{ccc|ccc} -3 & 0 & 2/3 & 0 & 1 & 0 \\ 1 & -3 & 0 & 0 & 0 & 2/3 \\ 0 & 0 & -3 & 0 & 0 & 0 \\ \hline 0 & 1 & 0 & -3 & 0 & 2/3 \\ 0 & 0 & 2/3 & 1 & -3 & 0 \\ 0 & 0 & 0 & 0 & 0 & -3 \end{array}\right],$$

$$(\widehat{\boldsymbol{\alpha}}^+, \widehat{\boldsymbol{\alpha}}^-) = (1, 0, 0, 0, 0, 0), \qquad \begin{bmatrix} \boldsymbol{s} \\ -\boldsymbol{s} \end{bmatrix} = \left[\begin{array}{c} 4/3 \\ 2/3 \\ 1 \\ \hline -4/3 \\ -2/3 \\ -1 \end{array}\right],$$

and $w^+ = 1$, $w^- = 0$. We can then verify that

$$\frac{(w^+ + w^-)}{\boldsymbol{\alpha}(\lambda I - T)^{-1}\boldsymbol{s}} \left((\widehat{\boldsymbol{\alpha}}^+, \widehat{\boldsymbol{\alpha}}^-) \exp\left(\begin{bmatrix} T^+ - \lambda I & T^- \\ T^- & T^+ - \lambda I \end{bmatrix} x \right) \begin{bmatrix} \boldsymbol{s} \\ -\boldsymbol{s} \end{bmatrix} \right)$$
$$= \frac{1}{\boldsymbol{\alpha}(\lambda I - T)^{-1}\boldsymbol{s}} \left(\frac{2}{3} e^{-3x}(1 + \cos(x)) \right) = \frac{e^{-2x}}{\boldsymbol{\alpha}(\lambda I - T)^{-1}\boldsymbol{s}} \left(\frac{2}{3} e^{-x}(1 + \cos(x)) \right),$$

the latter corresponding to the exponentially tilted matrix-exponential density function $f(x) = \frac{2}{3} e^{-x}(1 + \cos(x))$.

A probabilistic interpretation of f_λ alternative to that of (14) is the following.

Corollary 3.1. *Define* $\boldsymbol{d} = (d_1, \ldots, d_p)^\intercal := -(T^+ - \lambda I)\mathbf{1} - T^-\mathbf{1}$ *to be the termination intensities vector from* $\mathcal{E}^o$ *or* $\mathcal{E}^a$, *and define* $\boldsymbol{q}^\pm = (q_1^\pm, \ldots, q_p^\pm)^\intercal$ *by*

$$q_i^\pm = \begin{cases} \frac{s_i^\pm}{d_i} & \textit{if} \quad d_i > 0, \\ 0 & \textit{if} \quad d_i = 0. \end{cases}$$

Let $\bar{q} : \mathcal{E}^o \cup \mathcal{E}^a \mapsto \mathbb{R}$ *be defined by*

$$\bar{q}(i^o) = q_i^+ - q_i^- \quad \textit{and} \quad \bar{q}(i^a) = q_i^- - q_i^+ \quad \text{for} \quad i \in \{1, \ldots, p\}.$$

https://doi.org/10.1017/jpr.2022.25 Published online by Cambridge University Press

Then

$$f_\lambda(x)\mathrm{d}x = \left(\frac{w^+ + w^-}{\boldsymbol{\alpha}(\lambda I - T)^{-1}\boldsymbol{s}}\right)\mathbb{E}\left[\mathbb{1}\{\tau \in [x, x+\mathrm{d}x]\}\bar{q}\big(\varphi^\lambda_{\tau-}\big)\right], \tag{17}$$

where $\{\varphi^\lambda_t\}_{t\geq 0}$ *and* τ *are defined as in Theorem* 3.2.

Proof. First, notice that the jump mechanism of $\{\varphi^\lambda_t\}_{t\geq 0}$ described in (6) implies that for $i \in \{1, \ldots, p\}$,

$$\mathbb{P}(\varphi^\lambda_\tau = \Delta^o \mid \tau, \varphi^\lambda_{\tau-} = i^o) = q^+_i,$$
$$\mathbb{P}(\varphi^\lambda_\tau = \Delta^a \mid \tau, \varphi^\lambda_{\tau-} = i^o) = q^-_i,$$
$$\mathbb{P}(\varphi^\lambda_\tau = \Delta^o \mid \tau, \varphi^\lambda_{\tau-} = i^a) = q^-_i,$$
$$\mathbb{P}(\varphi^\lambda_\tau = \Delta^a \mid \tau, \varphi^\lambda_{\tau-} = i^a) = q^+_i,$$

which in turn implies that

$$\mathbb{E}\big[\beta(\varphi^\lambda_\tau) \mid \tau, \varphi^\lambda_{\tau-}\big] = \mathbb{P}\big(\varphi^\lambda_\tau = \Delta^o \mid \tau, \varphi^\lambda_{\tau-}\big) - \mathbb{P}\big(\varphi^\lambda_\tau = \Delta^a \mid \tau, \varphi^\lambda_{\tau-}\big) = \bar{q}(\varphi^\lambda_{\tau-}).$$

Consequently,

$$\begin{aligned}\mathbb{E}\big[\mathbb{1}\{\tau \in [x, x+\mathrm{d}x]\}\beta(\varphi^\lambda_\tau)\big] &= \mathbb{E}\big[\mathbb{E}\left[\mathbb{1}\{\tau \in [x, x+\mathrm{d}x]\}\beta(\varphi^\lambda_\tau) \mid \tau, \varphi^\lambda_{\tau-}\right]\big]\\ &= \mathbb{E}\big[\mathbb{1}\{\tau \in [x, x+\mathrm{d}x]\}\mathbb{E}\big[\beta(\varphi^\lambda_\tau) \mid \tau, \varphi^\lambda_{\tau-}\big]\big]\\ &= \mathbb{E}\big[\mathbb{1}\{\tau \in [x, x+\mathrm{d}x]\}\bar{q}\big(\varphi^\lambda_{\tau-}\big)\big],\end{aligned}$$

and the result follows from (14). □

Though closely related, the interpretation provided by Corollary 3.1 is more suitable than that of Theorem 3.2 for Monte Carlo applications. Indeed, a realization of $\{\varphi^\lambda_t\}_{t\geq 0}$ may get absorbed in Δ^o, Δ^a or terminated. If termination occurs, such a realization contributes nothing to the term in the right-hand side of (14). In contrast, by observing the process until its exit time from $\mathcal{E}^o \cup \mathcal{E}^a$ and disregarding its landing point as in Corollary 3.1, we make sure that each realization contributes towards the mass in the right-hand side of (17).

4. Recovering the untilted distribution

Once the exponentially tilted density f_λ of a matrix-exponential distribution of parameters $(\boldsymbol{\alpha}, T, \boldsymbol{s})$ has a tractable known form, say as in (15), in principle it is straightforward to recover the original untilted density f by taking

$$\begin{aligned}f(x) &= (\boldsymbol{\alpha}(\lambda I - T)\boldsymbol{s})e^{\lambda x}f_\lambda(x)\\ &= (w^+ + w^-)(\widehat{\boldsymbol{\alpha}}^+, \widehat{\boldsymbol{\alpha}}^-)\exp\left(\begin{bmatrix} T^+ & T^- \\ T^- & T^+\end{bmatrix}x\right)\begin{bmatrix}\boldsymbol{s} \\ -\boldsymbol{s}\end{bmatrix}, \qquad x \geq 0.\end{aligned} \tag{18}$$

While (18) is a legitimate matrix-exponential representation of f, it has two drawbacks:

1. The matrix $\begin{bmatrix} T^+ & T^- \\ T^- & T^+ \end{bmatrix}$ may no longer be a subintensity matrix.
2. The dominant eigenvalue of $\begin{bmatrix} T^+ & T^- \\ T^- & T^+ \end{bmatrix}$ may be nonnegative.

The first item may affect the probabilistic interpretation of f, while the second one may make integration of certain functions (with respect to the density f) more difficult to handle. For instance, in the context of Example 3.1, the matrix

$$\begin{bmatrix} T^+ & T^- \\ T^- & T^+ \end{bmatrix} = \left[\begin{array}{ccc|ccc} -1 & 0 & 2/3 & 0 & 1 & 0 \\ 1 & -1 & 0 & 0 & 0 & 2/3 \\ 0 & 0 & -1 & 0 & 0 & 0 \\ \hline 0 & 1 & 0 & -1 & 0 & 2/3 \\ 0 & 0 & 2/3 & 1 & -1 & 0 \\ 0 & 0 & 0 & 0 & 0 & -1 \end{array}\right]$$

is not a subintensity matrix since some row sums are strictly positive, and it has 0 as its dominant eigenvalue. Having 0 as an eigenvalue implies that some entries of $\exp\left(\begin{pmatrix} T^+ & T^- \\ T^- & T^+ \end{pmatrix} x\right)$ may potentially be of order $e^{0\cdot x} = 1$, meaning that the matrix integral

$$\int_0^\infty h(x) \exp\left(\begin{bmatrix} T^+ & T^- \\ T^- & T^+ \end{bmatrix} x\right) \mathrm{d}x \tag{19}$$

may only be well-defined for functions $h : \mathbb{R}_+ \mapsto \mathbb{R}_+$ that decrease to 0 fast enough. In comparison, $\exp(Tx)$ with T as in (5) has entries of order $e^{\sigma_0 x} = e^{-x}$ or less, so that

$$\int_0^\infty h(x) \exp(Tx)\, \mathrm{d}x \tag{20}$$

is well-defined and finite for every function $h : \mathbb{R}_+ \mapsto \mathbb{R}_+$ of order $e^{-\rho x}$ for any $\rho > -1$. This apparent disagreement between the applicability of (19) and (20) vanishes when we multiply $\exp\left(\begin{bmatrix} T^+ & T^- \\ T^- & T^+ \end{bmatrix} x\right)$ by $\begin{bmatrix} \boldsymbol{s} \\ -\boldsymbol{s} \end{bmatrix}$. Indeed, in the context of Example 3.1 it can be verified that the elements of the vector $\exp\left(\begin{bmatrix} T^+ & T^- \\ T^- & T^+ \end{bmatrix} x\right) \begin{bmatrix} \boldsymbol{s} \\ -\boldsymbol{s} \end{bmatrix}$ are at most of order e^{-x}, with the higher-order terms of $\exp\left(\begin{bmatrix} T^+ & T^- \\ T^- & T^+ \end{bmatrix} x\right)$ cancelling each other when we multiply the matrix function by $\begin{bmatrix} \boldsymbol{s} \\ -\boldsymbol{s} \end{bmatrix}$.

In the general case, this cancellation of higher-order terms is implied by Theorem 3.1 via the following arguments. If σ_0 is the dominant eigenvalue of T and has multiplicity m_0, then the order of $\boldsymbol{e}_i^\intercal e^{(T-\lambda I)x}\boldsymbol{s}$ is at most $x^{m_0} e^{(\sigma_0-\lambda)x}$. By (9) and (11), the elements of $\exp\left(\begin{bmatrix} T^+ - \lambda I & T^- \\ T^- & T^+ - \lambda I \end{bmatrix} x\right) \begin{bmatrix} \boldsymbol{s} \\ -\boldsymbol{s} \end{bmatrix}$ are also of order less than or equal to $x^{m_0} e^{(\sigma_0-\lambda)x}$. Finally, if we multiply the previous by $e^{\lambda x}$, then we get that $\exp\left(\begin{bmatrix} T^+ & T^- \\ T^- & T^+ \end{bmatrix} x\right) \begin{bmatrix} \boldsymbol{s} \\ -\boldsymbol{s} \end{bmatrix}$ is of order at most $x^{m_0} e^{\sigma_0 x}$, which coincides with the order of e^{Tx}.

In terms of expectations, (14) and (17) provide alternative ways to recover properties of any matrix-exponential density f of parameters $(\boldsymbol{\alpha}, T, \boldsymbol{s})$ in terms of the exponentially tilted

https://doi.org/10.1017/jpr.2022.25 Published online by Cambridge University Press

density f_λ. Indeed, for any function $h:\mathbb{R}_+\mapsto\mathbb{R}_+$ of order $e^{-\rho x}$ or less, $\rho>\sigma_0$, we have that

$$\int_0^\infty h(x)f(x)\mathrm{d}x=(\boldsymbol{\alpha}(\lambda I-T)\boldsymbol{s})\int_0^\infty h(x)e^{\lambda x}f_\lambda(x)\mathrm{d}x \tag{21}$$

$$=(w^+ + w^-)\mathbb{E}\left[h(\tau)e^{\lambda\tau}\beta(\varphi^\lambda_\tau)\right] \tag{22}$$

$$=(w^+ + w^-)\mathbb{E}\left[h(\tau)e^{\lambda\tau}\bar{q}(\varphi^\lambda_{\tau-})\right], \tag{23}$$

where $\{\varphi^\lambda_t\}_{t\geq 0}$ and τ are as in Theorem 3.2. Existence and finiteness of the first moment of $h(\tau)e^{\lambda\tau}\beta(\varphi^\lambda_\tau)$ in (22) is guaranteed by noting that the order $e^{-\rho+\lambda}$ of $h(x)e^{\lambda x}$ is dominated by that of f_λ. Notice that, as opposed to the formula in (18), the representations (22) and (23) still have probabilistic interpretations in terms of the Markov jump process $\{\varphi^\lambda_t\}_{t\geq 0}$.

Acknowledgements

I wish to thank the editor and anonymous reviewer for their constructive comments, which helped me to improve the manuscript.

Funding information

The author acknowledges funding from the Australian Research Council Discovery Project DP180103106 and the Swiss National Science Foundation Project 200021.191984.

Competing interests

There were no competing interests to declare which arose during the preparation or publication process for this article.

References

[1] Asmussen, S. and Bladt, M. (1996). Renewal theory and queueing algorithms for matrix-exponential distributions. In *Matrix-Analytic Methods in Stochastic Models*, CRC Press, Boca Raton, pp. 313–341.
[2] Asmussen, S. and Bladt, M. (1999). Point processes with finite-dimensional conditional probabilities. *Stoch. Process. Appl.* **82**, 127–142.
[3] Bean, N. G., Nguyen, G. T., Nielsen, B. F. and Peralta, O. (2022). RAP-modulated fluid processes: first passages and the stationary distribution. *Stoch. Process. Appl.* **149**, 308–340.
[4] Bean, N. G. and Nielsen, B. F. (2010). Quasi-birth-and-death processes with rational arrival process components. *Stoch. Models* **26**, 309–334.
[5] Bladt, M. and Neuts, M. F. (2003). Matrix-exponential distributions: calculus and interpretations via flows. *Stoch. Models* **19**, 113–124.
[6] Bladt, M. and Nielsen, B. F. (2017). *Matrix-Exponential Distributions in Applied Probability*. Springer, New York.
[7] Commault, C. and Mocanu, S. (2003). Phase-type distributions and representations: some results and open problems for system theory. *Internat. J. Control* **76**, 566–580.
[8] Cox, D. R. (1955). A use of complex probabilities in the theory of stochastic processes. *Math. Proc. Camb. Phil. Soc*. **51**, 313–319.
[9] Cox, D. R. (1955). The analysis of non-Markovian stochastic processes by the inclusion of supplementary variables. *Math. Proc. Camb. Phil. Soc*. **51**, 433–441.
[10] Cramér, H. (2013). Collective risk theory: a survey of the theory from the point of view of the theory of stochastic processes. In *Collected Works II*, Springer, Berlin, Heidelberg, pp. 1028–1115.
[11] Escher, F. (1932). On the probability function in the collective theory of risk. *Scand. Actuarial J.* **15**, 175–195.
[12] Farina, L. and Rinaldi, S. (2000). *Positive Linear Systems: Theory and Applications*. John Wiley, New York.

https://doi.org/10.1017/jpr.2022.25 Published online by Cambridge University Press

[13] GERBER, H. AND SHIU, E. (1994). Option pricing by Esscher transforms. *Trans. Soc. Actuaries* **46**, 99–140.
[14] JENSEN, A. (1954). *A Distribution Model Applicable to Economics*. Munksgaard, Copenhagen.
[15] NEUTS, M. (1975). Probability distributions of phase type. In *Liber Amicorum Prof. Emeritus H. Florin*, Department of Mathematics, University of Louvain, pp. 173–206.
[16] NEUTS, M. (1981). *Matrix-Geometric Solutions in Stochastic Models: an Algorithmic Approach*. Johns Hopkins University Press, Baltimore.
[17] O'CINNEIDE, C. A. (1990). Characterization of phase-type distributions. *Stoch. Models* **6**, 1–57.
[18] VÖLLERING, F. (2020). Markov process representation of semigroups whose generators include negative rates. *Electron. Commun. Prob.* **25**, 7 pp.