



UvA-DARE (Digital Academic Repository)

Extreme Value Analysis for a Markov Additive Process Driven by a Nonirreducible Background Chain

van Kreveld, L.; Mandjes, M.; Dorsman, J.L.

DOI

[10.1287/stsy.2021.0086](https://doi.org/10.1287/stsy.2021.0086)

Publication date

2022

Document Version

Final published version

Published in

Stochastic Systems

License

CC BY

[Link to publication](#)

Citation for published version (APA):

van Kreveld, L., Mandjes, M., & Dorsman, J. L. (2022). Extreme Value Analysis for a Markov Additive Process Driven by a Nonirreducible Background Chain. *Stochastic Systems*, 12(3), 293-317. <https://doi.org/10.1287/stsy.2021.0086>

General rights

It is not permitted to download or to forward/distribute the text or part of it without the consent of the author(s) and/or copyright holder(s), other than for strictly personal, individual use, unless the work is under an open content license (like Creative Commons).

Disclaimer/Complaints regulations

If you believe that digital publication of certain material infringes any of your rights or (privacy) interests, please let the Library know, stating your reasons. In case of a legitimate complaint, the Library will make the material inaccessible and/or remove it from the website. Please Ask the Library: <https://uba.uva.nl/en/contact>, or a letter to: Library of the University of Amsterdam, Secretariat, Singel 425, 1012 WP Amsterdam, The Netherlands. You will be contacted as soon as possible.

UvA-DARE is a service provided by the library of the University of Amsterdam (<https://dare.uva.nl>)

Extreme Value Analysis for a Markov Additive Process Driven by a Nonirreducible Background Chain

Lucas van Kreveld,^a Michel Mandjes,^{a,b,c} Jan-Pieter Dorsman^a

^aKorteweg-de Vries Institute for Mathematics, University of Amsterdam, 1098 XH Amsterdam, Netherlands; ^bEurandom, Eindhoven University of Technology, 5612 AZ Eindhoven, Netherlands; ^cAmsterdam Business School, Faculty of Economics and Business, University of Amsterdam, 1098 XH Amsterdam, Netherlands

Contact: l.r.vankreveld@uva.nl,  <https://orcid.org/0000-0001-9588-6083> (LvK); m.r.h.mandjes@uva.nl (MM); j.l.dorsman@uva.nl (J-PD)

Received: June 18, 2021

Revised: October 6, 2021


Accepted: October 13, 2021

Published Online in *Articles in Advance*:
January 26, 2022

<https://doi.org/10.1287/stsy.2021.0086>

Copyright: © 2022 The Author(s)

Abstract. A common assumption in the vast literature on the extremes of spectrally one-sided Markov additive processes (MAPs) is that the continuous-time Markov chain that serves as the background process is irreducible. In the present paper, we consider, motivated by, for example, applications in credit risk, the case in which the irreducibility condition has been lifted, thus allowing the presence of one or more transient classes. More specifically, we consider the distribution of the maximum when the MAP under study has only positive jumps (the spectrally positive case) or negative jumps (the spectrally negative case). The methodology used relies on two crucial previous results: (i) the Wiener–Hopf decomposition for Lévy processes and, in particular, its explicit form in spectrally one-sided cases and (ii) a result on the number of singularities of the matrix exponent of a spectrally one-sided MAP. In both the spectrally positive and negative cases, we derive a system of linear equations of which the solution characterizes the distribution of the maximum of the process. As a by-product of our results, we develop a procedure for calculating the maximum of a spectrally one-sided Lévy process over a phase-type distributed time interval.

 **Open Access Statement:** This work is licensed under a Creative Commons Attribution 4.0 International License. You are free to copy, distribute, transmit and adapt this work, but you must attribute this work as “*Stochastic Systems*. Copyright © 2022 The Author(s). <https://doi.org/10.1287/stsy.2021.0086>, used under a Creative Commons Attribution License: <https://creativecommons.org/licenses/by/4.0/>.”

Funding: This research is funded by the NWO Gravitation project Networks [Grant 024.002.003].

Keywords: Markov additive processes • fluctuation theory • background process

1. Introduction

The *Markov additive process* (in this paper abbreviated to MAP) can be seen as the Markov-modulated version of the Lévy process. Indeed, when an independently evolving continuous-time Markov chain on $d \in \mathbb{N}$ states, usually referred to as the *background process*, is in state i , the MAP locally behaves as a Lévy process $X_i(\cdot)$. Additionally, a MAP allows for jumps at transition epochs of the background process. As such, MAPs offer a natural modeling framework to study stochastic processes of which the dynamics change over time, with broad applications in, for example, credit and risk theory, queueing, inventory management, and finance; early references on MAPs include Çinlar (1972) and Neveu (1961).

A key object of study concerns the *extreme values* attained by the MAP over a finite or infinite horizon. With $Y(\cdot)$ denoting the MAP under consideration, the focus is on the analysis of the distribution of its *running maximum* process $\bar{Y}(t) := \sup_{s \in [0,t]} Y(s)$ (as well as the corresponding running minimum process). Besides being interesting in its own right, the running maximum process can be directly translated in terms of the first-passage process $\tau(y) := \inf\{s \geq 0 : Y(s) > y\}$ because of the known duality relation between the events $\{\bar{Y}(t) > y\}$ and $\{\tau(y) < t\}$. Building upon related results for Lévy processes, a wide range of characterizations is derived, typically in terms of transforms or so-called scale functions. We refer to Ivanovs (2011), chapter II, for an extensive account of the main results on extremes of MAPs as well as the corresponding first-passage process. Particularly noteworthy are the results obtained by Asmussen and Kella (2000), who use martingale methods to effectively extend the Pollaczek–Khinchine formula for spectrally one-sided Lévy processes to the MAP setting. We, in addition, mention the work by Dieker and Mandjes (2011) as well as D’Auria et al. (2010), the latter being predominantly in terms of the first-passage process.

In the literature on extremes for MAPs, one typically assumes that the background process is *irreducible*. This assumption is convenient as it guarantees existence and uniqueness of its invariant distribution and allows the use of time-reversal arguments. On the other hand, the assumed irreducibility obviously restricts the general applicability of the model. There may be events, such as a crash of a stock market, that cause permanent changes in the fluctuations of the price levels. The resulting process can naturally be modeled using a nonirreducible background process. Second, in the context of credit risk, one could think of companies paying interest to an obligor until they go into default, thus causing a loss to the obligor, after which they effectively leave the system—another setting that can be modeled using a nonirreducible background process. This credit-related example motivated Delsing and Mandjes (2021) to consider the extreme values attained by a MAP of a specific Cramér–Lundberg type endowed with a nonirreducible background process of a specific structure. Importantly, however, the area of extremes of more general MAPs under a nonirreducible background process is so far largely unexplored.

The main objective of this paper is to extend existing results for extremes of MAPs to the case of an arbitrary Markovian background process, thus covering the situation that the background process is nonirreducible. Concretely, we describe the distribution of the maximum attained by the MAP until the background process reaches an absorbing state (which covers the running maximum over an exponentially distributed horizon and, in particular, also the all-time maximum). With i representing the initial state of the background chain, we let this maximum be denoted by Z_i . The main results in this paper provide the distribution of Z_i for all i for a spectrally one-sided MAP (i.e., the direction of all jumps is either positive or negative). As a by-product of this result, we also succeed in deriving the distribution of the maximum of a spectrally one-sided Lévy process over a phase-type distributed time interval.

The way in which our results are obtained is remarkably straightforward and transparent. Our approach is based on two key results. First, the Wiener–Hopf decomposition for spectrally one-sided Lévy processes (Kyprianou 2006, Debicki and Mandjes 2015) allows us to describe the dynamics of the MAP between two successive transition epochs of the background process. Standard analytic techniques are then used to transform resulting expressions into a linear system of equations from which the distribution of the MAP’s maximum follows. The second key result, as established by Ivanovs et al. (2010), characterizes the number of singularities with positive real parts of the matrix exponent corresponding to a spectrally one-sided MAP. Using this result, we find a procedure of obtaining the solution to the system of equations, consequently determining the distribution of Z_i for all i . Notably, the approach does not rely on the use of martingale methods.

This paper is structured as follows. In Section 2, we describe our model, introduce the necessary notation, and state the two important previous results mentioned earlier. For the spectrally positive case (no downward jumps, that is), as treated in Section 3, we derive a system of linear equations for the transforms of Z_i . To solve these equations, we introduce an ordering on the communicating classes of the background chain, which allows us to recursively determine the distributions of Z_i for all initial states i . In Section 4, covering the spectrally negative case (no upward jumps, that is), we derive a similar system of equations for the complementary distribution function of Z_i . By using the fact that the solution is of a specific form, we show how to solve this system. Whereas in most of the literature on extreme values of Lévy processes, one considers extreme values over an exponentially distributed interval, in Section 5, it is pointed out how we can use our MAP-based framework to extend such results to extreme values of Lévy processes over a phase-type time interval. Section 6 presents a series of numerical experiments illustrating the use of our results and some implementation guidelines. Concluding remarks are made in Section 7.

2. Model and Preliminaries

In this section, we first describe in detail the model that we consider in this paper and introduce the corresponding notation. We then discuss the Wiener–Hopf decomposition for spectrally one-sided Lévy processes and the result on the number of singularities of a spectrally one-sided MAP, that is, the results (i) and (ii) mentioned in the introduction, which play a crucial role in our reasoning. We conclude this section by explicitly outlining the approach that is followed.

2.1. Model

We start by giving the formal definition of a MAP. Let the background process $(J(t))_{t \geq 0}$ be a continuous-time Markov chain with $d \in \mathbb{N}$ states. Associated with every state i , let $(X_i(t))_{t \geq 0}$ be a Lévy process and assume that these Lévy processes are mutually independent. With t_n denoting the time of the n th transition of the background process, suppose that $t \in [t_n, t_{n+1})$ and that, at t_n , the background chain jumped from state i to state j .

Then, a MAP is a stochastic process $(Y(t))_{t \geq 0}$ for which

$$Y(t) = Y(t_n-) + L_{ij}^n + (X_j(t) - X_j(t_n)),$$

where $(L_{ij}^n)_{n \in \mathbb{N}}$ is a sequence of independent copies of the random variable L_{ij} , representing the size of the jump at the time of a transition from background state i to background state j (where $i \neq j$). Because jumps at self-transitions, say from background state i to itself, can be incorporated in the Lévy process $X_i(\cdot)$, we assume without loss of generality that there are no such self-transitions. An example of a MAP is shown in Figure 1.

In this paper, we consider a spectrally one-sided MAP $Y(\cdot)$ with the following characteristics. First, assume $Y(0) = 0$ by convention. The generator matrix of the background process $J(\cdot)$, which evolves independently of the Lévy processes upon which $Y(\cdot)$ is based, is $(q_{ij})_{i,j=1}^d$ with $q_i := -q_{ii}$. When there is a transition of $J(\cdot)$ from i to j , there is a jump of which the size is distributed as the random variable L_{ij} ; for a given pair (i, j) of background states, these jumps are assumed independent of all of the driving Lévy processes as well as the background process.

Finally, we incorporate (state-dependent) *killing*, which happens with rate $\vartheta_i \geq 0$ when the background process is in state i . At the moment the MAP is killed, it remains constant indefinitely such that the running maximum becomes the all-time maximum of the process. Alternatively, killing can be thought of as reaching an absorbing background state that corresponds to a Lévy process that is identical to zero. Various specific choices of the killing rates ϑ_i are of interest. When choosing $\vartheta_i = \vartheta > 0$ for all i , for instance, we consider the running maximum over an exponentially distributed horizon with mean $1/\vartheta$. In addition, the choice $\vartheta_i = 0$ for all i corresponds to the all-time maximum. We also note that, with a specific choice of the rates ϑ_i , we can analyze the maximum of a Lévy process over a phase-type interval as argued in Section 5.

Denoting by Δ the killing time of the MAP, its distribution is characterized by the following system of equations:

$$\mathbb{E}(e^{-\alpha\Delta} \mid J(0) = i) = \frac{q_i + \vartheta_i}{q_i + \vartheta_i + \alpha} \left(\sum_{j \neq i} \frac{q_{ij}}{q_i + \vartheta_i} \mathbb{E}(e^{-\alpha\Delta} \mid J(0) = j) + \frac{\vartheta_i}{q_i + \vartheta_i} \right),$$

for $\alpha \geq 0$. This system of equations follows by observing that the time till the first event (being either a transition of the background process $J(\cdot)$ or killing) is exponentially distributed with rate $q_i + \vartheta_i$; then, one needs to distinguish between the background state becoming j (for $j \neq i$) and killing.

2.2. Notation

Each spectrally positive Lévy process $X_i(\cdot)$ is characterized by its *Laplace exponent* $\varphi_i(\cdot)$ as given by

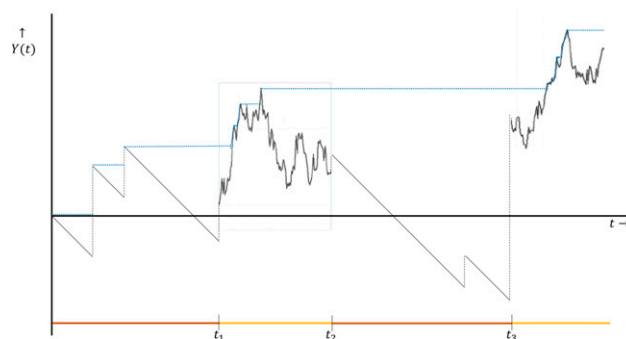
$$\varphi_i(\alpha) := \log \mathbb{E}(e^{-\alpha X_i(1)})$$

for $\alpha \geq 0$, where its right inverse is denoted by $\psi_i(\cdot)$. Similarly, for a spectrally negative Lévy process $X_i(\cdot)$, we consider the *cumulant generating function*

$$\Phi_i(\alpha) := \log \mathbb{E}(e^{\alpha X_i(1)})$$

for $\alpha \geq 0$ with $\Psi_i(\cdot)$ denoting its right inverse. To be interpreted as the MAP counterpart of the Laplace exponent defined earlier, let $M(\alpha) = (m_{ij}(\alpha))_{i,j=1,\dots,d}$ be the matrix with coefficients

Figure 1. Example of a MAP with Two Background States and Its Running Maximum Process



$$m_{ij}(\alpha) := q_{ij} \mathbb{E}(e^{-\alpha L_{ij}}) + \varphi_i(\alpha) \mathbf{1}_{\{i=j\}} - \vartheta_i \mathbf{1}_{\{i=j\}};$$

here, all $X_i(\cdot)$ are assumed spectrally positive, and the jumps L_{ij} are assumed nonnegative almost surely (a.s.). Later, we also work with a similar object for the spectrally negative case (in which the jumps L_{ij} are assumed non-positive almost surely); this MAP counterpart of the cumulant generating function is introduced at the beginning of Section 4.

Our aim is to analyze the distribution of Z_i , that is, the maximum of the MAP under state-dependent killing, conditional on the initial background state being i :

$$Z_i := \sup\{Y(s) : s \in [0, \Delta] \mid J(0) = i\}.$$

As mentioned, we wish to do this without a priori assuming that $J(\cdot)$ corresponds to an irreducible continuous-time Markov chain. A central role is played by the probabilities for $u \geq 0$,

$$p_i(u) := \mathbb{P}(Z_i \geq u).$$

2.3. Preliminaries

Before moving on to the analytic part of the paper, we elaborate on two important results that are essential in our approach. The first of these results, the Wiener–Hopf decomposition, shows that the state of the Lévy process at an exponentially distributed epoch can be written as the difference between two independent nonnegative quantities. In case the Lévy process is spectrally one-sided, these distributions can be characterized explicitly in terms of the model primitives; notably, one of the two quantities is exponentially distributed.

To state this Wiener–Hopf decomposition, denote, for a given Lévy process $X(\cdot)$, its running maximum process by $(\bar{X}(t))_{t \geq 0}$ and its running minimum process by $(\underline{X}(t))_{t \geq 0}$. Let T_ν be an exponentially distributed random variable with mean ν^{-1} , sampled independently of anything else.

Proposition 1 (Wiener–Hopf Decomposition). *Let $(X(t))_{t \geq 0}$ be a Lévy process. Then, $X(T_\nu)$ can be decomposed as the difference between the two independent nonnegative quantities $\bar{X}(T_\nu)$ and $X(T_\nu) - \bar{X}(T_\nu)$. Moreover, the second component $X(T_\nu) - \bar{X}(T_\nu)$ is distributed as $\underline{X}(T_\nu)$.*

If $X(\cdot)$ is spectrally positive with Laplace exponent $\varphi(\cdot)$ and corresponding right inverse $\psi(\cdot)$, then $\underline{X}(T_\nu)$ is distributed as $-T_{\psi(\nu)}$ and

$$\mathbb{E}\left(e^{-\gamma \bar{X}(T_\nu)}\right) = \frac{\nu}{\nu - \varphi(\gamma)} \left(1 - \frac{\gamma}{\psi(\nu)}\right).$$

If $X(\cdot)$ is spectrally negative with cumulant generating function $\Phi(\cdot)$ and corresponding right inverse $\Psi(\cdot)$, then $\bar{X}(T_\nu)$ is distributed as $T_{\Psi(\nu)}$ and

$$\mathbb{E}\left(e^{\gamma \underline{X}(T_\nu)}\right) = \frac{\nu}{\nu - \Phi(\gamma)} \left(1 - \frac{\gamma}{\Psi(\nu)}\right).$$

This decomposition shows that, when $X(\cdot)$ is spectrally one-sided, the (transforms of the) two components can be expressed explicitly in terms of the underlying Laplace exponent (in the spectrally positive case) or cumulant generating function (in the spectrally negative case) and their right inverses. For more background and proofs, we refer to, for example, Kyprianou (2006, chapter VI).

The second result concerns a characterization of the zeroes of the determinant of the matrix exponent $M(\alpha)$ of a spectrally positive MAP. A special role is played by Lévy processes $X_i(\cdot)$ that are monotone a.s. (also referred to as *subordinators*). Let S^\uparrow (S^\downarrow) represent the set of background states corresponding to increasing (decreasing, respectively) subordinators. The result is a slight restatement of Ivanovs et al. (2010), theorem 1 and remark 2.1.

Proposition 2. *Let $\gamma \geq 0$, and suppose the background chain $J(\cdot)$ consists of a single (hence, recurrent) class. If $X_1(\cdot), \dots, X_d(\cdot)$ are spectrally positive Lévy processes and the jump sizes at transition epochs L_{ij} are nonnegative a.s. for all i, j , then the equation $\det(M(\gamma)) = 0$ has $d - |S^\uparrow|$ solutions in \mathbb{C} with a positive real part.*

If $X_1(\cdot), \dots, X_d(\cdot)$ are spectrally negative Lévy processes and the jump sizes at transition epochs L_{ij} are nonpositive, then the equation $\det(M(-\gamma)) = 0$ has $d - |S^\downarrow|$ solutions in \mathbb{C} with a positive real part.

Next to these two results, we often exploit a standard relation between two transform types: for $\alpha \geq 0$ and Y a nonnegative random variable,

$$\beta_Y(\alpha) = \frac{1 - \eta_Y(\alpha)}{\alpha}, \quad (1)$$

where $\beta_Y(\alpha) := \int_0^\infty e^{-\alpha x} \mathbb{P}(Y > x) dx$ and $\eta_Y(\alpha) := \int_0^\infty e^{-\alpha x} \mathbb{P}(Y \in dx)$, the latter representing the Laplace–Stieltjes transform of Y . This relation trivially follows as an application of integration by parts.

2.4. Approach

Now that we have the essential notation and previous results at our disposal, we proceed by summarizing our approach. In both spectrally one-sided cases, the starting point is to use Proposition 1 to find a relationship between characteristics of the MAP at two successive transition epochs of the background chain. This relationship can be transformed to a system of equations for (transforms related to) $p_1(u), \dots, p_d(u)$, involving the matrix $M(\cdot)$; here, we recall that we defined $p_i(u) := \mathbb{P}(Z_i > u)$. In the spectrally positive case (Section 3), this system contains unknown constants that can be determined exploiting Proposition 2. In the spectrally negative case (Section 4), the solution is directly expressed in terms of the zeroes of $\det M(\cdot)$, entailing that, again, Proposition 2 can be used.

3. Spectrally Positive Case

Throughout this section, we assume that the MAP $Y(\cdot)$ is spectrally positive. As pointed out, this entails that, for each $i, j \in \{1, \dots, d\}$, $X_i(\cdot)$ has no downward jumps and the random variable L_{ij} is nonnegative a.s.

We point out how to identify

$$P_i(\gamma) := \int_0^\infty e^{-\gamma u} p_i(u) du,$$

which is the transform of the tail distribution of Z_i , and the corresponding Laplace–Stieltjes transform

$$\zeta_i(\gamma) := \mathbb{E}(e^{-\gamma Z_i}).$$

Note that $\zeta_i(\gamma) = 1 - \gamma P_i(\gamma)$ so that either of these transforms uniquely characterizes the distribution of Z_i , the random variable of our interest. Once the Laplace–Stieltjes transform $\zeta_i(\gamma)$ is evaluated, a numerical inversion algorithm, for example, the one developed in Abate and Whitt (2006), can be used to obtain the distribution of Z_i .

A few observations can be made.

- Recalling that T_γ denotes an exponentially distributed random variable with rate γ , it holds that $\gamma P_i(\gamma) = \mathbb{P}(Z_i > T_\gamma)$. In other words, $\gamma P_i(\gamma)$ can be interpreted as the probability of $Y(\cdot)$ reaching an exponentially distributed level (with mean γ^{-1}) before the process is killed.

- Furthermore, bearing in mind that killing occurs at rate ϑ_i when the background process $J(\cdot)$ is in state i , it is worth noting that, when $\vartheta_i = \vartheta$ for all i , (numerical) inversion of

$$\begin{aligned} \frac{P_i(\gamma)}{\vartheta} &= \frac{1}{\vartheta} \int_0^\infty \int_0^\infty e^{-\gamma u} \mathbb{P}(\bar{Y}(T_\vartheta) > u \mid J(0) = i) dt du \\ &= \int_0^\infty \int_0^\infty e^{-\gamma u} e^{-\vartheta t} \mathbb{P}(\bar{Y}(t) > u \mid J(0) = i) dt du \end{aligned}$$

with respect to both γ and ϑ yields $\mathbb{P}(\bar{Y}(t) > u \mid J(0) = i)$, that is, the tail probability of the running maximum of the unkilld MAP at time t .

- Finally, we note that $P_i(0) = \mathbb{E}(Z_i)$, the expected maximum that the MAP attains before being killed.

Throughout this section, we analyze the behavior of $P_i(\gamma)$ for a fixed initial state i . Because the running maximum of a nondecreasing process necessarily equals the current value of the process, the analysis turns out to be slightly different depending on whether the Lévy process $X_i(\cdot)$ is a nondecreasing subordinator. We consider, in Section 3.1, the case in which the fixed state i does not correspond to a nondecreasing subordinator and, in Section 3.2, the case in which it does. In the analysis of these sections, unknown constants appear; Section 3.3 points out how to determine these constants.

3.1. Nonsubordinator Case

In this section, we focus on the case in which $J(0) = i$, where $i \notin S^\uparrow$ so that the spectrally positive Lévy process $X_i(\cdot)$ may move downward in any interval with positive probability. Recall that $\vartheta_i + q_i$ is the rate of the exponentially distributed time until the first event; this first event corresponds to killing with probability $\pi_i^\circ := \vartheta_i / (\vartheta_i + q_i)$ and to a transition of the background process to state j with probability $\pi_{ij} := q_{ij} / (\vartheta_i + q_i)$. We decompose $p_i(u)$ by distinguishing between the case that the value of the MAP’s running maximum $\bar{X}_i(T_{\vartheta_i+q_i})$ at time $T_{\vartheta_i+q_i}$ is above or below u . In the former case, we have that $Z_i > u$ so that we are done; in the latter case with probability π_i° we do not exceed u before killing, whereas with probability π_{ij} , we are left with the probability of Z_j exceeding level

$u - X_i(T_{\vartheta_i+q_i}) - L_{ij}$ before killing. Formalizing this reasoning and applying Proposition 1 to decompose $X_i(T_{\vartheta_i+q_i})$ into the nonnegative independent random variables $\bar{X}_i(T_{\vartheta_i+q_i})$ and $X_i(T_{\vartheta_i+q_i}) - \bar{X}_i(T_{\vartheta_i+q_i})$, we obtain

$$P_i(\gamma) = \int_{u=0}^{\infty} e^{-\gamma u} \int_{w=u}^{\infty} \mathbb{P}(\bar{X}_i(T_{\vartheta_i+q_i}) \in dw) du + \int_{u=0}^{\infty} e^{-\gamma u} \int_{w=0}^u \mathbb{P}(\bar{X}_i(T_{\vartheta_i+q_i}) \in dw) \sum_{j \neq i} \frac{q_{ij}}{\vartheta_i + q_i} \mathbb{P}(X_i(T_{\vartheta_i+q_i}) + L_{ij} + Z_j \geq u - w) du,$$

where we have also used that $X_i(T_{\vartheta_i+q_i})$ has the same distribution as $X_i(T_{\vartheta_i+q_i}) - \bar{X}_i(T_{\vartheta_i+q_i})$. We continue by evaluating these two terms, which, in the sequel, we refer to by $P_i^+(\gamma)$ and $P_i^-(\gamma)$, separately. Evaluation of the first term is relatively straightforward; an interchange of the order of integration readily yields

$$P_i^+(\gamma) = \int_{w=0}^{\infty} \int_{u=0}^w e^{-\gamma u} du \mathbb{P}(\bar{X}_i(T_{\vartheta_i+q_i}) \in dw) = \int_{w=0}^{\infty} \frac{1 - e^{-\gamma w}}{\gamma} \mathbb{P}(\bar{X}_i(T_{\vartheta_i+q_i}) \in dw) = \frac{1 - \kappa_i(\gamma)}{\gamma},$$

where

$$\kappa_i(\gamma) := \mathbb{E}(e^{-\gamma \bar{X}_i(T_{\vartheta_i+q_i})}) = \frac{\vartheta_i + q_i}{\vartheta_i + q_i - \varphi_i(\gamma)} \left(1 - \frac{\gamma}{\psi_i(\vartheta_i + q_i)} \right) \tag{2}$$

by virtue of Proposition 1. We now focus on the evaluation of the second term, which is considerably more involved. As a first step, we interchange the order of the sum and the integrals:

$$P_i^-(\gamma) = \sum_{j \neq i} \frac{q_{ij}}{\vartheta_i + q_i} P_{ij}^-(\gamma),$$

where

$$P_{ij}^-(\gamma) := \int_{u=0}^{\infty} e^{-\gamma u} \int_{w=0}^u \mathbb{P}(\bar{X}_i(T_{\vartheta_i+q_i}) \in dw) \mathbb{P}(X_i(T_{\vartheta_i+q_i}) + L_{ij} + Z_j \geq u - w) du.$$

The quantities $P_{ij}^-(\gamma)$ can be evaluated separately as follows. Realize that, by Proposition 1, $-\bar{X}_i(T_{\vartheta_i+q_i})$ is exponentially distributed with rate $\mu_i := \psi_i(\vartheta_i + q_i)$. We, thus, obtain the triple integral

$$P_{ij}^-(\gamma) = \int_{u=0}^{\infty} e^{-\gamma u} \int_{w=0}^u \int_{y=0}^{\infty} \mu_i e^{-\mu_i y} \mathbb{P}(\bar{X}_i(T_{\vartheta_i+q_i}) \in dw) \mathbb{P}(L_{ij} + Z_j \geq u - w + y) dy du,$$

which, after replacing y by $x - u + w$, can be rewritten as

$$\int_{u=0}^{\infty} e^{-\gamma u} \int_{w=0}^u \int_{x=u-w}^{\infty} \mu_i e^{-\mu_i(x-u+w)} \mathbb{P}(\bar{X}_i(T_{\vartheta_i+q_i}) \in dw) \mathbb{P}(L_{ij} + Z_j \geq x) dx du.$$

Our strategy is to interchange the order of the integrals so as to be able to do the (easy) integration over u first. By first swapping the integrals over u and w and then those over u and x , we find

$$P_{ij}^-(\gamma) = \int_{w=0}^{\infty} \int_{x=0}^{\infty} \left(\int_{u=w}^{x+w} e^{-(\gamma-\mu_i)u} du \right) \mu_i e^{-\mu_i(x+w)} \mathbb{P}(\bar{X}_i(T_{\vartheta_i+q_i}) \in dw) \mathbb{P}(L_{ij} + Z_j \geq x) dx = \int_{w=0}^{\infty} \int_{x=0}^{\infty} \frac{\mu_i e^{-\gamma w}}{\gamma - \mu_i} (e^{-\mu_i x} - e^{-\gamma x}) \mathbb{P}(\bar{X}_i(T_{\vartheta_i+q_i}) \in dw) \mathbb{P}(L_{ij} + Z_j \geq x) dx,$$

where the second equality follows by performing the integration over u and reorganizing the resulting expression. We can rewrite this expression as the difference of two terms in each of which the double integral factorizes into the product of two single integrals. In particular, with

$$\eta_{ij}(\gamma) := \int_0^{\infty} e^{-\gamma x} \mathbb{P}(L_{ij} + Z_j \geq x) dx,$$

some rearranging of terms leads to the expression

$$P_{ij}^-(\gamma) = \frac{\mu_i}{\gamma - \mu_i} \int_0^{\infty} e^{-\gamma w} \mathbb{P}(\bar{X}_i(T_{\vartheta_i+q_i}) \in dw) (\eta_{ij}(\mu_i) - \eta_{ij}(\gamma)) = \frac{\mu_i \kappa_i(\gamma)}{\gamma - \mu_i} (\eta_{ij}(\mu_i) - \eta_{ij}(\gamma)).$$

To separate L_{ij} and Z_j in the expression for $\eta_{ij}(\gamma)$, we rely on a probabilistic argument for nonnegative and independent random variables A and B . That is, using the memoryless property of the exponential distribution,

$$\begin{aligned} \int_0^\infty e^{-\gamma x} \mathbb{P}(A + B \geq x) \, dx &= \frac{1}{\gamma} \mathbb{P}(A + B > T_\gamma) = \frac{1}{\gamma} \left(\mathbb{P}(A > T_\gamma) + \mathbb{P}(A < T_\gamma) \mathbb{P}(A + B > T_\gamma | A < T_\gamma) \right) \\ &= \frac{1}{\gamma} \left(\mathbb{P}(A > T_\gamma) + \mathbb{P}(A < T_\gamma) \mathbb{P}(B > T_\gamma) \right) \\ &= \frac{1 - \mathbb{E}(e^{-\gamma A})}{\gamma} + \mathbb{E}(e^{-\gamma A}) \int_0^\infty e^{-\gamma x} \mathbb{P}(B \geq x) \, dx, \end{aligned} \tag{3}$$

where we use (1) in the last step. Furthermore, let $\lambda_{ij}(\cdot)$ be the Laplace–Stieltjes transform of L_{ij} , the size of the (nonnegative) jump at a transition by the background chain from state i to state j (in other words, $\lambda_{ij}(\gamma) := \mathbb{E}(e^{-\gamma L_{ij}})$). Recalling the definition of $P_i(\gamma)$, we conclude from Identity (3) that

$$\eta_{ij}(\gamma) = \Lambda_{ij}(\gamma) + \lambda_{ij}(\gamma) P_j(\gamma), \quad \Lambda_{ij}(\gamma) := \frac{1 - \lambda_{ij}(\gamma)}{\gamma}.$$

We now combine all these findings, which enables us to express $P_i(\gamma)$ in terms of $P_j(\gamma)$ with $j \neq i$. Recalling that

$$P_i(\gamma) = P_i^+(\gamma) + \sum_{j \neq i} \frac{q_{ij}}{\vartheta_i + q_i} P_{ij}^-(\gamma)$$

and substituting the obtained expressions for $P_i^+(\gamma)$ and $P_{ij}^-(\gamma)$, we obtain, for any $i \notin S^\uparrow$,

$$P_i(\gamma) = \frac{1 - \kappa_i(\gamma)}{\gamma} + \frac{\mu_i \kappa_i(\gamma)}{\gamma - \mu_i} \sum_{j \neq i} \frac{q_{ij}}{\vartheta_i + q_i} (\Lambda_{ij}(\mu_i) + \lambda_{ij}(\mu_i) P_j(\mu_i) - \Lambda_{ij}(\gamma) - \lambda_{ij}(\gamma) P_j(\gamma)).$$

By using (2), recalling that $\mu_i = \psi_i(\vartheta_i + q_i)$, and defining

$$\bar{\kappa}_i(\gamma) := \frac{1 - \kappa_i(\gamma)}{\gamma} = \frac{1}{\vartheta_i + q_i - \varphi_i(\gamma)} \left(\frac{\vartheta_i + q_i}{\psi_i(\vartheta_i + q_i)} - \frac{\varphi_i(\gamma)}{\gamma} \right),$$

we can compactly summarize this result as follows.

Lemma 1. For $i \notin S^\uparrow$ and any $\gamma \geq 0$, the transform of the tail probability $p_i(u)$ is given by

$$P_i(\gamma) = \bar{\kappa}_i(\gamma) + \sum_{j \neq i} \frac{q_{ij}}{\vartheta_i + q_i - \varphi_i(\gamma)} (\Lambda_{ij}(\gamma) + \lambda_{ij}(\gamma) P_j(\gamma) - \Lambda_{ij}(\mu_i) - \lambda_{ij}(\mu_i) P_j(\mu_i)). \tag{4}$$

So far, we have been working with the transform $P_i(\gamma)$ of the tail probability $p_i(u)$. In the remainder of this section, we rewrite the above lemma in terms of $\zeta_i(\gamma) := \mathbb{E}(e^{-\gamma Z_i})$, which takes a particularly nice form. To this end, first note that, as a consequence of (1),

$$P_i(\gamma) = \frac{1 - \zeta_i(\gamma)}{\gamma}. \tag{5}$$

Substituting this in (4) and rewriting leads to, for $\gamma \geq 0$,

$$\frac{1 - \zeta_i(\gamma)}{\gamma} = \frac{1 - \kappa_i(\gamma)}{\gamma} + \sum_{j \neq i} \frac{q_{ij}}{\vartheta_i + q_i - \varphi_i(\gamma)} \left(\frac{1 - \lambda_{ij}(\gamma)}{\gamma} + \lambda_{ij}(\gamma) \frac{1 - \zeta_j(\gamma)}{\gamma} - \frac{1 - \lambda_{ij}(\mu_i)}{\mu_i} - \lambda_{ij}(\mu_i) \frac{1 - \zeta_j(\mu_i)}{\mu_i} \right),$$

or, equivalently,

$$\begin{aligned} \zeta_i(\gamma) &= \kappa_i(\gamma) + \sum_{j \neq i} \frac{q_{ij}}{\vartheta_i + q_i - \varphi_i(\gamma)} \left(\lambda_{ij}(\gamma) \zeta_j(\gamma) - \lambda_{ij}(\mu_i) \zeta_j(\mu_i) \frac{\gamma}{\mu_i} \right) - \frac{q_i}{\vartheta_i + q_i - \varphi_i(\gamma)} \left(1 - \frac{\gamma}{\mu_i} \right) \\ &= \sum_{j \neq i} \frac{q_{ij}}{\vartheta_i + q_i - \varphi_i(\gamma)} \left(\lambda_{ij}(\gamma) \zeta_j(\gamma) - \lambda_{ij}(\mu_i) \zeta_j(\mu_i) \frac{\gamma}{\mu_i} \right) + \frac{\vartheta_i}{\vartheta_i + q_i - \varphi_i(\gamma)} \left(1 - \frac{\gamma}{\mu_i} \right). \end{aligned} \tag{6}$$

Multiplying (6) by $\vartheta_i + q_i + \varphi_i(\gamma)$ yields the identity

$$(\vartheta_i + q_i - \varphi_i(\gamma)) \zeta_i(\gamma) = \vartheta_i \left(1 - \frac{\gamma}{\mu_i} \right) + \sum_{j \neq i} q_{ij} \left(\lambda_{ij}(\gamma) \zeta_j(\gamma) - \lambda_{ij}(\mu_i) \zeta_j(\mu_i) \frac{\gamma}{\mu_i} \right). \tag{7}$$

We continue by considering the case that none of the states correspond to a nondecreasing subordinator. Then, the system of Equations (7) that characterizes $\zeta_1(\gamma), \dots, \zeta_d(\gamma)$ can be written in a considerably more compact form. To this end, recall that the matrix $M(\gamma) \equiv (m_{ij}(\gamma))_{i,j=1}^d$ is defined by

$$m_{ij}(\gamma) = q_{ij}\lambda_{ij}(\gamma) + \varphi_i(\gamma) \mathbf{1}_{\{i=j\}} - \vartheta_i \mathbf{1}_{\{i=j\}}. \tag{8}$$

Furthermore, using that $m_{ii}(\mu_i) = 0$, we define the quantity $b_i(\gamma)$ as follows:

$$\begin{aligned} b_i(\gamma) &:= \sum_{j \neq i} m_{ij}(\mu_i) \zeta_j(\mu_i) \frac{\gamma}{\mu_i} - \vartheta_i \left(1 - \frac{\gamma}{\mu_i}\right) \\ &= \left(\sum_{j=1}^d m_{ij}(\mu_i) \zeta_j(\mu_i) + \vartheta_i \right) \frac{\gamma}{\mu_i} - \vartheta_i = \gamma \frac{\omega_i + \vartheta_i}{\mu_i} - \vartheta_i, \end{aligned} \tag{9}$$

with the constants ω_i defined by

$$\omega_i := \sum_{j=1}^d m_{ij}(\mu_i) \zeta_j(\mu_i). \tag{10}$$

Upon combining these, we obtain the equation $\sum_{j=1}^d m_{ij}(\gamma) \zeta_j(\gamma) = b_i(\gamma)$, for $i = 1, \dots, d$. In evident vector/matrix notation, we have, thus, rewritten (7) as follows.

Theorem 1. *If no state corresponds to a nondecreasing subordinator (i.e., $i \notin S^\uparrow$ for all $i = 1, \dots, d$), then, for any $\gamma \geq 0$,*

$$M(\gamma) \boldsymbol{\zeta}(\gamma) = \mathbf{b}(\gamma). \tag{11}$$

It is important to note that, throughout the analysis, no assumptions on the chain structure of the background process $J(\cdot)$ are imposed. Also observe that we still need to identify the constants ω_i that appear in (9), which we do in Section 3.3.

As an aside, we mention that Identity (6) can alternatively be derived using a probabilistic argumentation, which is, for completeness, provided in Appendix A.1.

3.2. Subordinator Case

The previous section deals with the case in which the initial state i is such that the spectrally positive Lévy process $X_i(\cdot)$ does not correspond to a nondecreasing subordinator ($i \notin S^\uparrow$, that is). The analysis led to the matrix Equation (11) for the case in which no state corresponds to a nondecreasing subordinator. In the present section, we address the case in which $i \in S^\uparrow$ and point out how (11) should be adjusted if some of the background states correspond to nondecreasing subordinators.

To this end, let, for a given $i = 1, \dots, d$, the Lévy process $X_i(\cdot)$ be nondecreasing almost surely. It is important to note that in this case, necessarily, $\varphi_i(\gamma) \leq 0$ and $\psi_i(\gamma) = \infty$ for all $\gamma \geq 0$. Our method for analyzing $P_i(\gamma)$ is largely the same as in the previous section, but is somewhat simpler because of the evident fact that any nondecreasing process attains its maximum at the end of the interval under consideration. Concretely, we could mimic the approach of the previous section while replacing $\bar{X}(T_{\vartheta_i+q_i})$ by $X(T_{\vartheta_i+q_i})$, but it turns out to be convenient to condition on the value of $Y(\cdot)$ at the minimum of the killing time and the first transition of the background process. This yields

$$\begin{aligned} P_i(\gamma) &= \int_0^\infty e^{-\gamma u} p_i(u) \, du \\ &= \int_0^\infty e^{-\gamma u} \mathbb{P}(X_i(T_{\vartheta_i}) \geq u, T_{\vartheta_i} \leq T_{q_i}) \, du + \int_0^\infty e^{-\gamma u} \sum_{j \neq i} \frac{q_{ij}}{q_i} \mathbb{P}(X_i(T_{q_i}) + L_{ij} + Z_j \geq u, T_{\vartheta_i} > T_{q_i}) \, du. \end{aligned} \tag{12}$$

With $P_i^+(\gamma)$ and $P_i^-(\gamma)$, respectively, representing the two terms in the right-hand side of (12), we use Proposition 1 and (1) to obtain

$$\begin{aligned} P_i^+(\gamma) &= \int_0^\infty e^{-\gamma u} \frac{\vartheta_i}{\vartheta_i + q_i} \mathbb{P}(X_i(T_{\vartheta_i+q_i}) \geq u) \, du \\ &= \frac{\vartheta_i}{\vartheta_i + q_i} \frac{1}{\gamma} \left(1 - \mathbb{E}\left(e^{-\gamma X_i(T_{\vartheta_i+q_i})}\right)\right) = \frac{\vartheta_i}{\vartheta_i + q_i} \frac{1}{\gamma} \left(1 - \frac{\vartheta_i + q_i}{\vartheta_i + q_i - \varphi_i(\gamma)}\right) \end{aligned}$$

and

$$\begin{aligned}
 P_i^-(\gamma) &= \int_0^\infty e^{-\gamma u} \sum_{j \neq i} \frac{q_{ij}}{\vartheta_i + q_i} \mathbb{P}(X_i(T_{\vartheta_i+q_i}) + L_{ij} + Z_j \geq u) \, du \\
 &= \sum_{j \neq i} \frac{q_{ij}}{\vartheta_i + q_i} \frac{1}{\gamma} \left(1 - \mathbb{E}(e^{-\gamma X_i(T_{\vartheta_i+q_i})}) \lambda_{ij}(\gamma) \zeta_j(\gamma) \right) \\
 &= \sum_{j \neq i} \frac{q_{ij}}{\vartheta_i + q_i} \frac{1}{\gamma} \left(1 - \frac{\vartheta_i + q_i}{\vartheta_i + q_i - \varphi_i(\gamma)} \lambda_{ij}(\gamma) (1 - \gamma P_j(\gamma)) \right) \\
 &= \frac{q_i}{\vartheta_i + q_i} \frac{1}{\gamma} - \frac{1}{\gamma} \sum_{j \neq i} \frac{q_{ij}}{\vartheta_i + q_i - \varphi_i(\gamma)} \left(1 - \gamma \Lambda_{ij}(\gamma) - \gamma \lambda_{ij}(\gamma) P_j(\gamma) \right),
 \end{aligned}$$

recalling the identities $\zeta_j(\gamma) = 1 - \gamma P_j(\gamma)$ and $\lambda_{ij}(\gamma) = 1 - \gamma \Lambda_{ij}(\gamma)$ in the last two steps.

After collecting these intermediate results, we obtain the following characterization.

Lemma 2. For $i \in S^\uparrow$ and any $\gamma \geq 0$,

$$P_i(\gamma) = \frac{-\varphi_i(\gamma)}{\vartheta_i + q_i - \varphi_i(\gamma)} \frac{1}{\gamma} + \sum_{j \neq i} \frac{q_{ij}}{\vartheta_i + q_i - \varphi_i(\gamma)} (\Lambda_{ij}(\gamma) + \lambda_{ij}(\gamma) P_j(\gamma)). \tag{13}$$

Observe that (13) can also be obtained by taking the limit $\mu_i = \psi_i(\vartheta_i + q_i) \rightarrow \infty$ in Lemma 1, which is consistent with the fact that $\psi_i(\cdot) = \infty$ for subordinator processes $X_i(\cdot)$. Similar to the nonsubordinator case, we can again present a vector/matrix version for the Laplace–Stieltjes transforms $\zeta_i(\gamma)$ of (the distributions of) the random variables Z_i . To this end, define $b_i^\circ(\gamma) := -\vartheta_i$ for $i \in S^\uparrow$ and $b_i^\circ(\gamma) := b_i(\gamma)$ for $i \notin S^\uparrow$ with $b_i(\gamma)$ as defined in (9). Then, using similar steps as before, we eventually find the following counterpart of Theorem 1.

Theorem 2. For any $\gamma \geq 0$, the vectors $\zeta(\gamma) = (\zeta_1(\gamma), \dots, \zeta_d(\gamma))$ and $\mathbf{b}^\circ(\gamma) = (b_1^\circ(\gamma), \dots, b_d^\circ(\gamma))$ satisfy

$$M(\gamma) \zeta(\gamma) = \mathbf{b}^\circ(\gamma). \tag{14}$$

Note that, also in this set of equations, the vector $\mathbf{b}^\circ(\gamma)$ still contains unknowns. These constants ω_i , one for each $i \notin S^\uparrow$, are identified in the next section.

3.3. Evaluation of the Unknowns

So far, we have established that, in the spectrally positive case, the Laplace–Stieltjes transforms of Z_1, \dots, Z_d are given by the solutions of (14) (which simplifies to (11) in case none of the Lévy processes $X_i(\cdot)$ is a nondecreasing subordinator process). This section settles the complication that (11) contains unknown constants ω_i . As we see, the number of such constants equals the number of states that do not correspond to nondecreasing subordinators, which we denote by d° (i.e., $d^\circ := d - |S^\uparrow|$). To identify these d° unknowns and, ultimately, the solution $\zeta(\gamma)$ of (14), we subsequently analyze three cases:

- The background chain has no transient classes.
- The background chain has exactly one transient class.
- The background chain has more than one transient class.

We proceed by studying each of these cases separately.

3.3.1. No Transient Classes. In case the background chain has no transient classes, all classes of the chain are necessarily recurrent. To analyze Z_i , it evidently suffices to restrict ourselves to the recurrent class in which the background state i is. As a consequence, without loss of generality, we may assume that the background process $J(\cdot)$ is irreducible. In this case, which has been studied extensively (see, e.g., the results in D’Auria et al. 2010, Dieker and Mandjes 2011), the following procedure can be used to identify the ω_i . Note that, using the linear equations given in (14), one may express the vector $\zeta(\gamma)$ by relying on Cramer’s rule. More concretely, with the matrix $M_{b,i}(\gamma)$ denoting the matrix $M(\gamma)$ in which the i th column is replaced by the vector $\mathbf{b}^\circ(\gamma)$, we have that

$$\zeta_i(\gamma) = \frac{\det M_{b,i}(\gamma)}{\det M(\gamma)}.$$

Because $\zeta_i(\gamma)$ is finite, any zero of the denominator should be a zero of the numerator. According to Proposition 2, in case $J(\cdot)$ consists of a single class, $\det M(\gamma) = 0$ has d° zeroes in the right half of the complex plane. For ease of exposition, we make the assumption that these zeroes have multiplicity one (and we call them, say, $\gamma_1, \dots, \gamma_{d^\circ}$). In the special case this assumption does not hold, a reasoning similar to the one that follows still

applies, but one needs to resort to the concept of Jordan chains. We do not discuss this procedure in detail and instead refer to the in-depth treatment in D'Auria et al. (2010).

Having distinct zeroes guarantees that we have d° equations to identify the ω_i . That is, for $i = 1, \dots, d$ and $j = 1, \dots, d^\circ$,

$$\det M_{b,i}(\gamma_j) = 0; \quad (15)$$

in other words, the zeroes of $\det M$ (in the right half of the complex plane, that is) are also zeroes of $\det M_{b,i}$ for each $i = 1, \dots, d$. For any given $j = 1, \dots, d^\circ$, this seemingly yields d equations, but it can be seen easily that each of these d equations effectively provides the same information. Indeed, with $\mathbf{m}_k(\gamma)$ denoting the k th column of $M(\gamma)$, suppose, for any fixed i , that $\det M(\gamma) = 0$ and $\det M_{b,i}(\gamma) = 0$ for some $\gamma \geq 0$. This implies that both $M(\gamma)$ and $M_{b,i}(\gamma)$ are singular, and as a consequence, there are nontrivial vectors \mathbf{u} and \mathbf{v} such that

$$\sum_{j=1}^d \mathbf{m}_j(\gamma) v_j = \mathbf{0}, \quad \sum_{j \neq i} \mathbf{m}_j(\gamma) u_j + \mathbf{b}^\circ(\gamma) u_i = \mathbf{0}.$$

As a consequence, for any $i' \neq i$,

$$\begin{aligned} \mathbf{0} &= -u_{i'} \sum_{j=1}^d \mathbf{m}_j(\gamma) v_j + v_{i'} \sum_{j \neq i} \mathbf{m}_j(\gamma) u_j + v_{i'} \mathbf{b}^\circ(\gamma) u_i \\ &= -u_{i'} v_i \mathbf{m}_i(\gamma) + \sum_{j \neq i, i'} (v_{i'} u_j - u_{i'} v_j) \mathbf{m}_j(\gamma) + v_{i'} u_i \mathbf{b}^\circ(\gamma), \end{aligned}$$

entailing that there is a linear combination of the columns of $M_{b,i'}(\gamma)$ that equals $\mathbf{0}$. In other words, $M_{b,i'}(\gamma)$ is singular, and hence, $\det M_{b,i'}(\gamma) = 0$ as well.

Now that we know that (15), for any given index $j = 1, \dots, d^\circ$, provides us with just a single equation, we study this equation in more detail. Let us focus on $\det M_{b,1}(\gamma_j) = 0$ for $j = 1, \dots, d^\circ$ (we take $i = 1$, that is). With $\bar{M}_{ij}(\gamma)$ representing the $(d-1) \times (d-1)$ matrix that results after deleting the i th column and the j th row from $M(\gamma)$ and recalling that $b_i^\circ(\gamma) = \gamma(\omega_i + \vartheta_i)/\mu_i - \vartheta_i$ for $i \notin S^\uparrow$ and $b_i^\circ(\gamma) = -\vartheta_i$ for $i \in S^\uparrow$, this equation can be rewritten as

$$\sum_{i \notin S^\uparrow} \left(\gamma_j \frac{\omega_i + \vartheta_i}{\mu_i} - \vartheta_i \right) (-1)^{1+i} \det \bar{M}_{i1}(\gamma_j) + \sum_{i \in S^\uparrow} \vartheta_i (-1)^i \det \bar{M}_{i1}(\gamma_j) = 0.$$

We, thus, obtain d° equations (one for each γ_j) that are linear in the unknowns $\omega_1, \dots, \omega_{d^\circ}$, which can be dealt with in the standard manner, thus yielding a solution for the ω_i .

3.3.2. A Single Transient Class. We now consider the case in which the background chain has a single transient class, say $T \subset \{1, \dots, d\}$, next to one or more recurrent classes. In this case, note that the $\zeta_i(\gamma)$ for all recurrent states i , that is, $i \notin T$, can be computed by the procedure pointed out earlier. Subsequently, for $i \in T$, we rewrite the i th equation of (14) as

$$\sum_{j \in T} m_{ij}(\gamma) \zeta_j(\gamma) = b_i^\circ(\gamma) - \sum_{j \notin T} m_{ij}(\gamma) \zeta_j(\gamma). \quad (16)$$

Observe that the right-hand side is known; we denote it by $\bar{b}_i^\circ(\gamma)$. Define $\bar{d} := |T|$ as the number of transient states and $\bar{d}^\circ := |T \setminus S^\uparrow|$ as the number of transient states that do not correspond to nondecreasing subordinators. In addition, we define the $\bar{d} \times \bar{d}$ matrix $\bar{M}(\gamma) := (m_{ij}(\gamma))_{i,j \in T}$, and we let the \bar{d} -dimensional vector $\bar{\zeta}(\gamma) := (\zeta_i(\gamma))_{i \in T}$ represent the entries of $\zeta(\gamma)$ that correspond to the states in T . Likewise, $\bar{\mathbf{b}}^\circ(\gamma) = (\bar{b}_i^\circ(\gamma))_{i \in T}$ represents the vector of right-hand sides of (16). Using these definitions, (16) can be written as

$$\bar{M}(\gamma) \bar{\zeta}(\gamma) = \bar{\mathbf{b}}^\circ(\gamma).$$

Clearly, suppose that we could prove that $\det \bar{M}(\gamma) = 0$ has \bar{d}° zeroes in the right half of the complex plane; then, we could identify the constants ω_i by following the same approach as the one we develop for the case of no transient classes. This is why we now verify that the entries of $\bar{M}(\gamma)$ can be written in the form (8) with transition rates that correspond to a single recurrent class so that we can apply Proposition 2 to establish the desired property for the number of zeroes of $\det \bar{M}(\gamma) = 0$ in the right half of the complex plane. By rewriting the diagonal elements of $\bar{M}(\gamma)$ as

$$m_{ii}(\gamma) = q_{ii} + \varphi_i(\gamma) - \vartheta_i = \bar{q}_{ii} + \varphi_i(\gamma) - \bar{\vartheta}_i, \quad (17)$$

with

$$\bar{q}_{ii} := - \sum_{j \in T \setminus \{i\}} q_{ij} \quad \text{and} \quad \bar{\vartheta}_i := \vartheta_i + \sum_{j \in T} q_{ij}, \tag{18}$$

we conclude that the row sums of transition rates $\bar{q}_{ii} + \sum_{j \in T \setminus \{i\}} q_{ij}$ equal zero for all $i \in T$. This means that $\bar{M}(\gamma)$ indeed has the desired form: the entries are of the form (8) with transition rates that correspond to a single recurrent class. Applying Proposition 2, we have that $\det \bar{M}(\gamma) = 0$ has \bar{d}° zeroes in the right half of the complex plane so that we can identify the ω_i for $i \in T \setminus S^\uparrow$ (repeating the remark on roots with the multiplicity larger than one as made earlier in relation to the case with recurrent states only).

3.3.3. Multiple Transient Classes. We now consider the case in which there are $K > 1$ transient classes (say T_1, \dots, T_K). We let R be the union of all remaining recurrent classes. Furthermore, we write $T_k \rightsquigarrow T_{k'}$ if there is a direct transition from a state in T_k to a state in $T_{k'}$, that is, there is a state $i \in T_k$ and a state $j \in T_{k'}$ such that $q_{ij} > 0$. To handle the case of multiple transient classes, we order the transient classes in “layers” as follows. Let $C_0 := R$, and for $n = 1, 2, \dots$, let the n th layer set be given by

$$C_n := C_{n-1} \cup \{T_k : \text{for all } k' \text{ such that } T_k \rightsquigarrow T_{k'} \text{ it holds that } k' \in C_{n-1}\}.$$

It is worth noting that, if a background state i is element of the layer set C_j but not of C_{j-1} , then the background chain can reach a recurrent state in minimally j transitions. In addition, we can observe that the number of non-empty layer sets (including C_0) is at most $K + 1$. See Figure 2 for a pictorial illustration.

In the previous two cases, we already explained how to compute $\zeta_i(\gamma)$ for $i \in R$ and $i \in C_1$, respectively. We now point out how we can evaluate $\zeta_i(\gamma)$ for $i \in C_n$, having $\zeta_i(\gamma)$ for $i \in R, C_1, \dots, C_{n-1}$ at our disposal so that we can recursively determine all $\zeta_i(\gamma)$. Suppose that $T_k \subseteq C_n \setminus C_{n-1}$ (where it is noted that there are potentially multiple transient classes in $C_n \setminus C_{n-1}$). As states in T_k cannot have direct transitions to classes outside C_{n-1} , we have, for $i \in T_k$, that

$$\sum_{j \in T_k} m_{ij}(\gamma) \zeta_j(\gamma) = b_i^\circ(\gamma) - \sum_{j \in C_{n-1}} m_{ij}(\gamma) \zeta_j(\gamma).$$

From this point, the analysis follows that of the case with a single transient class. More specifically, the number of zeroes of the determinant of the matrix $(m_{ij}(\gamma))_{i,j \in T_k}$ in the right half of the complex plane equals the number of states in T_k that do not correspond to nondecreasing subordinators, using the same argument as in the case of a single transient class. This allows us to identify the ω_i for $i \in T_k \setminus S^\uparrow$.

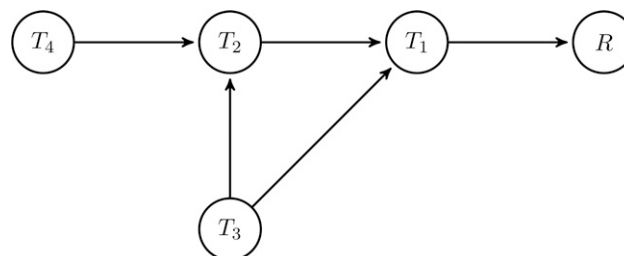
4. Spectrally Negative Case

The model we analyze in this section can be seen as the spectrally negative counterpart of the one considered in the previous section. This concretely means that now the Lévy processes $X_i(\cdot)$ are assumed to be spectrally negative and the jumps L_{ij} are nonpositive. In addition, we replace our earlier definition of the entries of the matrix $M(\nu) = (m_{ij}(\nu))_{i,j=1}^d$ by

$$m_{ij}(\nu) := q_{ij} \lambda_{ij}(-\nu) + \Phi_i(\nu) \mathbf{1}_{\{i=j\}} - \vartheta_i \mathbf{1}_{\{i=j\}}, \tag{19}$$

for $\nu \geq 0$, to account for the nonpositive jumps. As in the spectrally positive case, the matrix $M(\nu)$ is helpful in establishing the main result of this section.

Figure 2. Example of Layer Sets with $K = 4$ Transient Classes



Notes. In this case, $C_0 = R$, $C_1 = R \cup T_1$, $C_2 = R \cup T_1 \cup T_2$, and $C_3 = R \cup T_1 \cup T_2 \cup T_3 \cup T_4$. In this figure, an arrow between the classes U and V means that $U \rightsquigarrow V$.

Downloaded from infoms.org by [146.50.150.90] on 02 November 2022, at 03:33 . For personal use only, all rights reserved.

Unlike in the previous section, throughout the present section, we focus directly on $p_i(u) = \mathbb{P}(Z_i \geq u)$ rather than on its Laplace transform $P_i(\gamma)$; as it turns out, Laplace transforms are not required in the analysis of the spectrally negative case. A convenient feature, made more precise later, is that, in the spectrally negative setting, the *form* of the distribution of the Z_i is known.

Somewhat comparably to the setup of Section 3, to make the presentation as transparent as possible, we first treat the case in which none of the Lévy processes $X_i(\cdot)$ is a nonincreasing subordinator (Section 4.1), after which we point out how to adapt the analysis to the case in which some of the $X_i(\cdot)$ are (Section 4.2).

The following claim plays a crucial role in this section.

Lemma 3. *The equation $\det M(v) = 0$ has d° zeroes with a positive real part, say v_1, \dots, v_d .*

By Proposition 2, we already know that Lemma 3 holds if the background process is irreducible. In Section 4.3, we provide a proof for the case that the background process has a general chain structure.

4.1. Nonsubordinator Case

In this section, we consider the situation in which none of the states corresponds to a nonincreasing subordinator. This means that, for all $i = 1, \dots, d$, Proposition 1 implies that the running maximum $\bar{X}_i(T_{\vartheta_i+q_i})$ has an exponential distribution with rate $\mu_i := \Psi_i(\vartheta_i + q_i)$. Recalling that the time until either the process is killed or the background chain transitions are distributed exponentially with rate $\vartheta_i + q_i$, we, thus, obtain the identity

$$p_i(u) = e^{-\mu_i u} + \sum_{j \neq i} \frac{q_{ij}}{\vartheta_i + q_i} p_{ij}^-(u), \tag{20}$$

with

$$p_{ij}^-(u) := \int_0^u \mu_i e^{-\mu_i w} \mathbb{P}(\underline{X}_i(T_{\vartheta_i+q_i}) + L_{ij} + Z_j \geq u - w) dw. \tag{21}$$

To streamline our analysis, we impose Property (A). By Lemma 3, we know that, in this setting, without nonincreasing subordinators, the equation $\det M(v) = 0$ has d zeroes with a positive real part.

The d zeroes of $\det M(v)$ with a positive real part are distinct. (A)

Importantly, however, imposing (A) effectively does not impose any restriction: as discussed in Remark 2, the argumentation can be adapted to cover zeroes with higher multiplicities.

A crucial fact is that the first-passage process pertaining to $Y(\cdot)$ is a MAP itself irrespective of whether the background process is irreducible; cf. the discussion in Ivanovs (2011), section 2.6. This implies that the random variables Z_i have phase-type distributions; see, for example, Asmussen (2003), section III.4, for background on this class of distributions. More specifically, one obtains their Laplace transforms by plugging in $\alpha = 0$ in the expression of the first statement of Ivanovs (2011), corollary 4.21. The result concretely entails that, in our setting, without nonincreasing subordinators, for a $d \times d$ transition rate matrix Λ , a vector $\boldsymbol{\lambda} := -\Lambda \mathbf{1} \geq \mathbf{0}$ with at least one positive entry, and initial distributions $\mathbf{a}_1, \dots, \mathbf{a}_d$,

$$\mathbb{E}(e^{-\gamma Z_i}) = \mathbf{a}_i(\gamma I - \Lambda)^{-1} \boldsymbol{\lambda}.$$

Recalling the definition of $M(v)$ in (19), the zeroes of $\det M(v)$ coincide with those of $\det(-vI - \Lambda)$; cf. Ivanovs (2011), theorem 4.7, and again the first statement of Ivanovs (2011), corollary 4.21. Because of this, the matrix $(\gamma I - \Lambda)$ is singular in $\gamma = -v_1, \dots, -v_d$; hence, $\mathbb{E}(e^{-\gamma Z_i})$ can be written as a linear combination of the terms $1/(\gamma + v_1), \dots, 1/(\gamma + v_d)$. This means that, under (A), we can write, for $k = 1, \dots, d$ and $u \geq 0$,

$$p_i(u) = \sum_{k=1}^d c_{ik} e^{-v_k u}, \tag{22}$$

where $C = (c_{ik})_{i,k=1}^d$ is a matrix of unknown coefficients whose rows add up to one.

Remark 1. As mentioned, the first statement of Ivanovs (2011), corollary 4.21, already provides a characterization of the distribution of the random variables Z_i (for $i = 1, \dots, d$) under a possibly nonirreducible background chain. It is noted, though, that in Ivanovs (2011), corollary 4.21, the distribution of the Z_i is given in terms of a Laplace–Stieltjes transform, which contains unknown matrices (viz. in the terminology of Ivanovs (2011), the matrices $\Lambda(q)$ and $\Pi(q)$), which can be numerically computed, for example, using Ivanovs (2011), theorem 4.14. Our contribution is that we obtain a more explicit result in Theorem 3: our result concerns the probabilities $p_i(u)$,

corresponding to the tail of Z_i , rather than their transforms. For each i , we succeed in expressing $p_i(u)$ in terms of the solutions of an eigensystem.

We now exploit the structure as given in (22) to generate equations by which we can determine the coefficients c_{ik} . To this end, we define $a_{ik} := c_{ik}v_k$. By conditioning on the value of Z_j in (21) using (22), we, thus, obtain

$$p_{ij}^-(u) = \int_0^u \mu_i e^{-\mu_i w} \int_0^\infty \sum_{k=1}^d a_{jk} e^{-v_k v} \mathbb{P}(\underline{X}_i(T_{\vartheta_i+q_i}) + L_{ij} \geq u - w - v) \, dv \, d\omega. \quad (23)$$

We then substitute v by $u - w - x$ and recall that $\underline{X}_i(T_{\vartheta_i+q_i})$ and L_{ij} are nonpositive random variables, leading to

$$\begin{aligned} p_{ij}^-(u) &= \int_0^u \mu_i e^{-\mu_i w} \int_{-\infty}^{u-w} \sum_{k=1}^d a_{jk} e^{-v_k(u-w-x)} \mathbb{P}(\underline{X}_i(T_{\vartheta_i+q_i}) + L_{ij} \geq x) \, dx \, d\omega \\ &= \int_0^u \mu_i e^{-\mu_i w} \int_{-\infty}^0 \sum_{k=1}^d a_{jk} e^{-v_k(u-w-x)} \mathbb{P}(\underline{X}_i(T_{\vartheta_i+q_i}) + L_{ij} \geq x) \, dx \, d\omega. \end{aligned}$$

Pulling the sum in front of the integrals leads to a sum in which the two integrals factorize. That is, we obtain

$$p_{ij}^-(u) = \sum_{k=1}^d a_{jk} \int_{-\infty}^0 e^{v_k x} (1 - \mathbb{P}(\underline{X}_i(T_{\vartheta_i+q_i}) + L_{ij} < x)) \, dx \cdot \int_0^u \mu_i e^{-\mu_i w} e^{-v_k(u-w)} \, d\omega. \quad (24)$$

Now, we can rewrite the first integral in this expression using (1) and Proposition 1, yielding

$$\begin{aligned} \int_{-\infty}^0 e^{v_k x} (1 - \mathbb{P}(\underline{X}_i(T_{\vartheta_i+q_i}) + L_{ij} < x)) \, dx &= \frac{1}{v_k} \mathbb{E}\left(e^{v_k \underline{X}_i(T_{\vartheta_i+q_i})}\right) \lambda_{ij}(-v_k) \\ &= \frac{1}{v_k} \left(\frac{\Psi_i(\vartheta_i + q_i) - v_k}{\vartheta_i + q_i - \Phi_i(v_k)} \frac{\vartheta_i + q_i}{\Psi_i(\vartheta_i + q_i)} \right) \lambda_{ij}(-v_k) \\ &= \frac{1}{v_k} \left(\frac{\vartheta_i + q_i}{\vartheta_i + q_i - \Phi_i(v_k)} \frac{\mu_i - v_k}{\mu_i} \right) \lambda_{ij}(-v_k). \end{aligned}$$

Furthermore, we note for the second integral of (24) that

$$\int_0^u \mu_i e^{-\mu_i w} e^{-v_k(u-w)} \, d\omega = \frac{\mu_i}{\mu_i - v_k} (e^{-v_k u} - e^{-\mu_i u}).$$

Combining these, we conclude that

$$p_{ij}^-(u) = \sum_{k=1}^d c_{jk} \frac{\vartheta_i + q_i}{\vartheta_i + q_i - \Phi_i(v_k)} (e^{-v_k u} - e^{-\mu_i u}) \lambda_{ij}(-v_k). \quad (25)$$

It can be seen that the μ_i differ from the v_k because, if they were equal for some pair (i, k) , then $p_i(u)$ would have a term that is constant in u , thus violating its form given in (22). In Appendix A.2, an alternative, probabilistic proof of (25) is given.

We now focus on finding the values of the coefficients c_{ik} for $i, k = 1, \dots, d$. Observe that we have two alternative ways of writing $p_i(u)$: Representation (22) and a representation based on (20) and (25). Note that both are linear combinations of $e^{-\mu_i u}$ and $e^{-v_1 u}, \dots, e^{-v_d u}$. The weights corresponding to each of these $d + 1$ exponentials should match, thus providing equations that impose constraints on the c_{ik} .

- Focusing on the terms corresponding to $e^{-v_k u}$ for $k = 1, \dots, d$, we, thus, obtain the equations

$$c_{ik} = \sum_{j \neq i} \left(q_{ij} \frac{1}{\vartheta_i + q_i - \Phi_i(v_k)} \lambda_{ij}(-v_k) \right) c_{jk}, \quad (26)$$

where, as observed earlier, $\sum_{k=1}^d c_{ik} = 1$.

- Regarding the terms corresponding to $e^{-\mu_i u}$, recalling that μ_i differs from all the v_k , we should have

$$1 - \sum_{k=1}^d \sum_{j \neq i} \left(q_{ij} \frac{1}{\vartheta_i + q_i - \Phi_i(v_k)} \lambda_{ij}(-v_k) \right) c_{jk} = 0.$$

This equation holds true if (26) applies, which can be seen by recognizing the left-hand side as $1 - \sum_{k=1}^d c_{ik}$ (as the obvious consequence of $\sum_{k=1}^d c_{ik} = 1$). In other words, this equation does not provide any additional information.

We now observe that (26) is equivalent to, for $i, k = 1, \dots, d$,

$$\sum_{j=1}^d m_{ij}(v_k) c_{jk} = 0. \tag{27}$$

We reassuringly notice from (27) that the matrix $M(v_k)$ is singular for all $k = 1, \dots, d$, and hence, that the v_k are indeed the solutions to $\det M(v) = 0$.

As was done in the spectrally positive case, our result can be rewritten in a more compact vector/matrix form. In particular, to find the c_{jk} , it is enough to solve, for $k = 1, \dots, d$, the matrix-vector equation

$$M(v_k) c_k = \mathbf{0},$$

where $c_k := (c_{1k}, \dots, c_{dk})^\top$, subject to $C \mathbf{1} = \mathbf{1}$. The following theorem summarizes these findings.

Theorem 3. *Under Property (A), $p_i(u)$ satisfies*

$$p_i(u) = \sum_{k=1}^d c_{ik} e^{-v_k u},$$

for $i = 1, \dots, d$. Here, for $k = 1, \dots, d$, the vectors c_k solve $M(v_k) c_k = \mathbf{0}$ subject to $C \mathbf{1} = \mathbf{1}$.

Remark 2. We briefly comment on the case in which some solutions of $\det M(v) = 0$ have multiplicity larger than one. For instance in the case of a root with multiplicity two, suppose that, for some $k_1 \neq k_2$, $v_{k_1} = v_{k_2} = v$, giving rise to terms in (22) proportional to e^{-vu} and $u e^{-vu}$. Finding the associated weights works effectively as pointed out: use Identity (20) to find two alternative expressions for $p_i(u)$ and then equate the terms proportional to $u e^{-vu}$ so as to obtain linear equations for these coefficients (in addition to equating all terms proportional to $e^{-v_k u}$). For an in-depth treatment of these multiplicity issues, we again refer to D’Auria et al. (2010).

4.2. Subordinator Case

We now consider the case in which some of the states of the background process correspond to a nonincreasing subordinator. Let i be in the set of states corresponding to nonincreasing subordinators, denoted by S^\downarrow . For $i \in S^\downarrow$, $Z_i = 0$ with positive probability.

The structure of this section is similar to that of the nonsubordinator case, the main difference being that now the MAP cannot cross positive levels (in the upward direction, that is) while the background process is in $i \in S^\downarrow$. Therefore, the following decomposition applies for $u > 0$:

$$p_i(u) = \sum_{j \neq i} \frac{q_{ij}}{\vartheta_i + q_i} p_{ij}^-(u), \tag{28}$$

with

$$p_{ij}^-(u) := \mathbb{P}(X_i(T_{\vartheta_i + q_i}) + L_{ij} + Z_j \geq u).$$

Regarding the zeroes of $\det M(v)$, we make a similar claim and assumption as in Section 4.1. Let $d^\circ := |S \setminus S^\downarrow|$ be the number of states that do not correspond to a nonincreasing subordinator. Then, by Lemma 3, we know that $\det M(v)$ has d° zeroes with a positive real part, say $(v_k)_{k \in S^\downarrow}$. In our analysis, we impose Property (A’).

$$\text{The } d^\circ \text{ zeroes of } \det M(v) \text{ with a positive real part are distinct.} \tag{A’}$$

The case of zeroes with higher multiplicities can be dealt with as discussed in Remark 2, and the case that $J(\cdot)$ is not irreducible is covered by Section 4.3.

Relying on the same reasoning as in Section 4.1, under (A’), we again have, because of the first-passage process being a MAP and the first statement of Ivanovs (2011), corollary 4.21, that $p_i(u)$ is a linear combination of exponential terms, whereas the number of such terms now equals d° . Concretely, for $u > 0$ and $i = 1, \dots, d$,

$$p_i(u) = \sum_{k \in S^\downarrow} c_{ik} e^{-v_k u}. \tag{29}$$

To identify the coefficients c_{ik} , it proves worthwhile to further study $p_{ij}^-(u)$. In particular, for any $j = 1, \dots, d$, conditioning on the value of Z_j yields

$$p_{ij}^-(u) = \sum_{k \in S^\downarrow} \int_u^\infty c_{jk} v_k e^{-v_k v} \mathbb{P}(X_i(T_{\vartheta_i + q_i}) + L_{ij} \geq u - v) dv.$$

Then, we subsequently use Relation (1) and Proposition 1 to obtain

$$p_{ij}^-(u) = \sum_{k \in S^\downarrow} c_{jk} e^{-v_k u} \mathbb{E}\left(e^{v_k X_i(T_{\vartheta_i+q_i})}\right) \lambda_{ij}(-v_k) = \sum_{k \in S^\downarrow} c_{jk} \frac{\vartheta_i + q_i}{\vartheta_i + q_i - \Phi_i(v_k)} e^{-v_k u} \lambda_{ij}(-v_k),$$

such that, in combination with (28), we have

$$p_i(u) = \sum_{j=i} \sum_{k \in S^\downarrow} c_{jk} \frac{q_{ij}}{\vartheta_i + q_i - \Phi_i(v_k)} e^{-v_k u} \lambda_{ij}(-v_k). \tag{30}$$

By equating (29) and (30), we, thus, obtain equations that the coefficients should satisfy. As it turns out, doing this for any $i \in S^\downarrow$ and $k \in S^\downarrow$, we again obtain (26). Following the same steps as the ones leading to Theorem 3, we obtain the following result. However, the matrix C now consists of entries c_{ik} with $k \in S^\downarrow$, whereas $c_k := (c_{1k}, \dots, c_{dk})^\top$ as before.

Theorem 4. *Under Property (A'), the tail probability $p_i(u)$ satisfies*

$$p_i(u) = \sum_{k \in S^\downarrow} c_{ik} e^{-v_k u},$$

for $i = 1, \dots, d$. Here, for $k \in S^\downarrow$, the vectors c_k solve $M(v_k) c_k = \mathbf{0}$ subject to $\sum_{k \in S^\downarrow} c_{ik} = 1$ for all $i \in S^\downarrow$.

We note that, because the rows i of C such that $i \in S^\downarrow$ do not add up to one, we have that $\mathbb{P}(Z_i = 0) = 1 - \sum_{k \in S^\downarrow} c_{ik} > 0$.

4.3. Number of Roots with a Positive Real Part

In the previous sections, we provide a recipe to compute the tail probabilities $p_i(u)$ using Lemma 3. The objective of this section is to prove this lemma. To this end, we partition the state space of the background chain in K transient classes (say T_1, \dots, T_K) and L recurrent classes (say R_1, \dots, R_L). We label the classes such that, for $\ell \in \{1, \dots, K\}$, class ℓ refers to T_ℓ , and for $\ell \in \{K+1, \dots, K+L\}$, class ℓ refers to $R_{\ell-K}$. We also order the transient classes as is done in Section 3.3: for any ℓ , T_ℓ has no transitions to other classes $T_{\ell'}$ such that $\ell' \leq \ell$. Furthermore, we let d_ℓ° be the number of states in class ℓ that do not correspond to nonincreasing subordinators, $\ell = 1, \dots, K+L$.

With the introduced ordering of class, the transition rate matrix of the background chain can now be written in the following form:

$$Q = \begin{pmatrix} \bar{Q}_1 & S_{12} & \cdots & S_{1K} & S_{1,K+1} & \cdots & S_{1,K+L} \\ 0 & \bar{Q}_2 & \cdots & S_{2K} & S_{2,K+1} & \cdots & S_{2,K+L} \\ \vdots & \vdots & \ddots & \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & \cdots & \bar{Q}_K & S_{K,K+1} & \cdots & S_{K,K+L} \\ 0 & 0 & \cdots & 0 & Q_{K+1} & \cdots & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & 0 & 0 & \cdots & Q_{K+L} \end{pmatrix}. \tag{31}$$

The block matrices Q_{K+1}, \dots, Q_{K+L} correspond to the recurrent classes and can be interpreted as “true” transition rate matrices of Markov chains of lower dimension in that they have nonnegative entries except on their diagonals and their row sums are all zero. This does not hold for the block matrices $\bar{Q}_1, \dots, \bar{Q}_K$: because they correspond to transient classes, their off-diagonal entries are still nonnegative, but they have at least one strictly negative row sum. The matrices $S_{k,\ell}$ with $k = 1, \dots, K$ and $\ell = K+1, \dots, K+L$ contain nonnegative entries and correspond to transitions from T_k into a different class.

The next step is to construct the matrix $M(v)$ that corresponds to the “rearranged transition matrix” Q . This matrix is “block upper triangular,” which is inherited from the matrix Q . It concretely means that, for appropriately constructed matrices $\bar{M}_1(v), \dots, \bar{M}_K(v)$ and $M_{K+1}(v), \dots, M_{K+L}(v)$ (based on, respectively, $\bar{Q}_1(v), \dots, \bar{Q}_K(v)$ and $Q_{K+1}(v), \dots, Q_{K+L}(v)$),

$$\det M(v) = \det \bar{M}_1(v) \cdots \det \bar{M}_K(v) \det M_{K+1}(v) \cdots \det M_{K+L}(v); \tag{32}$$

here, the matrices $\bar{M}_\ell(v)$ (for $\ell = 1, \dots, K$) correspond to transient classes, whereas the matrices $M_\ell(v)$ (for $\ell = K+1, \dots, K+L$) correspond to recurrent classes. It is clear that $\det M_\ell(v)$, where $\ell = K+1, \dots, K+L$, has d_ℓ° roots with a positive real part as an immediate consequence of Proposition 2. This also holds for $\det \bar{M}_\ell(v)$, where $\ell = 1, \dots, K$, which follows by rewriting the diagonal entries as we did in (17) and (18) in such a way that $\bar{M}_\ell(v)$

has the desired form to apply Proposition 2. Upon combining these, we conclude that $\det M(v) = 0$ has $\sum_{\ell=1}^{K+L} d_{\ell}^{\circ} = d^{\circ}$ zeroes as desired.

Remark 3. As indicated, the computation of the coefficients c_{ik} amounts to solving an eigensystem; see Theorems 3 and 4. However, using the structure of the background chain, in specific cases, this computation can be simplified considerably. Appealing to the factorization of $\det M(v)$ provided in Equation (32), it can be argued that some of the c_{ik} are necessarily equal to zero. In the first place, let v_k be a root of $\det M(v)$ such that $\det M_{\ell}(v_k) = 0$ for some $\ell \in \{K+1, \dots, K+L\}$ (i.e., ℓ corresponds to a recurrent class). If the roots of $\det M(v)$ are simple, this means that v_k does not solve $\det M_{\ell'}(v) = 0$ for $\ell' \in \{K+1, \dots, K+L\}$ with $\ell' \neq \ell$, nor $\det \bar{M}_{\ell'}(v) = 0$ for $\ell' \in \{1, \dots, K\}$. By virtue of the structure of the matrix $M(v)$, which is inherited from the transition rate matrix Q (as given in (31)), we, thus, conclude that $c_{ik} = 0$ for all states i from which the recurrent class ℓ cannot be reached. In the second place, analogously, if v_i is such that $\det \bar{M}_{\ell}(v_i) = 0$ for some $\ell \in \{1, \dots, K\}$ (i.e., ℓ corresponds to a transient class), then $c_{ik} = 0$ for all states k that cannot be reached from this transient class. This reduction procedure makes intuitive sense: informally, the distribution of Z_i cannot be affected by properties of the MAP that correspond to states that cannot be reached from state i .

5. Maximum of a Spectrally One-Sided Lévy Process over a Phase-Type Period

In Lévy fluctuation theory, the focus is predominantly on the evaluation of the distribution of extreme values over exponentially distributed intervals; see, for instance, Proposition 1 for a key result in this context. In the present section, we use our results on the maximum of a killed MAP to determine the distribution of a spectrally one-sided Lévy process over a phase-type distributed time interval.

The practical relevance of working with the class \mathcal{P} of phase-type distributions lies in the fact that any distribution on the positive half-line can be approximated arbitrarily closely by a distribution in \mathcal{P} (Asmussen 2003, theorem III.4.2). The proof of this property reveals that, actually, any distribution on the positive half-line can be approximated arbitrarily closely by elements from a smaller class, namely, the class of mixtures of Erlang distributions. In particular, a deterministic positive number can be approximated by an Erlang distributed random variable with a large number of phases.

This section has two main goals. In Section 5.1, we show how our results on the maximum of a killed spectrally one-sided MAP can be applied to derive the distribution of the maximum of a spectrally one-sided Lévy process over a phase-type distributed time interval. Then, in Section 5.2, we obtain more specific results for the practically relevant class of mixtures of Erlang distributions.

5.1. Translation into the MAP Framework

We start our exposition by interpreting a phase-type distributed random variable as an absorption time in a continuous-time Markov chain. Each element in the class \mathcal{P} is characterized by (i) a finite state space $\{1, \dots, d\}$; (ii) an initial distribution $\alpha \in \mathbb{R}^d$; (iii) a $d \times d$ matrix $T = (t_{ij})_{i,j=1}^d$ with nonpositive diagonal entries, nonnegative off-diagonal entries, and nonpositive row sums; and (iv) a nonnegative exit vector $\mathbf{t} := -T\mathbf{1}$. Note that the $(d+1) \times (d+1)$ matrix

$$\bar{T} := \begin{pmatrix} T & \mathbf{t} \\ \mathbf{0}^{\top} & 0 \end{pmatrix},$$

is a genuine transition rate matrix of a $(d+1)$ -state Markov chain in that its diagonal entries are nonpositive, its off-diagonal entries are nonnegative, and its row sums are equal to zero. The $(d+1)$ st column and row in this matrix correspond to a newly added state $d+1$, which we refer to as the *absorbing state*. Observe that this chain can hit state $d+1$ from any other state according to the exit vector \mathbf{t} . Now, the phase-type random variable corresponding to the preceding instance is the time it takes the expanded Markov chain (with transition rate matrix \bar{T}) to reach the absorbing state, at which the initial state has been sampled according to the distribution α .

We now consider the distribution of the maximum of the spectrally one-sided Lévy process $X(\cdot)$ over a phase-type distributed time interval (being characterized by the initial distribution α and the transition rate matrix \bar{T}). To use the MAP framework that we have been working with in the previous sections, we let $X_1(\cdot), \dots, X_d(\cdot)$ be independent copies of a common spectrally one-sided Lévy process $X(\cdot)$ such that the resulting MAP evolves as this Lévy process. We write $\varphi(\cdot)$ for the Laplace exponent of $X(\cdot)$ in case it is spectrally positive, and we write $\Phi(\cdot)$ for the cumulant generating function of $X(\cdot)$ in case it is spectrally negative. In addition, we let $X_{d+1}(t) \equiv 0$ for all $t \geq 0$. Furthermore, we choose $Q = T + \text{diag}(\mathbf{t})$ and $\vartheta = \mathbf{t}$ such that absorption in state $d+1$ corresponds to killing. In addition, let the jumps of the MAP at transition epochs of the background process, as represented by the random variables L_{ij} , be equal to zero. Observe that, under this construction, with \bar{Z} denoting the maximum

of the Lévy process $X(\cdot)$ over the phase-type interval,

$$\mathbb{P}(\bar{Z} \geq u) = \sum_{i=1}^d \alpha_i \mathbb{P}(Z_i \geq u),$$

where $\mathbb{P}(Z_i \geq u)$ can be analyzed using the techniques for extremes of MAPs as developed earlier in the paper.

5.2. Mixtures of Erlang Distributions

Because, with this distribution class, we can approximate any nonnegative random variable arbitrarily closely, we are particularly interested in the case in which the time interval is a mixture of Erlang distributions. This concretely means that, for some $k \in \mathbb{N}$ and $i = 1, \dots, k$, with probability $p_i \in [0, 1]$ the length of the interval is sampled from an Erlang distribution with shape parameter $d_i \in \mathbb{N}$ and scale parameter $\tau_i > 0$ (obviously requiring $\sum_{i=1}^k p_i = 1$). It takes little thought to conclude that, in order to evaluate the maximum of the Lévy process of such an interval, it suffices to be able to evaluate its maximum over an Erlang distributed time interval (say with parameters $d \in \mathbb{N}$ and $\tau > 0$). This requires us to extend the result of an example from Asmussen and Ivanovs (2018), which focuses on the maximum of Brownian motion (with a given drift and variance parameter) over an Erlang(d, τ) distributed time interval. Specifically, we generalize this result to any spectrally one-sided Lévy process (in which, to avoid trivial cases, we assume that the underlying Lévy process is not a subordinator). Related results on maxima over an Erlang horizon include Boxma and Mandjes (2021), section 5; Dębicki and Mandjes (2015), section IV.1; and Starreveld et al. (2016). In the remainder of this section, we treat both spectrally one-sided cases separately.

5.2.1. Maximum of a Spectrally Positive Lévy Process over an Erlang-Distributed Time Interval. Recall from Theorem 1 that the Laplace–Stieltjes transform of the maximum, $\zeta(\gamma)$, satisfies the linear system $M(\gamma)\zeta(\gamma) = \mathbf{b}(\gamma)$, where the $(d \times d)$ -dimensional matrix $M(\cdot)$ is given by

$$M(\gamma) = \begin{pmatrix} -\tau + \phi(\gamma) & \tau & 0 & \dots & 0 & 0 \\ 0 & -\tau + \phi(\gamma) & \tau & \dots & 0 & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots & \vdots \\ 0 & 0 & 0 & \dots & -\tau + \phi(\gamma) & \tau \\ 0 & 0 & 0 & \dots & 0 & -\tau + \phi(\gamma) \end{pmatrix}.$$

The direct implication of the matrix $M(\cdot)$ being upper triangular is that $\zeta(\gamma)$ as well as the unknown constants ω_i can be solved recursively. A concrete recipe for this could be the following. Defining $m(\gamma) := -\tau + \phi(\gamma)$, we first find $\zeta_d(\gamma) = b_d(\gamma)/m(\gamma)$. Note that the numerator contains the constant ω_d (see (9)), which can be identified using the observation that the (single) zero, say $\gamma^* := \psi(\tau)$, from the denominator should be a zero of the numerator as well. As a next step, we identify $\zeta_{d-1}(\gamma)$ observing that

$$\zeta_{d-1}(\gamma) = \frac{b_{d-1}(\gamma) - \tau \zeta_d(\gamma)}{m(\gamma)} = \frac{m(\gamma) b_{d-1}(\gamma) - \tau b_d(\gamma)}{m^2(\gamma)}.$$

This expression contains the (by now known) constant ω_d as well as the (still unknown) constant ω_{d-1} through the function $b_{d-1}(\gamma)$. However, ω_{d-1} can again be found, noting that the double zero from the denominator (which is again γ^*) is also a double zero of the numerator. Thus, noting that $m(\gamma^*) = 0$, we obtain

$$m'(\gamma^*) b_{d-1}(\gamma^*) - \tau b'_d(\gamma^*) = 0,$$

from which we find the unknown constant ω_{d-1} . We can continue along these lines until we have identified all Laplace–Stieltjes transforms $\zeta_i(\gamma)$ and corresponding constants ω_i for $i = 1, \dots, d$. After a number of computations, one then obtains

$$\zeta_i(\gamma) = \left(1 - \frac{\gamma}{\psi(\tau)}\right) \left(-\frac{\tau}{m(\gamma)}\right)^{d-i+1} - \frac{\gamma}{\psi(\tau)} \sum_{j=1}^{d-i} \left(-\frac{\tau}{m(\gamma)}\right)^j \lim_{v \rightarrow \psi(\tau)} \zeta_{i+j}(v).$$

Because $\zeta_j(\psi(\tau))$ is not well defined for any $j = 1, \dots, d$, we note that $\lim_{v \rightarrow \psi(\tau)} \zeta_j(v)$ can be derived from $\zeta_j(\gamma)$ by L'Hôpital's rule.

5.2.2. Maximum of a Spectrally Negative Lévy Process over an Erlang-Distributed Time Interval. For the spectrally negative case, we follow the line of reasoning used in Section 4. The first observation is that, in this case,

$M(v)$ has a single positive root with multiplicity d , namely, $v := \Psi(\tau)$. Because the geometric multiplicity of $M(v)$ is one, its exponent contains multiples of $u^k e^{-vu}$ for $k = 0, \dots, d - 1$ (cf. Remark 2). Let Z_i be the maximum of the process $X(\cdot)$ when starting in phase i , that is, with still $d - i + 1$ phases ahead (which is equivalent to setting $\alpha_i = 1$). We can represent the density of Z_i by a mixture of $d - i + 1$ Erlang densities, that is,

$$\mathbb{P}(Z_i \in du) = \sum_{k=1}^{d-i+1} a_{ik} \frac{v^k}{(k-1)!} u^{k-1} e^{-vu} du, \tag{33}$$

where the a_{ik} are coefficients that we determine as follows. Evidently, the quantity $v^k/(k-1)!$ could have been incorporated in the coefficient a_{ik} , but as it turns out, not doing so makes the formulas slightly cleaner. We already mentioned that $a_{d1} = 1$ as the maximum of a spectrally negative Lévy process over an exponentially distributed interval is exponentially distributed with parameter $\Psi(\tau)$ (see Proposition 1). As a consequence of (33), we find that

$$p_i(u) = \int_u^\infty \mathbb{P}(Z_i \in dx) = \sum_{k=1}^{d-i+1} a_{ik} e^{-vu} \sum_{m=0}^{k-1} \frac{v^m u^m}{m!}. \tag{34}$$

Similar to the strategy followed in Section 4, we now derive a second expression for $p_i(u)$, also in terms of the a_{ik} , which, in combination with (34), allows us to determine these unknown coefficients. To this end, note that, in our setting, (20) implies, for $i = 1, \dots, d - 1$, that

$$p_i(u) = e^{-vu} + p_{i+1}^-(u), \tag{35}$$

where

$$p_{i+1}^-(u) = \int_0^u v e^{-vw} \mathbb{P}(\underline{X}(T_\tau) + Z_{i+1} \geq u - w) dw. \tag{36}$$

Conditioning on the value v of Z_{i+1} , using (33), and substituting $x = u - w - v$, we, thus, obtain

$$\begin{aligned} p_{i+1}^-(u) &= \int_0^u v e^{-vw} \int_{u-w}^\infty \sum_{k=1}^{d-i} a_{i+1,k} \frac{v^k}{(k-1)!} v^{k-1} e^{-vv} \mathbb{P}(\underline{X}(T_\tau) \geq u - w - v) dv dw \\ &= \int_0^u e^{-vw} \int_{-\infty}^0 \sum_{k=1}^{d-i} a_{i+1,k} \frac{v^{k+1}}{(k-1)!} (u - w - x)^{k-1} e^{-v(u-w-x)} \mathbb{P}(\underline{X}(T_\tau) \geq x) dx dw. \end{aligned}$$

It then follows from an application of the binomial theorem that

$$\begin{aligned} p_{i+1}^-(u) &= \int_0^u \int_{-\infty}^0 \sum_{k=1}^{d-i} a_{i+1,k} \frac{v^{k+1}}{(k-1)!} \sum_{\ell=0}^{k-1} \binom{k-1}{\ell} (u-w)^\ell (-x)^{k-1-\ell} e^{-v(u-x)} \mathbb{P}(\underline{X}(T_\tau) \geq x) dx dw \\ &= \sum_{k=1}^{d-i} a_{i+1,k} \sum_{\ell=0}^{k-1} e^{-vu} \int_0^u \frac{v^{\ell+1} (u-w)^\ell}{\ell!} dw \cdot \int_{-\infty}^0 \frac{v^{k-\ell} (-x)^{k-1-\ell}}{(k-1-\ell)!} e^{vx} \mathbb{P}(\underline{X}(T_\tau) \geq x) dx \\ &= \sum_{k=1}^{d-i} a_{i+1,k} \sum_{\ell=0}^{k-1} G_\ell(u) H_{k-1-\ell}, \end{aligned} \tag{37}$$

with

$$G_m(u) := e^{-vu} \int_0^u \frac{v^{m+1} w^m}{m!} dw \text{ and } H_m := \int_{-\infty}^0 \frac{v^{m+1} (-x)^m}{m!} e^{vx} \mathbb{P}(\underline{X}(T_\tau) \geq x) dx.$$

For an alternative, probabilistic derivation of (37), we refer to Appendix A.3. We proceed by evaluating the objects $G_m(u)$ and H_m . The former can be rewritten as

$$G_m(u) = e^{-vu} \frac{v^{m+1}}{m!} \int_0^u w^m dw = e^{-vu} \frac{v^{m+1}}{(m+1)!} u^{m+1}.$$

Deriving a closed-form expression for H_m is significantly more involved. The following lemma proves useful in this regard.

Lemma 4. For any nonnegative random variable A and $m = 1, 2, \dots$,

$$\int_{-\infty}^0 (-x)^m e^{vx} \mathbb{P}(-A \geq x) dx = m! v^{-m-1} - h_m,$$

where $h_m := \int_0^\infty x^m e^{-\nu x} \mathbb{P}(A > x) dx$. In addition, h_m satisfies the recursion

$$h_m = \frac{m}{\nu} h_{m-1} - \frac{1}{\nu} h_m^\circ,$$

with $h_m^\circ := \mathbb{E}(A^m e^{-\nu A})$ and $h_0 = \frac{1}{\nu}(1 - \mathbb{E}(e^{-\nu A}))$.

Proof. The first claim is verified by observing that $\mathbb{P}(-A \geq x) = 1 - \mathbb{P}(A > -x)$. The second claim, the recursive relation for h_m , follows by applying integration by parts. Namely, for $m = 1, 2, \dots$,

$$\begin{aligned} h_m &:= \int_0^\infty x^m e^{-\nu x} \mathbb{P}(A > x) dx = -\frac{1}{\nu} \int_0^\infty x^m \mathbb{P}(A > x) d(e^{-\nu x}) \\ &= \int_0^\infty \frac{e^{-\nu x}}{\nu} d(x^m \mathbb{P}(A > x)) = \frac{m}{\nu} h_{m-1} - \frac{1}{\nu} h_m^\circ. \end{aligned}$$

Finally, the expression for h_0 results from the definition of h_m in combination with (1). This completes the proof. \square

In our analysis, we apply this lemma to the case in which $A = -\underline{X}(T_\tau)$ so that

$$h_m^\circ = \mathbb{E}(A^m e^{-\nu A}) = (-1)^m \frac{d^m}{d\nu^m} \mathbb{E}(e^{-\nu A}) = (-1)^m \frac{d^m}{d\nu^m} \left(\frac{\Psi(\tau) - \nu}{\tau - \Phi(\nu)} \frac{\tau}{\Psi(\tau)} \right).$$

In conclusion,

$$H_m = 1 - \frac{\nu^{m+1}}{m!} h_m,$$

where, for $n = 1, \dots, m$,

$$h_n = \frac{n}{\nu} h_{n-1} - \frac{1}{\nu} h_n^\circ, \quad \text{and} \quad h_0 = \frac{1}{\nu} \left(1 - \frac{\Psi(\tau) - \nu}{\tau - \Phi(\nu)} \frac{\tau}{\Psi(\tau)} \right).$$

At this point, we have obtained two expressions for $p_i(u)$, both in terms of a sum whose summands are proportional to $e^{-\nu u}$, $u e^{-\nu u}$, \dots , $u^{d-i} e^{-\nu u}$. Equating these gives a linear system from which the coefficients a_{ik} can be solved. In this context, it is noted that, conveniently, solving this linear system allows a *recursive* solution procedure. To see this, first recall that $a_{d1} = \nu$. Then, $a_{d-1,1}$ and $a_{d-1,2}$, appearing in (34) for $i = d - 1$, are, by (37), expressed in terms of a_{d1} . Then, along similar lines, $a_{d-2,1}$, $a_{d-2,2}$ and $a_{d-2,3}$ are expressed in terms of $a_{d-1,1}$ and $a_{d-1,2}$ and so on.

6. Numerical Experiments

In the previous sections, we develop theory on the distribution of the maximum of a spectrally one-sided MAP. We now discuss some practical issues concerning the implementation of our findings, the determination of the role of the model parameters, and the application of our results in a practical context. So as to cover these three issues, we consider three experiments: the first highlights the impact of the structure of the background process, the second focuses on the maximum of a Lévy process in Erlang-distributed time intervals, and the third is motivated by a problem in risk theory.

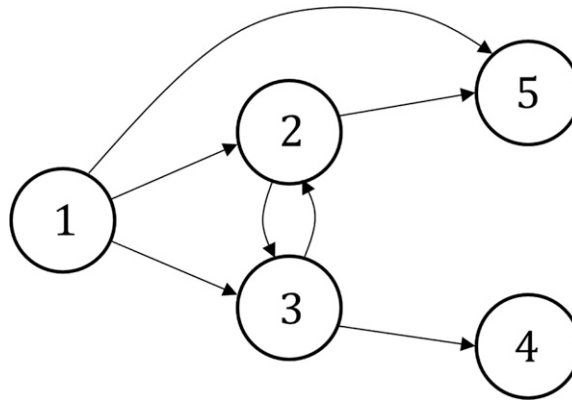
6.1. Impact of the Chain Structure of the Background Process

In this first experiment, we consider a spectrally positive MAP in which the background chain $J(\cdot)$ has the structure shown in Figure 3 (where $q_{ij} > 0$ when there is an arrow from i to j and $q_{ij} = 0$ otherwise). With states 4 and 5 being absorbing, the background process is clearly not irreducible. As a consequence, we cannot use results from the existing literature and have to rely on the results found in Section 3.

Our goal is to evaluate $\zeta(\gamma)$, that is, the vector of Laplace–Stieltjes transforms of the Z_i . This vector is the solution of the matrix Equation (14), in which we follow the procedure developed in Section 3.3 to determine the unknown constants ω_i . We first categorize the communicating classes of $J(\cdot)$ in layers: using the notation from Section 3.3, we have $C_0 = R = \{4, 5\}$, $C_1 = \{2, 3, 4, 5\}$ and $C_2 = \{1, 2, 3, 4, 5\}$. Note that, even though the communicating class $\{1\}$ has a transition to the recurrent state 5, it belongs only to C_2 because it also has transitions into C_1 . Following our procedure, we consecutively evaluate $\zeta_i(\gamma)$ for $i \in \{4, 5\}$, then for $i \in \{2, 3\}$, and finally for $i = 1$.

Using this approach, we now consider an example MAP with the background chain structure given in Figure 3. We let $X_1(\cdot), X_2(\cdot), X_3(\cdot), X_5(\cdot)$ correspond to standard Brownian motions, and $X_4(\cdot)$ to a gamma process (i.e., Lévy process with independent gamma distributed increments) with jump intensity two and jump size

Figure 3. Example Background Chain



parameter two so that

$$\varphi_4(\alpha) = 2 \log\left(\frac{2}{\alpha + 2}\right).$$

Additionally, we let the background be governed by the transition rate matrix

$$Q = \begin{pmatrix} -4 & 1 & 2 & 0 & 1 \\ 0 & -20 & 1 & 0 & 19 \\ 0 & 1 & -20 & 19 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{pmatrix},$$

and we set $L_{ij} = 0$ for all i, j , so there are no jumps at transition epochs of the background process. Finally, we consider the setting in which $\vartheta = (0,0,0,0.5,0.5)^\top$, meaning that killing only happens in states 4 and 5. For these model parameters, we plot in Figure 4 the density functions $f_i(\cdot)$ of the Z_i .

We comment on a few aspects pertaining to Figure 4 that illustrate the impact of the chain structure. First, from the given transition rate matrix, it is clear that, if $J(0) = 2$, then the process likely ends up in state 5. This explains why the densities of Z_2 and Z_5 behave similarly. A similar reasoning applies to Z_3 and Z_4 . Also, Z_4 and Z_5 are “closer to being killed” than Z_2 and Z_3 and, therefore, have more probability mass close to zero. Finally, notice that, from initial state 1, absorption in state 4 or 5 is about equally likely, resulting in a density function that roughly behaves as the average of the two pairs mentioned.

Figure 4. Probability Density Functions $f_1(\cdot), \dots, f_5(\cdot)$, Corresponding to Z_1, \dots, Z_5 with the Model Parameters as Specified in Section 6.1

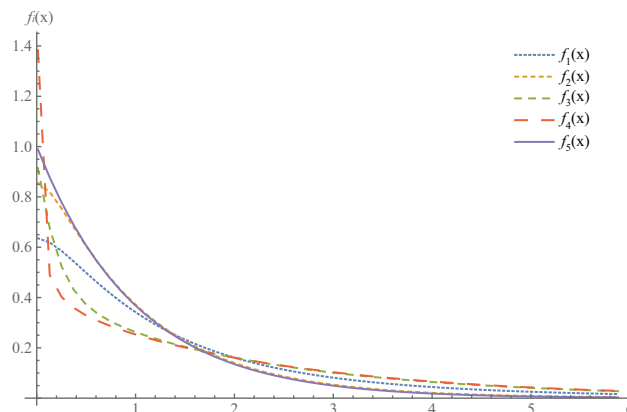
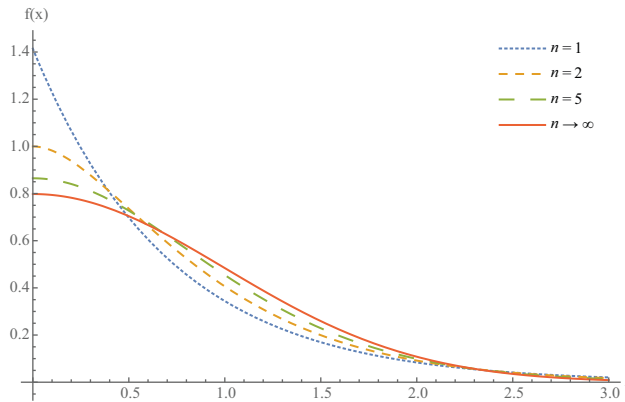


Figure 5. Density Functions $f(\cdot)$ of the Maximum of a Standard Brownian Motion over an n -Phase Erlang-Distributed Interval with Mean One (for $n = 1, 2, 5$) with the Solid Line Representing Their Counterpart for a Deterministic Interval of Length One (as $n \rightarrow \infty$)



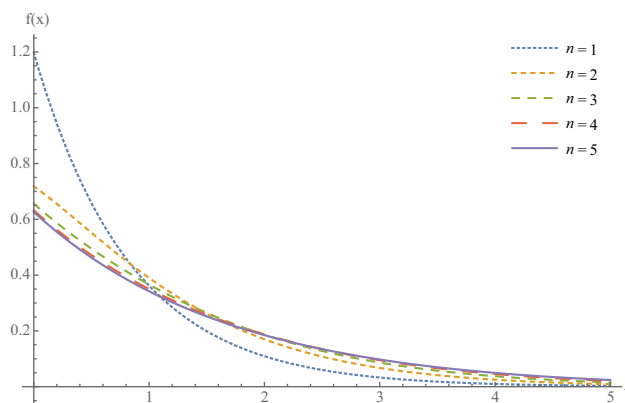
6.2. Maximum of a Lévy Process in an Erlang-Distributed Time Interval

This example focuses on the distribution of the maximum of a Lévy process over a time interval of Erlang-distributed length, applying the theory of Section 5. We choose the model parameters as pointed out in Section 5.1. That is, we let background state i represent the i th phase of the interval. With n denoting the total number of phases, we let $q_{i,i+1} = \frac{1}{n}$ for all $i = 1, \dots, n-1$ and $\vartheta = (0, 0, \dots, 0, \frac{1}{n})^T$. This way, the mean interval length equals unity, and killing occurs only in the last phase. Because phase transitions should not affect the Lévy process, we take $L_{i,i+1} = 0$ for all $i = 1, \dots, n-1$. We particularly study the impact of the number of phases on the distribution of the maximum of the Lévy process, bearing in mind the Erlang distribution’s capability of approximating a deterministic number. Indeed, our Erlang random variable converges to the deterministic value one as $n \rightarrow \infty$, and this experiment serves to get insight into the maximum of a Lévy process during a deterministic time interval.

We first consider the case of $X(\cdot)$ being a standard Brownian motion, noting that, for this instance, we know that its maximum in a deterministic interval has a half-normal distribution (i.e., the distribution of the absolute value of a normally distributed random variable). Figure 5 shows the corresponding density functions for $n = 1, 2, 5$ phases as well as its limiting counterpart. The figure confirms that the densities converge to the limit, at which the curve for $n = 5$ already produces a reasonable fit.

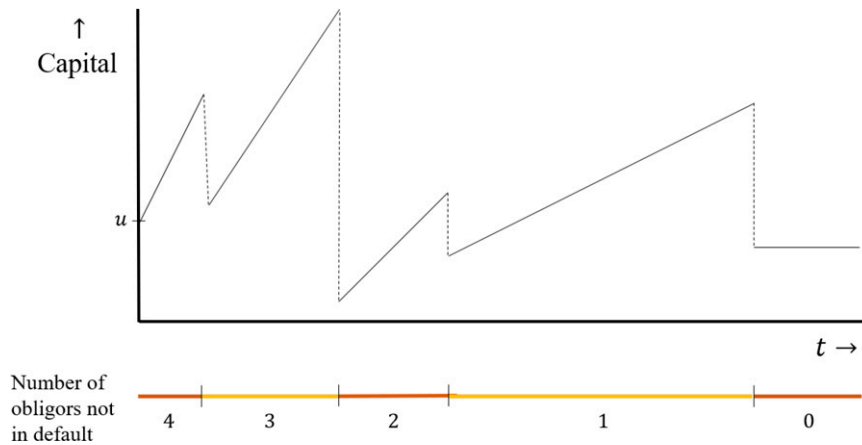
We proceed with an example in which the distribution of the maximum over a deterministic time horizon is not known. Let $X(\cdot)$ be the independent sum of (i) a standard Brownian motion that is increased by a positive

Figure 6. Density Functions $f(\cdot)$ of the Maximum of the Lévy Process $X(\cdot)$ over an n -Phase Erlang Distributed Interval with Mean One (for $n = 1, \dots, 5$)



Notes. Here, $X(\cdot)$ is the sum of a standard Brownian motion with positive drift and a compound Poisson process with negative Erlang-distributed jump sizes.

Figure 7. Sample Path of the Capital Surplus of the Insurance Company with Four Obligor



drift one and (ii) a compound Poisson process with arrival rate one and Erlang(2, 2) distributed jumps in the negative direction (thus, rendering the process spectrally negative). Figure 6 illustrates the (fast) convergence of the density functions as n grows, thus providing us with a way to approximate the distribution of the maximum of $X(\cdot)$ evaluated over a deterministic interval.

6.3. Risk Model

The last example is a special case of the model discussed in Delsing and Mandjes (2021) and is motivated by applications in credit risk. The process of interest is the capital of an insurance company over time with a finite number of obligors n . Each obligor independently goes into default after an exponentially distributed time with mean one. When going into default, the obligor makes a claim of exponential size with mean one and immediately ends the contract with the insurance company (i.e., leaves the system). Each obligor not gone into default pays premiums at rate r per time unit. Figure 7 shows a possible sample path of the process. We wish to quantify the ruin probability, that is, the probability that the capital of the insurance company eventually hits zero given some initial reserve $u \geq 0$.

This model can be cast in our framework as follows. Let background state i represent the number of obligors that have not yet gone into default. Then, the transition rates of the background chain are $q_{i,i-1} = i$ for $i = 1, \dots, n$ (all other transition rates are zero). As we are interested in the all-time ruin probability, we let $\vartheta = 0$. Observe that the ruin probability depends on the minimum of the process, in which the results in this paper are in terms of the maximum, but this is easily remedied by flipping the sign. Concretely, we choose $X_i(t) = -irt$ for $t \geq 0, i = 0, \dots, n$, and we have positive jumps $L_{i,i-1}$ of exponentially distributed size with mean one for $i = 1, \dots, n$. The ruin probability with initial capital u is now given by $\mathbb{P}(Z_n \geq u)$.

In an example with (initially) four obligors, Figure 8 shows how the ruin probability depends on the initial capital u and premium rate r . The ruin probability is decreasing in both u and r , and our techniques can be used to assess the corresponding sensitivities. Clearly, one can trade off u and r : when reducing the premium rate r , a higher initial surplus u is needed to guarantee a given ruin probability. This trade-off is illustrated in Figure 9.

Figure 8. Ruin Probability per Initial Capital for a Few Different Premium Rates r

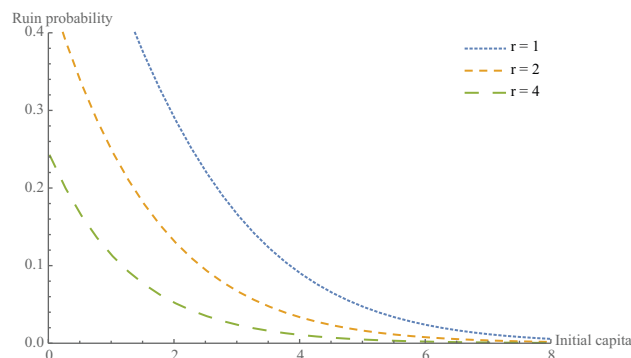
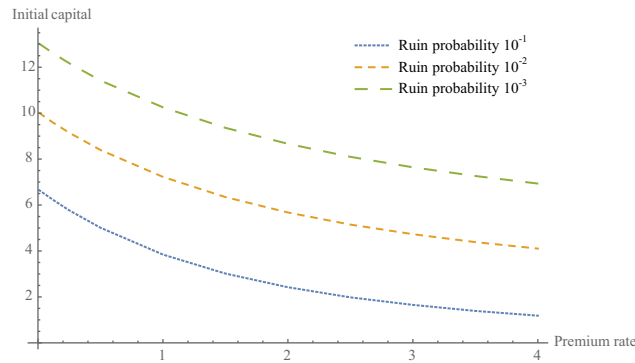


Figure 9. Relation Between Premium Rate and Initial Capital for a Few Fixed Ruin Probabilities



7. Directions for Further Research

In this final section, we discuss two directions for further research.

7.1. Adding Jumps in the Opposite Direction

A possible extension to our model could be the inclusion of phase-type distributed jumps in the “opposite” direction to the model. More concretely, in our spectrally positive setup, we would allow phase-type negative jumps (in both the Lévy processes and the jumps at transition epochs of the background process), and analogously, in our spectrally negative setup we would allow phase-type positive jumps. Alternatively, one could consider jumps in the opposite direction whose distribution has a rational Laplace transform (rather than being a phase-type distribution). Extensions of this type are in line with earlier analyses, such as Lewis and Mordecki (2008), being conceptually relatively straightforward but requiring a substantial amount of additional notation. Relative to the spectrally one-sided MAPs that we consider in the present paper, models that include phase-type jumps in the opposite direction are significantly more general; we recall that any positive random variable can be approximated arbitrarily closely by a phase-type distributed random variable (Asmussen 2003, theorem III.4.2).

7.2. Wiener–Hopf-Type Results Under Nonirreducibility

In the present paper, the focus is on spectrally one-sided MAPs, the underlying aim being the derivation of computable quantities. As a result, our analysis provides expressions for the distribution of the maximum in terms of the model primitives. In Ivanovs (2017), general MAPs (i.e., without any assumptions on the direction of the jumps but still requiring that $J(\cdot)$ is irreducible) are considered, leading to a Wiener–Hopf-type decomposition. The price to be paid, however, is that the characterization of the distribution of the maximum is considerably more implicit than in the spectrally one-sided cases. An interesting question is to what extent the results in Ivanovs (2017) for the extrema carry over to the case that $J(\cdot)$ is not irreducible.

Acknowledgments

The authors thank Onno Boxma (Eindhoven University of Technology, the Netherlands) and Jevgenijs Ivanovs (Aarhus University, Denmark) for providing useful remarks and references.

Appendix. Probabilistic Arguments

In this appendix, we present alternative, probabilistic derivations of some equations in this paper.

A.1. Alternative Derivation of Equation (6)

Starting from state i , we once again condition on whether the MAP is killed before leaving state i or the background process jumps to some state $j \neq i$. This leads to

$$Z_i = \begin{cases} \bar{X}_i(T_{\vartheta_i+q_i}) & \text{with probability } \frac{\vartheta_i}{\vartheta_i + q_i}, \\ \bar{X}_i(T_{\vartheta_i+q_i}) + [\underline{X}_i(T_{\vartheta_i+q_i}) + L_{ij} + Z_j]^+ & \text{with probability } \frac{q_{ij}}{\vartheta_i + q_i}, j \neq i. \end{cases}$$

In terms of Laplace–Stieltjes transforms, this is equivalent to

$$\zeta_i(\gamma) = \frac{\vartheta_i}{\vartheta_i + q_i} \kappa_i(\gamma) + \sum_{j \neq i} \frac{q_{ij}}{\vartheta_i + q_i} \kappa_i(\gamma) \mathbb{E} \left(e^{-\gamma(\underline{X}_i(T_{\vartheta_i+q_i}) + L_{ij} + Z_j)^+} \right), \quad (\text{A.1})$$

where, by Proposition 1, the random variable $-\underline{X}_i(T_{\vartheta_i+q_i})$ is exponentially distributed with rate μ_i . We now present a lemma enabling us to evaluate the rightmost Laplace–Stieltjes transform.

Lemma A.1. *Let A be a nonnegative random variable. Then, for $\gamma, \mu \geq 0$, we have*

$$\mathbb{E} \left(e^{-\gamma(A - T_\mu)^+} \right) = \frac{\mu}{\mu - \gamma} \mathbb{E} \left(e^{-\gamma A} \right) - \frac{\gamma}{\mu - \gamma} \mathbb{E} \left(e^{-\mu A} \right).$$

Proof. Applying the standard identity $e^{-x^+} + e^{x^-} = e^{-x} + 1$ to $x = \gamma(A - T_\mu)$, it holds that, using the memoryless property of T_μ ,

$$\begin{aligned} \mathbb{E} \left(e^{-\gamma(A - T_\mu)^+} \right) &= \mathbb{E} \left(e^{-\gamma(A - T_\mu)} \right) + 1 - \mathbb{E} \left(e^{\gamma(A - T_\mu)^-} \right) \\ &= \mathbb{E} \left(e^{-\gamma A} \right) \mathbb{E} \left(e^{\gamma T_\mu} \right) + 1 - \left(\mathbb{P}(A > T_\mu) + \mathbb{P}(A < T_\mu) \mathbb{E} \left(e^{\gamma T_\mu} \right) \right). \end{aligned}$$

The result follows from the fact that $\mathbb{P}(A < T_\mu) = \mathbb{E} \left(e^{-\mu A} \right)$. \square

Equation (6) can now immediately be obtained from (A.1) by using the preceding lemma with $A = L_{ij} + Z_j$ and recalling (2).

A.2. Alternative Derivation of Equation (25)

Observe that Definition (21) is equivalent to

$$p_{ij}^-(u) := \mathbb{P} \left(\bar{X}_i(T_{\vartheta_i+q_i}) < u, \bar{X}_i(T_{\vartheta_i+q_i}) + \underline{X}_i(T_{\vartheta_i+q_i}) + L_{ij} + Z_j \geq u \right).$$

For $k = 1, \dots, d$, let Z_{jk} be independent exponentially distributed random variables with rate ν_k . An alternative formulation of (23) is that, recalling that $\bar{X}_i(T_{\vartheta_i+q_i})$ is exponentially distributed with parameter μ_i ,

$$p_{ij}^-(u) = \sum_{k=1}^d c_{jk} \mathbb{P} \left(\bar{X}_i(T_{\vartheta_i+q_i}) < u, \bar{X}_i(T_{\vartheta_i+q_i}) + \underline{X}_i(T_{\vartheta_i+q_i}) + L_{ij} + Z_{jk} \geq u \right).$$

Because $\underline{X}_i(T_{\vartheta_i+q_i})$ and L_{ij} are nonpositive, the memoryless property of Z_{jk} implies that

$$p_{ij}^-(u) = \sum_{k=1}^d c_{jk} \mathbb{P} \left(\bar{X}_i(T_{\vartheta_i+q_i}) < u, Z_{jk} + \bar{X}_i(T_{\vartheta_i+q_i}) \geq u \right) \cdot \mathbb{P}(Z_{jk} > -(\underline{X}_i(T_{\vartheta_i+q_i}) + L_{ij})).$$

Because Z_{jk} has an exponential distribution with rate ν_k , we have the equality $\mathbb{P}(Z_{jk} > A) = \mathbb{E}(e^{-\nu_k A})$ for any nonnegative random variable A . Combining this observation with the fact that $\bar{X}_i(T_{\vartheta_i+q_i})$ is exponentially distributed with rate μ_i , we then arrive at

$$\begin{aligned} p_{ij}^-(u) &= \sum_{k=1}^d c_{jk} \frac{\mu_i}{\mu_i - \nu_k} (e^{-\nu_k u} - e^{-\mu_i u}) \mathbb{E}(e^{\nu_k \bar{X}_i(T_{\vartheta_i+q_i})}) \lambda_{ij}(-\nu_k) \\ &= \sum_{k=1}^d c_{jk} \frac{\vartheta_i + q_i}{\vartheta_i + q_i - \Phi_i(\nu_k)} (e^{-\nu_k u} - e^{-\mu_i u}) \lambda_{ij}(-\nu_k), \end{aligned}$$

where the transform of $\bar{X}_i(T_{\vartheta_i+q_i})$ is taken from Proposition 1. This completes the alternative derivation of Equation (25).

A.3. Alternative Derivation of Equation (37)

Let U_k be an Erlang random variable with k phases of rate ν , and let $(V_n)_{n \leq k}$ denote the first n phases of U_k . Equation (36) can be written as

$$\begin{aligned} p_{i+1}^-(u) &= \mathbb{P}(\bar{X}_i(T_\tau) \leq u, \bar{X}_i(T_\tau) + \underline{X}_i(T_\tau) + Z_{i+1} > u) \\ &= \sum_{k=1}^{d-i-1} a_{i+1,k} \mathbb{P}(\bar{X}_i(T_\tau) \leq u, \bar{X}_i(T_\tau) + \underline{X}_i(T_\tau) + U_k > u) \end{aligned}$$

because Z_{i+1} is a mixture of $d-i-1$ Erlang random variables (see (33)). Suppose we need to add to $\bar{X}_i(T_\tau)$ exactly $\ell \in \{1, \dots, k\}$ exponential phases of U_k in order to exceed u , that is, $u - \bar{X}_i(T_\tau) \in [V_{\ell-1}, V_\ell)$. The remainder of the ℓ phase is again exponential, so there are $k - \ell + 1$ phases left to negate the nonpositive variable $\underline{X}_i(T_\tau)$. Therefore, we write

$$p_{i+1}^-(u) = \sum_{k=1}^{d-i-1} a_{i+1,k} \sum_{\ell=1}^k \mathbb{P}(\bar{X}_i(T_\tau) + V_{\ell-1} \leq u < \bar{X}_i(T_\tau) + V_\ell, \underline{X}_i(T_\tau) + (V_k - V_\ell + T_\nu) > 0), \quad (\text{A.2})$$

where, by Proposition 1, $\bar{X}_i(T_\tau)$ is exponentially distributed with rate ν . Now note that these two events are independent, and their respective probabilities are

$$\begin{aligned} \mathbb{P}(\bar{X}_i(T_\tau) + V_{\ell-1} \leq u < \bar{X}_i(T_\tau) + V_\ell) &= \int_0^u \mathbb{P}(V_\ell - V_{\ell-1} > u - x) \mathbb{P}(\bar{X}_i(T_\tau) + V_{\ell-1} \in dx) \\ &= \int_0^u e^{-\nu(u-x)} \frac{\nu^\ell x^{\ell-1}}{(\ell-1)!} e^{-\nu x} dx = G_{\ell-1}(u) \end{aligned} \quad (\text{A.3})$$

and

$$\begin{aligned} \mathbb{P}(\underline{X}_i(T_\tau) + (V_k - V_\ell + T_\nu) > 0) &= \int_0^\infty \mathbb{P}(\underline{X}_i(T_\tau) > -x) \mathbb{P}(V_k - V_\ell + T_\nu \in dx) \\ &= \int_0^\infty \mathbb{P}(\underline{X}_i(T_\tau) > -x) \frac{\nu^{k-\ell+1} x^{k-\ell}}{(k-\ell)!} e^{-\nu x} dx = H_{k-\ell}. \end{aligned} \quad (\text{A.4})$$

Combining (A.2), (A.3), and (A.4) finally leads to Equation (37).

References

- Abate J, Whitt W (2006) A unified framework for numerically inverting Laplace transforms. *INFORMS J. Comput.* 18(4):408–421.
- Asmussen S (2003) *Applied Probability and Queues*, 2nd ed. (Springer, New York).
- Asmussen S, Ivanovs J (2018) A factorization of a Lévy process over a phase-type horizon. *Stochastic Models* 34(4):397–408.
- Asmussen S, Kella O (2000) A multi-dimensional martingale for Markov additive processes and its applications. *Adv. Appl. Probab.* 32(2): 376–393.
- Boxma O, Mandjes M (2021) Affine storage and insurance risk models. *Math. Oper. Res.* 46(4):1282–1302.
- Çınlar E (1972) Markov additive processes I. *Zeitschrift für Wahrscheinlichkeitstheorie und Verwandte Gebiete* 24:85–93.
- D’Auria B, Ivanovs J, Kella O, Mandjes M (2010) First passage of a Markov additive process and generalized Jordan chains. *J. Appl. Probab.* 47(4):1048–1057.
- Dębicki K, Mandjes M (2015) *Queues and Lévy Fluctuation Theory* (Springer, New York).
- Delsing G, Mandjes M (2021) A transient Cramér-Lundberg model with applications to credit risk. *J. Appl. Probab.* 58(3):721–745.
- Dieker A, Mandjes M (2011) Extremes of Markov-additive processes with one-sided jumps, with queueing applications. *Methodology Comput. Appl. Probab.* 13:221–267.
- Ivanovs J (2011) One-sided Markov additive processes and related exit problems. Unpublished PhD thesis, University of Amsterdam, Netherlands.
- Ivanovs J (2017) Splitting and time reversal for Markov additive processes. *Stochastic Processes Their Appl.* 127(8):2699–2724.
- Ivanovs J, Boxma O, Mandjes M (2010) Singularities of the matrix exponent of a Markov additive process with one-sided jumps. *Stochastic Processes Their Appl.* 120(9):1776–1794.
- Kyprianou A (2006) *Introductory Lectures on Fluctuations of Lévy Processes with Applications* (Springer, New York).
- Lewis A, Mordecki E (2008) Wiener-Hopf factorization for Lévy processes having positive jumps with rational transforms. *J. Appl. Probab.* 45(1):118–134.
- Neveu J (1961) Une généralisation des processus à accroissements positifs indépendents. *Abhandlungen aus dem Mathematischen Seminar der Universität Hamburg* 25:36–61.
- Starrevel N, Bekker R, Mandjes M (2016) Transient analysis of one-sided Lévy-driven queues. *Stochastic Models* 32(3):481–512.