THE TIME TO RUIN FOR A CLASS OF MARKOV ADDITIVE RISK PROCESS WITH TWO-SIDED JUMPS

MARTIN JACOBSEN,* *University of Copenhagen*

Abstract

We consider risk processes that locally behave like Brownian motion with some drift and variance, these both depending on an underlying Markov chain that is also used to generate the claims arrival process. Thus, claims arrive according to a renewal process with waiting times of phase type. Positive claims (downward jumps) are always possible but negative claims (upward jumps) are also allowed. The claims are assumed to form an independent, identically distributed sequence, independent of everything else. As main results, the joint Laplace transform of the time to ruin and the undershoot at ruin, as well as the probability of ruin, are explicitly determined under the assumption that the Laplace transform of the positive claims is a rational function. Both the joint Laplace transform and the ruin probability are decomposed according to the type of ruin: ruin by jump or ruin by continuity. The methods used involve finding certain martingales by first finding partial eigenfunctions for the generator of the Markov process composed of the risk process and the underlying Markov chain. We also use certain results from complex function theory as important tools.

Keywords: Probability of ruin; time to ruin; undershoot; passage time; martingale; optional sampling; additive process; Rouché's theorem

2000 Mathematics Subject Classification: Primary 60G40; 60J25; 60K15
Secondary 60G44; 60J35; 60J60

1. Introduction

Consider a real-valued process $X = (X_t)_{t \geq 0}$ and assume that $X_0 > 0$. The time to ruin for X, T_r, is the time at which X first becomes strictly negative, and the undershoot is the absolute value of X at the time of ruin: $Y_r = -X_{T_r}$. If $Y_r > 0$ then ruin has occurred through a jump, while if $Y_r = 0$ then the ruin is by continuity.

The main purpose of this paper is to derive explicit expressions for the joint Laplace transform of T_r and Y_r when X is a special type of Markov additive process with two-sided jumps (see (4), below), which is also a special kind of regime-switching Lévy process. By 'explicit' we here mean that the joint Laplace transform, although not available in closed analytic form, is described completely by the relevant complex roots of the relevant – and explicitly given – Cramér–Lundberg equation and the solution to a set of linear equations.

The technique we shall use is similar to that developed in [11]. In that paper, the simple risk model

$$X_t = x_0 + \beta t - \sum_{n=1}^{N_t} U_n \tag{1}$$

was considered, assuming that claims arrive according to a renewal counting process $N = (N_t)_{t \geq 0}$ with interclaim waiting times that are of phase type, while the U_n form an independent, identically distributed sequence of strictly positive claims, also independent of N. Assuming the Laplace transform for the U_n to be a rational function, explicit expressions for the Laplace transform of the time to ruin and the ruin probability were found using a certain family of martingales. These were determined not only by the risk process X itself, but also involved the Markov chain J used to generate the claims arrival process N in an essential manner.

Finding the relevant martingales amounts to finding partial (*not* ordinary) eigenfunctions (see (17), below) corresponding to the eigenvalue θ, $\theta \geq 0$, of the generator of the piecewise-deterministic Markov process (X, J). What will be shown in this paper is that the structure of the relevant eigenfunctions and martingales found in [11] pertains also to the much more general model to be discussed below. Informally, these partial eigenfunctions may be described as linear combinations of exponentials on part of the state space; this structure is a consequence of only considering processes that are locally Lévy. However, the partial eigenfunction method can in principle be used to study passage time problems for a risk process defined as just one coordinate of a general time-homogeneous Markov process.

Many of the existing results in the literature on ruin problems involve an 'extra' Laplace transform: if x_0 is the initial state of the risk process and $p_r(x_0)$ is the corresponding probability of ruin, one does not determine $p_r(x_0)$ directly, but finds the Laplace transform

$$\int_0^\infty e^{-vx} p_r(x) \, dx$$

instead. We stress that in this paper these 'extra' Laplace transforms are avoided. Another difference with much of the literature is that we do not require the use of, e.g. ladder height distributions, Wiener–Hopf factorization methods, or excursion theory.

The model itself (see (4)) is an example of a Markov additive process X (see, e.g. [1, Section II.5, pp. 39–47] for the definition and basic properties; some recent work on passage time distributions for general Markov additive processes by Kyprianou and Palmowski [14] is commented on below), behaving as a Brownian motion with a drift and variance determined by an underlying Markov chain J, which is also used to generate the times at which claims arrive – in particular, just as in (1), claims arrive according to a renewal process with phase-type waiting times.

The simple model (1) with a renewal process for the arrivals of one-sided claims has been studied recently in a number of papers: a particular case was discussed by Dickson and Hipp [7], while Avram and Usábel [3] obtained general distribution results concerning the time to ruin and the undershoot using a method entirely different from that of [11] and the present paper. For earlier work, also see [1, Chapter 5, pp. 131–144].

With Poisson arrivals, (1) is of course a Lévy process (a compound Poisson process plus linear drift). Adding an independent Brownian motion yields another well-studied Lévy process that is also a special case of (4). For this model, as in (4) with both positive and negative jumps allowed, Asmussen *et al.* [2] determined the joint Laplace transform of the time to ruin and the undershoot. For general Lévy processes (with two-sided jumps), some of the earliest results concerning the joint Laplace transform of T_r and the characteristic function of Y_r can be found in [10]. Emery [9] considered the time to ruin for spectrally negative Lévy processes (downward jumps) and also exit times for bounded intervals. Bertoin [5] contains, in particular, a more modern exposition of the classical theory, with a description of the Laplace transform for the time to exit from a finite interval for spectrally negative Lévy processes that are not minus

a subordinator. His description is in terms of the *scale function,* which is defined implicitly in terms of its Laplace transform, and it is this description that has now been generalized by Kyprianou and Palmowski [14] to a general class of Markov additive process with downward jumps, using a matrix-valued scale function determined, of course, via its matrix-valued Laplace transform. Another natural continuation of [5] is found in Winkel [18], who (in the terminology used here), for Lévy processes that are minus general subordinators, described not only the joint distribution of the time to ruin and the undershoot, but also other quantities related to the time of ruin, such as the size of the claim causing ruin and the time elapsed since the last claim.

A variation of the ruin problem considered here involves the study of exit times for a process obtained by reflecting X around its past maximum; see [4], where X is a spectrally negative Lévy process, and [2], where X is a certain regime-switching process with two-sided jumps. We do not treat these reflected processes here.

In the model (4), downward jumps are always present and upward jumps are allowed. It should be noted that there is no difference in the methods we use in the two-sided and the one-sided jump cases – although, admittedly, proving our results in the two-sided case is technically more difficult. At present, there are not very many results when X has two-sided jumps: they include the classical work by Gusak [10], the work by Asmussen *et al.* [2] quoted above, Kou and Wang [13] (who considered Brownian motion with jumps that follow a double exponential distribution), and the very recent work by Tolver Jensen [17] (who considered a shot noise process with double exponential jumps).

In order to treat the general model (4), it is, as already noted, vital that we can determine partial eigenfunctions of the generator of the Markov process (X, J), which in turn yields the martingales required for the main results. The idea of using partial eigenfunctions has certainly appeared before; see, e.g. Paulsen and Gjessing [16], who studied a risk model of a form other than (4). The martingales are martingales for the filtration generated by the Markov process (X, J), but *not* for that generated by X alone. The idea of involving an enlarged filtration is standard and was used by, e.g. Embrechts *et al.* [8] in their study of risk processes of the form (1) with N a Cox process, where they coined the phrase *Markovization* for this useful device.

The general model (4) studied in this paper is introduced in Section 2. The joint Laplace transform

$$\mathrm{E}^{x_0, i_0} \, \mathrm{e}^{-\theta T_\mathrm{r} - \zeta Y_\mathrm{r}}$$

of the time to ruin T_r and the undershoot Y_r, corresponding to an arbitrary initial state (x_0, i_0) for (X, J), is determined in Section 3 for $\theta > 0$ and $\zeta \geq 0$. Taking $\zeta = 0$ and letting $\theta \downarrow 0$ would then yield the probability of ruin, but in Section 4 it is shown how to find the ruin probability directly, a result that is in some sense more difficult than finding the joint Laplace transform! A numerical example is used to illustrate the ease with which ruin probabilities may be computed.

Our main results are Theorem 1 and Theorem 2. In these results, the joint Laplace transform and the ruin probability are respectively decomposed according to whether ruin is caused by a jump or by continuity. In the case of ruin by continuity, we also give a further decomposition according to the state J_{T_r} of the Markov chain J at the time of ruin.

2. The model

We consider a risk process $X = (X_t)_{t \geq 0}$ that is a real-valued Markov additive process defined as follows. Suppose that we are given a time-homogeneous Markov chain $J = (J_t)_{t \geq 0}$ with a finite state space E, and a counting process $N = (N_t)_{t \geq 0}$, with $N_0 \equiv 0$, such that (J, N) is a homogeneous Markov chain with state space $E \times \mathbb{N}_0$ and transition intensities $q_{(i,n),(i',n')}$ for

$i, i' \in E$, $n, n' \in \mathbb{N}_0$, and $(i, n) \neq (i', n')$. The transition intensities are greater than 0 only if $n' = n$ or $n' = n + 1$, in which case

$$q_{(i,n),(j,n)} = q_{ij}, \qquad i \neq j, \tag{2}$$

$$q_{(i,n),(j,n+1)} = \lambda_i a_j, \tag{3}$$

with all $q_{ij} \geq 0$, all $\lambda_i \geq 0$, and all $a_j \geq 0$, with $\sum_j a_j = 1$. Suppose that we are also given a one-dimensional standard Brownian motion $B = (B_t)_{t \geq 0}$, independent of (J, N), and a sequence $(U_n)_{n \geq 1}$ of independent, identically distributed claims that are independent of (J, N, B) and assumed to be nonzero with $U_n > 0$ always possible and $U_n < 0$ allowed. Then, for given constants $(\beta_i)_{i \in E}$, $\beta_i \in \mathbb{R}$, and $(\sigma_i^2)_{i \in E}$, $\sigma_i^2 \geq 0$, subject to the initial condition $X_0 \equiv x_0$, X is given by

$$X_t = x_0 + \int_0^t \beta_{J_s} \, \mathrm{d}s + \int_0^t \sigma_{J_{s-}} \, \mathrm{d}B_s - \sum_{n=1}^{N_t} U_n. \tag{4}$$

Thus, given that $J_t \equiv i$ and N does not jump, X behaves as a Brownian motion with drift β_i and variance σ_i^2 (with $\sigma_i^2 = 0$ allowed, corresponding to $t \mapsto X_t$ being a straight line with slope β_i for $J_t \equiv i$, and there being no jumps in N). The process X is continuous except at the times when N jumps; at these times a claim arrives, forcing a matching jump in X that is downwards if the claim is greater than 0 and upwards if the claim is less than 0. The Markov chain J may jump simultaneously with N but may also have jumps between the jumps in N. The latter jumps are governed by the intensities q_{ij} from (2), while, for each i, λ_i is the intensity for a claim to be triggered when J is in state i, a_i is the probability that J remains in i, and a_j, $j \neq i$, is the probability that J jumps to j simultaneously with the arrival of the claim. Note that, for the chain (J, N), the total intensity for a jump from (i, n) is $q_i + \lambda_i$, where we write $q_i = \sum_{j \neq i} q_{ij}$, $i \in E$.

The model studied in [11] corresponds to the special case in which $\beta_i \equiv \beta$ and $\sigma_i^2 = 0$ for all $i \in E$.

Remark 1. It follows from the above that J is indeed a Markov chain with transition intensities $\tilde{q}_{ij}$, $i \neq j$ (corresponding to true jumps in J), given by

$$\tilde{q}_{ij} = q_{ij} + \lambda_i a_j \tag{5}$$

and

$$\tilde{q}_{ii} = -q_i - \lambda_i (1 - a_i). \tag{6}$$

The process X given by (4) is an example of a Markov additive process. A more general class of process is obtained by allowing X to behave like an arbitrary Lévy process $X^{(i)}$ whenever J is in state i, and allowing the claims U_n to occur whenever J jumps from some i to some $j \neq i$ with a distribution F_{ij} depending on (i, j): with F_U the distribution of the independent, identically distributed nonzero claims from (4), this corresponds to taking

$$F_{ij} = \frac{\lambda_i a_j}{q_i + \lambda_i (1 - a_i)} F_U + \frac{q_i}{q_i + \lambda_i (1 - a_i)} \varepsilon_0$$

for $i \neq j$ (with ε_0 the probability measure degenerate at 0, so that F_{ij} may have an atom at 0), where we introduce $F_{ii} = F_U$.

In the recent paper [14], the authors studied (among other things) the time to ruin T_r for the class of Markov additive process just described, under the assumption that the Lévy processes $X^{(i)}$ are all spectrally negative and the claims are all greater than 0 (downward jumps). They then determined the Laplace transform for T_r in terms of a matrix-valued version of the scale function used by Bertoin [5], implicitly described in terms of its matrix-valued Laplace transform. For the special model given by (4) we obtain more explicit results, in that this and other 'extra' Laplace transforms are avoided.

For the risk process X, we define the time to ruin as

$$T_r = \inf\{t \geq 0 \colon X_t < 0\},$$

with $T_r = \infty$ if $X_t \geq 0$ for all t. Assuming that $X_0 \equiv x_0 > 0$, we shall below determine in particular the Laplace transform for T_r and also describe the joint distribution of $(T_r, Y_r, 1_r)$, where $Y_r = -X_{T_r}$ is the size of the *undershoot* at the time of ruin and 1_r is an indicator specifying whether there is *ruin by jump* (ruin caused by a sufficiently large claim, i.e. $Y_r > 0$) or *ruin by continuity* (ruin caused by X moving continuously through the level 0, i.e. $Y_r = 0$).

The ruin problem will be discussed subject either to the condition $(X_0, J_0) \equiv (x_0, i_0)$ for given, but arbitrary, $x_0 > 0$ and $i_0 \in E$, in which case we respectively write P^{x_0, i_0} and E^{x_0, i_0} for the underlying probability and matching expectation, or subject to $X_0 \equiv x_0 > 0$ with J_0 having distribution $\boldsymbol{a} = (a_i)$ and being independent of B and the U_n, in which case we write $\mathrm{P}^{x_0, \boldsymbol{a}}$ and $\mathrm{E}^{x_0, \boldsymbol{a}}$. If a formula applies to either situation, we just write P and E. It is unproblematic to set up all the probabilities P^{x_0, i_0} on the same space; we may then simply define $\mathrm{P}^{x_0, \boldsymbol{a}} = \sum_{i_0} a_{i_0} \mathrm{P}^{x_0, i_0}$.

Let us comment further on the model. The multiplicative structure of the intensities (3) used is essential in the proof of the main results, Theorems 1 and 2. It implies, in particular, that if $T_n = \inf\{t \geq 0 \colon N_t = n\}$ is the time of arrival of the nth claim (with $T_0 \equiv 0$), then the sequence $(T_n)_{n \geq 1}$ is a (possibly delayed) renewal sequence such that the waiting times $V_n = T_n - T_{n-1}$ are independent and, for $n \geq 2$, independent, identically distributed, and of phase type, with

$$\mathrm{P}(V_n > v) = \boldsymbol{a}^\top \mathrm{e}^{\boldsymbol{Q}_V v} \mathbf{1}, \tag{7}$$

where $\boldsymbol{a}^\top$ is the row vector with elements a_i, $\boldsymbol{Q}_V$ is the *subintensity* matrix with elements

$$q_{V,ij} = \begin{cases} q_{ij}, & i \neq j, \\ -(q_i + \lambda_i), & i = j, \end{cases} \tag{8}$$

and $\mathbf{1}$ is a column vector of 1s. Under $\mathrm{P}^{x_0, \boldsymbol{a}}$, V_1 has the same law as the V_n, $n \geq 2$.

Phase-type distributions are usually described as the distribution of the time to absorption for a Markov chain on a finite state space E with an additional absorbing state. For us the intensity for 'absorption' from i is λ_i, but of course J is not absorbed but returned instantly to E using the entrance law (a_j).

Once a claim has arrived, J will move only through states i such that either $a_i > 0$ or i can be reached by $q_{jj'}$-transitions from some $i' \in E$ with $a_{i'} > 0$. The following basic assumption requires all $i \in E$ to have this property and also for ruin to be possible from any state.

Assumption 1. (i) For any $i \in E$, either $a_i > 0$ or there exist an $n \geq 1$ and $i_0, \ldots, i_n \in E$, with $i_n \equiv i$ and $i_k \neq i_{k-1}$ for all k, such that $a_{i_0} > 0$ and $q_{i_{k-1}, i_k} > 0$ for all k.

(ii) For any $i \in E$, either $\lambda_i > 0$ or there exist an $n \geq 1$ and $i_0, \ldots, i_n \in E$, with $i_0 = i$ and $i_k \neq i_{k-1}$ for all k, such that $\lambda_{i_n} > 0$ and $q_{i_{k-1}, i_k} > 0$ for all k.

An alternative formulation of Assumption 1 is that the Markov chain on E with transition intensities $\tilde{q}_{ij}$ given by (5) and (6) is irreducible for some $\lambda_i > 0$. The assumption implies that there will be an infinity of claims: for any $n \geq 1$, we have $P(V_n = \infty) = 0$.

The reader is reminded that, under Assumption 1, the subintensity matrix $\boldsymbol{Q}_V$ in (8) is nonsingular (see [11, Lemma 1] and also Lemma 1 below), and that, since $\boldsymbol{Q}_V \mathbf{1} = -\boldsymbol{\lambda}$, where $\boldsymbol{\lambda}$ denotes the column vector $(\lambda_i)_{i \in E}$, we have

$$\boldsymbol{Q}_V^{-1}\boldsymbol{\lambda} = -\mathbf{1}. \tag{9}$$

The renewal structure of the claims arrival process is somewhat restrictive. It does, however, allow for a strong dependence between X and N, since, e.g. the behaviour of X between the jumps in N may indicate that J is in a state i with large λ_i, meaning that there is high probability of the short-term arrival of a new claim. (In general, the behaviour of X allows us to distinguish between equivalence classes of states in E according to the equivalence relation '$\sim$', where $i \sim j$ if either $\sigma_i^2 = \sigma_j^2 > 0$ or $\sigma_i^2 = \sigma_j^2 = 0$ and $\beta_i = \beta_j$: to determine the class to which J_t belongs, we look at the path of X in a sufficiently small neighbourhood to the right of t. If X follows a straight line, we deduce that $\sigma_{J_t}^2 = 0$ and read off β_{J_t} as the slope of the line. If X does not follow a straight line, we compute the quadratic variation $[X]$ of X, which necessarily satisfies $[X]_{t''} - [X]_{t'} = \sigma_{J_t}^2(t'' - t')$ in this small neighbourhood, and then read off the value of $\sigma_{J_t}^2$. Thus, if all equivalence classes contain just one state, then J_t is completely determined from the behaviour of X_s for $s \in [t, t + \varepsilon]$ for any $\varepsilon > 0$, say, while if all states belong to the same class, then X contains no precise information about J.)

A final comment on the model for X is that it shares with Lévy processes and general Markov additive processes the following *additivity property*: for arbitrary x_0, x_1, and i_0, under the probability P^{x_0,i_0} the distribution of the process $X + x_1 - x_0$ is the same as the distribution of X itself under P^{x_1,i_0}.

We now introduce some further notation. Let $p = |E| \geq 1$ be the number of states for J (with $p = 1$ allowed). Also, let E_j denote the set of $i \in E$ such that ruin by jump is possible when $J_t = i$, i.e.

$$E_j = \{i \in E : \lambda_i > 0\}, \tag{10}$$

with $E_j \neq \varnothing$ by Assumption 1. Similarly, let E_c denote the states from which ruin by continuity is possible, i.e.

$$E_c = \{i \in E : \sigma_i^2 > 0 \text{ or } \beta_i < 0\}. \tag{11}$$

Here $E_c = \varnothing$ is possible, occurring when $\sigma_i^2 = 0$ and $\beta_i \geq 0$ for all i. We write p_j and p_c for the numbers of elements in E_j and E_c, respectively, i.e.

$$p_j = \sum_{i \in E} 1_{\{\lambda_i > 0\}}, \qquad p_c = \sum_{i \in E} 1_{\{\sigma_i^2 > 0 \text{ or } \beta_i < 0\}}. \tag{12}$$

Now consider the joint process (X, J). This is a homogeneous Markov process, with state space $\mathbb{R} \times E$, adapted to the filtration $(\mathcal{F}_t)$ generated by (B, J, C), where $C_t = \sum_{n=1}^{N_t} U_n$. For sufficiently 'nice' functions $f : \mathbb{R} \times E \to \mathbb{R}$, it has an infinitesimal generator of the form

$$Af(x, i) = \beta_i D_x f(x, i) + \frac{1}{2}\sigma_i^2 D_{xx}^2 f(x, i) + \sum_{j \neq i} q_{ij}(f(x, j) - f(x, i))$$

$$+ \lambda_i \sum_j a_j \int_{-\infty}^{\infty} F_U(\mathrm{d}y)(f(x - y, j) - f(x, i)), \tag{13}$$

for $x \in \mathbb{R}$ and $i \in E$, with F_U denoting the distribution of the claims U_n, as before. (In (13), D_x and D_{xx} denote differentiation once and, respectively, twice with respect to x.)

Next, let $\mathcal{D}$ denote the domain of bounded functions $f : \mathbb{R} \times E \to \mathbb{R}$ such that, for all $i \in E$, $x \mapsto f(x, i)$ is twice continuously differentiable for $x \geq 0$ (but not necessarily for $x < 0$), with both $D_x f(x, i)$ and $D_{xx}^2 f(x, i)$ bounded for $x \geq 0$. Then by Itô's formula, assuming that $X_0 > 0$, for any $t \geq 0$ and $f \in \mathcal{D}$ we have

$$f(X_{T_r \wedge t}, J_{T_r \wedge t}) = f(X_0, J_0) + \int_0^{T_r \wedge t} A f(X_s, J_s) \, \mathrm{d}s + M_t, \qquad (14)$$

where M is an $\mathcal{F}_t$-martingale with $M_0 \equiv 0$. Note that, in the integral, $X_s \geq 0$ and $A f(X_s, J_s)$ is well defined except possibly at the single point $s = T_r$.

Remark 2. Identities like (14) are of course true quite generally for time-homogeneous Markov processes, with M typically a local martingale and, if f and Af are bounded, a true martingale. The particular form of M in (14) is

$$\mathrm{d}M_t = Z_t \, \mathrm{d}B_t + \sum_{j \in E} \int_{\mathbb{R} \setminus 0} S_t^{j,y} (\mu(\mathrm{d}t \times (\{j\} \times \mathrm{d}y)) - \Lambda(\mathrm{d}t \times (\{j\} \times \mathrm{d}y))), \qquad (15)$$

where the processes

$$Z_t = \sigma_{J_{t-}} D_x f(J_{t-}, X_{t-}), \qquad S_t^{j,y} = f(X_{t-} - y, j) - f(X_{t-}, J_{t-})$$

are $\mathcal{F}_t$-predictable. Furthermore, in (15), μ is the random counting measure on $\mathbb{R}_0 \times (E \times \mathbb{R})$ given by

$$\mu([0, t] \times (\{j\} \times \{0\})) = \sum_{0 < s \leq t} 1_{\{J_{s-} \neq J_s = j\}}$$

(counting jumps of J to j with no jump in C) and

$$\mu([0, t] \times (\{j\} \times H)) = \sum_{n=1}^{N_t} 1_{\{J_{T_n} = j, \, U_n \in H\}}$$

(counting jumps in C of sizes belonging to the Borel set H not containing 0, with a possible simultaneous jump of J to j), while Λ is the corresponding $\mathcal{F}_t$-compensating measure, i.e.

$$\Lambda(\mathrm{d}t \times (\{j\} \times \{0\})) = q_{J_{t-}, j} 1_{\{J_{t-} \neq j\}} \, \mathrm{d}t,$$
$$\Lambda(\mathrm{d}t \times (\{j\} \times H)) = \lambda_{J_{t-}} a_j F_U(H) \, \mathrm{d}t.$$

It is standard result from marked point process theory (see, e.g. [6], [15], or [12]) that the last term in (15) is a local martingale. In (14) itself, it may be verified by a direct check that the jumps on the left- and right-hand sides are the same (including, as is necessary, the possible jump at time T_r), while between jumps the verification is a consequence of the standard Itô formula for Brownian motion.

From (14), it follows directly that, for any $\theta \in \mathbb{R}$,

$$\mathrm{e}^{-\theta(T_r \wedge t)} f(X_{T_r \wedge t}, J_{T_r \wedge t}) = f(X_0, J_0) + \int_0^{T_r \wedge t} \mathrm{e}^{-\theta s} (A f(X_s, J_s) - \theta f(X_s, J_s)) \, \mathrm{d}s$$
$$+ \int_0^{T_r \wedge t} \mathrm{e}^{-\theta s} \, \mathrm{d}M_s.$$

If $\theta \geq 0$ then the last term is again a mean-0 martingale, meaning that, for $x_0 > 0$ and $i_0 \in E$,

$$
\begin{aligned}
\mathrm{E}^{x_0, i_0} \, \mathrm{e}^{-\theta(T_r \wedge t)} & f(X_{T_r \wedge t}, J_{T_r \wedge t}) \\
&= f(x_0, i_0) + \mathrm{E}^{x_0, i_0} \int_0^{T_r \wedge t} \mathrm{e}^{-\theta s} (Af(X_s, J_s) - \theta f(X_s, J_s)) \, \mathrm{d}s.
\end{aligned} \tag{16}
$$

Now suppose that $f \equiv f_\theta \in \mathcal{D}$ is a *partial eigenfunction* for A in the sense that

$$
Af(x, i) = \theta f(x, i), \qquad x \geq 0, \ i \in E. \tag{17}
$$

The integral in the last term of (16) then vanishes, the process

$$
(\mathrm{e}^{-\theta(T_r \wedge t)} f(X_{T_r \wedge t}, J_{T_r \wedge t}))_{t \geq 0}
$$

becomes a martingale, and (16) reduces to

$$
\mathrm{E}^{x_0, i_0}[\mathrm{e}^{-\theta T_r} f(X_{T_r}, J_{T_r}); \ T_r \leq t] + \mathrm{E}^{x_0, i_0}[\mathrm{e}^{-\theta t} f(X_t, J_t); \ T_r > t] = f(x_0, i_0). \tag{18}
$$

If we now assume not only that $\theta \geq 0$ but also that $\theta > 0$, and take the limit as $t \to \infty$, dominated convergence yields

$$
\mathrm{E}^{x_0, i_0}[\mathrm{e}^{-\theta T_r} f(X_{T_r}, J_{T_r}); \ T_r < \infty] = \mathrm{E}^{x_0, i_0} \, \mathrm{e}^{-\theta T_r} f(X_{T_r}, J_{T_r}) = f(x_0, i_0), \tag{19}
$$

which is a key identity to be exploited in the sequel.

It is important to stress that these identities rely on finding partial rather than true eigenfunctions for the generator (for which the identity (17) holds for *all* $x \in \mathbb{R}$ and $i \in E$): the choices for f (see (39), below) are such that, although $f(x, i)$ is defined for all $x \in \mathbb{R}$ and $i \in E$, $Af(x, i)$ (which involves $f(y, j)$ with $y < 0$) makes sense and is used only for $x \geq 0$, with the possibility of defining $f(x, i)$ freely and suitably for $x < 0$ then available; something that would be impossible with the true eigenfunctions. That (14) is valid only up to the stopping time T_r is also a consequence of choosing partial eigenfunctions.

As will be argued below, in order to find the partial eigenfunctions we must solve the Cramér–Lundberg equation (30) (and, in the case of downward jumps only, the simpler form (32)) and it is of some interest to see how this relates to the more standard forms from the literature. If $\tilde{X}$ is a Lévy process starting from 0, the Laplace exponent ψ is defined according to the identity

$$
\mathrm{E} \, \mathrm{e}^{z \tilde{X}_t} = \mathrm{e}^{t \psi(z)},
$$

for any $z \in \tilde{A} := \{z \in \mathbb{C} \colon \mathrm{E} \, |\mathrm{e}^{z \tilde{X}_t}| < \infty\}$. The Cramér–Lundberg equation then reads

$$
\psi(z) = \theta \tag{20}
$$

for a given $\theta \geq 0$, and, when using optional sampling and the fact that $(\mathrm{e}^{z \tilde{X}_t - t \psi(z)})_{t \geq 0}$ is a martingale, solving this with $z \equiv z(\theta) \in \tilde{A}$ and $\mathrm{Re}(z(\theta)) \geq 0$ gives

$$
\mathrm{E}[\mathrm{e}^{z(\theta)(\tilde{X}_{\tilde{T}_r} + x_0) - \theta \tilde{T}_r}; \ \tilde{T}_r < \infty] = \mathrm{e}^{z(\theta) x_0} \tag{21}
$$

for any $x_0 > 0$, where $\tilde{T}_r = \inf\{t : \tilde{X}_{\tilde{T}_r} + x_0 < 0\}$. To use this to derive the Laplace transform for $\tilde{T}_r$ requires, first, that we solve (20) with $z(\theta) \in \tilde{A}$ and $\mathrm{Re}(z(\theta)) \geq 0$ and, second, that we deal somehow with the random undershoot $\tilde{X}_{\tilde{T}_r} + x_0$. If $\tilde{X}$ is spectrally negative then $\tilde{A}$

contains all those z with $\mathrm{Re}(z) \geq 0$, which is helpful in solving (20); if, by contrast, $\tilde{X}$ also has upward jumps (exponential, say), to have $z \in \tilde{A}$ may require that we have $\mathrm{Re}(z) < \delta$ for some $\delta > 0$, and the precise structure of these upward jumps may affect whether (20) can be solved at all inside the required domain.

If, in our model (4), X is a spectrally negative Lévy process then we have $\beta_i \equiv \beta$ and $\sigma_i^2 \equiv \sigma^2 \geq 0$ for all i and $U_n > 0$ for all n. By introducing

$$L_U(z) = \int_{(0,\infty)} F_U(\mathrm{d}y)\mathrm{e}^{-zy},$$

which is well defined for those $z \in \mathbb{C}$ with $\mathrm{Re}(z(\theta)) \geq 0$, we find that

$$\psi(z) = \beta z + \tfrac{1}{2}\sigma^2 z^2 - \lambda(1 - L_U(z)). \tag{22}$$

It is a fundamental assumption made below that L_U be a rational function. This, in particular, allows us to define an analytic continuation $\bar{L}_U$ of L_U (analytic except for at finitely many poles z with $\mathrm{Re}(z) < 0$) and we then proceed to solve (20) for z with $\mathrm{Re}(z) < 0$. Of course, we must then argue that it is relevant to solve this 'extended' Cramér–Lundberg equation, which is nontrivial. One advantage is that, for z with $\mathrm{Re}(z) < 0$, the upward jumps in X are no longer a concern as regards the integrability of e^{zX_t}. It should be noted that the Cramér–Lundberg equation (32) reduces to (20) when X is spectrally negative Lévy as just described and ψ in (20) is replaced by the analytic continuation (22).

A final comment we shall make concerns the relation between (21) and (19): the use of the partial eigenfunctions allows us to replace the integrand in (21) with, e.g. the more appealing expression $\tilde{K}\mathrm{e}^{-\theta\tilde{T}_r}$, where $\tilde{K}$ is a constant depending on whether ruin occurs by continuity or by jump.

3. The joint Laplace transform

Consider the risk process given by (4) with fixed initial state $x_0 > 0$, and recall the definitions, (10) and (11), of the sets of states E_j and E_c from which ruin by jump and, respectively, by continuity are possible. Also recall that $E_j \neq \varnothing$ while $E_c = \varnothing$ can occur, and that the number of elements in the two sets are denoted by p_j and p_c, respectively. To prepare for Theorem 1, we shall distinguish $p_c + 1$ different types of ruin corresponding to the events A_j and $A_{c,i}$, $i \in E_c$, where

$$A_j = \{X_{T_r} < 0, \ T_r < \infty\} \tag{23}$$

is the event that ruin occurs by jump, and

$$A_{c,i} = \{X_{T_r} = 0, \ J_{T_r} = i, \ T_r < \infty\} \tag{24}$$

is the event that ruin occurs by continuity with the Markov chain J in state i.

We shall also need the following notation: for $z \in \mathbb{C}$ and $\theta \geq 0$, by $\boldsymbol{Q}(z,\theta) = (q_{ij}(z,\theta))_{i,j\in E}$ we denote the matrix given by

$$q_{ij}(z,\theta) = \begin{cases} \phi_i(z) - q_i - \lambda_i - \theta, & \text{if } i = j, \\ q_{ij}, & \text{if } i \neq j, \end{cases} \tag{25}$$

where

$$\phi_i(z) = \beta_i z + \tfrac{1}{2}\sigma_i^2 z^2$$

is a polynomial of degree less than or equal to 2 associated with the scaled Brownian motion with drift that X follows when J is in state i.

The matrix $\boldsymbol{Q}_V = \boldsymbol{Q}(0,0)$ is the subintensity matrix used in the description of the phase-type distribution of the V_n, $n \geq 2$; cf. (8). Thus, the Laplace transform of the waiting times between claims is

$$L_V(\nu) = \mathrm{E}\,\mathrm{e}^{-\nu V_n} = -\boldsymbol{a}^\top (\boldsymbol{Q}_V - \nu \boldsymbol{I})^{-1} \boldsymbol{\lambda}, \qquad \nu \geq 0, \, n \geq 2. \tag{26}$$

(It is a consequence of Assumption 1 that $\boldsymbol{Q}_V - \nu\boldsymbol{I}$ is nonsingular; see Lemma 1 of [11] or Lemma 1 below. Of course, with the notation used here, $\boldsymbol{Q}_V - \nu\boldsymbol{I} = \boldsymbol{Q}(0,\nu)$.)

The final assumptions that we require concern the distribution F_U of the claims. We must have $F_U(\{0\}) = 0$ to ensure that the $U_n \neq 0$ almost surely (a.s.). It is then natural to split F_U and write

$$F_U(\mathrm{d}u) = \begin{cases} p_U\, F_+(\mathrm{d}u), & u > 0, \\ (1 - p_U) F_-(\mathrm{d}u), & u < 0, \end{cases}$$

where we assume that

$$p_U := F_U((0,\infty))$$

satisfies $0 < p_U \leq 1$. The case $p_U = 1$ is of particular interest and corresponds to the more standard situation in which there are only positive claims, i.e. one-sided downward jumps in X.

Next, write

$$L_+(\nu) = \mathrm{E}[\mathrm{e}^{-\nu U_n} \mid U_n > 0] = \int_{(0,\infty)} F_+(\mathrm{d}u)\, \mathrm{e}^{-\nu u}, \qquad \nu \geq 0,$$

for the Laplace transform of the positive claims and, similarly,

$$L_-(\nu) = \mathrm{E}[\mathrm{e}^{-\nu U_n} \mid U_n < 0] = \int_{(-\infty,0)} F_-(\mathrm{d}u)\, \mathrm{e}^{-\nu u}, \qquad \nu \leq 0,$$

which, like F_-, is relevant only if $p_U < 1$.

A fundamental assumption concerning L_+ is that it be a rational function, i.e. that

$$L_+(\nu) = \mathrm{E}\,\mathrm{e}^{-\nu U_n} = \frac{P_+(\nu)}{R_+(\nu)}, \qquad \nu \geq 0, \tag{27}$$

where P_+ and R_+ are polynomials standardized in such a way that they have no common complex roots and the leading coefficient of R_+ is 1. We write $m \geq 1$ for the degree of R_+ and note that P_+ is necessarily of degree less than or equal to $m - 1$. Below we shall need $P_+(z)$ and $R_+(z)$ for all $z \in \mathbb{C}$, but we remind the reader that the resulting extension of L_+ to $\bar{L}_+(z) = P_+(z)/R_+(z)$, for $z \in \mathbb{C}$, is meaningless as an expectation: the identity

$$\mathrm{E}[\mathrm{e}^{-zU_n} \mid U_n > 0] = \int_{(0,\infty)} \mathrm{e}^{-zu}\, F_+(\mathrm{d}u) = \frac{P_+(z)}{R_+(z)} \tag{28}$$

is guaranteed to hold only for those z with $\mathrm{Re}(z) > -\varepsilon$ for some sufficiently small $\varepsilon > 0$. The fact that (28) is always true if $\mathrm{Re}(z) \geq 0$ is important: it implies that the m roots z of $R_+(z)$ (counted with multiplicity) must satisfy $\mathrm{Re}(z) < 0$, an observation used frequently below. It may be noted that $z \mapsto \bar{L}_+(z)$ is analytic in $\mathbb{C}$ except at finitely many poles, located where R_+ has its roots.

If $p_U < 1$ then the extension

$$L_-(z) = \mathrm{E}[\mathrm{e}^{-zU_n} \mid U_n < 0] = \int_{(-\infty,0)} F_-(\mathrm{d}u)\,\mathrm{e}^{-zu}$$

of $L_-(v)$ is well defined for those $z \in \mathbb{C}$ with $\mathrm{Re}(z) \le 0$, and is analytic in the region

$$\{z \in \mathbb{C} \colon \mathrm{Re}(z) < 0\}.$$

In principle, F_- may be arbitrary, but for technical reasons we impose a condition on the distribution of the negative claims in the statement of Theorem 1.

Let us establish more notation. For any $z \in \mathbb{C}$ and $\theta \ge 0$, we denote by

$$\boldsymbol{Q}^*(z, \theta) = (q_{ij}^*(z, \theta))_{i,j \in E}$$

the matrix with

$$q_{ij}^*(z, \theta) = (-1)^{i+j} m_{ji}, \tag{29}$$

where m_{ji} denotes the minor (subdeterminant) of $\boldsymbol{Q}(z, \theta)$ obtained by deleting its jth row and ith column. In particular, if $\boldsymbol{Q}(z, \theta)$ is nonsingular then

$$\boldsymbol{Q}^*(z, \theta) = \det \boldsymbol{Q}(z, \theta)\,\boldsymbol{Q}^{-1}(z, \theta).$$

Note that if $p = 1$ and $E = \{1\}$, then

$$\boldsymbol{Q}(z, \theta) = Q(z, \theta) = \phi_1(z) - \lambda_1 - \theta$$

is scalar and $Q^*(z, \theta) = 1$ for all z and θ.

Suppose that Assumption 1 holds. Then, since for $\theta \ge 0$, $z \mapsto \det \boldsymbol{Q}(z, \theta)$ is a polynomial that is not identically equal to 0 (cf. Lemmas 1 and 2, below), $\boldsymbol{Q}(z, \theta)$ is nonsingular for all but finitely many z. It is a consequence of Lemma 1 that $\boldsymbol{Q}(z, \theta)$ is nonsingular whenever $\mathrm{Re}(\phi_i(z)) \le 0$ for all $i \in E$.

In the statement of the main theorem (Theorem 1), we shall consider two versions of the Cramér–Lundberg equation. We shall refer to

$$R_+(\gamma)(1 + (1 - p_U)\boldsymbol{a}^\top \boldsymbol{Q}^{-1}(\gamma, \theta)\boldsymbol{\lambda}L_-(\gamma)) = -p_U P_+(\gamma)\boldsymbol{a}^\top \boldsymbol{Q}^{-1}(\gamma, \theta)\boldsymbol{\lambda} \tag{30}$$

as the *Cramér–Lundberg equation*, and, for a given $\theta \ge 0$, call $\gamma \in \mathbb{C}$ a solution to this equation if $\boldsymbol{Q}(\gamma, \theta)$ is nonsingular and (30) holds. By the *modified Cramér–Lundberg equation* we shall mean the equation

$$R_+(\gamma)(\det \boldsymbol{Q}(\gamma, \theta) + (1 - p_U)\boldsymbol{a}^\top \boldsymbol{Q}^*(\gamma, \theta)\boldsymbol{\lambda}L_-(\gamma)) = -p_U P_+(\gamma)\boldsymbol{a}^\top \boldsymbol{Q}^*(\gamma, \theta)\boldsymbol{\lambda}. \tag{31}$$

Note that if $p_U = 1$ then both the left- and right-hand sides are polynomials in γ, and that if $p_U < 1$ then we may rewrite (31) as an equation involving polynomials in γ, provided that L_- is a rational function.

For $p_U = 1$, (30) and (31) respectively simplify to

$$R_+(\gamma) = -P_+(\gamma)\boldsymbol{a}^\top \boldsymbol{Q}^{-1}(\gamma, \theta)\boldsymbol{\lambda}, \tag{32}$$

$$R_+(\gamma)\det \boldsymbol{Q}(\gamma, \theta) = -P_+(\gamma)\boldsymbol{a}^\top \boldsymbol{Q}^*(\gamma, \theta)\boldsymbol{\lambda}.$$

Theorem 1 deals with solutions to (30) or (31) with $\mathrm{Re}(\gamma) < 0$. It is clear that any solution to (30) is also a solution to (31), but it is entirely possible that (31) may have more solutions than (30). In particular, this happens if $\mathrm{Re}(\gamma) < 0$ and

$$\det \boldsymbol{Q}(\gamma, \theta) = 0, \qquad \boldsymbol{a}^{\top} \boldsymbol{Q}^*(\gamma, \theta) \boldsymbol{\lambda} = 0; \tag{33}$$

see Remark 5, below, for further discussion.

Recall that the size of the undershoot at the time of ruin is denoted by $Y_{\mathrm{r}} = -X_{T_{\mathrm{r}}}$.

Theorem 1. *Consider the risk process X given by (4) and assume that the Laplace transform L_+ for the distribution of the positive claims is given by (27) with the degree of R_+ equal to m. Furthermore, assume that if $p_U < 1$, the Laplace transform L_- of the distribution of the negative claims is analytic in an open region containing $\{z \in \mathbb{C}: \mathrm{Re}(z) \leq 0\}$. Finally, assume that Assumption 1 holds.*

(i) *For any $\theta > 0$, when counted with multiplicity the modified Cramér–Lundberg equation (31) has precisely $m + p_{\mathrm{c}}$ solutions $(\gamma_\ell)_{1 \leq \ell \leq m + p_{\mathrm{c}}} = (\gamma_\ell(\theta))$ with $\mathrm{Re}(\gamma_\ell) < 0$.*

(ii) *For any $\theta > 0$, a γ with $\mathrm{Re}(\gamma) < 0$ is a solution to the Cramér–Lundberg equation (30) if and only if it is a solution to the modified equation (31) with $\boldsymbol{Q}(\gamma, \theta)$ nonsingular.*

(iii) *For a $\theta > 0$ such that $(\tilde{\gamma}_k)_{1 \leq k \leq m}$ are any m distinct solutions to (30) with $\mathrm{Re}(\tilde{\gamma}_k) < 0$ and the matrices $\boldsymbol{Q}(\tilde{\gamma}_k, \theta)$ are nonsingular, for all $x_0 > 0$, $i_0 \in E$, and all $\zeta \geq 0$ we have*

$$\sum_{i \in E_{\mathrm{c}}} \sum_{k=1}^{m} r_k (\boldsymbol{Q}^{-1}(\tilde{\gamma}_k, \theta) \boldsymbol{\lambda})_i \, \mathrm{E}^{x_0, i_0}[\mathrm{e}^{-\theta T_{\mathrm{r}}}; \ A_{\mathrm{c}, i}]$$

$$+ \frac{1}{L_+(\zeta)} \sum_{k=1}^{m} \frac{P_+(\tilde{\gamma}_k)}{(\tilde{\gamma}_k - \zeta) \prod_{k' \neq k}(\tilde{\gamma}_k - \tilde{\gamma}_{k'})} \, \mathrm{E}^{x_0, i_0}[\mathrm{e}^{-\theta T_{\mathrm{r}} - \zeta Y_{\mathrm{r}}}; \ A_{\mathrm{j}}]$$

$$= \sum_{k=1}^{m} r_k (\boldsymbol{Q}^{-1}(\tilde{\gamma}_k, \theta) \boldsymbol{\lambda})_{i_0} \mathrm{e}^{\tilde{\gamma}_k x_0}, \tag{34}$$

where $r_k \equiv r_k(\theta, \zeta)$ is given by

$$r_k = \frac{R_+(\tilde{\gamma}_k)}{\boldsymbol{a}^{\top} \boldsymbol{Q}^{-1}(\tilde{\gamma}_k, \theta) \boldsymbol{\lambda}(\tilde{\gamma}_k - \zeta) \prod_{k' \neq k}(\tilde{\gamma}_k - \tilde{\gamma}_{k'})}. \tag{35}$$

(iv) *If all the solutions $(\gamma_\ell)_{1 \leq \ell \leq m + p_{\mathrm{c}}}$ to (30) with $\mathrm{Re}(\gamma_\ell) < 0$ are distinct and all the matrices $\boldsymbol{Q}(\gamma_\ell, \theta)$ nonsingular, by using (34) $p_{\mathrm{c}} + 1$ times with, say,*

$$(\tilde{\gamma}_k)_{1 \leq k \leq m} = (\tilde{\gamma}_1, \ldots, \tilde{\gamma}_{m-1}, \tilde{\gamma}_{m+s})$$

for $s = 0, \ldots, p_{\mathrm{c}}$, a system of linear equations is obtained that can be solved uniquely, provided that the matrix of coefficients of the $p_{\mathrm{c}} + 1$ unknowns

$$\mathrm{E}^{x_0, i_0}[\mathrm{e}^{-\theta T_{\mathrm{r}} - \zeta Y_{\mathrm{r}}}; \ A_{\mathrm{j}}] \quad \text{and} \quad \mathrm{E}^{x_0, i_0}[\mathrm{e}^{-\theta T_{\mathrm{r}}}; \ A_{\mathrm{c}, i}], \quad i \in E_{\mathrm{c}}, \tag{36}$$

is nonsingular.

Remark 3. It is quite possible that the system of equations in part (iv) always has a unique solution (except, possibly, for a few values of (θ, ζ)), but we have no proof of this.

In part (iv) there are a multitude of ways in which to choose $p_c + 1$ equations. Of course, the solutions do not depend on this choice, as may be verified directly and is verified numerically in Example 1, below. When choosing the equations it is vital that the γ_ℓ, $1 \leq \ell \leq m + p_c$, are *all* used.

Remark 4. The technique used to prove Theorem 1 was first used in [11]. The main result there, Theorem 6 (dealing with the Laplace transform of T_r only), corresponds to the model with $\beta_i = \beta > 0$ and $\sigma_i^2 = 0$ for all $i \in E$, which is the traditional risk model with a fixed premium rate and positive claims only that arrive according to a renewal process with interarrival times of phase type. We thus have $E_c = \varnothing$ and see that, since $\phi_i(z) = \beta z$ for all $i \in E$, meaning that

$$\boldsymbol{Q}(z, \theta) = \boldsymbol{Q}_V - (\theta - \beta z)\boldsymbol{I},$$

(31) becomes $L_V(\theta - \beta\gamma)\bar{L}_+(\gamma) = 1$ (cf. (26)), one of the forms presented in [11].

Remark 5. Theorem 1 is really intended for the situation assumed in part (iv), where (31) has $m + p_c$ distinct solutions γ with $\mathrm{Re}(\gamma) < 0$ and the matrices $\boldsymbol{Q}(\gamma, \theta)$ are nonsingular. The main purpose of this remark is to discuss what happens when the assumptions in part (iv) fail.

Consider the roots of (31) as θ varies. It could occur that, except for at finitely many values of θ, the $m + p_c$ roots are distinct, while moving θ across an exceptional value might cause two real roots to collapse into one, say, and then split into two complex conjugate roots. In this case, Theorem 1 can still be used to find the partial Laplace transforms (36), since they are all continuous functions of θ.

However, the assumptions made in part (iv) could also fail systematically, because for all $\theta > 0$ (33) holds for some $\gamma = \gamma(\theta)$ solving (31) with, in particular, $\boldsymbol{Q}(\gamma, \theta)$ singular; in this case, Theorem 1 is useless. One essential safeguard against this is Assumption 1, but it is also vital that the model (4) for X is parametrized in a suitably minimal fashion: suppose, for simplicity, that $p = 1$. The same model may then be obtained with an arbitrary $p \geq 2$ using the following trivial parametrization: simply set $\sigma_i^2 = \sigma^2$, $\beta_i = \beta$, and $\lambda_i = \lambda$ for all i and set $q_{ij} = 0$ for all $i \neq j$. Then it is easy to verify that (33) holds and Theorem 1(iv) cannot be used.

Remark 6. It is worth emphasizing that, for Theorem 1, the solutions to the Cramér–Lundberg equations with strictly *negative* real parts are required. The work of [3] and [2] involved special cases of the model (4) and used the solutions with *positive* real parts.

Remark 7. The assumption that L_- be suitably analytic is probably not necessary. It is made to simplify the proof of part (i).

Before giving the proof of Theorem 1, we present two lemmas, which are proved in Appendix A. Below, the imaginary unit is denoted by $i = \sqrt{-1}$.

Lemma 1. *Assume that Assumption 1 holds. For $\theta \geq 0$ and $y \in \mathbb{R}$, the matrix $\boldsymbol{Q}(iy, \theta)$ is nonsingular and, for any $\theta \geq 0$, $y \in \mathbb{R}$, and $j \in E$, we have*

$$|(\boldsymbol{Q}^{-1}(iy, \theta)\boldsymbol{\lambda})_j| \leq \frac{q_j + \lambda_j}{q_j + \lambda_j + \theta};$$

in particular,

$$|(\boldsymbol{Q}^{-1}(iy, \theta)\boldsymbol{\lambda})_j| < 1 \quad if \, \theta > 0. \tag{37}$$

Lemma 2. *Assume that Assumption 1 holds. For any $\theta \geq 0$, the polynomial $z \mapsto \det \boldsymbol{Q}(z, \theta)$ is of degree*

$$d = 2 \sum_{j \in E} 1_{\{\sigma_j^2 > 0\}} + \sum_{j \in E} 1_{\{\sigma_j^2 = 0, \, \beta_j \neq 0\}} \tag{38}$$

and has exactly p_c roots z with $\mathrm{Re}(z) < 0$.

Proof of Theorem 1. Parts (ii) and (iv) are obvious. We focus first on part (iii) and then prove part (i). For ease of notation, we shall write

$$L_+^* = p_U L_+, \qquad L_-^* = (1 - p_U) L_-.$$

Let $\theta > 0$ and $\zeta \geq 0$ be given. Consider a function $f : \mathbb{R} \times E \to \mathbb{R}$ of the form

$$f(x, i) = \begin{cases} \sum_{k=1}^{m} c_{ik} e^{\gamma_k x}, & x \geq 0, \\ K e^{\zeta x}, & x < 0. \end{cases} \tag{39}$$

If $\mathrm{Re}(\gamma_k) \leq 0$ for all k then $f \in \mathcal{D}$, the function space relevant for our use of (14). Suppose that $\gamma_1, \ldots, \gamma_m$ are distinct solutions to (30) with $\mathrm{Re}(\gamma_k) < 0$. We shall show that

$$Af(x, i) = \theta f(x, i) \tag{40}$$

for $x \geq 0$ and $i \in E$, where the kth column $\boldsymbol{c}_{|k}$ of the matrix $(c_{ik})_{i \in E, \, 1 \leq k \leq m}$ and the constant K are respectively given by the expressions

$$\boldsymbol{c}_{|k} = r_k \boldsymbol{Q}^{-1}(\gamma_k, \theta) \boldsymbol{\lambda} \tag{41}$$

and

$$K = -\frac{1}{L_+^*(\zeta)} \sum_{k=1}^{m} \left(r_k + \sum_j a_j c_{jk} L_-^*(\gamma_k) \right), \tag{42}$$

where r_k is as in (35). (Note that the γ_k depend on θ but not on ζ, while r_k, $\boldsymbol{c}_{|k}$, and K depend on both θ and ζ.)

From (40) and (19), it follows that, for arbitrary initial states $x_0 > 0$ and $i_0 \in E$,

$$\mathrm{E}^{x_0, i_0}\left[e^{-\theta T_r} \sum_{k=1}^{m} c_{J_{T_r}, k}; \; Y_r = 0 \right] + \mathrm{E}^{x_0, i_0}[K e^{-\theta T_r - \zeta Y_r}; \; Y_r > 0] = \sum_{k=1}^{m} c_{i_0 k} e^{\gamma_k x_0},$$

which, for $\tilde{\gamma}_k = \gamma_k$, $1 \leq k \leq m$, is precisely (34), since

$$K = -\frac{1}{L_+^*(\zeta)} \sum_{k=1}^{m} r_k (1 + \boldsymbol{a}^\top \boldsymbol{Q}^{-1}(\gamma_k, \theta) \boldsymbol{\lambda} L_-^*(\gamma_k))$$

$$= \frac{1}{L_+^*(\zeta)} \sum_{k=1}^{m} \frac{p_U P_+(\gamma_k)}{(\gamma_k - \zeta) \prod_{k' \neq k}(\gamma_k - \gamma_{k'})},$$

by (42), (41), and (30).

Thus, part (iii) follows by verifying that (40) holds when r_k, $\boldsymbol{c}_{|k}$, and K are given by (35), (41), and (42), respectively; this we now show.

With f as in (39), (13) implies that

$$Af(x,i) = \sum_{k=1}^{m} \Bigg[c_{ik}\phi_i(\gamma_k)e^{\gamma_k x} + \sum_{j \neq i} q_{ij}(c_{jk} - c_{ik})e^{\gamma_k x}$$
$$+ \lambda_i \sum_j a_j \Bigg\{ \int_{(-\infty,x]} F_U(\mathrm{d}y)c_{jk}e^{\gamma_k(x-y)} + \int_{(x,\infty)} F_U(\mathrm{d}y)Ke^{\zeta(x-y)} - c_{ik}e^{\gamma_k x} \Bigg\} \Bigg].$$

Hence, (40) is equivalent to

$$0 = \sum_k e^{\gamma_k x}(\boldsymbol{Q}(\gamma_k,\theta)\boldsymbol{c}_{|k})_i$$
$$+ \lambda_i \sum_j a_j \Bigg[\int_{(-\infty,x]} F_U(\mathrm{d}y)\sum_k c_{jk}e^{\gamma_k(x-y)} + \int_{(x,\infty)} F_U(\mathrm{d}y)Ke^{\zeta(x-y)} \Bigg],$$

which must hold for all $x \geq 0$ and $i \in E$. However, using (41) allows us to eliminate the dependence on i, since a common factor of λ_i appears, and we are left with

$$\sum_k r_k e^{\gamma_k x} + \sum_j a_j \Bigg[\int_{(-\infty,x]} F_U(\mathrm{d}y)\sum_k c_{jk}e^{\gamma_k(x-y)} + \int_{(x,\infty)} F_U(\mathrm{d}y)Ke^{\zeta(x-y)} \Bigg] = 0, \quad (43)$$

which must hold for all $x \geq 0$. Taking $x = 0$ shows that

$$\sum_k \Bigg(r_k + \sum_j a_j c_{jk} L_-^*(\gamma_k) \Bigg) + K L_+^*(\zeta) = 0,$$

in agreement with (42).

Since $\mathrm{Re}(\gamma_k) < 0$ for all k, (43) holds for all $x \geq 0$ if and only if it holds for the Laplace transform: multiplying by $e^{-\nu x}$ and integrating x from 0 to ∞ implies, using (42), that (43) is equivalent to

$$\sum_k r_k \Bigg(\frac{1}{\nu - \gamma_k} - \frac{1}{\nu - \zeta}\Bigg(1 - \frac{L_+^*(\nu)}{L_+^*(\zeta)}\Bigg)\Bigg)$$
$$+ \sum_k s_k \Bigg\{ \frac{L_+^*(\nu)}{\nu - \gamma_k} + L_-^*(\gamma_k)\Bigg(\frac{1}{\nu - \gamma_k} - \frac{1}{\nu - \zeta}\Bigg(1 - \frac{L_+^*(\nu)}{L_+^*(\zeta)}\Bigg)\Bigg) \Bigg\} = 0 \quad (44)$$

for $\nu \geq 0$, with

$$s_k = \sum_j a_j c_{jk} = r_k \boldsymbol{a}^\top \boldsymbol{Q}^{-1}(\gamma_k,\theta)\boldsymbol{\lambda}. \quad (45)$$

Note that to deduce (44) from (43), we use (42) and the elementary formulae

$$\int_0^\infty \mathrm{d}x\, e^{-\nu x} \int_{(0,x]} F_+(\mathrm{d}y)e^{\gamma_k(x-y)} = \frac{L_+(\nu)}{\nu - \gamma_k}$$

and

$$\int_0^\infty \mathrm{d}x\, e^{-\nu x} \int_{(x,\infty]} F_+(\mathrm{d}y)\, e^{\zeta(x-y)} = \frac{1}{\nu - \zeta}(L_+(\zeta) - L_+(\nu)).$$

Solving (44) for $L_+^*(\nu)$ gives

$$
\begin{aligned}
L_+^*(\nu) = & -\sum_k (r_k + s_k L_-^*(\gamma_k)) \left(\frac{1}{\nu - \gamma_k} - \frac{1}{\nu - \zeta} \right) \\
& \times \left[\sum_k \left(s_k \left\{ \frac{1}{\nu - \gamma_k} + \frac{L_-^*(\gamma_k)}{(\nu - \zeta) L_+^*(\zeta)} \right\} + \frac{r_k}{(\nu - \zeta) L_+^*(\zeta)} \right) \right]^{-1} \\
= & -\sum_k (\gamma_k - \zeta)(r_k + s_k L_-^*(\gamma_k)) \pi_{\backslash k}(\nu) \\
& \times \left[(\nu - \zeta) \sum_k \left\{ s_k \left(1 + \frac{L_-^*(\gamma_k)}{L_+^*(\zeta)} \right) + \frac{r_k}{L_+^*(\zeta)} \right\} \pi_{\backslash k}(\nu) \right. \\
& \left. - \frac{1}{L_+^*(\zeta)} \sum_k (\gamma_k - \zeta)(r_k + s_k L_-^*(\gamma_k)) \pi_{\backslash k}(\nu) \right]^{-1},
\end{aligned}
\tag{46}
$$

using the notation $\pi_{\backslash k}(\nu) = \prod_{k' \neq k} (\nu - \gamma_{k'})$.

At this stage, we remind the reader that if $\mathcal{P}$ is a polynomial of degree less than or equal to $m - 1$, then

$$
\mathcal{P}(z) = \sum_k \frac{\mathcal{P}(\gamma_k)}{\pi_{\backslash k}(\gamma_k)} \pi_{\backslash k}(z), \qquad z \in \mathbb{C};
\tag{47}
$$

see, e.g. Lemma 4 of [11]. Thus, to identify the numerator of the ratio in the last line of (46) with the numerator $p_U P_+(\nu)$ of $L_+^*(\nu)$, we must have

$$
-(\gamma_k - \zeta)(r_k + s_k L_-^*(\gamma_k)) = \frac{p_U P_+(\gamma_k)}{\pi_{\backslash k}(\gamma_k)},
\tag{48}
$$

which follows from (45), (35), and (30). To identify the denominator of (46) with $R_+(\nu)$, first note that

$$
S(\nu) := \frac{1}{\nu - \zeta} \left(R_+(\nu) - \frac{P_+(\nu)}{L_+(\zeta)} \right) = \frac{1}{\nu - \zeta} \left(R_+(\nu) - \frac{p_U P_+(\nu)}{L_+^*(\zeta)} \right), \qquad \nu \geq 0,
$$

defines a polynomial of degree less than or equal to $m-1$, simply because $R_+(\nu) - P_+(\nu)/L_+(\zeta)$ is a polynomial of degree m that has $\nu = \zeta$ as a root; hence, by (47),

$$
S(\nu) = \sum_k \frac{1}{(\gamma_k - \zeta)\pi_{\backslash k}(\gamma_k)} \left(R_+(\gamma_k) - \frac{p_U P_+(\gamma_k)}{L_+^*(\zeta)} \right) \pi_{\backslash k}(\nu).
\tag{49}
$$

Furthermore, the denominator equals $R_+(\nu)$ if and only if

$$
S(\nu) = \sum_k \left\{ s_k \left(1 + \frac{L_-^*(\gamma_k)}{L_+^*(\zeta)} \right) + \frac{r_k}{L_+^*(\zeta)} \right\} \pi_{\backslash k}(\nu).
$$

By using (48), it is seen that the coefficients of $\pi_{\backslash k}(\nu)$ here and in (49) are the same provided that

$$
s_k = \frac{R_+(\gamma_k)}{(\gamma_k - \zeta)\pi_{\backslash k}(\gamma_k)}.
$$

By (45), this follows directly from (35). This completes the proof of part (iii).

We proceed with the proof of part (i). Let $\theta > 0$ be given. By Lemma 2 and the earlier observation that $\mathrm{Re}(z) < 0$ for all roots z of $R_+(z)$, the polynomial

$$z \mapsto \mathscr{P}_1(z) := R_+(z) \det \boldsymbol{Q}(z, \theta)$$

is of the degree d given by (38) and has exactly $m + p_c$ roots z with $\mathrm{Re}(z) < 0$. Furthermore, none of the d roots are of the form $\mathrm{i}y$, $y \in \mathbb{R}$.

For $\rho > 0$, let Γ_ρ denote the interior of the subset of $\mathbb{C}$ determined by the outer boundary

$$\partial\Gamma_\rho = \{z : |z| = \rho, \ \mathrm{Re}(z) < 0\} \cup \{z : z = \mathrm{i}y, \ -\rho \le y \le \rho\}. \tag{50}$$

We first claim that the polynomial

$$z \mapsto \mathscr{P}_1(z) + p_U \mathscr{P}_2(z),$$

where

$$\mathscr{P}_2(z) = P_+(z)\boldsymbol{a}^\top \boldsymbol{Q}^*(z, \theta)\boldsymbol{\lambda},$$

has (counting with multiplicity) the same number of zeros in Γ_ρ as $\mathscr{P}_1$. This follows from Rouché's theorem in complex function theory if

$$p_U |\mathscr{P}_2(z)| < |\mathscr{P}_1(z)|, \qquad z \in \partial\Gamma_\rho, \tag{51}$$

and here it obviously suffices to consider the case $p_U = 1$. However, since, as can be directly verified, every minor m_{ij} of $\boldsymbol{Q}(z, \theta)$ is a polynomial of degree less than or equal to d, $\mathscr{P}_2(z)$ is of degree less than or equal to $(m-1) + d < m + d$. Equation (51) is then obvious for $z \in \partial\Gamma_\rho$ with $|z| = \rho$ if only ρ is large enough. Furthermore, for $z = \mathrm{i}y$ with $-\rho \le y \le \rho$, since $\boldsymbol{Q}(\mathrm{i}y, \theta)$ is nonsingular by Lemma 1, for $p_U = 1$ (51) is equivalent to

$$|P_+(\mathrm{i}y)\boldsymbol{a}^\top \boldsymbol{Q}^{-1}(\mathrm{i}y, \theta)\boldsymbol{\lambda}| < |R_+(\mathrm{i}y)|.$$

This follows from (37) (which implies that $|\boldsymbol{a}^\top \boldsymbol{Q}^{-1}(\mathrm{i}y, \theta)\boldsymbol{\lambda}| \le \sum_j a_j |(\boldsymbol{Q}^{-1}(\mathrm{i}y, \theta)\boldsymbol{\lambda})_j| < 1$) and the fact that

$$\frac{P_+(\mathrm{i}y)}{R_+(\mathrm{i}y)} = \bar{L}_+(\mathrm{i}y) = L_+(\mathrm{i}y) = \int_0^\infty \mathrm{e}^{\mathrm{i}yu} F_+(\mathrm{d}u),$$

where the last term is less than or equal to 1 in absolute value.

Having thus shown that the polynomial $\mathscr{P}_1 + p_U \mathscr{P}_2$ has exactly $m + p_c$ roots with $\mathrm{Re}(z) < 0$, we can show that the same holds for the function

$$z \mapsto R_+(z)(\det \boldsymbol{Q}(z, \theta) + (1 - p_U)\boldsymbol{a}^\top \boldsymbol{Q}^*(z, \theta)\boldsymbol{\lambda} L_-(z)) + p_U P_+(z)\boldsymbol{a}^\top \boldsymbol{Q}^*(z, \theta)\boldsymbol{\lambda}$$

(cf. (31)), by yet another application of Rouché's theorem: it suffices to argue that, for ρ sufficiently large,

$$(1 - p_U)|R_+(z)|\,|\boldsymbol{a}^\top \boldsymbol{Q}^*(z, \theta)\boldsymbol{\lambda} L_-(z)| < |\mathscr{P}_1(z) + p_U \mathscr{P}_2(z)|, \qquad z \in \partial\Gamma_\rho, \tag{52}$$

which, since $|L_-(z)| \le 1$ if $\mathrm{Re}(z) \le 0$, will follow from

$$(1 - p_U)|R_+(z)|\,|\boldsymbol{a}^\top \boldsymbol{Q}^*(z, \theta)\boldsymbol{\lambda}| < |\mathscr{P}_1(z) + p_U \mathscr{P}_2(z)|, \qquad z \in \partial\Gamma_\rho. \tag{53}$$

If $z \in \partial\Gamma_\rho$ with $|z| = \rho$ large, this is true whenever the polynomial

$$z \mapsto R_+(z)\boldsymbol{a}^\top \boldsymbol{Q}^*(z, \theta)\boldsymbol{\lambda} \tag{54}$$

is of degree less than $m + d$, which is the degree of $\mathcal{P}_1$ and also obviously of $\mathcal{P}_1 + p_U\mathcal{P}_2$. Now (see (29))

$$\boldsymbol{a}^\top \boldsymbol{Q}^*(z, \theta)\boldsymbol{\lambda} = \sum_{i, j \in E} a_i(-1)^{i+j} m_{ji}\lambda_j$$

and, since $m_{ji} = m_{ji}(z)$ is a polynomial of maximal degree d only if both i and j are such that removal of the jth row and the ith column from $\boldsymbol{Q}(z, \theta)$ does not remove any entries depending on z, it is seen that m_{ji} is of degree d only if $i, j \in E \setminus E'$, where

$$E' = \{i' \in E : \sigma_{i'}^2 > 0 \text{ or } \beta_{i'} \neq 0\}. \tag{55}$$

Consequently, (53) holds with $|z| = \rho$ (provided that ρ is large enough) whenever $E' = E$. The case $E \setminus E' \neq \varnothing$ is more delicate, since then the polynomial (54) may actually be of degree $m + d$ and it becomes necessary to compare the coefficients of z^{m+d} in the polynomials (54) and $\mathcal{P}_1 + p_U\mathcal{P}_2$. For the latter polynomial, this coefficient is the same as that in $\mathcal{P}_1$, and equals

$$\kappa \det \boldsymbol{Q}_{E \setminus E'}(\theta), \tag{56}$$

where

$$\kappa = \prod_{\{i \in E' : \sigma_i^2 > 0\}} \left(\frac{1}{2}\sigma_i^2\right) \prod_{\{i \in E' : \sigma_i^2 = 0\}} \beta_i$$

and $\boldsymbol{Q}_{E \setminus E'}(\theta)$ is the matrix obtained from $\boldsymbol{Q}(z, \theta)$ by deleting the rows and columns corresponding to states in E'; in particular, $\boldsymbol{Q}_{E \setminus E'}(\theta)$ does not depend on z. This is shown in the proof of Lemma 2; see (84), below.

Now, by inspection, the coefficient of z^{m+d} in (54) can be written

$$\kappa \sum_{i, j \in E \setminus E'} a_i(-1)^{i+j} m_{E \setminus E', ji}\lambda_j, \tag{57}$$

where $m_{E \setminus E', ji}$ is the minor of $\boldsymbol{Q}_{E \setminus E'}(\theta)$ obtained by deleting row j and column i. As a special case of Lemma 1, the subintensity matrix $\boldsymbol{Q}_{E \setminus E'}(\theta)$ is nonsingular with

$$|(\boldsymbol{Q}_{E \setminus E'}^{-1}(\theta)\boldsymbol{\lambda}_{E \setminus E'})_j| < 1;$$

hence, the expression in (57) equals

$$\kappa (\det \boldsymbol{Q}_{E \setminus E'}(\theta))(\boldsymbol{a}_{E \setminus E'}^\top \boldsymbol{Q}_{E \setminus E'}^{-1}(\theta)\boldsymbol{\lambda}_{E \setminus E'}),$$

which is less than (56) in absolute value. Because of this, (53) holds with $|z| = \rho$, for a sufficiently large ρ, and also if $E \setminus E' \neq \varnothing$.

It remains to verify (53) for $z = iy$ with $-\rho \leq y \leq \rho$. However, since $\det \boldsymbol{Q}(iy, \theta) \neq 0$ (by Lemma 1), in this case (53) is equivalent to the inequality

$$(1 - p_U)|R_+(iy)| \, |\boldsymbol{a}^\top \boldsymbol{Q}^{-1}(iy, \theta)\boldsymbol{\lambda}| < |R_+(iy) + p_U P_+(iy)\boldsymbol{a}^\top \boldsymbol{Q}^{-1}(iy, \theta)\boldsymbol{\lambda}|.$$

By Lemma 1, the expression on the left is less than $(1 - p_U)|R_+(iy)|$ and, since that on the right is greater than or equal to

$$|R_+(iy)| \, |1 - p_U(|L_+(iy)\boldsymbol{a}^\top \boldsymbol{Q}^{-1}(iy, \theta)\boldsymbol{\lambda}|)| > (1 - p_U)|R_+(iy)|,$$

which follows from Lemma 1 and the fact that $|L_+(iy)| \leq 1$, the proof of part (i) is complete.

4. The probability of ruin

The probability of ultimate ruin is

$$
\begin{aligned}
p_{\mathrm{r}} &= \mathrm{P}^{x_0,i_0}(T_{\mathrm{r}} < \infty) \\
&= \mathrm{P}^{x_0,i_0}(A_{\mathrm{j}}) + \sum_{i \in E_{\mathrm{c}}} \mathrm{P}^{x_0,i_0}(A_{\mathrm{c},i}).
\end{aligned}
\tag{58}
$$

Each term may be determined from the expressions for the quantities in (36) by taking $\zeta = 0$ and letting $\theta \downarrow 0$, since

$$
\mathrm{P}^{x_0,i_0}(T_{\mathrm{r}} < \infty, \, A) = \lim_{\theta \downarrow 0, \, \theta > 0} \mathrm{E}^{x_0,i_0}[\mathrm{e}^{-\theta T_{\mathrm{r}}}; \, A]
$$

for any event A. Thus, Theorem 1 makes it possible to determine each of the terms in (58). However, for direct calculation it is of course preferable to avoid taking limits, and we now discuss how this may be done. Compared with Theorem 1 and its proof, the discussion is more intricate and involves a number of subtleties of an analytic nature.

It is natural that we should try to use the Cramér–Lundberg equations (30) or (31) with $\theta = 0$. However, allowing $\theta = 0$ invalidates the strict inequality in (37) (for $y = 0$, the inequality does not hold for $\theta = 0$; cf. (9)), which was essential in applying Rouché's theorem in the proof of Theorem 1. It is thus necessary to be careful; in particular, as we shall see, Theorem 2, the analogue of Theorem 1, which is used to compute the terms in (58), has two versions: one applies when $p_{\mathrm{r}} < 1$ and the other when $p_{\mathrm{r}} = 1$.

Write $\xi = \mathrm{E}\, U_n$ for the expected claim size (which is well defined with $-\infty \leq \xi < \infty$) and $\mu = \mathrm{E}\, V_n$, $n \geq 2$, for the expected times between claims. Then (see, e.g. (7))

$$
\mu = -\boldsymbol{a}^{\top} \boldsymbol{Q}_V^{-1} \mathbf{1}.
\tag{59}
$$

Also, let $\boldsymbol{\alpha} = (\alpha_i)_{i \in E}$ denote the unique invariant probability for the irreducible Markov chain J; then $\alpha_i > 0$, $\sum \alpha_i = 1$, and

$$
\sum_{i \in E} \alpha_i \tilde{q}_{ij} = 0, \qquad j \in E,
\tag{60}
$$

where the $\tilde{q}_{ij}$ are as given in (5) and (6). The row vector $\boldsymbol{\alpha}^{\top}$ is in fact given by

$$
\boldsymbol{\alpha}^{\top} = -\frac{1}{\mu} \boldsymbol{a}^{\top} \boldsymbol{Q}_V^{-1}.
\tag{61}
$$

To see this, we use (8) to rewrite (60) as

$$
\sum_i \alpha_i q_{V,ij} = -a_j \sum_i \alpha_i \lambda_i.
$$

Thus, $\boldsymbol{\alpha}^{\top} \boldsymbol{Q}_V = -C \boldsymbol{a}^{\top}$, with $C = \sum_i \alpha_i \lambda_i$, and (61) follows from $\boldsymbol{\alpha}^{\top} \mathbf{1} = 1$ together with (59).

Proposition 1. *For $x_0 > 0$ and $i_0 \in E$, the ruin probability $p_{\mathrm{r}} = \mathrm{P}^{x_0,i_0}(T_{\mathrm{r}} < \infty)$ equals 1 if and only if*

$$
\sum_{i \in E} \alpha_i \beta_i \leq \frac{\xi}{\mu}.
\tag{62}
$$

Proof. By (4),

$$\frac{1}{t}(X_t - x_0) = \frac{1}{t}\int_0^t \beta_{J_s}\, ds + \frac{1}{t}\int_0^t \sigma_{J_{s-}}\, dB_s - \frac{1}{t}\sum_{n=1}^{N_t} U_n. \tag{63}$$

Since J is ergodic,

$$\lim_{t\to\infty}\frac{1}{t}\int_0^t \beta_{J_s}\, ds = \sum_{i\in E}\alpha_i\beta_i \quad P^{x_0,i_0}\text{-a.s.}$$

By standard properties of Brownian motion, the second term on the right-hand side of (63) converges to 0 a.s., while, because of the renewal structure of the claims arrival process (T_n) and the independence between this and (U_n),

$$\lim_{t\to\infty}\frac{1}{t}\sum_{n=1}^{N_t} U_n = \frac{\xi}{\mu} \quad \text{a.s.}$$

Thus,

$$X_t \to -\infty \quad P^{x_0,i_0}\text{-a.s. if} \quad \sum_{i\in E}\alpha_i\beta_i < \frac{\xi}{\mu}$$

and

$$X_t \to \infty \quad P^{x_0,i_0}\text{-a.s. if} \quad \sum_{i\in E}\alpha_i\beta_i > \frac{\xi}{\mu},$$

which shows both that $p_r = 1$ if (62) holds with strict inequality and, since the drift of X to ∞ may begin with positive probability before any claim has arrived, that $p_r < 1$ if (62) does not hold. If there is equality in (62), we may, for instance, argue directly using (4) that $p_r = p_r(\varepsilon)$ is a continuous function of $\varepsilon \geq 0$, where $p_r(\varepsilon)$ is the ruin probability for the process obtained by replacing β_i by $\beta_i - \varepsilon$ but retaining all the other parameters. Then, of course, $p_r(0) = \lim_{\varepsilon\downarrow 0,\,\varepsilon>0} p_r(\varepsilon) = 1$ since, for this '$(\beta_i - \varepsilon)$-process', (62) holds with strict inequality.

For $\theta = 0$, the Cramér–Lundberg equations (30) and (31) take the respective forms

$$R_+(\gamma)(1 + (1 - p_U)\boldsymbol{a}^\top \boldsymbol{Q}^{-1}(\gamma,0)\boldsymbol{\lambda}L_-(\gamma)) = -p_U P_+(\gamma)\boldsymbol{a}^\top \boldsymbol{Q}^{-1}(\gamma,0)\boldsymbol{\lambda} \tag{64}$$

and

$$R_+(\gamma)(\det \boldsymbol{Q}(\gamma,0) + (1 - p_U)\boldsymbol{a}^\top \boldsymbol{Q}^*(\gamma,0)\boldsymbol{\lambda}L_-(\gamma)) = -p_U P_+(\gamma)\boldsymbol{a}^\top \boldsymbol{Q}^*(\gamma,0)\boldsymbol{\lambda}, \tag{65}$$

where the elements of $\boldsymbol{Q}^*(\gamma,0)$ are as given in (29) with $\theta = 0$. If $p_U = 1$, the equations simplify to

$$R_+(\gamma) = -P_+(\gamma)\boldsymbol{a}^\top \boldsymbol{Q}^{-1}(\gamma,0)\boldsymbol{\lambda}$$

and

$$R_+(\gamma) \det \boldsymbol{Q}(\gamma,0) = -P_+(\gamma)\boldsymbol{a}^\top \boldsymbol{Q}^*(\gamma,0)\boldsymbol{\lambda}. \tag{66}$$

The first thing to note is that $\gamma = 0$ is a solution to both (64) and (65): $\boldsymbol{Q}(0,0) = \boldsymbol{Q}_V$ is nonsingular, by (9) we have $\boldsymbol{a}^\top \boldsymbol{Q}_V^{-1}\boldsymbol{\lambda} = -1$, while $R_+(0) = P_+(0)$ and $L_-(0) = 1$. However, $\tilde{\gamma}_k = 0$ cannot be used to define the r_k (see (35)) when $\zeta = 0$, so the solution $\gamma = 0$ is to be avoided when describing the terms in (58). What we shall show is that if $p_r < 1$ then (65) has

precisely $m + p_c$ solutions γ_ℓ with $\mathrm{Re}(\gamma_\ell) < 0$ (as in the case $\theta > 0$), while if $p_r = 1$ and $p_c \geq 1$ then there are only $m + p_c - 1$ solutions (one of the solutions from the case $\theta > 0$ has moved to the origin). If $p_r < 1$ then the $p_c + 1$ terms in (58) may be found in complete analogy with Theorem 1, parts (iii) and (iv), while if $p_r = 1$ and $p_c \geq 1$ then, since the terms of (58) now sum to 1, only p_c of these have to be found. For this, it suffices to exploit the $m + p_c - 1$ solutions γ to (65) with $\mathrm{Re}(\gamma) < 0$. (The trivial case $p_r = 1$ and $p_c = 0$ is ignored in the statement of Theorem 2.)

Recall that the events of ruin, A_j and $A_{c,i}$, respectively defined by (23) and (24), are subsets of the set $\{T_r < \infty\}$.

Theorem 2. *Consider the risk process X given by (4) and assume that the Laplace transform for the distribution of the claims is given by (27) with the degree of R_+ equal to m. Assume that Assumption 1 holds.*

(i) *If $p_r < 1$ then, when counted with multiplicity, the modified Cramér–Lundberg equation (65) for the ruin probabilities has precisely $m + p_c$ solutions $(\gamma_\ell)_{1 \leq \ell \leq m + p_c} = (\gamma_\ell(\theta))$ with $\mathrm{Re}(\gamma_\ell) < 0$. If $p_r = 1$ and $p_c \geq 1$ then, when counted with multiplicity, (65) has precisely $m + p_c - 1$ solutions $(\gamma_\ell)_{1 \leq \ell \leq m + p_c - 1} = (\gamma_\ell(\theta))$ with $\mathrm{Re}(\gamma_\ell) < 0$.*

(ii) *A γ with $\mathrm{Re}(\gamma) < 0$ is a solution to the Cramér–Lundberg equation (64) if and only if γ is a solution to the modified equation (65) with $\boldsymbol{Q}(\gamma, 0)$ nonsingular.*

(iii) *In both cases ($p_r < 1$ and $p_r = 1$, $p_c \geq 1$), if $(\tilde{\gamma}_k)_{1 \leq k \leq m}$ are any m distinct solutions to (64) with $\mathrm{Re}(\tilde{\gamma}_k) < 0$ and the matrices $\boldsymbol{Q}(\tilde{\gamma}_k, 0)$ are nonsingular, then for all $x_0 > 0$ and $i_0 \in E$ we have*

$$\sum_{i \in E_c} \sum_{k=1}^{m} r_k (\boldsymbol{Q}^{-1}(\tilde{\gamma}_k, 0)\boldsymbol{\lambda})_i \, \mathrm{P}^{x_0, i_0}(A_{c,i}) + \sum_{k=1}^{m} \frac{P_+(\tilde{\gamma}_k)}{\tilde{\gamma}_k \prod_{k' \neq k}(\tilde{\gamma}_k - \tilde{\gamma}_{k'})} \, \mathrm{P}^{x_0, i_0}(A_j)$$
$$= \sum_{k=1}^{m} r_k (\boldsymbol{Q}^{-1}(\tilde{\gamma}_k, 0)\boldsymbol{\lambda})_{i_0} \mathrm{e}^{\tilde{\gamma}_k x_0}, \tag{67}$$

where r_k is given by

$$r_k = \frac{R_+(\tilde{\gamma}_k)}{\boldsymbol{a}^\top \boldsymbol{Q}^{-1}(\tilde{\gamma}_k, 0)\boldsymbol{\lambda}\tilde{\gamma}_k \prod_{k' \neq k}(\tilde{\gamma}_k - \tilde{\gamma}_{k'})}. \tag{68}$$

(iv) *If $p_r < 1$ and the solutions $(\gamma_\ell)_{1 \leq \ell \leq m + p_c}$ to (30) with $\mathrm{Re}(\gamma_\ell) < 0$ are distinct and the matrices $\boldsymbol{Q}(\gamma_\ell, 0)$ are nonsingular, by using (67) $p_c + 1$ times with, say, $(\tilde{\gamma}_k)_{1 \leq k \leq m} = (\tilde{\gamma}_1, \ldots, \tilde{\gamma}_{m-1}, \tilde{\gamma}_{m+s})$ for $s = 0, \ldots, p_c$, a system of linear equations is obtained that can be solved uniquely, provided that the matrix of coefficients of the $p_c + 1$ unknowns*

$$\mathrm{P}^{x_0, i_0}(A_j) \qquad \text{and} \qquad \mathrm{P}^{x_0, i_0}(A_{c,i}), \quad i \in E_c, \tag{69}$$

is nonsingular.

If $p_r = 1$ and $p_c \geq 1$, the solutions $(\gamma_\ell)_{1 \leq \ell \leq m + p_c - 1}$ to (30) with $\mathrm{Re}(\gamma_\ell) < 0$ are distinct, and the matrices $\boldsymbol{Q}(\gamma_\ell, 0)$ are nonsingular, by using (67) p_c times with, say, $(\tilde{\gamma}_k)_{1 \leq k \leq m} = (\tilde{\gamma}_1, \ldots, \tilde{\gamma}_{m-1}, \tilde{\gamma}_{m+s})$ for $s = 0, \ldots, p_c - 1$, a system of linear equations is obtained that, provided the relevant matrix of coefficients is nonsingular, can be solved uniquely for p_c of the $p_c + 1$ unknowns in (69). The remaining unknown is equal to 1 minus the sum of the p_c others.

Proof. We do not give all the details, but rather focus on those parts of the proof that differ from the arguments used in the proof of Theorem 1. Also, we shall concentrate on the one-sided case $p_U = 1$ and, when treating the case $p_r = 1$, shall assume that there is strict inequality in (62).

Claims (ii) and (iv) of the theorem are obvious, so we proceed to discuss part (iii). Let $\gamma_1, \ldots, \gamma_m$ be m distinct roots of (64) and consider the function f given by (39) with

$$c_{|k} = r_k \boldsymbol{Q}^{-1}(\gamma_k, 0)\boldsymbol{\lambda}, \qquad K = -\frac{1}{p_U} \sum_{k=1}^{m} \left(r_k + \sum_{j \in E} a_j c_{jk} L_-^*(\gamma_k) \right)$$

(cf. (41) and (42)), where r_k is given by (68). By arguing as in the proof of Theorem 1, we find that

$$A f(x, i) = 0$$

for $x \geq 0$ and $i \in E$ (cf. (40)) and, consequently, by (18), that

$$\mathrm{E}^{x_0, i_0}[f(X_{T_r}, J_{T_r}); \ T_r \leq t] + \mathrm{E}^{x_0, i_0}[f(X_t, J_t); \ T_r > t] = f(x_0, i_0) \tag{70}$$

for any $t \geq 0$. We now claim that

$$\lim_{t \to \infty} \mathrm{E}^{x_0, i_0}[f(X_t, J_t); \ T_r > t] = 0. \tag{71}$$

If $p_r = 1$, since f is bounded this is obvious. If $p_r < 1$, it was argued in the proof of Proposition 1 that $\lim_{t \to \infty} X_t = \infty$ a.s., which, since having $\mathrm{Re}(\gamma_k) < 0$ for all k implies that $\lim_{t \to \infty} f(X_t, J_t) = 0$ a.s., gives (71) by dominated convergence.

Thus, (71) holds and letting $t \to \infty$ in (70) now yields

$$\mathrm{E}^{x_0, i_0}[f(X_{T_r}, J_{T_r}); \ T_r < \infty] = f(x_0, i_0),$$

the analogue of (19), from which (67) follows directly. This proves part (iii).

The main difference between the arguments given here and those yielding Theorem 1 is in the proof of part (i). First, suppose that $p_U = 1$. It is then still true that the left-hand side, say $\mathcal{P}_l$, and the right-hand side, say $\mathcal{P}_r$, of (66) are both polynomials, with $\mathcal{P}_l$ of degree $m + d$ (and d given by (38)). By Lemma 2, $\mathcal{P}_l(z)$ has exactly $m + p_c$ roots with $\mathrm{Re}(z) < 0$ and $\mathcal{P}_r$ is of degree less than $m + d$, so to complete the proof of part (i) it remains to apply Rouché's theorem. To do so it is, however, necessary to adjust and refine the argument in the proof of Theorem 1: because $\gamma = 0$ is always a solution to (65), the strict inequality

$$|P_+(\mathrm{i}y)\boldsymbol{a}^\top \boldsymbol{Q}^*(\mathrm{i}y, \theta)\boldsymbol{\lambda}| < |R_+(\mathrm{i}y) \det \boldsymbol{Q}(\mathrm{i}y, \theta)|,$$

established there for $\theta > 0$, fails for $\theta = 0$ and $y = 0$.

Instead of using the open set Γ_ρ determined by the boundary (50), let $\rho > 0$ be given and, for ε, $0 < \varepsilon < \rho$, so small that both $\boldsymbol{Q}(z, 0)$ is nonsingular for $|z| \leq \varepsilon$ (which is possible since $\boldsymbol{Q}(0, 0) = \boldsymbol{Q}_V$ is nonsingular) and

$$\bar{L}_+(z) = \int_0^\infty \mathrm{e}^{-zu} F_+(\mathrm{d}u) \tag{72}$$

(which is possible by the comment following (28)), define $\Gamma_{\rho, \varepsilon}$ as the interior of the subset of $\mathbb{C}$ determined by the outer boundary

$$\partial\Gamma_{\rho, \varepsilon} = \{z \colon |z| = \rho, \ \mathrm{Re}(z) < 0\} \cup \{z \colon z = \mathrm{i}y, \ y \in \mathbb{R}, \ \varepsilon \leq |y| \leq \rho\}$$
$$\cup \{z \colon |z| = \varepsilon, \ \mathrm{Re}(z) < 0\}$$

if $p_r < 1$, and

$$\partial\Gamma_{\rho,\varepsilon} = \{z\colon |z| = \rho,\ \mathrm{Re}(z) < 0\} \cup \{z\colon z = \mathrm{i}y,\ y \in \mathbb{R},\ \varepsilon \le |y| \le \rho\}$$
$$\cup \{z\colon |z| = \varepsilon,\ \mathrm{Re}(z) > 0\}$$

if $p_r = 1$. Thus, $0 \notin \Gamma_{\rho,\varepsilon}$ if $p_r < 1$ and $0 \in \Gamma_{\rho,\varepsilon}$ if $p_r = 1$.

Rouché's theorem will imply the claim of part (i) if we show that

$$|-P_+(z)\boldsymbol{a}^\top \boldsymbol{Q}^*(z,0)\boldsymbol{\lambda}| < |R_+(z)\det \boldsymbol{Q}(z,0)|, \tag{73}$$

for $z \in \partial\Gamma_{\rho,\varepsilon}$, when ρ is sufficiently large and ε is sufficiently small.

Here there is no problem if $|z| = \rho$ with ρ large, since the polynomial on the left is of lower degree than is that on the right. If $z = \mathrm{i}y$ with $|y| \ge \varepsilon$, since $\boldsymbol{Q}(\mathrm{i}y, 0)$ is nonsingular and $|\boldsymbol{a}^\top \boldsymbol{Q}^{-1}(\mathrm{i}y)\boldsymbol{\lambda}| \le 1$ by Lemma 1, it suffices to argue that

$$|\bar{L}_+(\mathrm{i}y)| < 1.$$

However, if

$$|\bar{L}_+(\mathrm{i}y_0)| = \left|\int_0^\infty \mathrm{e}^{-\mathrm{i}y_0 u} F_+(\mathrm{d}u)\right| = 1$$

for some $y_0 \in \mathbb{R} \setminus \{0\}$, then it follows readily that the probability with distribution function F_+ is concentrated on a subset of $(c + (2\pi/y_0)\mathbb{Z}) \cap \mathbb{R}_+$ for some $c \in \mathbb{R}$, in which case we also have $|\bar{L}_+(m\mathrm{i}y_0)| = 1$ for all $m \in \mathbb{Z}$. Since

$$\lim_{|m|\to\infty} |\bar{L}_+(m\mathrm{i}y_0)| = \lim_{|m|\to\infty} \left|\frac{P_+(m\mathrm{i}y_0)}{R_+(m\mathrm{i}y_0)}\right| = 0,$$

this yields a contradiction.

It remains to consider those $z \in \partial\Gamma_{\rho,\varepsilon}$ with $|z| = \varepsilon$. For a sufficiently small ε, $\boldsymbol{Q}(z, 0)$ is nonsingular and $P_+(z) \ne 0$; hence, in this case, (73) is equivalent to

$$|\boldsymbol{a}^\top \boldsymbol{Q}^{-1}(z,0)\boldsymbol{\lambda}| < \frac{1}{|\bar{L}_+(z)|},$$

with $\bar{L}_+(z)$ given by (72). The functions

$$z \mapsto g_l(z) := -\boldsymbol{a}^\top \boldsymbol{Q}^{-1}(z,0)\boldsymbol{\lambda}$$

and

$$z \mapsto g_r(z) := \frac{1}{\bar{L}_+(z)}$$

are analytic in a neighbourhood of 0, with $g_l(0) = g_r(0) = 1$. However, if a function g is analytic in a neighbourhood of 0 and $g(0)$, $g'(0)$, and $g''(0)$ are $\mathbb{R}$-valued, then for $z = x + \mathrm{i}y$ close to 0 we have

$$|g(z)|^2 = g^2(0) + 2xg(0)g'(0) + y^2(g'^2(0) - g(0)g''(0)) + o(x) + o(y^2),$$

where a prime denotes differentiation. Hence, the desired inequality $|g_l(z)| < |g_r(z)|$, for $z \in \partial\Gamma_{\rho,\varepsilon}$ with $|z| = \varepsilon$, will follow if we show that, for sufficiently small $x, y \ne 0$,

$$xg_l'(0) < xg_r'(0), \qquad y^2(g_l'^2(0) - g_l''(0)) < y^2(g_r'^2(0) - g_r''(0)),$$

i.e. we must show that

$$g_1'(0) > g_r'(0) \tag{74}$$

in the case $p_r < 1$, that

$$g_1'(0) < g_r'(0) \tag{75}$$

in the case $p_r = 1$, and that

$$g_1'^2(0) - g_1''(0) < g_r'^2(0) - g_r''(0) \tag{76}$$

in both cases.

For a sufficiently small z, we can differentiate with respect to z under the integral sign in (72), to find that

$$g_r'(0) = \xi, \qquad g_r''(0) = -(\mathrm{E}\, U_1^2 - 2\xi^2),$$

meaning that

$$g_r'^2(0) - g_r''(0) = \mathrm{var}(U_1) > 0.$$

Next, differentiating with respect to z in the matrix identity $\boldsymbol{I} = \boldsymbol{Q}^{-1}(z,0)\,\boldsymbol{Q}(z,0)$, and recalling (25), yields

$$\frac{\mathrm{d}}{\mathrm{d}z}\boldsymbol{Q}^{-1}(z,0)\bigg|_{z=0} = -\boldsymbol{Q}_V^{-1}\overline{\boldsymbol{\beta}}\,\boldsymbol{Q}_V^{-1},$$

$$\frac{\mathrm{d}^2}{\mathrm{d}z^2}\boldsymbol{Q}^{-1}(z,0)\bigg|_{z=0} = 2\boldsymbol{Q}_V^{-1}\overline{\boldsymbol{\beta}}\,\boldsymbol{Q}_V^{-1}\overline{\boldsymbol{\beta}}\,\boldsymbol{Q}_V^{-1} - \boldsymbol{Q}_V^{-1}\boldsymbol{S}\,\boldsymbol{Q}_V^{-1},$$

where $\overline{\boldsymbol{\beta}}$ is the diagonal matrix with entries β_i and $\boldsymbol{S}$ the diagonal matrix with entries σ_i^2. Thus, by using (9), we have

$$g_1'(0) = -\boldsymbol{a}^\top \boldsymbol{Q}_V^{-1}\boldsymbol{\beta},$$

$$g_1''(0) = 2\boldsymbol{a}^\top \boldsymbol{Q}_V^{-1}\overline{\boldsymbol{\beta}}\,\boldsymbol{Q}_V^{-1}\boldsymbol{\beta} - \boldsymbol{a}^\top \boldsymbol{Q}_V^{-1}\boldsymbol{\sigma}^2,$$

where $\boldsymbol{\beta}$ and $\boldsymbol{\sigma}^2$ denote the column vectors (β_i) and (σ_i^2), respectively.

We first show (74) and (75): by (61) $g_1'(0) = \mu \sum_i \alpha_i \beta_i$, which is strictly greater than $g_r'(0) = \xi$ when $p_r < 1$ and strictly less than ξ when $p_r = 1$, and there is strict inequality in (62).

The proof of (76) is more difficult. Since the right-hand side is greater than 0, it is enough to show that

$$g_1'^2(0) - g_1''(0) \le 0. \tag{77}$$

Suppose first that $\beta_i = \beta$ and $\sigma_i^2 = \sigma^2$ for all i. Then, for a sufficiently small $|z|$,

$$g_1(z) = L_V(-\beta z - \tfrac{1}{2}\sigma^2 z^2) = \mathrm{E}^{x_0, \boldsymbol{a}} \exp(\beta z + \tfrac{1}{2}\sigma^2 z^2) V_1$$

(cf. (26)), which implies that $g_1'(0) = \beta\mu$, $g_1''(0) = \sigma^2\mu + \beta^2\, \mathrm{E}^{x_0, \boldsymbol{a}}\, V_1^2$, and, therefore,

$$g_1'^2(0) - g_1''(0) = -\beta^2\, \mathrm{var}^{x_0, \boldsymbol{a}}\, V_1 - \sigma^2\mu \le 0.$$

Hence, (77) holds in this particular case. Inspired by this, for arbitrary β_i it is natural to consider the function $\tilde{g}$ given by

$$z \mapsto \tilde{g}(z) = \mathrm{E}^{x_0, \boldsymbol{a}} \exp\left(\int_0^{V_1} \left(\beta_{J_s} z + \frac{1}{2}\sigma_{J_s}^2 z^2\right) \mathrm{d}s\right),$$

which is well defined when $|z|$ is sufficiently small. Differentiation with respect to z may be performed under the expectation sign, and yields

$$
\begin{aligned}
\tilde{g}'(0) &= \mathrm{E}^{x_0,a} \int_0^{V_1} \beta_{J_s} \, \mathrm{d}s \\
&= \mathrm{E}^{x_0,a} \int_0^{\infty} \beta_{J_s} 1_{\{V_1 > s\}} \, \mathrm{d}s \\
&= \int_0^{\infty} a^{\top} \mathrm{e}^{s Q_V} \boldsymbol{\beta} \, \mathrm{d}s \\
&= -a^{\top} Q_V^{-1} \boldsymbol{\beta} \\
&= g_1'(0),
\end{aligned}
$$

where here and below we use the fact that, prior to the first claim, the transition matrix for the Markov chain J over a time interval of length s is $\mathrm{e}^{s Q_V}$. Differentiating once more gives

$$
\tilde{g}''(0) = -a^{\top} Q_V^{-1} \boldsymbol{\sigma}^2 + \mathrm{E}^{x_0,a} \left(\int_0^{V_1} \beta_{J_s} \, \mathrm{d}s \right)^2
$$

and, by straightforward computation, we have

$$
\begin{aligned}
\mathrm{E}^{x_0,a} \left(\int_0^{V_1} \beta_{J_s} \, \mathrm{d}s \right)^2 &= 2 \, \mathrm{E}^{x_0,a} \left(\int_0^{\infty} \beta_{J_s} 1_{\{V_1 > s\}} \, \mathrm{d}s \right) \int_s^{\infty} \beta_{J_t} 1_{\{V_1 > t\}} \, \mathrm{d}t \\
&= 2 \int_0^{\infty} \mathrm{d}s \int_s^{\infty} \mathrm{d}t \, a^{\top} \mathrm{e}^{s Q_V} \overline{\boldsymbol{\beta}} \mathrm{e}^{(t-s) Q_V} \boldsymbol{\beta} \\
&= 2 a^{\top} Q_V^{-1} \overline{\boldsymbol{\beta}} Q_V^{-1} \boldsymbol{\beta},
\end{aligned}
$$

implying that $\tilde{g}''(0) = g_1''(0)$. Therefore,

$$
g_1'^2(0) - g_1''(0) = a^{\top} Q_V^{-1} \boldsymbol{\sigma}^2 - \mathrm{var}^{x_0,a} \int_0^{V_1} \beta_{J_s} \, \mathrm{d}s
$$

and, since $a^{\top} Q_V^{-1} \boldsymbol{\sigma}^2 = -\mu \sum_i \alpha_i \sigma_i^2 \le 0$, by (61), inequalities (77) and, therefore, (76) have been shown.

It remains to consider the case $p_U < 1$. As in the proof of Theorem 1(i), we shall apply Rouché's theorem: it suffices to show that

$$
(1 - p_U)|R_+(z)| \, |a^{\top} Q^*(z,0) \boldsymbol{\lambda} L_-(z)| < |\mathcal{P}_1(z) + p_U \mathcal{P}_2(z)|, \qquad z \in \partial \Gamma_{\rho,\varepsilon} \tag{78}
$$

(cf. (52)), where

$$
\mathcal{P}_1(z) = R_+(z) \det Q(z,0), \qquad \mathcal{P}_2(z) = P_+(z) a^{\top} Q^*(z,0) \boldsymbol{\lambda}.
$$

For $z \in \partial \Gamma_{\rho,\varepsilon}$ with $|z| = \rho$ or $z = \mathrm{i}y$ with $\varepsilon \le |y| \le \rho$, we proceed exactly as in the proof of Theorem 1(i). If $z \in \partial \Gamma_{\rho,\varepsilon}$, with $|z| = \varepsilon$ sufficiently small, (78) is equivalent to $|g_l(z)| < |g_r(z)|$, where

$$
g_l(z) = (1 - p_U) h(z) L_-(z), \qquad g_r(z) = 1 - p_U h(z) L_+(z),
$$

and

$$
h(z) = -a^{\top} Q^{-1}(z,0) \boldsymbol{\lambda}.
$$

It is now quite straightforward to verify (74) and (75), as well as (76).

Example 1. We shall illustrate Theorem 2 in an example: take $E = \{0, 1\}$, $\beta_0 = 2$, $\beta_1 = 0$, $\sigma_0 = 0$, $\sigma_1 = 10$, $\lambda_0 = 1$, $\lambda_1 = 0$, $q_{01} = q_{10} = 1$, $a_0 = 1$, and $a_1 = 0$. When J is in state 0, X follows a straight line with slope 2, and when J is in state 1, X behaves as $\sigma_1 B$. A claim can only arrive when J is in state 0, and J then stays in 0 for a while. Clearly, ruin by continuity is possible only when J is in state 1 and ruin by jump is possible only when J is in state 0: $p = 2$ and $p_c = p_j = 1$.

We shall assume that $p_U = 1$, and for the claims distribution we take a mixture of two exponentials,

$$L_+(\nu) = (1 - \delta)\frac{1}{1 + \nu} + \delta\frac{\eta}{\eta + \nu} = \frac{\eta + (1 - \delta + \delta\eta)\nu}{(1 + \nu)(\eta + \nu)}, \tag{79}$$

with the idea that both δ and η should be so small that on rare occasions a huge claim will appear; below, we take $\delta = 0.01$ but allow η to vary.

We have $\xi = 1 - \delta + \delta/\eta$. Simple calculations yield $\mu = q_{01} + q_{10} = 2$ and the stationary distribution $\alpha_0 = \alpha_1 = \frac{1}{2}$ for J (essentially because $q_{01} = q_{10}$ and $a_0 = 1$). Thus, the condition (62) necessary and sufficient for $p_r = 1$ becomes $\xi \geq \alpha_0\beta_0\mu = 2$, which for $\delta = 0.01$ translates into $\eta \leq \frac{1}{101}$.

The matrix $Q(z, 0)$ has the form

$$Q(z, 0) = \begin{pmatrix} 2z - 2 & 1 \\ 1 & \frac{1}{2}\sigma_1^2 z^2 - 1 \end{pmatrix}$$

(cf. (25)), which means that the modified Cramér–Lundberg equation (66) becomes

$$R_+(\gamma)((2\gamma - 2)(\tfrac{1}{2}\sigma_1^2\gamma^2 - 1) - 1) = -P_+(\gamma)(\tfrac{1}{2}\sigma_1^2\gamma^2 - 1) \tag{80}$$

with (see (79))

$$R_+(\gamma) = (1 + \gamma)(\eta + \gamma), \qquad P_+(\gamma) = \eta + (0.99 + 0.01\eta)\gamma.$$

Solving (80) amounts to finding the roots of a polynomial of degree five for which 0 is always a root. In the cases considered below, the five roots are real, with of course precisely three roots being less than 0 if and only if $p_r < 1$, and precisely two roots being less than 0 if and only if $p_r = 1$. For $\eta = \frac{1}{101}$, the critical value 0 becomes a root of multiplicity two.

Table 1, below, collects the ruin probabilities by continuity and by jump for $\eta = 0.0001$, 0.001, 0.008, 0.012, 0.1, 0.99 for initial values $x_0 = 0.1, 1, 10, 100$ and $i_0 = 0, 1$. (The numbers 0.008 and 0.012 were picked for η to represent two values fairly close to and on either side of the critical value $\frac{1}{101}$.) The table illustrates, in particular, the impact that a high volatility has on the probability of ruin by continuity when x_0 is small, and how the presence of rare huge claims affects the ruin probability by jump. In each cell, the left-hand number is the probability of ruin by continuity and the right-hand number the probability of ruin by jumps.

In the case $p_r = 1$, having found the three negative roots of (80), we must solve two linear equations with two unknowns. There are three choices of equation, and it has been checked numerically that the solutions for the ruin probabilities do not depend on which two of these are chosen.

We conclude this section with some comments on how the Laplace transform for the undershoot alone may be determined, i.e. we find the quantities $E^{x_0, i_0}[e^{-\zeta Y_r}; A_j]$ jointly with

TABLE 1: The ruin probabilities p_c (*left*) and p_j (*right*) for Example 1.

		η					
x_0	i_0	0.0001	0.001	0.008	0.012	0.1	0.99
0.1	0	0.692 0.308	0.692 0.308	0.698 0.302	0.699 0.297	0.700 0.282	0.702 0.277
0.1	1	0.997 0.003	0.997 0.003	0.997 0.003	0.997 0.002	0.997 0.001	0.997 0.0008
1	0	0.793 0.207	0.794 0.206	0.804 0.196	0.806 0.187	0.807 0.159	0.811 0.151
1	1	0.969 0.031	0.969 0.031	0.974 0.026	0.975 0.021	0.975 0.007	0.975 0.006
10	0	0.734 0.266	0.741 0.259	0.789 0.211	0.802 0.165	0.803 0.024	0.800 0.010
10	1	0.761 0.239	0.767 0.233	0.811 0.189	0.823 0.146	0.820 0.020	0.815 0.010
100	0	0.072 0.928	0.100 0.900	0.310 0.690	0.345 0.468	0.170 0.006	0.142 0.002
100	1	0.074 0.926	0.103 0.897	0.312 0.688	0.346 0.467	0.173 0.006	0.144 0.002

the ruin probabilities $\mathrm{P}^{x_0,i_0}(A_{c,i}) = \mathrm{E}^{x_0,i_0}[e^{-\zeta Y_r}; A_{c,i}]$. For $\zeta > 0$, these $p_c + 1$ unknowns always require a system of $p_c + 1$ linear equations for their solution – in the case $p_r = 1$, we no longer know that they sum to 1. The idea is now to solve the Cramér–Lundberg equation (30) or (31) (that do not depend on ζ) for $\theta = 0$, i.e. to solve (64) or (65). If $p_r < 1$ then we know that we obtain $m + p_c + 1$ solutions γ_k with $\mathrm{Re}(\gamma_k) < 0$. We then proceed as in part (iv) of Theorem 1, using (34) with $\theta = 0$ and the given ζ. If $p_r = 1$, we use the $m + p_c$ roots of (65) *and* the root 0 substituted into (34), which is possible because the r_k given by (35) are well defined for both $\zeta > 0$ and $\tilde{\gamma}_k = 0$.

Appendix A. Proofs of Lemma 1 and Lemma 2

A.1. Proof of Lemma 1

For any given complex numbers z_i, $i \in E$, define the matrix $\hat{Q} = (\hat{q}_{ij})$ by

$$\hat{q}_{ij} = \begin{cases} z_i - q_i - \lambda_i, & \text{if } i = j, \\ q_{ij}, & \text{if } i \neq j, \end{cases}$$

meaning that $Q(z, \theta) = \hat{Q}$ if $z_i = \phi_i(z) - \theta$ for all i; see (25). Since $\mathrm{Re}(\phi_i(\mathrm{i}y)) \leq 0$, in order to prove the lemma it suffices to show that if $\mathrm{Re}(z_i) \leq 0$ for all i, then $\hat{Q}$ is nonsingular and

$$|u_j| \leq \frac{q_j + \lambda_j}{q_j + \lambda_j + c}, \qquad j \in E, \tag{81}$$

where $c = \min | \mathrm{Re}(z_i)|$ and

$$u_j = (\hat{Q}^{-1}\lambda)_j.$$

This we now proceed to do, assuming, from now on, that $\mathrm{Re}(z_i) \leq 0$ for all i.

To argue that $\hat{Q}$ is nonsingular, suppose that $v = (v_j)$ is a column vector such that $\hat{Q}v = 0$. Then

$$\sum_{j \neq i} q_{ij}v_j + \hat{q}_{ii}v_i = 0$$

for all i, and, since $\mathrm{Re}(\hat{q}_{ii}) < 0$ (because $\mathrm{Re}(z_i) \leq 0$ and $q_i + \lambda_i > 0$ by Assumption 1), we have $\hat{q}_{ii} \neq 0$ and, hence,

$$v_i = \frac{\sum_{j \neq i} q_{ij}v_j}{-\hat{q}_{ii}},$$

which implies that

$$|v_i| \leq \frac{q_i \max |v_j|}{q_i + \lambda_i + c}.$$

If $i \equiv i_0$ is now chosen such that $|v_{i_0}| = \max |v_j|$, and if we assume that $|v_{i_0}| > 0$, then we see that $q_{i_0} > 0$, $\lambda_{i_0} = c = 0$, and $|v_{j_0}| = |v_{i_0}|$ for all $j_0 \neq i_0$ such that $q_{i_0 j_0} > 0$. We then also have $q_{j_0} > 0$ and $\lambda_{j_0} = 0$ and it follows that, for all states j reachable from i_0 through q_{ij}-transitions alone, we must have $q_j > 0$ and $\lambda_j = 0$. However, by the irreducibility inherent in Assumption 1, there is some j_1 with $\lambda_{j_1} > 0$ that is reachable from i_0 by q_{ij}-transitions; from this contradiction we deduce that $\max |v_j| = 0$, as required.

To show (81), first rewrite it as

$$u_i = \frac{1}{-\hat{q}_{ii}} \left(\lambda_i + \sum_{j \neq i} q_{ij} u_j \right),$$

which implies that

$$|u_i| \leq \frac{\lambda_i + q_i \max |u_j|}{\lambda_i + q_i + c} \tag{82}$$

for all i. Now consider the states i_0 with $|u_{i_0}| = \max |u_j|$ (which is greater than 0 since $\lambda \neq \mathbf{0}$ and $(u_j) =: \boldsymbol{u} = -\hat{\boldsymbol{Q}}^{-1} \lambda$). We first argue that, for some such i_0, $\lambda_{i_0} + c > 0$. Were $\lambda_{i_0} + c = 0$ for all these i_0, we would have

$$u_{i_0} = \frac{1}{q_{i_0}} \sum_{j \neq i} q_{i_0 j} u_j,$$

implying that $|u_j| = |u_{i_0}|$ for all $j \neq i_0$ with $q_{i_0 j} > 0$; hence, by the assumption that all $\lambda_{i_0} + c = 0$, we also have $\lambda_j + c = 0$. The irreducibility in Assumption 1 would then yield a contradiction, exactly as in the first part of the proof. Thus, for some i_0 with $|u_{i_0}| = \max |u_j|$ we have $\lambda_{i_0} + c > 0$, and from

$$|u_{i_0}| \leq \frac{\lambda_{i_0} + q_{i_0} |u_{i_0}|}{\lambda_{i_0} + q_{i_0} + c},$$

which is a trivial consequence of (82), it follows that $|u_{i_0}| \leq \lambda_{i_0} / (\lambda_{i_0} + c) \leq 1$. Using this in (82) gives

$$|u_i| \leq \frac{\lambda_i + q_i}{\lambda_i + q_i + c}$$

for all i, proving (81).

Note that, if we set $z_i = 0$ for all i, the proof shows, in particular, that the subintensity matrix $\boldsymbol{Q}_V$ is nonsingular, as was also shown in [11, Lemma 1].

A.2. Proof of Lemma 2

Define $E' = \{i \in E : \sigma_i^2 > 0 \text{ or } \beta_i \neq 0\}$, as in (55). We have

$$\det \boldsymbol{Q}(z, \theta) = \sum_\pi \text{sgn}(\pi) \prod_{i \in E} q_{i, \pi(i)}(z, \theta), \tag{83}$$

where π ranges over the set of all permutations of the states in E. Because z appears only in the diagonal elements $q_{ii}(z, \theta)$, and then only if $i \in E'$, it is clear that a term in the sum in (83) is always a polynomial of degree less than or equal to d, with d given by (38), and that this

polynomial is of degree d if and only if $\pi(i) = i$ for all $i \in E'$. Adding all the terms of degree d gives a coefficient of z^d equal to

$$\det \boldsymbol{Q}_{E \setminus E'}(\theta) \prod_{\{i \in E' : \sigma_i^2 > 0\}} \left(\frac{1}{2}\sigma_i^2\right) \prod_{\{i \in E' : \sigma_i^2 = 0\}} \beta_i, \tag{84}$$

where, recall, $\boldsymbol{Q}_{E \setminus E'}(\theta)$ is the square matrix obtained from $\boldsymbol{Q}(z, \theta)$ by deleting all rows and columns corresponding to states in E', and does not depend on z. However, the matrix $\boldsymbol{Q}_{E \setminus E'}(\theta)$ is a subintensity matrix; hence, by Lemma 1 of [11], it is nonsingular and we have shown that $\det \boldsymbol{Q}(z, \theta)$ is a polynomial exactly of degree d.

To show that the number of roots z of $\det \boldsymbol{Q}(\cdot, \theta)$ with $\mathrm{Re}(z) < 0$ equals p_{c} (given in (12)), we argue as follows. With E' defined as above, first consider the matrix $\bar{\boldsymbol{Q}}(z, \theta)$ obtained from $\boldsymbol{Q}(z, \theta)$ by replacing all the off-diagonal elements q_{ij}, $i \in E'$, $j \in E$, by 0. Then

$$\det \bar{\boldsymbol{Q}}(z, \theta) = \det \boldsymbol{Q}_{E \setminus E'}(\theta) \prod_{i \in E'} (\phi_i(z) - q_i - \lambda_i - \theta). \tag{85}$$

For those $i \in E'$ with $\sigma_i^2 > 0$, the factor $\phi_i(z) - q_i - \lambda_i - \theta$ is a polynomial of degree 2 and, since this polynomial takes the value $-q_i - \lambda_i - \theta < 0$ for $z = 0$ (remember that, by Assumption 1, $\lambda_i + q_i > 0$ for all i), with the limit ∞ as $z \to \pm\infty$, it follows that it has two real roots, one less than 0 and one greater than 0. For those $i \in E'$ with $\sigma_i^2 = 0$, the factor $\phi_i(z) - q_i - \lambda_i - \theta$ is a polynomial of degree 1 and it is immediate that the root of this polynomial is less than 0 if $\beta_i < 0$ and greater than 0 if $\beta_i > 0$. Thus, all roots of the polynomial (85) are real and precisely p_{c} of them are less than 0.

For the general case in which $\boldsymbol{Q}(z, \theta)$ is given by (25), consider the map $s \mapsto \det \boldsymbol{Q}_s(\cdot, \theta)$, defined for $s \in [0, 1]$, where the elements of the matrix $\boldsymbol{Q}_s(z, \theta)$ are given by

$$q_{s,ij}(z, \theta) = \begin{cases} q_{ij}(z, \theta), & \text{if } i \in E \setminus E', \ j \in E, \\ \phi_i(z) - q_i - \lambda_i - \theta, & \text{if } i = j \in E', \\ s q_{ij}, & \text{if } i \in E', \ j \in E, \ i \neq j, \end{cases}$$

meaning that $\boldsymbol{Q}_1(z, \theta) = \boldsymbol{Q}(z, \theta)$ and $\det \boldsymbol{Q}_0(z, \theta) = \det \bar{\boldsymbol{Q}}(z, \theta)$. As s varies, the leading coefficient of the polynomial $\det \boldsymbol{Q}_s(\cdot, \theta)$ is always the same, and is given by (84); it hence follows that the roots of $\det \boldsymbol{Q}_s(\cdot, \theta)$ (when ordered lexicographically, say) are continuous functions of s. By Lemma 1, however, for every s, $\det \boldsymbol{Q}_s(\cdot, \theta)$ has no roots on the line $i\mathbb{R}$. Thus, as s varies from 0 to 1 none of the roots can cross from the strictly negative half of the complex plane ($\mathrm{Re}(z) < 0$) to the strictly positive half ($\mathrm{Re}(z) > 0$). However, as shown at the start of the proof, $\det \boldsymbol{Q}_0(\cdot, \theta)$ has precisely p_{c} roots z with $\mathrm{Re}(z) < 0$ and, hence, so also does $\det \boldsymbol{Q}_1(\cdot, \theta)$.

Acknowledgements

References

[1] ASMUSSEN, S. (2000). *Ruin Probabilities* (Adv. Ser. Statist. Sci. Appl. Prob. **2**). World Scientific, River Edge, NJ.

[2] ASMUSSEN, S., AVRAM, F. AND PISTORIUS, M. R. (2004). Russian and American put options under exponential phase-type Lévy models. *Stoch. Process. Appl.* **109,** 79–111.

[3] AVRAM, F. AND USÁBEL, M. (2004). Ruin probabilities and deficit for the renewal risk model with phase-type interarrival times. *Astin Bull.* **34,** 315–332.

[4] AVRAM, F., KYPRIANOU, A. E. AND PISTORIUS, M. R. (2004). Exit problems for spectrally negative Lévy processes and applications to (Canadized) Russian options. *Ann. Appl. Prob.* **14,** 215–238.

[5] BERTOIN, J. (1997). Exponential decay and ergodicity of completely asymmetric Lévy processes in a finite interval. *Ann. Appl. Prob.* **7,** 156–169.

[6] DAVIS, M. H. A. (1993). *Markov Models and Optimization* (Monogr. Statist. Appl. Prob. **49**). Chapman and Hall, London.

[7] DICKSON, D. C. M. AND HIPP, C. (1998). Ruin probabilities for Erlang(2) risk processes. *Insurance Math. Econom.* **22,** 251–262.

[8] EMBRECHTS, P., GRANDELL, J. AND SCHMIDLI, H. (1993). Finite-time Lundberg inequalities in the Cox case. *Scand. Actuarial J.* **1993,** 17–41.

[9] EMERY, D. J. (1973). Exit problem for a spectrally positive process. *Adv. Appl. Prob.* **5,** 498–520.

[10] GUSAK, D. V. (1969). On the joint distribution of the first exit time and exit value for homogeneous processes with independent increments. *Theory Prob. Appl.* **14,** 14–23.

[11] JACOBSEN, M. (2003). Martingales and the distribution of the time to ruin. *Stoch. Process. Appl.* **107,** 29–51.

[12] JACOBSEN, M. (2005). *Point Processes, Theory and Applications: Marked Point and Piecewise Deterministic Processes*. Birkhäuser, Boston, MA.

[13] KOU, S. G. AND WANG, H. (2003). First passage times of a jump diffusion process. *Adv. Appl. Prob.* **35,** 504–531.

[14] KYPRIANOU, A. E. AND PALMOWSKI, Z. (2003). Fluctuations of spectrally negative Markov additive processes. Submitted.

[15] LAST, G. AND BRANDT, A. (1995). *Marked Point Processes on the Real Line. The Dynamic Approach* (Prob. Appl. (NY)). Springer, New York.

[16] PAULSEN, J. AND GJESSING, H. K. (1997). Ruin theory with stochastic return on investments. *Adv. Appl. Prob.* **29,** 965–985.

[17] TOLVER JENSEN, A. (2004). The distribution of various hitting times for a shot noise process with two-sided jumps. Preprint 6, Department of Applied Mathematics and Statistics, University of Copenhagen.

[18] WINKEL, M. (2005). Electronic foreign-exchange markets and passage events of independent subordinators. *J. Appl. Prob.* **42,** 138–152.