Dipl.-Ing. Simon Pojer, BSc

Analysis of risk models driven by certain Poisson cluster processes

DOCTORAL THESIS

to achieve the university degree of

Doktor der Naturwissenschaften

submitted to

Graz University of Technology

Supervisor

Stefan Thonhauser, Univ.-Prof. Dipl.-Ing. Dr.techn.

Institute of Statistics

Graz, April 2023

AFFIDAVIT

I declare that I have authored this thesis independently, that I have not used other than the declared sources/resources, and that I have explicitly indicated all material which has been quoted either literally or by content from the sources used. The text document uploaded to TUGRAZonline is identical to the present doctoral thesis.

Date, Signature

Abstract

Since 1903, the field of ruin theory has provided a time-dynamic point of view on the problem of insolvency of an insurance portfolio. In his ground-breaking introduction of the classical risk process, Fillip Lundberg used a Poisson process to model the occurrence times of claims. This assumption implies that the inter-claim times are i.i.d. exponentially distributed, which is hardly supported by data.

If claims appear due to natural disasters, it is more realistic to use a counting process driven by a Markovian shot-noise intensity. Regarding the behaviour of the ruin probability in this model, existing results imply the existence of an upper and lower bound, both exponentially decreasing at the same rate. Assuming a recurrent structure of the intensity process, we use a renewal argument to show that the ruin probability has a Lundberg-type convergence behaviour. Exploiting a specific version of the second Borel-Cantelli lemma, we further show that the needed recurrence of the Markovian shot-noise intensity exists as long as the shock events are light-tailed.

Moreover, we consider risk models driven by linear marked Hawkes processes. The main feature of these processes is their self-exciting structure, a behaviour which can be observed in the occurrence of claims due to cyber risks. Interested in improving the existing result of logarithmic convergence, we show that the corresponding ruin probabilities are bounded from above and below by ruin probabilities of a related Cramér-Lundberg process. Given the claim size distribution is strongly subexponential, these inequalities imply the convergence of the ruin probabilities of the Hawkes model. To derive the exact asymptotics in the light-tailed case, we restrict ourselves to the Markovian model, i.e. a Hawkes model with an exponential decay function. At first, we show that the intensity is positive Harris recurrent and that the corresponding recurrence times are light-tailed. Then, we use similar renewal arguments as in the Markovian shot-noise model to get the exact asymptotic behaviour of the ruin probabilities.

By its simplicity, the ruin probability gives only limited insights into the structure of the event of ruin. To resolve this, we extend the concept of Gerber-Shiu functions to the Markovian shot-noise model. Unfortunately, these risk functionals can hardly be expressed explicitly. Since simulation is time-consuming, we develop a numerical scheme, which provides an efficient way to calculate the desired values. It is based on calculating expected values of appropriate Markov chains, by solving a system of linear equations. By showing that these Markov chains converge weakly against the piecewise deterministic Markov processes specifying our model, we prove that the numerical approximations converge to the required Gerber-Shiu functions.

Kurzfassung

Seit 1903 liefert das Gebiet der Ruintheorie einen zeitdynamischen Blickwinkel auf das Problem der Insolvenz von Versicherungsportfolios. In seiner bahnbrechenden Einführung des klassischen Risikoprozesses nutzte Fillip Lundberg einen Poisson-Prozess, um das Erscheinen von Schadensereignissen zu modellieren. Diese Annahme impliziert, dass die Zwischenschadenszeiten i.i.d. exponentialverteilt sind, eine Eigenschaft, die kaum durch reale Daten gestützt wird.

Wenn Schäden durch Naturkatastrophen verursacht werden, ist es realistischer Zählprozesse mit Markowscher Shot-Noise Intensität zu nutzen. Bisherige Resultate bezüglich des asymptotischen Verhaltens der Ruinwahrscheinlichkeit in diesem Modell implizieren die Existenz einer oberen und unteren Schranke, die beide mit der gleichen Rate exponentiell abfallen. Unter der Annahme einer rekurrenten Struktur des Intensitätsprozesses nutzen wir Erneuerungsargumente, um zu zeigen, dass die Ruinwahrscheinlichkeit ein Lundberg-artiges Konvergenzverhalten besitzt. Durch eine spezifische Version des zweiten Borel-Cantelli-Lemmas zeigen wir weiters, dass die benötigte Rekurrenz der Markowschen Shot-Noise Intensität existiert, solange die momenterzeugende Funktion der Schockereignisse existiert.

Weiters betrachten wir Risikoprozesse, die durch lineare Hawkes-Prozesse angetrieben werden. Das Hauptmerkmal dieser Prozesse ist ihre selbstanregende Struktur, ein Verhalten, das auch bei Schäden durch Cyberrisiken beobachtet werden kann. Interessiert daran, die bestehenden Ergebnisse der logarithmischen Konvergenz zu verbessern, zeigen wir, dass die dazugehörigen Ruinwahrscheinlichkeiten von oben und unten durch die Ruinwahrscheinlichkeiten eines verwandten Cramér-Lundberg-Prozesses beschränkt werden können. Gegeben, die Schadenshöhenverteilung ist stark subexponentiell, implizieren diese Ungleichungen die Konvergenz der Ruinwahrscheinlichkeiten des Hawkes-Modells. Um das exakte asymptotische Verhalten im Fall einer Schadenshöhenverteilung mit exponentiell abfallender Überlebensfunktion zu erlangen, beschränken wir uns auf das Markowsche Modell, also ein Hawkes-Modell mit exponentieller Anregungsfunktion. Zuerst zeigen wir, dass die Intensität Harris-rekurrent ist und die entsprechenden Rekurrenzzeiten endliche momenterzeugende Funktion besitzen. Dann nutzen wir ähnliche Erneuerungsargumente wie im Shot-Noise Fall, um das genaue asymptotische Verhalten der Ruinwahrscheinlichkeiten zu erhalten.

Durch ihre Einfachheit gibt die Ruinwahrscheinlichkeit nur limitierten Einblick in die Struktur des Ruinereignisses. Um dies zu beheben, erweitern wir das Konzept der Gerber-Shiu Funktionen auf das Markowsche Shot-Noise Modell. Unglücklicherweise

können diese Risikofunktionale kaum explizit ausgedrückt werden. Da Simulationen zeitaufwendig sind, entwickeln wir eine numerische Methode, die eine effiziente Berechnung der gewünschten Werte ermöglicht. Sie basiert auf der Bestimmung von Erwartungswerten geeigneter Markow-Ketten, durch Lösen eines Systems linearer Gleichungen. Wir zeigen, dass diese Markow-Ketten schwach gegen stückweise deterministische Markow-Prozesse konvergieren, die unser Modell beschreiben und erhalten dadurch die Konvergenz der numerischen Approximation gegen die gefragten Gerber-Shiu Funktionen.

Contents

List of Figures

List of Tables

1 Introduction

To prevent insolvencies of insurance companies and the associated economic shocks, the European Union enacted the Solvency II Directive 2009, which came into effect in 2016. Based on this directive and the corresponding national laws, the insurers must hold enough capital to satisfy a certain Solvency Capital Requirement based on their 99.5% one-year Value at Risk. By the static nature of the VaR, this method might not always be satisfactory, especially if we consider a larger time horizon. A time-dynamic approach to the estimation of insolvency risk is provided by the field of ruin theory. Here, stochastic processes are used to model the surplus of an insurance portfolio, which allows us to determine the theoretical properties of corresponding risk functionals. In this work, we extend some of the existing results of ruin theory to models driven by specific counting processes, namely the Markovian shot-noise process and the linear marked Hawkes process. By the cumulative structure of this thesis, it is composed of the following four self-sufficient scientific papers:

- Chapter 2 is based on the published paper: Simon Pojer and Stefan Thonhauser (2023a). "Ruin probabilities in a Markovian shot-noise environment". In: *Journal of Applied Probability* 60.2, pp. 542–556. DOI: `10.1017/jpr.2022.63`,
- Chapter 3 is based on the submitted paper: Simon Pojer (2022). "Level crossings of the Markovian shot-noise process". DOI: `10.2139/ssrn.4285543`. Submitted,
- Chapter 4 is based on the submitted paper: Zbigniew Palmowski, Simon Pojer, and Stefan Thonhauser (2023). "Exact asymptotics of ruin probabilities with linear Hawkes arrivals". DOI: `10.48550/arXiv.2304.03075`. Submitted,
- Chapter 5 is based on the published paper: Simon Pojer and Stefan Thonhauser (2023b). "The Markovian shot-noise risk model: a numerical method for Gerber-Shiu functions". In: *Methodology and Computing in Applied Probability* 25.1, p. 17. DOI: `10.1007/s11009-023-10001-w`.

Preliminary, we give a short introduction to ruin theory to motivate our research and contextualize our findings. A more detailed discussion of existing results can be found in Asmussen and Albrecher (2010), Grandell (1991), Rolski et al. (1999), Schmidli (2017) and Asmussen and Steffensen (2020).

An introduction to ruin theory

The pioneering definition of the Cramér-Lundberg process in Lundberg (1903) laid the foundations of modern ruin theory. In this model, the stochastic process describing the surplus of an insurance portfolio evolves as follows: the starting value is a non-negative initial capital u, and the value increases through premium payments, which are assumed to be linear in time with a constant rate $c > 0$. At random times, governed by a Poisson process $\{N_t\}_{t\geq 0}$ with constant jump intensity $\rho > 0$, claims occur which have to be paid for by the insurance company. The claim sizes $\{U_i\}_{i\in\mathbb{N}}$ are assumed to be independent and identically distributed (i.i.d.) and independent of the driving Poisson process. By this, the surplus process at time $t \geq 0$ can be written as

$$X_t = u + ct - \sum_{i=1}^{N_t} U_i.$$

As a basic measure for the risk of insolvency in this model, we consider the ruin probability, i.e. the probability that the surplus process $\{X_t\}_{t\geq 0}$ gets eventually negative. To stress the dependence on the initial capital u, we write $\tau_u = \{\, t \geq 0 \mid X_t < 0 \,\}$ for the time of ruin and $\psi(u) = \mathbb{P}[\tau_u < \infty]$ for the corresponding ruin probability. To avoid the trivial case $\psi(u) = 1$ for all $u \geq 0$, we assume that the net profit condition $c > \rho\,\mathbb{E}[U_1]$ holds.

Unfortunately, the ruin probability can hardly be expressed in a closed form. Therefore, it is convenient to use approximations, which are easy to calculate and behave asymptotically like the ruin probability as the initial capital tends to infinity. If the distribution of the claim sizes is light-tailed, i.e. if there is a positive constant r such that $\mathbb{E}\left[e^{rU_1}\right] < \infty$, and the moment-generating function is *well-behaved*, then, there exists a constant $C > 0$ and an adjustment coefficient $R > 0$ such that $\lim_{u\to\infty} \psi(u)e^{Ru} = C$. Exploiting this behaviour, we get for large initial capital u the so-called Cramér-Lundberg approximation

$$\psi(u) \sim Ce^{-Ru}.$$

One of the main assumptions of the classical risk model is that the claims arrive at the jump times of a Poisson process. Consequently, the inter-claim times are i.i.d. exponentially distributed. As an intuitive generalization of this, the Sparre Andersen model, also called the renewal model, was introduced by Sparre Andersen (1957). In this model, claims arrive according to a renewal process, a jump process whose inter-jump times are i.i.d., but not necessarily exponentially distributed. Further examples of risk models with different arrival processes are the Markov modulated model considered in Asmussen (1989), where the jump rate depends on the state of an underlying Markov chain, and the Björk-Grandell model defined in Björk and Grandell (1988), where the intensity of the arrival process is an i.i.d. sequence of positive random variables. In all of these models, there exist conditions under which the ruin probabilities have the same asymptotic behaviour as in the classical model, i.e. there exists also an adjustment coefficient $R > 0$ such that $\lim_{u\to\infty} \psi(u)e^{Ru} = C$ and the Cramér-Lundberg approximation is reasonable.

Over the last 30 years, the insured losses due to natural catastrophes had an average growth rate of 5 to 7 percent per year, and reached USD 125 billion in 2022. This increase of the loss size and the corresponding thread for the insurers is assumed to continue, driven by several factors such as the concentration of assets in vulnerable regions, the rising inflation, and climate change due to global warming, see Banerjee et al. (2023). One way to translate these risks in a mathematical model is to use counting processes with shot-noise intensity to describe the occurrence of claims due to natural disasters, as it was done by Dassios and Jang (2003) and Schmidt (2014). This family of Cox processes have the following dynamics: at random times, governed by an underlying Poisson process, a catastrophe, like an earthquake or hurricane, occurs and increases the corresponding intensity process by a random quantity. This shock triggers several jumps of the counting process, which do not appear simultaneously but are delayed over time. As in reality, the effects of a disaster decline over time, e.g. an earthquake from ten years ago will hardly cause new claims.

Motivated by this, we aim in Chapter 2 to extend the existing convergence results to a risk model whose arrivals are driven by a Markovian shot-noise process. In this setting, the structure of the surplus process is the same as in the classical model, but the arrival process has an intensity $\{\lambda_t\}_{t\geq 0}$ of the following form:

$$\lambda_t := \lambda\, e^{-\beta t} + \sum_{i=1}^{N_t^\rho} Y_i e^{-\beta(t-T_i^\rho)}.$$

Here, $\{N_t^\rho\}_{t\geq 0}$ is a Poisson process with constant rate $\rho > 0$ and corresponding jump times $\{T_i^\rho\}_{i\in\mathbb{N}}$. Further, $\lambda > 0$ is a deterministic initial value, and $\{Y_i\}_{i\in\mathbb{N}}$ an i.i.d. sequence of positive random variables independent of the Poisson process. The asymptotic behaviour in a general shot-noise environment was already studied by Albrecher and Asmussen (2006), where the authors were able to find an adjustment coefficient $R > 0$ and two positive constants C and C_- such that $Ce^{-Ru} \geq \psi(u) \geq C_- e^{-Ru}$. We enhance this result for the Markovian model with light-tailed claims and show convergence of the corresponding ruin probabilities as in the classical model, i.e. there exists again a constant C such that $\psi(u)e^{Ru} \to C$ as $u \to \infty$. To obtain this result, we assume the existence of a recurrent structure of the underlying intensity process, which exists if the shock events are exponentially distributed, as showed by Orsingher and Battaglia (1982). In Chapter 3, we refine this result and prove that the Markovian shot-noise process behaves recurrently as long as the shock events are light-tailed.

In Chapter 4, we consider a risk model whose arrival process is a linear marked Hawkes process $\{N_t\}_{t\geq 0}$. This process has a stochastic intensity of the form

$$\lambda_t := a + \sum_{i=1}^{N_t} h(t - T_i, Y_i),$$

where h is a positive function, which decays fast enough to 0. Further, $\{Y_i\}_{i\in\mathbb{N}}$ is again assumed to be an i.i.d. sequence of random variables and $a > 0$ is a constant baseline. As before, shocks appear, which increase the intensity process at random times and trigger claims. In contrast to the shot-noise process, the shock events do not follow an external Poisson process. Instead, the Hawkes process increases its own intensity, a feature called self-exciting. This property can be observed in social media, as studied, for example, by Lukasik et al. (2016), and Rizoiu et al. (2017), the dynamics of a pandemic, see exemplary Chiang et al. (2022), and Garetto et al. (2021), and the occurrence of cyber risks, as described by Bessy-Roland et al. (2021), and Hillairet et al. (2023).

Using a large deviation argument, Karabash and Zhu (2015) showed that the logarithm of the ruin probabilities in a linear Hawkes model with light-tailed claims converge, i.e. that there exists a constant R such that $\lim_{u\to\infty} \frac{1}{u} \log(\psi(u)) = -R$. By similar arguments as in Albrecher and Asmussen (2006), we prove that there exists a constant l_1 such that

$$\tilde{\psi}(u) \geq \psi(u) \geq \tilde{\psi}(u + cl_1),$$

where $\tilde{\psi}(u)$ denotes the ruin probability in a related Cramér-Lundberg model. Under the assumption of a strongly subexponential claim size distribution, the derived inequality is equivalent to

$$\lim_{u\to\infty} \frac{\psi(u)}{1 - F^s_{\tilde{U}}(u)} = \frac{\rho}{1-\rho},$$

where $F^s_{\tilde{U}}(u)$ denotes the integrated tail distribution of the claim size distribution of a whole cluster. In the light-tailed case, the above inequality is equivalent to the existence of positive constants C and C_- such that

$$Ce^{-Ru} \geq \psi(u) \geq C_- e^{-Ru}.$$

To derive stronger results, we restrict ourselves to the Markovian setting,

$$\lambda_t = (\lambda - a)e^{-\beta t} + a + \sum_{i=1}^{N_t} Y_i e^{-\beta(t-T_i)}.$$

In this case, we can exploit a Harris recurrent structure of the intensity process to derive the exact exponential asymptotic behaviour of the ruin probabilities.

As a basic measure of risk, the ruin probability gives only limited insights into the structure of the event of ruin. To resolve this, the concept of discounted penalty functions, also called Gerber-Shiu functions, was introduced by Gerber and Shiu (1998). For this, they chose suitable penalty functions $w : \mathbb{R}^2_+ \to \mathbb{R}$ and a discounting factor κ and defined the corresponding discounted penalty function for some initial capital u as

$$g(u) := \mathbb{E}\left[w(X_{\tau_u-}, -X_{\tau_u})\, e^{-\kappa\tau_u} I_{\{\,\tau_u<\infty\,\}}\right].$$

These functions allow for a more flexible and detailed description of the event of ruin, using the surplus before ruin $X_{\tau_u -}$, the deficit at ruin X_{τ_u}, and the time of ruin τ_u itself.

Since they are generalizations of the ruin probability, Gerber-Shiu functions share one disadvantage: generally, there is no explicit way to calculate them. Since simulation methods are time-consuming, there is an increasing effort to find efficient numerical methods to determine suitable approximations of the desired values. Recently, Strini and Thonhauser (2020) developed a numerical scheme to calculate approximations of certain discounted penalty functions in the Sparre Andersen model by exploiting a Markovian structure obtained by backward Markovization.

Based on this approach, we develop in Chapter 5 a method to numerically determine Gerber-Shiu functions in the Markovian shot-noise model. For this, we exploit the structure of the underlying model, which is fully specified by piecewise deterministic Markov processes (PDMPs). We identify suitable functionals of certain Markov chains, which can be determined by solving a system of linear equations and approximating the desired Gerber-Shiu functions. Eventually, we show that the underlying processes converge weakly against the original PDMPs; hence, the corresponding expectations converge too.

2 Ruin probabilities in a Markovian shot-noise environment

The following chapter was published as Pojer and Thonhauser (2023a) and is adopted almost verbatim. Some changes have been made to ensure consistency of notation throughout all chapters of this thesis.

2.1 Introduction

The theory of doubly stochastic Poisson processes described in Brémaud (1981), allows the generalization of the well-known Cramér-Lundberg model to the broad class of Cox models, which are discussed e.g. in Grandell (1991). Members of this family are for example the Markov-modulated risk model, where the intensity is modelled by a continuous-time Markov chain which can be found in Chapter VII of Asmussen and Albrecher (2010) and Chapter 8 of Rolski et al. (1999), the Björk-Grandell model considered in Schmidli (1997) and diffusion-driven models studied in Grandell and Schmidli (2011).
Especially, arrivals of claims caused by catastrophic events can be realistically modelled using shot-noise intensity. This has been done in Albrecher and Asmussen (2006), Dassios, Jang, and Zhao (2015) and Macci and Torrisi (2011) where the asymptotic behaviour of the ruin probability in general shot-noise environments was studied. In these settings, upper and lower bounds could be derived. The idea of applying the theory of piecewise deterministic Markov processes to a Cox model with Markovian shot-noise intensity was used in Dassios and Jang (2003) and Dassios and Jang (2005) in the context of pricing reinsurance contracts.

Interested in the behaviour of the ruin probability in this model, we follow the PDMP approach to find suitable alternative probability measures. Further, we take advantage of the properties of the process under these measures to obtain an exponentially decreasing upper bound. Exploiting a recurrent behaviour of the shot-noise process and applying the extended renewal theory obtained in Schmidli (1997) we eventually derive the exact asymptotic behaviour of the ruin probability.

2.2 The Markovian Shot-Noise Ruin Model

We assume for the rest of this paper the existence of a complete probability space $(\Omega, \mathcal{F}, \mathbb{P})$ which is big enough to contain all mentioned stochastic processes and random variables. For some stochastic process Z we denote the right continuous natural filtration by $\{\mathcal{F}_t^Z\}_{t\geq 0}$. For the shot-noise environment, we consider the following four objects: A Poisson process $\{N_t^\rho\}_{t\geq 0}$ with constant intensity $\rho > 0$ and jump times $\{T_i^\rho\}_{i\in\mathbb{N}}$, a sequence $\{Y_i\}_{i\in\mathbb{N}}$ of positive i.i.d. random variables with distribution function F_Y, a non-negative function w, and a positive starting value λ. With these components we define the multiplicative shot-noise process by

$$\lambda_t := \lambda w(t) + \sum_{i=1}^{N_t^\rho} Y_i w(t - T_i^\rho).$$

Since we want to exploit the theory of PDMPs, it would be preferable if the process $\{\lambda_t\}_{t\geq 0}$ satisfies the Markov property. As shown by Schmidt (2017), this is equivalent to the existence of some $\beta > 0$ such that $w(t) = e^{-\beta t}$. Due to this, we define the Markovian shot-noise process in the following way.

Definition 2.1. Let $\{N_t^\rho\}_{t\geq 0}$ be a Poisson process with intensity $\rho > 0$ and jump times $\{T_i^\rho\}_{i\in\mathbb{N}}$, $\{Y_i\}_{i\in\mathbb{N}}$ i.i.d. copies of a positive random variable Y with distribution function F_Y and independent of the process $\{N_t^\rho\}_{t\geq 0}$, $\lambda > 0$ and $\beta > 0$ constant. Then, we define the Markovian shot-noise process by

$$\lambda_t = \lambda e^{-\beta t} + \sum_{i=1}^{N_t^\rho} Y_i e^{-\beta(t - T_i^\rho)}.$$

As shown in Dassios and Jang (2005) the Markovian shot-noise process is a piecewise-deterministic Markov process with generator

$$\mathcal{A}^\lambda f(\lambda) = -\beta\lambda \frac{\partial f(\lambda)}{\partial \lambda} + \rho \int_0^\infty \left(f(\lambda + y) - f(\lambda)\right) F_Y(\mathrm{d}y).$$

Further information about PDMPs can be found in Davis (1993) or Chapter 11 of Rolski et al. (1999). To fully specify our model we will now define the surplus process.

Definition 2.2. Let $\{\lambda_t\}_{t\geq 0}$ be a Markovian shot-noise process, $\{N_t\}_{t\geq 0}$ a Cox process with intensity $\{\lambda_t\}_{t\geq 0}$ and $\{U_i\}_{i\in\mathbb{N}}$ a sequence of i.i.d. copies of a positive random variable U with continuous distribution F_U, which are independent of $\{N_t\}_{t\geq 0}$ and $\{\lambda_t\}_{t\geq 0}$. For some initial capital u and constant premium rate $c > 0$ we define the surplus process by

$$X_t = u + ct - \sum_{i=1}^{N_t} U_i.$$

Now define $\mathcal{F}_t := \mathcal{F}_t^X \vee \mathcal{F}_t^\lambda$; hence, $\{\mathcal{F}_t\}_{t\geq 0}$ is the combined filtration of the Markovian shot-noise process and the surplus process. If not mentioned differently, we will from now on consider the filtered probability space $\left(\Omega, \mathcal{F}, \{\mathcal{F}_t\}_{t\geq 0}, \mathbb{P}_{(u,\lambda)}\right)$, where we define the measure $\mathbb{P}_{(u,\lambda)}$ as the measure $\mathbb{P}$ under the conditions that the initial capital of the surplus process is u and the starting intensity is λ. We will denote the expectation of a random variable Z under this measure by $\mathbb{E}_{(u,\lambda)}[Z]$ or $\mathbb{E}[Z]$ if Z is independent of the initial values.

The multivariate process $\{(X_t, \lambda_t, t)\}_{t\geq 0}$ is a càdlàg PDMP without active boundary and generator

$$\begin{aligned}\mathcal{A}f(x,\lambda,t) &= c\frac{\partial f(x,\lambda,t)}{\partial x} - \beta\lambda\frac{\partial f(x,\lambda,t)}{\partial \lambda} + \frac{\partial f(x,\lambda,t)}{\partial t} \\ &+ \lambda\int_0^\infty \left(f(x-u,\lambda,t) - f(x,\lambda,t)\right)\,F_U(\mathrm{d}u) \\ &+ \rho\int_0^\infty \left(f(x,\lambda+y,t) - f(x,\lambda,t)\right)\,F_Y(\mathrm{d}y).\end{aligned}$$

Its domain consists of all functions f which are absolutely continuous and satisfy the integrability condition

$$\mathbb{E}_{(u,\lambda)}\left[\sum_{i=1}^{\tilde{N}_t} |f(X_{T_i},\lambda_{T_i},T_i) - f(X_{T_i-},\lambda_{T_i-},T_i-)|\right] < \infty$$

for all $t \geq 0$, where $\{\tilde{N}_t\}_{t\geq 0}$ denotes the process counting the random jumps of the PDMP $\{(X_t, \lambda_t, t)\}_{t\geq 0}$. Similar to the Cramér-Lundberg model, we want to state a net profit condition, which is necessary to ensure that ruin does not occur with probability 1.

Lemma 2.1. *The surplus process satisfies*

$$\lim_{t\to\infty}\frac{\mathbb{E}_{(u,\lambda)}\left[X_t\right]}{t} = c - \frac{\rho}{\beta}\mathbb{E}\left[U\right]\mathbb{E}\left[Y\right].$$

Proof. The function $\bar{f}(x,\lambda,t) := x$ is in the domain of the generator. Consequently,

$$\mathbb{E}_{(u,\lambda)}\left[X_t\right] = u + \mathbb{E}_{(u,\lambda)}\left[\int_0^t \mathcal{A}\bar{f}(X_s,\lambda_s,s)\,\mathrm{d}s\right] = u + ct - \mathbb{E}_{(u,\lambda)}\left[\int_0^t \lambda_s\mathbb{E}\left[U\right]\,\mathrm{d}s\right].$$

The process $\{\lambda_t\}_{t\geq 0}$ is positive so we can use Tonelli's theorem and interchange expectation and integration, which leads to

$$\mathbb{E}_{(u,\lambda)}\left[X_t\right] = u + ct - \mathbb{E}\left[U\right]\int_0^t \mathbb{E}_{(u,\lambda)}\left[\lambda_s\right]\,\mathrm{d}s. \tag{2.1}$$

Now we use the same procedure to obtain an equation for $\mathbb{E}_{(u,\lambda)}\left[\lambda_s\right]$. Defining the function $\tilde{f}(x,\lambda,t) := \lambda$ we get

$$\mathbb{E}_{(u,\lambda)}\left[\lambda_s\right] = \lambda - \beta\int_0^s \mathbb{E}_{(u,\lambda)}\left[\lambda_u\right]\,\mathrm{d}u + \rho s\mathbb{E}\left[Y\right].$$

Differentiating both sides with respect to s gives us that $\mathbb{E}_{(u,\lambda)}[\lambda_s]$ is the solution to the differential equation $g'(s) = -\beta g(s) + \rho\,\mathbb{E}[Y]$, with initial value $g(0) = \lambda$. The solution of the ODE is

$$\mathbb{E}_{(u,\lambda)}[\lambda_s] = \lambda e^{-\beta s} + \frac{\rho}{\beta}\mathbb{E}[Y](1 - e^{-\beta s}). \tag{2.2}$$

Using (2.2) in (2.1), leads to

$$\begin{aligned}\mathbb{E}_{(u,\lambda)}[X_t] = u + ct - \mathbb{E}[U]\frac{\rho}{\beta}\mathbb{E}[Y]\,t \\ + \mathbb{E}[U]\left(\frac{\lambda}{\beta} - \frac{\rho}{\beta^2}\mathbb{E}[Y]\right)\left(1 - e^{-\beta t}\right).\end{aligned}$$

Now, let us divide by t and let it tend to infinity to obtain

$$\lim_{t\to\infty}\frac{\mathbb{E}_{(u,\lambda)}[X_t]}{t} = c - \frac{\rho}{\beta}\mathbb{E}[U]\,\mathbb{E}[Y].$$

□

Motivated by this result we assume the following:

Assumption 2.1. *From now on we assume that the net profit condition*

$$c > \frac{\rho}{\beta}\mathbb{E}[U]\,\mathbb{E}[Y],$$

is satisfied.

2.3 Martingales and Change of Measure

To obtain the asymptotic behaviour of the ruin probability in this model, we want to exploit the following result derived in Schmidli (1997).

Theorem 2.2. *Theorem 2 of Schmidli (1997)*
Assume that $z(u)$ is directly Riemann integrable, that $0 \le p(u,x) \le 1$ is continuous in u and that $\int_0^u p(u,y)\,B(\mathrm{d}y)$ is directly Riemann integrable. Denote by $Z(u)$ the solution to

$$Z(u) = \int_0^u Z(u-y)(1 - p(u,y))\,B(\mathrm{d}y) + z(u),$$

which is bounded on bounded intervals. Then, the limit

$$\lim_{u\to\infty} Z(u)$$

exists and is finite provided $B(u)$ is not arithmetic. If $B(u)$ is arithmetic with span γ, then

$$\lim_{n\to\infty} Z(x + n\gamma)$$

exists and is finite for all x fixed.

Unfortunately, we cannot apply this theorem directly to our model because of two problems. The first issue is, that the ruin probability depends on the initial intensity level λ. To bypass this, we have to choose appropriate renewal times such that $\{\lambda_t\}_{t\geq 0}$ has always the same level, which we will do in Section 2.4. The second problem is, that suitable choices of B are defective under the original measure $\mathbb{P}_{(u,\lambda)}$. This is a common issue and can be solved through change of measure techniques.

To do so we have to find martingales of the form $M_t = h(X_t, \lambda_t, t)$. Our approach is a function of the form

$$h(x, \lambda, t) := K \exp(-\theta(r)t - \alpha(r)\lambda - rx).$$

To motivate the explicit choice of our parameters let us assume that h is in the domain of the generator and apply $\mathcal{A}$ to h. This gives us

$$\begin{aligned}\mathcal{A}h(x,\lambda,t) &= -\theta h(x,\lambda,t) - crh(x,\lambda,t) + \beta\lambda\alpha h(x,\lambda,t)\\ &\quad + \lambda h(x,\lambda,t)\int_0^\infty (e^{ru}-1)\ F_U(\mathrm{d}u) + \rho h(x,\lambda,t)\int_0^\infty \left(e^{-\alpha y}-1\right)\ F_Y(\mathrm{d}y) \stackrel{!}{=} 0.\end{aligned}$$

Since h is strictly positive, we can reformulate the equation to

$$\beta\lambda\alpha - cr - \theta + \lambda(M_U(r)-1) + \rho(M_Y(-\alpha)-1) = 0.$$

Here $M_U(s)$ and $M_Y(s)$ denote the moment-generating functions of the random variables U and Y, which we assume to be finite. The equation above has to hold for any $\lambda > 0$; hence, this is equivalent to

$$\begin{aligned}&\beta\alpha + M_U(r) - 1 = 0,\\ &-cr - \theta + \rho(M_Y(-\alpha)-1) = 0.\end{aligned}$$

Solving the above equations for some fixed r we get the unique solutions

$$\alpha(r) = \frac{1-M_U(r)}{\beta},$$

and

$$\theta(r) = -cr + \rho\left(M_Y\left(\frac{M_U(r)-1}{\beta}\right) - 1\right).$$

Now we still have to show that for this explicit choice of the parameters, the process $\{h(X_t,\lambda_t,t)\}_{t\geq 0}$ is indeed a martingale.

Lemma 2.3. *Let r be constant such that $M_U(r)$ is finite and define $\alpha(r) := \frac{1-M_U(r)}{\beta}$. Assume further that $M_Y(-\alpha(r))$ is finite. If $\theta(r) := -cr + \rho\left(M_Y(-\alpha(r)) - 1\right)$ and $K = \exp(ru + \alpha(r)\lambda)$, then $h(X_t,\lambda_t,t)$ is integrable and has expectation 1 for all $t \geq 0$.*

Proof. The expectation can be rewritten as

$$\mathbb{E}_{(u,\lambda)}\left[K\exp\left(-rX_t-\alpha(r)\lambda_t-\theta(r)t\right)\right]=\exp(-rct-\theta(r)t+\alpha(r)\lambda)$$
$$\mathbb{E}_{(u,\lambda)}\left[\exp\left(-r\sum_{i=1}^{N_t}U_i-\alpha(r)\lambda_t\right)\right].$$

Conditioned on $\mathcal{F}_t^\lambda$, the counting process $\{N_t\}_{t\geq 0}$ is an inhomogeneous Poisson process and as shown in Albrecher and Asmussen (2006) its integrated compensator has the form

$$\Lambda_t=\int_0^t\lambda_s\,\mathrm{d}s=\frac{1}{\beta}\left(\lambda+\sum_{j=1}^{N_t^\rho}Y_j-\lambda_t\right).$$

Using this we get

$$\begin{aligned}&\exp\left(-rct-\theta(r)t+\alpha(r)\lambda\right)\mathbb{E}_{(u,\lambda)}\left[\exp\left(-r\sum_{i=1}^{N_t}U_i-\alpha(r)\lambda_t\right)\right]=\\&\exp\left(-rct-\theta(r)t+\alpha(r)\lambda\right)\mathbb{E}_{(u,\lambda)}\left[\exp\Big(\left(M_U(r)-1\right)\Lambda_t-\alpha(r)\lambda_t\Big)\right]=\\&\exp\left(-rct-\theta(r)t\right)\mathbb{E}\left[\exp\left(-\alpha(r)\sum_{j=1}^{N_t^\rho}Y_j\right)\right].\end{aligned}$$

The process $\{\sum_{j=1}^{N_t^\rho}Y_j\}_{t\geq 0}$ is a compound Poisson process, whose moment-generating function is $\exp(\rho t(M_Y(-\alpha(r))-1)$. By this and the definition of $\theta(r)$ we get that $h(X_t,\lambda_t,t)$ has expectation 1. ☐

This result leads immediately to the following theorem.

Theorem 2.4. *Under the assumptions of Lemma 2.3, the process $M_t^r:=h(X_t,\lambda_t,t)$ is a martingale with expectation 1.*

Proof. By Lemma 2.3, the process is integrable and has constant expectation 1. Consequently, we just have to show that for all $t>s$

$$\mathbb{E}_{(u,\lambda)}\left[h(X_t,\lambda_t,t)\mid\mathcal{F}_s\right]=h(X_s,\lambda_s,s).$$

The function h is strictly positive for all values x,λ and t, hence we can simply expand the conditional expectation above by $\frac{h(X_s,\lambda_s,s)}{h(X_s,\lambda_s,s)}$. Consequently, we get

$$\begin{aligned}\mathbb{E}_{(u,\lambda)}\left[h(X_t,\lambda_t,t)\mid\mathcal{F}_s\right]&=h(X_s,\lambda_s,s)\mathbb{E}_{(u,\lambda)}\left[\frac{h(X_t,\lambda_t,t)}{h(X_s,\lambda_s,s)}\,\middle|\,\mathcal{F}_s\right]\\&=h(X_s,\lambda_s,s)\mathbb{E}_{(u,\lambda)}\left[\exp\left(-\theta(r)(t-s)-r(X_t-X_s)-\alpha(r)(\lambda_t-\lambda_s)\right)\mid\mathcal{F}_s\right]\\&=h(X_s,\lambda_s,s)\mathbb{E}_{(X_s,\lambda_s)}\left[h(X_{t-s},\lambda_{t-s},t-s)\right]=h(X_s,\lambda_s,s).\end{aligned}$$

□

Using these martingales, we can define a family of measures $\mathbb{Q}^{(r)}$ such that

$$\left.\frac{\mathrm{d}\mathbb{Q}^{(\mathrm{r})}}{\mathrm{d}\mathbb{P}_{(u,\lambda)}}\right|_{\mathcal{F}_t} = M_t^r.$$

The exponential form of the change of measure allows us to exploit the results shown in Palmowski and Rolski (2002) to derive the behaviour of the combined process under the new measures $\mathbb{Q}^{(r)}$.

Lemma 2.5. *Let r be such that $\{M_t^r\}_{t\geq 0}$ is well defined. Then, under the measure $\mathbb{Q}^{(r)}$, the process $\{(X_t, \lambda_t, t)\}_{t\geq 0}$ is again a PDMP with generator*

$$\begin{aligned}\mathcal{A}^{(r)} f(x,\lambda,t) =& c\frac{\partial f(x,\lambda,t)}{\partial x} - \beta\lambda\frac{\partial f(x,\lambda,t)}{\partial \lambda} + \frac{\partial f(x,\lambda,t)}{\partial t} \\ &+ \lambda\int_0^\infty (f(x-u,\lambda,t) - f(x,\lambda,t))e^{ru}\, F_U(\mathrm{d}u) \\ &+ \rho\int_0^\infty (f(x,\lambda+y,t) - f(x,\lambda,t))e^{-\alpha(r)y}\, F_Y(\mathrm{d}y),\end{aligned}$$

So far, we have found a new family of measures but we have to identify a measure that fits our needs. Motivated by the definition of the adjustment coefficient in the classical model we consider the function $\theta(r) = -cr + \rho(M_Y(-\alpha(r)) - 1)$.

Lemma 2.6. *The function $\theta(r)$ is convex on $\{r \mid M_U(r) < \infty,\, M_Y(-\alpha(r)) < \infty\}$ and satisfies $\theta(0) = 0$.*

Proof. To show convexity we use the fact, that moment-generating functions are log-convex, and therefore convex. Moreover, they are twice differentiable. Consequently, θ is twice differentiable too and its derivatives are

$$\theta'(r) = -c + \frac{\rho}{\beta}M_Y'\left(\frac{M_U(r)-1}{\beta}\right)M_U'(r),$$

$$\theta''(r) = \frac{\rho}{\beta^2}M_Y''\left(\frac{M_U(r)-1}{\beta}\right)M_U'(r)^2 + \frac{\rho}{\beta}M_Y'\left(\frac{M_U(r)-1}{\beta}\right)M_U''(r).$$

By convexity of the moment-generating functions, we know that their second derivatives are non-negative. To ensure that θ is convex, we have to check if the first derivative of the MGF of Y is non-negative too. Equivalently we show that the MGF of Y is monotone increasing. Let now $r > s$ then $\mathbb{E}\left[e^{rY}\right] = \mathbb{E}\left[e^{sY}e^{(r-s)Y}\right]$. The random variable Y is almost surely positive and $r - s$ is positive too. Hence, $e^{(r-s)Y} > 1$ almost surely. This gives us

$$M_Y(r) = \mathbb{E}\left[e^{sY}e^{(r-s)Y}\right] > \mathbb{E}\left[e^{sY}\right] = M_Y(s).$$

Consequently, the first derivative of $M_Y(r)$ is non-negative. Therefore, θ is convex and since $M_U(0) = M_Y(0) = 1$ we get that $\theta(0) = 0$. $\square$

Lemma 2.7. *Let r be such that the measure $\mathbb{Q}^{(r)}$ is well defined and assume there is some $\varepsilon > 0$ such that $M_U(r+\varepsilon)$ and $M_Y(-\alpha(r)+\varepsilon)$ are finite. Then,*

$$\lim_{t\to\infty} \frac{\mathbb{E}^{\mathbb{Q}^{(r)}}[X_t]}{t} = -\theta'(r).$$

Proof. To show this property, we can use the ideas of the proof of Lemma 2.1. The main difference is, that we apply the generator $\mathcal{A}^{(r)}$. Again we obtain

$$\mathbb{E}^{\mathbb{Q}^{(r)}}[X_t] = u + ct - M_U(r)\mathbb{E}^{\mathbb{Q}^{(r)}}[U] \int_0^t \mathbb{E}^{\mathbb{Q}^{(r)}}[\lambda_s] \, \mathrm{d}s.$$

The expectation of λ_t under $\mathbb{Q}^{(r)}$ satisfies

$$\mathbb{E}^{\mathbb{Q}^{(r)}}[\lambda_t] = \frac{\rho}{\beta} M_Y(-\alpha(r))\mathbb{E}^{\mathbb{Q}^{(r)}}[Y](1-e^{-\beta t}) + e^{-\beta t}\lambda.$$

The expectations $\mathbb{E}^{\mathbb{Q}^{(r)}}[U]$ and $\mathbb{E}^{\mathbb{Q}^{(r)}}[Y]$ can easily be obtained from

$$M_U^{\mathbb{Q}^{(r)}}(s) = \frac{M_U(s+r)}{M_U(r)}$$

and

$$M_Y^{\mathbb{Q}^{(r)}}(s) = \frac{M_Y(s-\alpha(r))}{M_Y(-\alpha(r))}.$$

Consequently, $\mathbb{E}^{\mathbb{Q}^{(r)}}[U] = \frac{M_U'(r)}{M_U(r)}$ and $\mathbb{E}^{\mathbb{Q}^{(r)}}[Y] = \frac{M_Y'(-\alpha(r))}{M_Y(-\alpha(r))}$.
Combining these results we get

$$\lim_{t\to\infty} \frac{\mathbb{E}^{\mathbb{Q}^{(r)}}[X_t]}{t} = c - \frac{\rho}{\beta} M_Y'(-\alpha(r))M_U'(r) = -\theta'(r).$$

$\square$

Assumption 2.2. *From now on we assume that there exists a positive solution R to the equation $\theta(R) = 0$, that $\mathbb{Q}^{(R)}$ is well defined and that for some $\varepsilon > 0$ both $M_U(R+\varepsilon)$ and $M_Y(\varepsilon - \alpha(R))$ are finite.*

This assumption ensures that the measure $\mathbb{Q}^{(R)}$ is well defined and that we can express the expectation of Y and U in terms of their original moment-generating functions. One example where this is satisfied is the following.

Example 2.1. Let μ and κ be positive constants. If $Y \sim Exp(\mu)$ and $U \sim Exp(\kappa)$, the net profit condition simplifies to $c > \frac{\rho}{\beta\kappa\mu}$. The moment-generating functions are given by $M_U(r) = \frac{\kappa}{\kappa - r}$ and $M_Y(-\alpha(r)) = \frac{\mu}{\mu+\alpha(r)}$, where $r < \kappa$ and $-\alpha(r) < \mu$. If we fix some

$r < \frac{\mu\beta\kappa}{1+\beta\mu}$ we can determine the functions $\alpha(r) = -\frac{r}{\beta(\kappa-r)}$ and

$$\theta(r) = -cr + \rho\left(\frac{r}{\mu\beta(\kappa-r)-r}\right).$$

Solving the equation $\theta(r) = 0$ gives us the solutions $r_1 = 0$ and

$$R := r_2 = \frac{\mu\beta\kappa c - \rho}{(1+\mu\beta)c},$$

which is positive by the net profit condition. Now we want to show that there is some $\varepsilon > 0$ such that $R + \varepsilon < \frac{\mu\beta}{1+\mu\beta}\kappa$ and $\varepsilon - \alpha(R) < \mu$. The first inequality is equivalent to

$$\varepsilon < \frac{\rho}{(1+\mu\beta)c},$$

which is a strictly positive upper bound. The second condition can be rewritten to

$$\varepsilon < \frac{\mu\rho\beta + \rho}{\beta\kappa c + \rho\beta},$$

which is positive too. Consequently, Assumption 2.2 is satisfied.

Lemma 2.8. *For every $u \geq 0$ and $\lambda > 0$ we have that $\mathbb{Q}^{(R)}[\tau_u < \infty] = 1$.*

Proof. We already know that $\lim_{t\to\infty} \frac{\mathbb{E}^{\mathbb{Q}^{(R)}}[X_t]}{t} = -\theta'(R)$ holds true. If we can show that $\theta'(R) > 0$ then ruin occurs almost surely under the new measure. The function θ is convex and satisfies $\theta(0) = \theta(R) = 0$. Further we have that

$$\theta'(0) = -c + \frac{\rho}{\beta}\mathbb{E}[Y]\,\mathbb{E}[U],$$

which is smaller than 0 by the net profit condition. Therefore, there exists $0 < r < R$ such that $\theta(r) < 0$. Since $\theta(R) > \theta(r)$, it follows by the mean-value theorem that there is a $\tilde{r} \in (r, R)$ such that

$$\theta'(\tilde{r}) = \frac{\theta(R) - \theta(r)}{R - r} > 0.$$

By convexity we know that θ' is a monotone increasing function and $\theta'(R) \geq \theta'(\tilde{r}) > 0$. □

Similar to the classical model and the Björk-Grandell model which is considered in Schmidli (1997), we have found a new measure under which ruin occurs almost surely. We can use this to get an upper bound for the ruin probability.

Theorem 2.9. *Under our assumptions*

$$\psi(u, \lambda) \leq e^{-\alpha(R)\lambda} e^{-Ru}.$$

Proof. The ruin probability can be rewritten as

$$\psi(u,\lambda) = \mathbb{E}_{(u,\lambda)}\left[I_{\{\tau_u<\infty\}}\right] = \mathbb{E}^{\mathbb{Q}(R)}\left[I_{\{\tau_u<\infty\}}\left(M^R_{\tau_u}\right)^{-1}\right]$$
$$= \exp(-Ru - \alpha(R)\lambda)\mathbb{E}^{\mathbb{Q}(R)}\left[\exp\left(RX_{\tau_u} + \alpha(R)\lambda_{\tau_u}\right)\right].$$

By definition of τ_u, the value X_{τ_u} is negative and since $R > 0$ we have that $M_U(R) > 1$. Consequently, $\alpha(R) < 0$. By this, we get that $\exp(RX_{\tau_u} + \alpha(R)\lambda_{\tau_u}) \le 1$ and

$$\psi(u,\lambda) \le \exp(-Ru - \alpha(R)\lambda).$$

□

2.4 The Renewal Equation

We now want to use Theorem 2.2 to get information about the asymptotic behaviour of the ruin probability $\psi(u,\lambda)$. Because of the dependence on λ, we have to choose the renewal times $\{S_+(i)\}_{i\in\mathbb{N}}$ such that $\lambda_{S_+(i)} = \lambda$. To exploit the renewal equation, we have to ensure that there are infinitely many renewal times and that they are almost surely finite. For this, we will use the ideas from Orsingher and Battaglia (1982) to get an intensity for the number of upcrossings of the process $\{\lambda_t\}_{t\ge0}$ through some level l.

Lemma 2.10. *Let $\{\lambda_t\}_{t\ge0}$ be the Markovian shot-noise process and $l > 0$ arbitrary. The process counting all upcrossings of $\{\lambda_t\}_{t\ge0}$ through l has intensity*

$$\nu_l^+(t) = \rho\int_0^l (1 - F_Y(l - z))\, F_\lambda(\mathrm{d}z, t),$$

where $F(z,t) = \mathbb{P}_{(u,\lambda)}[\lambda_t \le z]$ is the cumulative distribution function of λ_t.

Proof. Consider for some small $\beta t > 0$ the probability $\mathbb{P}_{(u,\lambda)}[\lambda_t \le l, \lambda_{t+\beta t} > l]$. The jumps of $\{\lambda_t\}_{t\ge0}$ are governed by a Poisson process with rate ρ; hence,

$$\mathbb{P}_{(u,\lambda)}[\lambda_t \le l, \lambda_{t+\beta t} > l] = \mathbb{P}_{(u,\lambda)}\left[N^\rho_{t+\beta t} - N^\rho_t = 1, \lambda_t \le l, \lambda_{t+\beta t} > l\right] + o(\beta t) =$$
$$\mathbb{P}_{(u,\lambda)}\left[N^\rho_{t+\beta t} - N^\rho_t = 1, \lambda_t \le l, \lambda_t e^{-\beta\beta t} + Ye^{-\beta(t+\beta t-T)} > l\right] + o(\beta t) =$$
$$\mathbb{P}_{(u,\lambda)}\left[N^\rho_{t+\beta t} - N^\rho_t = 1, \lambda_t \le l, \lambda_t + Ye^{-\beta(t-T)} > le^{\beta\beta t}\right] + o(\beta t).$$

Here T denotes the jump time occurring between t and $t+\beta t$ and Y is the corresponding shock. The random time $T - t$ can be represented as $\Theta\beta t$, where Θ is a random variable which takes values in the interval $(0,1)$. Consequently, we have that $Ye^{\beta\Theta\beta t} \in (Y, Ye^{\beta\beta t})$.

Using this we can bound the above probability by

$$\mathbb{P}_{(u,\lambda)}\left[N^\rho_{t+\beta t} - N^\rho_t = 1, \lambda_t \le l, \lambda_t + Ye^{\beta\beta t} > le^{\beta\beta t}\right] + o(\beta t) \ge$$
$$\mathbb{P}_{(u,\lambda)}\left[N^\rho_{t+\beta t} - N^\rho_t = 1, \lambda_t \le l, \lambda_t + Ye^{-\beta\Theta\beta t} > le^{\beta\beta t}\right] + o(\beta t) \ge$$
$$\mathbb{P}_{(u,\lambda)}\left[N^\rho_{t+\beta t} - N^\rho_t = 1, \lambda_t \le l, \lambda_t + Y > le^{\beta\beta t}\right] + o(\beta t).$$

Let us focus on the upper bound. The term $N^\rho_{t+\beta t} - N^\rho_t$ is independent of λ_t and Y so we can rewrite

$$\mathbb{P}_{(u,\lambda)}\left[N^\rho_{t+\beta t} - N^\rho_t = 1, \lambda_t \le l, \lambda_t + Ye^{\beta\beta t} > le^{\beta\beta t}\right] + o(\beta t) =$$
$$\rho\beta t\mathbb{P}_{(u,\lambda)}\left[\lambda_t \le l, \lambda_t + Ye^{\beta\beta t} > le^{\beta\beta t}\right] + o(\beta t) =$$
$$\rho\beta t\, \mathbb{E}_{(u,\lambda)}\left[\mathbb{E}_{(u,\lambda)}\left[I_{\{\lambda_t \le l\}} I_{\left\{Y > l - \lambda_t e^{-\beta\beta t}\right\}} \mid \lambda_t\right]\right] + o(\beta t) =$$
$$\rho\beta t \int_0^l (1 - F_Y(l - ze^{-\beta\beta t}))\, F_\lambda(\mathrm{d}z, t) + o(\beta t).$$

Now, let us divide by βt and consider the limit of $\beta t \to 0$. Since $F_Y(l - ze^{-\beta\beta t})$ decreases as βt becomes smaller, we get by the right continuity of cumulative distribution function that this tends to

$$\rho \int_0^l (1 - F_Y(l - z))\, F_\lambda(\mathrm{d}z, t).$$

Using the same arguments we can show that the lower bound divided by βt converges to the same value. Hence, the term $\frac{1}{\beta t}\mathbb{P}_{(u,\lambda)}\left[\lambda_t \le l, \lambda_{t+\beta t} > l\right]$ converges too. □

Assumption 2.3. *From now on we assume that*

$$\int_0^\infty \int_0^\lambda (1 - F_Y^{\mathbb{Q}^{(R)}}(\lambda - z))\, F_\lambda^{\mathbb{Q}^{(R)}}(\mathrm{d}z, t)\,\mathrm{d}t = \infty,$$

where $F_\lambda^{\mathbb{Q}^{(R)}}(z, t) := \mathbb{Q}^{(R)}\left[\lambda_t \le z\right]$ *and* $F_Y^{\mathbb{Q}^{(R)}}(x) = \mathbb{Q}^{(R)}\left[Y \le x\right]$.

This assumption guarantees that there are infinitely many upcrossings of the process through λ under the measure $\mathbb{Q}^{(R)}$; hence, the intensity is Harris recurrent. The structure of our Markovian shot-noise process gives us, that upcrossings can only happen through shock events and downcrossings are due to the continuous drift. Consequently, there have to be infinitely many continuous downcrossings and recurrence times $\{S(i)\}_{i\in\mathbb{N}}$, such that $\lambda_{S(i)} = \lambda$.
One example which satisfies Assumption 2.3 is the following.

Example 2.2. Consider the same configuration as in Example 2.1. Under the new measure $\mathbb{Q}^{(R)}$, the shocks are again exponentially distributed with parameter $\mu + \alpha(R)$

and the new intensity of $\{\lambda_t\}_{t\geq 0}$ is

$$\tilde{\rho} = \rho M_Y(-\alpha(R)) = \frac{\mu\beta\kappa c + \mu\beta\rho}{\mu\beta + 1}.$$

Assume that $\frac{\tilde{\rho}}{\beta} = n \in \mathbb{N}$. Like in Orsingher and Battaglia (1982) we can determine the distribution of $Y(t)$ using its characteristic function

$$K_t(s) = \mathbb{E}^{\mathbb{Q}^{(R)}}\left[\exp(is\lambda(t))\right] = \left(e^{-\beta t} + (1 - e^{\beta t})\frac{\mu + \alpha(R)}{(\mu + \alpha(R)) - is}\right)^n.$$

This is the characteristic function of the random variable $\eta = \sum_{i=1}^{B_t} Y_i$, where

$$B_t \sim B\left(n, 1 - e^{-\beta t}\right).$$

Consequently, λ_t admits a density of the form

$$f(z,t) = \sum_{j=1}^{n} \binom{n}{j} e^{-\beta t(n-j)}(1 - e^{-\beta t})^j (\mu + \alpha(R))^j e^{-(\mu+\alpha(R))z} \frac{z^{j-1}}{(j-1)!}.$$

Using this, the intensity of the upcrossings is given by

$$\nu_\lambda^+(t) = \rho \sum_{j=0}^{n} \binom{n}{j} e^{-\beta t(n-j)}(1 - e^{-\beta t})^j \frac{(\mu + \alpha(R))^j \lambda^j}{j!} e^{-(\mu+\alpha(R))\lambda}.$$

Since $\frac{(\mu+\alpha(R))^j \lambda^j}{j!}$ has a positive lower bound $\tilde{c}$, we get that

$$\int_0^\infty \nu_\lambda^+(t)\,\mathrm{d}t \geq \int_0^\infty \rho\tilde{c}e^{-(\mu+\alpha(R))\lambda}\,\mathrm{d}t = \infty.$$

Using this, we can even show that there are infinitely many recurrence times if $\frac{\tilde{\rho}}{\beta}$ is any real number greater than 1. For this, we consider two auxiliary shot-noise processes

$$\underline{\lambda}_t = e^{-\underline{\beta}t}\lambda + \sum_{i=1}^{N_t^\rho} e^{-\underline{\beta}(t-T_i^\rho)} Y_i,$$

and

$$\overline{\lambda}_t = e^{-\overline{\beta}t}\lambda + \sum_{i=1}^{N_t^\rho} e^{-\overline{\beta}(t-T_i^\rho)} Y_i,$$

where $\overline{\beta}$ and $\underline{\beta}$ are chosen such that

$$\frac{\tilde{\rho}}{\overline{\beta}} = N_1 > \frac{\tilde{\rho}}{\beta} > \frac{\tilde{\rho}}{\underline{\beta}} = N_2,$$

with $N_1, N_2 \in \mathbb{N}$. By construction we have that, $\underline{\lambda}_t \leq \lambda_t \leq \overline{\lambda}_t$ and both auxiliary processes cross λ infinitely often. As a consequence $\{\lambda_t\}_{t\geq 0}$ crosses λ infinitely often too.

If Assumption 2.3 holds we have that under the measure $\mathbb{Q}^{(R)}$, the surplus process tends to $-\infty$ and $\{\lambda_t\}_{t\geq 0}$ returns to λ infinitely often. Hence, we can define a sequence of renewal times $\{S_+(i)\}_{i\in\mathbb{N}_0}$ via $S_+(0) = 0$ and $S_+(i) = \min\left\{S(i) > S_+(i-1) \;\middle|\; X_{S(i)} < X_{S_+(i-1)}\right\}$ which satisfies $\mathbb{Q}^{(R)}\left[S_+(i) < \infty\right] = 1$ for all i. We will use these renewal times similar to the ladder epochs in the classical ruin model.
Define

$$B(x) = \mathbb{P}_{(u,\lambda)}\left[S_+(1) < \infty,\, u - X_{S_+(1)} \leq x\right],$$

and

$$p(u,x) = \mathbb{P}_{(u,\lambda)}\left[\tau_u \leq S_+(1) \,|\, S_+(1) < \infty,\, X_{S_+(1)} = u - x\right].$$

Then, the ruin probability satisfies:

$$\psi(u,\lambda) = \int_0^u \psi(u-x,\lambda)(1-p(u,x))\, B(\mathrm{d}x) + \mathbb{P}_{(u,\lambda)}\left[\tau_u \leq S_+(1),\, \tau_u < \infty\right].$$

This may look like a renewal equation but the distribution B is defective. We solve this problem by multiplying both sides with e^{Ru}, which is equivalent to a measure change from $\mathbb{P}$ to $\mathbb{Q}^{(R)}$, and obtain:

$$\psi(u,\lambda)e^{Ru} = \int_0^u \psi(u-x,\lambda)e^{R(u-x)}(1-p(u,x))e^{Rx}\, B(\mathrm{d}x) + \mathbb{P}_{(u,\lambda)}\left[\tau_u \leq S_+(1),\, \tau_u < \infty\right] e^{Ru}. \quad (2.3)$$

Lemma 2.11. *The distribution $\tilde{B}$ defined by $\tilde{B}(\mathrm{d}x) = e^{Rx}B(\mathrm{d}x)$ is non-defective.*

Proof. Using the definition of $\tilde{B}$ we get

$$\int_{\mathbb{R}} \tilde{B}(\mathrm{d}x) = \int_{\mathbb{R}} e^{Rx}B(\mathrm{d}x) = \mathbb{E}_{(u,\lambda)}\left[e^{R(u-X_{S_+(1)})}I_{\{S_+(1)<\infty\}}\right].$$

Now focus on our martingale $\{M_t^R\}_{t\geq 0}$ at time $S_+(1)$ and observe that

$$M_{S_+(1)}^R = \exp\left(\alpha(R)\lambda + Ru - \alpha(R)\lambda_{S_+(1)} - RX_{S_+(1)}\right) = \exp\left(R(u - X_{S_+(1)})\right)$$

Using this leads to

$$\int_{\mathbb{R}} \tilde{B}(\mathrm{d}x) = \mathbb{E}^{\mathbb{Q}^{(R)}}\left[I_{\{S_+(1)<\infty\}}\right] = \mathbb{Q}^{(R)}\left[S_+(1) < \infty\right] = 1.$$

Consequently, $\tilde{B}$ is not defective. □

Even though we have found a renewal equation, we still have to show that all functions appearing in (2.3) satisfy the assumptions of Theorem 2.2.

Assumption 2.4. *From now on we assume that there exists an* $\varepsilon > 0$ *such that for* $r := (1+\varepsilon)R$ *the measure* $\mathbb{Q}^{(r)}$ *is well defined and*

$$\mathbb{E}_{(u,\lambda)}\left[e^{-r(X_{S_+(1)}-u)}I_{\{S_+(1)<\infty\}}\right] < \infty.$$

Since $S_+(1)$ depends on $\{X_t\}_{t\geq 0}$ and $\{\lambda_t\}_{t\geq 0}$, this assumption may be hard to check. Alternatively, we can use the following lemma, which allows us to focus on the first recurrence time $S(1)$.

Lemma 2.12. *Let* $\varepsilon > 0$ *such that for* $r := (1+\varepsilon)R$ *the measure* $\mathbb{Q}^{(r)}$ *is well defined. Then,*

$$\mathbb{E}_{(u,\lambda)}\left[\exp(-r(X_{S_+(1)}-u))I_{\{S_+(1)<\infty\}}\right] < \infty,$$

if and only if

$$\mathbb{E}_{(u,\lambda)}\left[\exp\left(-r(X_{S(1)}-u)\right)I_{\{S(1)<\infty\}}\right] < \infty.$$

Proof. At first assume that

$$\mathbb{E}_{(u,\lambda)}\left[\exp\left(-r(X_{S_+(1)}-u)\right)I_{\{S_+(1)<\infty\}}\right] < \infty,$$

holds. By definition $S_+(1) \geq S(1)$ and $\theta(r) > 0$. Consequently,

$$\begin{aligned}\mathbb{E}_{(u,\lambda)}\left[\exp\left(-r(X_{S(1)}-u)\right)I_{\{S(1)<\infty\}}\right] &= \mathbb{E}^{\mathbb{Q}^{(r)}}\left[\exp\left(\theta(r)S(1)\right)\right] \leq \mathbb{E}^{\mathbb{Q}^{(r)}}\left[\exp\left(\theta(r)S_+(1)\right)\right] \\ &= \mathbb{E}_{(u,\lambda)}\left[\exp\left(-r(X_{S_+(1)}-u)\right)I_{\{S_+(1)<\infty\}}\right] < \infty.\end{aligned}$$

Let us now assume that

$$\mathbb{E}_{(u,\lambda)}\left[\exp\left(-r(X_{S(1)}-u)\right)I_{\{S(1)<\infty\}}\right] =: C < \infty,$$

holds true. Then,

$$\begin{aligned}\mathbb{E}_{(u,\lambda)}\left[\exp\left(-r(X_{S_+(1)}-u)\right)I_{\{S_+(1)<\infty\}}\right] &= \mathbb{E}^{\mathbb{Q}^{(R)}}\left[\exp\left(-\varepsilon R(X_{S_+(1)}-u)\right)\right] \\ &= \sum_{i=1}^{\infty}\mathbb{E}^{\mathbb{Q}^{(R)}}\left[\exp\left(-\varepsilon R(X_{S(i)}-u)\right)I_{\{S_+(1)=S(i)\}}\right].\end{aligned}$$

The indicator can rewritten as

$$I_{\{S_+(1)=S(i)\}} = I_{\{S_+(1)>S(i-1)\}}I_{\{X_{S(i)}<u\}} = \prod_{j=1}^{i-1}I_{\{X_{S(j)}\geq u\}}I_{\{X_{S(i)}<u\}}.$$

With $S(0)=0$ we define the i.i.d. sequence of random variables

$$\{\xi_j\}_{j\geq 1} := \{X_{S(j)} - X_{S(j-1)}\}_{j\geq 1}.$$

Consequently, $X_{S(i-1)} - u = \sum_{j=1}^{i-1} \xi_j$ holds true for all i. Using this, we get

$$\mathbb{E}^{\mathbb{Q}^{(R)}} \left[\exp\left(-\varepsilon R(X_{S(i)} - u)\right) I_{\{S_+(1)=S(i)\}}\right] \leq \mathbb{E}^{\mathbb{Q}^{(R)}} \left[\exp\left(-\varepsilon R \sum_{j=1}^{i-1} \xi_j\right) I_{\left\{\sum_{j=1}^{i-1} \xi_j > 0\right\}} \right.$$
$$\left. \mathbb{E}^{\mathbb{Q}^{(R)}} \left[\exp\left(-\varepsilon R \xi_i\right) I_{\left\{X_{S(i)} < u\right\}} \middle| \sum_{j=1}^{i-1} \xi_j \right]\right].$$

Let us focus on the conditional expectation. The indicator is less or equal to 1 and ξ_i is independent of the condition. Hence,

$$\mathbb{E}^{\mathbb{Q}^{(R)}} \left[\exp\left(-\varepsilon R \xi_i\right) I_{\left\{X_{S(i)} < u\right\}} \middle| \sum_{j=1}^{i-1} \xi_j \right] \leq \mathbb{E}^{\mathbb{Q}^{(R)}} \left[\exp\left(-\varepsilon R \xi_i\right)\right] = C < \infty.$$

By this, we get that

$$\mathbb{E}^{\mathbb{Q}^{(R)}} \left[\exp\left(-\varepsilon R(X_{S(i)} - u)\right) I_{\{S_+(1)=S(i)\}}\right] \leq C\, \mathbb{E}^{\mathbb{Q}^{(R)}} \left[\exp\left(-\varepsilon R \sum_{j=1}^{i-1} \xi_j\right) I_{\left\{\sum_{j=1}^{i-1} \xi_j > 0\right\}}\right].$$

Now we want to bound the remaining expectation. For this we observe that for all $\tilde{\varepsilon} > 0$

$$\mathbb{E}^{\mathbb{Q}^{(R)}} \left[\exp\left(-\varepsilon R \sum_{j=1}^{i-1} \xi_j\right) I_{\left\{\sum_{j=1}^{i-1} \xi_j > 0\right\}}\right] \leq \mathbb{E}^{\mathbb{Q}^{(R)}} \left[\exp\left(\tilde{\varepsilon} R \sum_{j=1}^{i-1} \xi_j\right)\right].$$

To choose $\tilde{\varepsilon}$ in a suitable way, we focus on the properties of θ. This function is convex and satisfies $\theta(0) = \theta(R) = 0$ and $\theta'(0) < 0$. Consequently, there exists a $\tilde{r} \in (0, R)$ such that $\theta(\tilde{r}) < 0$ Choosing $\tilde{\varepsilon} = 1 - \frac{\tilde{r}}{R} \in (0, 1)$ we have that

$$\mathbb{E}^{\mathbb{Q}^{(R)}} \left[\exp\left(-\varepsilon R \sum_{j=1}^{i-1} \xi_j\right) I_{\left\{\sum_{j=1}^{i-1} \xi_j > 0\right\}}\right] \leq \mathbb{E}^{\mathbb{Q}^{(R)}} \left[\exp\left(\tilde{\varepsilon} R \sum_{j=1}^{i-1} \xi_j\right)\right] =$$
$$\mathbb{E}^{\mathbb{Q}^{(R)}} \left[\exp\left(\tilde{\varepsilon} R \xi_1\right)\right]^{i-1} = \mathbb{E}_{(u,\lambda)} \left[\exp\left((\tilde{\varepsilon} - 1) R(X_{S(1)} - u)\right) I_{\{S(1)<\infty\}}\right]^{i-1} =$$
$$\mathbb{E}_{(u,\lambda)} \left[\exp\left(-\tilde{r}(X_{S(1)} - u)\right) I_{\{S(1)<\infty\}}\right]^{i-1} = \mathbb{E}^{\mathbb{Q}^{(\tilde{r})}} \left[\exp\left(\theta(\tilde{r}) S(1)\right) I_{\{S(1)<\infty\}}\right]^{i-1}.$$

By construction we have that $\theta(\tilde{r}) < 0$ and $S(1) > 0$; hence,

$$\mathbb{E}^{\mathbb{Q}^{(\tilde{r})}} \left[\exp\left(\theta(\tilde{r}) S(1)\right) I_{\{S(1)<\infty\}}\right] = p < 1.$$

Finally we get

$$\mathbb{E}_{(u,\lambda)}\left[\exp(-r(X_{S_+(1)}-u))I_{\{S_+(1)<\infty\}}\right] \leq C\sum_{i=1}^{\infty} p^{i-1} = \frac{C}{1-p} < \infty.$$

□

Lemma 2.13. *The function $\mathbb{P}_{(u,\lambda)}\left[\tau_u \leq S_+(1), \tau_u < \infty\right] e^{Ru}$ is directly Riemann integrable in u.*

Proof. Let r be as in Assumption 2.4. Observe that $\alpha(r) < 0$ and $\theta(r) > 0$ since $r > R > 0$. At first, we show that $\mathbb{P}_{(u,\lambda)}\left[\tau_u \leq S_+(1), \tau_u < \infty\right] e^{ru}$ is uniformly bounded. Let $t > 0$ be arbitrary but fixed. Then,

$$\begin{aligned}
\mathbb{P}_{(u,\lambda)}\left[\tau_u \leq (S_+(1)\wedge t)\right] e^{ru} &= \mathbb{E}^{\mathbb{Q}^{(r)}}\left[I_{\{\tau_u\leq(S_+(1)\wedge t)\}}e^{\theta(r)\tau_u}e^{rX_{\tau_u}}e^{\alpha(r)\lambda_{\tau_u}}\right]e^{-\alpha(r)\lambda} \\
&\leq \mathbb{E}^{\mathbb{Q}^{(r)}}\left[I_{\{\tau_u\leq(S_+(1)\wedge t)\}}e^{\theta(r)\tau_u}\right]e^{-\alpha(r)\lambda} \\
&\leq \mathbb{E}^{\mathbb{Q}^{(r)}}\left[e^{\theta(r)S_+(1)}\right]e^{-\alpha(r)\lambda} \\
&= \mathbb{E}_{(u,\lambda)}\left[e^{-rX_{S_+(1)}+ru-\alpha(r)\lambda_{S_+(1)}+\alpha(r)\lambda}I_{\{S_+(1)<\infty\}}\right]e^{-\alpha(r)\lambda} \\
&= \mathbb{E}_{(u,\lambda)}\left[I_{\{S_+(1)<\infty\}}e^{-r(X_{S_+(1)}-u)}\right]e^{-\alpha(r)\lambda} < \infty.
\end{aligned}$$

The upper bound is independent of t, so by letting t tend to infinity we get

$$\mathbb{P}_{(u,\lambda)}\left[\tau_u \leq S_+(1), \tau_u < \infty\right] e^{ru} \leq \mathbb{E}_{(u,\lambda)}\left[I_{\{S_+(1)<\infty\}}e^{-r(X_{S_+(1)}-u)}\right]e^{-\alpha(r)\lambda}.$$

This bound is even independent of u. To see this, we consider the process $R_t = ct - \sum_{i=1}^{N_t} U_i$ and define the random time $T_+(1) := \min\left\{S(i) \,|\, R_{S(i)} < 0\right\}$. They are independent of u but under $\mathbb{P}_{(u,\lambda)}$ we have almost surely $R_t = X_t - u$ and $T_+(1) = S_+(1)$. By this, we see that $X_{S_+(1)} - u = R_{T_+(1)}$ does not depend on u.
Using the derived boundedness we get that there is some $K > 0$ such that

$$\mathbb{P}_{(u,\lambda)}\left[\tau_u \leq S_+(1), \tau_u < \infty\right] e^{Ru} \leq Ke^{-(r-R)u},$$

which is a directly Riemann integrable upper bound. Consequently,

$$\mathbb{P}_{(u,\lambda)}\left[\tau_u \leq S_+(1), \tau_u < \infty\right] e^{Ru},$$

is directly Riemann integrable too. □

Let us now focus on the properties of $p(u,x)$.

Lemma 2.14. *The function $p(u,x)$ is continuous in u for $u > 0$.*

Proof. To prove continuity, we will show that

$$\lim_{\varepsilon\to 0} p(u+\varepsilon,x) = \lim_{\varepsilon\to 0} p(u-\varepsilon,x) = p(u,x).$$

We start with the first limit. To do so we will consider a path of our surplus process $\{X_t\}_{t\geq 0}$ with initial capital u and exactly the same path of the process $\{X_t^\varepsilon\}_{t\geq 0}$ with initial capital $u+\varepsilon$. The premium rate c, the claim sizes U_i and the counting process N do not depend on the initial capital; hence, $X_t^\varepsilon = X_t + \varepsilon$. By the same line of arguments as in the proof of Lemma 2.13, we see that $S_+(1)$ and the condition in the definition of p do not depend on u.
To be precise, let $\omega \in \Omega$ be an arbitrary event and let us compare the fixed paths of our processes. If $\{X_t(\omega)\}_{t\geq 0}$ gets ruined before $S_+(1)(\omega)$, there is some $\tilde{\varepsilon} > 0$ such that for all $\varepsilon < \tilde{\varepsilon}$ the path $\{X_t^\varepsilon(\omega)\}_{t\geq 0}$ gets ruined in the same moment. If $\{X_t(\omega)\}_{t\geq 0}$ stays greater or equal to 0 then $\{X_t^\varepsilon\}_{t\geq 0}$ stays positive for all $\varepsilon > 0$. Consequently, we have that $\lim_{\varepsilon\to 0} I_{\{\tau_{u+\varepsilon}<S_+(1)\}}(\omega) = I_{\{\tau_u<S_+(1)\}}(\omega)$ and by dominated convergence also $p(u+\varepsilon,x) \to p(u,x)$.
If we can exclude the case, that $\{X_t\}_{t\geq 0}$ hits exactly the value 0, then the same arguments hold for $X_t^{-\varepsilon} := X_t - \varepsilon$.
The infimum of the surplus process can only occur at a jump time of our counting process $\{N_t\}_{t\geq 0}$. Let T be an arbitrary claim time, then,

$$\mathbb{P}_{(u,\lambda)}\left[X_T = 0\right] = \mathbb{P}_{(u,\lambda)}\left[X_{T-} - U_{N_T} = 0\right] = \mathbb{E}_{(u,\lambda)}\left[\mathbb{P}_{(u,\lambda)}\left[X_{T-} - U_{N_T} = 0 \mid \mathcal{F}_{T-}\right]\right].$$

The random variable U_{N_T} is independent of $\mathcal{F}_{T-}$ and its distribution is continuous. Hence, the probability of hitting exactly the value X_{T-} is 0. Consequently, $\mathbb{P}_{(u,\lambda)}\left[X_T = 0\right] = 0$. Since we have only countably many jump times, the event that the surplus process hits 0 at any jump time has measure 0 too. Hence, $p(u-\varepsilon,x) \to p(u,x)$. Combining these results we get that $p(u,x)$ is continuous in u. □

Lemma 2.15. *Under our assumptions* $\int_0^u p(u,x)e^{Rx}\,B(\mathrm{d}x)$ *is directly Riemann integrable.*

Proof. Again let r be as in Assumption 2.4. Then,

$$\begin{aligned}\int_0^u p(u,x)e^{Rx}\,B(\mathrm{d}x) &\leq e^{Ru}\int_0^u p(u,x)\,B(\mathrm{d}x) = e^{Ru}\mathbb{P}_{(u,\lambda)}\left[\tau_u \leq S_+(1) < \infty\right] \\ &\leq e^{Ru}\mathbb{P}_{(u,\lambda)}\left[\tau_u \leq S_+(1), \tau_u < \infty\right] \leq Ke^{-(r-R)u}.\end{aligned}$$

As before we have a directly Riemann integrable upper bound, and therefore

$$\int_0^u p(u,x)e^{Rx}\,B(\mathrm{d}x),$$

is directly Riemann integrable. □

The continuity of the distribution of U implies that B is not arithmetic. Consequently, all conditions of Theorem 2.2 are satisfied. Hence, we can apply it to the renewal equation satisfied by $\psi(u)e^{Ru}$ and obtain our main result.

Theorem 2.16. *Under our Assumptions* $\lim_{u\to\infty}\psi(u,\lambda)e^{Ru}$ *exists and is finite.*

Finally, we consider an example where all our assumptions are satisfied.

Example 2.3. Let Y and U be exponentially distributed with parameter 1, $\beta = 1$, $\rho = 1.5$, $\lambda = 1$ and $c = \frac{15}{4}$. The net profit condition is satisfied since

$$c = \frac{15}{4} > \frac{3}{2} = \frac{\rho}{\beta}\mathbb{E}[Y]\,\mathbb{E}[U].$$

Further, the moment-generating function of U is given by $M_U(r) = \frac{1}{1-r}$ and $\alpha(r) = 1 - M_U(r) = -\frac{r}{1-r}$.
Consequently, $M_Y(-\alpha(r)) = \frac{1-r}{1-2r}$ is well defined for all $r < \frac{1}{2}$ and the adjustment coefficient is given by $R = \frac{3}{10}$. The measure $\mathbb{Q}^{(R)}$ is well defined and the new intensity is $\tilde{\rho}^{(R)} = \frac{21}{8}$. Since this is greater than 1, we know from Example 2.2 that there are infinitely many recurrence times $\{S(i)\}_{i\in\mathbb{N}}$ under $\mathbb{Q}^{(R)}$.
Choosing $r = \frac{1}{3} > R$, we see that $\mathbb{Q}^{(r)}$ is well defined and $\tilde{\rho}^{(r)} = 3 \in \mathbb{N}$. Following the results shown in Orsingher and Battaglia (1982), we know that under the measure $\mathbb{Q}^{(r)}$, the recurrence times $\{S(i)\}_{i\in\mathbb{N}}$ have intensity

$$\nu(t) = \frac{1}{2e}\left(1 + 3e^{-t} - 3e^{-2t} - e^{-3t}\right) \le \frac{2(\sqrt{2}-1)}{e}.$$

Therefore,

$$\begin{aligned}\mathbb{E}_{(u,\lambda)}\left[\exp\left(r(X_{S(1)} - u)\right) I_{\{S(1)<\infty\}}\right] &= \mathbb{E}^{\mathbb{Q}^{(r)}}\left[\exp\left(\theta(r)S(1)\right)\right] \\ &= \int_0^\infty \exp(0.25s)\nu(s)\exp\left(-\int_0^s \nu(u)\,\mathrm{d}u\right)\mathrm{d}s.\end{aligned}$$

For $t \ge 1$ we have that $\nu(t) > 0.26$ which gives us the existence of some constant c such that

$$\int_0^\infty \exp(0.25s)\nu(s)\exp\left(-\int_0^s \nu(u)\,\mathrm{d}u\right)\mathrm{d}s < c\int_0^\infty e^{0.25s}e^{-0.26(s-1)}\,\mathrm{d}s < \infty.$$

By this, all assumptions made are satisfied. Hence, there exists some constant C such that

$$\lim_{u\to\infty}\psi(u,\lambda)e^{0.3u} = C.$$

Acknowledgments

This research was funded in whole, or in part, by the Austrian Science Fund (FWF) P 33317. For the purpose of open access, the author has applied a CC BY public copyright licence to any Author Accepted Manuscript version arising from this submission.

3 Level crossing of the Markovian shot-noise process

The following chapter is based on the paper Pojer (2022), which is submitted for publication.

3.1 Introduction

The main motivation for us to consider the Markovian shot-noise process and its recurrence behaviour are the results derived by Pojer and Thonhauser (2023a). In this work, a risk model with Markovian shot-noise intensity is considered, and it is shown that under certain conditions, the ruin probability converges at a fixed exponential rate. One of the main assumptions made there is the recurrent behaviour of the shot-noise process, i.e. that it crosses a certain level infinitely often almost surely.
A first result, ensuring this assumption is satisfied, was already shown by Orsingher and Battaglia (1982). There, an explicit formula for the intensity of the level crossings is derived for a Markovian shot-noise process with exponentially distributed shock events. For more general shot-noise environments and distributions, there are papers studying the expected number of crossings, see Biermé and Desolneux (2012b), and Biermé and Desolneux (2012a). However, the results derived therein are not strong enough to ensure that the process crosses the level almost surely infinitely often.

3.2 Level Crossings

At first, we briefly introduce the considered Markovian shot-noise model as it is used by Pojer and Thonhauser (2023a).

Definition 3.1. Let $(\Omega, \mathcal{F}, \mathbb{P})$ be a probability space, big enough to contain all future stochastic objects, λ and β positive constants and $\{N_t^\rho\}_{t\geq 0}$ a homogeneous Poisson process with intensity ρ and jump times $\{T_n^\rho\}_{n\geq 1}$. Let further $\{Y_n\}_{n\geq 1}$ be i.i.d. copies of a positive random variable Y with distribution F_Y, which satisfies $F_Y(y) < 1$ for all $y \in \mathbb{R}_+$ and which are independent of $\{N_t^\rho\}_{t\geq 0}$. Then, the process

$$\lambda_t := \lambda e^{-\beta t} + \sum_{n=1}^{N_t^\rho} Y_n e^{-\beta(t-T_n^\rho)},$$

is called Markovian shot-noise process.

The main question considered here is if we can give conditions under which the Markovian shot-noise process returns to some certain level b infinitely often with probability 1. To show such results, we will use Theorem 4.3.4 of Durrett (2019).

Theorem 3.1 (Second Borel-Cantelli lemma II)**.** *Let $\{\mathcal{F}_n\}_{n\geq 0}$ be a filtration with $\mathcal{F}_0 = \{\emptyset, \Omega\}$ and $\{B_n\}_{n\geq 1}$ a sequence of events with $B_n \in \mathcal{F}_n$. Then,*

$$\{B_n \, i.o.\} = \left\{ \sum_{n=1}^{\infty} \mathbb{P}\left[B_n \,|\mathcal{F}_{n-1}\right] = \infty \right\}.$$

This version of the second Borel-Cantelli lemma allows structural dependencies without the common restriction of increasing events.

Lemma 3.2. *Let $\{\lambda_t\}_{t\geq 0}$ be a Markovian shot-noise process and $b > 0$. Then, almost surely there are infinitely many $n \in \mathbb{N}$ such that $\lambda_{T_n^\rho} > b$.*

Proof. Let $n \in \mathbb{N}$ be arbitrary. Then, we have almost surely that $\lambda_{T_n^\rho} > Y_n$. Now, consider the independent events $B_n = \{Y_n > b\}$. Then

$$\sum_{n=1}^{\infty} \mathbb{P}\left[B_n\right] = \sum_{n=1}^{\infty} \bar{F}_Y(b) = \infty.$$

By the second Borel-Cantelli lemma, see Theorem 2.3.7. in Durrett (2019), this implies that this happens a.s. infinitely often. Consequently, almost surely there are infinitely many $\lambda_{T_n^\rho} > b$. □

Since all shock events a strictly positive, the process $\{\lambda\}_{t\geq 0}$ decreases only along the continuous exponential curve. This implies that if $\lambda_{T_n^\rho}$ is infinitely often above b and $\lambda_{T_n^\rho -}$ infinitely often below this level, there must be infinitely many crossings of the level b.

Theorem 3.3. *Let $\{\lambda_t\}_{t\geq 0}$ be a Markovian shot-noise process and $b > 0$. Further, let $\frac{\rho}{\beta} \leq 1$ and $\mathbb{E}\left[Y\right] < \infty$. Then, $\{\lambda_t\}_{t\geq 0}$ crosses the level b infinitely often with probability 1.*

Proof. As already mentioned before, we will use Borel-Cantelli II. For this, we consider the sets $B_n = \left\{\lambda_{T_n^\rho -} < b\right\}$ and filtration $\mathcal{F}_n = \sigma\left(\{\lambda_t \,|\, t \leq T_n^\rho\}\right)$. The set $\{B_n \, i.o.\}$ has the same measure as the set of infinitely many crossings of b since there are a.s. infinitely many $\lambda_{T_n^\rho}$ above b by Lemma 3.2. The conditional probability of the event B_n is $\mathbb{P}\left[B_n \,|\mathcal{F}_{n-1}\right] = I_{\left\{\lambda_{T_{n-1}^\rho} > b\right\}} \left(\frac{b}{\lambda_{T_{n-1}^\rho}}\right)^{\frac{\rho}{\beta}} + I_{\left\{\lambda_{T_{n-1}^\rho} \leq b\right\}} = \min\left(\left(\frac{b}{\lambda_{T_{n-1}^\rho}}\right)^{\frac{\rho}{\beta}}, 1\right)$. The random variable $\lambda_{T_{n-1}^\rho}$ is almost surely less or equal to the starting intensity plus the sum of the shock events $\lambda + \sum_{i=1}^{n-1} Y_i$, and $\frac{1}{n-1}\left(\lambda + \sum_{i=1}^{n-1} Y_i\right)$ converges a.s. to $\mathbb{E}\left[Y\right]$ by the strong law of large numbers. Therefore, for a.e. $\omega \in \Omega$ we have that there exists a $K(\omega)$ such that $\frac{1}{n}\left(\lambda + \sum_{i=1}^{n} Y_i(\omega)\right) < \mathbb{E}\left[Y\right] + b$ for all $n \geq K(\omega)$. Consequently, we have for almost every $\omega \in \Omega$ that

$$
\begin{aligned}
\sum_{n=1}^{\infty} \mathbb{P}\left[B_n \left| \mathcal{F}_{n-1}\right.\right](\omega) &\geq \sum_{n=1}^{\infty} \min\left(1, \frac{1}{n^{\frac{\rho}{\beta}}}\left(\frac{nb}{\sum_{k=1}^{n} Y_k(\omega)}\right)^{\frac{\rho}{\beta}}\right) \\
&\geq \sum_{n=K(\omega)}^{\infty} \min\left(1, \frac{1}{n^{\frac{\rho}{\beta}}}\left(\frac{b}{\mathbb{E}\left[Y\right]+b}\right)^{\frac{\rho}{\beta}}\right) \\
&= \left(\frac{b}{\mathbb{E}\left[Y\right]+b}\right)^{\frac{\rho}{\beta}} \sum_{n=K(\omega)}^{\infty} \frac{1}{n^{\frac{\rho}{\beta}}} = \infty.
\end{aligned}
$$

□

This is a first result but it restricts the possible choices of the parameters. The reason, why we cannot extend this to the case $\rho > \beta$ is, that the series $\sum_{n=1}^{\infty} \frac{1}{n^{\frac{\rho}{\beta}}}$ converges. To bypass this, we use the equation

$$
\min\left(\left(\frac{b}{\lambda_{T_{n-1}^{\rho}}}\right)^{\frac{\rho}{\beta}}, 1\right) = \min\left(\frac{1}{n-1}\left(\frac{(n-1)^{\frac{\beta}{\rho}} b}{\lambda_{T_{n-1}^{\rho}}}\right)^{\frac{\rho}{\beta}}, 1\right).
$$

If we can show that the term $\frac{\lambda_{T_n^{\rho}}}{n^{\frac{\beta}{\rho}}}$ converges almost surely, we can use the same ideas as before to prove that there are infinitely many crossings of the level b.

Lemma 3.4. *Assume, that the distribution F_Y is light-tailed, i.e. that there exists some $R > 0$ such that $M_Y(R) := \mathbb{E}\left[e^{RY}\right] < \infty$ and that $\frac{\rho}{\beta} > 1$. Then, $\frac{\lambda_{T_n^{\rho}}}{n^{\frac{\beta}{\rho}}}$ converges almost surely to* 0.

Proof. To show a.s. convergence, we consider a $K > \frac{2\rho}{\beta}$ and show that $\mathbb{E}\left[\lambda_{T_n^{\rho}}^{K}\right]$ is finite. Then, we use Markov's inequality to show that for every $\varepsilon > 0$

$$
\sum_{n=1}^{\infty} \mathbb{P}\left[\lambda_{T_n^{\rho}} > \varepsilon n^{\frac{\beta}{\rho}}\right] \leq \sum_{n=1}^{\infty} \frac{C}{n^2} < \infty.
$$

Having this, the first Borel-Cantelli lemma, see Theorem 2.3.1 in Durrett (2019) gives us that a.s. only finitely many events $\lambda_{T_n^{\rho}} > \varepsilon n^{\frac{\beta}{\rho}}$ happen, i.e. that $\frac{\lambda_{T_n^{\rho}}}{n^{\frac{\beta}{\rho}}}$ converges to 0 almost surely.
Let now K be as desired and $r > 0$ such that $M_Y(r) < (1 + \frac{\beta}{\rho})$ and $\frac{K}{re} > 1$. Such r exists since $M_Y(s)$ is continuous in $s = 0$ and well defined for some $R > 0$. Then we know that

for every $m \geq 0$ the following inequality holds true

$$\mathbb{E}\left[Y^m\right] \leq \left(\frac{m}{re}\right)^m M_Y(r).$$

Using the multinomial-theorem, we get that

$$\mathbb{E}\left[\lambda_{T_n^\rho}^K\right] = \mathbb{E}\left[\sum_{k_0+\ldots k_n=K} \binom{K}{k_0,\ldots k_n} \lambda^{k_0} \exp(-\beta k_0 T_n^\rho) \prod_{i=1}^{n} Y_i^{k_i} \exp(-\beta k_i (T_n^\rho - T_i^\rho))\right].$$

The random variable $Y_n^{k_n}$ is independent of the other stochastic objects, hence this is

$$\mathbb{E}\left[\lambda_{T_n^\rho}^K\right] = \sum_{k_0+\ldots k_n=K} \binom{K}{k_0,\ldots k_n} \lambda^{k_0} \mathbb{E}[Y^{k_n}] \mathbb{E}\left[\exp(-\beta k_0 T_n^\rho) \prod_{i=1}^{n-1} Y_i^{k_i} \exp(-\beta k_i (T_n^\rho - T_i^\rho))\right].$$

Now, the random variable $\exp(-\beta(K-k_n)(T_n^\rho - T_{n-1}^\rho))$ is independent from the remaining random variables and has expectation

$$\mathbb{E}\left[\exp(-\beta(K-k_n)(T_n^\rho - T_{n-1}^\rho))\right] = \frac{\rho}{\rho + \beta \sum_{j=0}^{n-1} k_j}.$$

Doing this iteratively, we get that

$$\mathbb{E}\left[\lambda_{T_n^\rho}^K\right] = \sum_{k_0+\ldots k_n=K} \binom{K}{k_0,\ldots k_n} \lambda^{k_0} \prod_{i=1}^{n} \frac{\mathbb{E}\left[Y^{k_i}\right]\rho}{\rho + \beta \sum_{j=0}^{i-1} k_i}.$$

If we define $l := \min\{i \,|k_i > 0\}$, then we have for every $i > l$, that

$$\mathbb{E}\left[Y^{k_i}\right] \frac{\rho}{\rho + \beta \sum_{j=0}^{i-1} k_j} \leq \left(\frac{K}{re}\right)^{k_i} \frac{\rho M_Y(r)}{\rho + \beta},$$

and for $i = l$

$$\mathbb{E}\left[Y^{k_l}\right] \frac{\rho}{\rho + \beta \sum_{j=0}^{l-1} k_j} \leq \left(\frac{K}{re}\right)^{k_l} M_Y(r).$$

Now we separate the sum for different values of the first non-zero index

$$\sum_{k_0+\ldots k_n=K} \binom{K}{k_0,\ldots k_n} \lambda^{k_0} \prod_{i=1}^{n} \frac{\mathbb{E}\left[Y^{k_i}\right]\rho}{\rho+\beta\sum_{j=0}^{i-1} k_i} =$$

$$\sum_{\substack{k_0+\cdots+k_n=K \\ k_0>0}} \binom{K}{k_0,\ldots k_n} \lambda^{k_0} \prod_{i=1}^{n} \frac{\mathbb{E}\left[Y^{k_i}\right]\rho}{\rho+\beta\sum_{j=0}^{i-1} k_i} +$$

$$\sum_{l=1}^{n} \sum_{\substack{k_l+\cdots+k_n=K \\ k_l>0}} \binom{K}{k_l,\ldots k_n} \prod_{i=l}^{n} \frac{\mathbb{E}\left[Y^{k_i}\right]\rho}{\rho+\beta\sum_{j=l}^{i-1} k_i}.$$

Applying the inequalities derived before, we get that

$$\sum_{\substack{k_0+\cdots+k_n=K \\ k_0>0}} \binom{K}{k_0,\ldots k_n} \lambda^{k_0} \prod_{i=1}^{n} \frac{\mathbb{E}\left[Y^{k_i}\right]\rho}{\rho+\beta\sum_{j=0}^{i-1} k_i}$$

$$\leq \left(\lambda\frac{K}{re}\right)^{K} \left(\frac{M_Y(r)\rho}{\rho+\beta}\right)^{n} \sum_{\substack{k_0+\cdots+k_n=K \\ k_0>0}} \binom{K}{k_0,\ldots k_n}$$

$$\leq \left(\lambda\frac{K}{re}\right)^{K} \left(\frac{M_Y(r)\rho}{\rho+\beta}\right)^{n} (n+1)^K,$$

and for general l

$$\sum_{\substack{k_l+\cdots+k_n=K \\ k_l>0}} \binom{K}{k_l,\ldots k_n} \prod_{i=l}^{n} \frac{\mathbb{E}\left[Y^{k_i}\right]\rho}{\rho+\beta\sum_{j=l}^{i-1} k_i}$$

$$\leq \left(\frac{K}{re}\right)^{K} M_Y(r) \left(\frac{\rho M_Y(r)}{\rho+\beta}\right)^{n-l} \sum_{\substack{k_l+\cdots+k_n=K \\ k_l>0}} \binom{K}{k_l,\ldots k_n}$$

$$\leq \left(\frac{K}{re}\right)^{K} M_Y(r) \left(\frac{\rho M_Y(r)}{\rho+\beta}\right)^{n-l} (n-l+1)^K.$$

Consequently, this yields

$$\sum_{\substack{k_0+\cdots+k_n=K\\k_0>0}}\binom{K}{k_0,\dots k_n}\lambda^{k_0}\prod_{i=1}^{n}\frac{\mathbb{E}\left[Y^{k_i}\right]\rho}{\rho+\beta\sum_{j=0}^{i-1}k_i}+\sum_{l=1}^{n}\sum_{\substack{k_l+\cdots+k_n=K\\k_l>0}}\binom{K}{k_l,\dots k_n}\prod_{i=l}^{n}\frac{\mathbb{E}\left[Y^{k_i}\right]\rho}{\rho+\beta\sum_{j=l}^{i-1}k_i}$$

$$\leq\left(\lambda\frac{K}{re}\right)^K\left(\frac{M_Y(r)\rho}{\rho+\beta}\right)^n(n+1)^K+\sum_{l=1}^{n}\left(\frac{K}{re}\right)^K M_Y(r)\left(\frac{\rho M_Y(r)}{\rho+\beta}\right)^{n-l}(n-l+1)^K$$

$$\leq C\sum_{i=0}^{n}\left(\frac{M_Y(r)\rho}{\rho+\beta}\right)^i(i+1)^K,$$

where $C>0$ depends on K, but is independent of n. This series converges for $n\to\infty$; hence, the K-th moments of $\lambda_{T_n^\rho}$ are bounded uniformly in n.
By Markov's inequality, we get for all $\varepsilon>0$

$$\sum_{n=1}^{\infty}\mathbb{P}\left[\lambda_{T_n^\rho}>\varepsilon n^{\frac{\beta}{\rho}}\right]\leq\sum_{n=1}^{\infty}\frac{\mathbb{E}\left[\lambda_{T_n^\rho}^K\right]}{\varepsilon^K n^{\frac{K\beta}{\rho}}}<\infty,$$

where the last inequality holds by the boundedness of the K-th moment and the fact that $\frac{K\beta}{\rho}>2$. □

Having this, we can follow the same steps as in the proof of Theorem 3.3 to show that there are almost surely infinitely many crossings of b.

Theorem 3.5. *Let $b>0$ and assume that the distribution F_Y is light-tailed, i.e. that there exists some $R>0$ such that $M_Y(R):=\mathbb{E}\left[e^{RY}\right]<\infty$. Then, the Markovian shot-noise process crosses the level b infinitely often probability 1.*

Proof. Since we have already considered the case $\frac{\rho}{\beta}<1$ in Theorem 3.3, we can focus on the case, where $\rho>\beta$. Again, we will use the second Borel-Cantelli lemma II, and consider for all $n\in\mathbb{N}$ the events $B_n=\left\{\lambda_{T_n^\rho-}\leq b\right\}$ and the filtration $\mathcal{F}_n=\sigma\left(\{\lambda_t\mid t\leq T_n^\rho\}\right)$. The equality $\{B_n\, i.o.\}=\{\sum_{n=1}^\infty\mathbb{P}\left[B_n\,|\mathcal{F}_{n-1}\right]=\infty\}$ holds still true. For fixed $n\in\mathbb{N}$, the conditional probability $\mathbb{P}\left[B_n\,|\mathcal{F}_{n-1}\right]$ is

$$\mathbb{P}\left[B_n\,|\mathcal{F}_{n-1}\right]=I_{\left\{\lambda_{T_{n-1}^\rho}\leq b\right\}}+I_{\left\{\lambda_{T_{n-1}^\rho}>b\right\}}\left(\frac{b}{\lambda_{T_{n-1}^\rho}}\right)^{\frac{\rho}{\beta}}=\min\left(1,\left(\frac{b}{\lambda_{T_{n-1}^\rho}}\right)^{\frac{\rho}{\beta}}\right)$$

$$=\min\left(1,\frac{1}{n-1}\left((n-1)^{\frac{\beta}{\rho}}\frac{b}{\lambda_{T_{n-1}^\rho}}\right)^{\frac{\rho}{\beta}}\right).$$

Let $\varepsilon>0$ be arbitrary, then there are for almost all $\omega\in\Omega$ only finitely many k such that $\frac{\lambda_{T_n^\rho}}{n^{\frac{\beta}{\rho}}} > \varepsilon$, i.e. there exists a $K(\omega)$ such that for all $n \geq K(\omega)$ it holds that $\frac{\lambda_{T_n^\rho}}{n^{\frac{\beta}{\rho}}} < \varepsilon$. Consequently,for almost all ω we have that

$$\sum_{n=1}^{\infty} \mathbb{P}\left[B_n \left|\mathcal{F}_{n-1}\right.\right](\omega) \geq \sum_{n=K(\omega)}^{\infty} \min\left(1, \frac{1}{n}(b\varepsilon)^{\frac{\rho}{\beta}}\right) = \infty.$$

By this, there are almost surely infinitely many n such that $\lambda_{T_{n-}^\rho} < b$ and infinitely many crossings of the level b. □

Acknowledgments

This research was funded in whole, or in part, by the Austrian Science Fund (FWF) P 33317. For the purpose of open access, the author has applied a CC BY public copyright licence to any Author Accepted Manuscript version arising from this submission. Declaration of interests

4 Exact asymptotics of ruin probabilities with linear Hawkes arrivals

The following chapter is based on the paper Palmowski, Pojer, et al. (2023), which is submitted for publication.

4.1 Introduction

The concept of a ruin probability lies in the center of risk theory since the classical Cramér-Lundberg model was introduced by Lundberg (1903). It is defined as the probability that the surplus $\{X_t\}_{t\geq 0}$ of an insurance portfolio falls below zero, i.e.

$$\psi(u) := \mathbb{P}_u[\tau < +\infty],$$

where

$$\tau := \inf\{t \geq 0 : X_t < 0\} \tag{4.1}$$

denotes the time of ruin. Usually, it is assumed that the so-called net profit condition holds, that is,

$$\lim_{t\to+\infty} X_t = +\infty \qquad \text{a.s..} \tag{4.2}$$

This requirement is necessary to ensure $\psi(u) < 1$.

The event of ruin is a technical term - it does not mean that the company goes bankrupt at the time of ruin. A downward crossing of some fixed level (that could be treated as level zero) of the surplus signals the insurance company to increase its profitability. Further, ruin theory can be used to determine solvency capital requirements.

In this paper, we consider the ruin probability of the risk process

$$X_t := u + ct - \sum_{i=1}^{N_t} U_i, \tag{4.3}$$

where u denotes the positive initial capital and c the premium rate. The claim sizes $\{U_i\}_{i\in\mathbb{N}}$ are positive, i.i.d. random variables with cumulative distribution function F_U, and independent of the arrival process $\{N_t\}_{t\geq 0}$, which is assumed to be a marked linear Hawkes process. We recall that this means that $\{N_t\}_{t\geq 0}$ is a simple point process with intensity process

$$\lambda_t := a + \sum_{T_i \le t} h(t - T_i, Y_i),$$

where a is a constant baseline intensity and the shock sizes $\{Y_i\}_{i\in\mathbb{N}}$ are a sequence of positive i.i.d. random variables with distribution function F_Y. Further, $\{T_i\}_{i\in\mathbb{N}}$ are the jump times of $\{N_t\}_{t\ge 0}$, and $h : \mathbb{R}_{\ge 0} \times \mathbb{R}_{\ge 0} \to \mathbb{R}_{\ge 0}$ is a decay function satisfying

$$\mathbb{E}\left[\int_0^\infty h(t, Y)\mathrm{d}t\right] =: \mu < 1 \quad \text{and} \quad \mathbb{E}\left[\int_0^\infty t\, h(t, Y)\mathrm{d}t\right] < +\infty. \tag{4.4}$$

We assume that the considered stochastic quantities are defined on a suitable filtered probability space $(\Omega, \mathcal{F}, \{\mathcal{F}_t\}_{t\ge 0}, \mathbb{P})$ where the filtration is assumed to be right continuous but not necessarily $\mathbb{P}$-complete. To stress the dependence of the distribution of certain stochastic objects on the initial capital u, we write $\mathbb{P}_u$ for the measure $\mathbb{P}$ with condition $X_0 = u$, and $\mathbb{E}_u[\cdot]$ for the corresponding expectation.

Incorporating Hawkes arrival processes into risk theory is important due to their self-exciting structure, which is, for example, observable in the occurrence of claims due to cyber risks, as described in Bessy-Roland et al. (2021). This behaviour can be described as follows: every jump of a Hawkes process increases its own intensity; hence, the probability of further jumps rises with every event. Hawkes processes have another crucial interpretation: they are Poisson cluster processes. Clusters of claims appear according to a homogeneous Poisson process $\{\tilde{N}_t\}_{t\ge 0}$ with the baseline intensity a. These clusters generate claims according to a Galton-Watson branching process with an offspring distribution whose mean is $\mu < 1$. In other words, one event that arrives according to the Poisson process $\{\tilde{N}_t\}_{t\ge 0}$ produces a chain of further claims that are reported to the insurance company delayed over time.

Generally, the ruin probability cannot be expressed in a closed form. To bypass this drawback, it is essential to find appropriate bounds or approximations which behave asymptotically like the ruin probability as the initial capital u becomes large. The main goal of this paper is to derive such results for the aforementioned linear Hawkes model.

In the beginning, we compare our process to a version of it, modified by shifting all claims of a cluster to its beginning. This approach leads to two statements. Under some light-tailed assumptions, we prove in Corollary 4.5 that there exist positive constants C_-, C_+ and an adjustment coefficient $R > 0$ such that

$$C_- e^{-Ru} \le \psi(u) \le C_+ e^{-Ru}. \tag{4.5}$$

We want to underline that this result implies Theorem 4.1 of Stabile and Torrisi (2010) who derived the logarithmic asymptotic behaviour by using large deviation theory to show that $\lim_{u\to+\infty} \frac{1}{u} \ln \psi(u) = -R$. A similar result as (4.5) was shown in Theorem 4.5 of Albrecher and Asmussen (2006) for the Cox claim arrival process with a Poisson shot noise intensity. Our argumentation relies on the ideas presented by these authors, but their proof has a petite technical flaw - the according result is still correct. They assumed that if one considers all claims of the first n clusters, the last occurring claim always belongs to the n-th cluster. This premise is incorrect since claims which belong to previous clusters might still appear with positive probability after the n-th cluster has finished. We bypass this problem by considering the remaining time it takes to work off all clusters instead.

Under the complementary assumption of strongly subexponential (hence heavy-tailed) claim sizes, we prove in Corollary 4.7 that

$$\lim_{u\to\infty} \frac{\psi(u)}{1-F_U^s(u)} = C_h,$$

for an explicitly identified constant C_h, where

$$F_U^s(u) := \frac{1}{\mathbb{E}\left[U_1\right]} \int_0^u (1-F_U(y))\mathrm{d}y, \tag{4.6}$$

denotes the integrated tail distribution of the claim sizes. The above results generalize the findings in Proposition 13 of Zhu (2013), where the authors considered a risk process driven by a non-marked linear Hawkes process.

To identify the exact asymptotics of the ruin probability $\psi(u)$ in the light-tailed case, we additionally assume that the intensity process is Markovian, i.e.

$$\lambda_t := a + (\lambda - a)e^{-\beta t} + \sum_{i=1}^{N_t} Y_i e^{-\beta(t-T_i)},$$

for some $\lambda > a > 0$ and $\beta > 0$. Using change of measure techniques and renewal arguments, we show that in this case

$$\lim u \to \infty \psi(u) e^{Ru} = C, \tag{4.7}$$

for a positive constant C.

In the proof, we show a few other facts that are of interest on their own. For example, we find sufficient conditions for the intensity process to be positive Harris recurrent. Further, we prove that the corresponding recurrence times are light-tailed. We use these recurrence epochs of the intensity process combined with ladder epochs of the risk process to construct new time points $\{\phi_i\}_{i\in\mathbb{N}}$ that allow us to formulate a renewal-type equation for the ruin probability multiplied by an appropriate exponential function appearing on the right-hand side of (4.7). To prove that the limit of the solution of this renewal-type equation exists and is finite we identify sufficient conditions for the directly Riemann integrability of some functions appearing in this equation. This is possible due to a detailed analysis of exponential moments of X_{ϕ_1}. We believe that our approach and the ideas of the proofs presented in this paper can be applied to other cluster arrival processes too.

This manuscript is organized as follows. In Section 4.2, we construct a modified version of the risk process to show two-sided exponential bounds for the ruin probability $\psi(u)$ in the light-tailed case and the heavy-tailed asymptotics of $\psi(u)$. In Section 4.3 we study the theoretical properties of Markovian Hawkes processes and prove the aforementioned Cramér-Lundberg asymptotics for $\psi(u)$ using renewal arguments and the change of measure technique. We finish our paper with a detailed analysis of the with exponentially distributed claims and shocks (see Section 4.4).

4.2 Cluster representation and a modified risk process

4.2.1 Cluster representation of the risk process

We consider a risk process driven by a linear marked Hawkes process $\{N_t\}_{t\geq 0}$ satisfying (4.4). For convenience, we will omit the properties 'linear' and 'marked' for the rest of the paper and refer to this process as Hawkes process. In this section, the so-called *Poisson cluster representation* or simply cluster representation is crucial. For this, we consider the influence of a single shock event due to the baseline intensity a, called a base event. If such a jump occurs, the counting process $\{N_t\}_{t\geq 0}$ increases by 1, and the intensity increases by $h(0,y)$, where y is a realisation of a random variable with distribution function F_Y. This increase of the intensity triggers $\mathrm{Poi}(\int_0^\infty h(t,y)\mathrm{d}t)$ further jumps, which we denote by *children*, where $\mathrm{Poi}(\alpha)$ denotes the random variable with Poisson distribution with a parameter a. The occurrence of a single child increases the counting process by 1 and triggers again $\mathrm{Poi}(\int_0^\infty h(t,y')\mathrm{d}t)$ new jumps, where y' is a new, independent realisation with the same distribution. We collect the offspring of a base event and call it *cluster*. An important question which appears naturally is if such a cluster consists of finitely many points or may explode. As Basrak et al. (2019) showed the number of points κ in such a generic cluster coincides with the number of points in a subcritical Galton-Watson branching process. By this, we have that κ is finite almost surely and has finite expectation $\mathbb{E}[\kappa] = \frac{1}{1-\mu}$.
By this, we can rewrite the intensity process as

$$\lambda_t = a + \sum_{i=1}^{\bar{N}_t} \sum_{j=0}^{K^{(i)}_{t-\bar{T}_i}} h\left(t - (\bar{T}_i + \sum_{k=1}^{j} T_{ik}), Y_{ij}\right),$$

where $\{\bar{N}_t\}_{t\geq 0}$ denotes a Poisson process with constant intensity a and jump times $\bar{T}_i$. For fixed i, the process $\{K_t^{(i)}\}_{t\geq 0}$ counts the offspring of the i-th base event and its jump times are given by $\bar{T}_i + T_{i1}, \bar{T}_i + T_{i1} + T_{i2}, \ldots$, where T_{ij} denotes the inter-jump time between the $j-1$-th and j-th jump of cluster i. The random variables Y_{ij} correspond to the shock of the j-th event in the i-th cluster, and Y_{i0} is the shock due to the base event of the i-th cluster. Since every cluster has the same distribution, we have that for all $j \geq 1$, the sequence $\{\sum_{k=1}^{j} T_{ik}\}_{i\geq 1}$ is i.i.d. By the same procedure, we can rewrite the surplus process as

$$X_t = u + ct - \sum_{i=1}^{\bar{N}_t} \sum_{j=0}^{K_{t-\bar{T}_i}^{(i)}} U_{ij},$$

where the random variable U_{ij} denotes the claim due to the j-th event of the i-th cluster. Since the number of events of a single cluster is finite almost surely, we have that $K_t^{(i)} \to K_\infty^{(i)} =: \kappa_i$, almost surely as $t \to \infty$. Since all clusters have the same distribution, we have that $\{\kappa_i\}_{i\in\mathbb{N}}$ is an i.i.d. sequence of random variables satisfying

$$\mathbb{E}\left[\kappa_i\right] = \frac{1}{1-\mu}; \tag{4.8}$$

see e.g. p. 203 of Daley and Vere-Jones (2003).

This representation shows us the main feature of our model. An underlying Poisson process triggers with every jump a cluster consisting of a random number κ_i of claims. These claims do not occur immediately but are delayed in time. We will use this representation to derive pathwise bounds for the surplus process.

4.2.2 Upper and lower bounds for the surplus

To derive a lower bound for the surplus process we ignore the mentioned delay in time. We define the clustered process $\{\tilde{X}_t\}_{t\geq 0}$ as

$$\tilde{X}_t = u + ct - \sum_{i=1}^{\bar{N}_t} \sum_{j=0}^{\kappa_i} U_{ij} =: u + ct - \sum_{i=1}^{\bar{N}_t} \tilde{U}_i, \tag{4.9}$$

where

$$\tilde{U}_i := \sum_{j=0}^{\kappa_i} U_{ij}, \qquad i \in \mathbb{N} \tag{4.10}$$

form an i.i.d. sequence independent of the counting Poisson process $\{\bar{N}_t\}_{t\geq 0}$, counting the number of clusters. Since $\kappa_i \geq K_t^{(i)}$ for all t almost surely, we have that

$$X_t \geq \tilde{X}_t. \tag{4.11}$$

for any realisation of the arrival and claim processes. To avoid trivial cases, we assume that the net profit condition

$$c > \mathbb{E}[\tilde{U}]a, \tag{4.12}$$

holds for the modified risk process $\{\tilde{X}_t\}_{t\geq 0}$. Observe that the clustered process $\{\tilde{X}_t\}_{t\geq 0}$ is now a Cramér-Lundberg process. Consequently, we can use standard results for this process to obtain an upper bound for the ruin probability $\psi(u)$ of our surplus process

$\{X_t\}_{t\geq 0}$. Let

$$\tilde{\psi}(u) := \mathbb{P}_u(\inf_{t\geq 0} \tilde{X}_t < 0) \tag{4.13}$$

be a ruin probability of the corresponding clustered process $\{\tilde{X}_t\}_{t\geq 0}$.

Lemma 4.1. *We have,*

$$\psi(u) \leq \tilde{\psi}(u).$$

Proof. This follows immediately by inequality (4.11). □

To derive a lower bound for the ruin probability, we follow the ideas of Albrecher and Asmussen (2006), who consider a general shot-noise model, and find a suitable constant l_1 such that $\psi(u - l_1) \geq \tilde{\psi}(u)$. For this, we first introduce some additional notation. We write L_i for the length of cluster i, i.e.

$$L_i = \inf \left\{ t \geq 0 \,\middle|\, K_t^{(i)} = K_\infty^{(i)} = \kappa_i \right\}.$$

By the i.i.d. structure of the clusters, we have that the sequence $\{L_i\}_{i\in\mathbb{N}}$ is also i.i.d. following some cumulative distribution function F_L. Further, we observe that $L_i = \sum_{j=1}^{\kappa_i} T_{ij}$, where $\{T_{ij}\}_{j\in\{1,\ldots\kappa_i\}}$ denotes the set of inter-jump times of the i-th cluster. Let

$$\tilde{\tau} := \inf\{t \geq 0 : \tilde{X}_t < 0\}$$

denote the time of ruin of the clustered process $\{\tilde{X}_t\}_{t\geq 0}$ and $\tilde{L}$ the time from $\tilde{\tau}$ until all clusters which appeared up to time $\tilde{\tau}$ finished, i.e.

$$\tilde{L} := \sup \left\{ \bar{T}_i + \sum_{j=1}^{\kappa_i} T_{ij} : \bar{T}_i \leq \tilde{\tau} \right\} - \tilde{\tau}.$$

In other words, this is the minimal random time such that for all i with $\bar{T}_i \leq \tilde{\tau}$ we have $K^{(i)}_{\tilde{\tau}+\tilde{L}} = K^{(i)}_\infty = \kappa_i$.
Then, we have whenever $\tilde{\tau}$ and $\tilde{L}$ are finite that

$$X_{\tilde{\tau}+\tilde{L}} = u + c(\tilde{\tau} + \tilde{L}) - \sum_{i=1}^{\bar{N}_{\tilde{\tau}+\tilde{L}}} \sum_{j=0}^{K^{(i)}_{\tilde{\tau}+\tilde{L}}} U_{ij} \tag{4.14}$$

$$= u + c\tilde{\tau} - \sum_{i=1}^{\bar{N}_{\tilde{\tau}}} \sum_{j=0}^{\kappa_i} U_{ij} + c\tilde{L} - \sum_{i=\bar{N}_{\tilde{\tau}}+1}^{\bar{N}_{\tilde{\tau}+\tilde{L}}} \sum_{j=0}^{K^{(i)}_{\tilde{\tau}+\tilde{L}}} U_{ij} \tag{4.15}$$

$$\leq \tilde{X}_{\tilde{\tau}} + c\tilde{L}. \tag{4.16}$$

Consequently,

$$X_{\tilde{\tau}+\tilde{L}} - c\tilde{L} \leq \tilde{X}_{\tilde{\tau}} < 0.$$

Hence, if we reduce the initial capital of our process by $c\tilde{L}$, then ruin of the clustered process also causes ruin of the original surplus. The main problem is that $\tilde{L}$ is random, could be infinite and is not measurable with respect to the filtration of the original surplus process $\{X_t\}_{t\geq 0}$ nor with respect to the filtration of the clustered process $\{\tilde{X}_t\}_{t\geq 0}$. To bypass these problems, we want to identify constants l_1 and C such that $\psi(u - cl_1) \geq C\tilde{\psi}(u)$.

Lemma 4.2. *Let $t > 0$ be deterministic such that $\mathbb{P}[L_1 \leq t] > 0$. Then there exists a constant $\zeta_t \in (0, 1]$ such that $\psi(u - ct) \geq \zeta_t \tilde{\psi}(u)$.*

Proof. At first, we set for convenience $\tilde{L}(\omega) = \infty$ whenever $\tilde{\tau}(\omega) = \infty$. Then, we observe that for all ω such that $\tilde{L}(\omega) < t$ we have that $\tilde{\tau}_u(\omega) < +\infty$ and $\tau_{u-ct}(\omega) < +\infty$. Here, τ_u denotes the time of ruin with initial capital u of the surplus process $\{X_t\}_{t\geq 0}$ and $\tilde{\tau}_u$ the corresponding ruin time of $\{\tilde{X}_t\}_{t\geq 0}$ with starting point u. The implication that $\tilde{\tau}_u$ has to be finite is clear since $\tilde{L} = \infty$ if $\tilde{\tau} = \infty$. On the other hand, we have by inequality (4.14) that in this case

$$\tau_{u-ct}(\omega) \leq \tilde{\tau}_u(\omega) + \tilde{L}(\omega) \leq \tilde{\tau}_u(\omega) + t < +\infty.$$

Using these implications, we get for fixed t that

$$\begin{aligned}\psi(u - ct) \geq \mathbb{P}\left[\tau_{u-ct} < +\infty, \tilde{L} \leq t\right] &= \mathbb{P}\left[\tilde{L} \leq t\right] \\ &= \sum_{n=1}^{\infty} \mathbb{P}\left[\tilde{L} \leq t \,\middle|\, \tilde{\tau}_u = \bar{T}_n\right] \mathbb{P}\left[\tilde{\tau}_u = \bar{T}_n\right].\end{aligned}$$

If we can now bound $\mathbb{P}\left[\tilde{L} \leq t \,\middle|\, \tilde{\tau}_u = \bar{T}_n\right]$ from below by a positive constant ζ_t, this inequality would imply that

$$\psi(u - ct) \geq \zeta_t \sum_{n=1}^{\infty} \mathbb{P}\left[\tilde{\tau}_u = \bar{T}_n\right] = \zeta_t \tilde{\psi}(u),$$

since ruin for the clustered process can only happen at jump times $\bar{T}_n$ of the Poisson process $\{\bar{N}_t\}_{t\geq 0}$.

To do this, we determine the distribution of a sequence of auxiliary variables $\tilde{L}_n := \max\left\{\bar{T}_i - \bar{T}_n + L_i : i \leq n\right\}$, i.e. the random variable $\tilde{L}$ conditioned on $\tilde{\tau} = \bar{T}_n$. Then, we have that

$$\mathbb{P}\left[\tilde{L}_n \leq t\right] = \mathbb{P}\left[\bar{T}_1 - \bar{T}_n + L_1 \leq t, \ldots, \bar{T}_{n-1} - \bar{T}_n + L_{n-1} \leq t, L_n \leq t\right].$$

The random variable L_n is independent of the random variables L_i for $i < n$ and of the Poisson process $\bar{N}_t$. Further, by the lack of memory of exponential law, we have that $E_n := \bar{T}_n - \bar{T}_{n-1}$ is exponentially distributed with parameter a and independent of the information up to $\bar{T}_{n-1}$. This yields

$$\mathbb{P}\left[\tilde{L}_n \leq t\right] = F_L(t)\mathbb{P}\left[\bar{T}_1 - \bar{T}_{n-1} \leq t + E_n, \ldots, L_{n-1} \leq t + E_n\right].$$

Conditional on E_n, we have the same structure as before and now L_{n-1} is independent of all other random variables and $E_{n-1} = \bar{T}_{n-1} - \bar{T}_{n-2}$ is again exponentially distributed. Consequently,

$$\mathbb{P}\left[\tilde{L}_n \leq t\right] = F_L(t) \times \mathbb{E}\Big[F_L(t + E_n) \\ \mathbb{P}\left[\bar{T}_1 - \bar{T}_{n-2} \leq t + E_n + E_{n-1}, \ldots, L_{n-2} \leq t + E_n + E_{n-1} \,|\, E_n\right]\Big].$$

Using the fact that $F_L(x) \leq 1$ for all $x > 0$, and continuing above procedure we get that

$$\mathbb{P}\left[\tilde{L}_n \leq t\right] = \mathbb{E}\left[\prod_{i=0}^{n-1} F_L\left(t + \tilde{T}_i\right)\right] \geq \mathbb{E}\left[\prod_{i=0}^{\infty} F_L\left(t + \tilde{T}_i\right)\right],$$

where

$$\tilde{T}_k := \sum_{j=1}^{k} E_{n+1-j}.$$

This property holds for all $n \in \mathbb{N}$, and even though the jump times $\tilde{T}_i$ depend on n, the expectations coincide since they have the same distribution and the same dependence structure for all n due to the stationarity of the Poisson process. Now, we still have to show that this expectation is strictly positive.

To do so, we use that an infinite product $\prod_{n=1}^{\infty}(1+a_n)$ converges absolutely to a nonzero real number if the series $\sum_{n=1}^{\infty}|a_n|$ converges and $1 + a_n > 0$ for all n. If we write $\bar{F}_L(x)$ for the tail $1 - F_L(x)$ we get for fixed $\omega \in \Omega$ that

$$\prod_{i=0}^{\infty} F_L\left(t + \tilde{T}_i(\omega)\right) = \prod_{i=0}^{\infty}\left(1 - \bar{F}_L\left(t + \tilde{T}_i(\omega)\right)\right).$$

Since for all i, we have that $\bar{F}_L(t + \tilde{T}_i) \leq \bar{F}_L(t) < 1$, we want to show that the series $\sum_{i=0}^{\infty} \bar{F}_L\left(t + \tilde{T}_i(\omega)\right)$ converges for almost all ω, to get that the product $\prod_{i=0}^{\infty} F_L\left(t + \tilde{T}_i\right)$ converges almost surely to a random variable $C \in (0, 1]$, which would give us positiveness of the corresponding expectation.
By the strong law of large numbers, we have that $\frac{1}{n}\tilde{T}_n \to a$ almost surely. Let now ω be such that this convergence holds. Then, there exists a N_ω such that $\bar{T}_n \geq n\frac{a}{2}$ for all

$n \geq N_\omega$. By this, and the monotone decreasing behaviour of $\bar{F}_L$, we get that

$$\sum_{i=0}^{\infty} \bar{F}_L\left(t + \tilde{T}_i(\omega)\right) \leq N_\omega + 1 + \sum_{i=\leq N_\omega+1}^{\infty} \bar{F}_L\left(t + \tilde{T}_i(\omega)\right)$$
$$\leq N_\omega + 1 + \sum_{i=N_\omega+1}^{\infty} \bar{F}_L\left(t + n\frac{a}{2}\right).$$

The remaining series converges if and only if the corresponding integral

$$\int_{N_\omega+1}^{\infty} \bar{F}_L\left(t + x\frac{a}{2}\right)\,\mathrm{d}x,$$

converges. For this, we have that

$$\int_{N_\omega+1}^{\infty} \bar{F}_L\left(t + x\frac{a}{2}\right)\,\mathrm{d}x \quad\leq\quad \int_{N_\omega+1}^{\infty} \bar{F}_L\left(x\frac{a}{2}\right)\,\mathrm{d}x \quad\leq\quad \frac{2}{a}\int_0^{\infty} \bar{F}_L(y)\,\mathrm{d}y \quad=\quad \frac{2}{a}\mathbb{E}\left[L_1\right].$$

Here, L_1 denotes the length of the first cluster. By the proof of Lemma 1 in Møller and Rasmussen (2005), we have that

$$\mathbb{E}[L_1] \leq \frac{1}{1-\mu}\mathbb{E}\left[\int_0^{\infty} t\,h(t,Y)\,\mathrm{d}t\right] < +\infty.$$

Consequently, we have that the series $\sum_{i=0}^{\infty} \bar{F}_L\left(t + \tilde{T}_i\right)$ converges a.s., which implies that the product $\prod_{i=0}^{\infty} F_L\left(t + \tilde{T}_i\right) \in (0,1]$ almost surely. This gives us finally that $\zeta_t := \mathbb{E}\left[\prod_{i=0}^{\infty} F_L\left(t + \tilde{T}_i\right)\right] \in (0,1]$.

□

From Lemmas 4.1 and 4.2 we can conclude the general case with the following theorem.

Theorem 4.3. *Let l_1 be deterministic such that $\mathbb{P}\left[L_1 \leq l_1\right] > 0$. Then, there exists a constant $\zeta_{l_1} \in (0,1]$ such that the ruin probability in the general Hawkes model satisfies*

$$\tilde{\psi}(u) \geq \psi(u) \geq \zeta_{l_1}\tilde{\psi}(u + l_1 c),$$

where $\tilde{\psi}(u)$ denotes the ruin probability (4.13) of the clustered process $\{\tilde{X}_t\}_{t\geq 0}$, i.e. the ruin probability of a Cramér-Lundberg process.

4.2.3 Cramér-Lundberg bounds and heavy-tailed asymptotics

Interested in the asymptotic behaviour of the ruin probability as the initial capital tends to infinity, we will see that this behaviour depends highly on the behaviour of the distribution of the generic claim size U of the risk process (4.3). To understand this behaviour we recall that the generic claim size of the Cramér-Lundberg process (4.9) is given in (4.10), that is, $\tilde{U} = \sum_{k=1}^{\kappa} U_k$ for a generic cluster size κ. We prove the following basic fact.

Lemma 4.4. *The generic cluster size κ is light-tailed, i.e. there exists $\theta > 0$ such that $\mathbb{E}\left[e^{\theta\kappa}\right] < +\infty$.*

Proof. Observe that κ has the same law as a total progeny in a Galton–Watson branching process with Poisson offspring distribution whose mean is μ. From Dwass (1969) it follows that

$$\mathbb{P}(\kappa = n) = \frac{1}{n}\mathbb{P}(S_n = -1),$$

where S_n is a random walk with i.i.d. increments $X_i \overset{\mathcal{D}}{=} \text{Poi}(\mu) - 1$ for a Poissonian random variable $\text{Poi}(\mu)$ with the parameter μ. Hence $\mathbb{P}(\kappa = n) \le \mathbb{P}(-\frac{S_n}{n} \le 0)$ and this probability decays exponentially to zero by the Cramér-Chernoff Theorem. This completes the proof. □

We are ready to state the first corollary.

Corollary 4.5. *Assume that the claim events $\{U_i\}_{i\in\mathbb{N}}$ of the risk process* (4.3) *are light-tailed and that there exists $R > 0$ such that*

$$a(\mathbb{E}\left[\mathbb{E}\left[e^{RU}\right]^{\kappa}\right] - 1) = cR.$$

Then, there exist positive constants C_- and C_+ such that

$$C_- e^{-Ru} \le \psi(u) \le C_+ e^{-Ru}.$$

Proof. By Lemma 4.4 it follows that the generic $\tilde{U}$ defined in (4.10) is light-tailed and that

$$a(\mathbb{E}\left[e^{R\tilde{U}}\right] - 1) = cR. \tag{4.17}$$

Under our assumptions, we have by Theorem 5.4.1 on p. 170 of Rolski et al. (1999) that there exist constants $\tilde{C}_-$, C_+ such that the ruin probability of the clustered surplus process $\tilde{\psi}(u)$ satisfies

$$\tilde{C}_- e^{-Ru} \le \tilde{\psi}(u) \le C_+ e^{-Ru},$$

for all $u \ge 0$. Having this, we get that

$$C_+ e^{-Ru} \ge \tilde{\psi}(u) \ge \psi(u) \ge \tilde{\psi}(u + cl_1) \ge \tilde{C}_- e^{-Rcl_1} e^{-Ru} =: C_- e^{-Ru}.$$

□

Remark 4.1. This behaviour implies convergence of the logarithm of the ruin probabilities in the Hawkes model, as it was already derived by Karabash and Zhu (2015).

We identify the asymptotic behaviour of the ruin probability also in the case when the generic claim size is heavy-tailed. We introduce now the appropriate class of distributions that we will work with. Let $F_{\tilde{U}}$ be the distribution of $\tilde{U}$ defined in (4.10). We denote

$$F^s_{\tilde{U}}(u) := \frac{1}{\mathbb{E}\left[\tilde{U}_1\right]} \int_0^u (1 - F_{\tilde{U}}(y))\,\mathrm{d}y.$$

We say that a distribution F on $\mathbb{R}_+$ with unbounded support is strongly subexponential (writing $F \in \mathcal{S}^*$) if

$$\lim_{u\to+\infty} \frac{\int_0^u (1 - F(u-y))(1 - F(y))\mathrm{d}y}{2\int_0^\infty (1 - F(y))\mathrm{d}y} = 1.$$

It is known (see Klüppelberg (1988)) that any distribution from the class $\mathcal{S}^*$ is subexponential (writing $F \in \mathcal{S}$), that is, that

$$\lim_{u\to+\infty} \frac{1 - F^{*2}(u)}{2(1 - F(u))} = 1.$$

Classical examples of distributions from the class $\mathcal{S}^*$ are Pareto, log-normal and Weibull distributions. The latter with a shape parameter from $(0, 1)$.

Corollary 4.6. *Assume that the integrated tail distribution $F^s_{\tilde{U}}$ of the generic clustered claim size $\tilde{U}$ is subexponential, that is, that $F^s_{\tilde{U}} \in \mathcal{S}$. Then,*

$$\lim_{u\to\infty} \frac{\psi(u)}{1 - F^s_{\tilde{U}}(u)} = \frac{\rho}{1-\rho},$$

where

$$\rho := \frac{a\mathbb{E}\left[\tilde{U}_1\right]}{c} < 1$$

by (4.12).

Proof. Since $F^s_{\tilde{U}}$ is subexponential, the ruin probability of the Cramér-Lundberg model satisfies by Theorem 5.4.3 on p. 175 of Rolski et al. (1999), which gives

$$\lim_{u\to\infty} \frac{\tilde{\psi}(u)}{1 - F^s_{\tilde{U}}(u)} = \frac{\rho}{1-\rho} = \lim_{u\to\infty} \frac{\tilde{\psi}(u + cl_1)}{1 - F^s_{\tilde{U}}(u + cl_1)}.$$

Further, we have that $\lim_{u\to\infty} \frac{1-F^s_{\tilde{U}}(u+y)}{1-F^s_{\tilde{U}}(u)} = 1$ for all $y \in \mathbb{R}$ since any subexponential distribution is also long-tailed (see e.g. Lemma 3.2 on p. 40 of Foss et al. (2011)).

Consequently,

$$\frac{\rho}{1-\rho} = \lim_{u\to\infty} \frac{\tilde{\psi}(u)}{1-F^s_{\tilde{U}}(u)} \geq \limsup_{u\to\infty} \frac{\psi(u)}{1-F^s_{\tilde{U}}(u)} \geq \liminf_{u\to\infty} \frac{\psi(u)}{1-F^s_{\tilde{U}}(u)}$$
$$\geq \lim_{u\to\infty} \frac{1-F^s_{\tilde{U}}(u+cl_1)}{1-F^s_{\tilde{U}}(u)} \frac{\tilde{\psi}(u+cl_1)}{1-F^s_{\tilde{U}}(u+cl_1)} = \frac{\rho}{1-\rho}$$

which completes the proof. □

It is more valuable to derive the asymptotics of the ruin probability $\psi(u)$ in terms of the original distribution of the claim sizes F_U, under the assumption that F_U is strongly subexponential (hence heavy-tailed by Lemma 3.2 on p. 40 of Foss et al. (2011)).

Corollary 4.7. *If $F_U \in \mathcal{S}^*$ then*

$$\lim_{u\to\infty} \frac{\psi(u)}{1-F^s_U(u)} = \frac{\rho}{1-\rho}. \tag{4.18}$$

Proof. Recall that $\tilde{U} = \sum_{k=1}^{\kappa} U_k$ for a generic cluster size κ and by Lemma 4.4 κ is light-tailed. Now the statement follows from Theorem 1 of Denisov et al. (2010), (4.8) and Corollary 4.6 since

$$\lim_{u\to\infty} \frac{1-F_{\tilde{U}}(u)}{1-F_U(u)} = \frac{1}{1-\mu},$$

and hence

$$\lim_{u\to\infty} \frac{1-F^s_{\tilde{U}}(u)}{1-F^s_U(u)} = 1.$$

□

Remark 4.2. We recover Proposition 13 of Zhu (2013). Our proof is different though not requiring tedious checking of Assumptions 1 of Zhu (2013).

Although the results of Theorem 4.3 are similar to the findings of Albrecher and Asmussen (2006), it faces two drawbacks. The first problem is that the asymptotic behaviour depends highly on the distribution of the sum of all claim sizes occurring in a single cluster, something that might be hard to handle. The second weakness is that we were not able to show if there are conditions under which $\lim_{u\to\infty} e^{Ru}\psi(u)$ converges. To achieve this, we restrict ourselves to the Markovian version of the Hawkes process in the following.

4.3 Cramér-Lundberg asymptotics for Markovian Hawkes model

4.3.1 The Markovian Hawkes model

The intensity of a Hawkes process is generally not Markovian. To resolve this, we have to choose the specific decay function $h(t, y) = e^{-\beta t} y$, for some positive decay parameter β. Further, for some positive constants $\lambda > a > 0$, we define the Markovian (marked) Hawkes process $\{N_t\}_{t \geq 0}$ by its intensity process

$$\lambda_t := a + (\lambda - a)e^{-\beta t} + \sum_{i=1}^{N_t} Y_i e^{-\beta(t-T_i)}. \tag{4.19}$$

Here, the random variables $\{Y_i\}_{i \in \mathbb{N}}$ are assumed to be i.i.d. copies of a positive random variable Y with cumulative distribution function F_Y and finite expectation. In contrast to the previous part, we allow for different initial values λ instead of restricting ourselves to the case $\lambda_0 = a$. This process is well-defined if

$$\int_0^\infty \mathbb{E}[h(t, Y)] \,\mathrm{d}t = \mathbb{E}[Y] \int_0^\infty e^{-\beta t} \,\mathrm{d}t = \frac{\mathbb{E}[Y]}{\beta} < 1,$$

which gives us the restriction that

$$\beta > \mathbb{E}[Y], \tag{4.20}$$

whereas the integrability condition

$$\int_0^\infty t \,\mathbb{E}[h(t, Y)] \,\mathrm{d}t < +\infty,$$

is always satisfied. This process is called Markovian, since the intensity process $\{\lambda_t\}_{t \geq 0}$ is a piecewise deterministic Markov process (PDMP) with extended generator

$$\mathcal{A}^\lambda f(\lambda) = \beta(a - \lambda)\frac{\partial}{\partial \lambda} f(x, \lambda, t) + \lambda \int_0^\infty f(\lambda + y) \,F_Y(\mathrm{d}y) - \lambda f(x, \lambda, t);$$

see Davis (1984) for more details on theory of PDMPs. We recall that for any Markov process $\{Z_t\}_{t \geq 0}$ we say that $\mathcal{A}$ is its extended generator and $D(\mathcal{A})$ is a domain of this generator if $f(Z_t) - f(Z_0) - \int_0^t g(Z_s) \,\mathrm{d}s$ is a zero mean local martingale with respect to its natural filtration for $f \in \mathcal{D}(\mathcal{A})$ and some function g. We then write $g = \mathcal{A}f$.

Further, the multivariate process $\{(X_t, \lambda_t, t)\}_{t \geq 0}$ is also a piecewise deterministic Markov process with full generator

$$\begin{aligned}\mathcal{A}f(x, \lambda, t) = c\frac{\partial}{\partial x} f(x, \lambda, t) + \beta(a - \lambda)\frac{\partial}{\partial \lambda} f(x, \lambda, t) + \frac{\partial}{\partial t} f(x, \lambda, t) \\ + \lambda \int_0^\infty \int_0^\infty f(x - u, \lambda + y, t) \,F_U(\mathrm{d}u) \,F_Y(\mathrm{d}y) - \lambda f(x, \lambda, t).\end{aligned} \tag{4.21}$$

Since we have now two different initial values, one for the surplus process and one for the intensity, we write $\mathbb{P}_{(u,\lambda)}$ for the measure $\mathbb{P}$ under the condition that $X_0 = u$ and $\lambda_0 = \lambda$ for $u \geq 0$ and $\lambda > a$, and $\mathbb{E}_{(u,\lambda)}[\cdot]$ for the corresponding expectation. If a stochastic object Z is independent of the initial values, we will omit them and write $\mathbb{E}[Z]$ instead of $\mathbb{E}_{(u,\lambda)}[Z]$. Having this, we have by Rolski et al. (1999, p. 449), that the domain $\mathcal{D}(\mathcal{A})$ of this generator $\mathcal{A}$ consists of all functions $f : \mathbb{R}^3 \to \mathbb{R}$ such that the mapping $t \to f(x + ct, (\lambda - a)e^{-\beta t} + a, s + t)$ is absolutely continuous for almost all (x, λ, s) and

$$\mathbb{E}_{(u,\lambda)}\left[\sum_{k=1}^{N_t} |f(X_{T_k}, \lambda_{T_k}, T_k) - f(X_{T_k-}, \lambda_{T_k-}, T_k)|\right] < +\infty,$$

for all $t \geq 0$.
We start by identifying a suitable net profit condition, that is, the condition under which (4.2) holds. In fact, we will identify the limiting value $\lim_{t\to\infty} \frac{\mathbb{E}_{(u,\lambda)}[X_t]}{t}$ and assume that this limit is strictly positive which gives (4.2).

Lemma 4.8. *The surplus process satisfies* $\lim_{t\to\infty} \frac{\mathbb{E}_{(u,\lambda)}[X_t]}{t} = c - \frac{a\beta\mathbb{E}[U]}{\beta-\mathbb{E}[Y]}$.

Proof. The function $f(x, \lambda, t) = x$ is in the domain of the generator and satisfies $\mathcal{A}f(x, \lambda, t) = c - \lambda\mathbb{E}[U]$. Therefore, we have that $\mathbb{E}_{(u,\lambda)}[X_t] = u + ct - \mathbb{E}[U] \int_0^t \mathbb{E}_\lambda[\lambda_s]\, \mathrm{d}s$. By Cui et al. (2020), we have that $\mathbb{E}_\lambda[\lambda_s] = \frac{\beta a}{\beta-\mathbb{E}[Y]} + \left(\lambda - \frac{\beta a}{\beta-\mathbb{E}[Y]}\right) e^{-s(\beta-\mathbb{E}[Y])}$. Consequently, $\int_0^t \mathbb{E}_\lambda[\lambda_s]\, \mathrm{d}s = \frac{\beta a}{\beta-\mathbb{E}[Y]} t + o(t)$. Using this, we get

$$\lim_{t\to\infty} \frac{\mathbb{E}_{(u,\lambda)}[X_t]}{t} = c - \mathbb{E}[U]\frac{\beta a}{\beta - \mathbb{E}[Y]}$$

which completes the proof. □

By this result, we propose the following net profit condition.

Assumption 4.1. *From now on we assume that the net profit condition*

$$c > \frac{a\beta\mathbb{E}[U]}{\beta - \mathbb{E}[Y]}$$

holds. Further, we assume that there exists some positive s_Y *such that the moment-generating function*

$$M_Y(s) := \mathbb{E}\left[e^{sY}\right] < +\infty\, \forall s < s_Y, \quad \textit{and} \quad \lim_{s\to s_Y} M_Y(s) = +\infty.$$

4.3.2 Steps of the proof of Cramér-Lundberg asymptotics

Our main goal of this section is to prove the so-called Cramér Lundberg asymptotics for the ruin probability if the intensity of the arrival process is given by(4.19), i.e. that there exists an adjustment coefficient R solving Lundberg equation formulated in (4.28) and constant $C^\lambda > 0$, depending on λ, such that

$$\lim_{u\to\infty} \psi(u,\lambda)e^{Ru} = C^\lambda;$$

see Theorem 4.33.
To prove this statement we split the whole proof into the following steps:

1. In the next subsection we prove in Theorem 4.17 that the intensity process $\{\lambda_t\}_{t\geq 0}$ is positive Harris recurrent, and, in Theorem 4.21, that the corresponding recurrence times $\{\phi_i\}_{i\in\mathbb{N}}$ are light-tailed.

2. Later, in Definition 4.6, we introduce the functions $\alpha(r)$ and $\theta(r)$ as special solutions of equations (4.25)-(4.26).

3. The adjustment coefficient $R > 0$ is defined as a maximal solution of the Lundberg equation
$$\theta(R) = 0.$$

4. We introduce in Definition 4.7 the new exponential measure
$$\left.\frac{\mathrm{d}\mathbb{Q}^{(R)}}{\mathrm{d}\mathbb{P}}\right|_{\mathcal{F}_t} = M_t := \exp\left(Ru + \alpha(R)\lambda\right)\exp(-RX_t - \alpha(R)\lambda_t)\,.$$

5. In Lemma 4.28 we prove that under the new measure $\mathbb{Q}^{(R)}$ ruin occurs almost surely and the intensity $\{\lambda_t\}_{t\geq 0}$ remains positive Harris recurrent (see Lemma 4.3.4).

6. We introduce the distribution
$$B(\mathrm{d}x) := e^{Rx}\tilde{B}(\mathrm{d}x) \quad \text{for} \quad \tilde{B}(x) := e^{Rx}\mathbb{P}_{(u,\lambda)}\left[\phi_1 < +\infty, u - X_{\phi_1} \leq x\right]$$
and
$$p(u,x) := \mathbb{P}_{(u,\lambda)}\left[\tau \leq \phi_1, |\phi_1 < +\infty, X_{\phi_1} = u - x\right],$$
where τ is the time of ruin defined in (4.1) and ϕ_1 the first recurrence epoch of $\{\lambda_t\}_{t\geq 0}$ to λ such that the process $\{X_t\}_{t\geq 0}$ is getting below the initial starting position u. Now, we can prove that
$$Z(u) = \psi(u)e^{Ru}$$
satisfies the following renewal equation (see (4.32))
$$Z(u) = \int_0^u Z(u-x)(1-p(u,x))B(\mathrm{d}x) + z(u),$$
where
$$z(u) = e^{Ru}\mathbb{P}_{(u,\lambda)}\left[\tau \leq \phi_1, \tau < +\infty\right].$$

7. To demonstrate that $\lim_{u\to+\infty} Z(u)$ exists and is finite, we use Theorem 2 of Schmidli (1997) and prove that $z(u)$ and $\int_0^u p(u,x)B(\mathrm{d}x)$ are directly Riemann integrable.

We start from the analysis of the behaviour of the Markovian intensity process $\{\lambda_t\}_{t\geq 0}$.

4.3.3 Harris recurrence of the intensity process

In this section, we investigate the behaviour of the intensity of the Markovian marked Hawkes process. Our goal is to show that the intensity process is positive Harris recurrent (see Theorem 4.17) and that the corresponding recurrence times are light-tailed (see Theorem 4.21), both properties are needed to determine the asymptotic behaviour of the ruin probability.
To define these properties properly, we consider a right-continuous, time-homogeneous, strong Markov process $\{Z_t\}_{t\geq 0}$ on $(\mathcal{E}, \mathcal{B}(\mathcal{E}))$. Here, $\mathcal{E}$ denotes a locally compact, separable metric space and $\mathcal{B}(\mathcal{E})$ its Borel σ-algebra.

Definition 4.1. The process $\{Z_t\}_{t\geq 0}$ is called Harris recurrent if there exists a σ-finite measure φ on $\mathcal{B}(\mathcal{E})$ such that $\varphi(B) > 0 \Rightarrow \mathbb{P}_z[\eta_B = \infty] = 1$ for all initial values z and $B \in \mathcal{B}(\mathcal{E})$, where

$$\eta_B := \int_0^\infty I_{\{Z_t \in B\}} \mathrm{d}t \tag{4.22}$$

denotes the occupation measure of the process $\{Z_t\}_{t\geq 0}$. It is called positive Harris recurrent if it is Harris recurrent with finite invariant measure π.

To show that our intensity process satisfies this property, we need the following definitions of a continuous component and a T-process as in Section 3.2 of Meyn and Tweedie (1993, pp. 495–496).

Definition 4.2. Let $\{Z_t\}_{t\geq 0}$ be our strong Markov process and $\sigma_1, \sigma_2, \ldots$ an i.i.d. sequence of positive random variables with distribution F and independent of $\{Z_t\}_{t\geq 0}$. Then, we define the embedded Markov chain $Y_n := Z_{\sigma_1+\cdots+\sigma_n}$ with one-step transition probability $K_F(z, A) := \int_0^\infty \mathbb{P}_z[Z_t \in A]\, F(\mathrm{d}t)$ for all $A \in \mathcal{B}(\mathcal{E})$. A kernel $T : (\mathcal{E}, \mathcal{B}(\mathcal{E})) \to \mathbb{R}_+$ is called a continuous component of K_F if $K_F(x, A) \geq T(x, A)$ for all x and A, and for fixed $A \in \mathcal{B}(\mathcal{E})$, the function $T(\cdot, A)$ is lower semi-continuous. We say T is non-trivial if, for all $x \in \mathcal{E}$, we have that $T(x, \mathcal{E}) > 0$.

A special case of such an embedded Markov chain is the resolvent chain, whose transition kernel is given by $R_\gamma(x, A) := \int_0^\infty \mathbb{P}_x[Z_t \in A]\, e^{-\gamma t} \mathrm{d}t$, i.e. where $\sigma_1 \sim \mathrm{Exp}(\gamma)$ for the exponential random variable $\mathrm{Exp}(\gamma)$ with the parameter $\gamma > 0$.

Definition 4.3. The process $\{Z_t\}_{t\geq 0}$ is called T-process if there is a probability distribution F such that K_F admits a non-trivial continuous component T.

Definition 4.4. Recall that η_B defined in (4.22) is the occupation measure. Let φ be a σ-finite measure. If $\varphi(B) > 0 \Rightarrow \mathbb{E}_z[\eta_B] > 0$ for all initial values z and all $B \in \mathcal{B}(\mathcal{E})$, then $\{Z_t\}_{t\geq 0}$ is called φ-irreducible.

Definition 4.5. A process $\{Z_t\}_{t\geq 0}$ is called bounded in probability on average if for every initial value z and $\varepsilon > 0$ there is a compact set K such that

$$\liminf_{t\to\infty} \frac{1}{t} \int_0^t \mathbb{P}_z \left[Z_s \in K\right] \mathrm{d}s \geq 1 - \varepsilon.$$

These three properties are related to positive Harris recurrence by Theorem 3.2 of Meyn and Tweedie (1993), which states the following.

Theorem 4.9. *Suppose that $\{Z_t\}_{t\geq 0}$ is a φ-irreducible T-process. Then $\{Z_t\}_{t\geq 0}$ is positive Harris recurrent if and only if it is bounded in probability on average.*

Now, we want to show that our process $\{\lambda_t\}_{t\geq 0}$ satisfies all conditions of Theorem 4.9. It is a time-homogeneous strong Markov process with right-continuous paths defined on $(\mathbb{R}, \mathcal{B}(\mathbb{R}))$. The space $\mathbb{R}$ is a locally compact and separable metric space. The next point is to show that the process is a T-process.

Lemma 4.10. *The intensity process $\{\lambda_t\}_{t\geq 0}$ is a T-process.*

Proof. For this, we show that the resolvent kernel $R_1(\lambda, A) = \int_0^\infty \mathbb{P}_\lambda \left[\lambda_t \in A\right] e^{-t}\mathrm{d}t$ is continuous in λ, for every fixed $A \in \mathcal{B}(\mathbb{R})$. For this, observe that by Tonelli's theorem, we can interchange expectation and integration to get that

$$R_1(\lambda, A) = \mathbb{E}_\lambda \left[\int_0^\infty I_A(\lambda_t) e^{-t} \mathrm{d}t\right] =: V(\lambda).$$

Further, by Theorem 31.9 of Davis (1993) we know that

$$V(\lambda) = \tilde{V}(\lambda) := \mathbb{E}_\lambda \left[\int_0^\infty I_A(\tilde{\lambda}_t) \mathrm{d}t\right],$$

where $\{\tilde{\lambda}_t\}_{t\geq 0}$ denotes the process $\{\lambda_t\}_{t\geq 0}$ but killed with constant rate 1. The function $\tilde{V}(\lambda)$ is bounded by 1, and the function $l(x) := I_A(x)$ is measurable and integrable. Therefore, we have by Theorem 32.2 of Davis (1993), that $\tilde{V}(\lambda)$ is absolutely continuous. The kernel R_1 is non-trivial since $R_1(\lambda, \mathbb{R}_+) = 1$ for all $\lambda \in \mathbb{R}_+$. Therefore, R_1 serves as a non-trivial component. □

To show that the intensity process is φ-irreducible, we have to identify a suitable σ-finite measure φ. For this, we will show in the first step that the Markovian Hawkes intensity converges in distribution to a stationary probability distribution ν. To do so, we introduce the process $\mu_t = a + \sum_{i=1}^{N_t} Y_i e^{-\beta(t-T_i)}$, i.e. our Hawkes intensity with initial condition $\lambda_0 = a$. By Brémaud et al. (2002, p. 133), the distribution of μ_t converges weakly against a stationary distribution as $t \to \infty$. If we can show that, independent of the initial intensity λ, the process $\{\lambda_t - \mu_t\}_{t\geq 0}$ converges in probability to 0. Then, we have by Slutsky's theorem that $\lambda_t = \lambda_t - \mu_t + \mu_t$ converges in distribution to ν too.

Lemma 4.11. *Let $\lambda > a$ be arbitrary but fixed. Then* $\lim_{t\to\infty} \mathbb{P}_\lambda \left[|\lambda_t - \mu_t| > \varepsilon\right] = 0$ *for all $\varepsilon > 0$.*

Proof. To show this, we want to use Markov's inequality. Therefore, we are interested in the behaviour of $\mathbb{E}[|\lambda_t - \mu_t|]$ as $t \to \infty$. For this, we take a look at the change of the intensity by the increase of the initial value by $\lambda - a$. For fixed λ, we can decompose the corresponding counting process $\{N_t\}_{t\geq 0}$ as

$$N_t = N_t^\mu + M_t,$$

where M_t counts all jumps due to the initial increase by $\lambda - a$ and N_t^μ is a marked Hawkes process with intensity μ_t.
In the time interval $(0, \infty)$, the initial increase by $\lambda - a$ will cause

$$Z := \text{Poi}\left((\lambda - a)\int_0^\infty e^{-\beta t}\mathrm{d}t\right) = \text{Poi}\left(\frac{\lambda - a}{\beta}\right)$$

jumps of the marked Hawkes process. We will call these jumps 'children'. Each child increases the intensity by a generic $\tilde{Y}$, hence triggers $\text{Poi}(\frac{\tilde{Y}}{\beta})$ additional jumps, i.e. 'grandchildren'. These grandchildren cause new jumps again, so we get a branching structure. We call the collection of all jumps caused by $\lambda - a$ offspring. Since $\mathbb{E}\left[\int_0^\infty Y e^{-\beta t}\mathrm{d}t\right] = \frac{\mathbb{E}[Y]}{\beta} < 1$, we have that the number of jumps in such a cluster is an integrable random variable, see Basrak et al. (2019). The random variable M_t corresponds to the number of offspring due to the increase by $\lambda - a$, which appeared up to time t and converges almost surely to the integrable random variable M_∞, which corresponds to the size of the cluster caused by the additional initial intensity.
By this, we have that

$$|\lambda_t - \mu_t| = (\lambda - a)e^{-\beta t} + \sum_{i=1}^{M_t} \tilde{Y}_i e^{-\beta(t - T_i^M)} \leq (\lambda - a)e^{-\beta t} + \sum_{i=1}^{M_\infty} \tilde{Y}_i,$$

where $\{T_i^M\}_{i \leq M_\infty}$ are the jump times of the counting process $\{M_t\}_{t\geq 0}$. The upper bound is integrable. Hence, by dominated convergence, we have that

$$\begin{aligned}\limsup_{t\to\infty} \mathbb{E}_\lambda\left[|\lambda_t - \mu_t|\right] &\leq \mathbb{E}_\lambda\left[\limsup_{t\to\infty}\,(\lambda - a)e^{-\beta t} + \sum_{i=1}^{M_\infty} \tilde{Y}_i e^{-\beta(t - T_i^M)}\right] \\ &= \mathbb{E}_\lambda\left[\sum_{i=1}^{M_\infty} \limsup_{t\to\infty} \tilde{Y}_i e^{-\beta(t - T_i^M)}\right] = 0.\end{aligned}$$

The statement follows from Markov's inequality. □

Theorem 4.12. *The Markovian Hawkes intensity converges in distribution to the stationary distribution ν.*

Proof. Let $\{\lambda_t\}_{t\geq 0}$ and $\{\mu_t\}_{t\geq 0}$ be as before. By the previous lemma we have that $\lambda_t - \mu_t \to 0$ in probability, and by Brémaud et al. (2002), we get that $\mu_t \to \nu$ in distribution. Slutsky's theorem gives us that

$$\lambda_t = \lambda_t - \mu_t + \mu_t \xrightarrow{\mathcal{D}} 0 + Z,$$

where $Z \sim \nu$. □

To show that the intensity process is ν-irreducible, we still need some smoothness of the stationary distribution ν.

Lemma 4.13. *The stationary distribution ν is absolutely continuous with respect to the Lebesgue measure.*

Proof. This is due to Proposition 1.9 of Löpker and Palmowski (2013). □

Let $B \in \mathcal{B}(\mathbb{R})$ and define $\nu_t(B) := \frac{1}{t}\int_0^t \mathbb{P}_\lambda\left[\lambda_s \in B\right] \mathrm{d}s$.

Lemma 4.14. *The family of measures $\{\nu_t\}_{t\geq 0}$ converges weakly to the stationary measure ν as $t \to \infty$.*

Proof. Since λ_t converges in distribution to the absolutely continuous measure ν, we have that for all open sets U that $\liminf_{t\to\infty} \mathbb{P}_\lambda\left[\lambda_t \in U\right] \geq \nu(U)$. Consequently, for all $\varepsilon > 0$ there is a $T > 0$ such that for all $t \geq T$, it holds that $\mathbb{P}_\lambda\left[\lambda_t \in U\right] \geq \nu(U) - \varepsilon$. Therefore,

$$\begin{aligned}
&\liminf_{t\to\infty} \frac{1}{t}\int_0^t \mathbb{P}_\lambda\left[\lambda_s \in B\right] \mathrm{d}s \\
&\quad\geq \liminf_{t\to\infty} \frac{1}{t}\left(\int_0^T \mathbb{P}_\lambda\left[\lambda_s \in B\right] \mathrm{d}s + (t-T)(\nu(U)-\varepsilon)\right) \\
&\quad= \nu(U) - \varepsilon.
\end{aligned}$$

If we let ε tend to 0, we see that $\liminf_{t\to\infty} \nu_t(B) \geq \nu(B)$ for all open sets B. By the Portmanteau theorem, this implies that $\{\nu_t\}_{t\geq 0}$ converges weakly to the stationary measure ν. □

Now we are ready to show that our process is ν-irreducible.

Lemma 4.15. *The Markovian intensity process $\{\lambda_t\}_{t\geq 0}$ is ν-irreducible.*

Proof. Let B be measurable with $\nu(B) > 0$ and ν_t as in Lemma 4.14. By the absolute continuity of the measure ν, we have that B is a continuity set of ν. Hence, $\lim_{t\to\infty} \nu_t(B) = \nu(B) > 0$. Consequently, we have that $\mathbb{E}_\lambda\left[\eta_b\right] = \int_0^\infty \mathbb{P}_\lambda\left[\lambda_s \in B\right] \mathrm{d}s = \infty > 0$. □

Lemma 4.16. *The Markovian intensity process $\{\lambda_t\}_{t\geq 0}$ is bounded in probability on average.*

Proof. Let $\varepsilon > 0$ be arbitrary and $K_\varepsilon \subset \mathbb{R}_+$ be compact such that $\nu(K) \geq 1 - \varepsilon$. Then we have by Lemma 4.13 and Lemma 4.14, that

$$\liminf_{t\to\infty} \frac{1}{t} \int_0^t \mathbb{P}_\lambda \left[\lambda_s \in K\right] \mathrm{d}s = \nu(K) \geq 1 - \varepsilon.$$

□

Theorem 4.17. *The Markovian intensity process $\{\lambda_t\}_{t\geq 0}$ is positive Harris recurrent.*

Proof. This follows directly from Theorem 4.9 and the previous lemmas. □

This gives us that our intensity process visits all sets with $\nu(B) > 0$ infinitely often, but we still have to check that the stationary distribution has support (a, ∞). First, we will show that the support is unbounded from above.

Lemma 4.18. *The support of ν is unbounded from above.*

Proof. Assume there is a finite bound $b > 0$ such that $\nu((a, b)) = 1$. We use that there is a stationary version of our intensity process, and we will denote it by $\lambda'_t = a + (\lambda'_0 - a)e^{-\beta t} + \sum_{i=1}^{N'_t} Y_i e^{-\beta(t-T'_i)}$, where $\lambda'_0 \sim \nu$. Choose $\delta > 0$ arbitrary. Then, we have

$$\begin{aligned} 0 = \nu((b,\infty)) = \mathbb{P}_\nu\left[\lambda'_\delta > b\right] &\geq \mathbb{P}_\nu\left[\lambda'_\delta > b, N'_\delta = 1\right] \\ &= \mathbb{P}_\nu\left[\lambda'_\delta > b \,\middle|\, N'_\delta = 1\right] \mathbb{P}_\nu\left[N'_\delta = 1\right] \geq \mathbb{P}\left[Y > be^{\delta\beta}\right] \mathbb{P}_\nu\left[N'_\delta = 1\right] > 0, \end{aligned}$$

which is a contradiction. Since the support of the shock events Y is unbounded, we consequently have that the support of the stationary distribution is unbounded. □

Lemma 4.19. *The support of ν is an open set of the form (b, ∞) for some $b \geq a$.*

Proof. We already know that the support is unbounded. Assume that the support of the stationary distribution is not an open interval. Since ν is absolutely continuous with respect to the Lebesgue measure, we have that there exists an interval $(c, d) \subset (a, \infty)$ such that $\nu((c, d)) = 0$, $\nu((a, c)) > 0$, and $\nu(d, \infty) > 0$. Now, we assume that this interval is maximal. In particular, we want that for all $\varepsilon > 0$ $\nu((d, d + \varepsilon)) > 0$. Let $-\frac{1}{\beta} \ln\left(\frac{c-a}{d-a+\varepsilon}\right) > \delta > -\frac{1}{\beta} \ln\left(\frac{d-a}{d-a+\varepsilon}\right)$ deterministic. Then we have that if $\lambda'_0 \in (d, d+\varepsilon)$ and no jump occurs between time 0 and δ, $\lambda'_\delta \in (c, d)$, which is a contradiction. Writing this down, we get that

$$\begin{aligned} \nu((c,d)) &= \mathbb{P}_\nu\left[\lambda'_\delta \in (c,d)\right] \geq \mathbb{P}_\nu\left[\lambda'_\delta \in (c,d), \lambda'_0 \in (d, d+\varepsilon)\right] \\ &= \mathbb{P}_\nu\left[\lambda'_\delta \in (c,d) \,\middle|\, \lambda'_0 \in (d, d+\varepsilon)\right] \mathbb{P}_\nu\left[\lambda'_0 \in (d, d+\varepsilon)\right] \\ &\geq \mathbb{P}_\nu\left[N'_\delta = 0\right] \nu((d, d+\varepsilon)) > 0. \end{aligned}$$

This is a contradiction. Hence, the support of ν is an interval. □

Theorem 4.20. *The support of ν is (a, ∞).*

Proof. This proof is similar to the proof of Lemma 4.19. Assume that the support of ν is not (a, ∞). Then there exists a $\varepsilon > 0$ such that the support is $(a + \varepsilon, \infty)$ and $\nu((a+\varepsilon, a+2\varepsilon)) > 0$. Let $\delta > -\frac{1}{\beta}\ln(\frac{1}{2})$. Then again, if $\lambda'_0 \in (a+\varepsilon, a+2\varepsilon)$ and $N'_\delta = 0$, which both happen with positive probability, then $\lambda'_\delta < a + \varepsilon$. This contradicts the assumption that $\nu((a, a+\varepsilon)) = 0$. $\square$

This gives us that our intensity process visits every open interval in (a, ∞) infinitely often. Since it decays only in a continuous way via its exponentially decaying drift, we even have that the process $\{\lambda_t\}_{t\geq 0}$ visits every single point $\lambda \in (a, \infty)$ infinitely often with probability 1.

Theorem 4.21. *Let $\lambda > a$ be arbitrary and S_1^λ the first positive time point such that $\lambda_{S_i^\lambda} = \lambda$. Then, there exists a $r > 0$ such that $\mathbb{E}_\lambda\left[e^{rS_1^\lambda}\right] < +\infty$.*

Proof. The Markovian Hawkes process satisfies Scenario 1.1 of Borovkov and Last (2008) and $\beta(a-\lambda) \neq 0$ for all $\lambda > a$. Therefore, we have that, under the stationary distribution, the number of continuous crossings of our process through λ has intensity $\mu_c(\lambda) := \beta(\lambda - a)p(\lambda) > 0$. Here, $p(\lambda)$ denotes the density of the stationary distribution ν. Consequently, we have

$$\mathbb{P}_\nu\left[S_1^\lambda > t\right] = \exp\left(-\int_0^t \beta(\lambda - a)p(\lambda)\,\mathrm{d}s\right) = \exp(-t\beta(\lambda - a)p(\lambda)).$$

This implies that $\mathbb{E}_\nu\left[e^{rS_1^\lambda}\right] < +\infty$ for all $r < \beta(\lambda - a)p(\lambda)$.
Using the fact that ν is absolutely continuous with respect to the Lebesgue measure, we have that

$$\mathbb{E}_\nu\left[e^{rS_1^\lambda}\right] = \int_a^\infty \mathbb{E}_x\left[e^{rS_1^\lambda}\right] p(x)\,\mathrm{d}x < +\infty,$$

which gives us that $\mathbb{E}_x\left[e^{rS_1^\lambda}\right] < +\infty$ for Lebesgue almost every $x > a$. Let now $\lambda < y$ be arbitrary and write $S_1^\lambda|_{\lambda_0=y}$ for the time of the first crossing of the level λ starting in y. Then, there exists a $x > y$ such that $\mathbb{E}\left[e^{rS_1^\lambda|_{\lambda_0=x}}\right] = \mathbb{E}_x\left[e^{rS_1^\lambda}\right] < +\infty$. The downward movement of the intensity process $\{\lambda_t\}_{t\geq 0}$ is continuous. Hence, if it starts in x and reaches the level λ it must cross y. By the strong Markov property, we can restart the process after hitting y and therefore

$$S_1^\lambda|_{\lambda_0=x} = S_1^y|_{\lambda_0=x} + S_1^\lambda|_{\lambda_0=y} \geq S_1^\lambda|_{\lambda_0=y}.$$

Thus,

$$\mathbb{E}_y\left[e^{rS_1^\lambda}\right] \leq \mathbb{E}_x\left[e^{rS_1^\lambda}\right] < +\infty.$$

This property holds for all λ and y as long $y > \lambda$ and r is chosen suitable small, depending on the choice of λ.

Consider now $S_1^\lambda|_{\lambda_0=\lambda}$. Then, there exists an $x < \lambda$ such that $\mathbb{E}_x\left[e^{rS_1^\lambda}\right] < +\infty$. Since $\lambda > x$, we know there exists a positive $\tilde{r} > 0$ such that $\mathbb{E}_\lambda\left[e^{\tilde{r}S_1^x}\right] < +\infty$. Now, there are almost surely two possibilities. Either the intensity hits the level x before it returns to λ, i.e. $S_1^x|_{\lambda_0=\lambda} \le S_1^\lambda|_{\lambda_0=\lambda}$, or it first returns to λ. If we have $\omega \in \Omega$ such that the path of $\{\lambda_t\}_{t\ge 0}$ hits the level x before returning to λ, we can use the restart argument as before and obtain the equality $S_1^\lambda|_{\lambda_0=\lambda}(\omega) = S_1^x|_{\lambda_0=\lambda}(\omega) + S_1^\lambda|_{\lambda_0=x}(\omega)$. For almost every other ω, we have $S_1^\lambda|_{\lambda_0=\lambda}(\omega) < S_1^x|_{\lambda_0=\lambda}(\omega) \le S_1^x|_{\lambda_0=\lambda}(\omega) + S_1^\lambda|_{\lambda_0=x}(\omega)$. This give us for q small enough that

$$\mathbb{E}_\lambda\left[e^{qS_1^\lambda}\right] \le \mathbb{E}_\lambda\left[e^{qS_1^x}\mathbb{E}_x\left[e^{qS_1^\lambda}\right]\right] = \mathbb{E}_\lambda\left[e^{qS_1^x}\right]\mathbb{E}_x\left[e^{qS_1^\lambda}\right] < +\infty.$$

This ends the proof. □

4.3.4 Exponential change of measure

We derive now the Cramér-Lundberg asymptotics under the assumption that claims are light-tailed. More precisely, we assume the following:

Assumption 4.2. *From now on, we assume that the distribution of the claim sizes F_U is absolutely continuous with respect to the Lebesgue measure. Further, we assume that there exists some $s_U \in (0,\infty]$ such that the corresponding moment-generating*

$$M_U(s) := \mathbb{E}\left[e^{sU}\right]$$

is finite for all $s < s_U$ and $\lim_{s\to s_U} M_U(s) = \infty$.

We are interested in the asymptotic behaviour of the ruin probability

$$\psi(u) = \psi(u,\lambda) := \mathbb{P}_{(u,\lambda)}\left(\tau < +\infty\right), \tag{4.23}$$

where

$$\tau := \inf\{t \ge 0 : X_t \le 0\}.$$

The main tool to show convergence of the ruin probability is Theorem 2 of Schmidli (1997) which gives us that the solution to the generalized renewal equation

$$Z(u) = \int_0^u Z(u-x)(1-p(u,x))B(\mathrm{d}x) + z(u) \tag{4.24}$$

converges as $u \to \infty$ if $B(x)$ is a probability distribution, $p(u,x) \in [0,1]$ is continuous in u, and both $z(u)$ and $\int_0^u p(u,x)B(\mathrm{d}x)$ are directly Riemann integrable.
The first problem that occurs in this approach is that this equation is univariate, whereas the probability of ruin $\psi(u,\lambda)$ depends on the initial values of the surplus and the intensity process. To resolve this, we use the results of Section 4.3.3, i.e. the intensity is Harris positive recurrent. To be precise, we exploit that it returns infinitely often to its initial value with probability 1. This allows us to choose renewal times so that they

coincide with intensity recurrence times. A second problem is that under our original measure $\mathbb{P}$, suitable choices of the distribution B are generally defective. We bypass this by identifying an alternative measure under which the ruin occurs almost surely and B is no longer defective.

Now, our main goal is to identify a martingale $\{M_t^{(r)}\}_{t\geq 0}$ and the corresponding alternative measure under which ruin almost surely occurs. For this, we follow the ansatz of Pojer and Thonhauser (2023a), that is, $M_t^{(r)} = \exp(-rX_t - \alpha\lambda_t - \theta t)$. This process is a local martingale if the function $h_r(x, \lambda, t) = \exp(-rx - \alpha\lambda - \theta t)$ is in the domain of the extended generator and satisfies

$$\mathcal{A}h_r(x, \lambda, t) = 0.$$

We start from the latter requirement, which by (4.21) is equivalent to

$$\begin{aligned}\mathcal{A}h_r(x, \lambda, t) = -crh_r(x, \lambda, t) - \alpha\beta(a - \lambda)h_r(x, \lambda, t) - \theta h_r(x, \lambda, t)\\ + \lambda h_r(x, \lambda, t)M_U(r)M_Y(-\alpha) - \lambda h_r(x, \lambda, t) = 0,\end{aligned}$$

for all choices of x, λ, t. Since h_r is positive, we can divide by h_r and get the following two equations

$$-cr - \alpha\beta a - \theta = 0, \tag{4.25}$$

$$\alpha\beta + M_U(r)M_Y(-\alpha) - 1 = 0. \tag{4.26}$$

For fixed r, we get two equations for two missing variables $\theta(r)$ and $\alpha(r)$. We focus on equation (4.26) defining $\alpha(r)$.

Lemma 4.22. *For $r \leq 0$ and some $r > 0$, there exist two distinct solutions to the equation* (4.26).

Proof. First, we consider the case $r < 0$. The function

$$f_r(\alpha) := \alpha\beta + M_U(r)M_Y(-\alpha) - 1$$

is convex and satisfies $f_r(0) = M_U(r)-1 < 0$. Furthermore, $\lim_{\alpha\to\infty} f_r(\alpha) = \lim_{\alpha\to-\infty} f_r(\alpha) = \infty$. By continuity, there exists at least one root in $(-\infty, 0)$ and one root in $(0, \infty)$. By convexity, the corresponding roots are unique. For $r = 0$, we have $f_0(0) = 0$ and $\frac{\partial}{\partial\alpha} f_0(0) = \beta - \mathbb{E}[Y] > 0$. By this, there exists some $\varepsilon > 0$ such that $f_0(-\varepsilon) < 0$ and, by the same argumentation as before, we have that there exists a unique negative root of f_0. For $r > 0$, we see that the function $f_r(\alpha)$ is also continuous in r (as long as it is well defined). Consequently, there exists some $\delta > 0$ such that for all $r < \delta$ we have $f_r(-\varepsilon) < 0$. Again, by continuity and convexity in α, we get the existence of two solutions to $f_r(\alpha) = 0$. □

Definition 4.6. For fixed r, we define $\alpha(r)$ as the maximal solution to equation (4.26).

This is well defined for all $r \leq r_{\max}$, where $r_{\max}$ satisfies

$$\min_{\alpha} \left(\alpha\beta + M_U(r_{\max})M_Y(-\alpha) - 1\right) = 0.$$

Further, we define the function

$$\theta(r) := -cr - \alpha(r)\beta a. \tag{4.27}$$

Lemma 4.23. *The mapping $r \to \alpha(r)$ is concave and differentiable on $(-\infty, r_{\max})$. Furthermore, it satisfies $\alpha(0) = 0$ and $r\alpha(r) < 0$ for all $r \neq 0$ such that $\alpha(r)$ is well defined.*

Proof. By the proof of Lemma 4.22, we have that f_0 has a negative root and satisfies $f_0(0) = 0$. Thus, $\alpha(0) = 0$. For $r > 0$, we have $f_r(0) = M_U(r) - 1 > 0$, which gives that all roots must be negative and for $r < 0$ $f_r(0) = M_U(r) - 1 < 0$. Therefore there exists a positive root, and for all $r \neq 0$ for which $\alpha(r)$ is well defined, we have $r\alpha(r) < 0$.
To show the concavity of $\alpha(r)$, we first show that the function $f(r, \alpha) := f_r(\alpha)$ is convex as a function of (r, α) from $(-\infty, s_U) \times (-s_Y, \infty)$ to $\mathbb{R}$, where s_Y and s_U are defined in Assumptions 4.1 and 4.2, respectively (it is even convex and proper as function from $\mathbb{R}^2 \to \mathbb{R} \cup \{\infty\}$ if we set $f(r, \alpha) = \infty$ for all (r, α) outside $(-\infty, s_U) \times (-s_Y, \infty)$). To show this, we consider the Hessian

$$\mathcal{H} = \begin{bmatrix} M_U''(r)M_Y(-\alpha) & -M_U'(r)M_Y'(-\alpha) \\ -M_U'(r)M_Y'(-\alpha) & M_Y''(-\alpha)M_U \end{bmatrix},$$

which has only non-negative eigenvalues by the log-convexity of the moment generating functions.
If we now take some $r \geq s$ and $\lambda \in [0, 1]$ such that $\alpha(r)$ is well defined, we find that also $\alpha(\lambda r + (1-\lambda)s)$ is well defined and satisfies $f(\lambda r + (1-\lambda)s, \alpha(\lambda r + (1-\lambda)s)) = 0$. Moreover, since it is the maximal root, for all $\alpha > \alpha(\lambda r + (1-\lambda)s)$ we have $f(\lambda r + (1-\lambda)s, \alpha) > 0$. By the convexity of the function f we get

$$f(\lambda r + (1-\lambda)s, \lambda\alpha(r) + (1-\lambda)\alpha(s)) \leq \lambda f(r, \alpha(r)) + (1-\lambda)f(s, \alpha(s)) = 0.$$

Consequently, $\lambda\alpha(r) + (1-\lambda)\alpha(s) \leq \alpha(\lambda r + (1-\lambda)s)$.
We still have to show that the function $r \to \alpha(r)$ is differentiable. By concavity, it is differentiable almost everywhere. To be specific, everywhere except some countable set and at every other point, the one-sided limits exist but do not coincide. Let r be such that the derivative of $\alpha(r)$ exists. Then, we get that

$$\beta\alpha'(r) + M_U'(r)M_Y(-\alpha(r)) - \alpha'(r)M_U(r)M_Y'(-\alpha(r)) = 0,$$

which is, if $\beta - M_U(r)M_Y'(-\alpha(r)) \neq 0$, equivalent to

$$\alpha'(r) = -\frac{M_U'(r)M_Y(-\alpha(r))}{\beta - M_U(r)M_Y'(-\alpha(r))}.$$

This is the case for all $r < r_{\max}$. In the case $r = r_{\max}$, the root $\alpha(r_{\max})$ also minimizes $f_{r_{\max}}(\alpha)$. Since this function is convex and differentiable in α, we have $0 = \frac{\partial}{\partial\alpha} f_{r_{\max}}(\alpha(r_{\max})) = \beta - M_U(r_{\max})M'_Y(-\alpha(r_{\max}))$. For all $r < r_{\max}$, the root $\alpha(r)$ is not the minimizer; therefore, $\alpha'(r)$ is well defined and continuous. By the continuity of the derivative, we find that $\alpha(r)$ is differentiable at every point $r < r_{\max}$. □

We recall that the function θ is defined in (4.27).

Lemma 4.24. *The function $\theta(r)$ is convex, differentiable and satisfies $\theta(0) = 0$ and $\theta'(0) < 0$.*

Proof. Since $\theta(r) = -cr - \alpha(r)a\beta$ is the sum of two differentiable and convex functions, it is differentiable and convex as well and $\theta(0) = 0$. The derivative at the point $r = 0$ is

$$\theta'(0) = -c - \alpha'(r)a\beta = -c + \frac{\beta a \mathbb{E}[U]}{\beta - \mathbb{E}[Y]},$$

which is negative by the net profit condition. This completes the proof. □

For further analysis, we will also need the following important assumption.

Assumption 4.3. *From now on, we assume that there exists a positive solution R of*

$$\theta(r) = 0. \tag{4.28}$$

for the function θ defined in (4.27). *Further, we assume that there exists an $\varepsilon > 0$ such that $M_U(R+\varepsilon)$, $M_Y(-\alpha(R+\varepsilon))$ are finite.*

Theorem 4.25. *The process*

$$M_t^{(r)} := \exp(ru + \alpha(r)\lambda)\exp(-rX_t - \alpha(r)\lambda_t - \theta(r)t) \tag{4.29}$$

is a non-negative local martingale for all $0 \le r \le R + \varepsilon$.

Proof. Fix $r \le R + \varepsilon$ arbitrary and define the function

$$h_r(x, \tilde{\lambda}, t) := \exp(-r(x-u) - \alpha(r)(\tilde{\lambda} - \lambda) - \theta(r)t).$$

We show that this function is in the domain of the extended generator of our PDMP and satisfies $\mathcal{A}h_r(x, \tilde{\lambda}, t) = 0$, which gives us that this is a local martingale. The function h_r is absolutely continuous. Hence, by Theorem (26.14) and Remark (26.16) of Davis (1993), we only have to show that for all $n \in \mathbb{N}$

$$\mathbb{E}_{(u,\lambda)}\left[\sum_{i=1}^{n} |h_r(X_{T_i}, \lambda_{T_i}, T_i) - h_r(X_{T_i-}, \lambda_{T_i-}, T_i)|\right] < +\infty.$$

This is obviously satisfied for $r \leq 0$. For the case $r > 0$, we observe that $\alpha(r) < 0$ and the compensator of the jump process $\{N_t\}_{t\geq 0}$ is given by

$$\begin{aligned}\Lambda_t := \int_0^t \lambda_s \, \mathrm{d}s &= at + \frac{1}{\beta}(\lambda_0 - a)(1 - e^{-\beta t}) + \frac{1}{\beta}\sum_{k=1}^{N_t} Y_k (1 - e^{-\beta(t-T_k)}) \\ &= at + \frac{1}{\beta}\sum_{k=1}^{N_t} Y_k + \frac{1}{\beta}\lambda_0 - \frac{1}{\beta}\lambda_t.\end{aligned}$$

Consequently, we have for all $i \leq n$,

$$\begin{aligned}&- r(X_{T_i} - u) - \alpha(r)(\lambda_{T_i} - \lambda) - \theta(r)T_i \\ &\qquad = -rcT_i + r\sum_{j=1}^{i} U_j - \alpha(r)\left(\sum_{j=1}^{i} Y_j + \beta a T_i - \beta \Lambda_{T_i}\right) + crT_i + a\alpha(r)\beta T_i \\ &\qquad \leq r\sum_{j=1}^{i} U_j - \alpha(r)\sum_{j=1}^{i} Y_j.\end{aligned}$$

Further, observe that

$$h_r(X_{T_i}, \lambda_{T_i}, T_i) = h_r(X_{T_i-}, \lambda_{T_i-}, T_i)\exp\left(rU_i - \alpha(r)Y_i\right) > h_r(X_{T_i-}, \lambda_{T_i-}, T_i).$$

Using this, we get that

$$\begin{aligned}&\mathbb{E}_{(u,\lambda)}\left[\sum_{i=1}^{n} |h_r(X_{T_i}, \lambda_{T_i}, T_i) - h_r(X_{T_i-}, \lambda_{T_i-}, T_i)|\right] \\ &\leq 2\,\mathbb{E}_{(u,\lambda)}\left[\sum_{i=1}^{n} h_r(X_{T_i}, \lambda_{T_i}, T_i)\right] \leq 2\sum_{i=1}^{n}\mathbb{E}_{(u,\lambda)}\left[\exp\left(r\sum_{j=1}^{i} U_j - \alpha(r)\sum_{j=1}^{i} Y_j\right)\right] \\ &= 2\sum_{i=1}^{n} M_U(r)^i M_Y(-\alpha(r))^i = 2M_U(r)M_Y(-\alpha(r))\frac{M_U(r)^n M_Y(-\alpha(r))^n - 1}{M_U(r)M_Y(-\alpha(r)) - 1} < +\infty.\end{aligned}$$

By this, the function h_r is in the domain of the extended generator and by the construction of $\alpha(r)$ and $\theta(r)$ it satisfies $\mathcal{A}h_r(x, \lambda, t) = 0$. This completes the proof. □

Theorem 4.26. *Let $r < R + \varepsilon$ for some $\varepsilon > 0$. Then, $M^{(r)} = \{M_t^{(r)}\}_{t\geq 0}$ defined in* (4.29) *is a true martingale with expectation* 1.

Proof. Fix $r < R + \varepsilon$ and let $\{\varrho_n\}_{n\in\mathbb{N}}$ be a localizing sequence of stopping times for the local martingale $\{M_t^{(r)}\}_{t\geq 0}$. Then, by Lemma 2.2.2 of Fleming and Harrington (1991), we have that $\{M_t^{(r)}\}_{t\geq 0}$ is a martingale if for any fixed t, the family $\mathcal{X} = \{M_{t\wedge\varrho_n}^{(r)}\}_{n\in\mathbb{N}}$ is uniformly integrable. By de La Vallée Poussin's Theorem, a family of random variables

$\{Y_n\}_{n\in A}$ is uniformly integrable if there exists a monotone increasing convex function $G(t)$, satisfying $\lim_{t\to\infty}\frac{G(t)}{t}=\infty$ and $\sup_{n\in A}\mathbb{E}\left[G(|Y_n|)\right]<+\infty$.

Since $r<R+\varepsilon$, there exists a $\delta>0$ such that $r(1+\delta)<R+\varepsilon$ and $\{M_t^{(r(1+\delta))}\}_{t\geq 0}$ is well-defined. Since every non-negative local martingale with integrable initial value is a supermartingale, we have that for all $t\geq 0$ that $M_t^{(r(1+\delta))}$ is integrable with expectation less or equal $M_0^{(r(1+\delta))}=1$.
Let now t and n be arbitrary but fixed. Then, we have that

$$\begin{aligned}\left(M_{t\wedge\varrho_n}^{(r)}\right)^{1+\delta} &= \exp\left(-r(1+\delta)(X_{t\wedge\varrho_n}-u)-\alpha(r)(1+\delta)(\lambda_{t\wedge\varrho_n}-\lambda)\right)\\ &\qquad\times\exp\left(cr(1+\delta)t\wedge\varrho_n+a\beta\alpha(r)(1+\delta)t\wedge\varrho_n\right)\\ &= M_{t\wedge\varrho_n}^{(r(1+\delta))}\exp\left((\alpha(r(1+\delta))-\alpha(r)(1+\delta))\lambda_{t\wedge\varrho_n}\right)\\ &\qquad\qquad\times\exp\left((\alpha(r)(1+\delta)-\alpha(r(1+\delta)))(\lambda+a\beta t\wedge\varrho)\right).\end{aligned}$$

If we can show that $\alpha(r(1+\delta))-\alpha(r)(1+\delta)\leq 0$, then we have by the positivity of $\lambda_{t\wedge\varrho_n}$ that

$$\exp\left((\alpha(r(1+\delta))-\alpha(r)(1+\delta))\lambda_{t\wedge\varrho_n}\right)\leq 1,$$

and, since $t\wedge\varrho_n\leq t$, we would have that

$$\begin{aligned}&\exp\left(\alpha(r)(1+\delta)-\alpha(r(1+\delta))\right)(\lambda_0+a\beta t\wedge\varrho))\\ &\quad\leq\exp\left((\alpha(r)(1+\delta)-\alpha(r(1+\delta)))\right)\exp\left((\lambda_0+a\beta t)\right),\end{aligned}$$

which is deterministic and finite.
To show this, we will use the fact that $\alpha(r)$ is concave, differentiable and satisfies $\alpha(0)=0$; see Lemma 4.23. By this, we get that

$$\begin{aligned}\alpha(r(1+\delta))-\alpha(r)(1+\delta) &= \alpha(r(1+\delta))-\alpha(r)-\alpha(r)\delta\leq\alpha'(r)r\delta-\alpha(r)\delta\\ &= \alpha'(r)r\delta+\delta(\alpha(0)-\alpha(r))\leq\delta\alpha'(r)r\delta+\delta\alpha'(r)(-r)=0.\end{aligned}$$

This gives that

$$\begin{aligned}\mathbb{E}_{(u,\lambda)}\left[\left(M_{t\wedge\varrho_n}^{(r)}\right)^{1+\delta}\right] &\leq\mathbb{E}_{(u,\lambda)}\left[\left(M_{t\wedge\varrho_n}^{(r(1+\delta))}\right)\right]\\ &\qquad\times\exp\left(\alpha(r)(1+\delta)-\alpha(r(1+\delta))\right)(\lambda_0+a\beta t))\\ &\leq\exp\left(\alpha(r)(1+\delta)-\alpha(r(1+\delta))\right)(\lambda_0+a\beta t)).\end{aligned}$$

This bound is independent of n and finite for fixed t. Hence, taking the supremum gives us that

$$\sup_{n\in\mathbb{N}}\mathbb{E}_{(u,\lambda)}\left[\left(M_{t\wedge\varrho_n}^{(r)}\right)^{1+\delta}\right]<+\infty.$$

Therefore, $\mathcal{X}$ is uniformly integrable and the process $\{M_t^{(r)}\}_{t\geq 0}$ is a true martingale with expectation $M_0^{(r)} = 1$. ☐

Definition 4.7. Let $R + \varepsilon > r \geq 0$ for some $\varepsilon > 0$. Then, we define the measure $\mathbb{Q}^{(r)}$ by

$$\mathbb{Q}^{(r)}[A] = \mathbb{E}_{(u,\lambda)}\left[I_A M_t^{(r)}\right], \; \forall A \in \mathcal{F}_t.$$

Lemma 4.27. *Under the new measure $\mathbb{Q}^{(r)}$, the multivariate process $\{(X_t, \lambda_t, t)\}_{t\geq 0}$ is again a PDMP with generator*

$$\begin{aligned}\mathcal{A}^{(r)} f(x, \lambda, t) &= c\frac{\partial}{\partial x} f(x, \lambda, t) + \beta(a - \lambda)\frac{\partial}{\partial \lambda} f(x, \lambda, t) + \frac{\partial}{\partial t} f(x, \lambda, t) \\ &\quad + \lambda \int_0^\infty \int_0^\infty e^{ru} e^{-\alpha(r)y} \left(f(x - u, \lambda + y, t) - f(x, \lambda, t)\right) F_U(\mathrm{d}u)\, F_Y(\mathrm{d}y).\end{aligned}$$

Proof. This follows directly from Example 5.2 of Palmowski and Rolski (2002), where exactly this kind of exponential change of measures for PDMPs is studied. ☐

Lemma 4.28. *Under the new measure $\mathbb{Q}^{(R)}$, ruin occurs almost surely.*

Proof. Using the same ideas as in Lemma 4.8 but with the alternative generator $\mathcal{A}^{(R)}$, it is easy to see that

$$\lim_{t\to\infty} \frac{\mathbb{E}^{(R)}_{(u,\lambda)}[X_t]}{t} = -\theta'(R).$$

By the convexity of θ proved in Lemma 4.24 and the fact that there is some $r < R$ with $\theta(r) < 0$, we have that $-\theta'(R) < 0$. Consequently, ruin occurs almost surely under the new measure $\mathbb{Q}^{(R)}$. ☐

We will that under the new measure $\mathbb{Q}^{(r)}$, $\{\lambda_t\}_{t\geq 0}$ is no longer the intensity of a Markovian marked Hawkes process $\{N_t\}_{t\geq 0}$. In fact its jumps have now intensity $\{\lambda_t M_U(r) M_Y(-\alpha(r))\}_{t\geq 0}$ that still preserves its recurrent behaviour.

Lemma 4.29. *The process $\{\lambda_t\}_{t\geq 0}$ is Harris recurrent under $\mathbb{Q}^{(r)}$.*

Proof. At first, we show that, under the measure $\mathbb{Q}^{(r)}$, the process $\{M_U(r) M_Y(-\alpha(r))\lambda_t\}_{t\geq 0}$ is the intensity of a Markovian marked Hawkes process $\{N_t\}_{t\geq 0}$. Indeed, from the form of the generator $\mathcal{A}^{(r)}$ given in Lemma 4.27 we can conclude that the univariate process $\{\lambda_t\}_{t\geq 0}$ is a Markov process with generator

$$\mathcal{A}^{(r),\lambda} f(\lambda) = \beta(a - \lambda) f'(\lambda) + \lambda M_u(r) M_Y(-\alpha(r)) \int_0^\infty \left(f(\lambda + y) - f(\lambda)\right) \tilde{F}_Y(\mathrm{d}y),$$

where the distribution $\tilde{F}_Y$ is given by $\tilde{F}_Y(\mathrm{d}y) = \frac{e^{-\alpha(r)y}}{M_Y(-\alpha(r))} F_Y(\mathrm{d}y)$. Hence, under $\mathbb{Q}^{(r)}$, the process $\{\lambda_t\}_{t\geq 0}$ has the form

$$\lambda_t = e^{-\beta t}(\lambda - a) + a + \sum_{i=1}^{N_t^{(r)}} \tilde{Y}_i e^{-\beta(t - T_i^{(r)})},$$

where $\{N_t^{(r)}\}_{t\geq 0}$ has the intensity process $\{M_U(r)M_Y(-\alpha(r))\lambda_t\}_{t\geq 0}$.
Observe that the PDMP $\{M_U(r)M_Y(-\alpha(r))\lambda_t\}_{t\geq 0}$ can represented as follows

$$M_U(r)M_Y(-\alpha(r))\lambda_t = e^{-\beta t}(M_U(r)M_Y(-\alpha(r))\lambda - M_U(r)M_Y(-\alpha(r))a) + M_U(r)M_Y(-\alpha(r))a + \sum_{i=1}^{N_t^{(r)}} M_U(r)M_Y(-\alpha(r))\tilde{Y}_i e^{-\beta(t-T_t^{(r)})},$$

that is, the process $\{N_t^{(r)}\}_{t\geq 0}$ is a Markovian Hawkes process. The parameters are given by decay parameter β, baseline intensity $aM_U(r)M_Y(-\alpha(r))$ and shock distribution $\tilde{F}_Y(z/(M_U(r)M_Y(-\alpha(r))))$.
By Theorem 4.17 and (4.20), if we can now show that

$$\beta > M_U(r)M_Y(-\alpha(r))\mathbb{E}^{\mathbb{Q}^{(r)}}[Y] = M_U(r)M_Y(-\alpha(r))\mathbb{E}\left[\tilde{Y}\right], \tag{4.30}$$

then $\{M_U(r)M_Y(-\alpha(r))\lambda_t\}_{t\geq 0}$ returns to every point in $(aM_U(r)M_Y(-\alpha(r)),\infty)$ infinitely often. This implies that $\{\lambda_t\}_{t\geq 0}$ visits every point in (a,∞) infinitely often.
To prove (4.30), observe first that the expectation of Y under our new measure is $\frac{M_Y'(-\alpha(r))}{M_Y(-\alpha(r))}$ and that the mapping $r\to\alpha(r)$ is monotone decreasing. Hence, by the proof of Lemma 4.23 we have that $-\frac{M_U'(r)M_Y(-\alpha(r))}{\beta-M_U(r)M_Y'(-\alpha(r))}\leq 0$, which implies that

$$\beta - M_U(r)M_Y(-\alpha(r))\mathbb{E}[\tilde{Y}] = \beta - M_U(r)M_Y'(-\alpha(r)) > 0.$$

By this, we have that our intensity process returns almost surely to every point in $(aM_U(r)M_Y(-\alpha(r)),\infty)$. □

4.3.5 Cramér-Lundberg asymptotics and renewal arguments

Now, fix an initial value λ and let $S_1^\lambda, S_2^\lambda,\dots$ denote the recurrence times of the intensity process to the level λ, i.e. $\lambda_{S_i^\lambda}=\lambda$ for all i. Further, we define the renewal times $\{\phi_i\}_{i\geq 1}$ by $\phi_1=\min\left\{S_i^\lambda : X_{S_i^\lambda}<u\right\}$ and $\phi_j=\min\left\{S_i^\lambda : X_{S_i^\lambda}<X_{\phi_{j-1}}\right\}$ for $j>1$.
These times are a mixture of the recurrence times of the intensity process and ladder times of the surplus process, i.e. ladder times of the random process $\{X_{S_i^\lambda}\}_{i\geq 1}$. As we can see, these renewal times might be infinite under our original measure since $X_t\to+\infty$ $\mathbb{P}$-a.s. But under our alternative measure $\mathbb{Q}^{(R)}$, the surplus process $\{X_t\}_{t\geq 0}$ tends to $-\infty$, and the intensity returns infinitely often to λ. Hence, these times are finite almost surely. Define now

$$\tilde{B}(x) := \mathbb{P}_{(u,\lambda)}\left[\phi_1<+\infty, u-X_{\phi_1}\leq x\right]$$

and

$$p(u,x) := \mathbb{P}_{(u,\lambda)}\left[\tau\leq\phi_1, |\phi_1<+\infty, X_{\phi_1}=u-x\right].$$

Then we have, by conditioning on the distribution of the surplus at time ϕ_1, that

$$\psi(u,\lambda) = \int_0^u \psi(u-x,\lambda)(1-p(u,x))\,\tilde{B}(\mathrm{d}x) + \mathbb{P}_{(u,\lambda)}\left[\tau \le \phi_1, \tau < +\infty\right].$$

As already mentioned, the distribution $\tilde{B}$ is defective. To work with a proper distribution, we multiply the equation by e^{Ru} to obtain

$$\psi(u,\lambda)e^{Ru} = \int_0^u \psi(u-x,\lambda)e^{R(u-x)}(1-p(u,x))\,\tilde{B}(\mathrm{d}x) + e^{Ru}\mathbb{P}_{(u,\lambda)}\left[\tau \le \phi_1, \tau < +\infty,\right] \tag{4.31}$$

where

$$B(\mathrm{d}x) := e^{Rx}\tilde{B}(\mathrm{d}x).$$

Lemma 4.30. *The distribution B is a proper probability distribution.*

Proof. By the definition of $\tilde{B}$, we have that

$$\int_{\mathbb{R}} e^{Rx}\,\tilde{B}(\mathrm{d}x) = \mathbb{E}_{(u,\lambda)}\left[e^{R(u-X_{\phi_1})}I_{\{\phi_1<+\infty\}}\right].$$

Our martingale $\{M_t^{(R)}\}_{t\ge 0}$ at time ϕ_1 has the form

$$M_{\phi_1}^{(R)} = \exp(-R(X_{\phi_1}-u) - \alpha(R)(\lambda_{\phi_1}-\lambda)) = \exp(R(u-X_{\phi_1})).$$

Therefore,

$$\int_{\mathbb{R}} B(\mathrm{d}x) = \int_{\mathbb{R}} e^{Rx}\,\tilde{B}(\mathrm{d}x) = \mathbb{Q}^{(R)}\left[\phi_1 < +\infty\right] = 1.$$

□

Observe that equation (4.31) is of the form of renewal equation (4.24), that is,

$$Z(u) = \int_0^u Z(u-x)(1-p(u,x))B(\mathrm{d}x) + z(u) \tag{4.32}$$

for

$$Z(u) := \psi(u,\lambda)e^{Ru} \quad \text{and} \quad z(u) := e^{Ru}\mathbb{P}_{(u,\lambda)}\left[\tau \le \phi_1, \tau < +\infty\right].$$

To show convergence of $Z(u)$, hence Cramér-Lundberg asymptotics, we have to verify that $z(u)$ and $\int_0^u p(u,x)B(\mathrm{d}x)$ are directly Riemann integrable. For the direct Riemann integrability of the above-mentioned functions, we need to introduce an additional assumption.

Assumption 4.4. *We assume there exists an $\varepsilon > 0$ such that*

$$\mathbb{E}_{(u,\lambda)}\left[e^{-(1+\varepsilon)R(X_{\phi_1}-u)}I_{\{\phi_1<+\infty\}}\right] < +\infty.$$

Remark 4.3. The random time ϕ_1 depends on the behaviour of the bivariate process $\{(X_t,\lambda_t)\}_{t\ge 0}$. Therefore, this assumption may be hard to check. An alternative to this is the condition

$$\mathbb{E}_{(u,\lambda)}\left[e^{-(1+\varepsilon)R(X_{S_1^\lambda}-u)}I_{\left\{S_1^\lambda<+\infty\right\}}\right]<+\infty,$$

which is equivalent to Assumption 4.4 by Lemma 10 of Pojer and Thonhauser (2023a). Changing measure, we see that this assumption is equivalent to

$$\mathbb{E}^{\mathbb{Q}^{(1+\varepsilon)R}}\left[e^{\theta((1+\varepsilon)R)\phi_1}\right]<+\infty, \tag{4.33}$$

which shows the main influence of this assumption. By the structure of our renewal times, we cannot observe ruin exactly when it happens. Assumption 4.4 ensures that these renewal times happen often enough, such that there is one of these close enough to the time of ruin such that we do not miss the event that the surplus process is negative.

Now we will show that the functions $z(u)$ and $\int_0^u p(u,x)B(\mathrm{d}x)$ are directly Riemann integrable. To do so, we will use Proposition V.4.1 on p. 154 of Asmussen (1995), which gives us that it is sufficient to prove that both considered functions are continuous and there exists bounded directly Riemann integrable upper bounds for these functions.

Lemma 4.31. *Under Assumptions 4.1-4.4, we have that*

$$z(u)=e^{Ru}\mathbb{P}_{(u,\lambda)}\left[\tau\leq\phi_1,\tau<+\infty\right]$$

is directly Riemann integrable.

Proof. We start the proof by showing that the function $z(u)$ is continuous. Indeed, from (4.32) it follows that it suffices to show continuity of the ruin probability $\psi(u,\lambda)$ as a function of $u>0$. For $h>0$, by Markov property of $\{(X_t,\lambda_t,t)\}_{t\geq0}$, we have,

$$\begin{aligned}\psi(u,\lambda)&=\mathbb{P}(N_h=0)\psi(u+ch,\lambda)\\&\quad+\sum_{k=1}^{\infty}\mathbb{P}(N_h=k)\,\mathbb{E}\left[\psi(u+c(T_1+\ldots+T_k)-U_1-\ldots-U_k)\right].\end{aligned} \tag{4.34}$$

Observe that $\lim_{h\to0}\mathbb{P}(N_h=k)=0$ for $k\in\mathbb{N}$ and $\lim_{h\to0}\mathbb{P}(N_h=0)=1$; see e.g. Hawkes (1971). Hence, by Lemma 4.1 and the dominated convergence theorem, we can conclude that $\lim_{h\to0}\sum_{k=1}^{\infty}\mathbb{P}(N_h=k)\mathbb{E}\left[\psi(u+c(T_1+\ldots+T_k)-U_1-\ldots-U_k)\right]=0$ and that $\psi(u,\lambda)$ is right-continuous. Plugging on the left-hand side of (4.34), $u-ch$ instead of u into the argument of ψ gives the left-continuity of this function.
Now, let $\varepsilon>0$ such that $\mathbb{E}_{(u,\lambda)}\left[e^{-(1+\varepsilon)R(X_{\phi_1}-u)}I_{\{\phi_1<+\infty\}}\right]<+\infty$. Let $r=(1+\varepsilon)R$. Observe that

$$\begin{aligned}e^{ru}\mathbb{P}_{(u,\lambda)}\left[\tau\leq\phi_1,\tau<+\infty\right]&=e^{ru}\mathbb{E}_{(u,\lambda)}\left[I_{\{\tau<\phi_1\}}I_{\{\tau<+\infty\}}\right]\\&=e^{ru}\mathbb{E}^{\mathbb{Q}^{(r)}}\left[I_{\{\tau<\phi_1\}}\exp\left(r(X_\tau-u)+\alpha(r)(\lambda_\tau-\lambda)+\theta(r)\tau\right)\right].\end{aligned}$$

Since $r>0$ and $X_\tau<0$, we have that $\exp(rX_\tau)<1$ and the same holds for $\exp(\alpha(r)\lambda_\tau)$.

Therefore,

$$
\begin{aligned}
e^{ru}\mathbb{P}_{(u,\lambda)}\left[\tau \le \phi_1, \tau < +\infty\right] &\le e^{-\alpha(r)\lambda}\mathbb{E}^{\mathbb{Q}^{(r)}}\left[I_{\{\tau<\phi_1\}}\exp\left(\theta(r)\phi_1\right)\right] \\
&\le e^{-\alpha(r)\lambda}\mathbb{E}^{\mathbb{Q}^{(r)}}\left[I_{\{\phi_1<+\infty\}}\exp\left(\theta(r)\phi_1\right)\right] \\
&= e^{-\alpha(r)\lambda}\mathbb{E}_{(u,\lambda)}\left[e^{-(1+\varepsilon)R(X_{\phi_1}-u)}I_{\{\phi_1<+\infty\}}\right] < +\infty.
\end{aligned}
$$

Thus, we have that there exists a positive constant K such that

$$
e^{ru}\mathbb{P}_{(u,\lambda)}\left[\tau \le \phi_1, \tau < +\infty\right] \le Ke^{-(r-R)u} = Ke^{-\varepsilon u}
$$

and the upper bound is bounded and directly Riemann integrable. This completes the proof. □

Lemma 4.32. *Under Assumptions 4.1-4.4, the function which maps u to $\int_0^u p(u,x)e^{Rx}B(\mathrm{d}x)$ is directly Riemann integrable.*

Proof. Observe that the function $u \to \int_0^u p(u,x)e^{Rx}B(\mathrm{d}x)$ is continuous. To identify a bounded directly Riemann integrable upper bound, we choose an arbitrary but fixed u. Then,

$$
\begin{aligned}
\int_0^u p(u,x)e^{Rx}B(\mathrm{d}x) \le e^{Ru}\int_0^u p(u,x)B(\mathrm{d}x) &= e^{Ru}\mathbb{P}_{(u,\lambda)}\left[\tau \le \phi_1, \phi_1 < +\infty\right] \\
&\le e^{Ru}\mathbb{P}_{(u,\lambda)}\left[\tau \le \phi_1, \tau < +\infty\right].
\end{aligned}
$$

By Lemma 4.31, we know that the upper bound is directly Riemann integrable and bounded, which completes the proof. □

We are now ready to prove our next main result.

Theorem 4.33. *Under Assumptions 4.1-4.4, there exists a constant $C^\lambda > 0$, depending on λ, such that*

$$
\lim_{u\to\infty}\psi(u,\lambda)e^{Ru} = C^\lambda,
$$

where the adjustment coefficient $R > 0$ solves Lundberg equation (4.28).

Proof. By the absolute continuity of the claim events and the proof of Lemma 12 of Pojer and Thonhauser (2023a), we have that $p(u,x)$ is continuous in u. By this and Lemmas 4.30-4.32, all assumptions of Theorem 2 of Schmidli (1997) are satisfied. □

Remark 4.4. By Corollary 4.5, the adjustment coefficient $R > 0$ defined via (4.28) equals to the adjustment coefficient of 'shifted' Cramér-Lunberg risk process defined in (4.17).

Remark 4.5. Theorem 4.33 gives a stronger statement than Theorem 4.1 of Stabile and Torrisi (2010) who derived only the logarithmic asymptotic showing that $\lim_{u\to+\infty}\frac{1}{u}\ln\psi(u) = R$.

4.4 Markovian Hawkes arrival process with exponentially distributed shocks and exponential claims

Here, we introduce an example where all Assumptions 4.1-4.4 are satisfied. For this, we consider a Markovian Hawkes process with the intensity process (4.19) and with exponentially distributed shocks

$$Y_i \sim \mathrm{Exp}(\gamma).$$

To ensure that the integrability condition $\frac{\mathbb{E}[Y]}{\beta} < 1$ given in (4.20) is satisfied, we assume that

$$\beta\gamma > 1.$$

The stationary distribution

We start from the following fact which is of own interest.

Theorem 4.34. *The stationary measure ν of the intensity process $\{\lambda_t\}_{t\geq 0}$ exists and it is shifted Gamma law, that is,*

$$\nu \sim a + \mathrm{Gamma}(a/\beta, (\beta\gamma - 1)/\beta). \tag{4.35}$$

Proof. By Theorem 34.19 on p. 118 of Davis (1984) (see also Prop. 34.7, p. 113 and Prop. 34.11, p. 115 of Davis (1984)) and the stationary distribution ν of the PDMP with density p satisfies

$$0 = \int_a^\infty \mathcal{A}f(x)\nu(\mathrm{d}x) = \int_a^\infty f(x)\mathcal{A}^* p(x)\,\mathrm{d}x,$$

for all $f \in D(\mathcal{A})$ in the domain of the generator $\mathcal{A}$, where $\mathcal{A}^*$ is an adjoint operator to $\mathcal{A}$. If we can find the unique solution to the equation

$$\mathcal{A}^* g(\lambda) = 0, \tag{4.36}$$

then, by the uniqueness of the stationary distribution, this solution must be a density of the stationary distribution.
We recall that

$$\begin{aligned}\mathcal{A}f(\lambda) &= \beta(a-\lambda)f'(\lambda) + \lambda\int_0^\infty \gamma e^{-\gamma y}\left(f(\lambda+y) - f(\lambda)\right)\mathrm{d}y \\ &= \beta(a-\lambda)f'(\lambda) + \lambda\int_\lambda^\infty \gamma e^{-\gamma(y-\lambda)} f(y)\,\mathrm{d}y - \lambda f(\lambda)\end{aligned}$$

and the adjoint operator $\mathcal{A}^*$ satisfies

$$\int_a^\infty (\mathcal{A}f(\lambda))g(\lambda)\,\mathrm{d}\lambda = \int_a^\infty (f(\lambda))\mathcal{A}^* g(\lambda)\,\mathrm{d}\lambda,$$

for all functions f and g from the domain of $\mathcal{A}$. Therefore, for $b > a$,

$$\int_a^b (\mathcal{A}f(\lambda))g(\lambda)\,\mathrm{d}\lambda = \int_a^b \beta(a-\lambda)f'(\lambda)g(\lambda)\,\mathrm{d}\lambda \\ + \int_a^b \lambda g(\lambda) \int_\lambda^\infty \gamma e^{-\gamma(y-\lambda)} f(y)\,\mathrm{d}y\,\mathrm{d}\lambda - \int_a^b \lambda f(\lambda)g(\lambda)\,\mathrm{d}y\lambda.$$

Furthermore,. if we use integration by parts in the first integral, we get

$$\int_a^b \beta(a-\lambda)f'(\lambda)g(\lambda) = f(b)\beta(a-b)g(b) - \int_a^b f(\lambda)\left(\beta(a-\lambda)g'(\lambda) - \beta g(\lambda)\right)\,\mathrm{d}\lambda.$$

In the second term, we interchange integrals and obtain

$$\int_a^b \lambda g(\lambda) \int_\lambda^\infty \gamma e^{-\gamma(y-\lambda)} f(y)\,\mathrm{d}y\,\mathrm{d}\lambda = \int_a^\infty f(y) \int_a^{\min(y,b)} \lambda g(\lambda)\gamma e^{-\gamma(y-\lambda)}\,\mathrm{d}\lambda\,\mathrm{d}y \\ = \int_a^b f(\lambda) \int_a^\lambda y g(y)\gamma e^{-\gamma(\lambda-y)}\,\mathrm{d}y\,\mathrm{d}\lambda + \int_b^\infty f(\lambda) \int_a^b y\,g(y)\,\gamma e^{-\gamma(\lambda-y)}\,\mathrm{d}y\,\mathrm{d}\lambda.$$

Plugging these together, we have that

$$\int_a^b (\mathcal{A}f(\lambda))g(\lambda)\,\mathrm{d}\lambda = f(b)\beta(a-b)g(b) + \int_b^\infty f(\lambda) \int_a^b y\,g(y)\,\gamma e^{-\gamma(\lambda-y)}\,\mathrm{d}y\,\mathrm{d}\lambda \\ + \int_a^b f(\lambda)\left(\beta(a-\lambda)g'(\lambda) - \beta g(\lambda) + \int_a^\lambda y\,g(y)\,\gamma e^{-\gamma(\lambda-y)}\,\mathrm{d}y - \lambda\,g(\lambda)\right)\,\mathrm{d}\lambda.$$

If we let b tend to infinity, the first two terms vanish and we find that the adjoint operator is given by

$$\mathcal{A}^* g(\lambda) = \beta(a-\lambda)g'(\lambda) - \beta g(\lambda) + \int_a^\lambda y\,g(y)\,\gamma e^{-\gamma(\lambda-y)}\,\mathrm{d}y - \lambda\,g(\lambda).$$

To solve equation (4.36), observe that

$$\frac{\partial}{\partial\lambda}\mathcal{A}^* g(\lambda) = \beta(\lambda-a)g''(\lambda) + (2\beta-\lambda)g'(\lambda) + (\beta\gamma-1)g(\lambda) - \int_a^\lambda \gamma^2 e^{-\gamma(\lambda-y)}\,y\,g(y)\,\mathrm{d}y$$

and the solution of (4.36) satisfies

$$0 = \gamma\mathcal{A}^* g(\lambda) + \frac{\partial}{\partial\lambda}\mathcal{A}^* g(\lambda) \\ = \beta(\lambda-a)g''(\lambda) - (\lambda + \beta(-2 + a\gamma - \gamma\lambda))g'(\lambda) + (\beta\gamma-1)g(\lambda).$$

This equation has solutions of the form

$$g(\lambda) = c_1 e^{-\frac{(\beta\gamma-1)}{\beta}\lambda}(\lambda-a)^{\frac{a}{\beta}-1} \\ + c_2 e^{-\frac{(\beta\gamma-1)}{\beta}\lambda}(\lambda-a)^{\frac{a}{\beta}-1}\Gamma\left(1-\frac{a}{\beta},(\frac{1}{\beta}-\gamma)(\lambda-a)\right), \quad (4.37)$$

where Γ denotes the incomplete gamma function. To get a proper distribution from the function g we have to set $c_2 = 0$. Hence $g(\lambda)$ is the density of a gamma distribution with parameters $\frac{a}{\beta}$ and $\frac{(\beta\gamma-1)}{\beta}$ and support shifted by a. This completes the proof. □

Remark 4.6. This coincides with Remark 4.3 of Dassios and Zhao (2011), where they derived the stationary distribution using the limit of the corresponding Laplace transformations.

Assumptions 4.1- 4.3 and the form of adjustment coefficient R

We now consider the surplus process

$$X_t = u + ct - \sum_{i=1}^{N_t} U_i.$$

where the claims have an exponential distribution with parameter $\mu > 0$, that is,

$$U_i \sim \text{Exp}(\mu).$$

In this case, the net profit condition simplifies to

$$c > \frac{a\beta\gamma}{\mu(\beta\gamma-1)},$$

and the moment generating function $M_U(r) = \frac{\mu}{\mu-r}$ is well defined for all $r < \mu$ and satisfies $\lim_{r\to\mu} M_U(r) = +\infty$. Since $Y_i \sim \text{Exp}(\gamma)$, we get $M_Y(-\alpha) = \frac{\gamma}{\gamma+\alpha}$, for $\alpha > -\gamma$ and $\lim_{\alpha\to-\gamma} M_U(\alpha) = +\infty$. Hence Assumptions 4.1- 4.2 are satisfied.
To verify that Assumption 4.3 is also true, observe that the equations for $\theta(r)$ and $\alpha(r)$ have the form

$$\alpha^2\beta + \alpha(\beta\gamma-1) + \frac{\mu\gamma}{\mu-r} - \gamma = 0,$$
$$\theta = -cr - \alpha a\beta.$$

We can solve the quadratic equation for α and obtain the following solutions

$$\alpha_{1,2} = \frac{1-\gamma\beta}{2\beta} \pm \frac{\sqrt{(-4r\beta\gamma + (-1+\beta\gamma)^2(\mu-r))(\mu-r)}}{2\beta(\mu-r)}.$$

As we expect from the theory already derived, there are two distinct real solutions for

α as long $r < r_{\max} = \frac{(\beta\gamma-1)^2}{(\beta\gamma+1)^2}\mu$, there is one single solution for $r = r_{\max}$ and no real solution if $r > r_{\max}$.
The larger solution is

$$\alpha(r) = \frac{1-\gamma\beta}{2\beta} + \frac{\sqrt{(-4r\beta\gamma + (-1+\beta\gamma)^2(\mu - r))(\mu - r)}}{2\beta(\mu - r)}.$$

Hence, the function θ is given by

$$\theta(r) = -cr + \frac{a(\beta\gamma - 1)}{2} - \frac{a\sqrt{(-4r\beta\gamma + (-1+\beta\gamma)^2(\mu - r))(\mu - r)}}{2(\mu - r)}.$$

Solving $\theta(r) = 0$ to obtain the adjustment coefficient R gives us three solutions. Namely,

$$\begin{aligned}
r_1 &= 0,\\
r_2 &= \frac{-a + a\beta\gamma + c\mu - \sqrt{(a(1+\beta\gamma))^2 - 2ac(-1+\beta\gamma)\mu + c^2\mu^2}}{2c},\\
r_3 &= \frac{-a + a\beta\gamma + c\mu + \sqrt{(a(1+\beta\gamma))^2 - 2ac(-1+\beta\gamma)\mu + c^2\mu^2}}{2c}.
\end{aligned}$$

This seems surprising since, by Lemma 4.24, we know that θ is convex; hence, we would expect two roots. To resolve this puzzle, we take a closer look at the third root r_3 and see that

$$\begin{aligned}
r_3 = \frac{-a + a\beta\gamma + c\mu + \sqrt{(a(1+\beta\gamma))^2 - 2ac(-1+\beta\gamma)\mu + c^2\mu^2}}{2c} &= \\
\frac{-a + a\beta\gamma + c\mu + \sqrt{(a(-1+\beta\gamma) - c\mu)^2 + 4a^2\beta\gamma}}{2c} & \\
\geq \frac{-a + a\beta\gamma + c\mu + |(a(-1+\beta\gamma) - c\mu)|}{2c} & \\
\geq \frac{-a + a\beta\gamma + c\mu + (-a(-1+\beta\gamma) + c\mu)}{2c} &= \mu.
\end{aligned}$$

In the previous parts, θ was only defined in the interval $(-\infty, r_{\max})$. Since $\mu > r_{\max} = \frac{(\beta\gamma-1)^2\mu}{(\beta\gamma+1)^2}$, we see that the third root is not in the domain under consideration.
To ensure that the second root

$$R = r_2 = \frac{-a + a\beta\gamma + c\mu - \sqrt{(a(1+\beta\gamma))^2 - 2ac(-1+\beta\gamma)\mu + c^2\mu^2}}{2c}$$

(which is our adjustment coefficient) is in the domain, we must assume the additional condition

$$c < \frac{a(\beta\gamma+1)^2}{2(\beta\gamma-1)\mu}. \tag{4.38}$$

This requirement (4.38) corresponds exactly to (4.5) in Karabash and Zhu (2015), which was needed to show the convergence of the logarithm of the probability of ruin in the general Hawkes case. Further, there exists an $\varepsilon > 0$ such that $M_U(R+\varepsilon)$, $M_Y(-\alpha(R+\varepsilon))$ are finite, and hence Assumption 4.3 is satisfied.

Integrability condition of the recurrence times: Assumption 4.4

Finally, we have to check if Assumption 4.4 is satisfied. By the definition of the stopping times $\{\phi_i\}_{\in\mathbb{N}}$ at the beginning of Subsection 4.3.5 and (4.33), it suffices to show that, for fixed level $\lambda > a$, there exists some $r > R$ such that

$$\mathbb{E}_{\lambda}^{\mathbb{Q}^{(r)}}\left[e^{\theta(r)S_1^{\lambda}}\right] < +\infty. \tag{4.39}$$

To prove (4.39), we will use the ideas of the proof of Theorem 4.21, that is, we identify some constant q and $\tilde{\lambda} < \lambda$ such that

$$\mathbb{E}_{\lambda}^{\mathbb{Q}^{(r)}}\left[e^{qS_1^{\tilde{\lambda}}}\right] < +\infty \quad \text{and} \quad \mathbb{E}_{\tilde{\lambda}}^{\mathbb{Q}^{(r)}}\left[e^{qS_1^{\lambda}}\right] < +\infty.$$

For this, using the proof of Theorem 4.21, we recall that, under the stationary regime, the recurrence time S_1^{λ} is light-tailed and $\mathbb{E}_{\nu}^{\mathbb{Q}^{(r)}}\left[e^{qS_1^{\lambda}}\right] < +\infty$ for all $q < \beta(\lambda - a)p^{(r)}(\lambda)$, where $p^{(r)}(\lambda)$ denotes the density of the stationary distribution under the measure $\mathbb{Q}^{(r)}$. Due to the proof of Theorem 4.21, we have that for almost all $\tilde{\lambda} < \lambda$, that $\mathbb{E}_{\lambda}^{\mathbb{Q}^{(r)}}\left[e^{qS_1^{\lambda}}\right] < +\infty$, where $q < \min\left(\beta(\lambda - a)p^{(r)}(\lambda), \beta(\tilde{\lambda} - a)p^{(r)}(\tilde{\lambda})\right)$. Unfortunately, we have to show that this holds for $q = \theta(r)$, a quantity depending on r. Furthermore, we know that the exponential moment is finite for almost all $\tilde{\lambda}$, but we do not know which $\tilde{\lambda}$ does not satisfy this property. To bypass these problems, we aim to identify a lower bound K for $\beta(\tilde{\lambda} - a)p^{(r)}(\tilde{\lambda})$ which is independent of r and holds uniformly for $\tilde{\lambda} \in I$, where I is an interval containing λ. This would give us $\mathbb{E}_{\lambda}^{\mathbb{Q}^{(r)}}\left[e^{KS_1^{\lambda}}\right] < +\infty$, for all r, which would allow us to choose $r > R$ such that $\theta(r) < K$. Consequently, the necessary integrability condition (4.39) will be satisfied.
From the proof of Lemma 4.29 it follows that, under a measure $\mathbb{Q}^{(r)}$ for some arbitrary $r < r_{\max}$, the intensity process $\{\lambda_t\}_{t\geq 0}$ of a Markovian Hawkes process $\{N_t\}_{t\geq 0}$ with the baseline intensity $aM_U(r)M_Y(-\alpha(r))$, decay parameter β and shocks of the form $M_U(r)M_Y(-\alpha(r))\tilde{Y}$, where $\tilde{Y}$ has distribution

$$\frac{e^{-\alpha(r)y}}{M_Y(-\alpha(r))}F_Y(\mathrm{d}y) = \frac{1}{\gamma + \alpha(r)}e^{-(\gamma+\alpha(r))y}\,\mathrm{d}y,$$

that is, with the shocks that are exponentially distributed with the parameter

$$\gamma^{(r)} := \frac{\gamma + \alpha(r)}{M_U(r)M_Y(-\alpha(r))} = \frac{(\gamma + \alpha(r))^2(\mu - r)}{\gamma\mu}.$$

Hence, we can use the already determined stationary distribution inn (4.37) for Markovian Hawkes intensities with exponentially distributed shocks to conclude that our process $\{\lambda_t\}_{t\geq 0}$ has stationary density

$$
\begin{aligned}
p^{(r)}(\lambda) &= (\lambda - a)^{\frac{M_U(r)M_Y(-\alpha(r))a}{\beta}-1} \\
&\times \exp\left(-\frac{(\beta\gamma^{(r)}-1)M_U(r)M_Y(-\alpha(r))}{\beta}(\lambda - a)\right) \\
&\times (M_U(r)M_Y(-\alpha(r)))^{\frac{M_U(r)M_Y(-\alpha(r))a}{\beta}-1}\left(\frac{\beta\gamma^{(r)}-1}{\beta}\right)^{\frac{aM_U(r)M_Y(-\alpha(r))}{\beta}} \\
&\times \Gamma\left(\frac{M_U(r)M_Y(-\alpha(r))a}{\beta}\right)^{-1}.
\end{aligned}
$$

Choose now some $\varepsilon > 0$ such that $r_{\max} > R + \varepsilon$. Then, we see that the function $p^{(r)}(\lambda)$ is well defined for all $r \in [R, R+\varepsilon]$, continuous as a bivariate function $p(r,\lambda) := p^{(r)}(\lambda)$, and strictly positive. Consequently, we find that this function is uniformly bounded from below on $[R, R+\varepsilon] \times [\frac{\lambda+a}{2}, \lambda]$ by some positive constant K.
Recall that the function $\theta(r)$ is continuous and the adjustment condition R satisfies $\theta(R) = 0$. By this, we can choose some $r \in [R, R+\varepsilon]$ such that $\theta(r) < K\beta\left(\frac{\lambda-a}{2}\right)$. Therefore, it holds for this specific r that

$$
\mathbb{E}_{\lambda}^{\mathbb{Q}^{(\bar{r})}}\left[e^{\theta(\bar{r})S_1^{\lambda}}\right] < +\infty.
$$

Consequently, we have that all our assumptions are satisfied and, assuming (4.38), from Theorem 4.33 we can conclude that

$$
\lim_{u\to+\infty} \psi(u,\tilde{\lambda})e^{Ru} = C^{\tilde{\lambda}},
$$

for a positive constant $C^{\tilde{\lambda}}$.

Acknowledgments

This research was funded in whole, or in part, by the Austrian Science Fund (FWF) P 33317. For the purpose of open access, the author has applied a CC BY public copyright licence to any Author Accepted Manuscript version arising from this submission.

5 The Markovian shot-noise risk model: a numerical method for Gerber-Shiu functions

The following chapter was published as Pojer and Thonhauser (2023b) and is adopted almost verbatim. Some changes have been made to ensure consistency of notation throughout all chapters of this thesis.

5.1 Introduction and overview

The introduction of the family of penalty functions by Gerber and Shiu in Gerber and Shiu (1998) had and still has a huge impact on the field of ruin theory. This unifying approach generalizes previously considered risk measures and allows a comprehensive analysis of the ruin event of an insurance portfolio. Since then, Gerber-Shiu functions were analysed in different types of risk models. For example in the renewal model in Gerber and Shiu (2005), Li and Garrido (2005) and Willmot and Dickson (2003), the Markov modulated model in Zhang (2008), and the Björk-Grandell model in Schmidli (2010). The case of spectrally negative Lévy risk processes was already considered in Garrido and Morales (2006) and resolved in a very general form by the so-called quintuple law derived in Doney and Kyprianou (2006).

Initially, the main aim was to establish explicit formulas, which allow for direct calculation of discounted penalty functions. This was successfully done in the classical and renewal models if the claim sizes are exponentially or phase-type distributed. Due to the increasing complexity of underlying models and considered penalty functions, this is generally hardly possible nowadays. Since simulation techniques like (quasi-)Monte Carlo methods are time-consuming and not always directly implementable, there is an increasing effort in finding efficient numerical procedures to determine suitable approximations of penalty functions for more complex models. Exemplary contributions are Chau et al. (2015), Diko and Usábel (2011), Lee et al. (2021), and Preischl et al. (2018). For the renewal risk model, Strini and Thonhauser (2020) introduced a numerical scheme based on a discretization of the corresponding generator to determine discounted penalty functions depending on a local cost functional and the deficit at ruin.

We consider Gerber-Shiu functions in the context of a Markovian shot-noise environment. The motivation for using the Markovian shot-noise model is the modelling of disasters, like earthquakes, as it was applied in Dassios and Jang (2003) in the context of pricing of reinsurance of catastrophic events. A generalized version of this model was considered by Albrecher and Asmussen (2006), who were interested in the asymptotic behaviour of the ruin probability in a general shot-noise model and derived exponentially decaying upper and lower bounds. Further extensions of this model are Stabile and Torrisi (2010), who considered heavy-tailed claim events, and Macci and Torrisi (2011), considering a non-constant premium rate. Recently, Pojer and Thonhauser (2023a) were able to show the convergence behaviour of the ruin probability in the Markovian model.

In this contribution, we are able to deal with Gerber-Shiu functions in their full generality. The introduction of an additional process, allows us to include functions depending on the surplus before ruin. By the underlying structure of piecewise-deterministic Markov processes, we can represent these discounted penalty functions as solutions to Feynman-Kac type partial integro-differential equations. Since there is no evidently explicit solution to the resulting equations, we develop a scheme to solve these equations numerically. First, we resolve the problem of the unboundedness of the involved intensity process. In the second step, we discretize the bounded version of the partial integro-differential equations and solve the corresponding system of linear equations. The obtained numerical solutions correspond to Gerber-Shiu functions of approximating Markov chains with finite state space. Eventually, we use weak convergence on the Skorokhod space of càdlàg functions to obtain a convergence result for the determined function values.

This paper is organized in the following way. In Section 2, we define the considered model, the concept of Gerber-Shiu functions, and their analytic properties. In Section 3, we introduce families of auxiliary processes used to approximate the original PDMPs of the Markovian shot-noise model and motivate the proposed numerical scheme. In Section 4, we show convergence of the numerical approximation by exploiting convergence in distribution of processes over the space of càdlàg functions. Finally, in Section 5, we give examples that show the performance of the proposed numerical scheme.

5.2 Risk model and Gerber-Shiu functions

At first, we briefly introduce the considered Markovian shot-noise model as it is also used in Pojer and Thonhauser (2023a). For this, we consider a probability space $(\Omega, \mathcal{F}, \mathbb{P})$, which is assumed to be big enough to carry all of the subsequently defined stochastic objects.

Definition 5.1. Let λ and β be positive constants and $\{N_t^\rho\}_{t\geq 0}$ a homogeneous Poisson process with intensity ρ and jump times $\{T_i^\rho\}_{i\geq 1}$. Let further $\{Y_i\}_{i\geq 1}$ be i.i.d. copies of a positive random variable Y with distribution F_Y independent of N^ρ. Then, the intensity process $\{\lambda_t\}_{t\geq 0}$ given by

$$\lambda_t := \lambda e^{-\beta t} + \sum_{i=1}^{N_t^\rho} Y_i e^{-\beta(t-T_i^\rho)},$$

is called Markovian shot-noise process.

Using this, we can define the surplus process in the following way.

Definition 5.2. Let N be a Cox process whose stochastic intensity is a Markovian shot-noise process $\{\lambda_t\}_{t\geq 0}$ and $\{U_i\}_{i\geq 1}$ an i.i.d. sequence of positive random variables, independent of $\{N_t\}_{t\geq 0}$ and $\{\lambda_t\}_{t\geq 0}$, with distribution function F_U. Let further c be a positive and x a non-negative constant. Then, the surplus process $\{X_t\}_{t\geq 0}$ is given by

$$X_t = x + ct - \sum_{i=1}^{N_t} U_i.$$

The Markovian shot-noise model was used by Dassios and Jang (2003) to model catastrophic events like earthquakes. A single catastrophic event, e.g. the earthquake, increases the intensity by a random quantity Y, called shock, and induces $Poi(Y/\beta)$ many claims, called a cluster, which do not occur immediately, but instead, they will be reported over a period of time. This allows us the following interpretation of the involved parameters and random variables. The parameter ρ is the inverse of the expected time between two catastrophic events. Since the total number of claims due to a single catastrophe is $Poi(Y/\beta)$, we have that the random variable Y/β determines the distribution of the number of claims in a single cluster. The decay parameter β determines how long it will take until all claims of the cluster are reported and paid. Despite the easy interpretation, it might be hard to estimate these quantities, especially the distribution of Y.
From now on, we will assume the following:

Assumption 5.1. *Assume that the net profit condition* $c > \frac{\rho}{\beta}\mathbb{E}[U_1]\,\mathbb{E}[Y_1]$ *holds.*

In many cases, the ruin probability itself is not a satisfying measure of the risk in the given model. One well-established much more general approach is to use Gerber-Shiu functions. For the precise setup, we follow the presentation used in Schmidli (2010). Let $w(x, y)$ be a continuous and bounded function and $\kappa > 0$. Then, the corresponding Gerber-Shiu function is defined by

$$g_\kappa(x, \lambda) := \mathbb{E}_{(x,\lambda)}\left[e^{-\kappa\tau} w(X_{\tau-}, -X_\tau) I_{\{\tau<\infty\}}\right],$$

for $(x, \lambda) \in [0, \infty) \times (0, \infty)$. Further, $g_\kappa(x, \lambda) = 0$ for $x < 0$ or $\lambda \leq 0$. Here, the expectation $\mathbb{E}_{(x,\lambda)}$ is the expectation with conditions $X_0 = x$ and $\lambda_0 = \lambda$. Even though this representation is commonly used, it is not satisfying in our case. Since we want to exploit weak convergence of càdlàg processes to justify our numerical scheme, we have to extend the definition of GS-functions.

Definition 5.3 (Gerber-Shiu function)**.** Let $\{\lambda_t\}_{t\geq 0}$ denote the Markovian shot-noise process and $\{X_t\}_{t\geq 0}$ the corresponding surplus process. As a third process define $\{m_t\}_{t\geq 0}$ as $m_t := U_{N_t}$, the process which remembers the size of the latest claim. Using these processes we define for a continuous and bounded function w, and a constant $\kappa > 0$ the Gerber-Shiu function

$$g_\kappa(x, m, \lambda) = \mathbb{E}_{(x,m,\lambda)}\left[w\left(X_\tau + m_\tau, -X_\tau\right) e^{-\kappa\tau} I_{\{\tau<\infty\}}\right].$$

As already mentioned before, the main advantage of this representation is, that we use the càdlàg process $\{m_t\}_{t\geq 0}$ instead of the làdcàg process $\{X_{t-}\}_{t\geq 0}$. Given the actual level of the surplus process, the distribution of m_τ does not depend on the current level m, i.e. the size of the latest claim. Therefore, we will omit m in future and write still $g_\kappa(x, \lambda)$ for the GS-function. For the sake of completeness, we define the filtration $\{\mathcal{F}_t\}_{t\geq 0}$ to be the natural filtration of the multivariate process $\{(X_t, m_t, \lambda_t)\}_{t\geq 0}$. In this setting, the multivariate process is a strong Markov process with respect to this filtration.

To obtain a partial integro-differential equation which is satisfied by the GS-functions, we use the Markovian structure of our model. The process $\{(X_t, m_t, \lambda_t)\}_{t\geq 0}$ is a piecewise-deterministic Markov process (PDMP) with generator

$$\begin{aligned}\mathcal{A}f(x, m, \lambda) = c\frac{\partial f(x, m, \lambda)}{\partial x} - \beta\lambda\frac{\partial f(x, m, \lambda)}{\partial \lambda} + \lambda\int_0^\infty f(x-u, u, \lambda)\, F_U(\mathrm{d}u) \\ + \rho\int_0^\infty f(x, m, \lambda + y)\, F_Y(\mathrm{d}y) - (\lambda+\rho) f(x, m, \lambda),\end{aligned}$$

which is certainly well-defined for all bounded and continuously differentiable functions f. Since our Gerber-Shiu functions are generally not continuously differentiable, we use the general definition of the generator of a PDMP from Rolski et al. (1999). For a function f, the path-derivative is defined by

$$\delta_\phi f(x, m, \lambda) := \lim_{h\to 0}\frac{f(x + ch, m, \lambda e^{-\beta h}) - f(x, m, \lambda)}{h}.$$

Then, the domain of the generator of our PDMP consists of all functions f, which are path-differentiable a.e. and satisfy that for all $t \geq 0$,

$$\begin{aligned}\mathbb{E}_{(x,\lambda)}\Bigg[\sum_{i=1}^{N_t}\left|f(X_{T_i}, m_{T_i}, \lambda_{T_i}) - f(X_{T_i-}, m_{T_i-}, \lambda_{T_i-})\right| \\ + \sum_{i=1}^{N_t^\rho}\left|f(X_{T_i^\rho}, m_{T_i^\lambda}, \lambda_{T_i^\rho}) - f(X_{T_i^\rho-}, m_{T_i^\rho-}, \lambda_{T_i^\rho-})\right|\Bigg] < \infty.\end{aligned}$$

For such a function f, the generator is characterized by

$$\mathcal{A}f(x,m,\lambda) = \delta_\phi f(x,m,\lambda) + \lambda \int_0^\infty f(x-u,u,\lambda)\,F_U(\mathrm{d}u) + \rho \int_0^\infty f(x,m,\lambda+y)\,F_Y(\mathrm{d}y) - (\lambda+\rho) f(x,m,\lambda).$$

Theorem 5.1. *The Gerber-Shiu functions are in the domain of the generator of the PDMP* $\{(X_t, m_t, \lambda_t)\}_{t\geq 0}$.

Proof. To show this, we prove that the Gerber-Shiu functions are path-differentiable and bounded.

For the path-differentiability, we follow the line of arguments as given in Strini and Thonhauser (2020). Define for some deterministic $r > 0$ the bounded stopping time $\nu = r \wedge T_1$, where T_1 denotes the first jump-time of the PDMP. Doing this we get

$$\begin{aligned} g_\kappa(x,\lambda) &= \mathbb{E}_{(x,\lambda)}\left[e^{-\kappa\tau} w(X_\tau + m_\tau, -X_\tau) I_{\{\tau<\infty\}}\right] \\ &= \mathbb{E}_{(x,\lambda)}\left[e^{-\kappa\nu}\mathbb{E}\left[e^{-\kappa(\tau-\nu)} w(X_\tau + m_\tau, -X_\tau) I_{\{\tau<\infty\}} \middle| \mathcal{F}_\nu\right]\right]. \end{aligned}$$

Now, there are two cases. Either $\nu = r$ or $\nu = T_1$. Using this, we get

$$g_\kappa(x,\lambda) =: e^{-\int_0^r (\lambda e^{-\beta s}+\rho)\,\mathrm{d}s} e^{-\kappa r} g_\kappa(x+cr, \lambda e^{-\beta r}) + \int_0^r H(s)\,\mathrm{d}s,$$

where

$$H(s) = (\lambda e^{-\beta s}+\rho) e^{-\int_0^s (\lambda e^{-\beta u}+\rho)\,\mathrm{d}u} e^{-\kappa s}\left(\frac{\rho}{\lambda e^{-\beta s}+\rho}\int_0^\infty g_\kappa(x+cs, \lambda e^{-\beta s}+y)\,F_Y(\mathrm{d}y) + \frac{\lambda e^{-\beta s}}{\lambda e^{-\beta s}+\rho}\left(\int_0^{x+cs} g_\kappa(x+cs-u, \lambda e^{-\beta s})\,F_U(\mathrm{d}u) + \int_{x+cs}^\infty w(x+cs, u-x-cs)\,F_U(\mathrm{d}u)\right)\right).$$

Adding and subtracting $e^{-\int_0^r (\lambda e^{-\beta s}+\rho+\kappa)\,\mathrm{d}s} g_\kappa(x,\lambda)$ and rearranging gives us

$$\frac{g_\kappa(x+cr,\lambda e^{-\beta r}) - g_\kappa(x,\lambda)}{r} = \frac{e^{\int_0^r (\lambda e^{-\beta s}+\rho+\kappa)\,\mathrm{d}s} - 1}{r} g_\kappa(x,\lambda) - \frac{e^{\int_0^r (\lambda e^{-\beta s}+\rho+\kappa)\,\mathrm{d}s}}{r}\int_0^r H(s)\,\mathrm{d}s.$$

The integral over H is differentiable in $r = 0$ from the right with derivative $H(0)$. Hence, for $r \to 0$ we have

$$\lim_{r\to 0} \frac{g_\kappa(x+cr, \lambda e^{-\beta r}) - g_\kappa(x,\lambda)}{r} = (\lambda+\rho+\kappa)g_\kappa(x,\lambda) + H(0)$$

$$=(\lambda+\rho+\kappa)g_\kappa(x,\lambda) - \rho\int_0^\infty g_\kappa(x,\lambda+y)\,F_Y(\mathrm{d}y)$$

$$-\lambda\left(\int_0^x g_\kappa(x-u,\lambda)\,F_U(\mathrm{d}u) + \int_x^\infty w(x,u-x)\,F_U(\mathrm{d}u)\right),$$

which gives us the differentiability of g along the paths of our PDMP. The integrability is an immediate consequence of the boundedness of w. □

If we use the derived form of the path-derivative of g_κ in the definition of the generator, we see that the GS-function solves the partial integro-differential equation

$$\mathcal{A}g_\kappa(x,\lambda) = \kappa g_\kappa(x,\lambda) - \lambda\int_x^\infty \left(w(x,u-x) - g_\kappa(x-u,\lambda)\right)\,F_U(\mathrm{d}u),$$

on $(x,\lambda)\in[0,\infty)\times(0,\infty)$. But, we still have to show that it is its unique solution.

Theorem 5.2. *Let g_κ be a GS-function with some $\kappa > 0$. Then g_κ is the unique bounded solution to the partial integro-differential equation (PIDE)*

$$\delta_\phi f(x,\lambda) + \rho\int_0^\infty f(x,\lambda+y)\,F_Y(\mathrm{d}y)$$

$$+\lambda\left(\int_0^x f(x-u,\lambda)\,F_U(\mathrm{d}u) + \int_x^\infty w(x,u-x)\,F_U(\mathrm{d}u)\right) - (\kappa+\lambda+\rho)f(x,\lambda) = 0,$$

for $(x,\lambda)\in[0,\infty)\times(0,\infty)$.

Proof. As already shown, the GS-function is bounded and solves the equation stated above. Now, observe that the PIDE can be rewritten in terms of the generator of the PDMP by

$$\mathcal{A}f(x,\lambda) - \kappa f(x,\lambda) - \lambda\int_x^\infty \left(f(x-u,\lambda) - w(x,u-x)\right)F_U(\mathrm{d}u) = 0,$$

for $(x,\lambda)\in[0,\infty)\times(0,\infty)$. Let $h:\mathbb{R}^2\to\mathbb{R}$ be an arbitrary bounded solution of this equation, $\kappa > 0$, and S a bounded stopping time. Since h is path-differentiable and bounded, it is in the domain of the generator $\mathcal{A}$, which we use, to get

$$h(x,\lambda) = \mathbb{E}_{(x,\lambda)}\left[e^{-\kappa S}h(X_S,\lambda_S) - \int_0^S e^{-\kappa v}(\mathcal{A}h(X_v,\lambda_v) - \kappa h(X_v,\lambda_v))\mathrm{d}v\right] =$$

$$\mathbb{E}_{(x,\lambda)}\left[e^{-\kappa S}h(X_S,\lambda_S) - \int_0^S e^{-\kappa v}\lambda_v\int_{X_v}^\infty (h(X_v-u,\lambda_v) - w(X_v,u-X_v))\,F_U(\mathrm{d}u)\,\mathrm{d}v\right].$$

Using this representation for the bounded stopping time $\tau \wedge t$, yields

$$\begin{aligned} h(x,\lambda) =& \mathbb{E}_{(x,\lambda)}\left[e^{-\kappa\tau\wedge t}h(X_{\tau\wedge t},\lambda_{\tau\wedge t})\right] \\ &- \mathbb{E}_{(x,\lambda)}\left[\int_0^{\tau\wedge t} e^{-\kappa v}\lambda_v \int_{X_v}^{\infty} h(X_v - u,\lambda_v)\,F_U(\mathrm{d}u)\,\mathrm{d}v\right] \\ &+ \mathbb{E}_{(x,\lambda)}\left[\int_0^{\tau\wedge t} e^{-\kappa v}\lambda_v \int_{X_v}^{\infty} w(X_v, u - X_v)\,F_U(\mathrm{d}u)\,\mathrm{d}v\right]. \end{aligned}$$

Let us now focus on the third term. Using that $\{N_t\}_{t\geq 0}$ is a counting process with intensity $\{\lambda_t\}_{t\geq 0}$, we can rewrite this to

$$\begin{aligned} &\mathbb{E}_{(x,\lambda)}\left[\int_0^{\tau\wedge t} e^{-\kappa v}\lambda_v \int_{X_v}^{\infty} w(X_v, u - X_v)\,F_U(\mathrm{d}u)\,\mathrm{d}v\right] = \\ &\mathbb{E}_{(x,\lambda)}\left[\int_{(0,\tau\wedge t]} e^{-\kappa v} \int_{X_{v-}}^{\infty} w(X_{v-}, u - X_{v-})\,F_U(\mathrm{d}u)\,\mathrm{d}N_v\right] = \\ &\mathbb{E}_{(x,\lambda)}\left[\sum_{i=1}^{N_{\tau\wedge t}} e^{-\kappa T_i} \int_{X_{T_i-}}^{\infty} w(X_{T_i-}, u - X_{T_i-})\,F_U(\mathrm{d}u)\right] = \\ &\mathbb{E}_{(x,\lambda)}\left[\sum_{i=1}^{N_{\tau\wedge t}} e^{-\kappa T_i} \mathbb{E}_{(x,\lambda)}\left[w(X_{T_i-}, U_i - X_{T_i-}) I_{\{U_i > X_{T_i-}\}} \,|\, \mathcal{F}_{T_i-}\right]\right]. \end{aligned}$$

Since $I_{\{U_i > X_{T_i-}\}} = 0$ for all $T_i < \tau$ and 1 for $T_i = \tau$, the sum is 0 if $\tau > t$ and $e^{-\kappa\tau}\mathbb{E}_{(x,\lambda)}\left[w(X_{\tau-}, -X_\tau) \,|\, \mathcal{F}_{\tau-}\right]$ if $\tau \leq t$. Consequently,

$$\begin{aligned} &\mathbb{E}_{(x,\lambda)}\left[\sum_{i=1}^{N_{\tau\wedge t}} e^{-\kappa T_i} \mathbb{E}_{(x,\lambda)}\left[w(X_{T_i-}, U_i - X_{T_i-}) I_{\{U_i > X_{T_i-}\}} \,|\, \mathcal{F}_{T_i-}\right]\right] = \\ &\mathbb{E}_{(x,\lambda)}\left[\sum_{i=1}^{N_{\tau\wedge t}} e^{-\kappa T_i} \mathbb{E}_{(x,\lambda)}\left[w(X_{T_i-}, U_i - X_{T_i-}) I_{\{T_i = \tau\}} \,|\, \mathcal{F}_{T_i-}\right]\right] = \\ &\mathbb{E}_{(x,\lambda)}\left[e^{-\kappa\tau} w(X_{\tau-}, -X_\tau) I_{\{\tau \leq t\}}\right]. \end{aligned}$$

The same arguments yield

$$\mathbb{E}_{(x,\lambda)}\left[\int_0^{\tau\wedge t} e^{-\kappa v}\lambda_v \int_{X_v}^{\infty} h(X_v - u,\lambda_v)\,F_U(\mathrm{d}u)\,\mathrm{d}v\right] = \mathbb{E}_{(x,\lambda)}\left[e^{-\kappa\tau} h(X_\tau,\lambda_\tau) I_{\{\tau \leq t\}}\right].$$

Hence,

$$\begin{aligned}h(x,\lambda) =&\mathbb{E}_{(x,\lambda)}\left[e^{-\kappa\tau\wedge t}h(X_{\tau\wedge t},\lambda_{\tau\wedge t})\right]-\mathbb{E}_{(x,\lambda)}\left[e^{-\kappa\tau}h(X_\tau,\lambda_\tau)I_{\{\tau\le t\}}\right]\\&+\mathbb{E}_{(x,\lambda)}\left[e^{-\kappa\tau}w(X_{\tau-},-X_\tau)I_{\{\tau\le t\}}\right]\\=&\mathbb{E}_{(x,\lambda)}\left[e^{-\kappa t}h(X_t,\lambda_t)I_{\{\tau>t\}}\right]+\mathbb{E}_{(x,\lambda)}\left[e^{-\kappa\tau}w(X_{\tau-},-X_\tau)I_{\{\tau\le t\}}\right].\end{aligned}$$

Since h is bounded, we can find a positive constant K such that

$$\left|\mathbb{E}_{(x,\lambda)}\left[e^{-\kappa t}h(X_t,\lambda_t)I_{\{\tau>t\}}\right]\right|\le Ke^{-\kappa t}.$$

Using this, we finally get that

$$\begin{aligned}h(x,\lambda)&=\lim_{t\to\infty}\mathbb{E}_{(x,\lambda)}\left[e^{-\kappa t}h(X_t,\lambda_t)I_{\{\tau>t\}}\right]+\mathbb{E}_{(x,\lambda)}\left[e^{-\kappa\tau}w(X_{\tau-},-X_\tau)I_{\{\tau\le t\}}\right]\\&=\lim_{t\to\infty}\mathbb{E}_{(x,\lambda)}\left[e^{-\kappa\tau}w(X_{\tau-},-X_\tau)I_{\{\tau\le t\}}\right]\\&=\mathbb{E}_{(x,\lambda)}\left[e^{-\kappa\tau}w(X_{\tau-},-X_\tau)I_{\{\tau<\infty\}}\right].\end{aligned}$$

□

5.3 Auxiliary processes

As already shown in the previous section, the function g_κ satisfies a partial integro-differential equation (PIDE). Generally, this equation cannot be solved explicitly. Hence, we need some numerical scheme that allows us to calculate an approximation of the desired value. An intuitive way to do so is to bound the state space and suitably discretize the PIDE on the bounded domain. This results in a system of linear equations which we can solve. Our approach is to approximate a bounded version of the PDMP by Markov chains, determine the corresponding Gerber-Shiu functions and show that these converge to the original ones.

5.3.1 Bounded processes

The first step of the approximation procedure is to bound some components of the processes in a suitable way.

Definition 5.4. Let $b>0$, $\lambda_{max}(b)>0$, $U_{max}(b)>0$ and $Y_{max}(b)>0$ such that $\lim_{b\to\infty}\lambda_{max}(b)=\lim_{b\to\infty}U_{max}(b)=\lim_{b\to\infty}Y_{max}(b)=\infty$. Then, we define the bounded intensity process by

$$\lambda_t^{(b)}=\lambda e^{-\beta t}+\sum_{i=1}^{N_t^\rho}Y_i^{(b)}e^{-\beta(t-T_i^\rho)}.$$

Here, the distribution of the random variable $Y_j^{(b)}$ depends on the original j-th shock and the pre-jump location of the process $\{\lambda_t^{(b)}\}_{t\geq 0}$ in the following way: define the random variable $\bar{Y}_j = Y_j I_{\{Y_j \leq Y_{max}(b)\}} + Y_{max}(b) I_{\{Y_j > Y_{max}(b)\}}$, then the bounded shocks are given by

$$Y_j^{(b)} = \bar{Y}_j I_{\left\{\lambda^{(b)}_{T_j^\rho -} + \bar{Y}_j \leq \lambda_{max}(b)\right\}} + (\lambda_{max}(b) - \lambda^{(b)}_{T_j^\rho -}) I_{\left\{\lambda^{(b)}_{T_j^\rho -} + \bar{Y}_j > \lambda_{max}(b)\right\}}.$$

This might seem complicated, but it ensures, that our new shocks have bounded support and the whole process $\{\lambda_t^{(b)}\}_{t\geq 0}$ does not leave the bounded state space $(0, \lambda_{max}(b)]$. Given this new process, we define our new counting process $\{N_t^{(b)}\}_{t\geq 0}$ using the acceptance-rejection method, also called thinning. Let T be a jump time of the original counting process $\{N_t\}_{t\geq 0}$ and $U \sim U[0,1]$. We accept the jump time for the new jump process if $\frac{\lambda_T^{(b)}}{\lambda_T} \geq U$. That gives us that $\{N_t^{(b)}\}_{t\geq 0}$ is a Cox process with intensity $\{\lambda_t^{(b)}\}_{t\geq 0}$, whose jump times coincide with jump times of our original process. Defining the sequence of bounded claims by $U_j^{(b)} = U_j I_{\{U_j \leq U_{max}(b)\}} + U_{max}(b) I_{\{U_j > U_{max}(b)\}}$, we can further define

$$X_t^{(b)} = x + ct - \sum_{i=1}^{N_t^{(b)}} U_i^{(b)},$$

and $m_t^{(b)} = U^{(b)}_{N_t^{(b)}}$.

Even though, $\{Y_i^{(b)}\}_{i\in\mathbb{N}}$ is no longer an i.i.d. sequence, this new triplet of processes is again a PDMP with generator

$$\begin{aligned}
\mathcal{A}^{(b)} f(x,m,\lambda) &= \delta_\phi f(x,m,\lambda) + \lambda \int_0^{U_{max}(b)} f(x-u,u,\lambda)\, F_U(\mathrm{d}u) \\
&\quad + \lambda f(x - U_{max}(b), U_{max}(b), \lambda)\, \mathbb{P}\left[U > U_{max}(b)\right] \\
&\quad + \rho \int_0^{Y_{max}(b)} f(x,m,\min\{\lambda_{max}(b), \lambda + y\})\, F_Y(\mathrm{d}y) \\
&\quad + \rho\, f(x,m,\min\{\lambda_{max}(b), \lambda + Y_{max}(b)\}) \mathbb{P}\left[Y > Y_{max}(b)\right] \\
&\quad - (\lambda+\rho) f(x,m,\lambda),
\end{aligned}$$

or alternatively we write for convenience

$$\begin{aligned}
\mathcal{A}^{(b)} f(x,m,\lambda) &= \delta_\phi f(x,m,\lambda) + \lambda \int_{(0,U_{max}(b)]} f(x-u,u,\lambda)\, F_{U^{(b)}}(\mathrm{d}u) \\
&\quad + \rho \int_{(0,Y_{max}(b)]} f(x,m,\lambda+y)\, F_{Y^{(b)}}(\mathrm{d}y, \lambda) - (\lambda+\rho) f(x,m,\lambda).
\end{aligned}$$

Having this, we can now define the GS-function of the bounded process.

Definition 5.5. Let $g_\kappa(x,\lambda) = \mathbb{E}_{(x,\lambda)}\left[w(X_\tau + m_\tau, -X_\tau) e^{-\kappa\tau} I_{\{\tau<\infty\}}\right]$ be an arbitrary

Gerber-Shiu function. Then, we define the corresponding GS-function of the bounded process by

$$g_\kappa^{(b)}(x,\lambda) = \mathbb{E}_{(x,\lambda)}\left[w(X^{(b)}_{\tau^{(b)}} + m^{(b)}_{\tau^{(b)}}, -X^{(b)}_{\tau^{(b)}})e^{-\kappa\tau^{(b)}} I_{\left\{\tau^{(b)}<\infty\right\}}\right],$$

where $\tau^{(b)} = \inf\left\{t \geq 0 \,\middle|\, X_t^{(b)} \leq 0\right\}$.

We call these processes bounded, since the intensity process $\{\lambda_t^{(b)}\}_{t\geq 0}$ and the random variables $\{Y_i^{(b)}\}_{i\in\mathbb{N}}$ and $\{U_i^{(b)}\}_{i\in\mathbb{N}}$ are a.s. bounded. Despite this denomination, the multivariate process $\{(X_t^{(b)}, m_t^{(b)}, \lambda_t^{(b)})\}_{t\geq 0}$ is not bounded. The surplus process $\{X_t^{(b)}\}_{t\geq 0}$ is left unbounded since any change there would disturb the strictly monotone increasing drift. We could resolve this problem, by using external states, which could be introduced to change the deterministic flow. This would create an active boundary, i.e. an area where jumps occur deterministically. Unfortunately, this causes several problems in the proofs of weak convergence in Section 5.4.

5.3.2 Discrete state processes

Let us now fix some b, and $h > 0$. We set $N_U = \left\lfloor \frac{U_{max}(b)}{h} \right\rfloor$ and N_λ such that $\lim_{h\to 0} N_\lambda h = \infty$. Then we introduce a continuous-time Markov chain with countable state space approximating the bounded process the following way:

Definition 5.6. Define the state space

$$\{(x_i, x_l, \lambda_j) \,|\, i,l \in \mathbb{Z},\, 1 \leq j \leq N_\lambda\},$$

where $x_i = cih$, and $\lambda_j := \lambda_{max}(b)\exp(-\beta(N_\lambda - j)h)$, and the probabilities $p_k^U = \mathbb{P}\left[x_{k-1} < U^{(b)} \leq x_k\right]$ for $k < N_U$ and $p_{N_U}^U = 1 - \sum_{k=1}^{N_U-1} p_k^U$.
Further, we set for some fixed λ_j

$$N_Y(j) = \#\{\lambda_{j+k} \,|\, k \geq 1,\, \lambda_{j+k} - \lambda_j \leq Y_{max}(b)\},$$

the number of points which we can reach from λ_j with a bounded shock event. The corresponding probabilities are

$$p_k^Y(j) = \mathbb{P}\left[\lambda_{j+k-1} - \lambda_j < Y^{(b)} \leq \lambda_{j+k} - \lambda_j\right] \text{ for } k < N_Y(j),$$

and $p^Y_{N^Y(j)}(j) := 1 - \sum_{k=1}^{N_Y(j)-1} p_k^Y(j)$.

Using this, we define the Markov chain on the discrete state space via its generator

$$
\begin{aligned}
\mathcal{A}^{(h,b)} f(x_i, x_l, \lambda_j) =& \frac{f(x_{i+1}, x_l, \lambda_{j-1}) - f(x_i, x_l, \lambda_j)}{h} \\
&+ \lambda_j \sum_{k=1}^{N_U} f(x_i - x_k, x_k, \lambda_j)\, p_k^U \\
&+ \rho \sum_{k=1}^{N_Y(j)} f(x_i, x_l, \lambda_{j+k})\, p_k^Y(j) - (\lambda_j + \rho) f(x_i, x_l, \lambda_j),
\end{aligned}
$$

where we set $\lambda_{1-1} = \lambda_1$ and $\lambda_n = \lambda_{max}(b)$ for all $n \geq N_\lambda$.

This generator consists still of infinitely many expressions, due to the unbounded state-space. To bypass this, we have to introduce a third family of processes with finite state-space.

Definition 5.7. Let $\bar{x}$ be a positive constant and define $N_x = \left\lfloor \frac{\bar{x}}{ch} \right\rfloor$. Then, we define the finite state space by

$$
\left\{ (x_i, x_l, \lambda_j) \mid -N_x \leq i, l \leq N_x,\ 1 \leq j \leq N_\lambda \right\},
$$

where x_i and λ_j are as in the countable case. On this grid, we define the Markov chain $\left\{ (X^{(\bar{x},h,b)}, m^{(\bar{x},h,b)}, \lambda^{(\bar{x},h,b)}) \right\}$ by its generator. For $i < N_x$ and $j > 1$ set

$$
\begin{aligned}
\mathcal{A}^{(\bar{x},h,b)} f(x_i, x_l, \lambda_j) =& \frac{f(x_{i+1}, x_l, \lambda_{j-1}) - f(x_i, x_l, \lambda_j)}{h} \\
&+ \lambda_j \sum_{k=1}^{\min(N_U, N_x + i)} f(x_{i-k}, x_k, \lambda_j)\, p_k^U \\
&+ \lambda_1 f(x_{-N_x}, \lambda_j) \left(1 - \sum_{k=1}^{\min(N_U, N_x + i)} p_k^U \right) \\
&+ \rho \sum_{k=1}^{N_Y(j)} f(x_i, x_l, \lambda_{j+k})\, p_k^Y(j) - (\lambda_j + \rho) f(x_i, x_l, \lambda_j).
\end{aligned}
$$

For $i = N_x$ and $j > 1$ set

$$\begin{aligned}\mathcal{A}^{(\bar{x},h,b)} f(x_{N_x}, x_l, \lambda_j) =& \frac{f(x_{N_x}, x_l, \lambda_{j-1}) - f(x_{N_x}, x_l, \lambda_j)}{h} \\ &+ \lambda_j \sum_{k=1}^{\min(N_U, 2N_x)} f(x_{N_x-k}, x_k, \lambda_j)\, p_k^U \\ &+ \lambda_j f(x_{-N_x}, \lambda_j) \left(1 - \sum_{k=1}^{\min(N_U, 2N_x)} p_k^U\right) \\ &+ \rho \sum_{k=1}^{N_Y(j)} f(x_{N_x}, x_l, \lambda_{j+k})\, p_k^Y(j) - (\lambda_j + \rho) f(x_{N_x}, x_l, \lambda_j).\end{aligned}$$

For $i = N_x$ and $j = 1$ set

$$\begin{aligned}\mathcal{A}^{(\bar{x},h,b)} f(x_{N_x}, x_l, \lambda_1) =& \lambda_1 \sum_{k=1}^{\min(N_U, 2N_x)} f(x_{N_x-k}, x_k, \lambda_1)\, p_k^U \\ &+ \lambda_1 f(x_{-N_x}, \lambda_1) \left(1 - \sum_{k=1}^{\min(N_U, 2N_x)} p_k^U\right) \\ &+ \rho \sum_{k=1}^{N_Y(j)} f(x_{N_x}, x_l, \lambda_{1+k})\, p_k^Y(1) - (\lambda_1 + \rho) f(x_{N_x}, x_l, \lambda_1).\end{aligned}$$

Here, we write again $\lambda_{1-1} = \lambda_1$ and $\lambda_n = \lambda_{max}(b)$ for all $n \geq N_\lambda$.

Now, we can introduce the corresponding GS-function of the Markov chain with finite state space.

Lemma 5.3. *Let $g_\kappa(x, \lambda) = \mathbb{E}_{(x,\lambda)} \left[e^{-\kappa\tau} w(X_\tau + m_\tau, -X_\tau) I_{\{\tau<\infty\}}\right]$ be an arbitrary discounted penalty function of our original model. Then, we define the corresponding GS-function of the Markov chain with finite state space by*

$$g_\kappa^{(\bar{x},h,b)}(x_i, \lambda_j) = \mathbb{E}_{(x_i,\lambda_j)} \left[e^{-\kappa\tilde{\tau}} w(X_{\tilde{\tau}}^{(\bar{x},h,b)} + m_{\tilde{\tau}}^{(\bar{x},h,b)}, -X_{\tilde{\tau}}^{(\bar{x},h,b)}) I_{\{\tilde{\tau}<\infty\}}\right],$$

where $\tilde{\tau} = \inf \left\{t \geq 0 \,\middle|\, X_t^{(\bar{x},h,b)} \leq 0\right\}$. This function $g_\kappa^{(\bar{x},h,b)}(x_i, \lambda_j)$ is the unique solution *of the following finite system of linear equations:*

$$\frac{f(x_{i+1},\lambda_{j-1})-f(x_i,\lambda_j)}{h}-(\lambda_j+\rho+\kappa)f(x_i,\lambda_j)+\lambda_j\sum_{k=1}^{\min(i-1,N_U)}f(x_i-x_k,\lambda_j)\,p_k^U$$
$$+\lambda_j I_{\{i\le N_U\}}\sum_{k=i}^{N_U}w(x_i,x_k-x_i)\,p_k^U+\rho\sum_{k=1}^{N_Y(j)}f(x_i,\lambda_{j+k})\,p_k^Y(j)=0 \text{ for } i<N_x,\ j>1,$$

$$\frac{f(x_{N_x},\lambda_{j-1})-f(x_{N_x},\lambda_j)}{h}+\lambda_j\sum_{k=1}^{N_U}f(x_{N_x}-x_k,\lambda_j)\,p_k^U$$
$$+\rho\sum_{k=1}^{N_Y(j)}f(x_{N_x},\lambda_{j+k})\,p_k^Y(j)-(\lambda_j+\rho+\kappa)f(x_{N_x},\lambda_j)=0 \text{ for } j>1,$$

$$\frac{f(x_{i+1},\lambda_1)-f(x_i,\lambda_1)}{h}+\lambda_1\sum_{k=1}^{\min(i-1,N_U)}f(x_i-x_k,\lambda_1)\,p_k^U+\lambda_1 I_{\{i\le N_U\}}\sum_{k=i}^{N_U}w(x_i,x_k-x_i)\,p_k^U$$
$$+\rho\sum_{k=1}^{N_Y(1)}f(x_i,\lambda_{1+k})\,p_k^Y(1)-(\lambda_1+\rho+\kappa)f(x_i,\lambda_1)=0 \text{ for } i<N_x,$$

and

$$\lambda_1\sum_{k=1}^{N_U}f(x_{N_x}-x_k,\lambda_1)\,p_k^U+\rho\sum_{k=1}^{N_Y(1)}f(x_{N_x},\lambda_{1+k})\,p_k^Y(1)-(\lambda_1+\rho+\kappa)f(x_{N_x},\lambda_1)=0.$$

Proof. Since $\kappa>0$, the matrix corresponding to the above system of equations is strict diagonally dominant, hence regular. Since the GS-function solves the system, it is the unique solution.

□

5.4 Convergence of Gerber-Shiu functions

In this section, we will prove that our numerical scheme converges as $h\to 0$ and $b\to\infty$. For this, we want to exploit the convergence in distribution of processes as random variables on the Skorokhod space of càdlàg functions. This convergence implies the convergence of Skorokhod-continuous and bounded functionals of the corresponding processes. For further details on this metric space see Chapter 3 of Ethier and Kurtz (2009).
Since our processes are Markov processes, the main idea is to reduce this to the convergence of the corresponding generators. For Feller processes, these properties are

equivalent as shown in Theorem 19.25 of Kallenberg (2002). Since our processes are not Feller, we will use Theorem 8.2 in Chapter 5 of Ethier and Kurtz (2009) to show the same. Consequently, we have to find a suitable subdomain of our generators such that the induced semigroup is strongly continuous on this set of functions. If this domain is convergence determining, e.g. if it contains C_c^∞, and the generators converge for all f from this domain, then the corresponding processes converge weakly.

5.4.1 Convergence of the bounded processes

Lemma 5.4. *The generator $\mathcal{A}$ of the original PDMP generates a strongly continuous contraction semigroup $\{T_t\}_{t\geq 0}$ on the $\|\cdot\|_\infty$-closure of the set*

$$\mathcal{D} = \{f \in C_b \mid \delta_\phi f \text{ is path-continuous and } \mathcal{A}f \in C_b\},$$

by $T_t f(x,m,\lambda) := \mathbb{E}_{(x,m,\lambda)}\left[f(X_t, m_t, \lambda_t)\right]$.

Proof. Since $\mathcal{A}$ is the generator of the Markov process $\{(X_t, m_t, \lambda_t)\}_{t\geq 0}$, we have to show that T_t maps this set into itself and is strongly continuous there.
By a small modification of the proof of Theorem 27.6 in Davis (1993), we can relax the needed assumption that the intensity is bounded. This gives us that for all bounded and continuous f, we have that $T_t f \in C_b$ too. By Theorem 7.7.4 of Jacobsen (2006), the operators map path-differentiable functions satisfying $\|\mathcal{A}f\|_\infty$ into itself and satisfy $\mathcal{A}T_t f = T_t\mathcal{A}f$. By this, we get for all $f \in \mathcal{D}$ that $\mathcal{A}T_t f = T_t\mathcal{A}f \in C_b$.
The strong continuity is an immediate consequence of the boundedness of $\mathcal{A}f$. Consider $|T_t f(x,m,\lambda) - f(x,m,\lambda)|$ for some fixed t. Then, it holds that

$$|T_t f(x,m,\lambda) - f(x,m,\lambda)| = \left|\int_0^t \mathbb{E}_{(x,m,\lambda)}\left[\mathcal{A}f(X_s, m_s, \lambda_s)\right] \mathrm{d}s\right| \leq t\|\mathcal{A}f\|_\infty.$$

Since this upper bound is independent of (x,m,λ), we can let t tend to 0, which gives us that the contraction semigroup T_t is strongly continuous in $t = 0$. ☐

Theorem 5.5. *Let $f \in \mathcal{D}$ be arbitrary and $g = \mathcal{A}f$. Then, for all $k \geq 0$, $0 \leq t_1 < t_2 < \cdots < t_k \leq t < t+s$ and $h_1, \cdots, h_k \in C_b$ we have that*

$$\lim_{b\to\infty} \mathbb{E}_{(x,m,\lambda)}\Bigg[\bigg(f(X_{t+s}^{(b)}, m_{t+s}^{(b)}, \lambda_{t+s}^{(b)}) \;-\; f(X_t^{(b)}, m_t^{(b)}, \lambda_t^{(b)}) - \int_t^{t+s} g(X_v^{(b)}, m_v^{(b)}, \lambda_v^{(b)})\,\mathrm{d}v\bigg) \times \prod_{i=1}^k h_i(X_{t_i}^{(b)}, m_{t_i}^{(b)}, \lambda_{t_i}^{(b)})\Bigg] = 0.$$

Proof. For convenience, we will write $Z_t := (X_t, m_t, \lambda_t)$, $Z_t^{(b)} := (X_t^{(b)}, m_t^{(b)}, \lambda_t^{(b)})$ and $z = (x,m,\lambda)$. At first, we will cover the case $k = 0$. Let $f \in \mathcal{D}$ arbitrary and $g = \mathcal{A}f$. It is easy to see, that for every $b > 0$, f is in the domain of the generator $\mathcal{A}^{(b)}$ too. Writing

$g^{(b)}$ for $\mathcal{A}^{(b)}f$ we have that

$$\begin{aligned}\mathbb{E}_z\left[f(Z_{t+s}^{(b)}) - f(Z_t^{(b)}) - \int_t^{t+s} g(Z_v^{(b)})\,\mathrm{d}v\right] =& \mathbb{E}_z\left[f(Z_{t+s}^{(b)}) - f(Z_t^{(b)}) - \int_t^{t+s} g^{(b)}(Z_v^{(b)})\,\mathrm{d}v\right] \\ &+ \mathbb{E}_z\left[\int_t^{t+s} g^{(b)}(Z_v^{(b)}) - g(Z_v^{(b)})\,\mathrm{d}v\right].\end{aligned}$$

The first term is the expectation of a zero mean martingale, hence 0. For the second term, we take a closer look at the difference of the generators $\mathcal{A}$ and $\mathcal{A}^{(b)}$ applied to the same function f in the same point (x, m, λ):

$$\begin{aligned}\mathcal{A}f(x,m,\lambda) =& \delta_\phi f(x,m,\lambda) + \lambda \int_0^\infty f(x-u,u,\lambda)\,F_U(\mathrm{d}u) \\ &+ \rho \int_0^\infty f(x,m,\lambda+y)\,F_Y(\mathrm{d}y) - (\lambda+\rho) f(x,m,\lambda)\end{aligned}$$

and

$$\begin{aligned}\mathcal{A}^{(b)}f(x,m,\lambda) =& \delta_\phi f(x,m,\lambda) + \lambda \int_0^{U_{max}(b)} f(x-u,u,\lambda)\,F_U(\mathrm{d}u) \\ &+ \rho \int_0^{Y_{max}(b)} f(x,m,\min\{\lambda_{max}(b), \lambda+y\})\,F_Y(\mathrm{d}y) \\ &+ \lambda\,\mathbb{P}\left[U > U_{max}(b)\right] f(x-U_{max}(b), U_{max}(b), \lambda) \\ &+ \rho\,\mathbb{P}\left[Y > Y_{max}(b)\right] f(x,m,\min\{\lambda_{max}(b), \lambda+Y_{max}(b)\}) \\ &- (\lambda+\rho) f(x,m,\lambda)\end{aligned}$$

The derivatives coincide and so do the integrals from 0 to $U_{max}(b)$ and 0 to $\min\{\lambda_{max}(b) - \lambda, Y_{max}(b)\}$ respectively. The absolute value of the remaining parts can be bounded by

$$\begin{aligned}&\left|\lambda \int_{U_{max}(b)}^\infty f(x-u,u,\lambda)\,F_U(\mathrm{d}u) - \lambda\,\mathbb{P}\left[U > U_{max}(b)\right] f(x-U_{max}(b), U_{max}(b), \lambda)\right| + \\ &\left|\rho \int_{\min\left\{\lambda_{max}^{(b)}-\lambda, Y_{max}(b)\right\}}^\infty f(x,m,\lambda+y) - \rho\,\mathbb{P}\left[Y > \min\left\{\lambda_{max}^{(b)} - \lambda, Y_{max}(b)\right\}\right] f(x,m,\lambda_{max}^{(b)})\right| \leq \\ &\qquad 2\lambda\|f\|_\infty \mathbb{P}\left[U > U_{max}(b)\right] + 2\rho\|f\|_\infty\,\mathbb{P}\left[Y > \min\left\{\lambda_{max}^{(b)} - \lambda, Y_{max}(b)\right\}\right].\end{aligned}$$

This upper bound tends to 0 as $b \to \infty$ since $U_{max}(b)$, $\lambda_{max}(b)$ and $Y_{max}(b)$ tend to infinity but not uniformly in λ.

If we now get back to our expectation we see that

$$\begin{aligned}
\left|\mathbb{E}_z\left[f(Z^{(b)}_{t+s}) - f(Z^{(b)}_t) - \int_t^{t+s} g(Z^{(b)}_v)\,\mathrm{d}v\right]\right| &\le \mathbb{E}_z\left[\int_t^{t+s}\left|g^{(b)}(Z^{(b)}_v) - g(Z^{(b)}_v)\right|\,\mathrm{d}v\right] \\
&\le 2\|f\|_\infty \rho \int_t^{t+s} \mathbb{P}_\lambda\left[Y > \min\left\{\lambda_{max}(b) - \lambda^{(b)}_v, Y_{max}(b)\right\}\right]\,\mathrm{d}v \\
&\quad + 2\|f\|_\infty \int_t^{t+s} \mathbb{E}_\lambda\left[\lambda^{(b)}_v\right]\mathbb{P}\left[U > U_{max}(b)\right]\,\mathrm{d}v \\
&\le 2\|f\|_\infty \rho \int_t^{t+s} \mathbb{P}_\lambda\left[Y > \min\left\{\lambda_{max}(b) - \lambda^{(b)}_v, Y_{max}(b)\right\}\right]\,\mathrm{d}v \\
&\quad + 2\|f\|_\infty \int_t^{t+s} \mathbb{E}_\lambda\left[\lambda_v\right]\mathbb{P}\left[U > U_{max}(b)\right]\,\mathrm{d}v \\
&\le 2\|f\|_\infty \rho \int_t^{t+s} \mathbb{P}_\lambda\left[Y > \min\left\{\lambda_{max}(b) - \lambda_v, Y_{max}(b)\right\}\right]\,\mathrm{d}v \\
&\quad + 2\|f\|_\infty \left(\lambda + \frac{\rho}{\beta}\mathbb{E}\left[Y\right]\right)\mathbb{P}\left[U > U_{max}(b)\right].
\end{aligned}$$

The second part tends to 0 as $b \to \infty$ but the first part still needs some work since it depends on λ_v. For this, we remember that, given $\lambda_0 = \lambda$, $\lambda_v - \lambda$ with $v \le s + t$ can be bounded from above by the compound Poisson distributed random variable $\sum_{i=1}^{N^\rho_{t+s}} Y_i$. By this we get that

$$\mathbb{P}_\lambda\left[Y > \min\left\{\lambda_{max}(b) - \lambda_v, Y_{max}(b)\right\}\right] \le \mathbb{P}\left[\lambda + Y + \sum_{i=1}^{N^\rho_{t+s}} Y_i > \min\left\{\lambda_{max}(b), Y_{max}(b)\right\}\right],$$

which is independent of v and tends to 0 as b tends to infinity. By this, we have that

$$\begin{aligned}
\lim_{b\to\infty}\left|\mathbb{E}_z\left[\int_t^{t+s} g^{(b)}(Z^{(b)}_v) - g(Z^{(b)}_v)\,\mathrm{d}v\right]\right| &\le \lim_{b\to\infty} 2\|f\|_\infty\left(\lambda + \frac{\rho}{\beta}\mathbb{E}\left[Y\right]\right)\mathbb{P}\left[U > U_{max}(b)\right] \\
&\quad + \lim_{b\to\infty} s\rho\,\mathbb{P}\left[\lambda + Y + \sum_{i=1}^{N^\rho_{t+s}} Y_i > \min\left\{\lambda_{max}(b), Y_{max}(b)\right\}\right] = 0.
\end{aligned}$$

For $k > 0$, we observe that the chosen time points $t_1, \ldots, t_k$ are prior to time t, hence

$h_i(Z^{(b)}_{t_i})$ is $\mathcal{F}^{Z^{(b)}}_t$ measurable. By this we have that

$$\mathbb{E}_z\left[\left(f(Z^{(b)}_{t+s}) - f(Z^{(b)}_t) - \int_t^{t+s} g^{(b)}(Z^{(b)}_v)\,\mathrm{d}v\right)\prod_{i=1}^k h_i(Z^{(b)}_{t_i})\right] =$$

$$\mathbb{E}_z\left[\mathbb{E}_z\left[\left(f(Z^{(b)}_{t+s}) - f(Z^{(b)}_t) - \int_t^{t+s} g^{(b)}(Z^{(b)}_v)\,\mathrm{d}v\right)\,\Big|\,\mathcal{F}^{Z^{(b)}}_t\right]\prod_{i=1}^k h_i(Z^{(b)}_{t_i})\right] = 0.$$

Therefore, similar to the case $k = 0$ we can rewrite the difference as

$$\mathbb{E}_z\left[\left(\int_t^{t+s} g^{(b)}(Z^{(b)}_v) - g(Z^{(b)}_v)\,\mathrm{d}v\right)\prod_{i=1}^k h_i(Z^{(b)}_{t_i})\right].$$

The functions h_i are in C_b, hence we can bound the absolute value of the product uniformly by some constant $\tilde{c}$ and get

$$\lim_{b\to\infty}\left|\mathbb{E}_z\left[\left(\int_t^{t+s} g^{(b)}(Z^{(b)}_v) - g(Z^{(b)}_v)\,\mathrm{d}v\right)\prod_{i=1}^k h_i(Z^{(b)}_{t_i})\right]\right|$$

$$\leq \lim_{b\to\infty}\tilde{c}\int_t^{t+s}\mathbb{E}_z\left[\left|g^{(b)}(Z^{(b)}_v) - g(Z^{(b)}_v)\right|\right]\,\mathrm{d}v = 0.$$

□

Corollary 5.6. *The process $\{(X^{(b)}_t, m^{(b)}_t, \lambda^{(b)}_t)\}_{t\geq 0}$ converges weakly against the original PDMP as $b \to \infty$.*

Proof. This is a direct consequence of Theorem 5.5 and Theorem 8.2 in Ethier and Kurtz (2009). □

5.4.2 Convergence of the discrete processes

Now we will use the same ideas as before, but on the set

$$\mathcal{D}^{(b)} := \left\{f \in \mathcal{D}\;\Big|\; f \text{ and } \delta_\phi f \text{ are Lipschitz and } \mathcal{A}^{(b)}f \in C_b\right\}.$$

Again we define a contraction semigroup $T^{(b)}_t f(x, m, \lambda) := \mathbb{E}_{(x,m,\lambda)}\left[f(X^{(b)}_t, m^{(b)}_t, \lambda^{(b)}_t)\right]$ and want to show that this semigroup is strongly continuous at 0 over the set $\mathcal{D}^{(b)}$.

Lemma 5.7. *Let f be in $\mathcal{D}^{(b)}$. Then, $T^{(b)}_t f$ and $\delta_\phi T^{(b)}_t f$ are Lipschitz continuous.*

Proof. Let $f \in \mathcal{D}^{(b)}$ be arbitrary and consider for $x \neq y$

$$\begin{aligned}&\left|T_t^{(b)} f(x,m,\lambda) - T_t^{(b)} f(y,m,\lambda)\right| \\ &\qquad = \left|\mathbb{E}_{(x,m,\lambda)}\left[f(X_t^{(b)}, m_t^{(b)}, \lambda_t^{(b)})\right] - \mathbb{E}_{(y,m,\lambda)}\left[f(X_t^{(b)}, m_t^{(b)}, \lambda_t^{(b)})\right]\right|.\end{aligned}$$

The altered initial condition in the first variable only affects the surplus process. Let $\{\tilde{X}_t^{(b)}\}_{t\geq 0}$ be the reserve process with initial capital y and $\{X_t^{(b)}\}_{t\geq 0}$ the corresponding process with starting value x. By the linear structure of the surplus process, we see that $\tilde{X}_t^{(b)}(\omega) = (y-x) + X_t^{(b)}(\omega)$ for all $t \geq 0$ and all $\omega \in \Omega$. By this we get that

$$\begin{aligned}&\left|\mathbb{E}_{(x,m,\lambda)}\left[f(X_t^{(b)}, m_t^{(b)}, \lambda_t^{(b)})\right] - \mathbb{E}_{(y,m,\lambda)}\left[f(X_t^{(b)}, m_t^{(b)}, \lambda_t^{(b)})\right]\right| \\ &= \left|\mathbb{E}_{(x,m,\lambda)}\left[f(X_t^{(b)}, m_t^{(b)}, \lambda_t^{(b)}) - f((y-x) + X_t^{(b)}, m_t^{(b)}, \lambda_t^{(b)})\right]\right| \\ &\leq \mathbb{E}_{(x,m,\lambda)}\left[\left|f(X_t^{(b)}, m_t^{(b)}, \lambda_t^{(b)}) - f((y-x) + X_t^{(b)}, m_t^{(b)}, \lambda_t^{(b)})\right|\right] \leq L\,|y-x|\,,\end{aligned}$$

where L denotes the Lipschitz constant of f with respect to $\|.\|_1$. The same idea leads to a preserved Lipschitz-continuity in the second variable.

Now we want to show that this holds for the third variable too. Here, things get a little more complicated, since small changes in the intensity process influence all three processes. Let us now consider $\{\lambda_t^{(b)}\}_{t\geq 0}$, the intensity process with initial condition $\lambda_0^{(b)} = \lambda$, and for some $h > 0$ the altered intensity process $\{\tilde{\lambda}_t^{(b)}\}_{t\geq 0}$ with starting value $\lambda + h$ and take a look at the difference of those processes. If no shock event appeared until time t, or shocks happened but $\{\tilde{\lambda}_s^{(b)}\}_{0\leq s\leq t}$ did not hit $\lambda_{max}(b)$, the relation between those processes is $\tilde{\lambda}_t^{(b)} = \lambda_t^{(b)} + he^{-\beta t}$. Otherwise, the difference decreases and may even become 0 if both, $\{\lambda_t^{(b)}\}_{t\geq 0}$ and $\{\tilde{\lambda}^{(b)}\}_{t\geq 0}$ hit $\lambda_{max}(b)$.

As already mentioned, the difference in the starting intensity leads to a change in the surplus process too. To be precise, we again consider two realizations, $\{\tilde{X}_t^{(b)}\}_{t\geq 0}$ with starting intensity $\lambda + h$ and $\{X_t^{(b)}\}_{t\geq 0}$ corresponding to $\lambda_0^{(b)} = \lambda$. They are related by $\tilde{X}_t^{(b)} = X_t^{(b)} - \sum_{i=1}^{\tilde{N}_t} \tilde{U}_i$, where $\{\tilde{N}_t\}_{t\geq 0}$ is a counting process with intensity $\tilde{\lambda}_t^{(b)} - \lambda_t^{(b)} \leq he^{-\beta t}$ and additional i.i.d. claims $\tilde{U}_i \sim U$ independent of all U_i.

Finally, the corresponding realizations $m_t^{(b)}$ and $m_t^{(b)}$ may relate in three different ways. The first case is that $N_t^{(b)} > 0$ and the last jump before time t is due to $\{\tilde{N}_t\}_{t\geq 0}$. In this case $m_t^{(b)}$ and $\tilde{m}_t^{(b)}$ are not equal but i.i.d. random variables. In the second case, $N_t^{(b)} = 0$ but $\tilde{N}_t$ is not. In this case $m_t^{(b)} = m$ and $\tilde{m}_t^{(b)} \sim \tilde{U}$. In the remaining case we have that $\tilde{m}_t^{(b)} = m_t^{(b)}$.

Having this in mind we now consider the following:

$$\begin{aligned}\left|T_t^{(b)}f(x,m,\lambda+h)-T_t^{(b)}f(x,m,\lambda)\right| &= \left|\mathbb{E}_{(x,m,\lambda)}\left[f(\tilde{X}_t^{(b)},\tilde{m}_t^{(b)},\tilde{\lambda}_t^{(b)})-f(X_t^{(b)},m_t^{(b)},\lambda_t^{(b)})\right]\right| \\ &\leq \mathbb{E}_{(x,m,\lambda)}\left[\left|f(\tilde{X}_t^{(b)},\tilde{m}_t^{(b)},\tilde{\lambda}_t^{(b)})-f(X_t^{(b)},\tilde{m}_t^{(b)},\tilde{\lambda}_t^{(b)})\right|\right] \\ &+ \left|\mathbb{E}_{(x,m,\lambda)}\left[f(X_t^{(b)},\tilde{m}_t^{(b)},\tilde{\lambda}_t^{(b)})-f(X_t^{(b)},m_t^{(b)},\tilde{\lambda}_t^{(b)})\right]\right| \\ &+ \mathbb{E}_{(x,m,\lambda)}\left[\left|f(X_t^{(b)},m_t^{(b)},\tilde{\lambda}_t^{(b)})-f(X_t^{(b)},m_t^{(b)},\lambda_t^{(b)})\right|\right].\end{aligned}$$

Since f is Lipschitz, the third term can be bounded by

$$\mathbb{E}_{(x,m,\lambda)}\left[\left|f(X_t^{(b)},m_t^{(b)},\tilde{\lambda}_t^{(b)})-f(X_t^{(b)},m_t^{(b)},\lambda_t^{(b)})\right|\right] \leq L\mathbb{E}_{(x,m,\lambda)}\left[\left|\tilde{\lambda}_t^{(b)}-\lambda_t^{(b)}\right|\right] \leq Lhe^{-\beta t}.$$

By the same arguments, we can bound the first term by

$$\mathbb{E}_{(x,m,\lambda)}\left[\left|f(\tilde{X}_t^{(b)},\tilde{m}_t^{(b)},\tilde{\lambda}_t^{(b)})-f(X_t^{(b)},\tilde{m}_t^{(b)},\tilde{\lambda}_t^{(b)})\right|\right] \leq L\mathbb{E}_{(x,m,\lambda)}\left[\left|\sum_{i=1}^{\tilde{N}_t}\tilde{U}_i\right|\right] \leq \frac{L}{\beta}\mathbb{E}\left[U\right]h.$$

The second term can be reduced to

$$\begin{aligned}&\left|\mathbb{E}_{(x,m,\lambda)}\left[f(X_t^{(b)},\tilde{m}_t^{(b)},\tilde{\lambda}_t^{(b)})-f(X_t^{(b)},m_t^{(b)},\tilde{\lambda}_t^{(b)})\right]\right| = \\ &\left|\mathbb{E}_{(x,m,\lambda)}\left[\left(f(X_t^{(b)},\tilde{U},\tilde{\lambda}_t^{(b)})-f(X_t^{(b)},m,\tilde{\lambda}_t^{(b)})\right)I_{\left\{N_t^{(b)}=0\right\}}I_{\{\tilde{N}_t>0\}}\right]\right| \\ &\leq 2\|f\|_\infty\mathbb{P}_\lambda\left[\tilde{N}_t>0\right] \leq 2\|f\|_\infty h\frac{1-\exp\left(-\frac{1-e^{-\beta t}}{\beta}h\right)}{h} \leq 2\|f\|_\infty h\frac{1-e^{-\beta t}}{\beta}.\end{aligned}$$

Using these results, we get that there exists a constant K such that

$$\left|T_t^{(b)}f(x,m,\lambda+h)-T_t^{(b)}f(x,m,\lambda)\right| \leq Kh,$$

for all positive λ and h. Consequently, $T_t^{(b)}f$ is Lipschitz for all $f\in\mathcal{D}^{(b)}$.

To show the Lipschitz continuity of the path-derivative $\delta_\phi T_t^{(b)}f$, we use the following representation derived in the proof of Theorem 7.7.4 in Jacobsen (2006):

$$\begin{aligned}\delta_\phi T_t^{(b)} f(x,m,\lambda) =& T_t^{(b)}\left(\mathcal{A}^{(b)} f\right)(x,m,\lambda)\\ &+\lambda\int_{(0,U_{max}(b)]} T_t^{(b)} f(x,m,\lambda) - T_t^{(b)} f(x-u,u,\lambda)\, F_{U^{(b)}}(\mathrm{d}u)\\ &+\rho\int_{(0,Y_{max}(b)]} T_t^{(b)} f(x,m,\lambda) - T_t^{(b)} f(x,m,\lambda+y)\, F_{Y^{(b)}}(\mathrm{d}y).\end{aligned}$$

Since $T_t^{(b)}$ preserves Lipschitz continuity, we know that the integral terms are indeed Lipschitz. Now we just have to show that for every $f \in \mathcal{D}^{(b)}$, the function $\mathcal{A}^{(b)} f$ is Lipschitz too. Let $z_1 := (x_1, m_1, \lambda_1)$ and $z_2 := (x_2, m_2, \lambda_2)$ two suitable points and consider

$$\begin{aligned}\left|\mathcal{A}^{(b)} f(z_1) - \mathcal{A}^{(b)} f(z_2)\right| = \Bigg|&\delta_\phi f(z_1) - \delta_\phi f(z_2) + \lambda_1\int_{(0,U_{max}(b)]} f(x_1-u,u,\lambda_1)\, F_{U^{(b)}}(\mathrm{d}u)\\ &-\lambda_2\int_{(0,U_{max}(b)]} f(x_2-u,u,\lambda_2)\, F_{U^{(b)}}(\mathrm{d}u)\\ &-\rho(f(z_1)-f(z_2)) - \lambda_1 f(z_1) + \lambda_2 f(z_2)\\ &+\rho\int_{(0,Y_{max}(b)]} f(x_1,m_1,\lambda_1+y) - f(x_2,m_2,\lambda_2+y)\, F_{Y^{(b)}}(\mathrm{d}y)\Bigg|\end{aligned}$$

Using the triangle inequality and the Lipschitz continuity of $\delta_\phi f$ and f we get that there is a constant K such that the above is less or equal to

$$\begin{aligned}&K\|z_1-z_2\|_1 + \lambda_1\left|\int_0^\infty f(x_1-u,u,\lambda_1) - f(x_2-u,u,\lambda_2)\, F_{U^{(b)}}(\mathrm{d}u)\right|\\ &+\left|(\lambda_1-\lambda_2)\int_0^\infty f(x_2-u,u,\lambda_2)\, F_{U^{(b)}}(\mathrm{d}u)\right| + \lambda_1\left|f(z_1)-f(z_2)\right| + \left|(\lambda_1-\lambda_2) f(z_2)\right|.\end{aligned}$$

For all u, the Lipschitz continuity of f gives us the existence of positive constants L and $\tilde{L}$ such that

$$\begin{aligned}\left|f(x_1-u,u,\lambda_1) - f(x_2-u,u,\lambda_2)\right| \le L\max\left\{\left|x_1-x_2\right|, \left|\lambda_1-\lambda_2\right|\right\} &\le L\|z_1-z_2\|_\infty\\ &\le \tilde{L}\|z_1-z_2\|_1,\end{aligned}$$

where the last inequality is given by the equivalence of norms in finite dimensional spaces. Further, we get that there is a constant $\tilde{c}$ with

$$|(\lambda_1-\lambda_2) f(z_2)| \le \|f\|_\infty \|z_1-z_2\|_\infty \le \tilde{c}\|z_1-z_2\|_1.$$

Using these inequalities and the boundedness of $\lambda^{(b)}$ by $\lambda_{max}(b)$ we get that

$$\left|\mathcal{A}^{(b)} f(z_1) - \mathcal{A}^{(b)} f(z_2)\right| \le K\|z_1 - z_2\|_1 + 2\tilde{L}\|z_1 - z_2\|_1 + 2\tilde{c}\|z_1 - z_2\|_1.$$

By this, the function $\mathcal{A}^{(b)} f$ is Lipschitz continuous and further, the same holds for $\delta_\phi T_t^{(b)} f$. □

Lemma 5.8. *The family $\{T_t^{(b)}\}_{t\ge 0}$ is a strongly continuous contraction semigroup on $\mathcal{D}^{(b)}$.*

Proof. By the results shown in Lemma 5.7 and the ideas of the proof of Lemma 5.4, we get that $T_t^{(b)}$ maps $\mathcal{D}^{(b)}$ into itself and by the boundedness of $\mathcal{A}^{(b)} f$ we get the strong continuity property. □

Lemma 5.9. *Let $f \in \mathcal{D}^{(b)}$ be arbitrary. Then, there exists a positive constant $\tilde{K}$ such that for every point (x_i, x_l, λ_j) in the state space of the bounded and discrete process*

$$\left|\mathcal{A}^{(h,b)} f(x_i, x_l, \lambda_j) - \mathcal{A} f(x_i, x_l, \lambda_j)\right| \le \tilde{K} h.$$

Proof. Let $f \in \mathcal{D}^{(b)}$ be arbitrary and (x_i, x_l, λ_j) a point in the state space of our bounded and discrete process. If we consider the difference between the two generators we get by the triangle inequality that

$$\begin{aligned}
&\left|\mathcal{A}^{(h,b)} f(x_i, x_l, \lambda_j) - \mathcal{A}^{(b)} f(x_i, x_l, \lambda_j)\right| \\
&\le \left|\frac{f(x_i + ch, x_l, \lambda_j e^{-\beta h}) - f(x_i, x_l, \lambda_j)}{h} - \delta_\phi f(x_i, x_l, \lambda_j)\right| \\
&\quad + \lambda_j \left|\sum_{k=1}^{N_U} f(x_i - x_k, x_k, \lambda_j) p_k^U - \int_{(0, U_{max}(b)]} f(x_i - u, u, \lambda_j)\, F_{U^{(b)}}(\mathrm{d}u)\right| \\
&\quad + \rho \left|\sum_{k=1}^{N_Y(j)} f(x_i, x_l, \lambda_{j+k}) p_k^Y(j) - \int_{(0, Y_{max}(b)]} f(x_i, m, \lambda_j + y)\, F_{Y^{(b)}}(\mathrm{d}y)\right|.
\end{aligned}$$

Let us first consider the second term. We can rewrite the difference as

$$\begin{aligned}
&\lambda_j \left|\sum_{k=1}^{N_U} f(x_i - x_k, x_k, \lambda_j) p_k^U - \int_{(0, U_{max}(b)]} f(x_i - u, u, \lambda_j)\, F_{U^{(b)}}(\mathrm{d}u)\right| \\
&= \lambda_j \left|\sum_{k=1}^{N_U} \int_{(x_{k-1}, x_k]} f(x_i - x_k, x_k, \lambda_j) - f(x_i - u, u, \lambda_j)\, F_{U^{(b)}}(\mathrm{d}y)\right|.
\end{aligned}$$

By the Lipschitz continuity of f and the boundedness of λ_j, we get that this is less or equal to $2L\lambda_{max}(b)ch$, where L denotes a Lipschitz constant of f. By the same idea, we can bound the third term by

$$L\rho\lambda_{max}(b)(1-e^{-\beta h}) \leq L\rho\lambda_{max}(b)\beta h.$$

For the second term we define the function $g:[0,\infty)\to\mathbb{R}$ by

$$g(t) = f(x_i + ct, x_l, \lambda_j e^{-\beta t}).$$

This is a Lipschitz continuous function in one real variable. Hence, it is differentiable almost everywhere and at every u, where g is differentiable the equality $g'(u) = \delta_\phi f(x + cu, x_l, \lambda e^{-\beta u})$ holds. By this we get that

$$\left|\frac{f(x_i + ch, x_l, \lambda_j e^{-\beta h}) - f(x_i, x_l, \lambda_j)}{h} - \delta_\phi f(x_i, x_l, \lambda)\right| = \frac{1}{h}\left|\int_0^h g'(u) - \delta_\phi f(x_i, x_l, \lambda_j)\,\mathrm{d}u\right|$$

$$= \frac{1}{h}\left|\int_0^h \delta_\phi f(x_i + cu, x_l, \lambda_j e^{-\delta u}) - \delta_\phi f(x_i, x_l, \lambda_j)\,\mathrm{d}u\right|$$

$$\leq \tilde{L}(ch + \lambda_{max}(b)(1-e^{-\beta h})) \leq \tilde{L}(c + \lambda_{max}(b)\beta)h,$$

where $\tilde{L}$ is a Lipschitz constant of $\delta_\phi f$. By this we get that

$$\left|\mathcal{A}^{(h,b)} f(x_i, x_l, \lambda_j) - \mathcal{A}^{(b)} f(x_i, x_l, \lambda_j)\right| \leq (\tilde{L}(c+\beta\lambda_{max}(b)) + 2L\lambda_{max}(b)c + L\rho\lambda_{max}(b)\beta)h.$$

□

Equivalently to Theorem 5.5, we prove the following lemma.

Lemma 5.10. *Let $f \in \mathcal{D}^{(b)}$ be arbitrary but fixed. Then, for all $t, s > 0$, $k \geq 0$, $h_1, \ldots, h_k \in C_b$ and $t_1 < t_2 < \ldots < t_k \leq t$ we have that*

$$\mathbb{E}_{(x_i,m,\lambda_j)}\left[\left(\int_t^{t+s}\left(\mathcal{A}^{(h,b)} f(X_v^{(h,b)}, m_v^{(h,b)}, \lambda_v^{(h,b)}) - \mathcal{A}^{(b)} f(X_v^{(h,b)}, m_v^{(h,b)}, \lambda_v^{(h,b)})\right)\mathrm{d}v\right) \times \prod_{l=1}^k h_l\left(X_{t_l}^{(h,b)}, m_{t_l}^{(h,b)}, \lambda_{t_l}^{(h,b)}\right)\right] \to 0,$$

as $h \to 0$.

Proof. The functions h_l are bounded, hence there is a constant L such that

$$\left|\left(\int_t^{t+s}\left(\mathcal{A}^{(h,b)}f(X_v^{(h,b)},m_v^{(h,b)},\lambda_v^{(h,b)})-\mathcal{A}^{(b)}f(X_v^{(h,b)},m_v^{(h,b)},\lambda_v^{(h,b)})\right)\,\mathrm{d}v\right)\cdot\right.$$
$$\left.\prod_{l=1}^{k}h_l\left(X_{t_l}^{(h,b)},m_{t_l}^{(h,b)},\lambda_{t_l}^{(h,b)}\right)\right|$$
$$\leq L\int_t^{t+s}\left|\mathcal{A}^{(h,b)}f(X_v^{(h,b)},m_v^{(h,b)},\lambda_v^{(h,b)})-\mathcal{A}^{(b)}f(X_v^{(h,b)},m_v^{(h,b)},\lambda_v^{(h,b)})\right|\,\mathrm{d}v.$$

By the boundedness derived in Lemma 5.9, we get that there is a constant $\tilde{K}$ such that

$$\mathbb{E}_{(x_i,m,\lambda_j)}\left[\int_t^{t+s}\left|\mathcal{A}^{(h,b)}f(X_v^{(h,b)},m_v^{(h,b)},\lambda_v^{(h,b)})-\mathcal{A}^{(b)}f(X_v^{(h,b)},m_v^{(h,b)},\lambda_v^{(h,b)})\right|\,\mathrm{d}v\right]\leq\tilde{K}hs,$$

which tends to 0 as $h\to 0$. □

Theorem 5.11. *The process* $\{(X_t^{(h,b)},m_t^{(h,b)},\lambda_t^{(h,b)})\}_{t\geq 0}$ *converges in distribution to the bounded process* $\{(X_t^{(b)},m_t^{(b)},\lambda_t^{(b)})\}_{t\geq 0}$ *as* $h\to 0$.

Proof. We obtain this by the result of Lemma 5.10 and Theorem 8.2 of Ethier and Kurtz (2009). □

Theorem 5.12. *The Markov chain* $\left\{(X^{(\bar{x},h,b)},m^{(\bar{x},h,b)},\lambda^{(\bar{x},h,b)})\right\}$ *converges weakly against the discrete process* $\{(X_t^{(h,b)},m_t^{(h,b)},\lambda_t^{(h,b)})\}_{t\geq 0}$ *as* $\bar{x}\to\infty$.

Proof. This can be proven as the weak convergence of the other processes using the convergence of the generators on the set

$$\mathcal{D}^{(h,b)}=\left\{f\in C_b\;\middle|\;\mathcal{A}^{(h,b)}f\in C_b,\ \lim_{x\to\infty}f(x,m,\lambda)=0\ \text{uniformly in } m \text{ and } \lambda\right\},$$

where $\mathcal{A}^{(h,b)}$ generates a strongly continuous contraction semigroup. □

5.4.3 Convergence of the Gerber-Shiu functions

Theorem 5.13. *Let* g_κ *be an arbitrary Gerber-Shiu function and* $g_\kappa^{(\bar{x},h,b)}$ *the corresponding GS-function of the Markov chain with finite state space. For* (x,λ) *let* $j=N_\lambda-\left\lfloor\frac{\ln(\lambda_{max})-\ln(\lambda)}{h\beta}\right\rfloor$, *and* $i=\left\lfloor\frac{x}{hc}\right\rfloor$. *Then, we have*

$$\lim_{b\to\infty}\lim_{h\to 0}\lim_{\bar{x}\to\infty}\left|g_\kappa^{(\bar{x},h,b)}(x_i,\lambda_j)-g_\kappa(x,\lambda)\right|=0.$$

Proof. By the proof of Lemma 5.14 in Kritzer et al. (2019), we have that our GS-function is a Skorokhod-continuous function of the process. By the weak convergence of the underlying processes, we know that the penalty functions converge too. Consequently,

$$\lim_{b\to\infty}\lim_{h\to 0}\lim_{\bar{x}\to\infty}\left|g_\kappa^{(\bar{x},h,b)}(x_i,\lambda_j)-g_\kappa(x,\lambda)\right|=0.$$

□

5.5 Examples

In this section, we give some explicit examples of Gerber-Shiu functions and corresponding numerical approximations in a Markovian shot-noise model with the following specific parameters. We choose decay parameter $\beta = 1$, intensity of the underlying Poisson process $\rho = 1.5$ and premium rate $c = \frac{15}{4}$. Further, we assume that the shock events Y_i and the claim events U_i are exponentially distributed with mean 1. All computations and simulations are made on a standard notebook with an Intel Core i5.10210U processor at 1.60 GHz and 16 GB of RAM.

5.5.1 Laplace-transform function of the time of ruin

The first example is the GS-function $g_\kappa^{(1)} := \mathbb{E}_{(x,\lambda)}\left[e^{-\kappa\tau}I_{\{\tau<\infty\}}\right]$, i.e. the Laplace transform of τ, with $\kappa = 0.1$, and for fixed $\lambda = 2.3$. In Figure 5.1, the function in black is a Monte Carlo simulation using 10000 sample paths and the red area is the corresponding 95-percent confidence interval. For the numerical approximations, we choose $\bar{x} = 50$, $\lambda_{max} = 4.5$, $U_{max} = 10$, $Y_{max} = 4.5$ and $h = \frac{1}{cm}$ for $m \in \{5, 10, 12\}$.

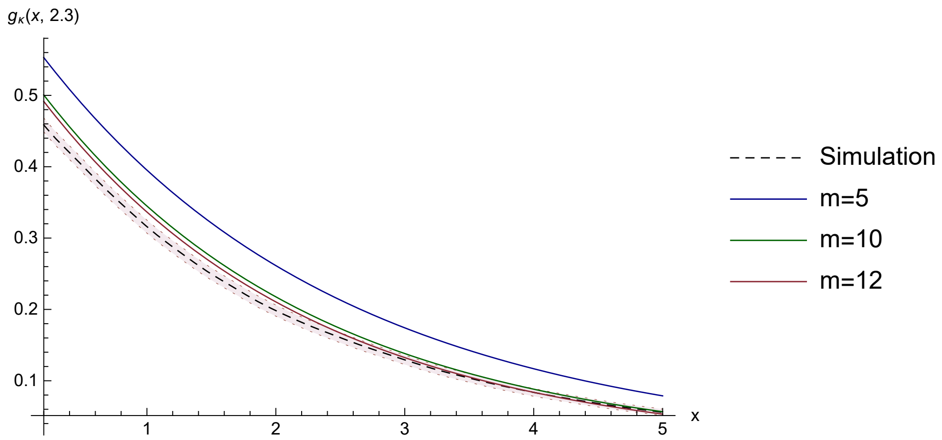

Figure 5.1: Laplace transform of the time of ruin and the corresponding numerical approximations.

As we can see in Table 5.1, the main advantage of the numerical method is the speed of computation. The scheme with $h \approx 0.02$ needed about 38 minutes whereas the computation of the corresponding simulation needed approximately 27 hours.

m	Minutes	Value	Abs. error	Rel. error
5	0.46	0.174	0.045	0.347
10	17	0.138	0.009	0.068
12	38	0.133	0.003	0.026
Sim.	1619	0.129	-	-

Table 5.1: Computation time of the surface and errors of the approximations at the point $x = 3$ and $\lambda = 2.3$.

5.5.2 Discounted surplus before ruin

Here, we consider the same setting as in the previous example, but now with penalty function $g_\kappa^{(2)}(x,\lambda) = \mathbb{E}_{(x,\lambda)}\left[e^{-\kappa\tau} X_{\tau-} I_{\{\tau<\infty\}}\right]$. Again, the black function in Figure 5.2 is a MC simulation from 10000 paths, which we will use as a reference solution. This time, the plot shows the behaviour of the GS-function in x for fixed $\lambda = 3.9$. As before, the numerical approximations are calculated with parameter $\bar{x} = 50$, $\lambda_{max} = 4.5$, $U_{max} = 10$, $Y_{max} = 4.5$ and $h = \frac{1}{cm}$ for $m \in \{5, 10, 12\}$. The function $w(x,y) = x$ is

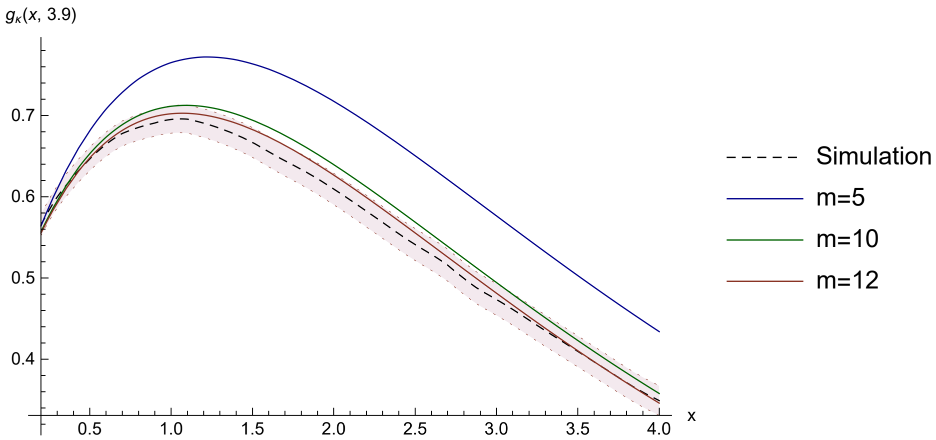

Figure 5.2: Discounted surplus before ruin and numerical approximations.

continuous but not bounded. We bypass this problem by considering penalty functions of the form $\tilde{w}(x,y) = \min(x,n)$ for $n \in \mathbb{N}$. These functions are continuous and bounded; hence, the theory derived before is applicable. Further, the sequence $\{e^{-\kappa\tau} \min(X_{\tau-}, n)\, I_{\{\tau<\infty\}}\}_{n\in\mathbb{N}}$ is monotone increasing. Consequently, the approximations converge for $n \to \infty$ by monotone convergence. As we can see in Table 5.2, the behaviour in terms of computing time and relative error is similar to the corresponding values in the example of the Laplace transform.

m	Minutes	Value	Abs. error	Rel. error
5	0.45	0.65	0.11	0.203
10	15	0.57	0.03	0.052
12	38	0.55	0.01	0.027
Sim.	1622	0.54	-	-

Table 5.2: Computation time of the surface, and errors of the approximations at the point $x = 2.5$ and $\lambda = 3.9$.

5.5.3 Ruin probability

As a third example, we consider the ruin probability $g_\kappa^{(3)}(x, \lambda) = \mathbb{E}_{(x,\lambda)}\left[I_{\{\tau<\infty\}}\right]$. The reference solution is again obtained by MC-simulation, and the bounds $\bar{x}$, λ_{max}, U_{max}, and Y_{max} are chosen as in the previous examples. An illustration of the simulation and the numerical approximations for fixed $\lambda = 2.3$ can be seen in Figure 5.3.

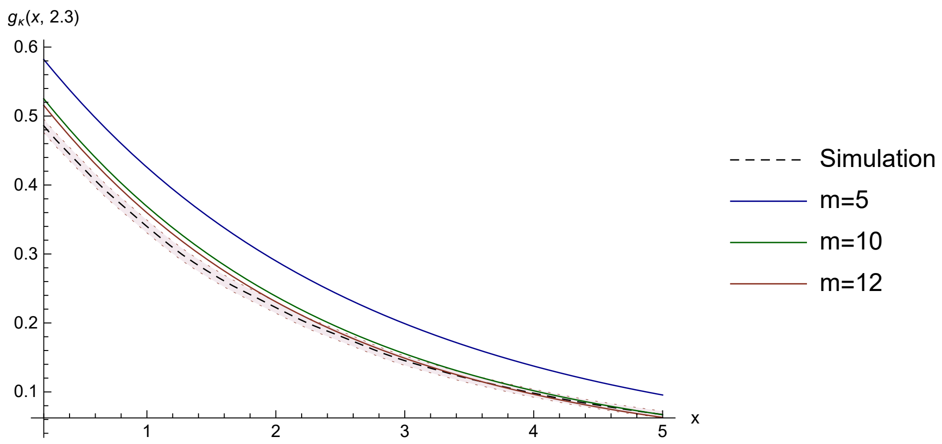

Figure 5.3: Ruin probability and numerical approximations.

In Table 5.3, we see that the run time and the relative error are very similar to the first two examples. Again, even in the finest step size considered, the numerical scheme beats the simulation by a factor of ≈ 40 in terms of computation time, which is the main advantage of our approach.

5.5.4 Empirical convergence order

Another topic of interest is the convergence order of numerical schemes. This has been studied for example by Chau et al. (2015), who considered GS-functions in a Lévy subordinator model. Numerical methods to solve integro-differential equations related to ours are also derived in Brunner (1988), who proposed spline collocation methods for ordinary Volterra integro-differential equations. He was able to achieve a convergence order up to order $2m$, given that the coefficients of the Volterra equation are $2m$ times

m	Minutes	Value	Abs. error	Rel. error
5	0.46	0.240	0.059	0.327
10	15	0.192	0.011	0.063
12	38	0.185	0.004	0.023
Sim.	1614	0.181	-	-

Table 5.3: Computation time of the surface and errors of the approximations at the point $x = 2.5$ and $\lambda = 2.3$.

continuously differentiable.
Since we consider partial integro-differential equations, we cannot use his results to obtain a theoretical convergence order. Alternatively, we compute the empirical order of convergence of our numerical scheme in the examples given before. For this, we consider two different approaches. The first is the estimated order of convergence (EOC) as defined in Steinbach (2008). For a sequence $\{x_n\}_{n\geq 0}$ with limit x, we define the sequence of absolute errors by $e_n := |x - x_n|$. Assuming that $e_n \approx Cn^{-\rho}$ for some fixed constant, i.e. that the sequence converges with order ρ, we divide by e_{n-1} and get $\frac{e_n}{e_{n-1}} \approx (\frac{n}{n-1})^{-\rho}$. Applying the logarithm and dividing by $\ln(\frac{n}{n-1})$ gives us the EOC

$$\hat{\rho}_n = \frac{\ln\left(\frac{e_n}{e_{n-1}}\right)}{\ln\left(\frac{n-1}{n}\right)}.$$

The second procedure uses the same assumption $e_n \approx Cn^{-\rho}$, or equivalently $\ln(e_n) \approx \ln(C) - \rho \ln(n)$. Having this form, we use a linear regression approach to get an estimator $\tilde{\rho}$ for the parameter ρ, as it is used by Chau et al. (2015).

We are interested in the convergence behaviour of the sequence of our numerical approximations at some fixed points (x_i, λ_j) as the fineness of the discretization tends to 0. To determine the error terms e_n correctly, we have to know the limit of this sequence, which is not the GS-function of our original process, but the GS-function of the bounded process, which we obtain by simulation.

In the following examples, we fix the bounds $\bar{x} = 50$, $Y_{max} = 4.5$, $U_{max} = 10$ and $\lambda_{max} = 4.5$ and the GS-functions $g_\kappa^{(1)}$, $g_\kappa^{(2)}$, and $g_\kappa^{(3)}$ as before. Then, we consider the sequence of numerical approximations with step size $h = \frac{1}{cm}$ for $m \in \{1, \dots, 12\}$ and compute the EOC and the regression estimate $\tilde{\rho}$ at the points $(0.4, 2.3)$, $(1.4, 2.3)$, and $(2.5, 2.3)$. As we can see in Table 5.4, it seems plausible that we observe a linear convergence behaviour.

In Figure 5.4 we see the linear function obtained by regression of the approach $-\ln(e_n) = \ln(C) + \rho \ln(n)$ for the ruin probability in the point $(0.4, 2.3)$ and the corresponding observed values in red. The coefficient of determination $R^2 = 0.9996$ indicates that there is indeed a linear relationship between $\ln(e_n)$ and $\ln(n)$, i.e. that the assumption

	Laplace transform			Surplus before ruin			Ruin probability		
m	$x = 0.4$	$x = 1.4$	$x = 2.5$	$x = 0.4$	$x = 1.4$	$x = 2.5$	$x = 0.4$	$x = 1.4$	$x = 2.5$
2	0.739	0.924	1.093	1.159	0.955	1.130	0.801	1.014	1.208
3	0.895	1.047	1.186	0.872	1.034	1.166	0.965	1.125	1.263
4	0.929	1.056	1.172	0.821	1.038	1.142	1.013	1.139	1.240
5	0.932	1.049	1.149	0.823	1.029	1.124	1.035	1.143	1.218
6	0.924	1.036	1.131	0.861	1.028	1.107	1.046	1.145	1.206
7	0.915	1.026	1.115	0.835	1.016	1.097	1.057	1.151	1.197
8	0.904	1.015	1.104	0.837	1.010	1.086	1.064	1.157	1.195
9	0.892	1.005	1.094	0.827	1.003	1.080	1.073	1.165	1.194
10	0.881	0.997	1.087	0.810	0.996	1.073	1.081	1.175	1.197
11	0.869	0.988	1.080	0.814	0.993	1.070	1.087	1.185	1.201
12	0.858	0.981	1.075	0.793	0.985	1.065	1.096	1.197	1.207
$\tilde{\rho}$	0.911	1.032	1.136	0.836	1.022	1.114	1.037	1.148	1.218

Table 5.4: Table of EOC and estimates done by regression approach.

$e_n \approx Cn^{-\rho}$ is reasonable, and that $\tilde{\rho} \approx 1.037$ is a good estimation of the true convergence order in this example.

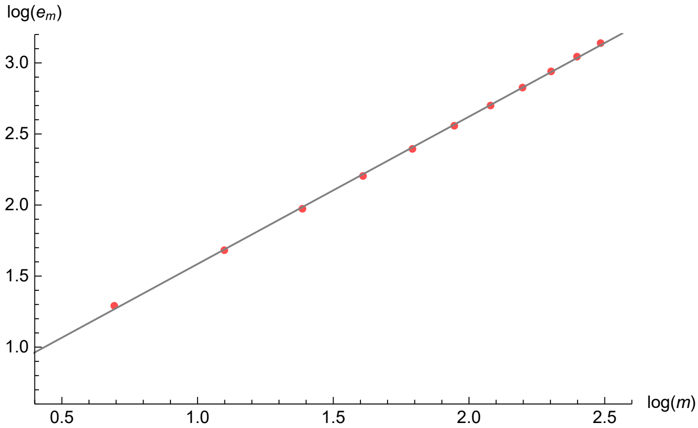

Figure 5.4: Linear regression line with slope $\tilde{\rho} = 1.037$ and observed errors.

Acknowledgments

This research was funded in whole, or in part, by the Austrian Science Fund (FWF) P 33317. For the purpose of open access, the author has applied a CC BY public copyright licence to any Author Accepted Manuscript version arising from this submission.
Declaration of interests

Bibliography

Albrecher, Hansjörg and Søren Asmussen (2006). "Ruin probabilities and aggregrate claims distributions for shot noise Cox processes". In: *Scandinavian Actuarial Journal* 2006.2, pp. 86–110. DOI: 10.1080/03461230600630395.

Asmussen, Søren (1989). "Risk theory in a Markovian environment". In: *Scandinavian Actuarial Journal* 1989.2, pp. 69–100. DOI: 10.1080/03461238.1989.10413858.

— (1995). *Applied Probability and Queues.* Wiley series in probability and mathematical statistics. Chichester [u.a.]: John Wiley.

Asmussen, Søren and Hansjörg Albrecher (2010). *Ruin probabilities.* Second. Vol. 14. Advanced Series on Statistical Science & Applied Probability. World Scientific Publishing Co. Pte. Ltd., Hackensack, NJ. DOI: 10.1142/9789814282536.

Asmussen, Søren and Mogens Steffensen (2020). *Risk and insurance—a graduate text.* Vol. 96. Probability Theory and Stochastic Modelling. Springer, Cham, pp. xv+505. DOI: 10.1007/978-3-030-35176-2.

Banerjee, Chandan et al. (2023). "A perfect storm: natural catastrophes and inflation in 2022". In: *sigma* 1/2023.

Basrak, Bojan, Olivier Wintenberger, and Petra Žugec (2019). "On the total claim amount for marked Poisson cluster models". In: *Advances in Applied Probability* 51.2, pp. 541–569. DOI: 10.1017/apr.2019.15.

Bessy-Roland, Yannick, Alexandre Boumezoued, and Caroline Hillairet (2021). "Multivariate Hawkes process for cyber insurance". In: *Annals of Actuarial Science* 15.1, pp. 14–39.

Biermé, Hermine and Agnès Desolneux (2012a). "A Fourier approach for the level crossings of shot noise processes with jumps". In: *Journal of Applied Probability* 49.1, pp. 100–113. DOI: 10.1239/jap/1331216836.

— (2012b). "Crossings of smooth shot noise processes". In: *The Annals of Applied Probability* 22.6, pp. 2240–2281. DOI: 10.1214/11-AAP807.

Björk, Tomas and Jan Grandell (1988). "Exponential inequalities for ruin probabilities in the Cox case". In: *Scandinavian Actuarial Journal* 1988.1-3, pp. 77–111. DOI: 10.1080/03461238.1988.10413839.

Borovkov, Konstantin and Günter Last (2008). "On level crossings for a general class of piecewise-deterministic Markov processes". In: *Advances in Applied Probability* 40.3, pp. 815–834. DOI: 10.1239/aap/1222868187.

Brémaud, Pierre (1981). *Point processes and queues.* Springer Series in Statistics. Springer-Verlag, New York - Berlin.

Brémaud, Pierre, Giovanna Nappo, and Giovanni Luca Torrisi (2002). "Rate of convergence to equilibrium of marked Hawkes processes". In: *Journal of Applied Probability* 39.1, pp. 123–136. DOI: 10.1017/s0021900200021562.

Brunner, Hermann (1988). "The approximate solution of initial-value problems for general Volterra integro-differential equations". In: *Computing. Archives for Scientific Computing* 40.2, pp. 125–137. DOI: 10.1007/BF02247941.

Chau, Ki Wai, Sheung Chi Phillip Yam, and Hailiang Yang (2015). "Fourier-cosine method for Gerber–Shiu functions". In: *Insurance: Mathematics and Economics* 61, pp. 170–180. DOI: 10.1016/j.insmatheco.2015.01.008.

Chiang, Wen-Hao, Xueying Liu, and George Mohler (2022). "Hawkes process modeling of COVID-19 with mobility leading indicators and spatial covariates". In: *International journal of forecasting* 38.2, pp. 505–520.

Cui, Lirong, Alan Hawkes, and He Yi (2020). "An elementary derivation of moments of Hawkes processes". In: *Advances in Applied Probability* 52.1, pp. 102–137. DOI: 10.1017/apr.2019.53.

Daley, D. J. and D. Vere-Jones (2003). *An introduction to the theory of point processes. Vol. I.* Second. Probability and its Applications (New York). Elementary theory and methods. New York: Springer-Verlag, pp. xxii+469.

Dassios, Angelos and Jiwook Jang (2003). "Pricing of catastrophe reinsurance and derivatives using the Cox process with shot noise intensity". In: *Finance and Stochastics* 7.1, pp. 73–95. DOI: 10.1007/s007800200079.

— (2005). "Kalman-Bucy filtering for linear systems driven by the Cox process with shot noise intensity and its application to the pricing of reinsurance contracts". In: *Journal of Applied Probability* 42.1, pp. 93–107. DOI: 10.1239/jap/1110381373.

Dassios, Angelos, Jiwook Jang, and Hongbiao Zhao (2015). "A risk model with renewal shot-noise Cox process". In: *Insurance Mathematics and Economics* 65, pp. 55–65. DOI: 10.1016/j.insmatheco.2015.08.009.

Dassios, Angelos and Hongbiao Zhao (2011). "A dynamic contagion process". In: *Advances in Applied Probability* 43.3, pp. 814–846. DOI: 10.1239/aap/1316792671.

Davis, Mark H. A. (1984). "Piecewise-Deterministic Markov Processes: A General Class of Non-Diffusion Stochastic Models". In: *Journal of the Royal Statistical Society: Series B (Methodological)* 46.3, pp. 353–376. DOI: 10.1111/j.2517-6161.1984.tb01308.x.

— (1993). *Markov models and optimization.* London and New York: Chapman & Hall.

Denisov, Denis, Serguei Foss, and Dmitry Korshunov (2010). "Asymptotics of randomly stopped sums in the presence of heavy tails". In: *Bernoulli* 16.4, pp. 971–994. DOI: 10.3150/10-BEJ251.

Diko, Peter and Miguel Usábel (2011). "A numerical method for the expected penalty–reward function in a Markov-modulated jump–diffusion process". In: *Insurance Mathematics and Economics* 49.1, pp. 126–131. DOI: 10.1016/j.insmatheco.2011.03.001.

Doney, R. A. and A. E. Kyprianou (2006). "Overshoots and undershoots of Lévy processes". In: *The Annals of Applied Probability* 16.1, pp. 91–106. DOI: 10.1214/105051605000000647.

Durrett, Rick (2019). *Probability—theory and examples.* Vol. 49. Cambridge Series in Statistical and Probabilistic Mathematics. Fifth edition of [MR1068527]. Cambridge University Press, Cambridge, pp. xii+419. DOI: 10.1017/9781108591034.

Dwass, Meyer (1969). "The total progeny in a branching process and a related random walk". In: *J. Appl. Probability* 6, pp. 682–686. DOI: 10.2307/3212112.

Ethier, Stewart N. and Thomas G. Kurtz (2009). *Markov Processes: Characterization and Convergence.* John Wiley & Sons.

Fleming, Thomas R. and David P. Harrington (1991). *Counting processes and survival analysis.* Wiley Series in Probability and Mathematical Statistics: Applied Probability and Statistics. John Wiley & Sons, Inc., New York, pp. xiv+429.

Foss, S., D. Korshunov, and S. Zachary (2011). *An Introduction to Heavy-Tailed and Subexponential Distributions.* Springer Series in Operations Research and Financial Engineering. Springer New York.

Garetto, Michele, Emilio Leonardi, and Giovanni Luca Torrisi (2021). "A time-modulated Hawkes process to model the spread of COVID-19 and the impact of countermeasures". In: *Annual reviews in control* 51, pp. 551–563.

Garrido, José and Manuel Morales (2006). "On The Expected Discounted Penalty function for Lévy Risk Processes". In: *North American Actuarial Journal* 10.4, pp. 196–216. DOI: 10.1080/10920277.2006.10597421.

Gerber, Hans U. and Elias S. W. Shiu (1998). "On the Time Value of Ruin". In: *North American Actuarial Journal* 2.1, pp. 48–72. DOI: 10.1080/10920277.1998.10595671.

— (2005). "The Time Value of Ruin in a Sparre Andersen Model". In: *North American Actuarial Journal* 9.2, pp. 49–69. DOI: 10.1080/10920277.2005.10596197.

Grandell, Jan (1991). *Aspects of risk theory.* Springer Series in Statistics: Probability and its Applications. Springer-Verlag, New York. DOI: 10.1007/978-1-4613-9058-9.

Grandell, Jan and Hanspeter Schmidli (2011). “Ruin probabilities in a diffusion environment”. In: *Journal of Applied Probability* 48.A, pp. 39–50. DOI: 10.1239/jap/1318940454.

Hawkes, Alan G. (Apr. 1971). “Spectra of some self-exciting and mutually exciting point processes”. In: *Biometrika* 58.1, pp. 83–90. DOI: 10.1093/biomet/58.1.83.

Hillairet, Caroline, Anthony Réveillac, and Mathieu Rosenbaum (2023). “An expansion formula for Hawkes processes and application to cyber-insurance derivatives”. In: *Stochastic Processes and their Applications*.

Jacobsen, Martin (2006). *Point process theory and applications: Marked point and piecewise deterministic processes / Martin Jacobsen*. Probability and its applications. Boston: Birkhäuser.

Kallenberg, Olav (2002). *Foundations of modern probability*. 2. ed. Probability and its applications. New York, NY: Springer.

Karabash, Dmytro and Lingjiong Zhu (2015). “Limit theorems for marked Hawkes processes with application to a risk model”. In: *Stochastic Models* 31.3, pp. 433–451. DOI: 10.1080/15326349.2015.1024868.

Klüppelberg, Claudia (1988). “Subexponential distributions and integrated tails”. In: *Journal of Applied Probability* 25.1, pp. 132–141. DOI: 10.2307/3214240.

Kritzer, Peter et al. (2019). “Approximation methods for piecewise deterministic Markov processes and their costs”. In: *Scandinavian Actuarial Journal* 2019.4, pp. 308–335. DOI: 10.1080/03461238.2018.1560357.

Lee, Wing Yan et al. (2021). “A Fourier-cosine method for finite-time ruin probabilities”. In: *Insurance Mathematics and Economics* 99, pp. 256–267. DOI: 10.1016/j.insmatheco.2021.03.001.

Li, Shuanming and José Garrido (2005). “On a general class of renewal risk process: analysis of the Gerber-Shiu function”. In: *Advances in Applied Probability* 37.3, pp. 836–856. DOI: 10.1239/aap/1127483750.

Löpker, Andreas and Zbigniew Palmowski (2013). “On time reversal of piecewise deterministic Markov processes”. In: *Electronic Journal of Probability* 18, no. 13, 29. DOI: 10.1214/EJP.v18-1958.

Lukasik, Michal et al. (2016). “Hawkes processes for continuous time sequence classification: an application to rumour stance classification in twitter”. In: *Proceedings of the 54th Annual Meeting of the Association for Computational Linguistics (Volume 2: Short Papers)*, pp. 393–398.

Lundberg, Fillip (1903). “Approximerad framställning av sannolikhetsfunktionen. aterförsäkring av kollektivrisker. Akad”. PhD thesis. Almqvist och Wiksell Uppsala, Sweden.

Macci, Claudio and Giovanni Luca Torrisi (2011). "Risk processes with shot noise Cox claim number process and reserve dependent premium rate". In: *Insurance Mathematics and Economics* 48.1, pp. 134–145. DOI: 10.1016/j.insmatheco.2010.10.007.

Meyn, Sean P. and R. L. Tweedie (1993). "Stability of Markovian processes. II. Continuous-time processes and sampled chains". In: *Advances in Applied Probability* 25.3, pp. 487–517. DOI: 10.2307/1427521.

Møller, Jesper and Jakob G. Rasmussen (2005). "Perfect simulation of Hawkes processes". In: *Advances in Applied Probability* 37.3, pp. 629–646. DOI: 10.1239/aap/1127483739.

Orsingher, Enzo and Francesco Battaglia (1982). "Probability distributions and level crossings of shot noise models". In: *Stochastics* 8.1, pp. 45–61. DOI: 10.1080/17442508208833227.

Palmowski, Zbigniew, Simon Pojer, and Stefan Thonhauser (2023). "Exact asymptotics of ruin probabilities with linear Hawkes arrivals". DOI: 10.48550/arXiv.2304.03075. Submitted.

Palmowski, Zbigniew and Tomasz Rolski (2002). "A technique for exponential change of measure for Markov processes". In: *Bernoulli. Official Journal of the Bernoulli Society for Mathematical Statistics and Probability* 8.6, pp. 767–785.

Pojer, Simon (2022). "Level crossings of the Markovian shot-noise process". DOI: 10.2139/ssrn.4285543. Submitted.

Pojer, Simon and Stefan Thonhauser (2023a). "Ruin probabilities in a Markovian shot-noise environment". In: *Journal of Applied Probability* 60.2, pp. 542–556. DOI: 10.1017/jpr.2022.63.

— (2023b). "The Markovian shot-noise risk model: a numerical method for Gerber-Shiu functions". In: *Methodology and Computing in Applied Probability* 25.1, p. 17. DOI: 10.1007/s11009-023-10001-w.

Preischl, Michael, Stefan Thonhauser, and Robert F. Tichy (2018). "Integral equations, quasi-Monte Carlo methods and risk modeling". In: *Contemporary computational mathematics—a celebration of the 80th birthday of Ian Sloan. Vol. 1, 2.* Springer, Cham, pp. 1051–1074.

Rizoiu, Marian-Andrei et al. (2017). "Hawkes processes for events in social media". In: *Frontiers of multimedia research*, pp. 191–218.

Rolski, Tomasz et al. (1999). *Stochastic processes for insurance and finance.* Wiley Series in Probability and Statistics. John Wiley & Sons, Ltd., Chichester. DOI: 10.1002/9780470317044.

Schmidli, Hanspeter (1997). "An extension to the renewal theorem and an application to risk theory". In: *The Annals of Applied Probability* 7.1, pp. 121–133. DOI: 10.1214/aoap/1034625255.

— (2010). "On the Gerber–Shiu function and change of measure". In: *Insurance Mathematics and Economics* 46.1, pp. 3–11. DOI: 10.1016/j.insmatheco.2009.04.004.

— (2017). *Risk theory.* Springer Actuarial. Springer, Cham, pp. xii+242. DOI: 10.1007/978-3-319-72005-0.

Schmidt, Thorsten (2014). "Catastrophe Insurance Modeled by Shot-Noise Processes". In: *Risks* 2.1, pp. 3–24. DOI: 10.3390/risks2010003.

— (2017). "Shot-noise processes in finance". In: *From statistics to mathematical finance.* Springer, pp. 367–385.

Sparre Andersen, Erik (1957). "On the collective theory of risk in case of contagion between claims". In: *Bulletin of the Institute of Mathematics and its Applications* 12.2, pp. 275–279.

Stabile, Gabriele and Giovanni Luca Torrisi (2010). "Large deviations of Poisson shot noise processes under heavy tail semi-exponential conditions". In: *Statistics & Probability Letters* 80.15-16, pp. 1200–1209. DOI: 10.1016/j.spl.2010.03.017.

Steinbach, Olaf (2008). *Numerical approximation methods for elliptic boundary value problems.* Finite and boundary elements, Translated from the 2003 German original. Springer, New York, pp. xii+386. DOI: 10.1007/978-0-387-68805-3.

Strini, Josef Anton and Stefan Thonhauser (2020). "On Computations in Renewal Risk Models—Analytical and Statistical Aspects". In: *Risks* 8.1, p. 24. DOI: 10.3390/risks8010024.

Willmot, Gordon E. and David C.M. Dickson (2003). "The Gerber–Shiu discounted penalty function in the stationary renewal risk model". In: *Insurance: Mathematics and Economics* 32.3, pp. 403–411. DOI: 10.1016/s0167-6687(03)00129-x.

Zhang, Xin (2008). "On the Ruin Problem in a Markov-Modulated Risk Model". In: *Methodology and Computing in Applied Probability* 10.2, pp. 225–238. DOI: 10.1007/s11009-007-9044-4.

Zhu, Lingjiong (2013). "Ruin probabilities for risk processes with non-stationary arrivals and subexponential claims". In: *Insurance Math. Econom.* 53.3, pp. 544–550. DOI: 10.1016/j.insmatheco.2013.08.008.