

# Optimal dividend strategies for a catastrophe insurer

Hansjörg Albrecher\*, Pablo Azcue<sup>†</sup> and Nora Muler<sup>†</sup>

## Abstract

In this paper we study the problem of optimally paying out dividends from an insurance portfolio, when the criterion is to maximize the expected discounted dividends over the lifetime of the company and the portfolio contains claims due to natural catastrophes, modelled by a shot-noise Cox claim number process. The optimal value function of the resulting two-dimensional stochastic control problem is shown to be the smallest viscosity supersolution of a corresponding Hamilton-Jacobi-Bellman equation, and we prove that it can be uniformly approximated through a discretization of the space of the free surplus of the portfolio and the current claim intensity level. We implement the resulting numerical scheme to identify optimal dividend strategies for such a natural catastrophe insurer, and it is shown that the nature of the barrier and band strategies known from the classical models with constant Poisson claim intensity carry over in a certain way to this more general situation, leading to action and non-action regions for the dividend payments as a function of the current surplus and intensity level. We also discuss some interpretations in terms of upward potential for shareholders when including a catastrophe sector in the portfolio.

## 1 Introduction

Optimal strategies for dividend payout from a surplus process of an insurance portfolio is a classical object of study in risk theory, starting with de Finetti [18] and Gerber [19]. Many of the existing results in the literature study variations of this problem (in terms of objective functions and constraints) under the assumption that the underlying risk process is a Brownian motion or a classical Cramér-Lundberg process, see e.g. [7, 9] for an overview. Over the years, it has been noted that the compound Poisson assumption of the Cramér-Lundberg process is too restrictive, and that claim number processes of doubly stochastic Poisson type (Poisson processes with stochastic intensity) are a more natural choice for certain application areas. The resulting Cox process can lead to tractable models, particularly under the assumption of a Poisson shot-noise intensity, see e.g. Dassios & Jang [15] and Albrecher & Asmussen [1]. Such a shot-noise dynamic is for instance a natural model for claim arrivals in the presence of catastrophes, where at Poissonian times a sudden jump of random size increases the intensity, leading to more claims for a certain period, and that additional intensity level then decreases over time as claims due to that catastrophe get reported and settled. In recent years, such models have been studied for various purposes, see e.g. Dassios & Zhao [17], Macci & Torrisi [24], Jang & Oh [22] and Pojer & Thonhauser [26, 27] in an insurance context, Boxma & Mandjes [14] for a related model in queueing and Schmidt [30] for applications in finance.

Stochastic control problems for such compound shot-noise Cox processes have, however, to the best of our knowledge not been addressed in the respective literature, with the notable recent exception of Liu & Cadenillas [23], who study a problem of optimal premiums, retention and prevention strategies which maximize expected utility of the insurer and policyholders in a certain way.

In this paper we will study the problem of maximizing expected aggregate discounted dividend payments up to ruin for an insurance risk process of compound shot-noise Cox type. In addition to the mathematical interest in studying such a problem, this allows to some extent to assess and establish profitable strategies

---

\*Corresponding Author. Department of Actuarial Science, Faculty of Business and Economics, University of Lausanne, CH-1015 Lausanne and Swiss Finance Institute. Email: hansjoerg.albrecher@unil.ch

<sup>†</sup>Departamento de Matemáticas, Universidad Torcuato Di Tella. Av. Figueroa Alcorta 7350 (C1428BIJ) Ciudad de Buenos Aires, Argentina. Email: pazcue@utdt.edu and nmuler@utdt.edu

*Keywords:* Optimal dividends; Shot-noise Cox process; Insurance; Viscosity solution; Hamilton-Jacobi-Bellman equation; Optimal control;

*MSC 2020 Classification:* 91G50, 49L25, 93E20

for investors of an insurance company which cover catastrophic risks, such as floods and storms, where one exogenous event leads to a sudden and time-transient increase of the claim arrival intensity. The latter is likely to become a more prominent topic in the future, not the least because of climate change and an increased frequency of catastrophes. While the setup of this paper is stationary, it may serve as a benchmark for future studies where also an anticipated increase of the frequency of catastrophes during the period of consideration and its effects on optimal strategies can be considered. Methodologically, the problem studied in this paper is a stochastic control problem in two dimensions. Due to the Markovian structure of the shot-noise process, the risk process is Markovian as a function of the current surplus and the current intensity level. Two-dimensional control problems have recently received quite some attention in risk theory, see e.g. [3, 4], [20] and [21]. However, the actual techniques needed in the present paper are somewhat different from the ones in the aforementioned papers, since the second dimension affects the dynamics of the process differently, and a substantial amount of technicalities will be needed to tackle the posed optimization problem in a rigorous way.

We will prove that the optimal value function of this stochastic control problem is the smallest viscosity supersolution of a Hamilton-Jacobi-Bellman equation and we will also show that it can be approximated uniformly by a sufficiently fine discretization of the surplus process and the intensity process. This will then allow us to determine numerically the value function of this optimal dividend problem together with the optimal dividend payment strategies. It will turn out that the additional variability of the claim intensity process leads to an upward potential for the shareholders of the insurance company. We assume here that the intensity process can be observed, which is not unrealistic, as the jumps in the shot-noise process of the intensity are the documented catastrophes and the used decay function may be assumed known as well as a consequence of a modelling approach on past claim settlement experiences, see e.g. [16] for its estimation using a Kalman-Bucy filter). In this case the company can in fact steer the dividend streams according to the present situation of claim intensity and surplus level, and benefit from the already received premiums from the underlying policyholders. The numerical results in this paper allow to quantify this effect, under the albeit somewhat simplistic model assumptions.

The structure of the remaining paper is as follows. Section 2 introduces the underlying insurance model and the concrete formulation of the considered stochastic control problem. Section 3 derives some basic properties of the optimal value function. Section 4 formulates the corresponding Hamilton-Jacobi-Bellman equation and shows that the optimal value function can be identified as its smallest viscosity supersolution, and Section 5 briefly discusses some asymptotic properties of the latter. Subsequently, in Section 6 we show that one can uniformly approximate the optimal value function through admissible strategies defined on a discretization grid of the surplus values. In Section 7 we then pave the way for actual numerical solutions of the control problem by discretizing also the intensity space and showing that the optimal value function can be approximated uniformly that way. In Section 8 we first describe how to set up the numerical scheme concretely, and afterwards we apply the procedure to a number of parameter settings from classical compound Poisson examples, so that we can study the deviations of the optimal strategies when introducing the shot-noise process for the claim number intensity. Section 9 concludes and gives some practical interpretations of the obtained results as well as possible directions for future research. All proofs are delegated to an extensive appendix.

## 2 The model

Consider a free surplus process of an insurance portfolio given by

$$X_t = x + pt - \sum_{j=1}^{N_t} U_j, \quad (1)$$

where  $x$  is the initial surplus,  $p$  is the premium rate and  $U_i$  is the size of the  $i$ -th claim (arriving at time  $\tau_i$ ), cf. [8]. All claims are assumed to be i.i.d. positive random variables with distribution function  $F_U$  and finite expectation. Let the process

$$N_t = \#\{j : \tau_j \leq t\}$$

be an inhomogeneous Poisson process with intensity  $\lambda_t$ , corresponding to the number of claims up to time  $t$ . The process  $N_t$  and the random variables  $U_i$  are independent of each other. We assume that the intensity  $\lambda_t$  is a shot-noise process, that is

$$\lambda_t = \lambda_t^c + \sum_{k=1}^{\tilde{N}_t} Y_k e^{-d(t-T_k)}, \quad (2)$$

where

$$\lambda_t^c = \underline{\lambda} + e^{-dt} (\lambda - \underline{\lambda}). \quad (3)$$

Here  $\lambda \geq \underline{\lambda}$  is the initial intensity and  $\tilde{N}_t = \#\{k : T_k \leq t\}$  is a Poisson process of constant intensity  $\beta$ . Note that  $\lambda_t \geq \underline{\lambda}$  for all  $t \geq 0$  and if the initial intensity is  $\lambda > \underline{\lambda}$ , then  $\lambda_t > \underline{\lambda}$  for all  $t \geq 0$ . See Figure 2.1 for an illustration of a sample path of  $\lambda_t$  for  $\lambda = \underline{\lambda}$ .  $T_k$  corresponds to the arrival times of catastrophes that produce

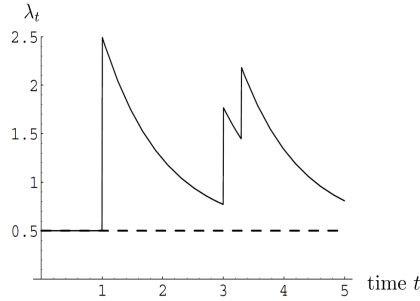


Figure 2.1: Sample path of  $\lambda_t$

jumps in the intensity  $\lambda_t$  and the upward jumps  $Y_k$  are assumed to be i.i.d. positive random variables with distribution function  $F_Y$  and finite expectation. In this context,  $\underline{\lambda}$  can also be interpreted as a base intensity for incoming claim occurrences, for instance also a "regular" non-catastrophic policies part in the portfolio.

Following Dassios and Jang [15], we can describe this model in a rigorous way by defining first the compound Cox filtered space. Given any intensity-filtered probability space  $(\Omega^0, \mathcal{F}^0, (\mathcal{F}_t^0)_{t \geq 0}, \mathbb{P}^0)$  and any intensity random process  $\lambda = (\lambda_t)_{t \geq 0}$  adapted to  $(\mathcal{F}_t^0)_{t \geq 0}$ , let us define the compound Cox filtered space  $(\Omega^\lambda, \mathcal{F}^\lambda, (\mathcal{F}_t^\lambda)_{t \geq 0}, \mathbb{P}^\lambda)$  conditional on the process  $\lambda$  in the following way. The sample set is

$$\Omega^\lambda = \Omega^0 \times \bar{\Omega},$$

where

$$\bar{\Omega} = \{(\tau_n, U_n)_{n \geq 1} \in [0, \infty) \times (0, \infty) : \tau_n < \tau_{n+1} \text{ and } \lim_{n \rightarrow \infty} \tau_n = \infty\}; \quad (4)$$

$\mathcal{F}^\lambda$  is the complete  $\sigma$ -field generated by the  $\sigma$ -field  $\mathcal{F}^0$  and the random variables  $\tau_n : \Omega^\lambda \rightarrow [0, \infty)$  and  $U_n : \Omega^\lambda \rightarrow (0, \infty)$ ; the filtration  $(\mathcal{F}_t^\lambda)_{t \geq 0}$ , where  $\mathcal{F}_t^\lambda$  is the complete  $\sigma$ -field generated by the  $\sigma$ -field  $\mathcal{F}_t^0$  and the random variables  $\tau_n : \Omega^\lambda \rightarrow [0, \infty)$  and  $U_n : \Omega^\lambda \rightarrow (0, \infty)$  for  $\tau_n \leq t$ ; and  $\mathbb{P}^\lambda$  the probability measure defined in  $\mathcal{F}^\lambda$  which satisfies

1.  $(U_n)_{n \geq 1}$  is a sequence of i.i.d. random variables with  $\mathbb{P}^\lambda(U_n \leq x) = F_U(x)$ ;
2. the counting process  $N_t^\lambda : \Omega^\lambda \rightarrow \mathbb{N}_0$  defined by  $N_t^\lambda = \#\{n : \tau_n \leq t\}$  satisfies

$$\mathbb{P}^\lambda(N_{t_2}^\lambda - N_{t_1}^\lambda = k \mid \lambda_s, t_1 \leq s \leq t_2) = e^{-\int_{t_1}^{t_2} \lambda_s ds} \frac{1}{k!} \left( \int_{t_1}^{t_2} \lambda_s ds \right)^k$$

for  $t_1 < t_2$ ;

3. the random variables  $U_n$  are independent of the counting process  $N_t^\lambda$ .

The shot-noise intensity filtered probability space is then defined as

$$(\Omega^{SN}, \mathcal{F}^{SN}, (\mathcal{F}_t^{SN})_{t \geq 0}, \mathbb{P}^{SN}),$$

where

$$\Omega^{SN} = \left\{ (T_k, Y_k)_{k \geq 1} \in [0, \infty) \times (0, \infty) : T_k < T_{k+1} \text{ and } \lim_{k \rightarrow \infty} T_k = \infty \right\}$$

and  $\mathcal{F}_t^{SN}$  is the  $\sigma$ -field generated by the set  $\{(T_k, Y_k)_{k \geq 1} : T_k \leq t\}$ . The Poisson process

$$\tilde{N}_t = \#\{k : T_k \leq t\}$$

of constant intensity  $\beta$  is independent of the random variables  $Y_k$ . The intensity shot-noise process  $\lambda_t$  with initial intensity  $\lambda$  is given by (2).

Given any  $\lambda \geq \underline{\lambda}$ , we consider in this paper the compound Cox filtered space  $(\Omega, \mathcal{F}, (\mathcal{F}_t)_{t \geq 0}, \mathbb{P})$  conditional on the intensity shot-noise process  $(\lambda_t)_{t \geq 0}$  with initial intensity  $\lambda$  given by (2). Here the intensity filtered

probability space  $(\Omega^0, \mathcal{F}^0, (\mathcal{F}_t^0)_{t \geq 0}, \mathbb{P}^0) = (\Omega^{SN}, \mathcal{F}^{SN}, (\mathcal{F}_t^{SN})_{t \geq 0}, \mathbb{P}^{SN})$ , the counting process  $N_t : \Omega \rightarrow \mathbb{N}_0$  is called the Cox process conditional on the intensity shot-noise process  $\lambda_t$  and the surplus given in (1) is called the compound Cox process with drift  $p$  generated by the shot-noise process  $\lambda_t$  with initial intensity  $\lambda$  and initial surplus  $x$ .

Let us define

$$\Lambda_t = \int_0^t \lambda_s ds. \quad (5)$$

Then,

$$\mathbb{P}(\tau_{j+1} - \tau_j \leq t | \mathcal{F}_{\tau_j}) = \mathbb{E} \left[ 1 - e^{-\int_0^t \lambda_{s+\tau_j} ds} \middle| \mathcal{F}_{\tau_j} \right] = \mathbb{E} \left[ 1 - e^{-(\Lambda(t+\tau_j) - \Lambda(\tau_j))} \middle| \mathcal{F}_{\tau_j} \right]$$

and in particular

$$\mathbb{P}(\tau_1 \leq t) = \mathbb{E} \left[ 1 - e^{-\Lambda_t} \right].$$

The insurance company uses part of the surplus to pay dividends to the shareholders. Let us define the dividend strategy  $L = (L_t)_{t \geq 0}$  where  $L_t$  denotes the cumulative dividends paid up to time  $t$ . The strategy  $(L_t)_{t \geq 0}$  is admissible if it is non-decreasing, càdlàg (right continuous with left limits), adapted with respect to the  $(\mathcal{F}_t)_{t \geq 0}$ , and if it satisfies  $L_0 \geq 0$  and  $L_t \leq X_t$  for  $t < \tau^L$ , where the ruin time  $\tau^L$  is defined as

$$\tau^L = \inf \{ t \geq 0 : X_t - L_{t-} < 0 \}. \quad (6)$$

This last condition means that the ruin time can only occur at the arrival of a claim and that no lump dividend payment can be made at the ruin time. We define the controlled surplus process as

$$X_t^L = X_t - L_t. \quad (7)$$

Denote by  $\Pi_{x, \lambda}$  the set of admissible dividend strategies starting with initial surplus level  $x \geq 0$  and initial intensity  $\lambda \geq \underline{\lambda}$ . For any initial surplus level  $x \geq 0$  and initial intensity  $\lambda \geq \underline{\lambda}$ , we can write the optimal value function as

$$V(x, \lambda) = \sup_{L \in \Pi_{x, \lambda}} J(L; x, \lambda), \quad (8)$$

where

$$J(L; x, \lambda) = \mathbb{E} \left( \int_{0-}^{\tau^L} e^{-qt} dL_t \right). \quad (9)$$

Here  $q > 0$  is a constant discount factor. We assume that the premium rate  $p$  is determined using an expected value principle w.r.t. the asymptotic distribution of  $\lambda_t$ . That is,

$$p = (1 + \eta) \mathbb{E}(U_1) \lim_{t \rightarrow \infty} \mathbb{E} \left( \frac{\Lambda_t}{t} \right) := (1 + \eta) \mathbb{E}(U_1) \lambda_{av}, \quad (10)$$

where  $\eta > 0$  is the relative safety loading (mathematically, one could also allow for negative  $\eta$  for the purposes of this paper as long as  $p$  remains positive, but only a positive safety loading is of practical interest).

Adapting results of [15, Cor.2.4], one can derive the following explicit expressions (we delegate the proof to Appendix A.1).

**Proposition 2.1** *We have that*

$$\begin{aligned} \mathbb{E}(\lambda_t) &= \underline{\lambda}(1 - e^{-dt}) + \lambda e^{-dt} + \frac{(1 - e^{-dt})}{d} \beta \mathbb{E}(Y_1), \\ \mathbb{E}(\Lambda_t) &= \underline{\lambda}t - \underline{\lambda} \left( \frac{1 - e^{-dt}}{d} \right) + \frac{(1 - e^{-dt})}{d} \lambda + \frac{(e^{-dt} - 1 + dt)}{d^2} \beta \mathbb{E}(Y_1) \end{aligned}$$

and

$$\lambda_{av} = \underline{\lambda} + \frac{\beta \mathbb{E}(Y_1)}{d}.$$

### 3 Basic results

The proofs of the propositions in this section are in Appendix A.2. First, we show that  $V(x, \cdot)$  is non-increasing on  $\lambda$  and uniformly continuous.

**Proposition 3.1** (1) *If  $\lambda_1 < \lambda_2$ , then  $V(x, \lambda_1) \geq V(x, \lambda_2)$ .* (2) *Given  $\varepsilon > 0$ , there exists  $\delta > 0$  (independent of  $x, \lambda_1$  and  $\lambda_2$ ) such that if  $\lambda_2 - \lambda_1 < \delta$ , then  $V(x, \lambda_1) - V(x, \lambda_2) \leq \varepsilon$ . So  $V$  is uniformly continuous in  $\lambda$ .*

Next, we prove that  $V(\cdot, \lambda) : [0, \infty) \rightarrow (0, \infty)$  is locally Lipschitz.

**Proposition 3.2** *Take  $x_2 > x_1$ , then*

$$0 \leq V(x_2, \lambda) - V(x_1, \lambda) \leq V(x_2, \underline{\lambda}) \frac{\beta\lambda + q}{p} (x_2 - x_1).$$

In the next proposition, we show a locally Lipschitz result for  $V(x, \cdot)$  in the open set  $(\underline{\lambda}, \infty)$ . At the lower boundary  $\underline{\lambda}$  we only have the uniform continuity result given in Proposition 3.1, because the Lipschitz constant obtained in the next proposition blows up at  $\lambda = \underline{\lambda}$ .

**Proposition 3.3** *Consider  $\underline{\lambda} < \lambda_1 < \lambda_2$ , then*

$$0 \leq V(x, \lambda_1) - V(x, \lambda_2) \leq V(x, \lambda_1) \frac{\beta\lambda_2 + q}{d(\lambda_1 - \underline{\lambda})} (\lambda_2 - \lambda_1).$$

**Remark 3.4** The dividend optimization problem with constant intensity was studied intensively in the literature (see e.g. Schmidli [29, Sec.2.4] and Azcue and Muler [10]). Unlike the shot-noise optimization problem, this constant-intensity problem is one-dimensional. Let us denote by  $v^\lambda(x)$  the optimal value function with constant intensity  $\lambda$  and premium rate  $p$ ; it is known that  $v^\lambda$  is non-decreasing with  $v^\lambda(x) - x$  positive, non-decreasing and bounded. Moreover,  $\lim_{\lambda \rightarrow \infty} v^\lambda(x) = x$ .

**Remark 3.5** From Proposition 3.1, we obtain that  $V(x, \lambda) \leq v^\lambda(x)$  for all  $x \geq 0$  and  $\lambda \in [\underline{\lambda}, \infty)$ . So, the optimal value function  $V$  satisfies  $V(x, \lambda) \leq v^\lambda(x) \leq x + K$  for some  $K > 0$ .

We now state the so-called Dynamic Programming Principle (DPP). The proof is similar to the one given in Lemma 1.2 of [11].

**Lemma 3.6** *For any initial surplus  $x \geq 0$  and any stopping time  $\tau$  adapted to  $\mathcal{F}^\lambda$ , we can write*

$$V(x, \lambda) = \sup_{L \in \Pi_{x, \lambda}} \left( \mathbb{E} \left( \int_0^{\tau \wedge \tau^L} e^{-qs} dL_s + e^{-q(\tau \wedge \tau^L)} I_{\{\tau \wedge \tau^L < \tau^L\}} V(X_{\tau \wedge \tau^L}^L, \lambda_{\tau \wedge \tau^L}) \right) \right).$$

Finally, the asymptotic behavior of  $V$  as  $\lambda$  goes to infinity can be determined.

**Proposition 3.7** *It holds that for all  $x \geq 0$ ,  $\lim_{\lambda \rightarrow \infty} V(x, \lambda) = x$ .*

## 4 Hamilton-Jacobi-Bellman equation

In this section we obtain the Hamilton-Jacobi-Bellman (HJB) equation associated to the optimization problem (8). These results are a generalization to the two-dimensional case of the ones given in [11, Sec.3] for the one-dimensional case. The proofs of the results of this section and the auxiliary results are deferred to Appendix A.3.

The HJB equation of this optimization problem is given by

$$\max\{\mathcal{L}(V)(x, \lambda), 1 - V_x(x, \lambda)\} = 0, \tag{11}$$

where

$$\begin{aligned} \mathcal{L}(V)(x, \lambda) &= pV_x(x, \lambda) - d(\lambda - \underline{\lambda})V_\lambda(x, \lambda) - (q + \lambda + \beta)V(x, \lambda) \\ &\quad + \lambda \int_0^x V(x - \alpha, \lambda) dF_U(\alpha) + \beta \int_0^\infty V(x, \lambda + \gamma) dF_Y(\gamma) \end{aligned} \tag{12}$$

and  $x \geq 0, \lambda \geq \underline{\lambda}$ . Since the optimal value function  $V$  is locally Lipschitz but could be not differentiable at some points, we cannot say that  $V$  is a solution of the HJB equation, we prove instead that  $V$  is a viscosity solution of the corresponding HJB equation. Let us define this notion.

**Definition 4.1** *A uniformly continuous function  $\bar{u} : [0, \infty) \times [\underline{\lambda}, \infty) \rightarrow \mathbf{R}$ , that is locally Lipschitz in  $(0, \infty) \times [\underline{\lambda}, \infty)$ , is a viscosity supersolution of (11) at  $(x, \lambda) \in (0, \infty) \times (\underline{\lambda}, \infty)$  if any continuously differentiable function  $\varphi : (0, \infty) \times [\underline{\lambda}, \infty) \rightarrow \mathbf{R}$  such that  $\varphi(y, \cdot)$  is bounded for any  $y \geq 0$ ,  $\varphi(x, \lambda) = \bar{u}(x, \lambda)$  and such that  $\bar{u} - \varphi$  reaches the minimum at  $(x, \lambda)$  satisfies*

$$\max\{\mathcal{L}(\varphi)(x, \lambda), 1 - \varphi_x(x, \lambda)\} \leq 0.$$

*A uniformly continuous function  $\underline{u} : [0, \infty) \times [\underline{\lambda}, \infty) \rightarrow \mathbf{R}$ , that is locally Lipschitz in  $(0, \infty) \times [\underline{\lambda}, \infty)$ , is a viscosity subsolution of (11) at  $(x, \lambda) \in (0, \infty) \times (\underline{\lambda}, \infty)$  if any continuously differentiable function  $\psi : (0, \infty) \times [\underline{\lambda}, \infty) \rightarrow \mathbf{R}$  such that  $\psi(y, \cdot)$  is bounded for any  $y \geq 0$ ,  $\psi(x, \lambda) = \underline{u}(x, \lambda)$  and such that  $\underline{u} - \psi$  reaches the maximum at  $(x, \lambda)$  satisfies*

$$\max\{\mathcal{L}(\psi)(x, \lambda), 1 - \psi_x(x, \lambda)\} \geq 0.$$

*A function  $u : (0, \infty) \times (\underline{\lambda}, \infty) \rightarrow \mathbf{R}$  which is both supersolution and subsolution at  $(x, \lambda) \in (0, \infty) \times (\underline{\lambda}, \infty)$  is called a viscosity solution of (11) at  $(x, \lambda)$ .*

**Proposition 4.1**  $V$  is a viscosity solution of the HJB equation (11) at any  $(x, \lambda) \in (0, \infty) \times (\underline{\lambda}, \infty)$ .

From Remark 3.5, the function  $V$  satisfies the growth condition

$$u(x, \lambda) \leq K + x \text{ for all } (x, \lambda) \in (0, \infty) \times (\underline{\lambda}, \infty). \quad (13)$$

We have the following characterization of the optimal value function.

**Proposition 4.2** The optimal value function  $V$  is the smallest viscosity supersolution of (11) among those that are non-increasing in  $\lambda$  and satisfy the growth condition (13).

From the previous proposition we deduce the usual viscosity verification result:

**Corollary 4.3** Consider a family of admissible strategies  $\{L^{x,\lambda} \in \Pi_{x,\lambda} : (x, \lambda) \in (0, \infty) \times (\underline{\lambda}, \infty)\}$ . If the function  $W(x, \lambda) := J(L^{x,\lambda}; x, \lambda)$  is a viscosity supersolution of (11), then  $W(x, \lambda)$  is the optimal value function. Also, if for each  $k \geq 1$  there exists a family of strategies  $\{L_k^{x,\lambda} \in \Pi_{x,\lambda} : (x, \lambda) \in (0, \infty) \times (\underline{\lambda}, \infty)\}$  such that  $W(x, \lambda) := \lim_{k \rightarrow \infty} J(L_k^{x,\lambda}; x, \lambda)$  is a viscosity supersolution of the HJB equation (11) in  $(0, \infty) \times (\underline{\lambda}, \infty)$ , then  $W$  is the optimal value function  $V$ .

As it is usual for these type of problems, the way in which the optimal value function  $V(x, \lambda)$  solves the HJB equation suggests that the state space  $[0, \infty) \times [\underline{\lambda}, \infty)$  is partitioned into two regions: a no-action region  $\mathcal{NA}$  in which no dividends are paid and an action region  $\mathcal{A}$  in which dividends are paid. Roughly speaking, the points in the  $\mathcal{NA}$  region satisfy  $\mathcal{L}(V) = 0$  and  $1 - V_x < 0$  and the points in  $\mathcal{A}$  satisfy  $\mathcal{L}(V) \leq 0$  and  $1 - V_x = 0$ .

## 5 Asymptotic properties of the optimal value function

In the following proposition, we use Corollary 4.3 in order to prove that there exists an explicit threshold  $p/q$ , such that if the surplus is above this explicit threshold, then one should pay at least the exceeding surplus as dividends, regardless of the current intensity of the shot-noise process. Hence  $[p/q, \infty) \times [\underline{\lambda}, \infty) \subset \mathcal{A}$  but this inclusion is strict.

The proofs of the results of this section and the auxiliary results are deferred to Appendix A.4.

**Proposition 5.1**  $V(x, \lambda) = x - p/q + V(p/q, \lambda)$  for  $x \geq p/q$ .

We have the following uniform convergence when the current intensity of the shot-noise process goes to infinity.

**Proposition 5.2**  $\lim_{\lambda \rightarrow \infty} \sup_{x \geq 0} (V(x, \lambda) - x) = 0$ .

**Remark 5.3** For technical reasons, in the development of the numerical scheme later on we will also need an extension of admissible strategies in which the insurance company can, at any time, pay all the surplus immediately as dividends and finish the insurance business (although this strategy is never optimal). More precisely, let us consider (as before)  $L_t$  as the cumulative dividend payment strategy up to time  $t$  and  $\tau^F$  is a stopping (finite or infinite) time at which the company pays all the surplus immediately as dividends and finishes the insurance business. We say that the strategy  $\tilde{\pi} = (L, \tau^F)$  is admissible if  $L$  is non-decreasing, càdlàg, adapted with respect to the filtration generated by the process  $(X_t, \lambda_t)$ , satisfies  $L_t \leq X_t$  up to ruin time and  $\tau^F$  is a stopping time with respect to the filtration generated by the process  $(X_t, \lambda_t)$ . We define  $\tilde{\Pi}_{x,\lambda}$  as the set of all  $\tilde{\pi}$ -admissible strategies. Take any  $L \in \Pi_{x,\lambda}$  then  $\tilde{\pi} = (L, \infty) \in \tilde{\Pi}_{x,\lambda}$  so we can think that  $\Pi_{x,\lambda}$  is contained in  $\tilde{\Pi}_{x,\lambda}$ . Let us define the value function of any strategy  $\tilde{\pi} \in \tilde{\Pi}_{x,\lambda}$  as

$$J(\tilde{\pi}; x, \lambda) = \mathbb{E} \left( \int_0^{\tau^L \wedge \tau^F} e^{-qt} dL_t + I_{\{\tau^F < \tau^L\}} e^{-q\tau^F} X_{\tau^F}^L \right).$$

It is straightforward to see that the optimal value function defined in (8) satisfies

$$V(x, \lambda) = \sup_{\tilde{\pi} \in \tilde{\Pi}_{x,\lambda}} J(\tilde{\pi}; x, \lambda) \quad (14)$$

because it is never optimal to pay all the current surplus and finish the business.

## 6 Approximation of the value function by a discretization of the surplus

In this section we show that it is possible to approximate uniformly the optimal value function  $V$  with value functions of admissible strategies that are constructed through a discrete set on the surplus level. These strategies are stationary in the sense that they only depend on the current surplus and intensity. Moreover, the only possible options are to either pay a lump sum of dividends or not paying any dividends. The result of this section follows closely the ones in [12]; in that paper, the intensity of the arrival of claims was constant and so this approach makes it possible to approximate numerically the optimal value function. However, in the model of the present paper, the situation is more complicated since the intensity of the arrival of claims changes over time. In contrast to [12], in this section we will only obtain a semi-discrete result that is not sufficient for the numerical results. In the next section, we will also discretize the intensity process arrivals in order to obtain a numerical scheme. The proofs of the results of this section can be found in Appendix A.5.

More precisely, we construct in this section a family of admissible strategies for any point in a subset of  $[0, \infty) \times [\underline{\lambda}, \infty)$  and then extend it to the whole set  $[0, \infty) \times [\underline{\lambda}, \infty)$ . We will show that the value functions of these strategies approximate uniformly the optimal value function  $V$ .

Given any approximation parameter  $\delta > 0$ , we define the grid domain

$$\mathcal{G}_\delta := \left\{ x_n^\delta = np\delta : n \geq 0 \right\} \quad (15)$$

in the surplus state space  $[0, \infty)$ ; we construct first a subfamily of admissible strategies  $\tilde{\Pi}_{x_n^\delta, \lambda}^\delta \subset \tilde{\Pi}_{x_n^\delta, \lambda}$  for any point  $(x_n^\delta, \lambda) \in \mathcal{G}_\delta \times [\underline{\lambda}, \infty)$ , and afterwards a subfamily of admissible strategies  $\tilde{\Pi}_{x, \lambda}^\delta \subset \tilde{\Pi}_{x, \lambda}$  for any point  $(x, \lambda) \in [0, \infty)$ .

Let us define the subfamily  $\tilde{\Pi}_{x, \lambda}^\delta \subset \tilde{\Pi}_{x, \lambda}$  for any point  $(x, \lambda) \in [0, \infty) \times [\underline{\lambda}, \infty)$  in a precise way.

Consider first the case in which  $(x, \lambda) \in \mathcal{G}_\delta \times [\underline{\lambda}, \infty)$ , so  $x = x_n^\delta$  for some  $n \geq 0$ . The idea of this construction is to find, at each point in  $\mathcal{G}_\delta \times [\underline{\lambda}, \infty)$ , the best local strategy among the ones suggested by the operators of the HJB equation (11). These possible local strategies are: either the company pays no dividends or pays immediately a lump sum  $p\delta$  as dividends; moreover, the company can finish the insurance activity at any time. We modify these local strategies in such a way that the controlled surplus always lies in  $\mathcal{G}_\delta$  immediately after the arrival of a claim. Let  $\tau$  and  $U$  be the arrival time and the size of the next claim, and  $T$  and  $Y$  be the arrival time and the size of the next intensity upward jump. We introduce the auxiliary function

$$\rho^\delta(x) := \max\{x_n^\delta : x_n^\delta \leq x\} \quad (16)$$

which gives the closest point of the grid  $\mathcal{G}_\delta$  below  $x$ . We first define the three possible control actions at any point of  $\mathcal{G}_\delta \times [\underline{\lambda}, \infty)$  as follows:

- Control action  $\mathbf{E}_0$ : Pay no dividends up to the time  $\delta \wedge \tau \wedge T$ .
  1. In the case that  $\delta < \tau \wedge T$ , the surplus at time  $\delta$  is  $x_{n+1}^\delta \in \mathcal{G}_\delta$ .
  2. If  $\delta \wedge T \geq \tau$ , the uncontrolled surplus at time  $\tau$  is  $x_n^\delta + \tau p - U$ ; if this value is positive, the company pays immediately the minimum amount of dividends in such a way that the controlled surplus lies in the closest point of the grid below  $x_n^\delta + \tau p - U$ ; this end surplus can be written  $\rho^\delta(x_n^\delta + \tau p - U)$  and the amount paid as dividends is equal to  $x_n^\delta + \tau p - U - \rho^\delta(x_n^\delta + \tau p - U)$ ; ruin occurs in case the surplus  $x_n^\delta + \tau p - U < 0$  at time  $\tau \leq \delta \wedge T$ .
  3. If  $T < \delta \wedge \tau$ , the end surplus is  $\rho^\delta(x_n^\delta + Tp)$  and the amount paid as dividends is  $x_n^\delta + Tp - \rho^\delta(x_n^\delta + Tp)$ .
- Control actions  $\mathbf{E}_1$ : The company pays immediately  $p\delta$  as dividends, so the controlled surplus becomes  $x_{n-1}^\delta \in \mathcal{G}_\delta$ . The control action  $\mathbf{E}_1$  can only be applied for current surplus  $x_n^\delta > 0$ .
- Control action  $\mathbf{E}_F$ : The manager opts to pay the current surplus  $x_n^\delta$  as dividends and to close the company.

We denote the space of control actions as

$$\mathcal{E} = \{\mathbf{E}_F, \mathbf{E}_1, \mathbf{E}_0\}. \quad (17)$$

Consider  $\tilde{\Pi}_{x_n^\delta, \lambda}^\delta \subset \tilde{\Pi}_{x_n^\delta, \lambda}$  as the set of all the admissible strategies with initial surplus  $x_n^\delta \in \mathcal{G}_\delta$  which can be obtained by a sequence of control actions in  $\mathcal{E}$ . Note that if  $\mathbf{E}_0$  is chosen, the next control action depends on the end surplus and intensity. The length of this sequence could be finite or infinite, in the case that is finite the last control action is either  $\mathbf{E}_0$  in the case that ruin occurs, or  $\mathbf{E}_F$  because the control action  $\mathbf{E}_F$  finishes the sequence.

Finally, we introduce  $\tilde{\Pi}_{x,\lambda}^\delta \subset \tilde{\Pi}_{x,\lambda}$  for any point  $(x, \lambda) \in [0, \infty)$  in the case that  $x \notin \mathcal{G}_\delta$ . Indeed,  $\tilde{\Pi}_{x,\lambda}^\delta$  is the subfamily of admissible strategies  $\tilde{\pi}$  which pays  $x - \rho^\delta(x)$  as dividends immediately and then follows any strategy  $\tilde{\pi}_1$  in  $\tilde{\Pi}_{\rho^\delta(x),\lambda}^\delta$  with  $\rho^\delta(x) \in \mathcal{G}_\delta$ , so we have that

$$J(\tilde{\pi}; x, \lambda) = J(\tilde{\pi}_1; \rho^\delta(x), \lambda) + x - \rho^\delta(x).$$

We define

$$V^\delta(x, \lambda) = \sup_{\tilde{\pi} \in \tilde{\Pi}_{x,\lambda}^\delta} J(\tilde{\pi}; x, \lambda) = V^\delta(\rho^\delta(x), \lambda) + x - \rho^\delta(x). \quad (18)$$

We will show that, in a certain sense,  $\lim_{\delta \rightarrow 0} V^\delta = V$  uniformly.

It is straightforward that  $V^\delta(x, \lambda)$  is non-decreasing in  $x$  and non-increasing in  $\lambda$ . In the next proposition, we find a bound on the variations of  $V^\delta$ , which show in particular that  $V^\delta(x, \cdot)$  is locally Lipschitz.

**Proposition 6.1** *The function  $V^\delta(x, \lambda)$  satisfies*

$$V^\delta(x_2, \lambda) - V^\delta(x_1, \lambda) \leq V^\delta(x_2, \underline{\lambda}) \frac{e^{(\beta+q)\delta + \int_0^\delta \lambda_u^c du} - 1}{p\delta} \left( \rho^\delta(x_2) - \rho^\delta(x_1) \right) + \delta p$$

for  $x_2 \geq x_1 \geq 0$ . Also

$$0 \leq V^\delta(x, \lambda_1) - V^\delta(x, \lambda_2) \leq V^\delta(x, \lambda_1) \frac{\beta\lambda_2 + q}{d(\lambda_1 - \underline{\lambda})} (\lambda_2 - \lambda_1)$$

for  $\lambda_2 \geq \lambda_1 > \underline{\lambda}$  and  $\lambda_2 - \lambda_1 \leq d(\lambda_1 - \underline{\lambda})$ .

As in [12], we will show that the function  $V^\delta$  restricted to  $\mathcal{G}_\delta \times [\underline{\lambda}, \infty)$  is a solution of a discrete version of the HJB equation (11), given in (23). In order to define this discrete HJB equation, let us introduce the operators related to the control actions in  $\mathcal{E}$ . Consider the operators  $\mathcal{T}_0$ ,  $\mathcal{T}_1$  and  $\mathcal{T}_F$  in the set of functions  $\mathcal{W}^\delta = \{w : \mathcal{G}_\delta \times [\underline{\lambda}, \infty) \rightarrow [0, \infty)\}$  which are Lebesgue measurable} defined as follows

$$\mathcal{T}_0(w)(x_n^\delta, \lambda) := \mathbb{P}(\delta \wedge T \wedge \tau = \delta) e^{-q\delta} w(x_{n+1}^\delta, \lambda_\delta^c) + \mathcal{I}^\delta(w)(x_n^\delta, \lambda), \quad (19)$$

$$\mathcal{T}_1(w)(x_n^\delta, \lambda) := w(x_{n-1}^\delta, \lambda) + \delta p \quad \text{and} \quad \mathcal{T}_F(w)(x_n^\delta, \lambda) := x_n^\delta. \quad (20)$$

Here,

$$\begin{aligned} \mathcal{I}^\delta(w)(x, \lambda) &:= \mathbb{E}(I_{\delta \wedge T \wedge \tau = \tau} e^{-q\tau} w(\rho^\delta(x_n^\delta + p\tau - U), \lambda_\tau^c)) \\ &+ \mathbb{E}(I_{\delta \wedge T \wedge \tau = T} e^{-qT} w(x_n^\delta, \lambda_T^c + Y)) \\ &+ \mathbb{E}(I_{\delta \wedge T \wedge \tau = \tau} e^{-q\tau} (x + p\tau - U - \rho^\delta(x_n^\delta + p\tau - U))) \\ &+ \mathbb{E}(I_{\delta \wedge T \wedge \tau = T} e^{-qT} pT), \end{aligned} \quad (21)$$

where  $\tau$  and  $U$  are the arrival time and the size of the next claim, and  $T$  and  $Y$  are the arrival time and the size of the next intensity upward jump. We also introduce the operator  $\mathcal{T}$  in  $\mathcal{W}^\delta$  as

$$\mathcal{T} := \max\{\mathcal{T}_0, \mathcal{T}_1, \mathcal{T}_F\}. \quad (22)$$

Given any family of admissible strategies

$$\tilde{\pi} = \left\{ \tilde{\pi}_{x_n^\delta, \lambda} \in \tilde{\Pi}_{x_n^\delta, \lambda}^\delta \text{ for } (x_n^\delta, \lambda) \in \mathcal{G}_\delta \times [\underline{\lambda}, \infty) \right\},$$

we define the value function  $W : \mathcal{G}_\delta \times [\underline{\lambda}, \infty) \rightarrow \mathbf{R}$  of  $\tilde{\pi}$  as

$$W(x_n^\delta, \lambda) := J(\tilde{\pi}_{x_n^\delta, \lambda}; x_n^\delta, \lambda).$$

$W \in \mathcal{W}^\delta$ ,  $\mathcal{T}_0(W)(x_n^\delta, \lambda)$  and  $\mathcal{T}_1(W)(x_n^\delta, \lambda)$  are the values of the strategies with initial surplus  $x_n^\delta \in \mathcal{G}_\delta$  and initial intensity  $\lambda$  which consist of applying first the control actions  $\mathbf{E}_0$  and  $\mathbf{E}_1 \in \mathcal{E}$  respectively, and afterwards applying the strategy in the family  $\tilde{\pi}$  corresponding to the end surplus and intensity. Also,  $\mathcal{T}_F(W)(x_n^\delta, \lambda)$  is the value function of the control action  $\mathbf{E}_F \in \mathcal{E}$ .

We define the *discrete HJB equation* in  $\mathcal{G}_\delta \times [\underline{\lambda}, \infty)$  as

$$\mathcal{T}(W) - W = 0. \quad (23)$$

Assume that there exists  $\tilde{\pi}_{x_n^\delta, \lambda} \in \tilde{\Pi}_{x_n^\delta, \lambda}^\delta$  such that  $V^\delta(x_n^\delta, \lambda) = J(\tilde{\pi}_{x_n^\delta, \lambda}; x_n^\delta, \lambda)$  for all  $(x_n^\delta, \lambda) \in \mathcal{G}_\delta \times [\underline{\lambda}, \infty)$ . Then, since  $V^\delta \in \mathcal{W}^\delta$  by Proposition 6.1, it is straightforward to see that  $\mathcal{T}(V^\delta)(x_n^\delta, \lambda) = V^\delta(x_n^\delta, \lambda)$ . Also we can find which is the optimal control action in  $\mathcal{E}$  at each  $(x_n^\delta, \lambda) \in \mathcal{G}_\delta \times [\underline{\lambda}, \infty)$ . In Proposition 6.4, we will show that  $\mathcal{T}(V^\delta) = V^\delta$  without the former assumption. In Proposition 6.8, we will show that indeed,  $V^\delta(x_n^\delta, \lambda)$  is the value function of an optimal strategy within  $\tilde{\Pi}_{x_n^\delta, \lambda}^\delta$  and this strategy is stationary.

By definitions (19), (20), (21) and (22), we obtain immediately the following result.



**Proposition 6.2** *The operators  $\mathcal{T}_0, \mathcal{T}_1, \mathcal{T}_F$  and  $\mathcal{T}$  are non-decreasing and  $\mathcal{T}$  satisfies*

$$\sup_{(x_n^\delta, \lambda) \in \mathcal{G}_\delta \times [\underline{\lambda}, \infty)} \left| \mathcal{T}(W_1)(x_n^\delta, \lambda) - \mathcal{T}(W_2)(x_n^\delta, \lambda) \right| \leq \sup_{(x_n^\delta, \lambda) \in \mathcal{G}_\delta \times [\underline{\lambda}, \infty)} \left| W_1(x_n^\delta, \lambda) - W_2(x_n^\delta, \lambda) \right|$$

for any  $W_1$  and  $W_2$  in  $\mathcal{W}^\delta$ .

Given  $l \geq 1$ , let us define  $\tilde{\Pi}_{x_n^\delta, \lambda}^{\delta, l}$  as the set of all the admissible strategies in  $\tilde{\Pi}_{x_n^\delta, \lambda}^\delta$  with initial surplus  $x_n^\delta \in \mathcal{G}_\delta$  and initial intensity  $\lambda \geq \underline{\lambda}$  which can be obtained by a sequence of exactly  $l$  local control actions in  $\mathcal{E}$ .

We define

$$V_l^\delta(x, \lambda) = \sup_{\tilde{\pi} \in \tilde{\Pi}_{x, \lambda}^{\delta, l}} J(\tilde{\pi}; x, \lambda). \quad (24)$$

Since  $\tilde{\Pi}_{x_n^\delta, \lambda}^{\delta, 1} = \{\mathbf{E}_F\}$  then  $V_1^\delta(x, \lambda) = x$ . By Proposition 6.2, we have that  $V_1^\delta \in \mathcal{W}^\delta$  and  $V_2^\delta(x, \lambda) = \mathcal{T}(V_1^\delta)(x_n^\delta, \lambda) \geq \mathcal{T}(V_l^\delta)(x_n^\delta, \lambda)$ . So, we can conclude, by a recursive argument, the following result.

**Proposition 6.3** *It holds that  $V^\delta \geq V_{l+1}^\delta \geq V_l^\delta$ ,  $V_l^\delta \in \mathcal{W}^\delta$  and  $\mathcal{T}(V_l^\delta)(x_n^\delta, \lambda) = V_{l+1}^\delta(x_n^\delta, \lambda)$  for  $l \geq 1$ .*

In the next proposition, we show that  $V^\delta : \mathcal{G}_\delta \times [\underline{\lambda}, \infty) \rightarrow [0, \infty]$  is a solution of the discrete HJB equation (23).

**Proposition 6.4** *It holds that  $\lim_{l \rightarrow \infty} V_l^\delta = V^\delta$  and  $\mathcal{T}(V^\delta)(x_n^\delta, \lambda) = V^\delta(x_n^\delta, \lambda)$ .*

**Remark 6.5** Since  $\mathcal{T}_F(V^\delta)(x_n^\delta, \lambda) = x_n^\delta < V^\delta(x_n^\delta, \lambda)$ , we can write the discrete HJB equation (23) as

$$\max\{\mathcal{T}_0(W) - W, \mathcal{T}_1(W) - W\} = 0,$$

disregarding the operator  $\mathcal{T}_F$ .

**Proposition 6.6** *Given any  $\tilde{\pi} \in \tilde{\Pi}_{x_n^\delta, \lambda}^\delta$  and any supersolution  $\bar{W} : \mathcal{G}_\delta \times [\underline{\lambda}, \infty) \rightarrow \mathbf{R}$  of (23), we have that  $J(\tilde{\pi}; x_n^\delta, \lambda) \leq \bar{W}(x_n^\delta, \lambda)$ .*

From the previous proposition, we deduce the next corollary.

**Corollary 6.7** *The  $\mathcal{G}_\delta$ -optimal value function  $V^\delta : \mathcal{G}_\delta \times [\underline{\lambda}, \infty) \rightarrow \mathbf{R}$  can be characterized as the smallest supersolution of the discrete HJB equation (23) with growth condition (13).*

**Definition 6.1** *Given any partition the set  $\mathcal{G}_\delta \times [\underline{\lambda}, \infty)$  into three measurable subsets  $\mathcal{P} = (\mathcal{A}, \mathcal{NA}, \mathcal{B})$ , we define for any point  $(x_n^\delta, \lambda) \in \mathcal{G}_\delta \times [\underline{\lambda}, \infty)$  the local control action  $S(x_n^\delta, \lambda) \in \mathcal{E}$  in the following way:*

- If  $(x_n^\delta, \lambda) \in \mathcal{B}$ , take  $S_{\mathcal{P}}(x_n^\delta, \lambda) = \mathbf{E}_F$ .  $\mathcal{B}$  is called finish-the-business set.
- If  $(x_n^\delta, \lambda) \in \mathcal{NA}$ , take  $S_{\mathcal{P}}(x_n^\delta, \lambda) = \mathbf{E}_0$ .  $\mathcal{NA}$  is called the non-action set.
- And if  $(x_n^\delta, \lambda) \in \mathcal{A}$ , take  $S_{\mathcal{P}}(x_n^\delta, \lambda) = \mathbf{E}_1$ . Note that  $(x_0^\delta, \lambda) = (0, \lambda) \notin \mathcal{A}$ .  $\mathcal{A}$  is called the action set.

The  $\mathcal{G}_\delta$ -strategy  $\pi_{x_n^\delta, \lambda}^{\mathcal{P}} \in \Pi_{x_n^\delta, \lambda}^\delta$  associated to  $\mathcal{P}$  in the initial surplus and intensity  $(x_n^\delta, \lambda) \in \mathcal{G}_\delta \times [\underline{\lambda}, \infty)$  is defined inductively as follows: Let us call  $m_1 = (x_n^\delta, \lambda)$  and  $s_1 = S_{\mathcal{P}}(x_n^\delta, \lambda)$ ; assuming that  $m_1, m_2, \dots, m_{k-1} \in \mathcal{G}_\delta \times [\underline{\lambda}, \infty)$  and  $s_1, s_2, \dots, s_{k-1} \in \mathcal{E}$  are defined and the process does not stop at step  $k-1$ , we define  $m_k \in \mathcal{G}_\delta \times [\underline{\lambda}, \infty)$  as the end surplus of  $s_{k-1}$  and  $s_k = S_{\mathcal{P}}(m_k) \in \mathcal{E}$ .

Note that the family  $\pi^{\mathcal{P}} = \left( \pi_{x_n^\delta, \lambda}^{\mathcal{P}} \right)_{(x_n^\delta, \lambda) \in \mathcal{G}_\delta \times [\underline{\lambda}, \infty)}$  is stationary in  $\mathcal{G}_\delta \times [\underline{\lambda}, \infty)$  in the sense that the local control actions depend only on the point of the grid at which the current surplus lies and also on the current intensity. Moreover, if we define the associated value function  $W_{\mathcal{P}}(x_n^\delta, \lambda) = J(\pi_{x_n^\delta, \lambda}^{\mathcal{P}}; x_n^\delta, \lambda)$ , then  $\mathcal{T}_F(W_{\mathcal{P}}) = W_{\mathcal{P}}$  in  $\mathcal{B}$ ,  $\mathcal{T}_1(W_{\mathcal{P}}) = W_{\mathcal{P}}$  in  $\mathcal{A}$  and  $\mathcal{T}_0(W_{\mathcal{P}}) = W_{\mathcal{P}}$  in  $\mathcal{NA}$ . We extend the definition of the value function  $W_{\mathcal{P}} : [0, \infty) \times [\underline{\lambda}, \infty) \rightarrow [0, \infty)$  as

$$W_{\mathcal{P}}(x, \lambda) = W_{\mathcal{P}}(\rho^\delta(x), \lambda) + x - \rho^\delta(x), \quad (25)$$

this corresponds to pay the minimum amount of dividends so the surplus lies in  $\mathcal{G}_\delta$ .

Given the function  $V^\delta$ , since  $V^\delta = \mathcal{T}(V^\delta)$  and  $V^\delta$  is Lebesgue measurable, we define the  $\mathcal{G}_\delta$ -partition  $\mathcal{P}_\delta^* = (\mathcal{A}_\delta^*, (\mathcal{NA})_\delta^*, \mathcal{B}_\delta^*)$  as

$$\begin{aligned} \mathcal{A}_\delta^* &= \{ (x_n^\delta, \lambda) \in \mathcal{G}_\delta \times [\underline{\lambda}, \infty) : \mathcal{T}_1(V^\delta)(x_n^\delta, \lambda) = V^\delta(x_n^\delta, \lambda) \}, \\ (\mathcal{NA})_\delta^* &= \{ (x_n^\delta, \lambda) \in (\mathcal{G}_\delta \times [\underline{\lambda}, \infty)) - \mathcal{A}_\delta^* : \mathcal{T}_0(V^\delta)(x_n^\delta, \lambda) = V^\delta(x_n^\delta, \lambda) \}, \\ \mathcal{B}_\delta^* &= \{ (x_n^\delta, \lambda) \in (\mathcal{G}_\delta \times [\underline{\lambda}, \infty)) - (\mathcal{A}_\delta^* \cup (\mathcal{NA})_\delta^*) : \mathcal{T}_F(V^\delta)(x_n^\delta, \lambda) = V^\delta(x_n^\delta, \lambda) \}. \end{aligned} \quad (26)$$

Note that, by Remark 6.5,  $\mathcal{B}_\delta^*$  is empty, so  $(x_0^\delta, \lambda) = (0, \lambda) \in (\mathcal{NA})_\delta^*$ . Since the value function  $W_{\mathcal{P}_\delta^*}(x_n^\delta, \lambda)$  is a supersolution of the discrete HJB equation (23) with growth condition (13), we deduce from Corollary 6.7, the following result.

**Proposition 6.8** *We have that  $V^\delta = W_{\mathcal{P}_\delta^*}$  in  $\mathcal{G}_\delta \times [\underline{\lambda}, \infty)$ , and so  $\pi^{\mathcal{P}_\delta^*} = \left( \pi_{x_n^\delta, \lambda}^{\mathcal{P}_\delta^*} \right)_{(x_n^\delta, \lambda) \in \mathcal{G}_\delta \times [\underline{\lambda}, \infty)}$  is the  $\mathcal{G}_\delta$ -optimal strategy.  $\mathcal{P}_\delta^*$  is called the optimal  $\mathcal{G}_\delta$ -partition.*

Let us prove now that the optimal value function  $V$  can be approximated uniformly by  $V^\delta$  for some  $\delta$  small enough. Since we have a monotonicity condition on the embedded grids  $\mathcal{G}_{\delta/2} \subset \mathcal{G}_\delta$ , so  $x_n^\delta = x_{2n}^{\delta/2}$ ,  $\tilde{\Pi}_{x_n^\delta, \lambda}^\delta \subset \tilde{\Pi}_{x_{2n}^{\delta/2}, \lambda}^{\delta/2}$ , and this implies that  $V^\delta(x_n^\delta, \lambda) \leq V^{\delta/2}(x_{2n}^{\delta/2}, \lambda)$ . Take  $\delta_k := \delta/2^k$  for  $k \geq 0$ . We will see that  $V^{\delta_k} \nearrow V$  locally uniformly as  $k$  goes to infinity. Consider the dense set in  $\mathbf{R}_+^n$ ,  $\mathcal{G} := \bigcup_{k \geq 0} \mathcal{G}_{\delta_k}$ . Note that  $\mathcal{G}_{\delta_k} \subset \mathcal{G}_{\delta_{k+1}}$ , so

$$V^{\delta_k} \leq V^{\delta_{k+1}} \leq V.$$

Now we conclude the main result of the paper. The proof is in Appendix A.5.

**Theorem 6.9** *For any  $\delta > 0$ , the functions  $V^{\delta_k} \nearrow V$  uniformly as  $k$  goes to infinity.*

## 7 Discretization on the intensity process

As we pointed out before, the construction of  $V^\delta$  given in the previous section only uses a discretization on the surplus space. In this section, we propose a numerical scheme using a discretization on the intensity space as well. We will find a partition into an action set, a non-action set and a finish-the-business set as introduced in Definition 6.1, depending on both the discretization in the surplus space and the discretization in the intensity space, whose value function approximates the value function  $V$  uniformly. The proofs of the results of this section are deferred to Appendix A.6.

For any parameter  $\Delta > 0$ , let us introduce the following discretization on the intensity space,

$$\mathcal{H}_\Delta = \left\{ \lambda_0^\Delta = \underline{\lambda}, \lambda_1^\Delta, \lambda_2^\Delta, \lambda_3^\Delta, \dots \right\} \subset [\underline{\lambda}, \infty). \quad (27)$$

where  $\lambda_m^\Delta = \underline{\lambda} + m\Delta$ , and consider the function

$$\sigma^\Delta(\lambda) = \min \left\{ \lambda_m^\Delta \in \mathcal{H}_\Delta : \lambda_m^\Delta \geq \lambda \right\} \in \mathcal{H}_\Delta. \quad (28)$$

**Definition 7.1** *Given  $\delta > 0$  and  $\Delta > 0$ , we say that  $\mathcal{P} = (\mathcal{A}, \mathcal{N}\mathcal{A}, \mathcal{B})$  is a  $(\delta, \Delta)$ -partition if the three subsets  $\mathcal{A}$ ,  $\mathcal{N}\mathcal{A}$  and  $\mathcal{B}$  satisfy the following condition: if  $(x_n^\delta, \lambda_m^\Delta)$  is in a subset for  $m \geq 1$ , then  $\{x_n^\delta\} \times (\lambda_{m-1}^\Delta, \lambda_m^\Delta]$  is also in this subset. So, the  $(\delta, \Delta)$ -partitions only depend on the discrete grid  $\mathcal{G}_\delta \times \mathcal{H}_\Delta$ .*

In order to find a  $(\delta, \Delta)$ -partition whose associated value function approximates the optimal value function, we modify the intensity process, defining a new intensity process in the grid domain  $\mathcal{G}_\delta \times \mathcal{H}_\Delta$ .

Given the process  $\lambda_t$  from (2) with initial value  $\lambda_0 = \lambda$ , let us define the auxiliary intensity process  $\hat{\lambda}_t$  for given parameters  $\delta, \Delta$  and initial  $\hat{\lambda}_0 = \sigma^\Delta(\lambda) = \lambda_m^\Delta \in \mathcal{H}_\Delta$  for some  $m \geq 0$  as

$$\hat{\lambda}_t := \sigma^\Delta(\lambda_t) \in \mathcal{H}_\Delta. \quad (29)$$

By definition,  $\Delta \geq \hat{\lambda}_t - \lambda_t \geq 0$ . We define  $\hat{\Pi}_{x_n^\delta, \lambda_m^\Delta}^{\delta, \Delta}$  as the set of all the admissible strategies with initial surplus  $x_n^\delta \in \mathcal{G}_\delta$ , initial intensity  $\hat{\lambda}_0 = \lambda_m^\Delta \in \mathcal{H}_\Delta$  and discrete intensity process  $\hat{\lambda}_t$  which can be obtained by a sequence of control actions in  $\mathcal{E}$  as in (17).

We can replicate the construction of  $V^\delta$  in the previous section, but considering the compound Cox filtered space of the auxiliary intensity process  $\hat{\lambda}_t \in \mathcal{H}_\Delta$  instead of the process  $\lambda_t$ . Consider for any  $(x_n^\delta, \lambda_m^\Delta) \in \mathcal{G}_\delta \times \mathcal{H}_\Delta$ ,

$$\hat{V}^{\delta, \Delta}(x_n^\delta, \lambda_m^\Delta) = \sup_{\pi \in \hat{\Pi}_{x_n^\delta, \lambda_m^\Delta}^{\delta, \Delta}} J_{\hat{\lambda}}(\pi; x_n^\delta, \lambda_m^\Delta), \quad (30)$$

where

$$J_{\hat{\lambda}}(\pi; x_n^\delta, \lambda_m^\Delta) = \mathbb{E} \left( \int_{0-}^{\tau^L} e^{-qt} dL_t + I_{\tau < \tau_F} e^{-q\tau_F} \hat{X}_{\tau_F}^L \right) \quad (31)$$

and

$$\hat{X}_t^L = \hat{X}_t - L_t \quad (32)$$

with

$$\hat{X}_t = x_n^\delta + pt - \sum_{j=1}^{\hat{N}_t} U_j, \quad (33)$$

and  $\hat{N}_t$  has intensity  $\hat{\lambda}_t$ .

One can extend the definition of  $\widehat{V}^{\delta,\Delta}$  to  $[0, \infty) \times [\underline{\lambda}, \infty)$  as

$$\widehat{V}^{\delta,\Delta}(x, \lambda) = (x - \rho^\delta(x)) + \widehat{V}^{\delta,\Delta}(\rho^\delta(x), \sigma^\Delta(\lambda))$$

for all  $x \geq 0$  and  $\lambda \geq \underline{\lambda}$ , where  $\rho^\delta$  is defined in (16).

Let us show now that  $\widehat{V}^{\delta,\Delta}$  converges uniformly to  $V^\delta$  as  $\Delta \rightarrow 0$ . The proof is in Appendix A.6.

**Proposition 7.1** *It holds that  $\widehat{V}^{\delta,\Delta} \leq V^\delta$  and  $\lim_{\Delta \rightarrow 0} \left( \sup_{x \geq 0, \lambda \geq \underline{\lambda}} V^\delta(x, \lambda) - \widehat{V}^{\delta,\Delta}(x, \lambda) \right) = 0$ .*

As in Section 6, but considering  $\widehat{\lambda}_t$  instead of  $\lambda_t$ , we introduce the following operators in  $\mathcal{W}^{\delta,\Delta} = \{w : \mathcal{G}_\delta \times \mathcal{H}_\Delta \rightarrow [0, \infty)\}$ :  $\mathcal{T}_1$  and  $\mathcal{T}_F$  as defined in (20); the operator

$$\widehat{\mathcal{T}}_0(w)(x_n^\delta, \lambda_m^\Delta) := \mathbb{P}(\delta \wedge T \wedge \tau = \delta) e^{-q\delta} w(x_{n+1}^\delta, \sigma^\Delta(\lambda_\delta^c)) + \widehat{\mathcal{I}}(w)(x_n^\delta, \lambda_m^\Delta),$$

where

$$\begin{aligned} \widehat{\mathcal{I}}(w)(x_n^\delta, \lambda_m^\Delta) &:= \mathbb{E}(I_{\delta \wedge T \wedge \tau = \tau} I_{x+p\tau-U \geq 0} e^{-q\tau} w(\rho^\delta(x_n^\delta + p\tau - U), \widehat{\lambda}_\tau)) \\ &+ \mathbb{E}(I_{\delta \wedge T \wedge \tau = T} e^{-qT} w(x_n^\delta, \widehat{\lambda}_T)) \\ &+ \mathbb{E}(I_{\delta \wedge T \wedge \tau = \tau} e^{-q\tau} (x_n^\delta + p\tau - U - \rho^\delta(x_n^\delta + p\tau - U))) \\ &+ \mathbb{E}(I_{\delta \wedge T \wedge \tau = T} e^{-qT} pT); \end{aligned}$$

and the operator

$$\widehat{\mathcal{T}} := \max\{\widehat{\mathcal{T}}_0, \mathcal{T}_1, \mathcal{T}_F\}.$$

In the next two propositions, we obtain the mirror results of Propositions 6.4 and 6.8, with similar proofs.

**Proposition 7.2** *It holds that  $\widehat{\mathcal{T}}(\widehat{V}^{\delta,\Delta}) = \widehat{V}^{\delta,\Delta}$  in  $\mathcal{G}_\delta \times \mathcal{H}_\Delta$ .*

Since  $\widehat{V}^{\delta,\Delta} = \mathcal{T}(\widehat{V}^{\delta,\Delta})$ , we can define, as in (26), the following partition  $\widehat{\mathcal{P}}^{\delta,\Delta}$  in the grid  $\mathcal{G}_\delta \times \mathcal{H}_\Delta$ :

$$\begin{aligned} \widehat{\mathcal{A}}_{\delta,\Delta} &= \left\{ (x_n^\delta, \lambda_m^\Delta) \in \mathcal{G}_\delta \times \mathcal{H}_\Delta : \mathcal{T}_1(\widehat{V}^{\delta,\Delta})(x_n^\delta, \lambda_m^\Delta) = \widehat{V}^{\delta,\Delta}(x_n^\delta, \lambda_m^\Delta) \right\}, \\ \widehat{\mathcal{N}}\widehat{\mathcal{A}}_{\delta,\Delta} &= \left\{ (x_n^\delta, \lambda_m^\Delta) \in (\mathcal{G}_\delta \times \mathcal{H}_\Delta) - \widehat{\mathcal{A}}_{\delta,\Delta} : \widehat{\mathcal{T}}_0(\widehat{V}^{\delta,\Delta})(x_n^\delta, \lambda_m^\Delta) = \widehat{V}^{\delta,\Delta}(x_n^\delta, \lambda_m^\Delta) \right\}, \\ \widehat{\mathcal{B}}_{\delta,\Delta} &= \left\{ (x_n^\delta, \lambda_m^\Delta) \in (\mathcal{G}_\delta \times \mathcal{H}_\Delta) - (\widehat{\mathcal{A}}_{\delta,\Delta} \cup \widehat{\mathcal{N}}\widehat{\mathcal{A}}_{\delta,\Delta}) : \mathcal{T}_F(\widehat{V}^{\delta,\Delta})(x_n^\delta, \lambda_m^\Delta) = \widehat{V}^{\delta,\Delta}(x_n^\delta, \lambda_m^\Delta) \right\}. \end{aligned} \quad (34)$$

Moreover,  $\widehat{\mathcal{B}}_{\delta,\Delta} = \emptyset$  and  $(x_0^\delta, \lambda_m^\Delta) \in \widehat{\mathcal{N}}\widehat{\mathcal{A}}_{\delta,\Delta}$ . And, as in Definition 6.1, we can define from this partition, a family of strategies

$$\widehat{\pi}^{\delta,\Delta} = \left\{ \widehat{\pi}_{x_n^\delta, \lambda_m^\Delta}^{\delta,\Delta} \in \widehat{\Pi}_{x_n^\delta, \lambda_m^\Delta}^{\delta,\Delta} \text{ for } (x_n^\delta, \lambda_m^\Delta) \in \mathcal{G}_\delta \times \mathcal{H}_\Delta \right\}.$$

These strategies are stationary in  $\mathcal{G}_\delta \times \mathcal{H}_\Delta$  in the sense that the local control actions depend only on the point of the grid at which the current surplus and current intensity lie. We obtain the existence of the optimal strategy in  $\widehat{\Pi}_{x_n^\delta, \lambda_m^\Delta}^{\delta,\Delta}$ .

**Proposition 7.3**  $\widehat{V}^{\delta,\Delta}(x_n^\delta, \lambda_m^\Delta) = J_{\widehat{\lambda}}(\widehat{\pi}_{x_n^\delta, \lambda_m^\Delta}^{\delta,\Delta}; x_n^\delta, \lambda_m^\Delta)$  for any  $(x_n^\delta, \lambda_m^\Delta) \in \mathcal{G}_\delta \times \mathcal{H}_\Delta$ .

Now we use the partition given in (26), to find a  $(\delta, \Delta)$ -partition in  $\mathcal{G}_\delta \times [\underline{\lambda}, \infty)$  as defined in Definition 7.1 whose value function approximates the value value function  $V$  uniformly. This is the main result of the section.

**Definition 7.2** *Given the partition (34) in  $\mathcal{G}_\delta \times \mathcal{H}_\Delta$ , we extend to the partition  $\mathcal{P}^{\delta,\Delta} = (\mathcal{A}_{\delta,\Delta}, (\mathcal{N}\mathcal{A})_{\delta,\Delta}, \mathcal{B}_{\delta,\Delta})$  in  $\mathcal{G}_\delta \times [\underline{\lambda}, \infty)$  as*

$$\begin{aligned} \mathcal{A}_{\delta,\Delta} &= \left( \bigcup_{\{(n,m):(x_n^\delta, \lambda_m^\Delta) \in \widehat{\mathcal{A}}_{\delta,\Delta}, m \geq 1\}} (\{x_n^\delta\} \times (\lambda_{m-1}^\Delta, \lambda_m^\Delta]) \right) \cup \left( \bigcup_{\{n: (x_n^\delta, \underline{\lambda}) \in \widehat{\mathcal{A}}_{\delta,\Delta}\}} \{(x_n^\delta, \underline{\lambda})\} \right) \\ (\mathcal{N}\mathcal{A})_{\delta,\Delta} &= \left( \bigcup_{\{(n,m):(x_n^\delta, \lambda_m^\Delta) \in \widehat{\mathcal{N}}\widehat{\mathcal{A}}_{\delta,\Delta}, m \geq 1\}} (\{x_n^\delta\} \times (\lambda_{m-1}^\Delta, \lambda_m^\Delta]) \right) \cup \left( \bigcup_{\{n: (x_n^\delta, \underline{\lambda}) \in \widehat{\mathcal{N}}\widehat{\mathcal{A}}_{\delta,\Delta}\}} \{(x_n^\delta, \underline{\lambda})\} \right) \\ \mathcal{B}_{\delta,\Delta} &= \emptyset. \end{aligned}$$

Let us consider the  $\mathcal{G}_\delta$ - strategy  $\pi_{x_n^\delta, \lambda}^{\delta,\Delta} \in \widetilde{\Pi}_{x_n^\delta, \lambda}^{\delta,\Delta}$  associated to  $\mathcal{P}^{\delta,\Delta}$  for each  $(x_n^\delta, \lambda) \in \mathcal{G}_\delta \times [\underline{\lambda}, \infty)$  and the associated function  $W_{\mathcal{P}^{\delta,\Delta}} : [0, \infty) \times [\underline{\lambda}, \infty) \rightarrow [0, \infty)$  defined in (25).

The next theorem states that  $V$  can be approximated uniformly by  $W_{\mathcal{P}^{\delta,\Delta}}$  for some  $\delta$  and  $\Delta$  small enough. In Section 8, we use this result in order to obtain numerically  $(\delta, \Delta)$ -partitions whose associated value functions approximate uniformly the optimal value function  $V$ .

**Theorem 7.4** *For any  $\varepsilon > 0$  there exists  $\delta$  and  $\Delta$  small enough so that  $0 \leq V - W_{\mathcal{P}^{\delta,\Delta}} \leq V - \widehat{V}^{\delta,\Delta} \leq \varepsilon$  in  $[0, \infty) \times [\underline{\lambda}, \infty)$ .*

## 8 Numerical results

In this section, we show some numerical results. Proposition 5.2 suggests that we can approximate the optimal value function with a value function of a partition with  $(x_n^\delta, \lambda_m^\Delta) \in \mathcal{B}_{\delta, \Delta}$  for  $m \leq m_1$  with  $m_1$  large enough and  $n \geq 0$ . From Proposition 5.1, we can assume that the partition satisfies  $(x_n^\delta, \lambda_m^\Delta) \in (\mathcal{NA})_{\delta, \Delta}$  for  $x_n^\delta > p/q$ . This allows to make a numerical computation because we stay in a finite grid.

So, given  $\delta, \Delta$  small enough and  $m_1$  large enough, we proceed in the following way: The numerical scheme should choose if the local action in each of the points in the grid

$$\bar{\mathcal{G}}_\delta \times \bar{\mathcal{H}}_\Delta = \left\{ (x_n^\delta, \lambda_m^\Delta) : x_n^\delta \leq p/q \text{ and } m \leq m_1 \right\}$$

is either  $\mathbf{E}_0$  or  $\mathbf{E}_1$  or  $\mathbf{E}_F$ . In order to do that, we proceed inductively, as follows:

1. We define the initial partition  $\hat{\mathcal{P}}_1$  as  $\mathcal{B}_1 = \bar{\mathcal{G}}_\delta \times \bar{\mathcal{H}}_\Delta$ ,  $\widehat{\mathcal{NA}}_1 = \hat{\mathcal{A}}_1 = \emptyset$  for the sake of simplicity. So, the value function is  $\widehat{W}_{\hat{\mathcal{P}}_1}(x_n^\delta, \lambda_m^\Delta) = x_n^\delta$ .
2. If the partition  $\hat{\mathcal{P}}_l$  and the associated value function  $\widehat{W}_{\hat{\mathcal{P}}_l}$  are given, we define the next partition  $\hat{\mathcal{P}}_{l+1}$  as

$$\begin{aligned} \hat{\mathcal{A}}_{l+1} &= \{(x_n^\delta, \lambda_m^\Delta) \in \bar{\mathcal{G}}_\delta \times \bar{\mathcal{H}}_\Delta : \hat{\mathcal{T}}(\widehat{W}_{\hat{\mathcal{P}}_l})(x_n^\delta, \lambda_m^\Delta) = \hat{\mathcal{T}}_1(\widehat{W}_{\hat{\mathcal{P}}_l})(x_n^\delta, \lambda_m^\Delta)\}, \\ \widehat{\mathcal{NA}}_{l+1} &= \{(x_n^\delta, \lambda_m^\Delta) \in \bar{\mathcal{G}}_\delta \times \bar{\mathcal{H}}_\Delta - \hat{\mathcal{A}}_{l+1} : \hat{\mathcal{T}}(\widehat{W}_{\hat{\mathcal{P}}_l})(x_n^\delta, \lambda_m^\Delta) = \hat{\mathcal{T}}_0(\widehat{W}_{\hat{\mathcal{P}}_l})(x_n^\delta, \lambda_m^\Delta)\}, \end{aligned}$$

and we put  $\hat{\mathcal{B}}_{l+1} = \emptyset$ , because  $\mathbf{E}_F$  is never optimal.

We have, as in Propositions 6.3 and 6.4, that  $\widehat{W}_{\hat{\mathcal{P}}_{l+1}} = \hat{\mathcal{T}}(\widehat{W}_{\hat{\mathcal{P}}_l}) \geq \widehat{W}_{\hat{\mathcal{P}}_l}$  in  $\bar{\mathcal{G}}_\delta \times \bar{\mathcal{H}}_\Delta$ ; that  $\hat{\mathcal{P}}_{l+1} = \hat{\mathcal{P}}_l = \hat{\mathcal{P}}$  for  $l$  large enough because here the grid  $\bar{\mathcal{G}}_\delta \times \bar{\mathcal{H}}_\Delta$  is finite; that  $\lim_{l \rightarrow \infty} \widehat{W}_{\hat{\mathcal{P}}_l} = \widehat{W}_{\hat{\mathcal{P}}}$  and that  $\hat{\mathcal{T}}(\widehat{W}_{\hat{\mathcal{P}}}) = \widehat{W}_{\hat{\mathcal{P}}}$  in  $\bar{\mathcal{G}}_\delta \times \bar{\mathcal{H}}_\Delta$ . In our numerical scheme, we also check that the limit action region of  $\hat{\mathcal{P}}$  does not change when  $m_1$  is enlarged.

By Definition 7.2 and Theorem 7.4, we know that when  $\delta$  and  $\Delta$  are small enough we obtain a near-optimal value function and  $\mathcal{G}_\delta$ - strategy. We check in our numerical scheme that  $\delta$  and  $\Delta$  are small enough by comparing the value functions  $\widehat{W}_{\hat{\mathcal{P}}}$  for  $(\delta, \Delta)$  with the one of  $(\delta/2, \Delta/2)$ . In the following examples, we show in gray the non-action region and in black the change region of the partition  $\mathcal{P}$  in  $\bar{\mathcal{G}}_\delta \times [\underline{\lambda}, \lambda_{m_1}^\Delta]$  associated to the partition  $\hat{\mathcal{P}}$  in  $\bar{\mathcal{G}}_\delta \times \bar{\mathcal{H}}_\Delta$  as in Definition 7.2. We also show the approximation of the optimal value function  $V$  (using the value function of the partition  $\hat{\mathcal{P}}$  in  $\bar{\mathcal{G}}_\delta \times \bar{\mathcal{H}}_\Delta$ ) and finally, we compare the approximation of our optimal value function with one of the classical dividend optimization problem with constant intensity, i.e. the situation without jumps in the claim arrival intensity.

### 8.1 Example 1: Exponential claim sizes

Let us first consider  $\underline{\lambda} = 1/4$ , the intensity jump distribution being exponential with cdf  $F_Y(x) = 1 - e^{-x/2}$ , the intensity of jump intensity arrivals being  $\beta = 1/2$  (so we expect a catastrophe every two years),  $d = 7/10$  (so that the additional claim arrival intensity due to a catastrophe is halved after one year), the claim size distribution being exponential with cdf  $F_U(x) = 1 - e^{-10x}$ , the discount rate being  $q = 2/10$  and the insurance safety loading applied in the policies equal to  $\eta = 2/10$ . We then obtain from (10) that  $p = 141/700$ . For the grid parameters, it turns out that  $\delta = 28/423$ ,  $\Delta = 23/240$  and  $m_1 = 60$  is appropriate here.

Figure 8.1a depicts the action and non-action region and Figure 8.1b the approximation of the optimal value function  $V$  as a function of initial surplus  $x$  and claim initial intensity level  $\lambda$ . In the absence of catastrophe jumps in the intensity process (that is, for the classical Cramér-Lundberg process with a compound Poisson claim process), it is well-known that for exponential claim sizes, a barrier strategy is optimal (cf. Gerber [19]). With the additional presence of a shot-noise component in the intensity process, one observes from Figure 8.1a that a barrier strategy is still optimal, but its value changes dynamically as a function of current claim intensity  $\lambda$ . Concretely, for small values of  $\lambda$ , a barrier strategy is optimal that pays all the surplus above the barrier as dividends, and this barrier first increases with  $\lambda$  and eventually decreases for larger values of  $\lambda$ . Finally, there is a critical value of  $\lambda$ , above which the situation becomes too risky, and the barrier level goes down to zero, i.e. all surplus is paid out as dividends. For comparison, in Figure 8.1a we also depict the optimal barrier level of the classical risk model for the same premium income  $p$ , but varying actual (constant in time) level  $\lambda$  (blue solid line), and one sees that there the optimal barrier level already goes down to zero for a considerably smaller value of  $\lambda$ . One can nicely see from this comparison how the dynamic change of the level of  $\lambda$  through time allows to be more adaptive in the strategy.

In Figure 8.1c we depict the resulting value function as a function of initial capital  $x$ , where we choose  $\lambda = \underline{\lambda} = 1/4$  as the initial level (solid line, which is in fact just the cross-section  $V(x, 0.4)$  from Figure 8.1b)

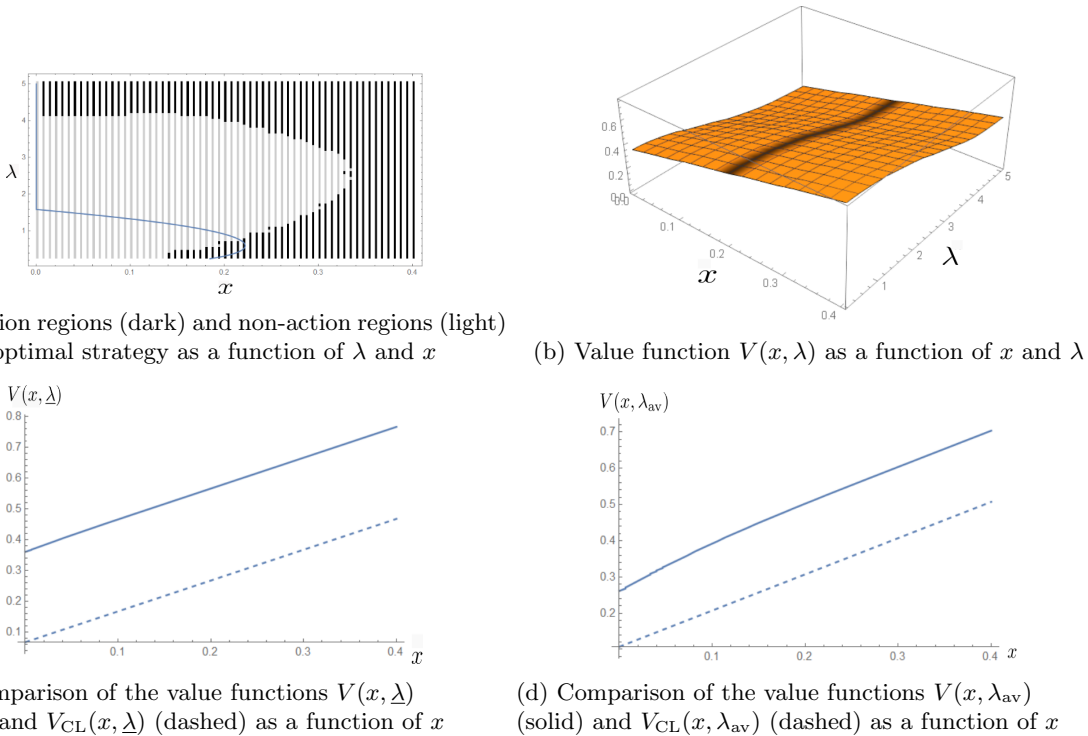


Figure 8.1: Optimal dividends for a compound shot-noise Cox process with exponential claim sizes

and compare it to the value function  $V_{CL}(x, \lambda)$  for the classical risk model with homogeneous intensity  $\lambda$  throughout (dashed line), which corresponds to the case  $\beta = 0$  of the shot-noise case here. Note that the same relative safety loading  $\eta$  is applied in the two cases, so the absolute value of  $p$  does not coincide for the two models (i.e.,  $V_{CL}(x, \lambda)$  does not correspond to  $v^\lambda(x)$ , for which that was the case). In other words, Figure 8.1c compares the economic value of the insurance company for the shareholders with and without additional catastrophe insurance business, and the plot shows that including that business line is an advantage. This can be interpreted as follows: the additional insurance business linked to catastrophe claims (above the baseline intensity  $\lambda$ ) leads to additional premium income, which is upfront (as potential claims will only appear later when the intensity indeed jumped up), so that one has additional degrees of freedom to steer the dividend streams according to the current level of intensity and surplus. In fact, the additional variability in the intensity process stays advantageous even when starting at higher levels of the initial intensity (and the comparison then is whether a volatile claim intensity can be preferable to a constant one, for the same number of policies). To see this, Figure 8.1d compares the two value functions for the long-term average value  $\lambda_{av} = 1/4 + 2/(2 \cdot 0.7) = 1.679$  as the initial intensity level (solid line) or as homogeneous intensity level throughout (dashed line), respectively (again with safety loading  $\eta = 0.2$  for both, which now also leads to the same  $p$ , so here  $V_{CL}(x, \lambda_{av}) = v^{\lambda_{av}}(x)$ ). Clearly, the additional variability of the claim intensity process (accompanied by the optimal associated dividend strategy) leads still to a significantly larger value function.

If one replaces the exponential intensity jump distribution by a deterministic jump of the same expected value (i.e.  $F_Y(x) = I_{[2, \infty)}$ ), the modified action/non-action regions are shown in Figure 8.2, which shows that the results are not really sensitive to the choice of the concrete intensity jump distribution (as long as the expected jump size is maintained). The resulting value function is even visually indistinguishable from Figure 8.1b, so that we do not include it here.

In order to assess the sensitivity of the action region with respect to changes of other parameters, we depict in Figure 8.3 the counterpart of Figure 8.1a for slight changes of the parameters in either direction. We also plot the optimal barrier of the classical Cramér-Lundberg model for the same premium level  $p$  for each case. One can see that increasing or decreasing the intensity  $\beta$  of catastrophe arrivals has a considerable effect on the optimal barrier as a function of  $x$  and  $\lambda$ . A smaller value of the decay rate  $d$  in the intensity function leads to a more dangerous situation, enlarging the barrier levels, where for increasing  $d$  the intensity growth due to a catastrophe disappears quicker and we get closer to the classical barrier level (in blue), which would be reached for  $d \rightarrow \infty$ . A higher safety loading  $\eta$  leads to larger barrier levels (one may interpret that this

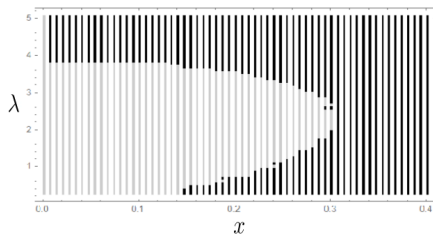


Figure 8.2: Counterpart of Figure 8.1a for deterministic intensity jumps of size 2

is due to the fact that one wants to benefit from the positive drift longer, exceeding the advantage of paying out profits early). Finally, an increased discount rate makes earlier payments more important and attractive, leading to lower barrier levels.

## 8.2 Example 2: Erlang(2) claim sizes

Next, we consider a situation of a certain Erlang(2) claim size distribution, for which we know from [11] that a two-band strategy maximizes the expected discounted dividend payments in the absence of shot-noise jumps in the Poisson intensity. Consider therefore the parameters choices of [11]  $\lambda = 10$ ,  $F_U(x) = 1 - (1+x)e^{-x}$ ,  $q = 1/10$  and  $\eta = 7/100$ , to which we add now a catastrophe shot-noise component with exponential intensity jumps with cdf  $F_Y(x) = 1 - e^{-x/2}$ ,  $\beta = 2/10$  and  $d = 2/10$ . We obtain from (10) that  $p = 642/25$ .

For the grid parameters, the values  $\delta = 25/963$ ,  $\Delta = 1/2$  and  $m_1 = 60$  turn out to be appropriate here. Figure 8.4a depicts the action and non-action region. One observes that the optimality of the two-band regime is in fact retained here also, but only for a small strip of  $\lambda$ -values, and in that region the corresponding band values vary with  $\lambda$  also. For values of  $\lambda$  below that regime, the two-band strategy collapses to a barrier strategy (the fact that this band strategy is very sensitive to the particular model assumptions is well-known, see e.g. also [5] for such an effect under discrete surplus observations). For values of  $\lambda$  above that regime, the optimal strategy is again of take-the-money-and-run type, i.e. pay all the surplus as dividends immediately (as the risk of facing many claims in the near future diminishing the surplus is too high). Figure 8.4b depicts the resulting value function as a function of current levels of  $x$  and  $\lambda$ . Finally, Figure 8.4c compares  $V(x, \lambda_{av})$  to the value function of the classical risk model with constant intensity  $\lambda_{av}$  and the same premium income. Again, we observe that the additional variability introduced in the portfolio by having catastrophe insurance business (rather than only non-catastrophic one with constant claim intensity) is in fact an advantage for the expected discounted dividend payments until ruin.

## 8.3 Example 3: Deterministic claim sizes

In addition to a fine interplay of many factors, one intuitive reason for the optimality of band (rather than barrier) strategies can be attributed to the presence of modes in the claim size density (see e.g. [13, 6] where in the latter reference even a situation with an optimal strategy consisting of 4 bands was identified). Following that line of thinking, one might expect a 2-band strategy to remain optimal when replacing the Erlang claim size distribution by a deterministic claim size equal to its expected value. We therefore reconsider the situation of Example 2, solely changing the cdf of the claim size distribution to  $F_U(x) = I_{[2, \infty)}$ . We use the same grid parameters as above:  $\delta = 25/963$ ,  $\Delta = 1/2$  and  $m_1 = 60$ . Figure 8.5a depicts the action and non-action region. Indeed, the result is very similar to Figure 8.4a, so for a certain range of  $\lambda$ -values, a two-band strategy is optimal. Yet, the lack of variability of the claim size is reflected in different numerical values of the  $\lambda$ -dependent bands. Figure 8.5b depicts the resulting value function as a function of current levels of  $x$  and  $\lambda$ , which again has a very similar shape, but different absolute values. As before, Figure 8.5c compares  $V(x, \lambda_{av})$  to the value function of the model without a catastrophe component.

If instead we replace the exponential claim size distribution in Example 1 above by its deterministic counterpart equal to the expected value  $1/10$  (i.e.  $F_U(x) = I_{[1/10, \infty)}$ ), but keep all other parameters as in Example 1, we arrive at Figure 8.6a. In this example one can nicely see the signature of the deterministic claim size equal to 0.1 in the shape of the action/non-action regions. For fixed  $\lambda$ , one can also see how multiple bands can appear here, leading to a surprisingly aesthetic butterfly shape. Note, however, that when the optimal strategy is applied dynamically, both the values of  $x$  and  $\lambda$  change instantaneously, so that one will in fact not literally apply a band strategy in the classical sense. Figure 8.6b gives the corresponding

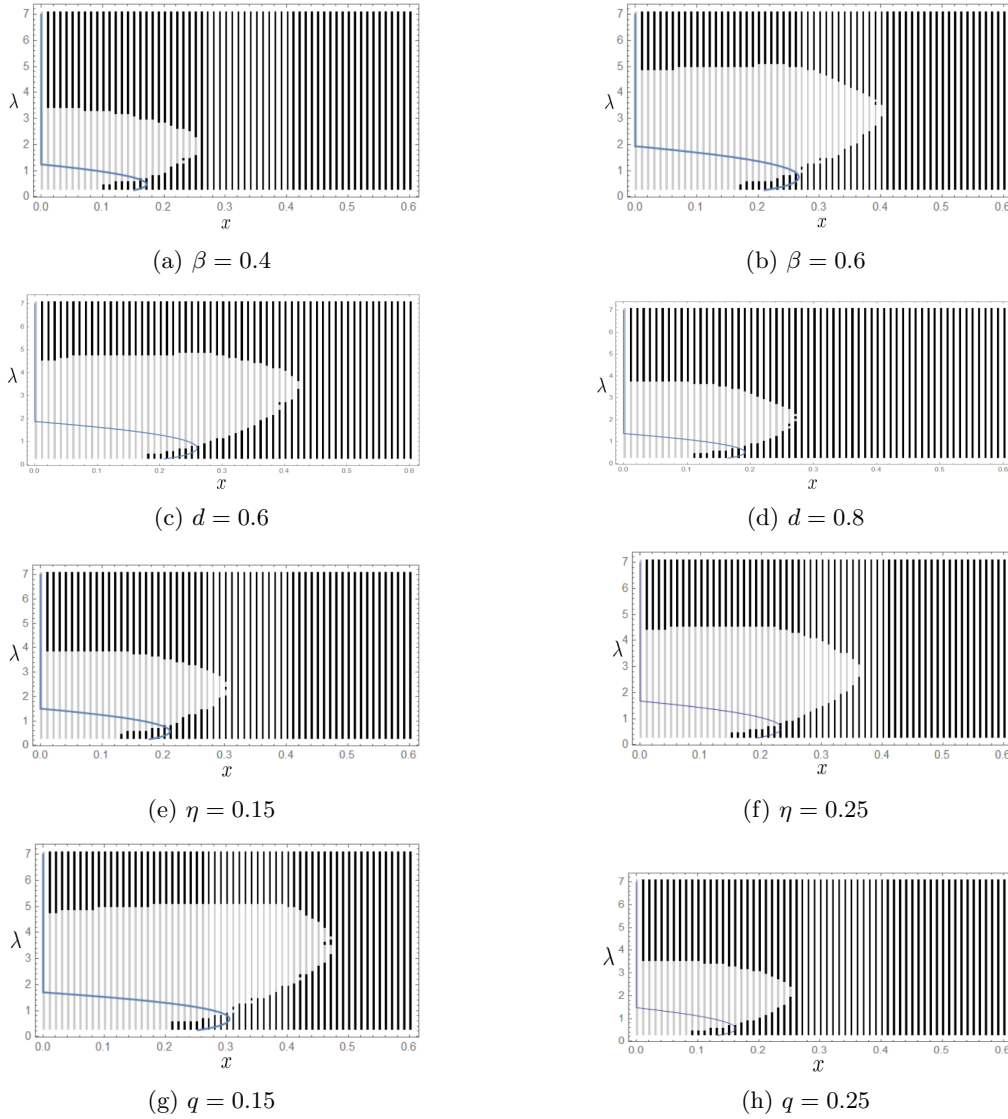
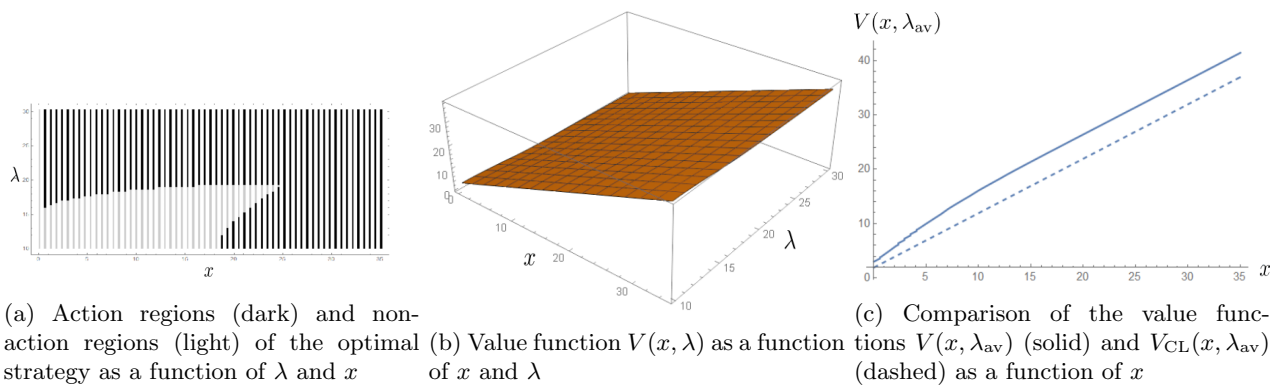


Figure 8.3: Counterpart of Figure 8.1a for modified  $\beta$  (first row), modified  $d$  (second row), modified  $\eta$  (third row) and modified  $q$  (last row)



(a) Action regions (dark) and non-action regions (light) of the optimal strategy as a function of  $\lambda$  and  $x$

(b) Value function  $V(x, \lambda)$  as a function of  $x$  and  $\lambda$

(c) Comparison of the value functions  $V(x, \lambda_{av})$  (solid) and  $V_{CL}(x, \lambda_{av})$  (dashed) as a function of  $x$

Figure 8.4: Optimal dividends for a compound shot-noise Cox process with Erlang(2) claim sizes

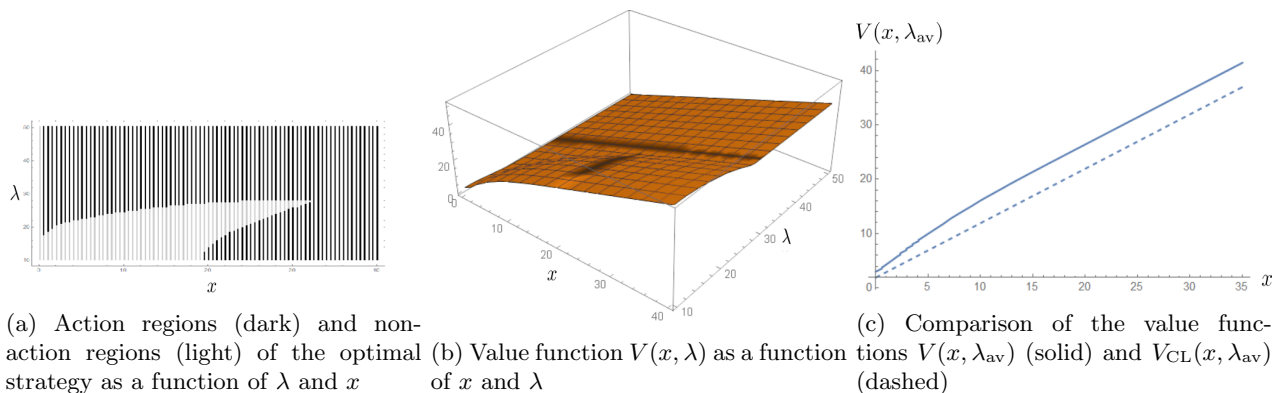


Figure 8.5: Optimal dividends for a compound shot-noise Cox process with deterministic claim sizes for the parameters of Example 2

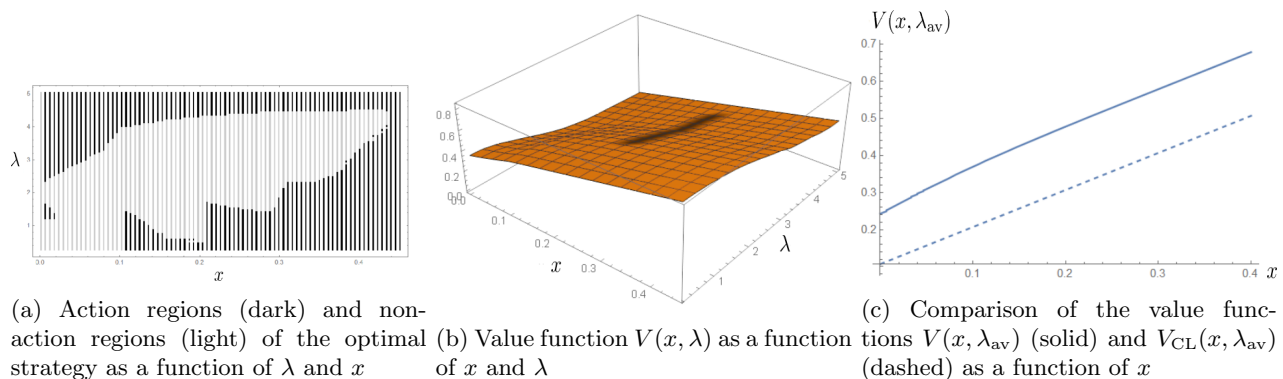


Figure 8.6: Optimal dividends for a compound shot-noise Cox process with deterministic claim sizes for the parameters of Example 1

value function as a function of initial  $x$  and  $\lambda$ , and Figure 8.6c compares it for  $\lambda = \lambda_{av}$  with the one of the classical risk model with homogeneous intensity. Notice that in all the examples of this section the additional variability due to the shot-noise component is in fact advantageous for the shareholders.

## 9 Conclusion and Outlook

In this paper we solved the two-dimensional stochastic control problem of optimizing expected discounted dividends until ruin of an insurance portfolio, when the claim number process is a Cox process with shot-noise intensity function. We identified the optimal value function as the smallest viscosity supersolution of a Hamilton-Jacobi-Bellman equation and provided a numerical scheme to uniformly approximate it via a discretization of the surplus space and the intensity space. In the numerical implementations, we then investigated to what extent the optimal dividend strategies deviate from the classical ones where the Poisson intensity is constant. As it turns out, the additional variability of the claim occurrence pattern can be used to the advantage of the insurer as far as a valuation of the company in terms of expected discounted future dividends is concerned. Concerning the concrete dividend strategies, the nature of barrier and band strategies remains in principle valid, although with barrier and band levels that now depend on the current level of the claim occurrence intensity  $\lambda$ , so that one needs to react adaptively to its change over time.

The fact that additional insurance business linked to catastrophe claims can be beneficial is somewhat promising in times when even reinsurers become more reluctant to include natural catastrophe claims in their portfolio and risk structure, and reinsurance premiums increase globally, see e.g. [31]. While the model assumptions employed in this paper are clearly somewhat too simple for practical implementation in catastrophe insurance practice, the results still demonstrate that a good understanding of the underlying actuarial risks in the catastrophe lines of business also has upward potential. Clearly, the present paper is only one first step into this direction, and many generalizations are possible and desirable. As already mentioned in the intro-



duction, a prominent one is to go beyond the stationary model world and allow for risk factors that change over time, a feature that will allow to more explicitly model the effects of climate change on the profitability of an insurance portfolio. Also, the profitability criterion of maximizing dividends over the lifetime of the company should be accompanied by some constraints on solvency, for instance along approaches that have been pursued in the classical optimal dividend literature already. Also, with respect to regulation, a more refined model as to when received and unused premiums can be paid out as company profits may be pursued. Moreover, other interpretations of the reason and nature of the shot noise Cox feature of the claim number model are naturally possible, and each one will lead to different or additional features of the model to be studied. Finally, it may be possible to apply and extend some features of the techniques developed in this paper to the problem of optimizing dividends for a cyber risk insurer, where in addition to the shot-noise feature one typically also has a self-exciting component in the intensity process, see e.g. Zeller & Scherer [33].

## A Appendices

### A.1 Proofs of Section 1

*Proof of Proposition 2.1.* Let us define  $\lambda_t^0$  and  $\Lambda_t^0$  as the processes defined in (2) and (5) in the particular case of  $\underline{\lambda} = 0$ , then we have

$$\lambda_t = \underline{\lambda}(1 - e^{-dt}) + \lambda_t^0, \Lambda_t = \underline{\lambda} \int_0^t (1 - e^{-ds}) ds + \mathbb{E}(\Lambda_t^0).$$

From [15, Cor.2.4], we get

$$\mathbb{E}(e^{-v \Lambda_t^0}) = e^{-\frac{v \underline{\lambda}}{d}(1 - e^{-dt})} e^{-\beta \int_0^t (1 - g(\frac{v}{d}(1 - e^{-d(t-s)}))) ds},$$

where

$$g(u) = \mathbb{E}(e^{-u Y_1}).$$

So we get

$$\begin{aligned} \mathbb{E}(e^{-v \Lambda_t}) &= \mathbb{E}(e^{-v \Lambda_t^0}) \cdot e^{-v \underline{\lambda} \int_0^t (1 - e^{-ds}) ds} \\ &= e^{-\frac{v \underline{\lambda}}{d}(1 - e^{-dt})} \cdot e^{-\beta \int_0^t (1 - g(\frac{v}{d}(1 - e^{-d(t-s)}))) ds} \cdot e^{-v(\underline{\lambda} t - \underline{\lambda} \left(\frac{1 - e^{-dt}}{d}\right))}. \end{aligned}$$

We also have

$$\begin{aligned} \mathbb{E}(\Lambda_t^0) &= -\partial_v \mathbb{E}(e^{-v \Lambda_t^0}) \Big|_{v=0} \\ &= \frac{1 - e^{-dt}}{d} \lambda + \frac{e^{-\frac{v}{d} t} - 1 + dt}{d^2} \beta \mathbb{E}(Y_1), \end{aligned}$$

so

$$\begin{aligned} \mathbb{E}(e^{-v \Lambda_t}) &= \mathbb{E}(e^{-v \Lambda_t^0}) \cdot e^{-v \underline{\lambda} \int_0^t (1 - e^{-ds}) ds} \\ &= e^{-\frac{v \underline{\lambda}}{d}(1 - e^{-dt})} \cdot e^{-\beta \int_0^t (1 - g(\frac{v}{d}(1 - e^{-d(t-s)}))) ds} \cdot e^{-v(\underline{\lambda} t - \underline{\lambda} \left(\frac{1 - e^{-dt}}{d}\right))}. \end{aligned}$$

In addition,

$$\begin{aligned} \mathbb{E}(\lambda_t^0) &= \partial_t \mathbb{E}(\Lambda_t) \\ &= \lambda e^{-dt} + \frac{(1 - e^{-dt})}{d} \beta \mathbb{E}(Y_1), \end{aligned}$$

and we conclude that

$$\begin{aligned} \mathbb{E}(e^{-v \Lambda_t}) &= \mathbb{E}(e^{-v \Lambda_t^0}) \cdot e^{-v \underline{\lambda} \int_0^t (1 - e^{-ds}) ds} \\ &= e^{-\frac{v \underline{\lambda}}{d}(1 - e^{-dt})} \cdot e^{-\beta \int_0^t (1 - g(\frac{v}{d}(1 - e^{-d(t-s)}))) ds} \cdot e^{-v(\underline{\lambda} t - \underline{\lambda} \left(\frac{1 - e^{-dt}}{d}\right))} \end{aligned}$$

and

$$\begin{aligned} \mathbb{E}(\lambda_t) &= \underline{\lambda}(1 - e^{-dt}) + \mathbb{E}(\lambda_t^0) \\ &= \underline{\lambda}(1 - e^{-dt}) + \lambda e^{-dt} + \frac{(1 - e^{-dt})}{d} \beta \mathbb{E}(Y_1), \end{aligned}$$

as well as

$$\begin{aligned} \mathbb{E}(\Lambda_t) &= \underline{\lambda} \int_0^t (1 - e^{-ds}) ds + \mathbb{E}(\Lambda_t^0) \\ &= \underline{\lambda} \int_0^t (1 - e^{-ds}) ds + \frac{(1 - e^{-dt})}{d} \lambda + \frac{(e^{-dt} - 1 + dt)}{d^2} \beta \mathbb{E}(Y_1) \\ &= \underline{\lambda} t - \underline{\lambda} \left(\frac{1 - e^{-dt}}{d}\right) + \frac{(1 - e^{-dt})}{d} \lambda + \frac{(e^{-dt} - 1 + dt)}{d^2} \beta \mathbb{E}(Y_1). \end{aligned}$$

■

## A.2 Proofs of Section 3

In order to prove Propositions 3.1, 3.2 and 3.3, we need to make a definition and give two lemmas.

**Definition A.1** *Given two sequences  $(\tau_i^1, U_i^1)_{i \geq 1} \in \bar{\Omega}$  and  $(\tau_j^2, U_j^2)_{j \geq 1} \in \bar{\Omega}$  introduced in (4), we define the ordered union of the two sequences as*

$$(\tau_i^1, U_i^1)_{i \geq 1} \amalg (\tau_j^2, U_j^2)_{j \geq 1} = (\tau_n, U_n)_{n \geq 1} \in \bar{\Omega},$$

where  $\{(\tau_n, U_n) : n \geq 1\} = \{(\tau_i^1, U_i^1) : i \geq 1\} \cup \{(\tau_j^2, U_j^2) : j \geq 1\}$  and  $\tau_n \leq \tau_{n+1}$ .

**Definition A.2** *Given any intensity-filtered probability space  $(\Omega^0, \mathcal{F}^0, (\mathcal{F}_t^0)_{t \geq 0}, \mathbb{P}^0)$  and two intensity random process  $\lambda^1 = (\lambda_t^1)_{t \geq 0}$  and  $\lambda^2 = (\lambda_t^2)_{t \geq 0}$  adapted to  $(\mathcal{F}_t^0)_{t \geq 0}$ , we define the superposition*

$$(\Omega^{\lambda^1 \oplus \lambda^2}, \mathcal{F}^{\lambda^1 \oplus \lambda^2}, (\mathcal{F}_t^{\lambda^1 \oplus \lambda^2})_{t \geq 0}, \mathbb{P}^{\lambda^1 \oplus \lambda^2})$$

of the two compound Cox filtered spaces  $(\Omega^{\lambda^1}, \mathcal{F}^{\lambda^1}, (\mathcal{F}_t^{\lambda^1})_{t \geq 0}, \mathbb{P}^{\lambda^1})$  and  $(\Omega^{\lambda^2}, \mathcal{F}^{\lambda^2}, (\mathcal{F}_t^{\lambda^2})_{t \geq 0}, \mathbb{P}^{\lambda^2})$  in the following way:

- $\Omega^{\lambda^1 \oplus \lambda^2} = \left\{ (\omega, \tau_i^1, U_i^1)_{i \geq 1} \amalg (\tau_j^2, U_j^2)_{j \geq 1} \text{ s.t. } (\omega, (\tau_n^i, U_n^i)_{n \geq 1}) \in \Omega^{\lambda^i} \text{ for } i = 1, 2 \right\}$ ;
- $\mathcal{F}^{\lambda^1 \oplus \lambda^2}$  is the complete  $\sigma$ -field generated by the  $\sigma$ -field  $\mathcal{F}^0$  and the random variables  $\tau_n^1, U_n^1, \tau_m^2, U_m^2$ ;
- $\mathcal{F}_t^{\lambda^1 \oplus \lambda^2}$  is the complete  $\sigma$ -field generated by the  $\sigma$ -field  $\mathcal{F}_t^0$  and the random variables  $\tau_n^1, U_n^1$  for  $\tau_n^1 \leq t$  and  $\tau_m^2$  and  $U_m^2$  for  $\tau_m^2 \leq t$ ;
- The probability measure  $\mathbb{P}^{\lambda^1 \oplus \lambda^2}$  is the probability which satisfies that
  1. the processes  $N_t^{\lambda^1} = \#\{k : \tau_k^1 \leq t\}$ ,  $N_t^{\lambda^2} = \#\{k : \tau_k^2 \leq t\}$  and the random variables  $(U_n^1)_{n \geq 1}$  and  $(U_n^2)_{n \geq 1}$  are independent,
  2. the random variables  $U_n^1$  and  $U_m^2$  with  $n, m \in \mathbb{N}$  are i.i.d.,  $\mathbb{P}^{\lambda^1 \oplus \lambda^2}(U_n^i \leq x) = F_U(x)$  for  $i = 1, 2$  and

$$\mathbb{P}^{\lambda^1 \oplus \lambda^2}(N_t^{\lambda^1} + N_t^{\lambda^2} = n) = \sum_{s=0}^n \mathbb{P}^{\lambda^1}(N_t^{\lambda^1} = s) \cdot \mathbb{P}^{\lambda^2}(N_t^{\lambda^2} = n - s).$$

We have the following lemma (cf. for instance [25, Sec.3]).

**Lemma A.1** *The set  $\Omega_0^{\lambda^1 \oplus \lambda^2} = \{\tau_n^1 \neq \tau_m^2 \text{ for any } n, m \in \mathbb{N}\} \in \mathcal{F}^{\lambda^1 \oplus \lambda^2}$  has full measure. Also, the restriction of the superposition probability space*

$$(\Omega^{\lambda^1 \oplus \lambda^2}, \mathcal{F}^{\lambda^1 \oplus \lambda^2}, (\mathcal{F}_t^{\lambda^1 \oplus \lambda^2})_{t \geq 0}, \mathbb{P}^{\lambda^1 \oplus \lambda^2})$$

to  $\Omega_0^{\lambda^1 \oplus \lambda^2}$  coincides with the compound Cox filtered space  $(\Omega^\lambda, \mathcal{F}^\lambda, (\mathcal{F}_t^\lambda)_{t \geq 0}, \mathbb{P}^\lambda)$  conditional on the intensity process  $\lambda = \lambda_1 + \lambda_2 = (\lambda_t^1 + \lambda_t^2)_{t \geq 0}$ .

Moreover, if  $P_1 : \Omega_0^{\lambda^1 \oplus \lambda^2} \rightarrow \Omega^{\lambda^1}$  and  $P_2 : \Omega_0^{\lambda^1 \oplus \lambda^2} \rightarrow \Omega^{\lambda^2}$  are the projections defined as

$$P_i((\omega, (\tau_n^1, U_n^1)_{n \geq 1} \amalg (\tau_n^2, U_n^2)_{n \geq 1})) = (\omega, (\tau_n^i, U_n^i)_{n \geq 1}) \in \Omega^{\lambda^i},$$

then

1. For any  $A \in \mathcal{F}^{\lambda^i}$ , we have that the preimage  $P_i^{-1}(A) \in \mathcal{F}^{\lambda^1 \oplus \lambda^2}$  and  $\mathbb{P}^{\lambda^1 \oplus \lambda^2}(P_i^{-1}(A)) = \mathbb{P}^{\lambda^i}(A)$ .
2. For any  $A \in \mathcal{F}_t^{\lambda^i}$ , we have that the preimage  $P_i^{-1}(A) \in \mathcal{F}_t^{\lambda^1 \oplus \lambda^2}$ .
3.  $N_t^{\lambda^1} \circ P_1 + N_t^{\lambda^2} \circ P_2 = N_t^{\lambda^1 + \lambda^2} : \Omega_0^{\lambda^1 \oplus \lambda^2} \rightarrow \mathbb{N}_0$ .

**Lemma A.2** *As a consequence of the previous lemma, consider any intensity-filtered probability space  $(\Omega^0, \mathcal{F}^0, (\mathcal{F}_t^0)_{t \geq 0}, \mathbb{P}^0)$  and two intensity random processes  $\lambda^1 = (\lambda_t^1)_{t \geq 0}$  and  $\lambda^2 = (\lambda_t^2)_{t \geq 0}$  adapted to  $(\mathcal{F}_t^0)_{t \geq 0}$  with  $\lambda^1 \leq \lambda^2$ ; writing  $\lambda^2 = \lambda^1 + (\lambda^2 - \lambda^1)$  we can consider the projections  $\bar{P}_1 : \Omega^{\lambda^2} \rightarrow \Omega^{\lambda^1}$  and  $\bar{P}_2 : \Omega^{\lambda^2} \rightarrow \Omega^{\lambda^2 - \lambda^1}$ . Also, the processes  $N_t^{\lambda^1} \circ \bar{P}_1$  and  $N_t^{\lambda^2 - \lambda^1} \circ \bar{P}_2$  are independent and satisfy  $N_t^{\lambda^1} \circ \bar{P}_1 + N_t^{\lambda^2 - \lambda^1} \circ \bar{P}_2 = N_t^{\lambda^2}$ . We also can write*

$$(\tau_n^2, U_n^2)_{n \geq 1} = (\tau_n^1, U_n^1)_{n \geq 1} \amalg (\hat{\tau}_m, \hat{U}_m)_{m \geq 1}.$$

*Proof of Proposition 3.1.* Proof of (1): Take  $\lambda_1 < \lambda_2$  and consider the shot-noise processes  $\lambda^i = (\lambda_t^i)_{t \geq 0}$  with initial intensity  $\lambda_i$  and the associated compound Cox process  $X_t^i$  with drift  $p$  generated by the shot-noise process  $\lambda_t^i$  with initial intensity  $\lambda_i$  and initial surplus  $x$  for  $i = 1, 2$  as defined in (1) and (2). Since

$$\lambda_t^2 = \lambda_t^1 + e^{-dt}(\lambda_2 - \lambda_1),$$

we have that  $\lambda^1 \leq \lambda^2$ . We use the results of Lemma A.2, calling  $N_t^1 = N_t^{\lambda^1} \circ \bar{P}_1$ ,  $N_t^2 = N_t^{\lambda^2}$  and  $\hat{N}_t = N_t^{\lambda^2 - \lambda^1} \circ \bar{P}_2$ , we have

$$X_t^2 = x + pt - \sum_{n=1}^{N_t^2} U_n^2 = x + pt - \left( \sum_{n=1}^{N_t^1} U_n^1 + \sum_{m=1}^{\hat{N}_t} \hat{U}_m \right) = X_t^1 - \sum_{m=1}^{\hat{N}_t} \hat{U}_m.$$

Now take  $(L_t)_{t \geq 0} \in \Pi_{x, \lambda_2}$  with

$$V(x, \lambda_2) \leq J(L; x, \lambda_2) + \varepsilon.$$

Let  $r_i = \inf \{t : X_t^i - L_t < 0\}$ ; since  $X_t^1 \geq X_t^2$ , we get  $r_1 \geq r_2$ ;

$$V(x, \lambda_1) \geq J(L; x, \lambda_1) = E\left(\int_{0^-}^{r_1} e^{-qs} dL_s\right) \geq E\left(\int_{0^-}^{r_2} e^{-qs} dL_s\right) = J(L; x, \lambda_2) \geq V(x, \lambda_2) - \varepsilon,$$

so we conclude (1).

Proof of (2): Take  $\lambda_1 < \lambda_2$  and  $L = (L_t)_{t \geq 0} \in \Pi_{x, \lambda_1}$  with

$$V(x, \lambda_1) \leq J(L; x, \lambda_1) + \varepsilon.$$

Note that

$$\mathbb{P}(\hat{N}_t = k) = \mathbb{P}(N_t^2 - N_t^1 = k) = e^{-\int_0^t e^{-ds}(\lambda_2 - \lambda_1) ds} \frac{1}{k!} \left(\int_0^t e^{-ds}(\lambda_2 - \lambda_1) ds\right)^k.$$

Let  $r_i = \inf \{t : X_t^i - L_t < 0\}$ , we have that  $r_1 \geq r_2$  and since  $X_t^1 = x + pt - \sum_{n=1}^{N_t^1} U_n^1$  and

$$\begin{aligned} X_t^2 &= x + pt - \sum_{n=1}^{N_t^2} U_n^2 \\ &= x + pt - \left( \sum_{n=1}^{N_t^1} U_n^1 + \sum_{m=1}^{\hat{N}_t} \hat{U}_m \right) \\ &= X_t^1 - \sum_{m=1}^{\hat{N}_t} \hat{U}_m \end{aligned}$$

we have that

$$X_{r_2}^1 = X_{r_2}^2 + \sum_{m=1}^{\hat{N}_{r_2}} \hat{U}_m - L_{r_2} < \sum_{m=1}^{\hat{N}_{r_2}} \hat{U}_m.$$

Take  $T$  such that  $e^{-qT} \left( \frac{1}{d} \mathbb{E}(U_1) + \frac{p}{q} \right) \leq \frac{\varepsilon}{2}$  and consider  $\lambda_2 \leq \lambda_1 + 1$  and

$$\lambda_2 - \lambda_1 \leq \min \left\{ \frac{-\log \left( 1 - \frac{\varepsilon}{2 \left( \frac{1}{d} \mathbb{E}(U_1) + p/q \right)} \right)}{T \frac{(1 - e^{-dT})}{dT}}, 1 \right\}.$$

Let us show that

$$\begin{aligned} J(L; x, \lambda_1) - J(L; x, \lambda_2) &= \mathbb{E} \left( \mathbb{E} \left( \int_{r_2}^{r_1} e^{-qs} dL_s \mid \mathcal{F}_{r_2}^{\lambda_2} \right) \right) \\ &\leq \mathbb{E} \left( \mathbb{E} \left( e^{-qr_2} V(X_{r_2}^1 - L_{r_2}, \lambda_{r_2}^1) \mid \mathcal{F}_{r_2}^{\lambda_2} \right) \right) \\ &\leq \mathbb{E} \left[ e^{-qr_2} V(X_{r_2}^1 - L_{r_2}, \lambda_{r_2}^1) \right] \\ &\leq \mathbb{E} \left[ I_{r_2 > T} e^{-qr_2} V(\sum_{m=1}^{\hat{N}_{r_2}} \hat{U}_m, \lambda_{r_2}^1) \right] + \mathbb{E} \left[ I_{r_2 \leq T} e^{-qr_2} V(X_{r_2}^1 - L_{r_2}, \lambda_{r_2}^1) \right] \\ &\leq \varepsilon. \end{aligned}$$

From  $V(y, \lambda_{r_2}^1) \leq V(y, \lambda) \leq y + \frac{p}{q}$  and

$$\mathbb{E}(\hat{N}_t) = (\lambda_2 - \lambda_1) \frac{1 - e^{-dt}}{d}$$

we get

$$\begin{aligned}
\mathbb{E} \left[ I_{r_2 > T} e^{-qr_2} V(\sum_{m=1}^{\hat{N}_{r_2}} \hat{U}_m, \lambda_{r_2}^1) \right] &\leq \mathbb{E} \left[ I_{r_2 > T} e^{-qr_2} \left( \sum_{m=1}^{\hat{N}_{r_2}} \hat{U}_m + \frac{p}{q} \right) \right] \\
&\leq e^{-qT} \mathbb{E} \left[ \left( \sum_{m=1}^{\hat{N}_{r_2}} \hat{U}_m + \frac{p}{q} \right) \right] \\
&\leq e^{-qT} \left( \frac{1}{d} (\lambda_2 - \lambda_1) \mathbb{E}(U_1) + \frac{p}{q} \right) \\
&\leq e^{-qT} \left( \frac{1}{d} \mathbb{E}(U_1) + \frac{p}{q} \right) + \frac{\varepsilon}{2} \\
&\leq \frac{\varepsilon}{2}.
\end{aligned}$$

Also, if  $\hat{N}_T = 0$  then  $X_{r_2}^1 - L_{r_2} = X_{r_2}^2 - L_{r_2} < 0$  and so  $V(X_{r_2}^1 - L_{r_2}, \lambda_{r_2}^1) = 0$ . Hence, taking

$$\lambda_2 - \lambda_1 \leq \frac{-\log(1 - \frac{\varepsilon}{2(\frac{1}{d}\mathbb{E}(U_1) + p/q)})}{T \frac{(1-e^{-dT})}{dT}},$$

we have

$$1 - \mathbb{P}(\hat{N}_T = 0) = 1 - e^{-(\lambda_2 - \lambda_1)t \frac{(1-e^{-dt})}{dt}} < \frac{\varepsilon}{2(\frac{1}{d}\mathbb{E}(U_1) + p/q)}$$

and so

$$\begin{aligned}
\mathbb{E} \left[ I_{r_2 < T} e^{-qr_2} V(X_{r_2}^1 - L_{r_2}, \lambda_{r_2}^1) \right] &\leq \mathbb{E} \left[ I_{r_2 < T} e^{-qr_2} V(\sum_{m=1}^{\hat{N}_{r_2}} \hat{U}_m, \lambda_{r_2}^1) I_{\{\hat{N}_T > 0\}} \right] \\
&\leq \mathbb{E} \left[ I_{r_2 < T} V(\sum_{m=1}^{\hat{N}_T} \hat{U}_m, \lambda_{r_2}^1) I_{\{\hat{N}_T > 0\}} \right] \mathbb{P}[\hat{N}_T > 0] \\
&\leq \mathbb{E} \left[ \left( \sum_{m=1}^{\hat{N}_T} \hat{U}_m + p/q \right) \right] \mathbb{P}[\hat{N}_T > 0] \\
&\leq \left( \frac{1}{d} \mathbb{E}(U_1) + p/q \right) \mathbb{P}[\hat{N}_T > 0] \\
&\leq \frac{\varepsilon}{2}.
\end{aligned}$$

Hence, the result follows. ■

*Proof of Proposition 3.2.* It is straightforward to show that  $V(\cdot, \lambda)$  is non-decreasing because

$$V(x_2, \lambda) \geq x_2 - x_1 + V(x_1, \lambda).$$

Let us prove the other inequality. Take  $t_0 = (x_2 - x_1)/p < 1$  and consider  $\lambda_{t_0} = \underline{\lambda} + e^{-dt_0}(\lambda - \underline{\lambda})$ . Given an initial surplus  $x \geq 0$  and  $\varepsilon > 0$ , consider an admissible strategy  $L = (L_t)_{t \geq 0} \in \Pi_{x_2, \lambda_{t_0}}$  such that  $J(L; x_2, \lambda_{t_0}) \geq V(x_2, \lambda_{t_0}) - \varepsilon$  for any  $x_2 > x_1$ . Take now the strategy  $\tilde{L} \in \Pi_{x_1, \lambda}$  that starts with surplus  $x_1$ , pays no dividends if  $X_t^{\tilde{L}} < x_2$  and follows strategy  $L$  after the current reserve reaches  $(x_2, \lambda_{t_0})$ , that is

$$\tilde{L}_t = \begin{cases} 0 & \text{if } t \leq t_0 \wedge \tau_1 \wedge T_1 \\ L_{t-t_0} & \text{if } t \geq t_0 \text{ and } \tau_1 \wedge T_1 > t_0 \\ 0 & \text{if } t > \tau_1 \wedge T_1 \text{ and } t_0 \geq \tau_1 \wedge T_1. \end{cases}$$

The strategy  $\tilde{L}$  is admissible, and we get

$$\begin{aligned}
V(x_1, \lambda) &\geq J(\tilde{L}; x_1, \lambda) \\
&\geq \mathbb{P}(\tau_1 \wedge T_1 > t_0) J(L; x_2, \lambda_{t_0}) e^{-qt_0} \\
&\geq (V(x_2, \lambda_{t_0}) - \varepsilon) e^{-qt_0} \mathbb{P}(\tau_1 \wedge T_1 > t_0),
\end{aligned}$$

and

$$V(x_2, \lambda) - V(x_1, \lambda) \leq V(x_2, \lambda) - (V(x_2, \lambda_{t_0}) - \varepsilon) \mathbb{P}(\tau_1 \wedge T_1 > t_0) e^{-qt_0}.$$

So, using that  $V$  is non-increasing on  $\lambda$  and  $t_0 < 1$ ,

$$\begin{aligned}
V(x_2, \lambda) - V(x_1, \lambda) &\leq V(x_2, \lambda) - V(x_2, \lambda_{t_0}) \mathbb{P}(\tau_1 \wedge T_1 > t_0) e^{-qt_0} \\
&\leq V(x_2, \lambda) - V(x_2, \lambda) \mathbb{P}(\tau_1 \wedge T_1 > t_0) e^{-qt_0} \\
&\leq V(x_2, \lambda) (1 - \mathbb{P}(\tau_1 \wedge T_1 > t_0) e^{-qt_0}) \\
&\leq V(x_2, \lambda) \frac{\beta \lambda_2 + q}{p} (x_2 - x_1) \\
&\leq V(x_2, \underline{\lambda}) \frac{\beta \lambda_2 + q}{p} (x_2 - x_1).
\end{aligned}$$
■

*Proof of Proposition 3.3.* From Proposition 3.1, we have  $0 \leq V(x, \lambda_1) - V(x, \lambda_2)$ . Let us prove the second inequality. Take  $L \in \Pi_{x, \lambda_1}$  such that  $V(x, \lambda_1) \leq J(L; x, \lambda_1) + \varepsilon/2$ , and the associated controlled process  $(X_t^L, \lambda_t)$  starting at  $(x, \lambda_1)$ . Consider  $\delta_0 = \frac{1}{d} \log(\frac{\lambda_2 - \underline{\lambda}}{\lambda_1 - \underline{\lambda}})$  so  $\underline{\lambda} + e^{-d\delta_0}(\lambda_2 - \underline{\lambda}) = \lambda_1$ , then we have that

$$\delta_0 \leq \frac{1}{d(\lambda_1 - \underline{\lambda})} (\lambda_2 - \lambda_1).$$

Define  $\tilde{L} \in \Pi_{x, \lambda_2}$  as

$$\tilde{L}_t = \begin{cases} pt & \text{if } t \leq \delta_0 \wedge \tau_1 \wedge T_1 \\ L_{t-\delta_0} & \text{if } t \geq \delta_0 \text{ and } \tau_1 \wedge T_1 > \delta_0 \\ p(\tau_1 \wedge T_1) & \text{if } t > \tau_1 \wedge T_1 \text{ and } \delta_0 \geq \tau_1 \wedge T_1. \end{cases}$$

We have that

$$\begin{aligned} J(\tilde{L}; x, \lambda_2) &\geq J(L; x, \lambda_1) e^{-q\delta_0} \mathbb{P}(\tau_1 \wedge T_1 > \delta_0) \\ &= J(L; x, \lambda_1) e^{-q\delta_0} (1 - \mathbb{P}(\tau_1 \wedge T_1 \leq \delta_0)) \\ &\geq J(L; x, \lambda_1) e^{-q\delta_0} (1 - (1 - e^{-\beta\delta_0})) (1 - e^{-\lambda_2\delta_0}) \\ &\geq J(L; x, \lambda_1) (1 - q\delta_0) (1 - \beta\lambda_2\delta_0^2) \end{aligned}$$

because

$$\int_0^{\delta_0} (\underline{\lambda} + e^{-ds} (\lambda_2 - \underline{\lambda})) ds \leq \lambda_2 \delta_0,$$

and so, taking  $\lambda_2$  and  $\lambda_1$  close enough so that  $\delta_0 < 1$ ,

$$\begin{aligned} V(x, \lambda_1) - V(x, \lambda_2) &\leq J(L; x, \lambda_1) - J(\tilde{L}; x, \lambda_2) + \varepsilon/2 \\ &\leq J(L; x, \lambda_1) (1 - (1 - q\delta_0) (1 - \beta\lambda_2\delta_0^2)) + \varepsilon/2 \\ &\leq V(x, \lambda_1) (1 - (1 - q\delta_0) (1 - \beta\lambda_2\delta_0^2)) + \varepsilon/2 \\ &\leq V(x, \lambda_1) \frac{(\beta\lambda_2 + q)}{d(\lambda_1 - \underline{\lambda})} (\lambda_2 - \lambda_1) + \varepsilon/2, \end{aligned}$$

and the result follows. ■

*Proof of Proposition 3.7.* Let us define

$$t^*(\lambda) := \frac{1}{d} \ln\left(\frac{\lambda - \underline{\lambda}}{\ln(\lambda) - \underline{\lambda}}\right)$$

so that

$$\underline{\lambda} + e^{-dt^*(\lambda)} (\lambda - \underline{\lambda}) = \ln(\lambda)$$

and  $t^*(\lambda) \rightarrow \infty$  as  $\lambda \rightarrow \infty$ . Take a near optimal strategy  $L = (L_t)_{t \geq 0} \in \Pi_{x, \lambda}$  such that  $V(x, \lambda) \leq J(L; x, \lambda) + \varepsilon$ , hence  $\lambda_t \geq \ln(\lambda)$  for  $t \leq t^*(\lambda)$ . Then, using Lemma 3.6, Proposition 3.1 and Remark 3.5,

$$\begin{aligned} V(x, \lambda) - \varepsilon &\leq J(L; x, \lambda) \\ &= \mathbb{E}(\int_0^{t^*(\lambda) \wedge \tau^L} e^{-qs} dL_s) + e^{-qt^*(\lambda)} \mathbb{E}(I_{t^*(\lambda) > \tau^L} V(X_{t^*(\lambda)}^L, \lambda_{t^*(\lambda)})) \\ &\leq v^{\ln(\lambda)}(x) + e^{-qt^*(\lambda)} v^{\underline{\lambda}}(x + pt^*(\lambda)). \end{aligned}$$

By Remark 3.4, we have that  $\lim_{\lambda \rightarrow \infty} v^{\ln(\lambda)}(x) = x$  and

$$\lim_{\lambda \rightarrow \infty} e^{-qt^*(\lambda)} v^{\underline{\lambda}}(x + pt^*(\lambda)) \leq \lim_{\lambda \rightarrow \infty} e^{-qt^*(\lambda)} (x + pt^*(\lambda) + K) = 0,$$

so we have the result. ■

### A.3 Proofs of Section 4

*Proof of Proposition 4.1.* Let us prove first that  $V$  is a viscosity supersolution. Given initial values  $(x, \lambda) \in (0, \infty) \times (\underline{\lambda}, \infty)$  and any  $l \geq 0$ , let us consider the admissible strategy  $L \in \Pi_{x, \lambda}$  where the company pays dividends with constant rate  $l$  and consider  $\tau^L$  defined as in (6). Let  $\varphi$  be a test function for supersolution of (11) at  $(x, \lambda)$ . We have  $X_t^L = x + (p - l)t$  and  $\lambda_t = \underline{\lambda} + e^{-dt}(\lambda - \underline{\lambda})$  for  $t < \tau_1 \wedge T_1$ . Applying Lemma 3.6

with stopping time  $\tau_1 \wedge T_1 \wedge h$ , we obtain

$$\begin{aligned}
0 &= V(x, \lambda) - \varphi(x, \lambda) \\
&\geq \mathbb{E} \left( \int_0^{\tau_1 \wedge T_1 \wedge h} e^{-qt} dt \right) \\
&\quad + \mathbb{E} (I_{\tau_1 < T_1 \wedge h} e^{-q\tau_1} V(x + (p-l)\tau_1 - U_1, \underline{\lambda} + e^{-d\tau_1} (\lambda - \underline{\lambda}))) \\
&\quad + \mathbb{E} (I_{T_1 < \tau_1 \wedge h} e^{-qT_1} V(x + (p-l)T_1, \underline{\lambda} + e^{-dT_1} (\lambda - \underline{\lambda}) + Y_1)) \\
&\quad + \mathbb{E} (I_{h < \tau_1 \wedge T_1} e^{-qh} V(x + (p-l)h, \underline{\lambda} + e^{-dh} (\lambda - \underline{\lambda}))) - \varphi(x, \lambda) \\
&\geq \mathbb{E} \left( \int_0^{\tau_1 \wedge T_1 \wedge h} e^{-qt} dt \right) \\
&\quad + \mathbb{E} (I_{\tau_1 < T_1 \wedge h} e^{-q\tau_1} \varphi(x + (p-l)\tau_1 - U_1, \underline{\lambda} + e^{-d\tau_1} (\lambda - \underline{\lambda}))) \\
&\quad + \mathbb{E} (I_{T_1 < \tau_1 \wedge h} e^{-qT_1} \varphi(x + (p-l)T_1, \underline{\lambda} + e^{-dT_1} (\lambda - \underline{\lambda}) + Y_1)) \\
&\quad + \mathbb{E} (I_{h < \tau_1 \wedge T_1} e^{-qh} \varphi(x + (p-l)h, \underline{\lambda} + e^{-dh} (\lambda - \underline{\lambda}))) - \varphi(x, \lambda).
\end{aligned}$$

So, dividing by  $h$  and taking  $h \rightarrow 0^+$ , we obtain

$$\mathcal{L}(\varphi)(x, \lambda) + l(1 - \varphi_x(x, \lambda)) \leq 0.$$

For  $l \rightarrow \infty$ , we obtain

$$\max\{\mathcal{L}(\varphi)(x, \lambda), 1 - \varphi_x(x, \lambda)\} \leq 0.$$

Hence  $V$  is a viscosity supersolution at  $(x, \lambda)$ .

Let us prove now that  $V$  is a viscosity subsolution. Arguing by contradiction, we assume that  $V$  is not a subsolution of (11) at  $(x_0, \lambda_0) \in (0, \infty) \times (\underline{\lambda}, \infty)$ . As in the proof of Proposition 3.8 of [10] but extended to two variables as in Proposition 3.2 in [2], we can find a continuously differentiable function  $\psi : (0, \infty) \times (\underline{\lambda}, \infty) \rightarrow \mathbb{R}$  such that  $\psi$  is a test function for subsolution of equation (11) at  $(x_0, \lambda_0)$ , for  $h < x_0 \wedge (\lambda_0 - \underline{\lambda})$  small enough so

$$\psi_x(x, \lambda) \geq 1 \tag{35}$$

for  $(x, \lambda) \in [0, x_0 + h] \times (\underline{\lambda}, \infty)$ ,

$$\mathcal{L}(\psi)(x, \lambda) \leq -\varepsilon q \tag{36}$$

for  $(x, \lambda) \in [x_0 - h, x_0 + h] \times [\lambda_0 - h, \lambda_0 + h]$ , and

$$V(x, \lambda) \leq \psi(x, \lambda) - \varepsilon \tag{37}$$

for  $(x, \lambda) \in [0, x_0 + h] \times (\underline{\lambda}, \infty) - [x_0 - h, x_0 + h] \times [\lambda_0 - h, \lambda_0 + h]$ .

Let us take any admissible strategy  $L \in \Pi_{x_0, \lambda_0}$ . Consider the corresponding controlled risk process  $(X_t^L, \lambda_t)$  starting at  $(x_0, \lambda_0)$ , and define the stopping time

$$\tau^* = \inf\{t > 0 : (X_t^L, \lambda_t) \in (0, \infty) \times (\underline{\lambda}, \infty) - [x_0 - h, x_0 + h] \times [\lambda_0 - h, \lambda_0 + h]\}.$$

From (37), we obtain that if  $\tau^* < \tau^L$ ,

$$V(X_{\tau^*}^L, \lambda_{\tau^*}) \leq \psi(X_{\tau^*}^L, \lambda_{\tau^*}) - 2\varepsilon. \tag{38}$$

We can write

$$d\lambda_t = -d(\lambda_t - \underline{\lambda}) + P_t^Y, \quad dX_t^L = pdt + P_t^U - dL_t,$$

where  $P_t^Y = \sum_{T_i} I_{T_i=t} Y_i$  and  $P_t^U = \sum_{\tau_i} I_{\tau_i=t} Y_i$ . Here we assume that  $P_t^Y$  and  $P_t^U$  do not jump simultaneously (because they jump at exponential times that are independent). Using Theorem 31 of [28], we get

$$\begin{aligned}
&\psi(X_{\tau^*}^L, \lambda_{\tau^*}) e^{-q\tau^*} - \psi(x_0, \lambda_0) \\
&\leq \int_0^{\tau^*} \mathcal{L}(\psi)(X_{s^-}^L, \lambda_{s^-}) e^{-qs} ds + M_{\tau^*}^1 + M_{\tau^*}^2 - \int_0^{\tau^*} e^{-qs} dL_s,
\end{aligned}$$

where

$$M_t^1 = \sum_{\substack{P_s^U \neq 0 \\ s \leq t}} \left( \psi(X_{s^-}^L - P_s^U, \lambda_s) - \psi(X_{s^-}^L, \lambda_{s^-}) \right) e^{-qs} - \int_0^t \lambda_{s^-} \int_0^{X_{s^-}^L} \psi(X_s^L - \alpha, \lambda_{s^-}) dF_U(\alpha) e^{-qs} ds$$

and

$$M_t^2 = \sum_{\substack{P_s^Y \neq 0 \\ s \leq t}} \left( \psi(X_{s^-}^L, \lambda_{s^-} - P_s^Y) - \psi(X_{s^-}^L, \lambda_{s^-}) \right) e^{-qs} - \int_0^t \beta \int_0^{X_{s^-}^L} \psi(X_{s^-}^L, \lambda_{s^-} + \gamma) dF_Y(\gamma) e^{-qs} ds$$

are martingales with zero expectation. Hence, we obtain

$$\mathbb{E}(\psi(X_{\tau^*}^L, \lambda_{\tau^*})e^{-q\tau^*} - \psi(x_0, \lambda_0)) \leq \mathbb{E}\left(\int_0^{\tau^*} \mathcal{L}(\psi)(X_{s^-}^L, \lambda_{s^-})e^{-qs} - \mathbb{E}\left(\int_0^{\tau^*} e^{-qs} dL_s\right)\right). \quad (39)$$

Using (36), we get

$$\int_0^{\tau^*} \mathcal{L}(\psi)(X_{s^-}^L, \lambda_{s^-})e^{-qs} ds \leq -\varepsilon q \int_0^{\tau^*} e^{-qs} ds. \quad (40)$$

From (41), Lemma 3.6, (38), (39) and (40) it follows that

$$\begin{aligned} V(x_0, \lambda_0) &= \sup_L \mathbb{E} \left( \int_0^{\tau^*} e^{-qs} dL_s + e^{-q\tau^*} V(X_{\tau^*}^L, \lambda_{\tau^*}) \right) \\ &\leq \sup_L \mathbb{E} \left( \int_0^{\tau^*} e^{-qs} dL_s + e^{-q\tau^*} (\psi(X_{\tau^*}^L, \lambda_{\tau^*}) - \varepsilon) I_{\tau^* < \tau L} \right) \\ &\leq \psi(x_0, \lambda_0) + \sup_L \mathbb{E} \left( \int_0^{\tau^*} \mathcal{L}(\psi)(X_{s^-}^L, \lambda_{s^-})e^{-qs} ds - \varepsilon e^{-q\tau^*} I_{\tau^* < \tau L} \right) \\ &\leq \psi(x_0, \lambda_0) + \sup_L \mathbb{E} \left( -\varepsilon(1 - e^{-q\tau^*}) - \varepsilon e^{-q\tau^*} I_{\tau^* < \tau L} \right) \\ &\leq \psi(x_0, \lambda_0) - \varepsilon + \varepsilon \mathbb{E}(e^{-q(\tau_1 \wedge T_1)}) \\ &\leq \psi(x_0, \lambda_0) - \varepsilon + \varepsilon(\lambda_0 + \beta) / (q + \lambda_0 + \beta) \\ &< \psi(x_0, \lambda_0). \end{aligned} \quad (41)$$

But the latter contradicts the assumption that  $V(x_0, \lambda_0) = \psi(x_0, \lambda_0)$ . Hence  $V$  is a viscosity subsolution at  $(x_0, \lambda_0)$  and this complete the proof.  $\blacksquare$

The next lemma will be used to prove Proposition 4.2.

**Lemma A.3** *Let  $\bar{u}$  be a non-negative supersolution of (11) satisfying the growth condition (13) and non-increasing in  $\lambda$ . We can find a sequence of positive functions  $\bar{u}^m : (0, \infty) \times (\underline{\lambda}, \infty) \rightarrow \mathbb{R}$  such that:*

- (a)  $\bar{u}^m$  is continuously differentiable.
- (b)  $\bar{u}^m(x, \lambda) \leq K + x$  and  $\bar{u}^m$  is non-increasing in  $\lambda$ .
- (c)  $1 \leq \bar{u}_x^m(x, \lambda) \leq ((q + \lambda + \beta) / p) \bar{u}(x + 1/m, \lambda)$  for  $(x, \lambda) \in (0, \infty) \times (\underline{\lambda}, \infty)$ .
- (d)  $\bar{u}^m \rightarrow \bar{u}$  uniformly on compact sets in  $(0, \infty) \times (\underline{\lambda}, \infty)$  and  $\nabla \bar{u}^m$  converges to  $\nabla \bar{u}$  a.e. in  $(0, \infty) \times (\underline{\lambda}, \infty)$ .
- (e) There exists a sequence  $c_m$  with  $\lim_{m \rightarrow \infty} c_m = 0$  such that

$$\sup_{(x, \lambda) \in A_0} \mathcal{L}(\bar{u}^m)(x, \lambda) \leq c_m, \text{ where } A_0 = [0, x_0] \times [\lambda_0, \lambda_1],$$

where  $x_0 > 0$  and  $\lambda_0, \lambda_1 \in (\underline{\lambda}, \infty)$ .

*Proof of Lemma A.3.* Let us define the set

$$D = \{(x, \lambda) \in (0, \infty) \times (\underline{\lambda}, \infty) \text{ s.t. } \bar{u} \text{ differentiable in } (x, \lambda)\}.$$

Since  $\bar{u}$  is a supersolution of (11), we have that

$$p\bar{u}_x(x, \lambda) - d(\lambda - \underline{\lambda})\bar{u}_\lambda(x, \lambda) \leq (q + \lambda + \beta)\bar{u}(x, \lambda) \quad (42)$$

and  $\bar{u}_x \geq 1$  for all  $(x, \lambda) \in D$  and so a.e. in  $(0, \infty) \times (\underline{\lambda}, \infty)$ .

Let  $\phi$  be a nonnegative continuously differentiable function with support included in  $(0, 1)$  such that  $\int_0^1 \phi(x) dx = 1$ , we define  $\bar{u}^m : (0, \infty) \times (\underline{\lambda}, \infty) \rightarrow \mathbb{R}$  as the convolution

$$\bar{u}^m(x, \lambda) = m^2 \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \bar{u}(x + s, \lambda + t) \phi(ms) \phi(mt) ds dt. \quad (43)$$

By definition,  $\bar{u}^m(x, \lambda)$  is a weighted average of values of  $\bar{u}$  in  $A_m = [x, x + 1/m] \times [\lambda, \lambda + 1/m]$ ; (a), (b) and (d) follow by standard techniques, because  $\bar{u}^m \geq \bar{u}$  and  $\bar{u}$  is absolutely continuous in  $(0, \infty) \times (\underline{\lambda}, \infty)$  and satisfies the growth condition (13) (see for instance [32]). From Equation (11) we have that  $\bar{u}_x \geq 1$  a.e., also since for all  $(x, \lambda) \in D$ ,  $\mathcal{L}(\bar{u})(x, \lambda) \leq 0$  we have

$$-(q + \lambda + \beta)\bar{u}(x, \lambda) \leq p\bar{u}_x(x, \lambda) - d(\lambda - \underline{\lambda})\bar{u}_\lambda(x, \lambda) \leq (q + \lambda + \beta)\bar{u}(x, \lambda),$$

so that we conclude (c).

Let us define for  $(x, \lambda) \in A_0$  the function

$$\xi_m(x, \lambda) = \sup_{(x, \lambda) \in A_m \cap D} (p\bar{u}_x(x, \lambda) - d(\lambda - \underline{\lambda})\bar{u}_\lambda(x, \lambda)). \quad (44)$$

We have that for all  $(x, \lambda) \in D$

$$p\bar{u}_x(x, \lambda) - d(\lambda - \underline{\lambda})\bar{u}_\lambda(x, \lambda) \leq \xi_m(x, \lambda) \leq (q + \lambda + \beta)\bar{u}(x + 1/m, \lambda). \quad (45)$$

From (44), for any  $(x, \lambda) \in A_0$  there exists  $(x_m, \lambda_m) \in A_m \cap D$  such that

$$p\bar{u}_x(x_m, \lambda_m) - d(\lambda - \underline{\lambda})\bar{u}_\lambda(x_m, \lambda_m) \geq \xi_m(x, \lambda) - \frac{1}{m}. \quad (46)$$

So,

$$\begin{aligned} & p\bar{u}_x^m(x, \lambda) - d(\lambda - \underline{\lambda})\bar{u}_\lambda^m(x, \lambda) - (p\bar{u}_x(x_m, \lambda_m) - d(\lambda - \underline{\lambda})\bar{u}_\lambda(x_m, \lambda_m)) \\ &= m^2 \iint_D (p\bar{u}_x(x + s, \lambda + t) - d(\lambda + t - \underline{\lambda})\bar{u}_\lambda(x + s, \lambda + t)) \phi(ms)\phi(mt) ds dt \\ & \quad - (p\bar{u}_x(x_m, \lambda_m) - d(\lambda_m - \underline{\lambda})\bar{u}_\lambda(x_m, \lambda_m)) \\ & \leq \xi_m(x, \lambda) - (\xi_m(x, \lambda) - \frac{1}{m}) \\ & \leq \frac{1}{m}. \end{aligned}$$

From (42), (45), (46) and using that  $\bar{u}$  and  $\bar{u}^m$  are uniformly continuous in compact sets and that it is non-decreasing in  $\lambda$ , we have the result.  $\blacksquare$

*Proof of Proposition 4.2.* Let  $\bar{u}$  be a non-negative supersolution of (11) satisfying the growth condition (13) and take an admissible strategy  $L \in \Pi_{x, \lambda}$ , define  $(X_t^L, \lambda_t)$  as the corresponding controlled risk process starting at  $(x, \lambda)$ . Let us consider the function  $\bar{u}^m$  as defined in Lemma A.3 in  $(0, \infty) \times (\underline{\lambda}, \infty)$  and let us extend this function as  $\bar{u}^m(x, \lambda) = 0$  otherwise, as in (39) in the proof of Proposition 4.1, we obtain (using  $\bar{u}_x^m \geq 1$ ),

$$\begin{aligned} & \mathbb{E} \left( \bar{u}^m(X_{t \wedge \tau^L}^L, \lambda_{t \wedge \tau^L}) e^{-q(t \wedge \tau^L)} \right) - \bar{u}^m(x, \lambda) \\ & \leq \mathbb{E} \left( \int_0^{t \wedge \tau^L} \mathcal{L}(\bar{u}^m)(X_{s^-}^L, \lambda_{s^-}) e^{-qs} ds \right) - \mathbb{E} \left( \int_0^{t \wedge \tau^L} e^{-qs} dL_s \right). \end{aligned} \quad (47)$$

Since  $L_t$  is a non-decreasing process we get, by the monotone convergence theorem, that

$$\lim_{t \rightarrow \infty} \mathbb{E} \left( \int_0^{t \wedge \tau^L} e^{-qs} dL_s \right) = J(L; x, \lambda).$$

From Lemma A.3(c), we have

$$-(q + \lambda + \beta)\bar{u}(x + \frac{1}{m}, \lambda) \leq \mathcal{L}(\bar{u}^m)(x, \lambda) \leq (q + \lambda + \beta)\bar{u}(x + \frac{1}{m}, \lambda) + \beta\bar{u}(x, \lambda). \quad (48)$$

But using Lemma A.3(b) and the inequalities  $X_s^L \leq x + ps$ ,  $\lambda_s > \underline{\lambda}$  we obtain

$$\bar{u}(X_s^L, \lambda_s) \leq K + ps. \quad (49)$$

So, by the bounded convergence theorem, we get

$$\lim_{t \rightarrow \infty} \mathbb{E} \left( \int_0^{t \wedge \tau^L} \mathcal{L}(\bar{u}^m)(X_{s^-}^L, \lambda_{s^-}) e^{-qs} ds \right) = \mathbb{E} \left( \int_0^{\tau^L} \mathcal{L}(\bar{u}^m)(X_{s^-}^L, \lambda_{s^-}) e^{-qs} ds \right). \quad (50)$$

From (47) and (50), we have

$$\begin{aligned} & \lim_{t \rightarrow \infty} \mathbb{E} \left( \bar{u}^m(X_{t \wedge \tau^L}^L, \lambda_{t \wedge \tau^L}) e^{-q(t \wedge \tau^L)} I_{\tau^L < t} \right) - \bar{u}^m(x, \lambda) \\ &= -\bar{u}^m(x, \lambda) \\ & \leq \mathbb{E} \left( \int_0^{\tau^L} \mathcal{L}(\bar{u}^m)(X_{s^-}^L, \lambda_{s^-}) e^{-qs} ds \right) - J(L; x, \lambda). \end{aligned} \quad (51)$$

Let us prove now that

$$\limsup_{m \rightarrow \infty} \mathbb{E} \left( \int_0^{\tau^L} \mathcal{L}(\bar{u}^m)(X_{s^-}^L, \lambda_{s^-}) e^{-qs} ds \right) \leq 0. \quad (52)$$

Because  $\bar{u}^m(x, \lambda) \leq K + x$ , Lemma A.3(b) and Lemma A.3(b)(c) give

$$\begin{aligned} |\mathcal{L}(\bar{u}^m)(X_{s^-}^L, \lambda_{s^-})| & \leq (q + \lambda + \beta)\bar{u}(X_{s^-}^L + \frac{1}{m}, \lambda_{s^-}) + \beta\bar{u}(X_{s^-}^L, \lambda_{s^-}) \\ & \leq (q + \lambda + \beta)\bar{u}(X_{s^-}^L + \frac{1}{m}, \underline{\lambda}) + \beta\bar{u}(X_{s^-}^L, \underline{\lambda}) \\ & \leq (q + \lambda + 2\beta)(x + \frac{1}{m} + ps). \end{aligned}$$

Then, given any  $\varepsilon > 0$ , we can find  $T$  such that

$$\int_T^\infty \mathcal{L}(\bar{u}^m)(X_{s^-}^L, \lambda_{s^-}) e^{-qs} ds < \frac{\varepsilon}{4} \quad (53)$$



for any  $m \geq 1$ . Note that for  $s \leq T$ , we have that  $X_{s-}^L \leq x_0 := x + pT$ ,  $\lambda_{s-} \geq \lambda_1 := \underline{\lambda} + (\lambda - \underline{\lambda})e^{-dt}$ . From (2), there exists  $M$  large enough such that

$$\mathbb{P}(\sup_{s \leq T} \lambda_s > M) < \frac{\varepsilon}{4(q + \lambda + 2\beta)(x + \frac{1}{m} + ps)}.$$

Now, take the compact set  $[0, x_0] \times [\lambda_1, M]$  and the set

$$A^{T,M} = \left\{ (\lambda_s)_{s \leq T} : \sup_{s \leq T} \lambda_s \leq M \right\}.$$

From Lemma A.3(e), we can find  $m_0$  large enough such that for any  $m \geq m_0$ ,

$$I_{A^{T,M}} \int_0^T \mathcal{L}(\bar{u}^m)(X_{s-}^L, \lambda_{s-}) e^{-qs} ds \leq c_m \int_0^T e^{-qs} ds \leq \frac{c_m}{q} \leq \frac{\varepsilon}{2},$$

and so taking the expectation we get (52). Then, from (51) and using Lemma A.3(d), we finally obtain

$$\bar{u}(x, \lambda) = \lim_{m \rightarrow \infty} \bar{u}^m(x, \lambda) \geq J(L; x, \lambda). \quad (54)$$

Since  $V$  is a viscosity solution of (11), the result follows.  $\blacksquare$

## A.4 Proofs of Section 5

*Proof of Proposition 5.1.* Let us define for  $M > 0$ ,

$$V_M(x, \lambda) = \begin{cases} V(x, \lambda) & \text{if } x \leq M, \\ V(M, \lambda) + x - M & \text{if } x > M. \end{cases}$$

$V_M(x, \lambda)$  is a limit of value functions of strategies. From Propositions 3.1, 3.2 and 3.3,  $V_M$  is locally Lipschitz, increasing on  $x$  and non-increasing on  $\lambda$ . For  $x < M$ ,  $V_M$  is a viscosity supersolution of (11). For  $x > M$ ,  $\partial_x V_M(x, \lambda) = 1$ . In order to see that  $V_M$  is a viscosity supersolution of (11) for  $x > M$  and  $M$  large enough, we need to show that  $\mathcal{L}(V_M)(x, \lambda) \leq 0$  for  $x > M$ . Indeed,

$$\begin{aligned} & \mathcal{L}(V_M)(x, \lambda) \\ & \leq p - (q + \lambda)V_M(x, \lambda) + \lambda \int_0^x V_M(x - \alpha, \lambda) dF_U(\alpha) + \beta \left( \int_0^\infty V_M(x, \lambda) dF_Y(\gamma) - V_M(x, \lambda) \right) \\ & \leq p - (q + \lambda)V_M(x, \lambda) + \lambda \int_0^x V_M(x - \alpha, \lambda) dF_U(\alpha) \\ & \leq p - (q + \lambda)(V(M, \lambda) + x - M) + \lambda \int_0^x (V(M, \lambda) + x - \alpha - M) dF_U(\alpha) \\ & \leq p - qV(M, \lambda) \\ & \leq 0 \end{aligned}$$

for any  $M \geq p/q$  because  $V(x, \lambda) \geq x$ . Consider the following dense set  $B$  in  $[0, \infty) \times [\underline{\lambda}, \infty)$

$$B := \{(x, \lambda) \in [0, \infty) \times [\underline{\lambda}, \infty) : V \text{ is differentiable}\}.$$

For any point  $(M, \lambda) \in B$  with  $M \geq \frac{p}{q}$ , we have that  $\partial_x V_M(M, \lambda) = 1$  and so, by Corollary 4.3,  $V_M$  coincides with  $V$ .  $\blacksquare$

*Proof of Proposition 5.2.* From Proposition 5.1,  $V(x, \lambda) = x + A(\lambda)$  for  $x \geq p/q$ , where

$$A(\lambda) := V(p/q, \lambda) - p/q,$$

and since from Proposition 3.7,  $\lim_{\lambda \rightarrow \infty} V(p/q, \lambda) - p/q = 0$  and  $V(p/q, \lambda)$  is non-increasing on  $\lambda$ , this implies  $A(\lambda) \searrow 0$  as  $\lambda \rightarrow \infty$  and so

$$\lim_{\lambda \rightarrow \infty} \sup_{x \geq p/q} (V(x, \lambda) - x) = 0.$$

Also, for all  $x \geq 0$  such that  $V_x(x, \lambda)$  exists,  $V_x(x, \lambda) \geq 1$  and so  $h(x, \lambda) := V(x, \lambda) - x$  is non-decreasing on  $x$ . Hence,

$$V(x, \lambda) - x \leq V(p/q, \lambda) - p/q \leq A(\lambda)$$

and that implies the result.  $\blacksquare$

## A.5 Proofs of Section 6

*Proof of Proposition 6.1.* We have that

$$V^\delta(x_{n+1}^\delta, \lambda) - V^\delta(x_n^\delta, \lambda) \leq V^\delta(x_n^\delta) e^{(\beta+q)\delta + \int_0^\delta \lambda_u^c du} - V^\delta(x_n^\delta).$$

Then, denoting  $x_n^\delta = \rho^\delta(x_2)$  and  $x_m^\delta = \rho^\delta(x_1)$ ,

$$\begin{aligned} V^\delta(x_2, \lambda) - V^\delta(x_1, \lambda) &\leq V^\delta(\rho^\delta(x_2), \lambda) - V^\delta(\rho^\delta(x_1), \lambda) + (x_2 - \rho^\delta(x_2)) \\ &\leq V^\delta(x_n^\delta, \lambda) - V^\delta(x_m^\delta, \lambda) + \delta p \\ &\leq \sum_{j=m}^{n-1} \left( e^{(\beta+q)\delta + \int_0^\delta \lambda_u^c du} - 1 \right) V^\delta(x_j^\delta, \lambda) + \delta p \\ &\leq (\rho^\delta(x_2) - \rho^\delta(x_1)) \frac{\left( e^{(\beta+q)\delta + \int_0^\delta \lambda_u^c du} - 1 \right)}{p\delta} V^\delta(x_2, \lambda) + \delta p. \end{aligned}$$

With a similar proof to the one of Proposition 3.1, we have that  $V^\delta(x, \cdot)$  is non-increasing and so  $0 \leq V^\delta(x, \lambda_1) - V^\delta(x, \lambda_2)$ . It suffices to prove

$$V^\delta(x_n^\delta, \lambda_1) - V^\delta(x_n^\delta, \lambda_2) \leq V^\delta(x_n^\delta, \lambda_1) \frac{\beta\lambda_2 + q}{d(\lambda_1 - \underline{\lambda})} (\lambda_2 - \lambda_1) \quad (55)$$

for  $0 < \lambda_2 - \lambda_1$  small enough.

In order to prove (55), we modify the set of admissible strategies  $\tilde{\Pi}_{x_n^\delta, \lambda}^\delta$ , adding new local strategies which are not optimal. Let us define  $\hat{\Pi}_{x_n^\delta, \lambda}^\delta$  as the set of all the strategies with initial surplus  $x_n^\delta \in \mathcal{G}_\delta$  which can be obtained by a sequence of control actions in

$$\hat{\mathcal{E}} = \{\mathbf{E}_F, \mathbf{E}_1, \mathbf{E}_0, \hat{\mathbf{E}}_\eta\},$$

where the controls  $\mathbf{E}_F$ ,  $\mathbf{E}_1$  and  $\mathbf{E}_0$  are defined in (17). For any time  $\eta > 0$ , the new control action  $\hat{\mathbf{E}}_\eta$  consists of throwing away the incoming premium  $p$  up to time  $\eta \wedge \tau_1 \wedge T_1$  (note that if  $\eta < \tau_1 \wedge T_1$  the final surplus and intensity of this local control action is  $(x_n^\delta, \underline{\lambda} + (\lambda - \underline{\lambda})e^{-d\eta}) \in \mathcal{G}_\delta \times [0, \infty)$ ). Since it is never optimal to throw away money,

$$\sup_{\pi \in \hat{\Pi}_{x_n^\delta, \lambda}^\delta} J(\pi; (x_n^\delta, \lambda)) = \sup_{\pi \in \tilde{\Pi}_{x_n^\delta, \lambda}^\delta} J(\pi; (x_n^\delta, \lambda)) = V^\delta(x_n^\delta, \lambda).$$

Given any  $\varepsilon > 0$ , take  $\pi = (L, \tau_F) \in \tilde{\Pi}_{x_n^\delta, \lambda_1}^\delta$  such that  $V^\delta(x_n^\delta, \lambda_1) - J((L, \tau_F); (x_n^\delta, \lambda_1)) < \varepsilon/2$ . By Remark 5.3 we can take  $\tau_F = \infty$ , and the associated controlled process

$$(X_t^L, \lambda_t) = \left( x_n^\delta + pt - \sum_{i=1}^{N_t} U_i - L_t, \underline{\lambda} + e^{-dt} (\lambda_1 - \underline{\lambda}) + \sum_{0 \leq T_k \leq t} Y_n e^{-d(t-T_k)} \right).$$

Consider  $\varkappa$  such that

$$\underline{\lambda} + e^{-d\varkappa} (\lambda_2 - \underline{\lambda}) = \lambda_1, \quad (56)$$

that is

$$\varkappa = \frac{1}{d} \log\left(\frac{\lambda_2 - \underline{\lambda}}{\lambda_1 - \underline{\lambda}}\right) = \frac{1}{d} \log\left(1 + \frac{\lambda_2 - \lambda_1}{\lambda_1 - \underline{\lambda}}\right) \leq \frac{1}{d(\lambda_1 - \underline{\lambda})} (\lambda_2 - \lambda_1).$$

Define now  $\hat{\pi}$  as the strategy to apply first the local control  $\hat{\mathbf{E}}_\varkappa$  and then either  $\pi = (L, \tau_F) \in \tilde{\Pi}_{x_n^\delta, \lambda_1}^\delta$  in case  $\varkappa < \tau_1 \wedge T_1$ , or  $\mathbf{E}_F$  otherwise. So  $\hat{\pi} \in \hat{\Pi}_{x_n^\delta, \lambda_2}^\delta$  and

$$\hat{L}_t = \begin{cases} 0 & \text{if } t \leq \varkappa \wedge \tau_1 \wedge T_1 \\ L_{t-\varkappa} & \text{if } t \geq \varkappa \text{ and } \tau_1 \wedge T_1 > \varkappa \end{cases}.$$

One has

$$\begin{aligned} J(\hat{\pi}; (x_n^\delta, \lambda_2)) &\geq J(\hat{\pi}; x_n^\delta, \lambda_1) e^{-q\varkappa} \mathbb{P}(\tau_1 \wedge T_1 \geq \varkappa) \\ &\geq J(\hat{\pi}; x_n^\delta, \lambda_1) e^{-q\varkappa} (1 - (1 - e^{-\beta\varkappa})(1 - e^{-\int_0^\varkappa (\underline{\lambda} + e^{-ds}(\lambda_2 - \underline{\lambda})) ds})) \\ &\geq J(\hat{\pi}; x_n^\delta, \lambda_1) e^{-q\varkappa} (1 - (1 - e^{-\beta\varkappa})(1 - e^{-\lambda_2\varkappa})) \\ &\geq J(\hat{\pi}; x_n^\delta, \lambda_1) (1 - q\varkappa) (1 - \beta\lambda_2\varkappa) \end{aligned}$$

because

$$\int_0^\varkappa (\underline{\lambda} + e^{-ds} (\lambda_2 - \underline{\lambda})) ds \leq \lambda_2 \varkappa.$$

Then, using that  $\lambda_2 - \lambda_1 \leq d(\lambda_1 - \underline{\lambda})$ , we get  $\varkappa < 1$ . So, from (56),

$$\begin{aligned} V^\delta(x_n^\delta, \lambda_1) - V^\delta(x_n^\delta, \lambda_2) &\leq J(\pi; x_n^\delta, \lambda_1) - J(\widehat{\pi}; x_n^\delta, \lambda_2) + \varepsilon/2 \\ &\leq J(\pi; x_n^\delta, \lambda_1) - J(\pi; x_n^\delta, \lambda_1)(1 - q\varkappa)(1 - \beta\lambda_2\varkappa) + \varepsilon/2 \\ &\leq V(x_n^\delta, \lambda_1)\varkappa(\beta\lambda_2\varkappa + q) + \varepsilon/2 \\ &\leq V(x, \lambda_1)\left(\beta\lambda_2\frac{(\lambda_2 - \lambda_1)}{d(\lambda_1 - \underline{\lambda})} + q\right)\frac{(\lambda_2 - \lambda_1)}{d(\lambda_1 - \underline{\lambda})} + \varepsilon/2, \end{aligned}$$

so that the result follows.  $\blacksquare$

In order to prove Propositions 6.4 and 6.6, we need the following lemma.

**Lemma A.4** *Given  $\pi \in \widetilde{\Pi}_{x_n^\delta, \lambda}^\delta$ ,  $\pi$  can be obtained by a sequence of control actions  $\mathbf{s} = (s_k)_{k=1, \dots, \tilde{k}}$ , where  $\tilde{k}$  could be finite or infinite. Let us define  $t_k$  as the time when the control action  $s_k$  is applied. We have that  $\lim_{k \rightarrow \infty} t_k = \infty$  a.s. within the subset  $\{\tilde{k} = \infty\} \subset \Omega$ .*

*Proof of Lemma A.4.* Suppose that  $\tilde{k} = \infty$ . Calling  $k_m = n + m$ , let  $i_m$  be the number of control actions  $\mathbf{E}_0$  in  $(s_1, s_2, \dots, s_{k_m})$ , then  $i_m \geq m$ . Given the two sequences of stopping times  $(\tau_i)_{i \geq 1}$  and  $(T_j)_{j \geq 1}$ , we define the ordered union of the two sequences as

$$(\tau_i)_{i \geq 1} \amalg (T_j)_{j \geq 1} := (r_h)_{h \geq 1},$$

where  $\{r_h : h \geq 1\} = \{\tau_i : i \geq 1\} \cup \{T_j : j \geq 1\}$  and  $r_h \leq r_{h+1}$ . We have that  $\lim_{h \rightarrow \infty} r_h = \infty$  a.s. Let us consider the non-decreasing sequence  $(j_m)$  defined as  $j_m = \max\{h : r_h \leq t_{k_m}\}$ , then we have that  $t_{k_m} \geq \max\{r_{j_m}, (i_m - j_m - 1)\delta\}$ . If  $\lim_{m \rightarrow \infty} i_m - j_m = \infty$ , then

$$\lim_{m \rightarrow \infty} t_{k_m} \geq \lim_{m \rightarrow \infty} (i_m - j_m - 1)\delta = \infty;$$

if not,  $\lim_{m \rightarrow \infty} j_m = \infty$  and so

$$\lim_{m \rightarrow \infty} t_{k_m} \geq \lim_{m \rightarrow \infty} r_{j_m}$$

and since  $\lim_{m \rightarrow \infty} r_{j_m} = \lim_{i \rightarrow \infty} r_i = \infty$  a.s., since  $t_{k+1} \geq t_k$ , we have  $\lim_{m \rightarrow \infty} t_k = \infty$  a.s.  $\blacksquare$

*Proof of Proposition 6.4.* Let us define  $W = \lim_{l \rightarrow \infty} V_l^\delta$  and let us show that  $W(x_n^\delta, \lambda) = V^\delta(x_n^\delta, \lambda)$ . Given  $(x_n^\delta, \lambda)$  and  $\varepsilon > 0$ , take  $\pi = (L, \tau_F) \in \widetilde{\Pi}_{x_n^\delta, \lambda}^\delta$  such that  $V^\delta(x_n^\delta, \lambda) - J(\pi; x_n^\delta, \lambda) < \varepsilon/2$ .  $\pi$  can be obtained by a sequence of control actions  $\mathbf{s} = (s_k)_{k=1, \dots, \tilde{k}}$ , where  $\tilde{k}$  could be finite or infinite. Let us define  $t_k$  as the time when the control action  $s_k$  is applied. By Lemma A.4,  $\lim_{k \rightarrow \infty} t_k = \infty$  a.s. within the subset  $\{\tilde{k} = \infty\} \subset \Omega$ . Let us take  $l$  large enough such that  $e^{-qt_l} \sup_{x \geq 0} (V(x, \lambda) - x) \leq e^{-qt_l} \frac{\underline{p}}{q} < \varepsilon/2$  and consider  $\pi^l \in \widetilde{\Pi}_{x_n^\delta, \lambda}^{\delta, l}$ , defined by the sequence  $(s_1, s_2, s_3, \dots, s_{l-1}, \mathbf{E}_F)$  if  $\tilde{k} \geq l$  and by  $\mathbf{s}$  otherwise. We have that

$$\begin{aligned} J(\pi; x_n^\delta, \lambda) - J(\pi^l; x_n^\delta, \lambda) &= \mathbb{E} \left( \mathbb{E}(I_{\tau^L \wedge \tau_F > t_l} \int_{t_l}^{\tau^L} e^{-qs} dL_s + I_{\tau^L \wedge \tau_F > t_l} I_{\{\tau^F < \tau^L\}} e^{-q\tau^F} X_{\tau^F}^L - e^{-qt_l} X_{t_l}^L) \middle| \mathcal{F}_{t_l} \right) \\ &\leq e^{-qt_l} \mathbb{E}(V(X_{t_l}^L, \lambda) - X_{t_l}^L) \\ &< \varepsilon/2, \end{aligned}$$

and so  $V^\delta(x_n^\delta, \lambda) - V_l^\delta(x_n^\delta, \lambda) \leq \varepsilon$ .

Finally, since  $V_l^\delta(x_n^\delta, \lambda) \nearrow V^\delta(x_n^\delta, \lambda)$ , from Proposition 6.3, we get that  $\mathcal{T}(V^\delta)(x_n^\delta, \lambda) = V^\delta(x_n^\delta, \lambda)$ .  $\blacksquare$

*Proof of Proposition 6.6.* Assume that  $\pi = (L, \tau_F) \in \Pi_{x_m^\delta, \lambda}^\delta$ . For any  $\omega = (\tau_i, U_i)_{i \geq 1}, (T_j, Y_j)_{j \geq 1}$ , consider the sequence  $\mathbf{s} = (s_k)_{k=1, \dots, \tilde{k}}$  with  $s_k \in \mathcal{E}$  corresponding to  $\pi$ . Let  $x_m^{\delta, k} \in \mathcal{G}_\delta$  and  $\lambda^k \geq \underline{\lambda}$  be the surplus and the intensity in which the control action  $s_k$  is applied,  $t_k$  be the time at which the control action  $s_k$  is chosen, and let  $y^k$  be the end surplus resulting from the control action  $s_k$ . Denote by  $(\kappa_l)_{l \geq 1}$  the indices of the sequence  $\mathbf{s} = (s_k)_{k=1, \dots, \tilde{k}}$ , where  $s_k$  is either  $\mathbf{E}_F$  or  $\mathbf{E}_0$ . If the sequence stops at  $\tilde{k} = \kappa_{l_0} < \infty$ , we define

$$\kappa_l = \kappa_{l_0} \text{ for } l \geq l_0, \quad t_{\kappa_{l_0} + j} = t_{\kappa_{l_0}} + \Delta_{\kappa_{l_0}} \text{ for } j \geq 1;$$

if  $\tilde{k} = \infty$  we put  $l_0 = \infty$ . We define, for  $l \geq 1$ ,

$$H(l) = W(x_m^{\delta, 1 + \kappa_l}, \lambda^{1 + \kappa_l}) I_{\{s_{\kappa_l} = \mathbf{E}_0\}} I_{\{y^{\kappa_l} \geq 0\}} + x_m^{\delta, \kappa_l} I_{\{s_{\kappa_l} = \mathbf{E}_F\}}.$$

If we put  $H(0) = W(x_m^\delta, \lambda)$ ,  $\kappa_0 = 0$  and  $t_0 = 0$ , we have, using  $T_1(W) - W \leq 0$ ,

$$\begin{aligned}
e^{-qt\kappa_{l+1}} H(l) - W(x_m^\delta, \lambda) &= \sum_{j=1}^l (e^{-qt\kappa_{j+1}} H(j) - e^{-qt\kappa_j} H(j-1)) \\
&= \sum_{j=1}^l I_{\{\kappa_{j+1} \neq \kappa_j\}} (e^{-qt\kappa_{j+1}} H(j) - e^{-qt\kappa_j} H(j-1)) \\
&= \sum_{j=1}^l I_{\{\kappa_{j+1} \neq \kappa_j\}} (e^{-qt_1 + \kappa_j - 1} (\sum_{k=1+\kappa_j-1}^{\kappa_j-1} (W(x_{m^k}^\delta - p\delta) - W(x_{m^k}^\delta))) \\
&\quad + \sum_{j=1}^l I_{\{\kappa_{j+1} \neq \kappa_j\}} (e^{-qt\kappa_{j+1}} H(j) - e^{-qt\kappa_j} W(x_{m^k}^\delta)) \\
&\leq \sum_{j=1}^l I_{\{\kappa_{j+1} \neq \kappa_j\}} (\sum_{k=1+\kappa_j-1}^{\kappa_j-1} e^{-qt_1 + \kappa_j - 1} (-p\delta I_{\{s_k = \mathbf{E}_i\}})) \\
&\quad + \sum_{j=1}^l I_{\{\kappa_{j+1} \neq \kappa_j\}} (e^{-qt\kappa_{j+1}} H(j) - e^{-qt\kappa_j} W(x_{m^k}^\delta)).
\end{aligned} \tag{57}$$

Since  $T_0(W) - W \leq 0$  and  $T_F(W) - W \leq 0$ , if  $\kappa_{j+1} \neq \kappa_j$ ,

$$\begin{aligned}
&\mathbb{E} \left( e^{-qt\kappa_{j+1}} H(j) - e^{-qt\kappa_j} W(x_{m^k}^\delta, \lambda^{k_j}) \middle| \mathcal{F}_{t\kappa_j} \right) \\
&= \mathbb{E} \left( (e^{-qt\kappa_{j+1}} H(j) - e^{-qt\kappa_j} W(x_{m^k}^\delta, \lambda^{k_j})) I_{\{s_{\kappa_j} = \mathbf{E}_0\}} \middle| \mathcal{F}_{t\kappa_j} \right) + I_{\{s_{\kappa_j} = \mathbf{E}_F\}} e^{-qt\kappa_j} (x_{m^k}^\delta - W(x_{m^k}^\delta, \lambda^{k_j})) \\
&\leq \mathbb{E} \left( e^{-qt\kappa_{j+1}} I_{\{s_{\kappa_j} = \mathbf{E}_0\}} (W(x_{m^k}^\delta, \lambda^{k_j+1}) I_{\{y^{\kappa_j} \geq 0\}}) \middle| \mathcal{F}_{t\kappa_j} \right) - e^{-qt\kappa_j} W(x_{m^k}^\delta, \lambda^{k_j}) I_{\{s_{\kappa_j} = \mathbf{E}_0\}} \\
&= e^{-qt\kappa_j} I_{\{s_{\kappa_j} = \mathbf{E}_0\}} (T_0(W)(x_{m^k}^\delta, \lambda^{k_j}) - W(x_{m^k}^\delta, \lambda^{k_j})) \\
&\leq 0.
\end{aligned} \tag{58}$$

From (57) and (58), we have

$$\limsup_{l \rightarrow \infty} \mathbb{E} \left( e^{-qt\kappa_{l+1}} H(l) - W(x_m^\delta, \lambda) \right) \leq -\mathbb{E} \left( \int_{0-}^{(\tau_L \wedge \tau_F)^-} e^{-qs} dL_s \right).$$

Consequently,

$$W(x_m^\delta, \lambda) \geq J(\pi; x_m^\delta, \lambda) + \limsup_{l \rightarrow \infty} \mathbb{E} \left( e^{-qt\kappa_{l+1}} (W(x_{m^{1+\kappa_l}}^\delta, \lambda^{1+\kappa_l}) I_{\{s_{\kappa_l} = \mathbf{E}_0\}} I_{\{y^k \geq 0\}}) \right).$$

Since  $W$  satisfies the growth condition (13), by Lemma A.4,

$$\limsup_{l \rightarrow \infty} \mathbb{E} \left( e^{-qt\kappa_{l+1}} (W(x_{m^{1+\kappa_l}}^\delta, \lambda^{1+\kappa_l}) I_{\{s_{\kappa_l} = \mathbf{E}_0\}} I_{\{y^k \geq 0\}}) \right) = 0,$$

and so we have the result.  $\blacksquare$

In order to prove Theorem 6.9, we need the next definition.

**Definition A.3** We define the auxiliary function  $\bar{V} : [0, \infty) \times [\underline{\lambda}, \infty) \rightarrow \mathbf{R}$  as

$$\bar{V}(x, \lambda) := \lim_{k \rightarrow \infty} V^{\delta k}(x, \lambda).$$

We will prove that  $\bar{V}$  is the optimal value function. In order to do that, we will show that  $\bar{V}$  is a viscosity supersolution of (11). It is straightforward to see that  $\bar{V}$  is a limit of value functions of admissible strategies in  $\Pi_{x, \lambda}$  for all  $(x, \lambda) \in [0, \infty) \times [\underline{\lambda}, \infty)$ , so the result will follow from Corollary 4.3. Since there is no uniqueness of the solution of the HJB equation, it is essential to show that this function is a limit of value functions of admissible strategies. In the next lemma, we find a bound on the variation of  $V^\delta$  and as a consequence we obtain that  $\bar{V}$  is locally Lipschitz in  $[0, \infty) \times [\underline{\lambda}, \infty)$  and so it is absolutely continuous.

**Lemma A.5** We have that  $\bar{V}$  is locally Lipschitz in  $[0, \infty) \times (\underline{\lambda}, \infty)$ . That is, for any  $x_2, x_1 \geq 0$  and for any  $\lambda_2, \lambda_1 > \underline{\lambda}$  with  $\lambda_2 - \lambda_1 \leq d(\lambda_1 - \underline{\lambda})$ ,

$$\begin{aligned}
&|\bar{V}(x_2, \lambda_2) - \bar{V}(x_1, \lambda_1)| \\
&\leq |\bar{V}(x_2, \lambda_2) - \bar{V}(x_1, \lambda_2)| + |\bar{V}(x_1, \lambda_2) - \bar{V}(x_1, \lambda_1)| \\
&\leq \bar{V}(x_2 \vee x_1, \underline{\lambda}) \frac{p+q+(\lambda_1 \vee \lambda_2)}{p} |x_2 - x_1| + \bar{V}(x_2 \vee x_1, \underline{\lambda}) \frac{(\beta(\lambda_1 \vee \lambda_2) + q)}{d((\lambda_1 \wedge \lambda_2) - \underline{\lambda})} |\lambda_2 - \lambda_1| \\
&\leq \bar{V}(x_2 \vee x_1, \underline{\lambda}) \left( \frac{p+q+(\lambda_1 \vee \lambda_2)}{p} + \frac{\beta(\lambda_1 \vee \lambda_2) + q}{d((\lambda_1 \wedge \lambda_2) - \underline{\lambda})} \right) (|x_2 - x_1| + |\lambda_2 - \lambda_1|).
\end{aligned}$$

*Proof of Lemma A.5.* We can write, from Proposition 6.1,

$$\begin{aligned}
&\bar{V}(x_2, \lambda) - \bar{V}(x_1, \lambda) \\
&= \bar{V}(x_2, \lambda) - V^{\delta k}(x_2, \lambda) + V^{\delta k}(x_2, \lambda) - V^{\delta k}(x_1, \lambda) + V^{\delta k}(x_1, \lambda) - \bar{V}(x_1, \lambda) \\
&\leq \bar{V}(x_2, \lambda) - V^{\delta k}(x_2, \lambda) + (\rho^{\delta k}(x_2) - \rho^{\delta k}(x_1)) \frac{e^{(\beta+q)\delta k} + \int_0^{\delta k} \lambda_u^\beta du - 1}{p\delta k} V^{\delta k}(x_2, \lambda) + \delta_k p \\
&\quad + V^{\delta k}(x_1, \lambda) - \bar{V}(x_1, \lambda).
\end{aligned}$$

Also,

$$\begin{aligned}
& \bar{V}(x, \lambda_1) - \bar{V}(x, \lambda_2) \\
&= \bar{V}(x, \lambda_1) - V^{\delta_k}(x, \lambda_1) + V^{\delta_k}(x, \lambda_1) - V^{\delta_k}(x, \lambda_2) + V^{\delta_k}(x, \lambda_2) - \bar{V}(x_1, \lambda_2) \\
&\leq \bar{V}(x, \lambda_1) - V^{\delta_k}(x, \lambda_1) + V^\delta(x, \lambda_1) \frac{(\beta\lambda_2 + q)}{d(\lambda_1 - \underline{\lambda})} (\lambda_2 - \lambda_1) + V^{\delta_k}(x, \lambda_2) - \bar{V}(x_1, \lambda_2).
\end{aligned}$$

Taking the limit as  $k$  goes to infinity, we obtain

$$|\bar{V}(x_2, \lambda) - \bar{V}(x_1, \lambda)| \leq \bar{V}(x_2 \vee x_1, \underline{\lambda}) \frac{p + q + \lambda}{p} |x_2 - x_1|$$

and

$$|\bar{V}(x, \lambda_1) - \bar{V}(x, \lambda_2)| \leq \bar{V}(x, \underline{\lambda}) \frac{(\beta\lambda_2 + q)}{d(\lambda_1 - \underline{\lambda})} |\lambda_2 - \lambda_1|.$$

■

In the next lemma, we show that the convergence of  $V^{\delta_k}$  to  $\bar{V}$  is locally uniformly.

**Lemma A.6**  $V^{\delta_k} \nearrow \bar{V}$  locally uniformly as  $k$  goes to infinity.

*Proof of Lemma A.6.* Consider a compact set  $K$  in  $[0, \infty) \times (\underline{\lambda}, \infty)$ ,  $(x_1, \lambda_1) \in K$  and  $\varepsilon > 0$ . Let us take  $M \in [0, \infty)$  such that  $M \geq x$ , for all  $(x, \lambda) \in K$ . We show first that there exists  $k_0$  large enough and  $\eta > 0$  small enough such that if  $|x - x_1| + |\lambda - \lambda_1| < \eta$ , and  $k \geq k_0$ , then

$$\bar{V}(x, \lambda) - V^{\delta_k}(x, \lambda) < \varepsilon. \quad (59)$$

Indeed, by pointwise convergence at  $(x_1, \lambda_1)$ , there exists  $k_1$  such that

$$\bar{V}(x_1, \lambda_1) - V^{\delta_k}(x_1, \lambda_1) < \varepsilon/3 \text{ for } k \geq k_1. \quad (60)$$

By Lemma A.5, there exists  $\eta_1$  such that if  $|x - x_1| + |\lambda - \lambda_1| < \eta_1$ , then

$$|\bar{V}(x, \lambda) - \bar{V}(x_1, \lambda_1)| < \varepsilon/3. \quad (61)$$

Also, from Proposition 6.1, there exists  $\eta_2$  and  $k_2$  such that if  $|x - x_1| + |\lambda - \lambda_1| < \eta_1$ , then

$$\begin{aligned}
|V^{\delta_k}(x, \lambda) - V^{\delta_k}(x_1, \lambda_1)| &\leq V^{\delta_k}(M, \underline{\lambda}) \left( \frac{e^{(\beta+q)\delta} + \int_0^\delta \lambda_u^c du - 1}{p\delta} \right. \\
&\quad \left. + \frac{\beta(\lambda_1 \vee \lambda_2) + q}{d((\lambda_1 \wedge \lambda_2) - \underline{\lambda})} \right) (\rho^\delta(x_2) - \rho^\delta(x_1) + |\lambda_2 - \lambda_1|) + \delta p \\
&< \varepsilon/3
\end{aligned} \quad (62)$$

for  $k \geq k_2$ . Therefore, taking  $\eta := \eta_1 \wedge \eta_2$ , for  $k \geq k_0 := k_1 \vee k_2$ , we obtain (59) from (60), (61) and (62).

Finally, we conclude the result by taking a finite covering of the compact set  $K$ . ■

The next lemma is the key argument in the proof of Theorem 6.9.

**Lemma A.7**  $\bar{V}$  is a viscosity supersolution of (11) in  $(0, \infty) \times (\underline{\lambda}, \infty)$ , and so  $\bar{V} = \lim_{k \rightarrow \infty} V^{\delta_k} = V$ .

*Proof of Lemma A.7.* Take  $(x^0, \lambda^0) \in (0, \infty) \times (\underline{\lambda}, \infty)$  and a differentiable test function  $\varphi : [0, \infty) \times [\underline{\lambda}, \infty) \rightarrow \mathbb{R}$  for a viscosity supersolution of (11) at  $(x^0, \lambda^0)$ , that is

$$\bar{V}(x, \lambda) \geq \varphi(x, \lambda) \text{ and } \bar{V}(x^0, \lambda^0) = \varphi(x^0, \lambda^0). \quad (63)$$

Consider the sets  $K_1 = [x^0, x^0 + \delta_1] \subset (0, \infty)$ ,  $K_2 = [\lambda^0, \lambda^0 + 1] \subset (\underline{\lambda}, \infty)$  and  $K^{\delta_k} = (\mathcal{G}_{\delta_k} \cap K_1) \times K_2 \subset (0, \infty) \times (\underline{\lambda}, \infty)$ . In order to prove that  $\mathcal{L}(\varphi)(x^0, \lambda^0) \leq 0$ , consider now, for  $\eta > 0$  small enough,

$$\varphi_\eta(x, \lambda) = \varphi(x, \lambda) - \eta \left( (x - x^0)^2 + (\lambda - \lambda^0)^2 \right).$$

Given  $k \geq 0$ , the set  $K^{\delta_k}$  is non-empty and compact, so we can define

$$a_k^\eta := \min_{K^{\delta_k}} \{V^{\delta_k}(x, \lambda) - \varphi_\eta(x, \lambda)\} \quad (64)$$

and

$$(x_k^\eta, \lambda_k^\eta) := \arg \min_{K^{\delta_k}} \{V^{\delta_k}(x, \lambda) - \varphi_\eta(x, \lambda)\} \in K^{\delta_k}. \quad (65)$$

Since  $V^{\delta_k} \leq \bar{V}$ , we have from (63), that  $a_k^\eta \leq 0$ . Taking

$$0 \leq b_k^\eta := \max_{K^{\delta_k}} \{\bar{V}(x, \lambda) - V^{\delta_k}(x, \lambda)\},$$

by Lemma A.6,  $a_k^\eta \rightarrow 0$  and  $b_k^\eta \rightarrow 0$  as  $k \rightarrow \infty$ . Moreover, for all  $(x, \lambda) \in K^{\delta_k}$ , we get from (63), (64) and (65) that

$$\begin{aligned} a_k^\eta &= V^{\delta_k}(x_k^\eta, \lambda_k^\eta) - \varphi_\eta(x_k^\eta, \lambda_k^\eta) \\ &= V^{\delta_k}(x_k^\eta, \lambda_k^\eta) - \bar{V}(x_k^\eta, \lambda_k^\eta) + \bar{V}(x_k^\eta, \lambda_k^\eta) - \varphi(x_k^\eta, \lambda_k^\eta) + \eta \left( (x_k^\eta - x^0)^2 + (\lambda_k^\eta - \lambda^0)^2 \right) \\ &\geq -b_k + \eta \left( (x_k^\eta - x^0)^2 + (\lambda_k^\eta - \lambda^0)^2 \right). \end{aligned}$$

Then, the minimum argument in (64) is attained at  $(x_k^\eta, \lambda_k^\eta) \in K^{\delta_k}$  such that

$$(x_k^\eta - x^0)^2 + (\lambda_k^\eta - \lambda^0)^2 \leq (b_k + a_k)/\eta.$$

Hence,  $(x_k^\eta, \lambda_k^\eta) \rightarrow ((x^0)^+, (\lambda^0)^+)$  as  $k$  goes to infinity. So

$$V^{\delta_k}(x, \lambda) \geq \varphi_\eta(x, \lambda) + a_k^\eta \text{ for } (x, \lambda) \in K^{\delta_k} \text{ and } V^{\delta_k}(x_k^\eta, \lambda_k^\eta) = \varphi_\eta(x_k^\eta, \lambda_k^\eta) + a_k^\eta. \quad (66)$$

Since

$$\mathcal{T}_0(V^{\delta_k})(x_k^\eta, \lambda_k^\eta) - V^{\delta_k}(x_k^\eta, \lambda_k^\eta) \leq 0,$$

we obtain

$$\begin{aligned} 0 &\geq \lim_{k \rightarrow \infty} \frac{\mathcal{T}_0^{\delta_k}(\varphi_\eta)(x_k^\eta, \lambda_k^\eta) - \varphi_\eta(x_k^\eta, \lambda_k^\eta) - a_k^\eta(1 - e^{-q\delta_k})}{\delta_k} \\ &= \lim_{k \rightarrow \infty} \frac{\mathcal{T}_0^{\delta_k}(\varphi_\eta)(x_k^\eta, \lambda_k^\eta) - \varphi_\eta(x_k^\eta, \lambda_k^\eta)}{\delta_k} \\ &= \mathcal{L}(\varphi_\eta)(x^0, \lambda^0). \end{aligned}$$

Since  $\partial_x(\varphi_\eta)(x^0, \lambda^0) = \partial_x(\varphi)(x^0, \lambda^0)$  and  $\partial_\lambda(\varphi_\eta)(x^0, \lambda^0) = \partial_\lambda(\varphi)(x^0, \lambda^0)$  and  $\varphi_\eta \nearrow \varphi$  as  $\eta \searrow 0$ , we obtain that  $\mathcal{L}(\varphi)(x^0, \lambda^0) \leq 0$  and the result follows.  $\blacksquare$

*Proof of Theorem 6.9.* From Lemmas A.6 and A.7 we have that for any  $\delta > 0$ , the functions  $V^{\delta_k} \nearrow \bar{V} = V$  locally uniformly as  $k$  goes to infinity. From Proposition 5.2, it is enough to show that  $V^{\delta_k} \nearrow V$  uniformly in  $[0, \infty) \times [\lambda, \lambda_1]$  for any  $\lambda_1 > \lambda$ . From Proposition 5.1, take  $x > x^* = p/q + \delta_k$ , then

$$V^{\delta_k}(x^*, \lambda) + (x - x^*) \leq V^{\delta_k}(x, \lambda) \leq V(x, \lambda) = V(x^*, \lambda) + (x - x^*),$$

so

$$\begin{aligned} V(x, \lambda) - V^{\delta_k}(x, \lambda) &= V(x^*, \lambda) + (x - x^*) - V^{\delta_k}(x, \lambda) \\ &\leq V(x^*, \lambda) + (x - x^*) - (V^{\delta_k}(x^*, \lambda) + (x - x^*)) \\ &= V(x^*, \lambda) - V^{\delta_k}(x^*, \lambda), \end{aligned}$$

and the result follows.  $\blacksquare$

## A.6 Proofs of Section 7

*Proof of Proposition 7.1.* Since  $\hat{\lambda}_t \geq \lambda_t$  for all  $t \geq 0$ , with a proof analogous to the one of Proposition 3.1-(1), one can prove that  $\hat{V}^{\delta, \Delta} \leq V^\delta$ . Let us prove now that  $\lim_{\Delta \rightarrow 0} \sup_{x \geq 0, \lambda \geq \lambda} \left( V^\delta(x, \lambda) - \hat{V}^{\delta, \Delta}(x, \lambda) \right) = 0$ . It is enough to do the proof for  $(x, \lambda) \in \mathcal{G}_\delta \times \mathcal{H}_\Delta$ . Take the optimal  $\mathcal{G}_\delta$ -strategy  $\pi = (L, \infty) \in \tilde{\Pi}_{x, \lambda}$  such that

$$V^\delta(x, \lambda) = J(\pi; x, \lambda)$$

and let us call the corresponding ruin time of the controlled process  $X_t^L = X_t - L_t$  as  $\tau^L$ .

Given  $\delta$  and  $\Delta$ , since  $0 \leq \hat{\lambda}_t - \lambda_t \leq \Delta$ , we can write the Poisson process  $\hat{N}_t$  as  $\hat{N}_t = N_t + \bar{N}_t$  where  $\bar{N}_t$  is a Poisson process independent of  $N_t$  and intensity  $\hat{\lambda}_t - \lambda_t$ . Therefore,

$$\hat{X}_t = X_t - \sum_{m=1}^{\bar{N}_t} \bar{U}_m. \quad (67)$$

Define

$$\begin{aligned} \hat{\tau}_1 &= \sup\{t : \hat{X}_t - L_t \geq 0\}, \\ \hat{L}_t &= L_t I_{\{t < \hat{\tau}_1\}} + L_{\hat{\tau}_1} I_{\{t \geq \hat{\tau}_1\}} \end{aligned}$$

and

$$\hat{\tau}^F = \hat{\tau}_1 I_{\{X_{\hat{\tau}_1} - L_{\hat{\tau}_1} \geq 0\}} + \infty I_{\{X_{\hat{\tau}_1} - L_{\hat{\tau}_1} < 0\}}.$$

Then, taking  $\hat{\pi} = (\hat{L}, \hat{\tau}^F)$ , it holds that  $\hat{\pi} \in \hat{\Pi}_{x,\lambda}^{\delta,\Delta}$ . That is, it is admissible for the discrete intensity process  $\hat{\lambda}_t$ . Denote the ruin time of as  $\hat{\tau}^L$  and  $\hat{\tau}_2 = \hat{\tau}^L \wedge \hat{\tau}^F$ . So we have from (67)  $\tau^L \geq \hat{\tau}_2$  and

$$X_{\hat{\tau}_2}^L = X_{\hat{\tau}_2} - L_{\hat{\tau}_2} = \hat{X}_{\hat{\tau}_2} + \sum_{m=1}^{\bar{N}_{\hat{\tau}_2}} \bar{U}_m - L_{\hat{\tau}_2} \leq \sum_{m=1}^{\bar{N}_{\hat{\tau}_2}} \bar{U}_m.$$

We therefore can write

$$\begin{aligned} J(\pi; x, \lambda) - J_{\hat{\lambda}}(\hat{\pi}; x, \lambda) &= \mathbb{E}(\int_{\hat{\tau}_2}^{\tau^L} e^{-qs} dL_s) - \mathbb{E}(I_{\{\hat{\tau}^F < \hat{\tau}^L\}} e^{-q\tau^F} (\hat{X}_{\hat{\tau}^F} - L_{(\hat{\tau}^F)-})) \\ &\leq \mathbb{E}(\int_{\hat{\tau}_2}^{\tau^L} e^{-qs} dL_s) \\ &\leq \mathbb{E}\left(e^{-q\hat{\tau}_2} I_{\{\hat{\tau}_2 < \tau^L\}} V(\sum_{m=1}^{\bar{N}_{\hat{\tau}_2}} \bar{U}_m, \lambda_{\hat{\tau}_2})\right) \\ &\leq \mathbb{E}\left(e^{-q\hat{\tau}_2} I_{\{\hat{\tau}_2 < \tau^L\}} (\sum_{m=1}^{\bar{N}_{\hat{\tau}_2}} \bar{U}_m + \frac{p}{q})\right). \end{aligned} \quad (68)$$

From  $\hat{\lambda}_t - \lambda_t \leq \Delta$ , we get that  $\mathbb{E}(\bar{N}_t) \leq t\Delta$  for any  $t \geq 0$ . So, given  $\varepsilon > 0$  and taking  $\Delta \leq 1$  and  $T$  large enough such that  $e^{-qT} \left(T\Delta \mathbb{E}(U_1) + \frac{p}{q}\right) \leq \frac{\varepsilon}{2}$ , we get

$$\begin{aligned} \mathbb{E}\left(I_{\{\hat{\tau}_2 < \tau^L\}} I_{\{\hat{\tau}_2 > T\}} e^{-q\hat{\tau}_2} \sum_{m=1}^{\bar{N}_{\hat{\tau}_2}} \bar{U}_m + \frac{p}{q}\right) &\leq e^{-qT} \mathbb{E}\left(\sum_{m=1}^{\bar{N}_{\hat{\tau}_2}} \bar{U}_m + \frac{p}{q}\right) \\ &\leq e^{-qT} \left(T\Delta \mathbb{E}(U_1) + \frac{p}{q}\right) \\ &\leq \frac{\varepsilon}{2}. \end{aligned}$$

Moreover,  $\hat{\tau}_2 < \tau^L$  implies  $\bar{N}_{\hat{\tau}_2} \geq 1$ , so choosing  $\Delta \leq \varepsilon / (4 \left(\mathbb{E}(U_1)T + \frac{p}{q}\right))$  and since  $\mathbb{P}(\bar{N}_t \geq 1) \leq 1 - e^{-\Delta t}$ , we obtain

$$\begin{aligned} &\mathbb{E}\left(I_{\{\hat{\tau}_2 < \tau^L\}} I_{\{\hat{\tau}_2 \leq T\}} e^{-q\hat{\tau}_2} (\sum_{m=1}^{\bar{N}_{\hat{\tau}_2}} \bar{U}_m + \frac{p}{q})\right) \\ &\leq \mathbb{E}\left(I_{\bar{N}_T \geq 1} (\sum_{m=1}^{\bar{N}_T} \bar{U}_m + \frac{p}{q})\right) \\ &\leq \mathbb{P}[\bar{N}_T \geq 1] \mathbb{E}\left(\sum_{m=1}^{\bar{N}_T} \bar{U}_m + \frac{p}{q}\right) \\ &\leq (1 - e^{-\Delta T}) \left(T\Delta \mathbb{E}(U_1) + \frac{p}{q}\right) \\ &\leq \frac{\varepsilon}{2}. \end{aligned}$$

So, from (68),

$$J(\pi; x, \lambda) - J_{\hat{\lambda}}(\hat{\pi}; x, \lambda) \leq \varepsilon \text{ for } \Delta \leq \frac{\varepsilon}{4 \left(T\mathbb{E}(U_1) + \frac{p}{q}\right)},$$

and we get the result.  $\blacksquare$

*Proof of Theorem 7.4.* On the one hand, since  $\hat{\lambda}_t \geq \lambda_t$  for all  $t \geq 0$ , one can prove that  $\hat{V}^{\delta,\Delta} \leq W_{\mathcal{P}^{\delta,\Delta}}$  in  $[0, \infty) \times [\lambda, \infty)$  with a proof analogous to the one of Proposition 3.1-(1). On the other hand, given any  $\varepsilon > 0$ , from Theorem 6.9 and from Proposition 7.1 there exists  $\delta$  and  $\Delta$  small enough so that

$$0 \leq V(x, \lambda) - \hat{V}^{\delta,\Delta}(x, \lambda) \leq \varepsilon$$

in  $[0, \infty) \times [\lambda, \infty)$ , which establishes the result.  $\blacksquare$

## References

- [1] Albrecher, H. and Asmussen, S. (2006). Ruin probabilities and aggregate claims distributions for shot noise Cox processes. *Scandinavian Actuarial Journal*, **2006**(2), 86-110.
- [2] Albrecher, H., Azcue, P., Muler, N. (2017). Optimal dividend strategies for two collaborating insurance companies. *Adv. Appl. Probab.* **49**(2), 515-548.
- [3] Albrecher, H., Azcue, P., and Muler, N. (2020). Optimal ratcheting of dividends in insurance. *SIAM Journal on Control and Optimization* **58**(4), 1822-1845.
- [4] Albrecher, H., Azcue, P., and Muler, N. (2023). Optimal dividends under a drawdown constraint and a curious square-root rule. *Finance and Stochastics* **27**(2), 341-400.

- [5] Albrecher, H., Bäuerle, N., and Thonhauser, S. (2011). Optimal dividend-payout in random discrete time. *Statistics and Risk Modeling* **28**(3), 251-276.
- [6] Albrecher, H., and Garcia Flores, B. (2023). Optimal dividend bands revisited: a gradient-based method and evolutionary algorithms. *Scandinavian Actuarial Journal*, 2023(8), 788-810.
- [7] Albrecher, H. and Thonhauser, S. (2009). Optimality results for dividend problems in insurance. *RACSAM-Revista de la Real Academia de Ciencias Exactas, Fisicas y Naturales. Serie A. Matematicas* **103**, No.2, 295–320.
- [8] Asmussen, S. and Albrecher, H.(2010). *Ruin probabilities*. World Scientific, Singapore.
- [9] Avanzi, B. (2009). Strategies for dividend distribution: A review. *North American Actuarial Journal* **13**, 2, 217-251.
- [10] Azcue, P. and Muler, N. (2005). Optimal reinsurance and dividend distribution policies in the Cramér-Lundberg model. *Math. Finance* **15**, 261-308.
- [11] Azcue P. and Muler N. (2014). *Stochastic Optimization in Insurance: a Dynamic Programming Approach*. Springer Briefs in Quantitative Finance. Springer.
- [12] Azcue, P. and Muler N. (2021). A multidimensional problem of optimal dividends with irreversible switching: a convergent numerical scheme. *Appl Math Optim* **83**, 1613–1649.
- [13] Berdel, J. (2014). *Optimale Bandstrategien für Dividendenzahlungen im Cramér-Lundberg-Modell*. PhD thesis, Karlsruher Institut für Technologie.
- [14] Boxma, O., and Mandjes, M. (2021). Shot-noise queueing models. *Queueing Systems* **99**(1-2), 121-159.
- [15] Dassios, A., Jang, J.W. (2003). Pricing of catastrophe reinsurance and derivatives using the Cox process with shot noise intensity. *Finance Stochastics* **7**, 73–95.
- [16] Dassios, A., Jang, J.W. (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. *Journal of Applied Probability* **42**, 93–107.
- [17] Dassios, A., and Zhao, H. (2012). Ruin by dynamic contagion claims. *Insurance: Mathematics and Economics* **51**(1), 93-106.
- [18] De Finetti, B. (1957). Su un’Impostazione Alternativa della Teoria Collettiva del Rischio. *Transactions of the 15th Int. Congress of Actuaries* **2**, 433–443.
- [19] Gerber, H.U. (1969). Entscheidungskriterien fuer den zusammengesetzten Poisson-Prozess. *Schweiz. Aktuarver. Mitt.* (1969), No.1, 185–227.
- [20] Grandits, P. (2019). A two-dimensional dividend problem for collaborating companies and an optimal stopping problem. *Scandinavian Actuarial Journal* **2019**(1), 80-96.
- [21] Gu, J. W., Steffensen, M., and Zheng, H. (2018). Optimal dividend strategies of two collaborating businesses in the diffusion approximation model. *Mathematics of Operations Research* **43**(2), 377-398.
- [22] Jang, J., and Oh, R. (2021). A review on Poisson, Cox, Hawkes, shot-noise Poisson and dynamic contagion process and their compound processes. *Annals of Actuarial Science* **15**(3), 623-644.
- [23] Liu, W., and Cadenillas, A. (2023). Optimal insurance contracts for a shot-noise Cox claim process and persistent insured’s actions. *Insurance: Mathematics and Economics* **109**, 69-93.
- [24] Macci, C., and Torrisi, G. L. (2011). Risk processes with shot noise Cox claim number process and reserve dependent premium rate. *Insurance: Mathematics and Economics* **48**(1), 134-145.
- [25] Møller, J. (2005). Properties of spatial Cox process models. *Journal of Statistical Research of Iran*, **2**(1), 1-18.
- [26] Pojer, S., and Thonhauser, S. (2023). Ruin probabilities in a Markovian shot-noise environment. *Journal of Applied Probability* **60**(2), 542-556.
- [27] Pojer, S., and Thonhauser, S. (2023). The Markovian shot-noise risk model: a numerical method for Gerber-Shiu functions. *Methodology and Computing in Applied Probability* **25**(1), 17.
- [28] Protter, P. (1992): *Stochastic integration and differential equations*. Berlin: Springer Verlag.
- [29] Schmidli, H. (2008). *Stochastic Control in Insurance*. Springer, New York.
- [30] Schmidt, T. (2017). Shot-noise processes in finance. *From Statistics to Mathematical Finance: Festschrift in Honour of Winfried Stute*, 367-385.
- [31] Schneider, R. (2023). Extreme events require new forms of financial collaboration to become more resilient. *Environment Systems and Decisions*, to appear.



- [32] Wheeden, R. L. and Zygmund, A. (1977). *Measure and Integral*. New York: Marcel Dekker.
- [33] Zeller, G., and Scherer, M. (2022). A comprehensive model for cyber risk based on marked point processes and its application to insurance. *European Actuarial Journal* **12**(1), 33-85.