A Multidimensional Problem of Optimal Dividends with Irreversible Switching: A Convergent Numerical Scheme

Pablo Azcue[1] · **Nora Muler**[1]

Abstract

In this paper we study the problem of optimal dividend payment strategy which maximizes the expected discounted sum of dividends to a multidimensional set up of n associated insurance companies where the surplus process follows an n-dimensional compound Poisson process. The general manager of the companies has the possibility at any time to exercise an irreversible switch into another regime; we also take into account an expected discounted value at ruin. This multidimensional dividend problem is a mixed singular control/optimal problem. We prove that the optimal value function is a viscosity solution of the associated HJB equation and that it can be characterized as the smallest viscosity supersolution. The main contribution of the paper is to provide a numerical method to approximate (locally uniformly) the optimal value function by an increasing sequence of sub-optimal value functions of admissible strategies defined in an n-dimensional grid. As a numerical example, we present the optimal time of merger for two insurance companies.

Keywords Mixed singular/switching control problem · Multidimensional compound Poisson process · Optimal dividends · Optimal switching · Hamilton–Jacobi–Bellman equation · Viscosity solutions · Convergence of numerical scheme

1 Introduction

In this paper we study the problem of the optimal dividend payment strategy which maximizes the expected discounted sum of dividends, in a multidimensional setup of n associated insurance companies. We assume that the surplus process follows a multidimensional compound Poisson process. The general manager of the companies

✉ Pablo Azcue
  pazcue@utdt.edu

1   Departamento de Matematicas, Universidad Torcuato Di Tella, Buenos Aires, Argentina

has the possibility to exercise an irreversible switch into another regime at any time; we also take into account an expected discounted value at ruin, which is due at the first time of ruin of one of the companies, and may depend on the value of the surplus of all the companies both at and before this time of ruin. This ruin value is a generalization to the multidimensional setting of the Gerber-Shiu penalty functions introduced in Gerber and Shiu [16].

The problem of optimal dividend payments in the one-dimensional case was proposed by de Finetti [12] and it was studied in different model setups. In the compound Poisson risk model, this problem was studied by Gerber [15] using a limit of an associated discrete problem, and by Azcue and Muler [5] using a dynamic programming approach; see also an overview on this problem in Schmidli [24] and in Azcue and Muler [6]. For the limit diffusion approximations, see for example Asmussen and Taksar [2] and for spectrally negative Lévy risk processes see, for instance, Avram et al. [3] and Loeffen [21].

In the one dimensional case, the final value of the portfolio at ruin is non-positive and it is called a penalty. Let us mention for instance Dickson and Waters [13], where the shareholders take care of the deficit at ruin; Gerber et al. [18] where the penalty is a function depending on the deficit at ruin; Thonhauser and Albrecher [26] where they address the optimal dividend problem with constant penalty. The optimal dividend problem in the spectrally negative Lévy setting was solved by Loeffen and Renaud [22] with an affine penalty function, and by Avram et al. [4] with a general penalty function depending on the deficit at ruin.

The one dimensional dividend problem with the possibility of an irreversible switch was addressed by Ly Vath et al. [23] in the Brownian motion setup and by Azcue and Muler [7] in the compound Poisson setting.

The problem of dividend payment in the case of two insurances companies in the compound Poisson risk model was studied by Czarna and Palmowski [11] for a particular dividend strategy of reflecting two-dimensional risk process from the line, and by Albrecher et al. [1] where they study the optimal dividend strategy for two collaborating companies.

In this paper, the multidimensional dividend problem is a mixed singular control/optimal problem. Its associated Hamilton–Jacobi–Bellman equation (HJB) involves a first-order integro-differential operator, an obstacle operator and n derivative constraints; the integro-differential operator corresponds to the discounted infinitesimal operator of the compound Poisson process, the obstacle is related to the value of the portfolio after switching and the derivative constraints are related to the dividend payments of the companies. We prove that the optimal value function is a viscosity solution of the HJB equation, that it can be characterized as the smallest viscosity supersolution and also that a convergent limit of a family of admissible strategies that is a viscosity solution of the associated HJB equation should be the optimal value function (verification result). These results are natural extensions of the results of [7] to the multidimensional setting.

The way in which the optimal value function solves the HJB equation in the n-dimensional state space suggests the optimal local control: in the closed set where the optimal value function coincides with the obstacle (*switch region*), an immediate switch should be done; in the interior of the set where the integro-differential operator

is zero (*non-action region*), no dividends are paid; and in the interior of the set in which one or more of the derivative constraints are tight (*dividend payment region*), the corresponding companies pay a lump sum of dividends. However, it is not clear what the optimal local control is in the *free boundaries* between the non-action region and the dividend payment region. In the one dimensional case the "free boundaries" are indeed "free points", and it can be seen that the optimal local control at these points is just to pay all the incoming premium as dividends, so the control surplus stays there until the arrival of the next claim. This is the reason why the optimal strategy has a band structure and these free points can be obtained by one-dimensional optimization techniques, see [7]. It is a hard task to obtain the free boundaries in the multidimensional setting and there is no hope of finding a closed-form solution for the optimal value function. The main contribution of this paper is to provide a numerical method to approximate (locally uniformly) the optimal value function by a sequence of sub-optimal value functions of admissible strategies defined in an n-dimensional grid. These sub-optimal value functions solve a discrete version of the HJB equation, and the corresponding sub-optimal strategies are constructed partitioning the grid in switch, non-action and dividend payment regions; so we also obtain numerical approximations of the optimal switch, non-action and dividend payment regions and the free boundaries between them.

For a convergence analysis of a numerical scheme for multidimensional singular control problems in the diffusion setting using Markov chain approximation methods, let us mention Kushner and Martins [20] and Budhiraja and Ross [9]; see also the book of Kushner and Dupuis [19] for an exhaustive survey. Regarding convergence of numerical schemes using the viscosity solution approach, let us mention for instance Souganidis [25] and Barles and Souganidis [8], where they propose a numerical scheme for non-singular control problems in the context of the diffusion setting; roughly speaking, they prove that the solutions of the numerical scheme converge to a viscosity solution of the associated HJB equation and then, using a uniqueness argument, they obtain the convergence result. In the numerical method of the present work, there is not uniqueness of viscosity solutions in the HJB equation; nevertheless, we construct numerically an increasing sequence of value functions of a family of admissible strategies whose limit is a viscosity solution of the associated HJB equation; then, using the verification result mentioned above, we deduce that this limit is the optimal value function.

As an application, we present the optimal time of merger (as change of regime at the switch time) for two insurance companies. We show examples where the non-action region could be non-connected even for exponential claim size distributions. For a criteria of merger being an advantage over keeping the two stand-alone companies under barrier strategies see Gerber and Shiu [17].

The rest of the paper is organized as follows. In Sect. 2, we introduce the model and derive some basic properties of the optimal value function. In Sect. 3, we show that the optimal value function is a viscosity solution of the corresponding (HJB) equation; we also characterize it as the smallest viscosity supersolution and give a verification result. In Sect. 4, we construct a family of admissible strategies at any point in a suitable grid. In Sect. 5, we show that the discrete scheme convergences locally uniformly by taking a suitable sequence of embedded grids. In Sect. 6, we present examples of the problem

of optimal merger time. Finally, in Sect. 7, there is an Appendix with the proofs of the technical lemmas.

We use the following notation: $\mathbf{R}_+^n = [0, \infty)^n$, $\leq$ refers to the element-wise order on $\mathbf{R}^n$, $\mathbf{1} = (1, 1, \ldots, 1) \in \mathbf{N}^n$, $(\mathbf{e}_i)_{i=1,\ldots,n}$ is the standard basis of $\mathbf{R}^n$, $[\mathbf{x}, \mathbf{y}] = \{\mathbf{z} \in \mathbf{R}^n : \mathbf{x} \leq \mathbf{z} \leq \mathbf{y}\}$, $\mathbf{x} \vee \mathbf{y} = (x_1 \vee y_1, \ldots, x_n \vee y_n)$, $\mathbf{x} \wedge \mathbf{y} = (x_1 \wedge y_1, \ldots, x_n \wedge y_n)$.

2 Model

Let us consider that the surplus process of n companies, or branches of the same company, follows an n-dimensional compound Poisson process with drift, that means that the uncontrolled process $\mathbf{X}_t \in \mathbf{R}_+^n$ can be written as

$$\mathbf{X}_t = \mathbf{x}^0 + \mathbf{p}t - \sum_{k=1}^{N_t} \mathbf{U}_k. \tag{2.1}$$

Here $\mathbf{x}^0 \in \mathbf{R}_+^n$ is the initial surplus, $\mathbf{p} = (p_1, \ldots, p_n)$ where $p_i > 0$ is the premium rate of company i, N_t is a Poisson process with intensity λ and the downward jumps $\mathbf{U}_k \in \mathbf{R}_+^n$ are i.i.d. random vectors with joint multivariate distribution function F. We also assume that $\mathbb{E}(\|\mathbf{U}_k\|) < \infty$ and $F(\mathbf{0}) = 0$. We call τ_k the time of arrival of the k-th jump of the process, so $N_t = \max\{k : \tau_k \leq t\}$.

We can describe this model in a rigorous way by defining its filtered probability space $(\Omega, \mathcal{F}, (\mathcal{F}_t)_{t \geq 0}, \mathbb{P})$, where

$$\Omega = \left\{ (\tau_k, \mathbf{U}_k)_{k \geq 1} \in \left([0, \infty) \times \mathbf{R}_+^n\right)^{\mathbf{N}} : \tau_k < \tau_{k+1} \right\}$$

and $\mathcal{F}_t$ is the σ-field generated by the set $\{(\tau_k, \mathbf{U}_k) : \tau_k \leq t\}$. The uncontrolled surplus process $\mathbf{X}_t$ is an $\mathcal{F}_t$-adapted càdlàg (right continuous with left limits) stochastic process. Each company pays dividends to the same shareholders, let $\mathbf{L}_t \in \mathbf{R}_+^n$ be the vector of cumulative amount of dividends paid out up to time t by each company; we say that the dividend payment strategy $\mathbf{L}_t$ is admissible if it is a non decreasing process, càdlàg, adapted with respect to the filtration $(\mathcal{F}_t)_{t \geq 0}$ and satisfies $\mathbf{L}_0 \geq 0$ and $\mathbf{L}_t \leq \mathbf{X}_t$ for any $0 \leq t < \tau^{\mathbf{L}}$, where $\tau^{\mathbf{L}}$ is the time at which the process exits the set $\mathbf{R}_+^n$ due to a jump, that is

$$\tau^{\mathbf{L}} := \inf\{\tau_k : \mathbf{X}_{\tau_k} - \mathbf{L}_{\tau_k-} \notin \mathbf{R}_+^n\}. \tag{2.2}$$

We define the *controlled* surplus process as

$$\mathbf{X}_t^{\mathbf{L}} := \mathbf{X}_t - \mathbf{L}_t. \tag{2.3}$$

It is not possible to pay any dividends once the controlled process $\mathbf{X}_t^{\mathbf{L}}$ exits $\mathbf{R}_+^n$ so we extend $\mathbf{L}_t = \mathbf{L}_{\tau^{\mathbf{L}}-}$ for $t \geq \tau^{\mathbf{L}}$. Note that $\mathbf{X}_{\tau^{\mathbf{L}}}^{\mathbf{L}} = \mathbf{X}_{\tau^{\mathbf{L}}-}^{\mathbf{L}} - \mathbf{U}_{k_0}$ if $\tau^{\mathbf{L}} = \tau_{k_0}$. At time $\tau^{\mathbf{L}}$, the shareholders pay a penalty $\upsilon(\mathbf{X}_{\tau^{\mathbf{L}}-}^{\mathbf{L}}, \mathbf{U}_{k_0})$ (or get a reward in the case that $\upsilon(\mathbf{X}_{\tau^{\mathbf{L}}-}^{\mathbf{L}}, \mathbf{U}_{k_0})$ is negative) depending on the surplus prior to ruin $\mathbf{X}_{\tau^{\mathbf{L}}-}^{\mathbf{L}}$ and the size $\mathbf{U}_{k_0}$ of the last jump of the uncontrolled process. Denote

$$B = \left\{ (\mathbf{x}, \boldsymbol{\alpha}) \in \mathbf{R}_+^n \times \mathbf{R}_+^n \text{ s.t. } \mathbf{x} - \boldsymbol{\alpha} \notin \mathbf{R}_+^n \right\}, \tag{2.4}$$

the function $\upsilon : B \to \mathbf{R}$ generalizes the concept of penalty at ruin. It is natural to assume that the penalty function $\upsilon(\mathbf{x}, \boldsymbol{\alpha})$ is non-increasing on $\mathbf{x}$ and non-decreasing on $\boldsymbol{\alpha}$; furthermore, we assume that $\mathbb{E}\left(|\upsilon(\mathbf{0}, \mathbf{U}_1)|\right) < \infty$. The manager of the company also has the possibility at any time $0 \le t < \tau^{\mathbf{L}}$ to exercise an irreversible switch whose value is associated to a given function $f : \mathbf{R}_+^n \to \mathbf{R}$. We assume that the function f is either right continuous and non decreasing or continuous.

Given an initial surplus $\mathbf{x} \ge 0$, let us denote by $\Pi_{\mathbf{x}}$ the set of all pairs $\pi = (\mathbf{L}, \overline{\tau})$ where $\mathbf{L}$ is an admissible dividend payment strategy and $\overline{\tau}$ is a switch time. We define

$$
\begin{aligned}
J(\pi; \mathbf{x}) = {} & \mathbb{E}_{\mathbf{x}} \left(\int_{0-}^{\overline{\tau} \wedge \tau^{\mathbf{L}}} e^{-cs} \mathbf{a} \cdot d\mathbf{L}_s + I_{\{\overline{\tau} < \tau^{\mathbf{L}}\}} e^{-c\overline{\tau}} f(\mathbf{X}_{\overline{\tau}}^{\mathbf{L}}) \right) \\
& - \mathbb{E}_{\mathbf{x}} \left(I_{\{\overline{\tau} \ge \tau^{\mathbf{L}}\}} e^{-c\tau^{\mathbf{L}}} \upsilon \left(\mathbf{X}_{\tau^{\mathbf{L}}-}^{\mathbf{L}}, \mathbf{X}_{\tau^{\mathbf{L}}-}^{\mathbf{L}} - \mathbf{X}_{\tau^{\mathbf{L}}}^{\mathbf{L}} \right) \right)
\end{aligned}
\tag{2.5}
$$

for any $\pi \in \Pi_{\mathbf{x}}$ and the optimal value function as

$$V(\mathbf{x}) = \sup_{\pi \in \Pi_{\mathbf{x}}} J(\pi; \mathbf{x}). \tag{2.6}$$

The value $c > 0$ is a constant discount factor, and $a_i > 0$ are the weights of the dividends paid by the i-th company. The integral in (2.5) is defined as

$$\int_{0-}^{t} e^{-cs} \mathbf{a} \cdot d\mathbf{L}_s = \mathbf{a} \cdot \mathbf{L}_0 + \int_{0}^{t} e^{-cs} \mathbf{a} \cdot d\mathbf{L}_s.$$

Note that we are allowing to make a lump dividend payment $\mathbf{L}_{\overline{\tau}} - \mathbf{L}_{\overline{\tau}-}$ at the switch time $\overline{\tau} < \tau^{\mathbf{L}}$ and also at time zero.

Remark 2.1 (*on the multivariate compound Poisson process*) The most important cases of multivariate compound Poisson process we are considering in the examples correspond to m independent sources of risk that are coinsured between the n insurance companies with different proportions. More precisely, let us assume that there are m independent (univariate) compound Poisson processes given by

$$C^l(t) = \sum_{k=1}^{N_t^l} u_k^l, \tag{2.7}$$

where N_t^l is a Poisson process with intensity λ_l and u_k^l with $k = 1, 2, \dots$ are i.i.d. random variables with distribution F_l. Assume that the total claim arrival process is given by

$$\sum_{j=1}^{N_t} u_j := \sum_{l=1}^{m} C^l(t)$$

and that the i-th company pays a proportion a_{il} of any claim of the l-th compound Poisson process C_t^l. We denote $A := (a_{il}) \in \mathbf{R}^{n \times m}$ with $\sum_{i=1}^n a_{il} = 1$ and $a_{il} \geq 0$. The compound Poisson process $\sum_{l=1}^m C^l(t)$ has intensity $\lambda = \sum_{l=1}^m \lambda_l$. Furthermore,

$$\sum_{k=1}^{N_t} \mathbf{U}_k = A \cdot \left(C^1(t), \ldots, C^m(t) \right)', \tag{2.8}$$

where $N_t = \sum_{l=1}^m N_t^l$ is a compound Poisson process with intensity $\lambda = \sum_{l=1}^m \lambda_l$ and multivariate distribution

$$F(\mathbf{x}) = \mathbb{P}(\mathbf{U} \leq \mathbf{x}) = \sum_{l=1}^m \frac{\lambda_l}{\lambda} F_l \left(\min_{1 \leq i \leq n,\, a_{il} \neq 0} \left\{ \frac{x_i}{a_{il}} \right\} \right).$$

Without loss of generality, we can assume that the columns $\mathbf{a}_l := (a_{il})_{i=1,\ldots,n}$ of the matrix A are different, because if $\mathbf{a}_{l_1} = \mathbf{a}_{l_2}$, we can regard $C^{l_1}(t) + C^{l_2}(t)$ as just one independent source of risk. For instance, in the special case in which the n uncontrolled one-dimensional surplus processes of the companies are independent compound Poisson processes with intensity λ_i and claim size distribution $F_i(x_i)$ ($i = 1, \ldots, n$), A would be the identity matrix and

$$F(\mathbf{x}) = \sum_{i=1}^n \lambda_i F_i(x_i) / \lambda. \tag{2.9}$$

Remark 2.2 *(on the penalty function υ)* Consider the multivariate uncontrolled Poisson process (2.1) described in the previous remark. Suppose that the penalty (or reward) function depends on two factors: (1) which of the m independent compound Poisson processes (2.7) make the controlled process exit $\mathbf{R}_+^n$ and (2) the deficit at this exit time. Let $\mathbf{a}_l = (a_{il})_{i=1,\ldots,n}$ be the l-th column of A, then we have that

$$\upsilon(\mathbf{x}, \boldsymbol{\alpha}) = \sum_{l=1}^m \upsilon_l(\mathbf{x} - \boldsymbol{\alpha}) I_{\{\boldsymbol{\alpha} = \beta_l \mathbf{a}_l \text{ with } \beta_l > 0\}}, \tag{2.10}$$

where $\upsilon_l(\mathbf{X}_{\tau^{\mathbf{L}}}^{\mathbf{L}})$ is the penalty (or reward) when the process $\mathbf{X}_t^{\mathbf{L}}$ exits $\mathbf{R}_+^n$ due to a jump of C^l.

If $n = m = 1$, this definition of penalty function υ includes: (1) the penalty function defined in Gerber and Shiu [16], taking $\upsilon(x, \alpha) = w(x, |x - \alpha|) \geq 0$; (2) the case in which the shareholders take care of the deficit at ruin, taking $\upsilon(x, \alpha) = \alpha - x > 0$ (Dickson and Waters [13]); (3) the case in which the insurer earns continuously Λ as long as the company is alive. This is equivalent to consider $\upsilon(x, \alpha) = \Lambda/c$ (Thonhauser and Albrecher [26]).

In the multidimensional framework, the function υ could be negative, and so considered as a reward. For example, in the case of two companies with independent compound Poisson processes as in (2.9), we can consider the situation in which if one

of the companies goes to ruin, the other survives and continues paying dividends with its own optimal policy. In this case, A is the 2×2 identity matrix and

$$\upsilon(\mathbf{x}, \boldsymbol{\alpha}) = - \left(V_2(x_2) I_{\{x_1 - \alpha_1 < 0\}} + V_1(x_1) I_{\{x_2 - \alpha_2 < 0\}} \right), \tag{2.11}$$

where V_i is the optimal dividend payment function of the i-th company. Note that $\upsilon(\mathbf{x}, \boldsymbol{\alpha})$ is non-increasing on $\mathbf{x}$ and non-decreasing on $\boldsymbol{\alpha}$.

Remark 2.3 *(on the switch-value function f)* The switch-value function $f(\mathbf{x})$ can be thought of as the price in which the shareholders can sell the shares when the controlled current surplus of the n companies is $\mathbf{x}$. It also can be though as the present value of all the dividends paid in the future after a change of regime is decided by the manager (this change of regime could have a cost); for instance, if the manager decides to merge the n companies (that is the n companies put together all their surpluses, pay all the claims and pay dividends until the joined surplus becomes negative). In the case of merger,

$$f(\mathbf{x}) = V_M(x_1 + x_2 + \ldots + x_n - c_M) I_{\{x_1 + x_2 + \ldots + x_n \geq c_M\}} \tag{2.12}$$

where the one-dimensional function V_M is the optimal dividend payment function of the merger of all the companies and $c_M \geq 0$ is the merger cost. So, f is right continuous and non decreasing. The case $n = 2$, A the 2×2 identity matrix, υ as in (2.11) and

$$f(x_1, x_2) = V_M(x_1 + x_2 - c_M) I_{\{x_1 + x_2 \geq c_M\}} \tag{2.13}$$

corresponds to the problem of optimal time of merger proposed by Gerber and Shiu [17]. The case where no switching is allowed is also included in this work, just consider f small enough (see Remark 2.6).

In the next proposition we give sufficient conditions under which the function V is well defined. We say that a function $h : \mathbf{R}^n_+ \to \mathbf{R}$ satisfies the growth condition **GC** if

$$h(\mathbf{x})/h_0(\mathbf{x}) \text{ is upper bounded in } \mathbf{R}^n_+, \tag{2.14}$$

where

$$h_0(\mathbf{x}) := e^{\frac{c}{2n} \sum_{i=1}^{n} \frac{x_i}{p_i}}. \tag{2.15}$$

Proposition 2.4 *If the functions f and $S(\mathbf{x}) := \sup_{\{\boldsymbol{\alpha}:(\mathbf{x},\boldsymbol{\alpha})\in B\}} (-\upsilon(\mathbf{x}, \boldsymbol{\alpha}))$ satisfy the growth condition **GC**, then V is well defined, satisfies the growth condition **GC** and $V \geq -\mathbb{E}(|\upsilon(\mathbf{0}, \mathbf{U}_1)|).$*

Proof Take any initial surplus $\mathbf{x} \geq \mathbf{0}$ and any admissible strategy $\pi = (\mathbf{L}, \overline{\tau}) \in \Pi_{\mathbf{x}}$, since $\mathbf{L}_t \leq \mathbf{X}_t \leq \mathbf{x} + \mathbf{p}t$, we have (using integration by parts),

$$
\mathbb{E}_{\mathbf{x}} \left(\int_{0-}^{\tau^{\mathbf{L}} \wedge \overline{\tau}} e^{-cs} dL_i(s) \right) = \mathbb{E}_{\mathbf{x}} \left(\int_{0}^{\tau^{\mathbf{L}} \wedge \overline{\tau}} e^{-cs} dL_i(s) \right) + L_i(0)
$$
$$
\leq \mathbb{E}_{\mathbf{x}} \left(\int_{0}^{\tau^{\mathbf{L}} \wedge \overline{\tau}} e^{-cs} d(x_i + p_i s) \right) + x_i \leq x_i + \frac{p_i}{c}.
$$

So

$$
\mathbb{E}_{\mathbf{x}} \left(\int_{0-}^{\overline{\tau} \wedge \tau^{\mathbf{L}}} e^{-cs} \mathbf{a} \cdot d\mathbf{L}_s \right) \leq \mathbf{a} \cdot \left(\mathbf{x} + \frac{\mathbf{p}}{c} \right) \leq d_1 e^{\frac{c}{2n} \sum_{i=1}^{n} \frac{x_i}{p_i}} = d_1 h_0(\mathbf{x})
$$

for $d_1 \geq 2n \max \{a_1, \ldots, a_n\} \max \{p_1, \ldots, p_n\} / c$ since

$$
e^{\frac{c}{2n} \sum_{i=1}^{n} \frac{x_i}{p_i}} \geq 1 + \frac{c}{2n} \sum_{i=1}^{n} \frac{x_i}{p_i}.
$$

Consider the processes $\mathbf{z}_s = \mathbf{X}_s^{\mathbf{L}}$ defined in (2.3) and let us call $\tau = \tau^{\mathbf{L}}$. We get that $\mathbf{z}(s) \leq \mathbf{x} + \mathbf{p}s$ and f satisfies (2.14) in $\mathbf{R}_+^n$, so

$$
\mathbb{E}_{\mathbf{x}} \left(e^{-c\overline{\tau}} f(\mathbf{z}_{\overline{\tau}}) I_{\{\tau > \overline{\tau}\}} \right) \leq d_2 e^{\sum_{i=1}^{n} \frac{cx_i}{2np_i}} = d_2 h_0(\mathbf{x})
$$

for d_2 large enough. Similarly,

$$
\mathbb{E}_{\mathbf{x}} \left(-e^{-c\tau} \upsilon(\mathbf{z}_{\tau-}, \mathbf{z}_{\tau-} - \mathbf{z}_\tau) I_{\{\tau \leq \overline{\tau}\}} \right) \leq \mathbb{E}_{\mathbf{x}} \left(e^{-c\tau} I_{\{\tau \leq \overline{\tau}\}} S(\mathbf{z}_{\tau-}) \right)
$$
$$
\leq d_3 e^{\sum_{i=1}^{n} \frac{cx_i}{2np_i}} = d_3 h_0(\mathbf{x})
$$

for d_3 large enough. Then $J(\pi; \mathbf{x})$ (and so $V(\mathbf{x})$) satisfy the growth condition (2.14) in $\mathbf{R}_+^n$. Finally, since τ is the first time that the controlled process $\mathbf{X}^{\mathbf{L}}$ leaves $\mathbf{R}_+^n$, calling $\mathbf{U}_{k_0}$ the jump size at τ, we have $\mathbf{z}_\tau = \mathbf{z}_{\tau-} - \mathbf{U}_{k_0} \geq \mathbf{0} - \mathbf{U}_{k_0}$. Since $-\upsilon(\mathbf{x}, \boldsymbol{\alpha})$ is non-decreasing on $\mathbf{x}$ and non-increasing on $\boldsymbol{\alpha}$, we obtain taking the strategy with no switching and no dividend payment, that

$$
V(\mathbf{x}) \geq \mathbb{E}_{\mathbf{x}} \left(-e^{-c\tau} \upsilon(\mathbf{z}_{\tau-}, \mathbf{U}_{k_0}) \right) \geq \mathbb{E}_{\mathbf{x}} \left(-e^{-c\tau} \upsilon(\mathbf{0}, \mathbf{U}_{k_0}) \right) \geq -\mathbb{E} \left(|\upsilon(\mathbf{0}, \mathbf{U}_1)| \right).
$$

$\square$

Remark 2.5 Let us extend the definition of υ to the closure of B as $\upsilon(\mathbf{x}, \boldsymbol{\alpha}) = \inf_{\boldsymbol{\beta} \geq \boldsymbol{\alpha}, (\mathbf{x}, \boldsymbol{\beta}) \in B} \upsilon(\mathbf{x}, \boldsymbol{\beta})$. Since $-\upsilon(\mathbf{x}, \boldsymbol{\alpha})$ is non-decreasing on $\mathbf{x}$ and non-increasing on $\boldsymbol{\alpha}$ then

$$
\sup_{\boldsymbol{\alpha} \geq \mathbf{0}, \mathbf{x} - \boldsymbol{\alpha} \notin \mathbf{R}_+^n} (-\upsilon(\mathbf{x}, \boldsymbol{\alpha})) \leq \max_{i=1, \ldots n} (-\upsilon(\mathbf{x}, x_i \mathbf{e}_i))
$$

and so the assumption on υ of Proposition 2.4 becomes that $\max_{i=1, \ldots n} (-\upsilon(\mathbf{x}, x_i \mathbf{e}_i))$ satisfies the growth condition **GC**.

Remark 2.6 By Proposition 2.4, taking any switch-value function $f < -\mathbb{E}\left(|\upsilon(\mathbf{0}, \mathbf{U}_1)|\right)$, it is never optimal to switch. So, the problem of maximizing the expected cumulative discounted dividend payments until $\tau^{\mathbf{L}}$ (without the possibility of switching) is a particular case of the problem (2.6).

Remark 2.7 Consider $f_1 \le f_2$ and $\upsilon_1 \ge \upsilon_2$. Let V_{f_1, υ_1} and V_{f_2, υ_2} be the corresponding optimal value functions, then it is straightforward to see that $V_{f_1, \upsilon_1} \le V_{f_2, \upsilon_2}$.

Remark 2.8 Since the optimal dividend payment function in the one-dimensional problem has linear growth, see for instance Proposition 1.2 in [6]; the functions (2.11) and (2.12) satisfy the conditions of Proposition 2.4.

In the next proposition, we show that V is increasing and locally Lipschitz (so it is absolutely continuous).

Proposition 2.9 *V is increasing, locally Lipschitz in $\mathbf{R}_+^n$ and satisfies for each $\mathbf{x} \in \mathbf{R}_+^n$, $h > 0$ and $1 \le i \le n$,*

$$a_i h \le V(\mathbf{x} + h\mathbf{e}_i) - V(\mathbf{x}) \le \left(e^{(c+\lambda)h/p_i} - 1\right) V(\mathbf{x}).$$

Proof Given $h > 0$ and $\mathbf{x} \in \mathbf{R}_+^n$, consider for each $\varepsilon > 0$ an admissible strategy $\pi_{\mathbf{x}} = (\mathbf{L}, \overline{\tau}) \in \Pi_{\mathbf{x}}$ such that $J(\pi; \mathbf{x}) \ge V(\mathbf{x}) - \varepsilon$. Let us define an strategy $\widetilde{\pi} \in \Pi_{\mathbf{x}+h\mathbf{e}_i}$ as follows: the i-th company pays immediately h as dividends and then follows the strategy $\pi \in \Pi_{\mathbf{x}}$. For each $\varepsilon > 0$, we get

$$V(\mathbf{x} + h\mathbf{e}_i) \ge J(\widetilde{\pi}; \mathbf{x} + h\mathbf{e}_i) = J(\pi; \mathbf{x}) + a_i h \ge V(\mathbf{x}) - \varepsilon + a_i h,$$

so we obtain the first inequality.

Now consider for each $\varepsilon > 0$ and $1 \le i \le n$, a strategy $\pi = (\mathbf{L}, \overline{\tau}) \in \Pi_{\mathbf{x}+h\mathbf{e}_i}$ such that

$$J(\pi; \mathbf{x} + h\mathbf{e}_i) \ge V(\mathbf{x} + h\mathbf{e}_i) - \varepsilon.$$

Take now the following admissible strategy $\widetilde{\pi} = (\widetilde{\mathbf{L}}, \widetilde{\tau}) \in \Pi_{\mathbf{x}}$ starting with surplus $\mathbf{x}$: the i-th company pays no dividends and the other companies pay all the incoming premium as dividends as long as $\mathbf{X}_t^{\widetilde{\mathbf{L}}} < \mathbf{x} + h\mathbf{e}_i$; after the current surplus reaches $\mathbf{x} + h\mathbf{e}_i$, follow strategy π. Let us call $\tau^{\widetilde{L}}$ the exit time of the process $\mathbf{X}_t^{\widetilde{\mathbf{L}}}$. If

$$\tau := \min\left\{t : \mathbf{X}_t^{\widetilde{\mathbf{L}}} = \mathbf{x} + h\mathbf{e}_i\right\},$$

then $\widetilde{\tau} = \tau + \overline{\tau}$ and we get that $p_i \tau \ge h$. So,

$$
\begin{aligned}
V(\mathbf{x}) &\ge J(\widetilde{\pi}; \mathbf{x}) \ge J(\pi; \mathbf{x} + h\mathbf{e}_i) \mathbb{E}\left(e^{-c\frac{h}{p_i}} I_{\{\tau < \tau\widetilde{L}\}}\right) \\
&\ge (V(\mathbf{x} + h\mathbf{e}_i) - \varepsilon) e^{-c\frac{h}{p_i}} \mathbb{P}\left(\tau < \tau^{\widetilde{L}}\right) \\
&\ge \left(V(\mathbf{x} + h\mathbf{e}_i) - \varepsilon\right) e^{-c\frac{h}{p_i}} \mathbb{P}\left(\tau_1 > \frac{h}{p_i}\right) = (V(\mathbf{x} + h\mathbf{e}_i) - \varepsilon) e^{-(c+\lambda)\frac{h}{p_i}},
\end{aligned}
$$

where τ_1 is the time of the first jump; so we get the second inequality. $\square$

In order to distinguish the jumps of the controlled process due to the jumps of the uncontrolled process from the ones due to lump dividend payments, let us define an auxiliary process which includes the jump of the uncontrolled process occurring at time t but excludes the lump dividend payment occurring at this time as

$$\check{\mathbf{X}}_t^{\mathbf{L}} = \mathbf{X}_t - \mathbf{L}_{t-} = \mathbf{X}_{t-}^{\mathbf{L}} - (\mathbf{X}_{t-} - \mathbf{X}_t). \tag{2.16}$$

Note that $\check{\mathbf{X}}_t^{\mathbf{L}} = \mathbf{X}_{t-}^{\mathbf{L}} - \mathbf{U}_k$ if $t = \tau_k$ and $\check{\mathbf{X}}_t^{\mathbf{L}} = \mathbf{X}_{t-}^{\mathbf{L}}$ otherwise. Also, $\mathbf{X}_{\tau^{\mathbf{L}}}^{\mathbf{L}} = \check{\mathbf{X}}_{\tau^{\mathbf{L}}}^{\mathbf{L}}$ because no dividends are paid at the exit time $\tau^{\mathbf{L}}$.

3 HJB Equation

In this section we show that the optimal value function V defined in (2.6) is a viscosity solution of the corresponding Hamilton–Jacobi–Bellman (HJB) equation; moreover we characterize the optimal value function as the smallest viscosity supersolution with growth condition **GC**. We also give a verification result for V. These results are a generalization to the multidimensional case of the ones given in Sect. 3 of [7] for the one dimensional case.

The HJB equation of problem (2.6) can be written as

$$\max\{\mathbf{a} - \nabla V(\mathbf{x}), \mathcal{L}(V)(\mathbf{x}), f(\mathbf{x}) - V(\mathbf{x})\} = 0, \tag{3.1}$$

where

$$\mathcal{L}(V)(\mathbf{x}) = \mathbf{p} \cdot \nabla V(\mathbf{x}) - (c + \lambda)V(\mathbf{x}) + \mathcal{I}(V)(\mathbf{x}) - \mathcal{R}(\mathbf{x}), \tag{3.2}$$

$$\mathcal{I}(W)(\mathbf{x}) := \lambda \int_{(\mathbf{x} - \boldsymbol{\alpha}) \in \mathbf{R}_+^n} W(\mathbf{x} - \boldsymbol{\alpha}) dF(\boldsymbol{\alpha}) \text{ and } \mathcal{R}(\mathbf{x}) := \lambda \int_{(\mathbf{x} - \boldsymbol{\alpha}) \notin \mathbf{R}_+^n} \upsilon(\mathbf{x}, \boldsymbol{\alpha}) dF(\boldsymbol{\alpha}). \tag{3.3}$$

As usual, the operator $\mathcal{L}$ is the discounted infinitesimal generator of the uncontrolled surplus process $\mathbf{X}_t$ defined in (2.1); that is, for any continuously differentiable function $W : \mathbf{R}_+^n \to \mathbf{R}$, we have

$$\mathcal{L}(W)(\mathbf{x}) = \lim_{t \searrow 0} \frac{\mathbb{E}_{\mathbf{x}}\left(e^{-ct} W(\mathbf{X}_t) - W(\mathbf{x})\right)}{t}.$$

Thus, if W is a solution of $\mathcal{L}(W) = 0$ in an open set in $\mathbf{R}_+^n$, then the process $e^{-ct} W(\mathbf{X}_t)$ is a martingale in this set.

The HJB equation implies that $\mathcal{L}(V) \leq 0$, the condition $\mathcal{L}(V) = 0$ in an open set in $\mathbf{R}_+^n$ would suggest that (locally) the optimal dividend strategy consists of paying no dividends as long as the current surplus is in this set. The HJB equation also implies that V is always above f, so f can be interpreted as *an obstacle* in equation (3.1).

Moreover, the condition $V_{x_i}(\mathbf{x}) = a_i$ in an open set means that (locally) the optimal dividend strategy should be the one in which the i-th company pays immediately a lump sum as dividends.

We prove in this section that, under the assumption

$$\mathcal{R} : \mathbf{R}_+^n \to \mathbf{R} \text{ is continuous,} \tag{3.4}$$

the value function V is a viscosity solution of the HJB equation (3.1). From now on, we assume that this assumption holds.

Crandall and Lions [10] introduced the concept of viscosity solutions for first-order Hamilton–Jacobi equations. It is the standard tool for studying HJB equations, see for instance Fleming and Soner [14]. In the context of using viscosity solutions for the problem of dividend payment optimization in the one-dimensional case, see for instance [6].

Definition 3.1 (a) A locally Lipschitz function $\underline{u} : \mathbf{R}_+^n \to \mathbf{R}$ is a viscosity subsolution of (3.1) at $\mathbf{x} \in \mathbf{R}_+^n$ if any continuously differentiable function ψ defined in $\mathbf{R}_+^n$ with $\psi(\mathbf{x}) = \underline{u}(\mathbf{x})$ such that $\underline{u} - \psi$ reaches the maximum at $\mathbf{x}$ satisfies

$$\max\{\mathbf{a} - \nabla\psi(\mathbf{x}), \mathcal{L}(\psi)(\mathbf{x}), f(\mathbf{x}) - \psi(\mathbf{x})\} \geq 0.$$

(b) A locally Lipschitz function $\overline{u} : \mathbf{R}_+^n \to \mathbf{R}$ is a viscosity supersolution of (3.1) at $\mathbf{x} \in \mathbf{R}_+^n$ if any continuously differentiable function φ defined in $\mathbf{R}_+^n$ with $\varphi(\mathbf{x}) = \overline{u}(\mathbf{x})$ such that $\overline{u} - \varphi$ reaches the minimum at $\mathbf{x}$ satisfies

$$\max\{\mathbf{a} - \nabla\varphi(\mathbf{x}), \mathcal{L}(\varphi)(\mathbf{x}), f(\mathbf{x}) - \varphi(\mathbf{x})\} \leq 0.$$

(c) A locally Lipschitz function $u : \mathbf{R}_+^n \to \mathbf{R}$ is a viscosity solution of (3.1) if it is both a viscosity subsolution and a viscosity supersolution at any $\mathbf{x} \in \mathbf{R}_+^n$.

In order to prove that V is a viscosity solution of the HJB equation we need to use the following two lemmas. The first one states the Dynamic Programming Principle (DPP); its proof follows from standard arguments, see for instance Lemma 1.2 of [6]. The proof of the second one is in the Appendix.

Lemma 3.2 *Given any* $\mathbf{x} \in \mathbf{R}_+^n$ *and any finite stopping time* $\widetilde{\tau}$, *we have that the function* V *defined in* (2.6) *satisfies* $V(\mathbf{x}) = \sup_{\pi=(\mathbf{L},\overline{\tau})\in\Pi_{\mathbf{x}}} v_{\pi,\widetilde{\tau}}(\mathbf{x})$, *where*

$$v_{\pi,\widetilde{\tau}}(\mathbf{x}) = \mathbb{E}_{\mathbf{x}}\left(\int_{0-}^{\overline{\tau}\wedge\tau^{\mathbf{L}}\wedge\widetilde{\tau}} e^{-cs}\mathbf{a}\cdot d\mathbf{L}_s + e^{-c(\overline{\tau}\wedge\tau^{\mathbf{L}}\wedge\widetilde{\tau})}\left(I_{\{\overline{\tau}\wedge\widetilde{\tau}<\tau^{\mathbf{L}}\}}V(\mathbf{X}_{\overline{\tau}\wedge\widetilde{\tau}}^{\mathbf{L}})\right)\right)$$
$$- \mathbb{E}_{\mathbf{x}}\left(I_{\{\tau^{\mathbf{L}}\leq\overline{\tau}\wedge\widetilde{\tau}\}}\upsilon\left(\mathbf{X}_{\tau^{\mathbf{L}}-}^{\mathbf{L}}, \mathbf{X}_{\tau^{\mathbf{L}}-}^{\mathbf{L}} - \mathbf{X}_{\tau^{\mathbf{L}}}^{\mathbf{L}}\right)\right).$$

Lemma 3.3 *Given any continuously differentiable function* $g : \mathbf{R}_+^n \to \mathbf{R}$, *any admissible strategy* $\pi = (\mathbf{L}, \overline{\tau}) \in \Pi_{\mathbf{x}}$ *and any finite stopping time* $\tau \leq \tau^{\mathbf{L}}$, *consider*

$$\mathbf{L}_t = \int_0^t d\mathbf{L}_s^c + \sum_{0\leq s\leq t}\Delta\mathbf{L}_s,$$

where $\Delta \mathbf{L}_s = \mathbf{L}_s - \mathbf{L}_{s-}$ and $\mathbf{L}_s^c$ is a continuous and non-decreasing process. Then we have

$$
\left(g\left(\mathbf{X}_\tau^{\mathbf{L}}\right) I_{\{\tau < \tau^{\mathbf{L}}\}} - \upsilon\left(\mathbf{X}_{\tau-}^{\mathbf{L}}, \mathbf{X}_{\tau-}^{\mathbf{L}} - \mathbf{X}_\tau^{\mathbf{L}}\right) I_{\{\tau = \tau^{\mathbf{L}}\}} \right) e^{-c\tau} - g(\mathbf{x})
$$
$$
= \int_0^\tau \mathcal{L}(g)\left(\mathbf{X}_{s-}^{\mathbf{L}}\right) e^{-cs} ds - \int_{0-}^\tau e^{-cs} \mathbf{a} \cdot d\mathbf{L}_s + \int_0^\tau e^{-cs}\left(\mathbf{a} - \nabla g\left(\mathbf{X}_{s-}^{\mathbf{L}}\right)\right) \cdot d\mathbf{L}_s^c
$$
$$
+ \sum_{\mathbf{L}_s \neq \mathbf{L}_{s-}, s \leq \tau} e^{-cs} \int_0^1 \left(\mathbf{a} - \nabla g\left(\check{\mathbf{X}}_s^{\mathbf{L}} - \gamma \Delta \mathbf{L}_s\right) \cdot \Delta \mathbf{L}_s\right) d\gamma + M(\tau);
$$

where $M(t)$ is a martingale with zero expectation.

Proposition 3.4 *The optimal value function V is a viscosity solution of the HJB equation (3.1) at any $\mathbf{x}$ in the interior of $\mathbf{R}_+^n$.*

Proof Let us show that V is a viscosity supersolution at any $\mathbf{x}$ in the interior of $\mathbf{R}_+^n$. The inequality $V \geq f$ follows from the definition (2.6) taking $\bar{\tau} = 0$. Given any initial surplus $\mathbf{x}$ in the interior of $\mathbf{R}_+^n$ and any $\mathbf{l} \in \mathbf{R}_+^n$, take $h > 0$ small enough such that $h(\mathbf{l} - \mathbf{p}) < \mathbf{x}$. Consider the dividend payment strategy $\mathbf{L}_t = \mathbf{l}t$ for $t < h \wedge \tau_1$ and $\mathbf{L}_t = \mathbf{l}(h \wedge \tau_1)$ for $t \geq h \wedge \tau_1$; also consider a switch time $\bar{\tau} > \tau^{\mathbf{L}}$. Using Lemma 3.2 with stopping time $\tilde{\tau} = h \wedge \tau_1$, we get

$$
V(\mathbf{x}) \geq \mathbb{E}_\mathbf{x} \left(\mathbf{a} \cdot \mathbf{l} \int_0^{h \wedge \tau_1} e^{-cs} ds + e^{-c(h \wedge \tau_1)} \left(I_{\{h \wedge \tau_1 < \tau^{\mathbf{L}}\}} V\left(\mathbf{X}_{h \wedge \tau_1}^{\mathbf{L}}\right) - I_{\{\tau^{\mathbf{L}} = \tau_1 \leq h\}} \upsilon\left(\mathbf{X}_{\tau_1-}^{\mathbf{L}}, \mathbf{U}_1\right) \right) \right).
$$

Let φ be a test function for supersolution of (3.1) at $\mathbf{x}$ as in Definition 3.1. We have,

$$
\varphi(\mathbf{x}) = V(\mathbf{x})
$$
$$
\geq \mathbb{E}_\mathbf{x} \left(\mathbf{a} \cdot \mathbf{l} \int_0^{h \wedge \tau_1} e^{-cs} ds \right)
$$
$$
+ \mathbb{E}_\mathbf{x} \left(e^{-c(h \wedge \tau_1)} \left(I_{\{h \wedge \tau_1 < \tau^{\mathbf{L}}\}} \varphi\left(\mathbf{X}_{h \wedge \tau_1}^{\mathbf{L}}\right) - I_{\{\tau^{\mathbf{L}} = \tau_1 \leq h\}} \upsilon\left(\mathbf{X}_{\tau_1-}^{\mathbf{L}}, \mathbf{U}_1\right) \right) \right).
$$

We can write

$$
\mathbb{E}_\mathbf{x} \left(e^{-c(h \wedge \tau_1)} \left(I_{\{h \wedge \tau_1 < \tau^{\mathbf{L}}\}} \varphi\left(\mathbf{X}_{h \wedge \tau_1}^{\mathbf{L}}\right) - I_{\{\tau^{\mathbf{L}} = \tau_1 \leq h\}} \upsilon(\mathbf{X}_{\tau_1-}^{\mathbf{L}}, \mathbf{U}_1) \right) \right)
$$
$$
= \mathbb{E}_\mathbf{x} \left(I_{\{h < \tau_1\}} e^{-ch} \varphi(\mathbf{x} + (\mathbf{p} - \mathbf{l}) h) \right)
$$
$$
+ \mathbb{E}_\mathbf{x} \left(I_{\{\tau_1 \leq h\}} I_{\{\mathbf{U}_1 \leq \mathbf{x} + (\mathbf{p}-\mathbf{l})\tau_1\}} e^{-c\tau_1} \varphi(\mathbf{x} + (\mathbf{p} - \mathbf{l}) \tau_1 - \mathbf{U}_1) \right)
$$
$$
- \mathbb{E}_\mathbf{x} \left(I_{\{\tau_1 \leq h\}} I_{\{\mathbf{U}_1 \not\leq \mathbf{x} + (\mathbf{p}-\mathbf{l})\tau_1\}} e^{-c\tau_1} \upsilon(\mathbf{x} + (\mathbf{p} - \mathbf{l}) \tau_1, \mathbf{U}_1) \right).
$$

Here $\not\leq$ means that at least one of the coordinates of the former vector is greater than the corresponding one of the latter. Therefore, using that $\mathcal{R}$ is continuous,

$$0 \geq (\mathbf{a} \cdot \mathbf{l}) \lim_{h \to 0+} \frac{1}{h} \mathbb{E}_{\mathbf{x}} \left(\int_0^{h \wedge \tau_1} e^{-cs} ds \right)$$
$$+ \lim_{h \to 0+} \frac{1}{h} \left(e^{-(\lambda+c)h} \varphi(\mathbf{x} + (\mathbf{p} - \mathbf{l}) h) - \varphi(\mathbf{x}) \right)$$
$$+ \lim_{h \to 0+} \frac{1}{h} \mathbb{E}_{\mathbf{x}} \left(I_{\{\tau_1 \leq h\}} I_{\{\mathbf{U}_1 \leq \mathbf{x} + (\mathbf{p} - \mathbf{l}) \tau_1\}} e^{-c\tau_1} \varphi(\mathbf{x} + (\mathbf{p} - \mathbf{l}) \tau_1 - \mathbf{U}_1) \right)$$
$$- \lim_{h \to 0+} \frac{1}{h} \mathbb{E}_{\mathbf{x}} \left(I_{\{\tau_1 \leq h\}} I_{\{\mathbf{U}_1 \leq \mathbf{x} + (\mathbf{p} - \mathbf{l}) \tau_1\}} e^{-c\tau_1} \upsilon(\mathbf{x} + (\mathbf{p} - \mathbf{l}) \tau_1, \mathbf{U}_1) \right)$$
$$= \mathbf{a} \cdot \mathbf{l} - (c + \lambda) \varphi(\mathbf{x}) + (\mathbf{p} - \mathbf{l}) \cdot \nabla \varphi(\mathbf{x}) + \mathcal{I}(\varphi)(\mathbf{x}) - \mathcal{R}(\mathbf{x}).$$

And so $\mathcal{L}(\varphi)(\mathbf{x}) + \mathbf{l} \cdot (\mathbf{a} - \nabla \varphi(\mathbf{x})) \leq 0$. Taking $\mathbf{l} = \mathbf{0}$, we get $\mathcal{L}(\varphi)(\mathbf{x}) \leq 0$; taking $\mathbf{l} = l \mathbf{e}_i$ with $l \to \infty$ ($1 \leq i \leq n$), we obtain $\mathbf{a} - \nabla \varphi(\mathbf{x}) \leq \mathbf{0}$. So V is a viscosity supersolution at the point $\mathbf{x}$.

We omit the proof that V is a viscosity subsolution in the interior of $\mathbf{R}_+^n$. This result follows from Lemma 3.3 and the proof is similar to the ones of Proposition 3.2 in [7] for the unidimensional case with switching and of Proposition 3.2 in [1] for the multidimensional case with no switching.   □

Remark 3.5 In general, we cannot expect to have uniqueness of viscosity solutions of the HJB equation (3.1). Take for instance the two dimensional case with independent companies, the switch function f given in (2.13) and the function υ defined in (2.11). Consider the function $W_k(\mathbf{x}) := x_1 + x_2 + k$ for $\mathbf{x} \in \mathbf{R}_+^2$, and take k_0 large enough such that,

$$k_0 > \frac{p_1 + p_2}{c}, \quad V_2(z) < z + k_0, \quad V_1(z) < z + k_0 \quad \text{and} \quad V_M(z) < z + c_M + k_0$$

for $z \geq 0$. Hence, we have for all $k \geq k_0$ that $\nabla W_k(\mathbf{x}) - \mathbf{1} = \mathbf{0}$; $f(\mathbf{x}) - W_k(\mathbf{x}) < 0$ and

$$\mathcal{L}(W_k)(\mathbf{x}) \leq p_1 + p_2 - (c + \lambda) W_k(\mathbf{x}) + \lambda_1 W_k(\mathbf{x}) F_1(x_1) + \lambda_2 W_k(\mathbf{x}) F_2(x_2)$$
$$+ \lambda_1(x_2 + k)(1 - F_1(x_1)) + \lambda_2(x_1 + k)(1 - F_2(x_2))$$
$$\leq p_1 + p_2 - (c + \lambda) W_k(\mathbf{x}) + \lambda_1 W_k(\mathbf{x}) + \lambda_2 W_k(\mathbf{x})$$
$$= -c(x_1 + x_2 + K) + p_1 + p_2$$
$$< 0.$$

Therefore, there are infinitely many viscosity solutions of the HJB equation (3.1).

The following lemma states that any viscosity supersolution with the appropriate growth condition is above the value function of a family of admissible strategies. The proof is in the Appendix.

Lemma 3.6 *Fix* $\mathbf{x} \in \mathbf{R}_+^n$, *let* $\overline{u}$ *be a viscosity supersolution of* (3.1) *satisfying growth condition* (2.14) *and take any admissible strategy* $\pi = (\mathbf{L}, \overline{\tau}) \in \Pi_{\mathbf{x}}$, *then* $\overline{u}(\mathbf{x}) \geq J(\pi; \mathbf{x})$ *(and so* $\overline{u}(\mathbf{x}) \geq V(\mathbf{x})$).

There are not natural boundary conditions for the optimal value function V (see for instance Sect. 1.6 of [6] for a discussion about it in the one dimensional case). As a consequence of the previous lemma, we get the following characterization result:

Theorem 3.7 *The optimal value function V can be characterized as the smallest viscosity supersolution of the HJB equation* (3.1) *satisfying growth condition* (2.14).

Also, we obtain immediately the next verification theorem.

Theorem 3.8 *Consider a family of strategies* $(\pi_{\mathbf{x}})_{\mathbf{x} \in \mathbf{R}_+^n}$ *where each* $\pi_{\mathbf{x}} \in \Pi_{\mathbf{x}}$. *If the function* $W(\mathbf{x}) = J(\pi_{\mathbf{x}}; \mathbf{x})$ *is a viscosity supersolution of the HJB equation* (3.1) *in the interior of* $\mathbf{R}_+^n$, *then W is the optimal value function. Also, if for each $k \geq 1$ there exists a family of strategies* $(\pi_{\mathbf{x}}^k)_{\mathbf{x} \in \mathbf{R}_+^n}$ *with* $\pi_{\mathbf{x}}^k \in \Pi_{\mathbf{x}}$ *such that* $W(\mathbf{x}) := \lim_{k \to \infty} J(\pi_{\mathbf{x}}^k; \mathbf{x})$ *is a viscosity supersolution of the HJB equation* (3.1) *in the interior of* $\mathbf{R}_+^n$, *then W is the optimal value function.*

Remark 3.9 It is easy to show that the function h_0 defined in (2.15) satisfies that bh_0 are supersolutions of the HJB equation (3.1) for all b large enough.

4 Discrete Scheme

In this section we construct a family of admissible strategies for any point in a suitable grid and then extend it to $\mathbf{R}_+^n$. We will show in the next section that the value function of these strategies converge to the optimal value function V as the mesh size goes to zero.

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

$$\mathcal{G}^\delta := \left\{ (m_1 \delta p_1, \ldots, m_n \delta p_n) : \mathbf{m} \in \mathbf{N}_0^n \right\}.$$

The idea of the numerical scheme is to find, at each point of the grid $\mathcal{G}^\delta$, the best local strategy among the ones suggested by the operators of the HJB equation (3.1); these possible local strategies are: none of the companies pay dividends, one of the companies pays a lump sum as dividends, or the manager of the company opts to switch immediately. We modify these local strategies in such a way that the controlled surplus lies in the grid after the arrival of a jump of the uncontrolled process. In order to do that, let us introduce the functions $g^\delta : \mathbf{N}_0^n \to \mathbf{R}_+^n$ which relates the indices with the corresponding points of the grid and $\rho^\delta : \mathbf{R}_+^n \to \mathbf{N}_0^n$ which assigns to each point points $\mathbf{x}$ in $\mathbf{R}_+^n$ the index of the closest point of the grid below $\mathbf{x}$. More precisely,

$$g^\delta(\mathbf{m}) = (m_1 \delta p_1, \ldots, m_n \delta p_n) \text{ and } \rho^\delta(\mathbf{x}) := \max\{\mathbf{m} \in \mathbf{N}_0^n : g^\delta(\mathbf{m}) \leq \mathbf{x}\};$$

we can also write

$$\rho^\delta(\mathbf{x}) = \left(\left[\frac{x_1}{\delta p_1} \right], \ldots, \left[\frac{x_n}{\delta p_n} \right] \right) \in \mathbf{N}_0^n$$

where [.] means the integer part in each coordinate. Note that ρ^δ is the left-inverse function of g^δ and that

$$\langle \mathbf{x} \rangle^\delta := g^\delta(\rho^\delta(\mathbf{x})) = \max\{\mathbf{y} \in \mathcal{G}^\delta : \mathbf{y} \leq \mathbf{x}\}.$$

Given any current surplus $g^\delta(\mathbf{m}) \in \mathcal{G}^\delta$, let τ and $\mathbf{U}$ be the arrival time and the size of the next jump of the uncontrolled process. We first define the $n + 2$ possible control actions at any point of the grid $\mathcal{G}^\delta$ as follows.

- Control action $\mathbf{E}_0$: Pay no dividends up to the time $\delta \wedge \tau$. In the case that $\delta < \tau$, the uncontrolled surplus at time δ is $g^\delta(\mathbf{m} + \mathbf{1}) \in \mathcal{G}^\delta$; and if $\delta \geq \tau$, the uncontrolled surplus at time τ is

$$g^\delta(\mathbf{m}) + \tau \mathbf{p} - \mathbf{U}.$$

  If this vector is in $\mathbf{R}_+^n$, the companies pay immediately the minimum amount of dividends in such a way that the controlled surplus lies in a point of the grid; this end surplus can be written as $g^\delta(\mathbf{k})$, where

$$\mathbf{k} = \rho^\delta(g^\delta(\mathbf{m}) + \tau \mathbf{p} - \mathbf{U}).$$

  The amount paid as dividends is equal to

$$g^\delta(\mathbf{m} - \mathbf{k}) + \tau \mathbf{p} - \mathbf{U}.$$

  In the case that the surplus $g^\delta(\mathbf{m}) + \tau \mathbf{p} - \mathbf{U} \notin \mathbf{R}_+^n$ at time $\tau \leq \delta$, the process stops.
- Control actions $\mathbf{E}_i$ with $i = 1, \ldots, n$: The i-th company pays immediately $p_i \delta$ as dividends, so the controlled surplus becomes $g^\delta(\mathbf{m} - \mathbf{e}_i) \in \mathcal{G}^\delta$. The control action $\mathbf{E}_i$ can only be applied for current surplus $g^\delta(\mathbf{m}) \in \mathcal{G}^\delta$ if $m_i > 0$.
- Control action $\mathbf{E}_s$: The manager opts to switch immediately and the process stops.

We denote the space of controls as

$$\mathcal{E} = \{\mathbf{E}_s, (\mathbf{E}_i)_{i=1,\ldots,n}, \mathbf{E}_0\}.$$

Consider $\Pi^\delta_{g^\delta(\mathbf{m})} \subset \Pi_{g^\delta(\mathbf{m})}$ as the set of all the admissible strategies with initial surplus $g^\delta(\mathbf{m}) \in \mathcal{G}^\delta$ which can be obtained by a sequence of control actions in $\mathcal{E}$ at each point of the grid. Let us describe the strategies $\pi = (\mathbf{L}, \overline{\tau}) \in \Pi^\delta_{g^\delta(\mathbf{m})}$; we take, for any $\omega = (\tau_j, \mathbf{U}_j)_{j \geq 1} \in \Omega$, a sequence $\mathbf{s} = (s_k)_{k=1,\ldots,\tilde{k}}$ with $s_k \in \mathcal{E}$ and $1 \leq \tilde{k} \leq \infty$, the first control action s_1 is applied at the point $g^\delta(\mathbf{m}) \in \mathcal{G}^\delta$, the second control action s_2 is applied at the end surplus in $\mathcal{G}^\delta$ resulting from the control action s_1, and so on. If the length of the sequence $\mathbf{s}$ is $\tilde{k} < \infty$, then $s_{\tilde{k}}$ should be either $\mathbf{E}_s$ or $\mathbf{E}_0$. In the last case, the end surplus resulting from the final control action $s_{\tilde{k}}$ is outside $\mathbf{R}_+^n$ due to the arrival of a jump.

Take $\mathbf{m}^k \in \mathbf{N}_0^n$ in such a way that $g^\delta(\mathbf{m}^k)$ is the point of $\mathcal{G}^\delta$ in which the control action s_k is applied; let t_k be the time at which the control action s_k is chosen; let Δ_k be the time elapsed for the control action s_k and let $\mathbf{y}^k \in \mathcal{G}^\delta \cup (\mathbf{R}_+^n)^c$ be the end surplus resulting from the control action s_k.

Remark 4.1 Let us describe in a precise way the values of $(\mathbf{m}^k, \Delta_k, t_k, \mathbf{y}^k)_{k=1,\ldots,\tilde{k}}$.

- In the case that $s_k = \mathbf{E}_i$, then $k < \tilde{k}$, $\Delta_k = 0$, $t_{k+1} = t_k$, $\mathbf{m}^{k+1} = \mathbf{m}^k - \mathbf{e}_i$ and $\mathbf{y}^k = g^\delta(\mathbf{m}^{k+1})$.
- In the case that $s_k = \mathbf{E}_s$, then

$$k = \tilde{k}, t_k = \overline{\tau}, \Delta_k = 0 \quad \text{and} \quad \mathbf{y}^k = g^\delta(\mathbf{m}^k).$$

- In the case that $s_k = \mathbf{E}_0$, take $j_k := \min\{j : \tau_j > t_k\}$ (so τ_{j_k} is the arrival time of the first jump after t_k); there are three possibilities:

  (a) If $\tau_{j_k} > t_k + \delta$, then

$$k < \tilde{k}, \quad \Delta_k = \delta, \quad t_{k+1} = t_k + \delta, \quad \mathbf{m}^{k+1} = \mathbf{m}^k + \mathbf{1} \quad \text{and } \mathbf{y}^k = g^\delta\left(\mathbf{m}^{k+1}\right).$$

  (b) If $\tau_{j_k} \leq t_k + \delta$ and $g^\delta(\mathbf{m}) + \left(\tau_{j_k} - t_k\right)\mathbf{p} - \mathbf{U}_j \in \mathbf{R}_+^n$, then

$$k < \tilde{k}, \quad \Delta_k = \tau_{j_k} - t_k, \quad t_{k+1} = \tau_{j_k}, \quad \mathbf{m}^{k+1}$$
$$= \rho^\delta(g^\delta(\mathbf{m}) + \left(\tau_{j_k} - t_k\right)\mathbf{p} - \mathbf{U}_j) \text{ and } \mathbf{y}^k = g^\delta\left(\mathbf{m}^{k+1}\right).$$

  (c) If $\tau_{j_k} \leq t_k + \delta$ and $\mathbf{y}^k = g^\delta(\mathbf{m}) + \left(\tau_{j_k} - t_k\right)\mathbf{p} - \mathbf{U}_j \notin \mathbf{R}_+^n$, then

$$k = \tilde{k}, \quad \Delta_k = \tau_{j_k} - t_k \quad \text{and} \quad t_k + \Delta_k = \tau_{j_k} = \tau^{\mathbf{L}}.$$

Defining $\Delta\mathbf{L}_k$ as the amount of dividends paid by the control action s_k, we have

$$\Delta\mathbf{L}_k = \begin{cases} p_i \delta \mathbf{e}_i & \text{if } s_k = \mathbf{E}_i \\ \mathbf{c}_k - \langle \mathbf{c}_k \rangle^\delta & \text{if } s_k = \mathbf{E}_0, \tau_{j_k} \in (t_k, t_k + \delta] \text{ and } \mathbf{c}_k \in \mathbf{R}_+^n \\ 0 & \text{otherwise,} \end{cases}$$

where j_k is defined in the previous remark and

$$\mathbf{c}_k = g^\delta(\mathbf{m}^k) + (\tau_{j_k} - t_k)\mathbf{p} - \mathbf{U}_j.$$

Therefore, if the strategy $\pi = (\mathbf{L}, \overline{\tau}) \in \Pi_{g^\delta(\mathbf{m})}^\delta$ then the cumulative dividend payment strategy is

$$\mathbf{L}_t = \sum_{k \leq \tilde{k}, t_k \leq t} \Delta\mathbf{L}_k,$$

and the switch time $\overline{\tau}$ is the time at which the control action $\mathbf{E}_s$ is chosen. By construction, if $\pi \in \Pi_{g^\delta(\mathbf{m})}^\delta$ then $\mathbf{X}_{t_k}^{\mathbf{L}} \in \mathcal{G}^\delta$ for all $k \leq \tilde{k}$, also the set of times

$$\{t_k : k \leq \tilde{k}\} \subseteq \left\{\tau_i + j\delta : i, j \in \mathbf{N}_0 \text{ and } j \leq \frac{\tau_{i+1} - \tau_i}{\delta}\right\};$$

here $\tau_0 = 0$. For the strategy $(\mathbf{L}, \overline{\tau})$ to be admissible, we need to assume the following condition: If the arrival times and sizes of the claims of two elements in Ω coincide up to time t, then the corresponding sequences of control actions $\mathbf{s} = (s_k)_{k=1,\ldots,\tilde{k}}$ must coincide for all k such that $t_k \leq t$.

The following lemma states that the sequences $(t_k)_{k\geq 1}$ do not have an accumulation point, the proof is in the Appendix.

Lemma 4.2 *Given $\pi \in \Pi^\delta_{g^\delta(\mathbf{m})}$, $\lim_{k\to\infty} t_k = \infty$ a.s. within the subset $\{\tilde{k} = \infty\} \subset \Omega$.*

We define the $\mathcal{G}^\delta$-*optimal function* v^δ as the supremum of the value functions of admissible strategies which are a combination of the control actions in $\mathcal{E}$, that is

$$v^\delta(\mathbf{m}) = \sup_{\pi \in \Pi^\delta_{g^\delta(\mathbf{m})}} J(\pi; g^\delta(\mathbf{m})). \tag{4.1}$$

Remark 4.3 We will show in Proposition 4.5 that the supremum in (4.1) is indeed a maximum, so there exists an optimal strategy in the class $\Pi^\delta_{g^\delta(\mathbf{m})}$.

4.1 Characterization of the $\mathcal{G}^\delta$-Optimal Function

In this subsection, we show that the $\mathcal{G}^\delta$-optimal function $v^\delta : \mathbf{N}_0^n \to \mathbf{R}$ is a solution of a discrete version of the HJB equation (3.1). We also see that v^δ can be characterized as the smallest supersolution of this discrete HJB equation. Moreover, we prove that there exists an optimal admissible strategy for the problem (4.1). This strategy, called the $\mathcal{G}^\delta$-*optimal strategy*, is stationary in the sense that the control actions depends only on the point of the grid at which the current surplus lies.

We now introduce some operators related to the control actions in $\mathcal{E}$, these operators will be involved in the discrete version of the HJB equation. Given any family of admissible strategies $\widetilde{\pi} = (\pi_{g^\delta(\mathbf{m})})_{\mathbf{m}\in\mathbf{N}_0^n}$ with $\pi_{g^\delta(\mathbf{m})} \in \Pi^\delta_{g^\delta(\mathbf{m})}$, we define the value function $\widetilde{w} : \mathbf{N}_0^n \to \mathbf{R}$ of $\widetilde{\pi}$ as

$$\widetilde{w}(\mathbf{m}) := J(\pi_{g^\delta(\mathbf{m})}; g^\delta(\mathbf{m})).$$

Let us consider the admissible strategies with initial surplus $g^\delta(\mathbf{m}) \in \mathcal{G}^\delta$ which consists of applying first one of the control actions in $\mathcal{E}$, and afterwards applying the strategy in the family $\widetilde{\pi}$ corresponding to the end surplus (if it is possible); the value functions of these strategies are given by

$$T_0(\widetilde{w})(\mathbf{m}) := \widetilde{w}(\mathbf{m}+\mathbf{1})e^{-(c+\lambda)\delta} + \mathcal{I}^\delta(\widetilde{w})(\mathbf{m}) - \int_0^\delta e^{-(c+\lambda)t}\mathcal{R}(g^\delta(\mathbf{m}) + t\mathbf{p})dt \,, \tag{4.2}$$

$$T_i(\widetilde{w})(\mathbf{m}) := \widetilde{w}(\mathbf{m}-\mathbf{e}_i) + \delta a_i \, p_i \quad \text{and} \quad T_s(\widetilde{w})(\mathbf{m}) := f(g^\delta(\mathbf{m})), \tag{4.3}$$

depending on which control action in $\mathcal{E}$ is chosen. Here,

$$
\begin{aligned}
&\mathcal{I}^\delta(w)(\mathbf{m}) \\
&:= \int_0^\delta \left(\int_0^{g^\delta(\mathbf{m})+t\mathbf{p}} \lambda e^{-(c+\lambda)t} w\left(\rho^\delta\left(g^\delta(\mathbf{m}) + t\mathbf{p} - \boldsymbol{\alpha}\right)\right) dF(\boldsymbol{\alpha}) \right) dt \\
&+ \int_0^\delta \left(\int_0^{g^\delta(\mathbf{m})+t\mathbf{p}} \lambda e^{-(c+\lambda)t} \mathbf{a} \cdot \left(g^\delta(\mathbf{m}) + t\mathbf{p} - \boldsymbol{\alpha} - \left\langle g^\delta(\mathbf{m}) + t\mathbf{p} - \boldsymbol{\alpha}\right\rangle^\delta\right) dF(\boldsymbol{\alpha}) \right) dt.
\end{aligned}
$$
(4.4)

We can consider T_0, T_i and T_s as operators in the set of functions $\left\{w : \mathbf{N}_0^n \to \mathbf{R}\right\}$; we also define the operator T as

$$
T := \max\{T_0, (T_i)_{i=1,\dots,n}, T_s\}.
$$
(4.5)

The following lemma is technical and the proof is in the Appendix.

Lemma 4.4 *The operators T_0, T_i, T_s and T are non-decreasing and T satisfies,*

$$
\sup_{\mathbf{m}\in\mathbf{N}_0^n} |T(w_1)(\mathbf{m}) - T(w_2)(\mathbf{m})| \le \sup_{\mathbf{m}\in\mathbf{N}_0^n} |w_1(\mathbf{m}) - w_2(\mathbf{m})|.
$$

Moreover, $T_0(w)$, $T_i(w)$ and $T_s(w)$ can be written as a linear combination of the values of $w(\mathbf{m})$ with $\mathbf{m} \in \mathbf{N}_0^n$ plus a constant.

We define the *discrete HJB equation* as

$$
(T(w) - w)(\mathbf{m}) = \max\{T_0(w) - w, (T_i(w) - w)_{i=1,\dots,n}, T_s(w) - w\}(\mathbf{m}) = 0
$$
(4.6)

for $\mathbf{m} \in \mathbf{N}_0^n$. Analogously to Definition 3.1, we say that a function $\overline{w} : \mathbf{N}_0^n \to \mathbf{R}$ is a *supersolution* of (4.6) if $T(\overline{w}) - \overline{w} \le 0$, and a function $\underline{w} : \mathbf{N}_0^n \to \mathbf{R}$ is a *subsolution* of (4.6) if $T(\underline{w}) - \underline{w} \ge 0$.

The following results are the discrete versions of Propositions 3.4, Lemma 3.6, Theorems 3.7 and 3.8. The discrete version of the growth condition (2.14) is given by

$$
w(\mathbf{m})e^{\frac{-c}{2n}\sum_{i=1}^n m_i \delta} \text{ is upper bounded in } \mathbf{N}_0^n.
$$
(4.7)

Proposition 4.5 *The function $v^\delta : \mathbf{N}_0^n \to \mathbf{R}$ is well defined and it is a solution of (4.6). Moreover, given an initial surplus $g^\delta(\mathbf{m}_0) \in \mathcal{G}^\delta$, there exists a $\mathcal{G}^\delta$-optimal strategy $\pi^\delta_{g^\delta(\mathbf{m}_0)} \in \Pi^\delta_{g^\delta(\mathbf{m}_0)}$ such that*

$$
v^\delta(\mathbf{m}_0) = J\left(\pi^\delta_{g^\delta(\mathbf{m}_0)}; g^\delta(\mathbf{m}_0)\right).
$$

This $\mathcal{G}^\delta$-optimal strategy is stationary in the following sense: the control action s_k in the sequence $\mathbf{s} = (s_k)_{k=1,\dots,\tilde{k}}$ depends only on the current surplus $g^\delta(\mathbf{m}^k) \in \mathcal{G}^\delta$.

Proof By definitions (2.6) and (4.1), we have

$$f(g^\delta(\mathbf{m})) \le v^\delta(\mathbf{m}) \le V(g^\delta(\mathbf{m})),$$

so v^δ is well defined.

Let us prove that $v^\delta = T(v^\delta)$. Take a sequence $(p_l)_{l \ge 1}$ of families of strategies $p_l = \left(\pi^l_{g^\delta(\mathbf{m})} \right)_{\mathbf{m} \in \mathbf{N}_0^n}$ with $\pi^l_{g^\delta(\mathbf{m})} \in \Pi^\delta_{g^\delta(\mathbf{m})}$ such that

$$v^\delta(\mathbf{m}) - J\left(\pi^l_{g^\delta(\mathbf{m})}; g^\delta(\mathbf{m}) \right) \le \frac{1}{l}$$

for all $\mathbf{m} \in \mathbf{N}_0^n$. Define $w_l : \mathbf{N}_0^n \to \mathbf{R}$ as $w_l(\mathbf{m}) = J(\pi^l_{g^\delta(\mathbf{m})}; g^\delta(\mathbf{m}))$, by Lemma 4.4, we have that

$$T(v^\delta)(\mathbf{m}) = \lim_{l \to \infty} T(w_l)(\mathbf{m}) \le v^\delta(\mathbf{m}).$$

On the other hand, since $\pi^l_{g^\delta(\mathbf{m})}$ can be obtained by a sequence of control actions $\mathbf{s} = (s_k)_{k=1,\dots,\tilde{k}}$ and at any point $g^\delta(\mathbf{m})$ of the grid all the value functions of strategies in $\Pi^\delta_{g^\delta(\mathbf{m})}$ are below $v^\delta(\mathbf{m})$, we have by definition of T given in (4.5), that $w_l(\mathbf{m}) \le T(v^\delta)(\mathbf{m})$. So taking the limit as $l \to \infty$, we obtain that

$$v^\delta(\mathbf{m}) \le T(v^\delta)(\mathbf{m}).$$

Finally, since $v^\delta = T(v^\delta)$, we can define for any $\mathbf{m} \in \mathbf{N}_0^n$, a control action $S(\mathbf{m}) \in \mathcal{E}$ in the following way:

- If $T_s(v^\delta)(\mathbf{m}) = v^\delta(\mathbf{m})$, take $S(\mathbf{m}) = \mathbf{E}_s$.
- If $T_0(v^\delta)(\mathbf{m}) = v^\delta(\mathbf{m})$, take $S(\mathbf{m}) = \mathbf{E}_0$.
- and if $T_i(v^\delta)(\mathbf{m}) = v^\delta(\mathbf{m})$ for some $i = 1, \dots, n$, take $S(\mathbf{m}) = \mathbf{E}_i$.

Given an initial surplus $g^\delta(\mathbf{m}_0) \in \mathcal{G}^\delta$, the $\mathcal{G}^\delta$-optimal strategy $\pi^\delta_{g^\delta(\mathbf{m}_0)} \in \Pi^\delta_{g^\delta(\mathbf{m}_0)}$ is defined inductively as follows: $s_1 = S(\mathbf{m}_0)$; assuming that $s_1, s_2, \dots, s_{k-1}$ are defined and the process does not stop at step $k-1$, we define $s_k = S(\mathbf{m}_0^k)$ where $g^\delta\left(\mathbf{m}_0^k\right) \in \mathcal{G}^\delta$ is the end surplus of s_{k-1}. $\qquad\square$

Analogously to Remark 3.5, we cannot expect in general to have uniqueness of viscosity solutions of the discrete HJB equation (4.6). For instance, in the two dimensional case with independent companies, the switch function f given in (2.13) and the function υ defined in (2.11), we have that

$$w(\mathbf{m}) := \sum_{i=1}^{n} p_i m_i \delta + k$$

is a solution of (4.6) for k large enough. The following lemma is the discrete version of Lemma 3.6, the proof is in the Appendix.

Lemma 4.6 *Given any* $\pi = (\mathbf{L}, \overline{\tau}) \in \Pi^{\delta}_{g^{\delta}(\mathbf{m})}$ *and any supersolution* $w : \mathbf{N}_0^n \to \mathbf{R}$ *of* (4.6) *with growth condition* (4.7)*, we have that* $J(\pi; g^{\delta}(\mathbf{m})) \leq w(\mathbf{m})$.

From Lemma 4.6, we obtain the following theorems.

Theorem 4.7 *The* $\mathcal{G}^{\delta}$*-optimal value function* $v^{\delta} : \mathbf{N}_0^n \to \mathbf{R}$ *can be characterized as the smallest supersolution of the discrete HJB equation* (4.6) *with growth condition* (4.7)*.*

Theorem 4.8 *If the function* $w : \mathbf{N}_0^n \to \mathbf{R}$ *with growth condition* (4.7) *is a supersolution of* (4.6)*, and also satisfies that for any* $\mathbf{m} \in \mathbf{N}_0^n$*,* $w(\mathbf{m})$ *is either* $J(\pi; g^{\delta}(\mathbf{m}))$ *with* $\pi \in \Pi^{\delta}_{g^{\delta}(\mathbf{m})}$ *or* $\lim_{l \to \infty} J(\pi_l; g^{\delta}(\mathbf{m}))$ *with* $\pi_l \in \Pi^{\delta}_{g^{\delta}(\mathbf{m})}$ *for any* $l \geq 1$*, then* $w = v^{\delta}$*.*

4.2 Construction of the $\mathcal{G}^{\delta}$-Optimal Strategy and the $\mathcal{G}^{\delta}$-Optimal Function

In this subsection we construct recursively the $\mathcal{G}^{\delta}$-optimal strategy and the corresponding $\mathcal{G}^{\delta}$-optimal function.

Since T defined in (4.5) is not a contraction operator, v^{δ} can not be obtained numerically as a fixed point; so we construct value functions v_l^{δ} of strategies in $\Pi^{\delta}_{g^{\delta}(\mathbf{m})}$ which can be calculated explicitly by (4.2), (4.3) and (4.5) such that $v_l^{\delta} \nearrow v^{\delta}$ as $l \to \infty$.

Let us define iteratively the families of strategies $\widetilde{\pi}_l = (\pi^l_{g^{\delta}(\mathbf{m})})_{\mathbf{m} \in \mathbf{N}_0^n}$ for each $l \geq 1$ in the following way:

(1) We start with the family of strategies $\widetilde{\pi}_1 = (\pi^1_{g^{\delta}(\mathbf{m})})_{\mathbf{m} \in \mathbf{N}_0^n}$ where $\pi^1_{g^{\delta}(\mathbf{m})} \in \Pi^{\delta}_{g^{\delta}(\mathbf{m})}$ consists of switching immediately; the value of this strategy is

$$v_1^{\delta}(\mathbf{m}) := f\left(g^{\delta}(\mathbf{m})\right).$$

(2) Given the family of strategies $\widetilde{\pi}_l = (\pi^l_{g^{\delta}(\mathbf{m})})_{\mathbf{m} \in \mathbf{N}_0^n}$ with $\pi^l_{g^{\delta}(\mathbf{m})} \in \Pi^{\delta}_{g^{\delta}(\mathbf{m})}$, we define the family $\widetilde{\pi}_{l+1} = (\pi^{l+1}_{g^{\delta}(\mathbf{m})})_{\mathbf{m} \in \mathbf{N}_0^n}$ as follows: We choose for any $\mathbf{m} \in \mathbf{N}_0^n$, the best strategy $\pi^{l+1}_{g^{\delta}(\mathbf{m})} \in \Pi^{\delta}_{g^{\delta}(\mathbf{m})}$ among the ones which initially follows one of control actions in $\mathcal{E}$ and then continues with the corresponding strategy in the family $\widetilde{\pi}_l$. The value of this new strategy is given by

$$v_{l+1}^{\delta}(\mathbf{m}) := T(v_l^{\delta})(\mathbf{m}) = T^l(v_1^{\delta})(\mathbf{m}) \quad \text{for } \mathbf{m} \in \mathbf{N}_0^n. \tag{4.8}$$

Remark 4.9 v_l^{δ} can be thought as the maximum of the value function of strategies $\pi \in \Pi^{\delta}_{g^{\delta}(\mathbf{m})}$ where the length $\tilde{k}$ of the corresponding sequence $\mathbf{s}$ is upper bounded by l and $s_l = \mathbf{E}_s$ in the case that $\tilde{k} = l$.

In the next proposition we use Theorem 4.8 to see that the limit of v_l^{δ} is indeed v^{δ}.

Proposition 4.10 *We have that* $v_{l+1}^{\delta} \geq v_l^{\delta}$ *for* $l \geq 1$ *and that* $\lim_{l \to \infty} v_l^{\delta} = v^{\delta}$*.*

Proof Take $\mathbf{m} \in \mathbf{N}_0^n$, it is straightforward to see by (4.8) that $v_2^\delta(\mathbf{m}) \geq v_1^\delta(\mathbf{m})$; on the other hand, the operator T is non-decreasing, so we obtain that $v_{l+1}^\delta \geq v_l^\delta$ for $l \geq 1$. Then, there exists $w_0 : \mathbf{N}_0^n \to \mathbf{R}$ such that

$$w_0(\mathbf{m}) := \lim_{l \to \infty} v_l^\delta(\mathbf{m}) \leq V(g^\delta(\mathbf{m})).$$

Note that all the functions v_l^δ are subsolutions (4.6) and that w_0 is a solution of (4.6) because $T(w_0) = w_0$. Since w_0 satisfies the growth condition (4.7), w_0 coincides with the value function v^δ by Theorem 4.8. □

4.3 Definition of the Value Function V^δ

In this subsection we define, using the $\mathcal{G}^\delta$-optimal functions and strategies, a family of admissible strategies for any point in $\mathbf{R}_+^n$ and the corresponding value function V^δ.

Definition 4.11 We use the $\mathcal{G}^\delta$-optimal function $v^\delta : \mathbf{N}_0^n \to \mathbf{R}$ to define a function $V^\delta : \mathcal{G}^\delta \to \mathbf{R}$ as

$$V^\delta(g^\delta(\mathbf{m})) := v^\delta(\mathbf{m})$$

for $\mathbf{m} \in \mathbf{N}_0^n$. Note that $V^\delta(g^\delta(\mathbf{m}))$ is the value of the $\mathcal{G}^\delta$-optimal admissible strategy $\pi_{g^\delta(\mathbf{m})}^\delta \in \Pi_{g^\delta(\mathbf{m})}^\delta$.

We construct now a family of strategies $\widetilde{\pi}^\delta = (\pi_{\mathbf{x}})_{\mathbf{x} \in \mathbf{R}_+^n}$, where $\pi_{\mathbf{x}} \in \Pi_{\mathbf{x}}$, such that the corresponding value function $V^\delta(\mathbf{x}) = J(\pi_{\mathbf{x}}; \mathbf{x})$ extends to $\mathbf{R}_+^n$ the function defined in Definition 4.11. Take the strategy $\pi_{\mathbf{x}} \in \Pi_{\mathbf{x}}$ which pays immediately $\mathbf{x} - \langle \mathbf{x} \rangle^\delta$ as dividends and then follows the $\mathcal{G}^\delta$-optimal strategy $\pi_{\langle \mathbf{x} \rangle^\delta}^\delta \in \Pi_{\langle \mathbf{x} \rangle^\delta}^\delta$. We obtain that $V^\delta : \mathbf{R}_+^n \to \mathbf{R}$ is given by

$$V^\delta(\mathbf{x}) := V^\delta(\langle \mathbf{x} \rangle^\delta) + \mathbf{a} \cdot (\mathbf{x} - \langle \mathbf{x} \rangle^\delta). \tag{4.9}$$

5 Convergence of the Discrete Scheme

In this section we show the locally uniformly convergence of the discrete scheme defined in the previous section by taking a suitable sequence of embedded grids.

In the next technical lemma, we show that the functions v^δ satisfy a δ-locally Lipschitz condition and a relation between $v^{2\delta}$ and v^δ which gives a monotonicity condition on the embedded grids; the proof is in the Appendix.

Lemma 5.1 *The functions v^δ defined in (4.1) satisfy:*
(1) $v^\delta(\mathbf{m} + \mathbf{e}_i) - v^\delta(\mathbf{m}) \geq a_i p_i \delta$ *and* $v^\delta(\mathbf{m} + \mathbf{1}) - v^\delta(\mathbf{m}) \leq v^\delta(\mathbf{m})(e^{(c+\lambda)\delta} - 1)$;
(2) $\Pi_{g^{2\delta}(\mathbf{m})}^{2\delta} \subset \Pi_{2g^\delta(\mathbf{m})}^\delta \subset \Pi_{2g^\delta(\mathbf{m})}$ *and so* $v^{2\delta}(\mathbf{m}) \leq v^\delta(2\mathbf{m})$.

Let us take $\delta_k := \delta/2^k$ for $k \geq 0$. In the remainder of the section we will prove 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 by Lemma 5.1–(2),

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

then we can define the function $\overline{V} : \mathbf{R}_+^n \to \mathbf{R}$ as

$$\overline{V}(\mathbf{x}) := \lim_{k\to\infty} V^{\delta_k}(\mathbf{x}). \tag{5.1}$$

Remark 5.2 We will prove that $\overline{V}$ is the optimal value function. In order to do that, we will show that $\overline{V}$ is a viscosity supersolution of (3.1). It is straightforward to see that $\overline{V}(\mathbf{x})$ is a limit of value functions of admissible strategies in $\Pi_{\mathbf{x}}$ for all $\mathbf{x} \in \mathbf{R}_+^n$ so the result will follow from Theorem 3.8. Since there is no uniqueness of 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^{δ_k} and we show that $\overline{V}$ is locally Lipschitz in $\mathbf{R}_+^n$ and so it is absolutely continuous; the proof is in the Appendix.

Lemma 5.3 *We have for each* $\mathbf{y} \geq \mathbf{x}$ *in* $\mathbf{R}_+^n$ *that*

$$\left| V^{\delta_k}(\mathbf{y}) - V^{\delta_k}(\mathbf{x}) \right| \leq \left\| \langle \mathbf{y} \rangle^{\delta_k} - \langle \mathbf{x} \rangle^{\delta_k} \right\|_1 \frac{2}{\hat{p}} V^{\delta_k}\left(\langle \mathbf{x} \vee \mathbf{y} \rangle^{\delta_k} \right) \left(\frac{e^{(c+\lambda)\delta_k} - 1}{\delta_k} \right) + 2\delta_k \mathbf{a} \cdot \mathbf{p},$$

and also

$$\mathbf{a} \cdot (\mathbf{y} - \mathbf{x}) \leq \overline{V}(\mathbf{y}) - \overline{V}(\mathbf{x}) \leq \overline{V}(\mathbf{y}) \frac{2(c+\lambda)}{\hat{p}} \|\mathbf{y} - \mathbf{x}\|_1,$$

where $\hat{p} := \min_{i=1,\dots,n} p_i$.

In the next two propositions we address the convergence of V^{δ_k} to $\overline{V}$ and we prove that $\overline{V}$ coincides with V.

Proposition 5.4 *For any* $\delta > 0$, $V^{\delta_k} \nearrow \overline{V}$ *locally uniformly as* k *goes to infinity.*

Proof Consider a compact set K in $\mathbf{R}_+^n$, $\mathbf{x}^1 \in K$ and $\varepsilon > 0$. Let us take an upper bound $\mathbf{z} \in \mathbf{R}_+^n$ of K. We show first that there exists k_0 large enough and $\eta > 0$ small enough such that if $\left\| \mathbf{x} - \mathbf{x}^1 \right\|_1 < \eta$ and $k \geq k_1$, then

$$\overline{V}(\mathbf{x}) - V^{\delta_k}(\mathbf{x}) < \varepsilon. \tag{5.2}$$

Indeed, by pointwise convergence at $\mathbf{x}^1$, there exists k_1 such that

$$\overline{V}(\mathbf{x}^1) - V^{\delta_k}(\mathbf{x}^1) < \varepsilon/3 \text{ for } k \geq k_1. \tag{5.3}$$

By Lemma 5.3, there exists η_1 such that if $\left\| \mathbf{x} - \mathbf{x}^1 \right\|_1 < \eta_1$, then

$$\left| \overline{V}(\mathbf{x}) - \overline{V}(\mathbf{x}^1) \right| < \varepsilon/3. \tag{5.4}$$

Also, from Lemma 5.3, there exists η_2 and k_2 such that if $\left\| \mathbf{x} - \mathbf{x}^1 \right\|_1 < \eta_1$, then

$$\left| V^{\delta_k}(\mathbf{x}) - V^{\delta_k}(\mathbf{x}^1) \right| \leq \left\| g^{\delta_k} \left(\rho^{\delta_k}(\mathbf{x}) - \rho^{\delta_k}(\mathbf{x}^1) \right) \right\|_1 \overline{V}(\mathbf{z}) 2 e^{(c+\lambda)} / \hat{p} + 2\delta_k \mathbf{a} \cdot \mathbf{p} < \varepsilon/3 \tag{5.5}$$

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 (5.2) from (5.3), (5.4) and (5.5).

Finally, we conclude the result taking a finite covering of the compact set K. □

Proposition 5.5 *The function $\overline{V}$ defined in* (5.1) *is the optimal value function V.*

Proof By Remark 5.2, it is enough to prove that $\overline{V}$ is a viscosity supersolution of (3.1) in the interior of $\mathbf{R}^n_+$. Take $\mathbf{x}^0$ in the interior of $\mathbf{R}^n_+$ and a differentiable test function $\varphi : \mathbf{R}^n_+ \to \mathbf{R}$ for viscosity supersolution of (3.1) at $\mathbf{x}^0$, that is

$$\overline{V}(\mathbf{x}) \geq \varphi(\mathbf{x}) \text{ and } \overline{V}(\mathbf{x}^0) = \varphi(\mathbf{x}^0). \tag{5.6}$$

Since $\mathcal{G}$ is a dense set in $\mathbf{R}^n_+$, we obtain by the continuity assumptions on the function f given in Sect. 2 and (5.1) that $f \leq \overline{V}$ in $\mathbf{R}^n_+$, so $f(\mathbf{x}^0) - \varphi(\mathbf{x}^0) \leq 0$. By Proposition 5.3,

$$\overline{V}(\mathbf{y}) - \overline{V}(\mathbf{x}) \geq \mathbf{a} \cdot (\mathbf{y} - \mathbf{x})$$

for all $\mathbf{y} \geq \mathbf{x}$, so it holds that $\mathbf{a} - \nabla \varphi(\mathbf{x}^0) \leq \mathbf{0}$. In order to prove that $\mathcal{L}(\varphi)(\mathbf{x}^0) \leq 0$, consider now for $\eta > 0$ small enough,

$$\varphi_\eta(\mathbf{x}) = \varphi(\mathbf{x}) - \eta \left(\mathbf{x} - \mathbf{x}^0 \right) \cdot (\mathbf{x} - \mathbf{x}^0).$$

Given $k \geq 0$, the set $\mathcal{G}^{\delta_k} \cap [\mathbf{0}, \mathbf{x}^0 + \mathbf{1}]$ is finite, so we can define

$$a^\eta_k := \min_{\mathcal{G}^{\delta_k} \cap [\mathbf{0}, \mathbf{x}^0 + \mathbf{1}]} \{ V^{\delta_k}(\mathbf{x}) - \varphi_\eta(\mathbf{x}) \}. \tag{5.7}$$

Since $V^{\delta_k} \leq \overline{V}$, we have from (5.6), that $a^\eta_k \leq 0$. Taking

$$0 \leq b_k := \max_{\mathcal{G}^{\delta_k} \cap [\mathbf{0}, \mathbf{x}^0 + \mathbf{1}]} \left(\overline{V} - V^{\delta_k} \right),$$

by Proposition 5.4, $b_k \to 0$ as $k \to \infty$. For all $\mathbf{x} \in \mathcal{G}^{\delta_k} \cap [\mathbf{0}, \mathbf{x}^0 + \mathbf{1}]$ we get from (5.6),

$$\begin{aligned} V^{\delta_k}(\mathbf{x}) - \varphi_\eta(\mathbf{x}) &= V^{\delta_k}(\mathbf{x}) - \overline{V}(\mathbf{x}) + \overline{V}(\mathbf{x}) - \varphi(\mathbf{x}) + \eta \left(\mathbf{x} - \mathbf{x}^0 \right) \cdot (\mathbf{x} - \mathbf{x}^0) \\ &\geq -b_k + \eta \left(\mathbf{x} - \mathbf{x}^0 \right) \cdot (\mathbf{x} - \mathbf{x}^0). \end{aligned}$$

Then, the minimum argument in (5.7) is attained at $\mathbf{x}^k \in \mathcal{G}^{\delta_k}$ such that

$$\left(\mathbf{x}^k - \mathbf{x}^0\right) \cdot (\mathbf{x}^k - \mathbf{x}^0) \leq b_k/\eta.$$

Then, we have $\mathbf{x}^k \to \mathbf{x}^0$ and $a_k^\eta \to 0$ as k goes to infinity. So

$$V^{\delta_k}(\mathbf{x}) \geq \varphi_\eta(\mathbf{x}) + a_k^\eta \text{ for } \mathbf{x} \in \mathcal{G}^{\delta_k} \cap [0, \mathbf{x}^0 + \mathbf{1}] \text{ and } V^{\delta_k}(\mathbf{x}^k) = \varphi_\eta(\mathbf{x}^k) + a_k^\eta.$$

Since

$$T_0(v^{\delta_k})\left(\left[\frac{x_1^k}{\delta_k p_1}\right], \ldots, \left[\frac{x_n^k}{\delta_k p_n}\right]\right) - v^{\delta_k}\left(\left[\frac{x_1^k}{\delta_k p_1}\right], \ldots, \left[\frac{x_n^k}{\delta_k p_n}\right]\right) \leq 0,$$

we obtain

$$\begin{aligned}
0 &\geq e^{-(c+\lambda)\delta_k}\left(V^{\delta_k}(\mathbf{x}^k + \delta_k\mathbf{p})\right) \\
&\quad + \int_0^{\delta_k} \lambda e^{-(c+\lambda)t}\left(\int_{0\leq\boldsymbol{\alpha}\leq\mathbf{x}^k+t\mathbf{p}} V^{\delta_k}\left(\mathbf{x}^k + t\mathbf{p} - \boldsymbol{\alpha}\right) dF(\boldsymbol{\alpha})\right) dt \\
&\quad - \int_0^{\delta_k} e^{-(c+\lambda)t}\mathcal{R}(\mathbf{x}^k + t\mathbf{p})dt - V^{\delta_k}(\mathbf{x}^k) \\
&\geq e^{-(c+\lambda)\delta_k}\left(\varphi_\eta(\mathbf{x}^k + \delta_k\mathbf{p}) - \varphi_\eta(\mathbf{x}^k)\right) \\
&\quad - \left(\varphi_\eta(\mathbf{x}^k) + a_k^\eta\right)\left(1 - e^{-(c+\lambda)\delta_k}\right) \\
&\quad + \int_0^{\delta_k} \lambda e^{-(c+\lambda)t}\left(\int_{0\leq\boldsymbol{\alpha}\leq\mathbf{x}^k+t\mathbf{p}}\left(\varphi_\eta(\rho^{\delta_k}\left(\mathbf{x}^k + t\mathbf{p} - \boldsymbol{\alpha}\right) + a_k^\eta\right) dF(\boldsymbol{\alpha})\right) dt \\
&\quad + \int_0^{\delta_k} \lambda e^{-(c+\lambda)t}(\int_{0\leq\boldsymbol{\alpha}\leq\mathbf{x}^k+t\mathbf{p}}\left(\mathbf{a} \cdot \left(\mathbf{x}^k + t\mathbf{p} - \boldsymbol{\alpha} - \langle\mathbf{x}^k + t\mathbf{p} - \boldsymbol{\alpha}\rangle^{\delta_k}\right)\right) dF(\boldsymbol{\alpha}))dt \\
&\quad - \int_0^{\delta_k} e^{-(c+\lambda)t}\mathcal{R}(\mathbf{x}^k + t\mathbf{p})dt.
\end{aligned}$$

Dividing by δ_k, taking k to infinity and using the continuity of $\mathcal{R}$, we get $\mathcal{L}(\varphi_\eta)(\mathbf{x}^0) \leq 0$. Finally, since $\nabla\varphi_\eta(\mathbf{x}^0) = \nabla\varphi(\mathbf{x}^0)$ and $\varphi_\eta \nearrow \varphi$ as $\eta \searrow 0$, we obtain that $\mathcal{L}(\varphi)(\mathbf{x}^0) \leq 0$ and the result follows. $\square$

From Propositions 5.4 and 5.5, we conclude the main result of the paper.

Theorem 5.6 *For any $\delta > 0$, the functions $V^{\delta_k} \nearrow \overline{V} = V$ locally uniformly as k goes to infinity.*

6 Optimal Merger Time

Let us assume that the uncontrolled bivariate surplus $\mathbf{X}_t$ of two insurance companies with the same shareholders follows the process (2.1). Both branches pay dividends up to the time of their respective ruin τ^{L_i} with $i = 1, 2$, but the shareholders have the possibility of *merging* the two branches at any time $\overline{\tau}$ prior to $\tau^{\mathbf{L}} = \tau^{L_1} \wedge \tau^{L_2}$ (as defined in (2.2)); at this time the branches put together all their surplus, pay the claims of both branches and pay dividends until the joined surplus becomes negative, see e.g. Gerber and Shiu [17]. The aim is to find both the dividend payment policy and the merging time which maximize the expected sum of all the discounted dividends paid
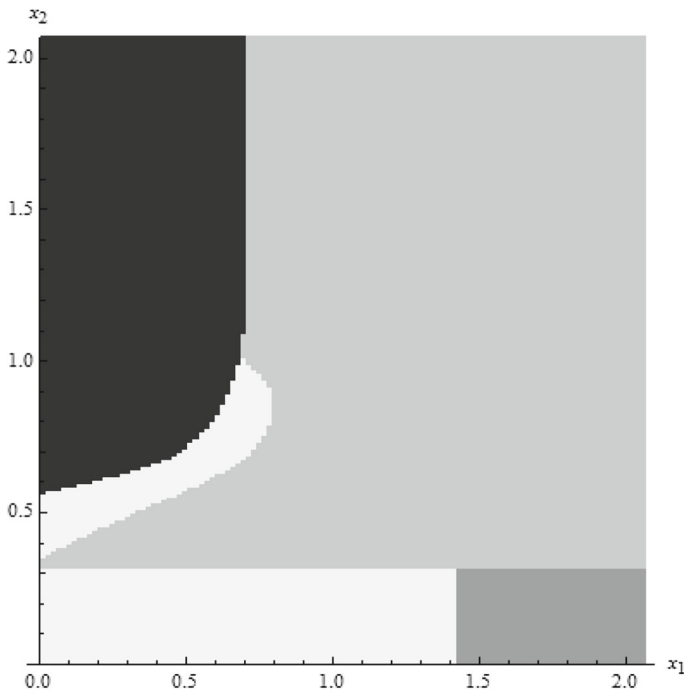
Fig. 1 .

to the shareholders. This problem corresponds to (2.6) where $n = 2$, $\mathbf{a} = (1, 1)$, A is the 2×2 identity matrix, the function υ is defined as in (2.11) and the the switch-value function f is defined as in (2.13). In the numerical examples, we consider

$$F(x_1, x_2) = \mathbb{P}(\alpha_1 \leq x_1, \alpha_2 \leq x_2) = \frac{\lambda_1}{\lambda}(1 - e^{-d_1 x_1}) + \frac{\lambda_2}{\lambda}(1 - e^{-d_2 x_2})$$

with $d_1 = 3$ and $d_2 = 3.5$. Note that the above formula for F corresponds to the case in which the surplus processes of the two branches are independent, as we pointed out in (2.9); so the function

$$\mathcal{R}(x_1, x_2) = \frac{\lambda_1}{\lambda} V_2(x_2) e^{-d_1 x_1} + \frac{\lambda_2}{\lambda} V_1(x_1) e^{-d_2 x_2}$$

is continuous in $\mathbf{R}_+^2$. The parameters of the merger company (that is a one dimension problem) are $\lambda_M = \lambda_1 + \lambda_2$, $p_M = p_1 + p_2$ and $F_M(x) = F(x, x)$.

In the first example, we consider $\lambda_1 = 2.4$, $\lambda_2 = 2$, $\lambda = \lambda_1 + \lambda_2$, $p_1 = 1.08$, $p_2 = 0.674$, $c = 0.11$, $\delta = 1/60$ and $c_M = 0$. In Fig. 1, we show the $\mathcal{G}^\delta$-optimal strategy: the merger region is in black, the non-action region in white, the dividend payment region for the first company in dark grey and the dividend payment region for the second company in light grey. Note that the non-action region has two connected components; in the one on the top, the optimal strategy is to withhold dividend payments in order

to reach the merger region, and in the white rectangle on the bottom the optimal strategy corresponds to the non-action region of the stand-alone problem (in which the companies never merge). This figure suggests that, as $\delta \to 0$, the optimal local control in the boundary between the non-action rectangle and the dividend payment region for the second company region (light grey), should be that the second company pay the incoming premium as dividends while the first company pays no dividends, so the bivariate control surplus stays on the top boundary $x_2 = 0.33$ of the rectangle and moves rightward at constant speed p_1 to the point $(1.42, 0.33)$, which corresponds to the right-top corner of the rectangle (until the arrival of the next claim). Analogously, the optimal strategy in the right boundary $x_1 = 1.42$ of the non action rectangle should be that the first company pay the incoming premium as dividends while the second company pay no dividends, in this case the bivariate control surplus stays on the right boundary of the rectangle and moves upward at constant speed p_2 to the righ-top corner (until the arrival of the next claim). At this corner, both companies pay their incoming premium as dividends and the surplus process remains constant (until the arrival of the next claim). It is more difficult to guess the optimal local control (as $\delta \to 0$) in the boundary between the upper connected component of the non-action region and the dividend payment region for the second company region (light grey). Our conjecture, assuming some regularity on this boundary, is the following: In the upper part of this boundary (up to the furthest point to the right), the second company should pay dividends with some rate in such a way that the bivariate control surplus stays in this part of the boundary (moving downwards), and in the lower part of this boundary, the second company should pay a lump sum in such a way that the bivariant surplus reaches the line $x_2 = 0.33$.

In the second example, we consider $\lambda_1 = 2.44$, $\lambda_2 = 2.22$, $\lambda = \lambda_1 + \lambda_2$, $p_1 = 1.100$, $p_2 = 0.825$, $c = 0.1$, $\delta = 1/50$ and $c_M = 0.364$. In Fig. 2, we show the $\mathcal{G}^\delta$-optimal strategy; the regions are described with the same colors as before. This figure suggests that, as $\delta \to 0$, the optimal local control in the boundary between the non-action region (white) and the dividend payment region for the second company (light grey region), would be -assuming some regularity on the boundary- that the second company pay dividends with some rate in such a way that, in the absence of a claim, the bivariate control surplus stays in the boundary. This control surplus would move downward (because the second company is paying more dividends than the incoming premium rate) and rightward (because the first company is receiving the incoming premium rate and not paying dividends). Eventually, in the absence of claims, the bivariate surplus reaches the point $(1.61, 1.06)$ at which the light grey, the dark grey and the white regions meet. At this point, both companies should pay the incoming premiums as dividends and the bivariate surplus process remains constant until the arrival of the next claim. Similarly, the optimal local control in the boundary between the non-action region (white) and the dividend payment region for the first company (dark grey region), would be -assuming some regularity on the boundary-that the first company pay dividends with some rate in such a way that, in the absence of a claim, the bivariate control surplus stays in the boundary. This control surplus would move leftward (because the first company is paying more dividends than the incoming premium rate) and upward (because the second company is receiving the incoming
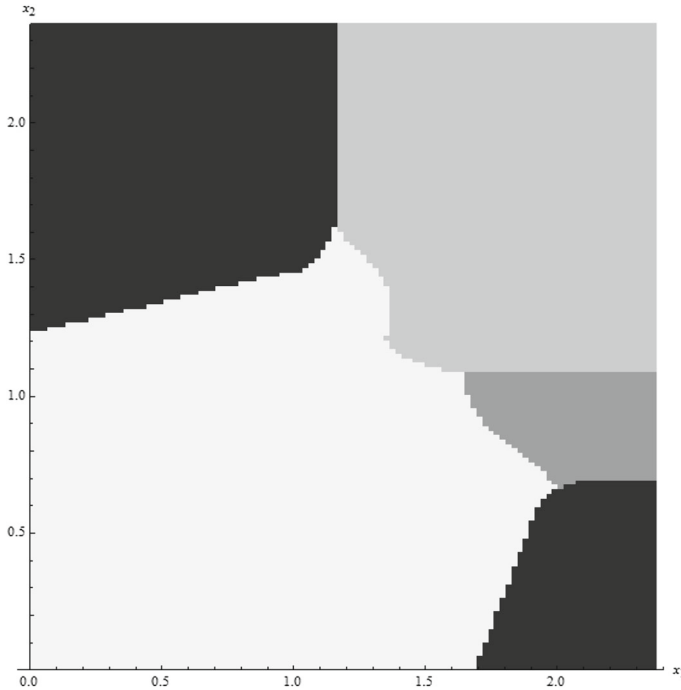
Fig. 2 .

premium rate and not paying dividends). Eventually, in the absence of claims, the bivariate surplus reaches the point (1.61, 1.06).

7 Appendix

This section contains the proofs of all the lemmas.

Proof of Lemma 3.3 Let us extend the function g to $\mathbf{R}^n$ as $g(\mathbf{x}) = 0$ for $\mathbf{x} \notin \mathbf{R}^n_+$ and the function υ to $\mathbf{R}^n \times \mathbf{R}^n_+$ as $\upsilon(\mathbf{x}, \boldsymbol{\alpha}) = 0$ for $\left(\mathbf{R}^n \times \mathbf{R}^n_+\right) \big/ B$, where B is defined in (2.4). Using the expressions (2.1) and the change of variables formula for finite variation processes, and calling $\mathbf{z}_s = \mathbf{X}^{\mathbf{L}}_s$ and $\check{\mathbf{z}}_s = \check{\mathbf{X}}^{\mathbf{L}}_s$, we can write

$$
\begin{aligned}
g(\mathbf{z}_\tau)&e^{-c\tau} - g(\mathbf{x}) \\
&= \int_0^\tau \mathbf{p} \cdot \nabla g(\mathbf{z}_{s-}) e^{-cs} ds - c \int_0^\tau g(\mathbf{z}_{s-}) e^{-cs} ds \\
&\quad - \int_0^\tau e^{-cs} \left(\nabla g(\mathbf{z}_{s-}) \cdot d\mathbf{L}^c_s \right) + \sum_{\mathbf{L}_s \neq \mathbf{L}_{s-}, \ s \leq \tau} \left(g(\mathbf{z}_s) - g(\check{\mathbf{z}}_s) \right) e^{-cs} \\
&\quad + \sum_{\check{\mathbf{z}}_s \neq \mathbf{z}_{s-}, \ s \leq \tau} \left(g(\check{\mathbf{z}}_s) - g(\mathbf{z}_{s-}) \right) e^{-cs}.
\end{aligned}
\tag{7.1}
$$

Note that $\mathbf{z}_s \in \mathbf{R}_+^n$ for $s \le \tau$ except in the case that $\tau = \tau^{\mathbf{L}}$. Since $\mathbf{z}_s = \check{\mathbf{z}}_s - \Delta \mathbf{L}_s$,

$$
-\int_0^\tau e^{-cs} \nabla g(\mathbf{z}_{s-}) \cdot d\mathbf{L}_s^c + \sum_{\mathbf{L}_s \ne \mathbf{L}_{s-}, s \le \tau} (g(\mathbf{z}_s) - g(\check{\mathbf{z}}_s)) e^{-cs}
$$

$$
= -\int_0^\tau e^{-cs} \nabla g(\mathbf{z}_{s-}) \cdot d\mathbf{L}_s^c - \sum_{\mathbf{L}_s \ne \mathbf{L}_{s-}, s \le \tau} e^{-cs} \left(\int_0^1 (\nabla g(\check{\mathbf{z}}_s - \gamma \Delta \mathbf{L}_s) \cdot \Delta \mathbf{L}_s) \, d\gamma \right)
$$

$$
= -\int_{0-}^\tau e^{-cs} \mathbf{a} \cdot d\mathbf{L}_s + \int_0^\tau e^{-cs} (\mathbf{a} - \nabla g(\mathbf{z}_{s-})) \cdot d\mathbf{L}_s^c
$$

$$
+ \sum_{\mathbf{L}_s \ne \mathbf{L}_{s-}, s \le \tau} e^{-cs} \int_0^1 (\mathbf{a} - \nabla g(\check{\mathbf{z}}_s - \gamma \Delta \mathbf{L}_s)) \cdot \Delta \mathbf{L}_s d\gamma. \tag{7.2}
$$

Since

$$
M_1(t) = \sum_{\check{\mathbf{z}}(s-) \ne \mathbf{z}_{s-}, s \le t} (g(\check{\mathbf{z}}_s) - g(\mathbf{z}_{s-})) e^{-cs}
$$

$$
-\lambda \int_0^t e^{-cs} \int_{\mathbf{R}_+^n} (g(\mathbf{z}_{s-} - \boldsymbol{\alpha}) - g(\mathbf{z}_{s-})) \, dF(\boldsymbol{\alpha}) ds \tag{7.3}
$$

and

$$
M_2(t) = \sum_{\check{\mathbf{z}}(s-) \ne \mathbf{z}_{s-}, s \le t} -\upsilon(\check{\mathbf{z}}_{s-}, \mathbf{z}(s-) - \check{\mathbf{z}}_s) e^{-cs} + \lambda \int_0^t e^{-cs} \int_{\mathbf{R}_+^n} \upsilon(\mathbf{z}_{s-}, \boldsymbol{\alpha}) dF(\boldsymbol{\alpha}) ds
$$

$$
\tag{7.4}
$$

are martingales with zero expectation, we have from (7.1) and (7.2)

$$
\begin{aligned}
& (g(\mathbf{z}_\tau) I_{\{\tau < \tau^{\mathbf{L}}\}} - \upsilon(\mathbf{z}_{\tau-}, \mathbf{z}_{\tau-} - \mathbf{z}_\tau) I_{\{\tau = \tau^{\mathbf{L}}\}}) e^{-c\tau} - g(\mathbf{x}) \\
& = (g(\mathbf{z}_\tau) - \upsilon(\mathbf{z}_{\tau-}, \mathbf{z}_{\tau-} - \mathbf{z}_\tau)) e^{-c\tau} - g(\mathbf{x}) \\
& = \int_0^\tau \mathcal{L}(g)(\mathbf{z}_{s-}) e^{-cs} ds - \int_{0-}^\tau e^{-cs} \mathbf{a} \cdot d\mathbf{L}_s \\
& \quad + \int_0^\tau e^{-cs} (\mathbf{a} - \nabla g(\mathbf{z}_{s-})) \cdot d\mathbf{L}_s^c \\
& \quad + \sum_{\mathbf{L}_s \ne \mathbf{L}_{s-}, s \le \tau} e^{-cs} \int_0^1 (\mathbf{a} - \nabla g(\check{\mathbf{z}}_s - \gamma \Delta \mathbf{L}_s) \cdot \Delta \mathbf{L}_s) \, d\gamma + M(\tau);
\end{aligned}
$$

where $M(t) = M_1(t) + M_2(t)$.   □

In order to prove Lemma 3.6, we will use a technical lemma in which we construct a sequence of smooth functions that approximate a (possible non-smooth) viscosity supersolution. This is done in order to apply Lemma 3.3 to an approximate smooth function instead of the viscosity supersolution; we have to do that because the amount of time the controlled process spends at non-differentiable points of the viscosity supersolution could have positive Lebesgue measure. We omit the proof of this lemma

because it is similar to the one-dimensional version given in Lemma 4.1 of [6]; the result is obtained by standard convolution arguments using that the function $\mathcal{R}$ is continuous.

Lemma 7.1 *Fix $\mathbf{x}^0$ in the interior of $\mathbf{R}^n_+$ and let $\overline{u}$ be a supersolution of* (3.1) *satisfying the growth condition* (2.14). *We can find a sequence of functions $\overline{u}_m : \mathbf{R}^n_+ \to \mathbf{R}$ such that:*

(a) *$\overline{u}_m$ is continuously differentiable and $\overline{u}_m \geq \overline{u} \geq f$.*

(b) *$\overline{u}_m$ satisfies the growth condition* (2.14).

(c) *$\mathbf{p} \cdot \nabla \overline{u}_m \leq (c + \lambda) \overline{u}_m + \lambda |\overline{u}(\mathbf{0})| + \lambda \mathbb{E}(|\upsilon(\mathbf{0}, \mathbf{U}_1)|)$ in $\mathbf{R}^n_+$ and $\mathbf{a} - \nabla \overline{u}_m \leq \mathbf{0}$.*

(d) *$\overline{u}_m \searrow \overline{u}$ uniformly on compact sets in $\mathbf{R}^n_+$ and $\nabla \overline{u}_m$ converges to $\nabla \overline{u}$ a.e. in $\mathbf{R}^n_+$.*

(e) *There exists a sequence c_m with $\lim\limits_{m \to \infty} c_m = 0$ such that*

$$\sup_{\mathbf{x} \in [0, \mathbf{x}^0]} \mathcal{L}(\overline{u}_m)(\mathbf{x}) \leq c_m.$$

Proof of Lemma 3.6 Consider the processes $\mathbf{z}_s = \mathbf{X}^{\mathbf{L}}_s$ defined in (2.3), let us call $\tau = \tau^{\mathbf{L}}$ and take $\widetilde{\tau} = \overline{\tau} \wedge \tau$. Let us consider the functions $\overline{u}_m$ defined in Lemma 7.1 in $\mathbf{R}^n_+$. Using Lemma 3.3 for $\widetilde{\tau} \wedge t$, we get from Lemma 7.1 (a) and (c) that

$$
\begin{aligned}
&\overline{u}_m(\mathbf{z}_t) e^{-ct} I_{\{t < \widetilde{\tau}\}} + e^{-c\overline{\tau}} f(\mathbf{z}_{\overline{\tau}}) I_{\{t \wedge \widetilde{\tau} = \overline{\tau}, \overline{\tau} < \tau\}} - e^{-c\overline{\tau}} \upsilon(\mathbf{z}_{\tau-}, \mathbf{z}_{\tau-} - \mathbf{z}_\tau) I_{\{t \wedge \widetilde{\tau} = \tau\}} - \overline{u}_m(\mathbf{x}) \\
&\leq \overline{u}_m(\mathbf{z}_t) e^{-ct} I_{\{t < \widetilde{\tau}\}} + e^{-c\overline{\tau}} \overline{u}_m(\mathbf{z}_{\overline{\tau}}) I_{\{t \wedge \widetilde{\tau} = \overline{\tau}, \overline{\tau} < \tau\}} - e^{-c\overline{\tau}} \upsilon(\mathbf{z}_{\tau-}, \mathbf{z}_{\tau-} - \mathbf{z}_\tau) I_{\{t \wedge \widetilde{\tau} = \tau\}} - \overline{u}_m(\mathbf{x}) \\
&\leq \int_0^{t \wedge \widetilde{\tau}} \mathcal{L}(\overline{u}_m)(\mathbf{z}_{s-}) e^{-cs} ds - \int_{0-}^{t \wedge \widetilde{\tau}} e^{-cs} \mathbf{a} \cdot d\mathbf{L}_s + M(t \wedge \widetilde{\tau}),
\end{aligned}
\tag{7.5}
$$

where $M(t)$ is a zero-expectation martingale. Since $\mathbf{L}_s$ is non-decreasing we get, using the monotone convergence theorem, that

$$
\begin{aligned}
&\lim_{t \to \infty} \mathbb{E}_{\mathbf{x}} \left(\int_{0-}^{t \wedge \widetilde{\tau}} e^{-cs} \mathbf{a} \cdot d\mathbf{L}_s + e^{-c\overline{\tau}} f(\mathbf{z}_{\overline{\tau}}) I_{\{t \wedge \widetilde{\tau} = \overline{\tau}, \overline{\tau} < \tau\}} - e^{-c\overline{\tau}} \upsilon(\mathbf{z}_{\tau-}, \mathbf{z}_{\tau-} - \mathbf{z}_\tau) I_{\{t \wedge \widetilde{\tau} = \tau\}} \right) \\
&= J(\pi; \mathbf{x}).
\end{aligned}
$$

From Lemma 7.1(c), we have

$$
\begin{aligned}
&-(c + \lambda) \overline{u}_m(\mathbf{x}) + \overline{u}_m(0) \lambda F(\mathbf{x}) - \lambda \mathbb{E}(|\upsilon(\mathbf{0}, \mathbf{U}_1)|) \\
&\leq \mathcal{L}(\overline{u}_m)(\mathbf{x}) \leq \lambda \overline{u}_m(\mathbf{x}) + \lambda |\overline{u}(\mathbf{0})| + \lambda \mathbb{E}(|\upsilon(\mathbf{0}, \mathbf{U}_1)|) - \mathcal{R}(\mathbf{x}).
\end{aligned}
\tag{7.6}
$$

By Lemma 7.1(b), (c) and the inequality $\mathbf{z}_s \leq \mathbf{x} + \mathbf{p}s$, there exists d_0 large enough such that

$$
\overline{u}_m(\mathbf{z}_s) \leq \overline{u}_m(\mathbf{x} + \mathbf{p}s) \leq d_0 e^{\frac{c}{2n} \sum_{i=1}^n \frac{x_i + p_i s}{p_i}} = d_0 h_0(\mathbf{x}) e^{\frac{c}{2}s}
\tag{7.7}
$$

and

$$
-\upsilon(\mathbf{z}_{s-}, \boldsymbol{\alpha}) \leq S(\mathbf{z}_{s-}) \leq d_0 h_0(\mathbf{x}) e^{\frac{c}{2}s} \quad \text{for} \quad (\mathbf{z}_{s-} - \boldsymbol{\alpha}) \notin \mathbf{R}^n_+,
\tag{7.8}
$$

where h_0 and S are defined in (2.15) and Proposition 2.4 respectively. Therefore, from (7.6), we obtain that there exists d_1 large enough such that,

$$e^{-cs} \left| \mathcal{L}(\overline{u}_m) \left(\mathbf{z}_{s-} \right) \right| \leq d_1 e^{-\frac{c}{2}s}. \tag{7.9}$$

And using the bounded convergence theorem,

$$\lim_{t \to \infty} \mathbb{E}_{\mathbf{x}} \left(\int_0^{t \wedge \widetilde{\tau}} \mathcal{L}(\overline{u}_m)(\mathbf{z}_{s-}) e^{-cs} ds \right) = \mathbb{E}_{\mathbf{x}} \left(\int_0^{\widetilde{\tau}} \mathcal{L}(\overline{u}_m)(\mathbf{z}_{s-}) e^{-cs} ds \right). \tag{7.10}$$

From (7.5) and (7.10), we get

$$\lim_{t \to \infty} \mathbb{E}_{\mathbf{x}} \left(\overline{u}_m(\mathbf{z}_t) e^{-ct} I_{\{t < \widetilde{\tau}\}} \right) - \overline{u}_m(\mathbf{x}) \leq \mathbb{E}_{\mathbf{x}} \left(\int_0^{\widetilde{\tau}} \mathcal{L}(\overline{u}_m)(\mathbf{z}_{s-}) e^{-cs} ds \right) - J(\pi; \mathbf{x}). \tag{7.11}$$

By (7.7),

$$\lim_{t \to \infty} \mathbb{E}_{\mathbf{x}} \left(\overline{u}_m(\mathbf{z}_t) e^{-ct} I_{\{t < \widetilde{\tau}\}} \right) = 0. \tag{7.12}$$

Let us prove now that

$$\limsup_{m \to \infty} \mathbb{E}_{\mathbf{x}} \left(\int_0^{\widetilde{\tau}} \mathcal{L}(\overline{u}_m)(\mathbf{z}_{s-}) e^{-cs} ds \right) \leq 0. \tag{7.13}$$

Given any $\varepsilon > 0$, from (7.9), we can find T large enough such that

$$\mathbb{E}_{\mathbf{x}} \left(\int_{T \wedge \widetilde{\tau}}^{\widetilde{\tau}} \left| \mathcal{L}(\overline{u}_m)(\mathbf{z}_{s-}) \right| e^{-cs} ds \right) \leq \frac{2d_1}{c} \left(e^{-\frac{c}{2}T} \right) < \frac{\varepsilon}{2}. \tag{7.14}$$

For $s \leq T$, we get $\mathbf{z}_{s-} \in [\mathbf{0}, \mathbf{x} + \mathbf{p}T \,]$, then from Lemma 7.1(e) we can find m_0 large enough such that for any $m \geq m_0$

$$\int_0^T \mathcal{L}(\overline{u}_m)(\mathbf{z}_{s-}) e^{-cs} ds \leq c_m \int_0^T e^{-cs} ds \leq \frac{c_m}{c} \leq \frac{\varepsilon}{2}$$

and so we have (7.13). Thus, from (7.11) and using (7.12) and (7.13), we obtain

$$\overline{u}(\mathbf{x}) = \lim_{m \to \infty} \overline{u}_m(\mathbf{x}) \geq J(\pi; \mathbf{x}). \tag{7.15}$$

$\square$

Proof of Lemma 4.2 Suppose that $\tilde{k} = \infty$, calling

$$k_l := \mathbf{m} \cdot \mathbf{1} + (l - 1)n + 1,$$

there are at least $i_l \geq l$ control actions $\mathbf{E}_0$ in $(s_1, s_2, \ldots, s_{k_l})$. Let us consider the non-decreasing sequence $(j_l)_l$ defined as

$$j_l := \max\{j : \tau_j \leq t_{k_l}\},$$

we have that $t_{k_l} \geq \tau_{j_l} + (i_l - j_l)\delta$. If $\lim_{l \to \infty} i_l - j_l = \infty$, then

$$\lim_{l \to \infty} t_{k_l} \geq \lim_{l \to \infty} \tau_{j_l} + (i_l - j_l)\delta \geq \lim_{l \to \infty} (i_l - j_l)\delta = \infty;$$

if not, $\lim_{l \to \infty} j_l = \infty$ and so

$$\lim_{l \to \infty} t_{k_l} \geq \lim_{l \to \infty} \tau_{j_l} + (i_l - j_l)\delta \geq \lim_{l \to \infty} \tau_{j_l}$$

and since $\lim_{l \to \infty} \tau_{j_l} = \lim_{i \to \infty} \tau_i = \infty$ a.s., we have the result. □

Proof of Lemma 4.4 It is straightforward that T_0, T_i, T_s and T are non-decreasing and that

$$\sup_{\mathbf{m} \in \mathbf{N}_0^n} |T(w_1)(\mathbf{m}) - T(w_2)(\mathbf{m})| \leq \sup_{\mathbf{m} \in \mathbf{N}_0^n} |w_1(\mathbf{m}) - w_2(\mathbf{m})|.$$

Also, given a function $w : \mathbf{N}_0^n \to \mathbf{R}$ it is immediate to see that $T_i(w)$ and $T_s(w)$ can be written as a linear combination of the values of $w(\mathbf{m})$ plus a constant. Let us prove now that

$$T_0(w)(\mathbf{m}) = e^{-(c+\lambda)\delta} w(\mathbf{m}+1) + \sum_{0 \leq \mathbf{k} \leq \mathbf{m}} a_1(\mathbf{k}, \mathbf{m}) w(\mathbf{k}) + a_2(\mathbf{m}),$$

Lemma 7.2 *where*

$$a_1(\mathbf{k}, \mathbf{m}) = I_{\{\mathbf{k} \leq \mathbf{m}-1\}} \int_0^\delta \lambda e^{-(c+\lambda)t} (F(g^\delta(\mathbf{m}-\mathbf{k})+t\mathbf{p}) - F(g^\delta(\mathbf{m}-\mathbf{k}-1)+t\mathbf{p}))dt$$

$$+ I_{\{\mathbf{k} \leq \mathbf{m}, \mathbf{k} \nleq \mathbf{m}-1\}} \int_0^\delta \lambda e^{-(c+\lambda)t} (F(g^\delta(\mathbf{m}-\mathbf{k})$$
$$+t\mathbf{p}) - F(\mathbf{0} \vee (g^\delta(\mathbf{m}-\mathbf{k}) + t\mathbf{p})))dt$$

and

$$a_2(\mathbf{m}) = \sum_{0 \leq \mathbf{k} < \mathbf{m}-1} \int_0^\delta \left(\lambda e^{-(c+\lambda)t} \int_{g^\delta(\mathbf{m}-\mathbf{k}-1)+t\mathbf{p}}^{g^\delta(\mathbf{m}-\mathbf{k})+t\mathbf{p}} \mathbf{a} \cdot (g^\delta(\mathbf{m}-\mathbf{k})+t\mathbf{p}-\boldsymbol{\alpha})dF(\boldsymbol{\alpha}) \right) dt$$

$$+ \sum_{\mathbf{k} \leq \mathbf{m}, \mathbf{k} \nleq \mathbf{m}-1} \int_0^\delta \left(\lambda e^{-(c+\lambda)t} \int_{\mathbf{0} \vee (g^\delta(\mathbf{m}-\mathbf{k})+t\mathbf{p})}^{g^\delta(\mathbf{m}-\mathbf{k})+t\mathbf{p}} \mathbf{a} \cdot (g^\delta(\mathbf{m}-\mathbf{k})+t\mathbf{p}-\boldsymbol{\alpha})dF(\boldsymbol{\alpha}) \right) dt$$

$$- \int_0^\delta e^{-(c+\lambda)t} \mathcal{R}(g^\delta(\mathbf{m})+t\mathbf{p})dt.$$

Given $\mathbf{m} \in \mathbf{N}_0^n, \boldsymbol{\alpha} \in \mathbf{R}_+^n$ and $0 < t \le \delta$ such that $\mathbf{0} \le g^\delta(\mathbf{m}) + t\mathbf{p} - \boldsymbol{\alpha}$, let us define

$$\mathbf{k} := \rho^\delta(g^\delta(\mathbf{m}) + t\mathbf{p} - \boldsymbol{\alpha}),$$

and so $\mathbf{k} \le \mathbf{m}$.

If $\mathbf{k} \le \mathbf{m} - \mathbf{1}$,

$$g^\delta(\mathbf{k}) \le g^\delta(\mathbf{m}) + t\mathbf{p} - \boldsymbol{\alpha} < g^\delta(\mathbf{k} + \mathbf{1}) \le g^\delta(\mathbf{m})$$

that implies

$$\mathbf{0} < g^\delta(\mathbf{m} - \mathbf{k} - \mathbf{1}) + t\mathbf{p} < \boldsymbol{\alpha} \le g^\delta(\mathbf{m} - \mathbf{k}) + t\mathbf{p}.$$

If $\mathbf{k} \le \mathbf{m}$ with $\mathbf{k} \not\le \mathbf{m} - \mathbf{1}$,

$$g^\delta(\mathbf{k}) \le g^\delta(\mathbf{m}) + t\mathbf{p} - \boldsymbol{\alpha} < g^\delta(\mathbf{k} + \mathbf{1}) \wedge \left(g^\delta(\mathbf{m}) + t\mathbf{p} \right)$$

and so

$$\left(g^\delta(\mathbf{m} - \mathbf{k} - \mathbf{1}) + t\mathbf{p} \right) \vee \mathbf{0} < \boldsymbol{\alpha} \le g^\delta(\mathbf{m} - \mathbf{k}) + t\mathbf{p}.$$

Then, we can write

$$
\begin{aligned}
&\mathcal{I}^\delta(w)(\mathbf{m}) \\
&= \sum_{\mathbf{0} \le \mathbf{k} \le \mathbf{m} - \mathbf{1}} w(\mathbf{k}) \int_0^\delta \lambda e^{-(c+\lambda)t} \left(\int_{g^\delta(\mathbf{m}-\mathbf{k}-\mathbf{1})+t\mathbf{p}}^{g^\delta(\mathbf{m}-\mathbf{k})+t\mathbf{p}} dF(\boldsymbol{\alpha}) \right) dt \\
&+ \sum_{\mathbf{0} \le \mathbf{k} \le \mathbf{m} - \mathbf{1}} \int_0^\delta \lambda e^{-(c+\lambda)t} \left(\int_{g^\delta(\mathbf{m}-\mathbf{k}-\mathbf{1})+t\mathbf{p}}^{g^\delta(\mathbf{m}-\mathbf{k})+t\mathbf{p}} \mathbf{a} \cdot \left(g^\delta(\mathbf{m} - \mathbf{k}) + t\mathbf{p} - \boldsymbol{\alpha} \right) dF(\boldsymbol{\alpha}) \right) dt \\
&+ \sum_{\mathbf{k} \le \mathbf{m}, \mathbf{k} \not\le \mathbf{m} - \mathbf{1}} w(\mathbf{k}) \int_0^\delta \lambda e^{-(c+\lambda)t} \left(\int_{(g^\delta(\mathbf{m}-\mathbf{k}-\mathbf{1})+t\mathbf{p})\vee\mathbf{0}}^{g^\delta(\mathbf{m}-\mathbf{k})+t\mathbf{p}} dF(\boldsymbol{\alpha}) \right) dt \\
&+ \sum_{\mathbf{k} \le \mathbf{m}, \mathbf{k} \not\le \mathbf{m} - \mathbf{1}} \int_0^\delta \lambda e^{-(c+\lambda)t} \left(\int_{(g^\delta(\mathbf{m}-\mathbf{k}-\mathbf{1})+t\mathbf{p})\vee\mathbf{0}}^{g^\delta(\mathbf{m}-\mathbf{k})+t\mathbf{p}} \mathbf{a} \cdot \left(g^\delta(\mathbf{m}-\mathbf{k}) + t\mathbf{p} - \boldsymbol{\alpha} \right) dF(\boldsymbol{\alpha}) \right) dt.
\end{aligned}
$$

Therefore, from (4.2), we have the result.   □

Proof of Lemma 4.6 The proof of this lemma is a discrete version of the one of Lemma 3.6. Assume that $\pi = (\mathbf{L}, \overline{\tau}) \in \Pi_{g^\delta(\mathbf{m})}^\delta$. For any $\omega = (\tau_i, \mathbf{U}_i)_{i\ge1}$, consider the sequence $\mathbf{s} = (s_k)_{k=1,\dots,\tilde{k}}$ with $s_k \in \mathcal{E}$ corresponding to π and $\mathbf{m}^k, \mathbf{y}^k$ and times t_k and Δ_k as defined in Section 4. Let $(\kappa_l)_{l\ge1}$ be the indices of the sequence $\mathbf{s} = (s_k)_{k=1,\dots,\tilde{k}}$ where s_k is either $\mathbf{E}_s$ or $\mathbf{E}_0$. If the sequence stops at $\tilde{k} = \kappa_{l_0} < \infty$, we define

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

and if $\tilde{k} = \infty$ we put $l_0 = \infty$. Consider the case in which the process goes to ruin at κ_l, that is $\mathbf{y}^{\kappa_l} \notin \mathbf{R}_+^n$; then the surplus prior to the ruin is $\mathbf{y}^{\kappa_l} + \mathbf{U}$ and the penalty paid

at ruin is $\upsilon(\mathbf{y}^{\kappa_l} + \mathbf{U}, \mathbf{U})$, where $\mathbf{U}$ is the last jump of the uncontrolled process. So we define, for $l \geq 1$,

$$H(l) = w(\mathbf{m}^{1+\kappa_l})I_{\{s_{\kappa_l}=\mathbf{E}_0\}}I_{\{\mathbf{y}^{\kappa_l}\in\mathbf{R}^n_+\}} - \upsilon(\mathbf{y}^{\kappa_l} + \mathbf{U}, \mathbf{U})I_{\{s_{\kappa_l}=\mathbf{E}_0\}}I_{\{\mathbf{y}^{\kappa_l}\notin\mathbf{R}^n_+\}} + f(g^\delta(\mathbf{m}^{\kappa_l}))I_{\{s_{\kappa_l}=\mathbf{E}_s\}}.$$

If we put $H(0) = w(\mathbf{m})$, $\kappa_0 = 0$ and $t_0 = 0$, we have using $(T_i(w) - w)_{i=1,\dots,n} \leq 0$,

$$
\begin{aligned}
e^{-ct_{\kappa_{l+1}}} H(l) - w(\mathbf{m}) &= \sum_{j=1}^{l} \left(e^{-ct_{\kappa_{j+1}}} H(j) - e^{-ct_{\kappa_j}} H(j-1) \right) \\
&= \sum_{j=1}^{l} I_{\{\kappa_{j+1}\neq\kappa_j\}} \left(e^{-ct_{\kappa_{j+1}}} H(j) - e^{-ct_{\kappa_j}} H(j-1) \right) \\
&= \sum_{j=1}^{l} I_{\{\kappa_{j+1}\neq\kappa_j\}} \left(e^{-ct_{1+\kappa_{j-1}}} \left(\sum_{k=1+\kappa_{j-1}}^{\kappa_j-1} \left(w(\mathbf{m}^{k+1}) - w(\mathbf{m}^k) \right) \right) \right) \\
&\quad + \sum_{j=1}^{l} I_{\{\kappa_{j+1}\neq\kappa_j\}} \left(e^{-ct_{\kappa_{j+1}}} H(j) - e^{-ct_{\kappa_j}} w(\mathbf{m}^{\kappa_j}) \right) \\
&\leq \sum_{j=1}^{l} I_{\{\kappa_{j+1}\neq\kappa_j\}} \left(\sum_{k=1+\kappa_{j-1}}^{\kappa_j-1} e^{-ct_{1+\kappa_{j-1}}} \left(\sum_{i=1}^{n} (-a_i p_i \delta) I_{\{s_k=\mathbf{E}_i\}} \right) \right) \\
&\quad + \sum_{j=1}^{l} I_{\{\kappa_{j+1}\neq\kappa_j\}} \left(e^{-ct_{\kappa_{j+1}}} H(j) - e^{-ct_{\kappa_j}} w(\mathbf{m}^{\kappa_j}) \right);
\end{aligned}
\tag{7.16}
$$

and since $T_0(w) - w \leq 0$ and $T_s(w) - w \leq 0$, if $\kappa_{j+1} \neq \kappa_j$,

$$
\begin{aligned}
&\mathbb{E}\left(e^{-ct_{\kappa_{j+1}}} H(j) - e^{-ct_{\kappa_j}} w(\mathbf{m}^{\kappa_j}) \Big| \mathcal{F}_{t_{\kappa_j}} \right) \\
&= \mathbb{E}\left((e^{-ct_{\kappa_{j+1}}} H(j) - e^{-ct_{\kappa_j}} w(\mathbf{m}^{\kappa_j}))I_{\{s_{\kappa_j}=\mathbf{E}_0\}} \Big| \mathcal{F}_{t_{\kappa_j}} \right) \\
&\quad + I_{\{s_{\kappa_j}=\mathbf{E}_s\}} e^{-ct_{\kappa_j}} \left(f(g^\delta(\mathbf{m}^{\kappa_j})) - w(\mathbf{m}^{\kappa_j}) \right) \\
&\leq \mathbb{E}\left(e^{-ct_{\kappa_{j+1}}} I_{\{s_{\kappa_j}=\mathbf{E}_0\}}(w(\mathbf{m}^{1+\kappa_j})I_{\{\mathbf{y}^{\kappa_j}\in\mathbf{R}^n_+\}} - \upsilon(\mathbf{y}^{\kappa_j} + \mathbf{U}, \mathbf{U})I_{\{\mathbf{y}_{\kappa_j}\notin\mathbf{R}^n_+\}}) \Big| \mathcal{F}_{t_{\kappa_j}} \right) \\
&\quad - e^{-ct_{\kappa_j}} w(\mathbf{m}^{\kappa_j})I_{\{s_{\kappa_j}=\mathbf{E}_0\}} \\
&= e^{-ct_{\kappa_j}} I_{\{s_{\kappa_j}=\mathbf{E}_0\}} (T_0(w)(\mathbf{m}^{\kappa_j}) - w(\mathbf{m}^{\kappa_j})) \\
&\quad - e^{-ct_{\kappa_j}} I_{\{s_{\kappa_j}=\mathbf{E}_0\}} \int_0^\delta \int_{\boldsymbol{\alpha}\in[\mathbf{0},\mathbf{z}_j(t)]} \lambda e^{-(c+\lambda)t} \mathbf{a} \cdot \left(\mathbf{z}_j(t) - \boldsymbol{\alpha} - \langle \mathbf{z}_j(t) - \boldsymbol{\alpha} \rangle^\delta \right) dF(\boldsymbol{\alpha})dt \\
&\leq -e^{-ct_{\kappa_j}} I_{\{s_{\kappa_j}=\mathbf{E}_0\}} \int_0^\delta \int_{\boldsymbol{\alpha}\in[\mathbf{0},\mathbf{z}_j(t)]} \lambda e^{-(c+\lambda)t} \mathbf{a} \cdot \left(\mathbf{z}_j(t) - \boldsymbol{\alpha} - \langle \mathbf{z}_j(t) - \boldsymbol{\alpha} \rangle^\delta \right) dF(\boldsymbol{\alpha})dt,
\end{aligned}
\tag{7.17}
$$

where $\mathbf{z}_j(t) = g^\delta(\mathbf{m}^{\kappa_j}) + t\mathbf{p}$. From (7.16) and (7.17), and calling the initial surplus $\mathbf{x} = g^\delta(\mathbf{m}) \in \mathcal{G}^\delta$ we have,

$$\limsup_{l\to\infty} \mathbb{E}_\mathbf{x}\left(e^{-ct_{\kappa_{l+1}}} H(l) - w(\mathbf{m}) \right) \leq -\mathbb{E}_\mathbf{x}\left(\int_{0-}^{\bar{\tau}\wedge\tau_L} e^{-cs} \mathbf{a} \cdot d\mathbf{L}_s \right).$$

Then,

$$w(\mathbf{m}) \geq J(\pi; g^\delta(\mathbf{m})) + \limsup_{l\to\infty} \mathbb{E}_\mathbf{x}\left(I_{\{l\leq l_0\}} e^{-ct_{1+\kappa_l}} w(\mathbf{m}^{1+\kappa_l})I_{\{\mathbf{y}^{\kappa_l}\in\mathbf{R}^n_+\}} \right).$$

Since

$$g^\delta(\mathbf{m}^{1+\kappa_l}) \leq g^\delta\left(\mathbf{m} + \rho^\delta(t_{1+\kappa_l}\mathbf{p})\right)$$

and w satisfies the growth condition (4.7), there exists d large enough such that

$$\limsup_{l\to\infty} \left(\mathbb{E}_\mathbf{x} I_{\{l\leq l_0\}} e^{-ct_{1+\kappa_l}} w(\mathbf{m}^{1+\kappa_l}) I_{\{\mathbf{y}^{\kappa_l}\in\mathbf{R}_+^n\}}\right)$$

$$\leq d \lim_{l\to\infty} \mathbb{E}_\mathbf{x}\left(I_{\{l\leq l_0\}} e^{-ct_{1+\kappa_l}} e^{c\delta\mathbf{m}\cdot\mathbf{1}/(2n)} e^{\frac{c}{2}t_{1+\kappa_l}}\right) = 0;$$

so we have the result.   □

Proof of Lemma 5.1 (1) Take the $\mathcal{G}^\delta$-optimal strategy $\pi^\delta_{g^\delta(\mathbf{m})} \in \Pi^\delta_{g^\delta(\mathbf{m})}$ and define $\overline{\pi}_{g^\delta(\mathbf{m}+\mathbf{e}_i)} \in \Pi^\delta_{g^\delta(\mathbf{m}+\mathbf{e}_i)}$ by applying first the control action $\mathbf{E}_i$ and then the $\mathcal{G}^\delta$-optimal strategy $\pi^\delta_{g^\delta(\mathbf{m})}$. The value function of this strategy is given by

$$a_i\, p_i\, \delta + v^\delta(\mathbf{m}),$$

so we obtain the the first inequality of this proposition. Now, take the $\mathcal{G}^\delta$-optimal strategy $\pi^\delta_{g^\delta(\mathbf{m}+1)} \in \Pi^\delta_{g^\delta(\mathbf{m}+1)}$ and define $\overline{\pi}_{g^\delta(\mathbf{m})} \in \Pi^\delta_{g^\delta(\mathbf{m})}$ by applying first the control action $\mathbf{E}_0$ and then the $\mathcal{G}^\delta$-optimal strategy $\pi^\delta_{g^\delta(\mathbf{m}+1)}$. Hence, we obtain the second inequality from

$$v^\delta(\mathbf{m}+1)\, e^{-(c+\lambda)\delta} \leq T_0(v^\delta)(\mathbf{m}) \leq T(v^\delta)(\mathbf{m}) = v^\delta(\mathbf{m}).$$

(2) In order to avoid any confusion, in the remainder of the proof we put a superindex δ to the control actions in $\mathcal{G}^\delta$. Note first that given any surplus in $\mathbf{R}_+^n$, the strategy of paying dividends in such a way that the surplus goes to the nearest smaller point in $\mathcal{G}^{2\delta}$ corresponds to go first to the nearest smaller point in $\mathcal{G}^\delta$ and then to apply (possibly) a combination of control actions $\mathbf{E}_i^{\delta'}s$. Consider $\pi_{2g^\delta(\mathbf{m})} \in \Pi^{2\delta}_{g^{2\delta}(\mathbf{m})}$ given by the random sequence $\mathbf{s} = (s_k)_{k=1,\ldots,\tilde{k}}$ with

$$s_k \in \mathcal{E}^{2\delta} = \left\{\mathbf{E}_s^{2\delta}, \left(\mathbf{E}_i^{2\delta}\right)_{i=1,\ldots,n}, \mathbf{E}_0^{2\delta}\right\}.$$

We can see that $\pi_{2g^\delta(\mathbf{m})}$ also belongs to $\Pi^\delta_{2g^\delta(\mathbf{m})}$ rewriting the sequence as follows: If $s_k = \mathbf{E}_i^{2\delta}$, we replace it by the pair $\mathbf{E}_i^\delta, \mathbf{E}_i^\delta$; if $s_k = \mathbf{E}_s^{2\delta}$, we replace it by $\mathbf{E}_s^\delta$; and if $s_k = \mathbf{E}_0^{2\delta}$, we replaces it

- either by $\mathbf{E}_0^\delta, \mathbf{E}_0^\delta$ if the next jump in the uncontrolled process arrives at time $\tau > 2\delta$;
- or by $\mathbf{E}_0^\delta, \mathbf{E}_0^\delta$, and a possible combination of $\mathbf{E}_i^{\delta'}s$, if it arrives at time $\tau \in (\delta, 2\delta]$, so the surplus goes to the nearest smaller point in $\mathcal{G}^{2\delta}$;

- or by $\mathbf{E}_0^{\delta}$, and a possible combination of $\mathbf{E}_i^{\delta'}s$, if it arrives at time $\tau \le \delta$, so again the surplus goes to the nearest smaller point in $\mathcal{G}^{2\delta}$.
  So we have the result.

$\square$

Proof of Lemma 5.3 Let us first prove that

$$
\begin{aligned}
&\left| V^{\delta_k}(\mathbf{y}) - V^{\delta_k}(\mathbf{x}) \right| \\
&\le \frac{2}{\hat{p}} V^{\delta_k}\left(\langle \mathbf{x} \vee \mathbf{y} \rangle^{\delta_k} \right) \left(\frac{e^{(c+\lambda)\delta_k} - 1}{\delta_k} \right) \left\| \langle \mathbf{y} \rangle^{\delta_k} - \langle \mathbf{x} \rangle^{\delta_k} \right\|_1 + 2\delta_k \mathbf{a} \cdot \mathbf{p},
\end{aligned}
\tag{7.18}
$$

for any $\mathbf{x}$ and $\mathbf{y}$ in $\mathbf{R}_+^n$. Let us assume first that $\mathbf{y} > \mathbf{x}$. We have from Lemma 5.1,

$$
\begin{aligned}
V^{\delta_k}&(g^{\delta_k}(\mathbf{m} + \mathbf{e}_i)) - V^{\delta_k}(g^{\delta_k}(\mathbf{m})) \\
&\le V^{\delta_k}(g^{\delta_k}(\mathbf{m} + \mathbf{1})) - V^{\delta_k}(g^{\delta_k}(\mathbf{m})) \le V^{\delta_k}(g^{\delta_k}(\mathbf{m}))(e^{(c+\lambda)\delta_k} - 1).
\end{aligned}
$$

Let us call $\mathbf{m_y} = \rho^{\delta_k}(\mathbf{y})$ and $\mathbf{m_x} = \rho^{\delta_k}(\mathbf{x})$. Then,

$$
\begin{aligned}
V^{\delta_k}(\mathbf{y}) - V^{\delta_k}(\mathbf{x}) &\le V^{\delta_k}\left(g^{\delta_k}(\mathbf{m_y}) \right) - V^{\delta_k}\left(g^{\delta_k}(\mathbf{m_x}) \right) + \mathbf{a} \cdot (\mathbf{y} - g^{\delta_k}(\mathbf{m_y})) \\
&\le \left(\frac{e^{(c+\lambda)\delta_k} - 1}{\delta_k} \right) V^{\delta_k}(\mathbf{y}) \sum_{i=1}^n \frac{g_i^{\delta_k}(\mathbf{m_y} - \mathbf{m_x})}{p_i} + \delta_k \mathbf{a} \cdot \mathbf{p} \\
&\le \left(\frac{e^{(c+\lambda)\delta_k} - 1}{\hat{p}\delta_k} \right) V^{\delta_k}(\mathbf{y}) \left\| g^{\delta_k}(\mathbf{m_y} - \mathbf{m_x}) \right\|_1 + \delta_k \mathbf{a} \cdot \mathbf{p}.
\end{aligned}
$$

Let us consider now $\mathbf{x}$ and $\mathbf{y}$ in $\mathbf{R}_+^n$, consider $\mathbf{m}_0 = \rho^{\delta_k}(\mathbf{x} \wedge \mathbf{y})$,

$$
\begin{aligned}
\left| V^{\delta_k}(\mathbf{y}) - V^{\delta_k}(\mathbf{x}) \right| \\
\le V^{\delta_k}(\mathbf{y}) - V^{\delta_k}(\mathbf{x} \wedge \mathbf{y}) + V^{\delta_k}(\mathbf{x}) - V^{\delta_k}(\mathbf{x} \wedge \mathbf{y}) \\
\le \frac{1}{\hat{p}} V^{\delta_k}(\mathbf{x} \vee \mathbf{y}) \left(\frac{e^{(c+\lambda)\delta_k} - 1}{\delta_k} \right) \left(\left\| g^{\delta_k}(\mathbf{m_y} - \mathbf{m}_0) \right\|_1 + \left\| g^{\delta_k}(\mathbf{m_x} - \mathbf{m}_0) \right\|_1 \right) + 2\delta_k \mathbf{a} \cdot \mathbf{p} \\
\le \frac{2}{\hat{p}} V^{\delta_k}(\mathbf{x} \vee \mathbf{y}) \left(\frac{e^{(c+\lambda)\delta_k} - 1}{\delta_k} \right) \left\| g^{\delta_k}(\mathbf{m_y} - \mathbf{m_x}) \right\|_1 + 2\delta_k \mathbf{a} \cdot p.
\end{aligned}
$$

Therefore we have (7.18).

By definitions (4.9) and (5.1), and since $T_i\left(v^{\delta_k} \right) \le v^{\delta_k}$,

$$
\begin{aligned}
\overline{V}(\mathbf{y}) - \overline{V}(\mathbf{x}) &\ge \overline{V}(\mathbf{y}) - V^{\delta_k}(\mathbf{y}) + \mathbf{a} \cdot g^{\delta_k}\left(\rho^{\delta_k}(\mathbf{y}) - \rho^{\delta_k}(\mathbf{x}) \right) \\
&\quad + \mathbf{a} \cdot (\mathbf{y} - g^{\delta_k}(\rho^{\delta_k}(\mathbf{y}) - \rho^{\delta_k}(\mathbf{x})) + \mathbf{x}) + V^{\delta_k}(\mathbf{x}) - \overline{V}(\mathbf{x});
\end{aligned}
$$

taking the limit as k goes to infinity, we obtain the first inequality of the Lipschitz inequality.

We can write, from (7.18),

$$
\begin{aligned}
\overline{V}(\mathbf{y}) - \overline{V}(\mathbf{x}) &= \overline{V}(\mathbf{y}) - V^{\delta_k}(\mathbf{y}) + V^{\delta_k}(\mathbf{y}) - V^{\delta_k}(\mathbf{x}) + V^{\delta_k}(\mathbf{x}) - \overline{V}(\mathbf{x}) \\
&\le \overline{V}(\mathbf{y}) - V^{\delta_k}(\mathbf{y}) + \frac{2}{\hat{p}} \overline{V}(\mathbf{y}) \left(\frac{e^{(c+\lambda)\delta_k} - 1}{\delta_k} \right) \left\| g^{\delta_k}\left(\rho^{\delta_k}(\mathbf{y}) - \rho^{\delta_k}(\mathbf{x}) \right) \right\|_1 \\
&\quad + 2\delta_k \mathbf{a} \cdot \mathbf{p} + V^{\delta_k}(\mathbf{x}) - \overline{V}(\mathbf{x});
\end{aligned}
$$

taking the limit as k goes to infinity, we obtain the second inequality of the Lipschitz inequality.   □

References

1. Albrecher, H., Azcue, P., Muler, N.: Optimal dividend strategies for two collaborating insurance companies. Adv. Appl. Probab. **49**(2), 515–548 (2017)
2. Asmussen, S., Taksar, M.: Controlled diffusion models for optimal dividend pay-out. Insur. Math. Econ. **20**, 1–15 (1997)
3. Avram, F., Palmowski, Z., Pistorius, M.: On the optimal dividend problem for a spectrally negative Lévy process. Ann. Appl. Probab. **17**, 156–180 (2007)
4. Avram, F., Palmowski, Z., Pistorius, M.: On Gerber–Shiu functions and optimal dividend distribution for a Levy risk-process in the presence of a penalty function. Ann. Appl. Probab. **25**(4), 1868–1935 (2015)
5. Azcue, P., Muler, N.: Optimal reinsurance and dividend distribution policies in the Cramér–Lundberg model. Math. Financ. **15**, 261–308 (2005)
6. Azcue, P., Muler, N.: Stochastic Optimization in Insurance: A Dynamic Programming Approach. Springer Briefs in Quantitative Finance. Springer, New York (2014)
7. Azcue, P., Muler, N.: Optimal dividend payment and regime switching in a compound Poisson risk model. SIAM J. Control Optim. **53**(5), 3270–3298 (2015)
8. Barles, G., Souganidis, P.E.: Convergence of approximation schemesfor fully nonlinear second order equations. Asymptot. Anal. **4**, 271–283 (1991)
9. Budhiraja, A., Ross, K.: Convergent numerical scheme for singular stochastic control with state constraints in a portfolio selection problem. SIAM J. Control Optim. **45**(6), 2169–2206 (2007)
10. Crandall, M.G., Lions, P.L.: Viscosity solution of Hamilton–Jacobi equations. Trans. Am. Math. Soc. **277**, 1–42 (1983)
11. Czarna, I., Palmowski, Z.: De Finetti's dividend problem and impulse control for a two-dimensional insurance risk process. Stoch. Models **27**, 220–250 (2011)
12. De Finetti, B.: Su Un'impostazione Alternativa Della Teoria Collettiva Del Rischio. Trans. XVth Int. Cong. Actuar. **2**, 433–443 (1957)
13. Dickson, D.C.M., Waters, H.R.: Some optimal dividends problems. ASTIN Bull. **34**, 49–74 (2004)
14. Fleming, W.H., Soner, H.M.: Controlled Markov Processes and Viscosity Solutions. Springer, New York (1993)
15. Gerber, H.: Entscheidungskriterien für den zusammengesetzten Poisson-Prozeß. Mitt. Ver. Schweiz. Vers. Math **69**, 185–228 (1969)
16. Gerber, H.U., Shiu, E.S.W.: On the time value of ruin. N. Am. Actuar. J. **2**(1), 48–78 (1998)
17. Gerber, H.U., Shiu, E.S.W.: On the merger of two companies. N. Am. Actuar. J. **10**(3), 60–67 (2006)
18. Gerber, H.U., Lin, X.S., Yang, H.: A note on the dividends-penalty identity and the optimal dividend barrier. Astin Bull. **36**, 489–503 (2006)
19. Kushner, H.J., Dupuis, P.: Numerical Methods for Stochastic Control Problems in Continuous Time. Springer, New York (2001)
20. Kushner, H.J., Martins, L.F.: Numerical methods for stochastic singular control problems. SIAM J. Control Optim. **29**, 1443–1475 (1991)
21. Loeffen, R.L.: On optimality of the barrier strategy in de Finetti's dividend problem for spectrally negative Levy processes. Ann. Appl. Probab. **18**, 1669–1680 (2008)
22. Loeffen, R.L., Renaud, J.F.: De Finetti's optimal dividends problem with an affine penalty function at ruin. Insur. Math. Econ. **46**, 98–108 (2010)
23. Ly Vath, V., Pham, H., Villeneuve, S.: A mixed singular/switching control problem for a dividend policy with reversible technology investment. Ann. Appl. Probab **18**, 1164–1200 (2008)
24. Schmidli, H.: Stochastic Control in Insurance. Springer, New York (2008)
25. Souganidis, P.E.: Approximation schemes for viscosity solutions of Hamilton–Jacobi equations. J. Differ. Equ. **59**, 1–43 (1985)
26. Thonhauser, S., Albrecher, H.: Dividend maximization under consideration of the time value of ruin. Insur. Math. Econ. **41**, 163–184 (2007)

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.