Optimal Singular Dividend Problem Under the Sparre Andersen Model

Linlin Tian · Lihua Bai · Junyi Guo

Received: 20 January 2019 / Accepted: 3 October 2019 / Published online: 31 October 2019
© Springer Science+Business Media, LLC, part of Springer Nature 2019

Abstract
In this paper, we study the optimal dividend problem assuming that the underlying reserve process follows the Sparre Andersen model. In this model, there is no constant restriction on the dividend rates, i.e., the optimization problem is of singular type. In this case, the value function is no longer bounded and the associated Hamilton–Jacobi–Bellman equation is a variational inequality involving a first-order integro-differential operator and a gradient constraint. We prove the regularity properties for the value function by constructing strategies and show that the value function is a constrained viscosity solution of the associated Hamilton–Jacobi–Bellman equation. In addition, we prove that the value function is the upper semicontinuous envelope of the supremum for a class of subsolutions.

Keywords
Singular control · Optimal dividend · Hamilton–Jacobi–Bellman equation · Constrained viscosity solution · Viscosity supersolution · Viscosity subsolution

Mathematics Subject Classification 49L20 · 49L25 · 93E20

Communicated by Kok Lay Teo.

LiHua Bai
lhbai@nankai.edu.cn

Linlin Tian
1120150031@mail.nankai.edu.cn

Junyi Guo
jyguo@nankai.edu.cn

1 School of Mathematical Sciences, Nankai University, Tianjin 300071, China
1 Introduction

The dividend problem was first posed by De Finetti [1] at the International Congress of Actuaries in 1957. He proposed that the insurance company has the option to pay out dividends to its shareholders, and the way of measuring the performance of the company is to look at the maximum possible dividend paid during its lifetime. For the optimal dividend problem, many works model the underlying surplus process by a Brownian motion with drift (see, for example Asmussen and Taksar [2] and Guo et al. [3]). The dividend optimization problem under the compound Poisson model is studied in Azcué and Muler [4], Belhaj [5] and Gerber and Shiu [6]. If the compound Poisson claim process is replaced by a renewal process, the model is known as the Sparre Andersen risk model [7]. Under this model, Li and Garrido [8] computed an integro-differential equation for the Gerber–Shiu function in the case of a barrier strategy, and Albrecher et al. [9] calculated the moments of the expected discounted dividend payments under a barrier strategy. Moreover, Albrecher and Hartinger [10] showed that even in the case of Erlang(2) distributed interclaim times and exponentially distributed claim amounts a horizontal barrier strategy is not optimal anymore, as it can be outperformed by a strategy that depends on the time elapsed since the last claim. Consequently, the optimal dividend problem under the Sparre Andersen models has since been listed as an open problem that requires attention (see Albrecher and Thonhauser [11]).

In this paper, we study the finite-time optimal singular dividend problem under the Sparre Andersen model we mentioned before. For Markov process, one can associate a Hamilton–Jacobi–Bellman (HJB) equation to the stochastic optimal control problem by dynamic programming principle (DPP). But for our problem, the Sparre Andersen model is no longer Markovian. To overcome this difficulty, we use the backward Markovization technique. More precisely, we recast the problem in a Markovian framework with expanded dimension representing the time elapsed after the last claim. This means there are three variables in our HJB equation. Another difficulty is that the dividend process is not necessarily absolutely continuous with respect to the Lebesgue measure. This gives rise to a gradient constraint in the HJB equation. All the above difficulties make it impossible to explore the classical solution for our HJB equation.

Since it is hard to conjecture the existence of continuously differentiable solutions for our HJB equations, it is natural to invoke the notion of viscosity solution as done by Azcué and Muler [12], Benth et al. [13] and Seydel [14]. Thus, we will explore this problem in the framework of viscosity solution. For past development on viscosity solutions, see Crandall and Lions [15], Crandall et al. [16], Lions [17,18] and Soner [19]. For detailed exposition of viscosity solution and the related DPP on optimal stochastic control, see Fleming and Soner [20] and Yong and Zhou [21].

Under the Sparre Andersen risk model, Bai et al. [22] studied the optimal dividend and investment problem with a ceiling on dividend rates. They showed that the value function is the unique constrained viscosity solution of the corresponding HJB equation. Based on [22], we explore the optimal singular dividend problem under the Sparre Andersen model, i.e., the jump dividend is allowed in our model. The difference between this paper and [22] primarily consists of three aspects:
In the context of showing the continuity of the value function with respect to initial wealth, Bai et al. [22] introduced a penalty function while we consider this question by constructing admissible strategies. This also leads to a difference in the proof of continuity of the value function with respect to the time elapsed since the last claim.

We extended the verification theorem because jump dividend is allowed in our model.

The structure of the HJB equation and the boundary condition of the value function are changed compared to those of Bai et al. [22]. Based on the HJB equation and the boundary condition, we construct a viscosity subsolution and a supersolution of the HJB equation and we provide a candidate for the value function.

The paper is organized as follows. In Sect. 2, we establish the basic setting, formulate the problem and introduce the backward Markovization technique. In Sect. 3, we study the properties of the value function and prove the continuity of the value function in the temporal variable. In Sect. 4, we validate the dynamic programming principle (DPP) and characterize the value function as a constrained viscosity solution of the associated HJB equation. In Sect. 5, we construct a viscosity supersolution and subsolution and show that the value function is the upper semicontinuous envelope of the supremum for a class of subsolutions.

2 Model and Assumptions

Let $(\Omega, \mathbb{P}, \mathcal{F})$ be a probability space and $(\mathcal{F}_t)$ a given filtration satisfying the usual conditions. We consider a renewal counting process $N = \{N_t\}_{t \geq 0}$ on $(\Omega, \mathbb{P}, \mathcal{F})$. Denote $\{\sigma_n\}_{n \geq 0}$ as the jump times of the renewal counting process $N$, and denote $T_i = \sigma_i - \sigma_{i-1}, i = 1, 2, \ldots$ as the waiting times between claims. We also assume that $\{T_i\}_{i=1}^{\infty}$ are independent and identically distributed with the common distribution $F : \mathbb{R}_+ \to [0, 1]$, and that there exists an intensity function $\lambda : [0, \infty[ \to [0, \infty[$ such that $\bar{F}(t) := \mathbb{P}\{T_1 > t\} = \exp\{-\int_0^t \lambda(u)du\}$. In other words, $\lambda(t) = f(t)/\bar{F}(t), t \geq 0$, where $f$ is the common density function of the $T_i$’s. An important feature of the Sparre Andersen model is the “compound renewal process” that describes the total claim process $\sum_{i=1}^{N_t} U_i$, where $N_t$ is the renewal process representing the frequency of claims up to time $t$, and $\{U_i\}_{i=1}^{\infty}$ is a sequence of random variables representing the size of incoming claims. We assume that the $U_i$ are independent, identically distributed with the common distribution $G : \mathbb{R}_+ \to [0, 1]$, and are independent of $N$. Denote $Q_t = \sum_{i=1}^{N_t} U_i, t \geq 0$. Since the process $Q_t$ is non-Markovian in general (unless the counting process is a Poisson process), we cannot use the dynamic programming approach directly. Instead, we use the so-called backward Markovization technique that was used in [22]. Thus, we need to introduce a new process

$$W_t = t - \sigma_{N_t}, \quad t \geq 0,$$

of the time elapsed since the last jump. We observe that $0 \leq W_t \leq t \leq T$. It is known that the process $(t, Q_t, W_t), t \geq 0$ is a piecewise deterministic Markov process (see,
e.g., Rolski et al. [23]). Throughout this paper, we consider the following filtration
\( \mathcal{F}_t \) for any time \( s \), where \( \mathcal{F} = \mathcal{F}^N \vee \mathcal{F}^W, t \geq 0 \). Here, \( \mathcal{F}^N, \mathcal{F}^W \) denote the natural filtrations generated by processes \( N, W \), respectively, with the usual \( \mathbb{P} \)-augmentation such that it satisfies the usual hypotheses; see, e.g., Protter [24].

An important feature of the dynamic optimal control theory is to allow the starting point to be any time \( s \in [0, T] \). In our Sparre Andersen model, we consider the initial time \( s \in [0, T] \) with the initial elapsed time \( W_s = w \) instead of \( t = 0 \). We consider the regular conditional probability distribution \( \mathbb{P}_s w(\cdot) = \mathbb{P}[\cdot | W_s = w] \) on \( (\Omega, \mathcal{F}) \) and consider the “shifted” version of process \( (Q, W) \) on space \( (\Omega, \mathcal{F}, \mathbb{P}_s w, \mathcal{F}_s) \), where \( \mathcal{F}_s = (\mathcal{F}_t)_{t \geq s} \). Next, we restart the clock at time \( s \in [0, T] \) by defining the new counting process

\[ N_t^s := N_t - N_s, t \in [s, T] \]

Then, under \( \mathbb{P}_s w, N^s \) is a “delayed” renewal process, in the sense that while its waiting times \( T_s \) satisfy the usual hypotheses

\[ T_s \sim \text{Exp}(\lambda(w)) \]

for simplicity. We call a control strategy \( \pi \) an \( \mathbb{F} \)-adapted Markov process defined on \( (\Omega, \mathcal{F}, \mathbb{P}_s w) \). We will rely on the following standing assumptions.

**Assumption 2.1** (a) The insurance premium \( p \) and discount factor \( c \) are all positive constants.
(b) The distribution function \( F \) (of \( T_i \)'s) and \( G \) (of \( U_i \)'s) are continuous on \( [0, +\infty] \).

The common density function of \( F \) is denoted by \( f(t) \); there exists a continuous and bounded intensity function such that \( \lambda(t) = f(t) / F(t) > 0, t \in [0, T] \). There exists a constant \( \Lambda > 0 \) such that for all \( w \in [0, T] \), it holds that \( \lambda(w) \leq \Lambda \).

From now on, we consider the optimal dividend problem. For a given \( s \in [0, T] \) and any dividend policy \( L_t^s, w \), where \( L_t^s, w \) denotes the cumulative dividend from time \( s \) to time \( t \), the stochastic differential equation (SDE) of the wealth process \( X_t^{\pi, x} \) satisfies

\[ X_t^{\pi, x} = x + p(t - s) - Q_t^s, w - L_t^s, w. \]

Denote \( (X^{\pi, x, w}, W, L) = (X^{\pi, s, x, w}, W^s, w, L^s, w) \) for simplicity. We call a control strategy \( \pi = \{L_t\}_{t \geq 0} \) admissible, if the following hold true:

(a) It is \( \mathcal{F} \)-predictable, nondecreasing and càglàd.
(b) For any time \( t \geq 0 \), the process \( L_t \) satisfies

\[ L_t \leq x + p(t - s) - Q_t^s, w, \quad (1) \]

which means that the dividend process \( L_t \) cannot cause bankruptcy. Denote \( U_{ad}^s, w [s, T] \) as the set of all admissible control strategies with initial wealth \( x \) and elapsed time \( w \) since the last jump at time \( s \). From (1), we observe that the set of admissible strategies
is related to the initial capital $x$, which is quite different from [22]. For the given initial data $(s, x, w)$, we define the cost function by
\[
J(s, x, w; \pi) = \mathbb{E}_{sxw} \left[ \int_s^{(\tau_{\pi,x} \wedge T)+} e^{-c(t-s)} dL_t \right] := \mathbb{E} \left[ \int_s^{(\tau_{\pi,x} \wedge T)+} e^{-c(t-s)} dL_t | X_s = x, W_s = w \right],
\]
where $\tau_{\pi,x} = \inf \{ t \geq s : X_{t,x}^\pi < 0 \}$ denotes the ruin time of the insurance company, and $c > 0$ is the discounting factor. From now on, denote $X_{t}^\pi = X_{t,x}^\pi$ and $\tau^\pi = \tau_{\pi,x}$ for simplicity, if such notation does not cause confusion. $J(s, x, w; \pi)$ is the expected total discounted amount of dividends before ruin. Our objective is to find the optimal strategy to maximize the expectation of cumulative discounted dividends. The value function is defined by:
\[
V(s, x, w) = \sup_{\pi \in U_{sxw}[s,T]} J(s, x, w; \pi).
\]
The function $V(s, x, w)$ should be defined on $\{ (s, x, w) : 0 \leq s \leq T, x \geq 0, 0 \leq w \leq s \}$. Thus, we introduce several notations for simplicity. Denote
\[
D := \{ (s, x, w) | 0 \leq s \leq T, 0 \leq x, 0 \leq w \leq s \},
\]
\[
\mathcal{D} := \text{int}D = \{ (s, x, w) \in D | 0 < s < T, 0 < x, 0 < w < s \},
\]
\[
\mathcal{D}^* := \{ (s, x, w) \in D | 0 \leq s < T, 0 \leq x, 0 \leq w \leq s \}.
\]
Here, we note that $\mathcal{D}^*$ does not include the boundary of $s = T$.

**Remark 2.1** As the singular dividend policy is admissible in our model, it may have a jump at any time $t \leq T$. Thus, the optimal choice at time $T$ is to pay the entire wealth as dividends.

### 3 Basic Properties of the Value Function

In this section, we present several propositions to characterize a number of regularity properties of the value function $V(s, x, w)$.

**Proposition 3.1** **Under Assumption 2.1**, the optimal return function $V$ is well defined on $D$, and for all $x \geq 0, s \in [0, T], w \in [0, s]$, it holds that $x \leq V(s, x, w) \leq x + p(T - s)$.

**Proof** From $E \left[ \int_s^{(\tau_{\pi,x} \wedge T)+} e^{-c(t-s)} dL_t \right] \leq L_{\tau_{\pi,x} \wedge T}$, we observe that for any strategy $\pi \in U_{sxw}[s,T]$, it holds that $J(s, x, w; \pi) \leq \mathbb{E}_{sxw} \left[ L_{(\tau_{\pi,x} \wedge T)+} \right]$. By (1), we observe that $L_{(\tau_{\pi,x} \wedge T)+} \leq x + p(T - s)$. On the other hand, the lower bound of $V$ is trivial. ∎
Proposition 3.2  Under Assumption 2.1, for all \((s, x, w), (s + h, x, w) \in D, h > 0\), it holds that \(0 \leq V(s, x, w) - V(s + h, x, w) \leq 2ph\).

Proof  The proof of the first inequality is identical to that of Proposition 3.3 in Bai et al. [22]. We omit it here. Now, we prove the second inequality. For any strategy \(\pi \in U_{ad}^{x, w}[s, T]\), we construct \(\pi^h \in U_{ad}^{x, w}[s, T]\) as follows:

\[
\pi^h_t = \begin{cases} 1 & \{t > T - h\} \pi^t_t + 1 & \{t = T - h\} \pi^p_t + 1 & \{t < T - h\} \pi^{all}_t \end{cases},
\]

where \(\pi^{all}\) means using the entire wealth as dividend immediately and \(\pi^p\) means paying dividend at rate \(p\). Denote \(\tau^h\) as the ruin time of strategy \(\pi^h\) and \(L^h_t\) as the corresponding cumulative dividend payments of \(\pi^h\). Thus,

\[
J(s, x, w; \pi^h) = \mathbb{E}_{sxw}[\int_{t = (T-h)^+}^{\tau_h^T} e^{-c(t-s)} \, dL^h_t]
\]

\[
= \mathbb{E}_{sxw}[\int_{s}^{t = (T-h)^+} e^{-c(t-s)} \, dL^h_t : \tau^h > T - h] + \mathbb{E}_{sxw}[\int_{s}^{t = (T-h)^+} e^{-c(t-s)} \, dL^h_t : \tau^h \leq T - h] + \mathbb{E}_{sxw}[\int_{s}^{T-h} e^{-c(t-s)} \, dL^h_t : \tau^h \leq T - h].
\]

(2)

Note that \(\pi^h |_{[s, T-h]}\) can be regarded as an admissible strategy of \(U_{ad}^{x, w}[s + h, T]\). Thus, we obtain

\[
\mathbb{E}_{sxw}[\int_{s}^{T-h} e^{-c(t-s)} \, dL^h_t : \tau^h > T - h] + \mathbb{E}_{sxw}[\int_{s}^{T-h} e^{-c(t-s)} \, dL^h_t : \tau^h \leq T - h] \leq V(s + h, x, w).
\]

(3)

Substituting (3) into (2), and considering that on \(\{\tau^\pi > T - h\}, X^h_{(T-h)^+} = 0\), and all dividends of \(\pi^h\) in the interval \([T - h, T]\) are less than \(ph\), we obtain

\[
J(s, x, w; \pi^h) \leq V(s + h, x, w) + \mathbb{E}_{sxw}[\int_{(T-h)^+}^{T-h} e^{-c(t-s)} \, dL^h_t : \tau^h > T - h]
\]

\[
\leq V(s + h, x, w) + ph.
\]

(4)
As there is no difference between $\pi^h$ and $\pi$ on $[s, T-h]$, we obtain
\[ J(s, x, w; \pi) - J(s, x, w; \pi^h) = \mathbb{E}_{sxw} \left[ \int_{T-h}^{(\tau^\pi \wedge T) +} e^{-c(t-s)} dL_t^\pi - \int_{T-h}^{(\tau^h \wedge T) +} e^{-c(t-s)} dL_t^h : \tau^\pi > T-h \right]. \] (5)

As the dividend strategy $\pi$ cannot cause bankruptcy, we observe that
\[ \mathbb{E}_{sxw} \left[ \int_{(T-h)-}^{(\tau^\pi \wedge T) +} e^{-c(t-s)} dL_t^\pi : \tau^\pi > T-h \right] \leq \mathbb{E}_{sxw} \left[ ph + X_{T-h}^\pi : \tau^\pi > T-h \right]. \] (6)

By definition, the dividend $\pi^h$ will have a jump at time $T-h$ on $\{\tau^\pi > T-h\}$; thus,
\[ \mathbb{E} \left[ \int_{(T-h)-}^{(\tau^h \wedge T) +} e^{-c(t-s)} dL_t^h : \tau^\pi > T-h \right] \geq \mathbb{E}_{sxw} \left[ X_{T-h}^\pi : \tau^\pi > T-h \right]. \] (7)

Substituting (6) and (7) into (5), we obtain
\[ J(s, x, w; \pi) - J(s, x, w; \pi^h) \leq \mathbb{E}_{sxw} \left[ ph + X_{T-h}^\pi - X_{T-h}^\pi : \tau^\pi > T-h \right] \leq ph. \] (8)

Combining (4) and (8), we obtain that for any strategy $\pi \in U_{ad}^{x,w}[s, T]$,
\[ J(s, x, w; \pi) - V(s+h, x, w) = J(s, x, w; \pi) - J(s, x, w; \pi^h) + J(s, x, w; \pi^h) - V(s+h, x, w) \leq 2ph. \]

The proof is complete by taking the supremum over all strategies $\pi \in U_{ad}^{x,w}[s, T]$. \hfill \Box

**Proposition 3.3** Under Assumption 2.1, the following holds:

(a) For any $x_1 \geq x_2 \geq 0$, $V(s, x_1, w) - V(s, x_2, w) \geq x_1 - x_2$;
(b) For any compact set $\mathcal{D}$, where $\mathcal{D}$ is a subset of $D$, the mapping $x \mapsto V(s, x, w)$ is uniformly continuous on $\mathcal{D}$.

**Proof** (a) For the first part, it is trivial to note that for any strategy $\pi_2 \in U_{ad}^{x_2,w}[s, T]$, we can construct $\pi_1$ such that $\pi_1$ has a jump dividend $x_1 - x_2$ at time $s$ and then follows the strategy $\pi_2$. Thus, we observe that $V(s, x_1, w) \geq J(s, x_1, w; \pi_1) = J(s, x_2, w; \pi_2) + x_1 - x_2$. As $\pi_2 \in U_{ad}^{x_2,w}[s, T]$ is arbitrary, taking the supremum over $U_{ad}^{x_2,w}[s, T]$, we obtain $V(s, x_1, w) - V(s, x_2, w) \geq x_1 - x_2$.

(b) Here, we borrow several arguments from Scheer and Schmidli [25]. Suppose that $h \geq 0$, and let $\pi$ be a strategy with initial capital $x$. Now, we construct another strategy $\tilde{\pi}$ that pays no dividend until some stopping time $\tau^h$ to be defined below. The
strategy \( \tilde{\pi} \in U_{\text{ad}}^{x-h,w} [s, T] \) is adjusted at time \( t^h \) so that two surplus processes \( X_t^{\pi,x} \) and \( X_t^{\pi,x-h} \) coincide from \( t^h \) on. Let \( t^h := \inf \{ t \geq s : L_t^\pi \geq h \} \) be the first time that \( h \) is compensated by paying dividends with strategy \( \pi \). We define the strategy \( \tilde{\pi} \) as follows:

\[
L_t^{\pi} = (L_t^\pi - h)^+ + \mathbf{1}_{\{t^h \leq t \leq T\}} \mathbf{1}_{\{t < \tilde{\tau}\}} + X_T^{\pi,x-h} \mathbf{1}_{\{t = T < t^h \land \tilde{\tau}\}},
\]

where \( (L_t^\pi - h)^+ = \max \{0, L_t^\pi - h\} \) and \( \tilde{\tau} = \inf \{ t \geq s : X_t^{\pi,x-h} < 0 \} \) is the ruin time of \( X_t^{\pi,x-h} \). Note that \( \tilde{\pi} \) is admissible. From the definition of \( \tilde{\pi} \), we obtain

\[
\mathbb{E}_{s,x,w} \left[ \int_s^{(\pi \land T)^+} e^{-c(t-s)} dL_t^\pi - \int_s^{t^h \land T} e^{-c(t-s)} dL_t^\pi : S \leq \tilde{\tau} \right] \\
\leq L_{(S \land T)^+}^\pi - L_{(S \land T)^+}^\pi \leq h. \tag{9}
\]

\[
\mathbb{E}_{s,x,w} \left[ \int_s^{(\pi \land T)^+} e^{-c(t-s)} dL_t^\pi - \int_s^{(\tilde{\tau} \land T)^+} e^{-c(t-s)} dL_t^{\tilde{\pi}} : S > \tilde{\tau}, \tilde{\tau} = \pi^\pi \right] \leq h. \tag{10}
\]

As for any strategy \( \pi \), the corresponding cost function is bounded, we obtain

\[
\mathbb{E}_{s,x,w} \left[ \int_s^{(\pi \land T)^+} e^{-c(t-s)} dL_t^\pi - \int_s^{(\tilde{\tau} \land T)^+} e^{-c(t-s)} dL_t^{\tilde{\pi}} : \tau^h > \tilde{\tau}, \tilde{\tau} \neq \pi^\pi \right] \\
\leq (x + pT) \int_0^\infty \mathbb{P} \left\{ \Delta X_t^{\pi,x-h} \in (y, y + h) | X_t^{\pi,x-h} = y \right\} F_{X_t^{\pi,x-h}}(dy) \\
= (x + pT) \int_0^\infty [G(y + h) - G(y)] F_{X_t^{\pi,x-h}}(dy), \tag{11}
\]

where \( G \) denotes the common distribution function of the claim size \( U_t \)’s. As \( G \) is uniformly continuous, we observe that for any \( \varepsilon > 0 \), there exists a constant \( \delta_1(x) > 0 \) such that for all \( h \leq \delta_1(x) \), we have \( G(y + h) - G(y) \leq \frac{\varepsilon}{2(x + pT)} \). Combining this with (9), (10) and (11), we obtain that for any \( \varepsilon > 0 \), there exists a \( \delta(x) = \min \{ \frac{\varepsilon}{2}, \delta_1(x) \} > 0 \) such that for all \( h < \delta(x) \) and any strategy \( \pi \in U_{\text{ad}}^{x,w} [s, T] \), it holds that

\[
J(s, x, w; \pi) \leq J(s, x - h, w; \tilde{\pi}) + \varepsilon \leq V(s, x - h, w) + \varepsilon.
\]

As \( \pi \) is arbitrary, we obtain that for a sufficiently small \( h \), \( V(s, x, w) - V(s, x - h, w) \leq \varepsilon \). This shows that \( V \) is continuous with respect to \( x \). Thus, for all compact sets, \( V \) is uniformly continuous with respect to \( x \).

\[\square\]

**Proposition 3.4** Under Assumption 2.1, for any \( h > 0 \) and \( 0 \leq s < s + h \leq T \), it holds that

(a) \( V(s + h, x, w + h) - V(s, x, w) \leq \left[ 1 - \exp \left\{ -ch - \int_w^{w+h} \lambda(u) du \right\} \right] V(s + h, x, w + h) \).
(b) \( V(s, x, w + h) - V(s, x, w) \leq 2ph + \left[1 - \exp \left\{ -ch - \int_{w}^{w+h} \lambda(u)du \right\} \right] V(s + h, x, w + h). \)

**Proof**  
(a) For all strategies \( \pi \in U_{\text{ad}}^{\pi, s, w + h}[s + h, T] \), \( \tilde{\pi}^{h} \in U_{\text{ad}}^{\pi, s, w}[s, T] \) is defined so that on the set \( \{ T_{1}^{s, w} > h \} \) the insurance company pays dividends at rate \( p \) during time \( [s, s + h] \), and then follows the strategy \( \pi \) during \( t \in [s + h, T] \); on set \( \{ T_{1}^{s, w} \leq h \} \), \( \tilde{\pi}^{h} \) pays no dividends before time \( T \) and then distributes the entire wealth as dividends at time \( T \) if \( X_{T}^{\tilde{\pi}^{h}} > 0 \). The rest of the proof is exactly the same as the proof of Proposition 5.1 of Bai et al. [22]. Depending on whether there are claims in the time interval \([s, s + h]\), we can deduce that

\[
V(s, x, w) \geq \exp \left\{ -ch - \int_{w}^{w+h} \lambda(u)du \right\} V(s + h, x, w + h). \tag{12}
\]

(b) We calculate the following directly:

\[
V(s, x, w + h) - V(s, x, w) = V(s, x, w + h) - V(s + h, x, w + h) + V(s + h, x, w + h) - V(s, x, w). \tag{13}
\]

Combining Proposition 3.2 and (12) with (13), we obtain the desired result (b). \( \square \)

**Proposition 3.5**  
Under Assumption 2.1, \( \lim_{h \downarrow 0} [V(s, x, w) - V(s, x, w + h)] = 0. \)

**Proof**  
From Proposition 3.4, we note that \( \lim_{h \downarrow 0} [V(s, x, w) - V(s, x, w + h)] \geq 0 \). We only need to prove the opposite inequality. By a direct calculation and Proposition 3.2, we obtain

\[
V(s, x, w) - V(s, x, w + h) \leq V(s, x, w) - V(s + h, x, w + h). \tag{14}
\]

For any strategy \( \pi \in U_{\text{ad}}^{s, w}[s, T] \), depending on whether there are claims between \([s, s + h]\), we obtain

\[
\mathbb{E}_{s, x, w} \left[ \int_{s}^{(\tau^{\pi} \wedge T)^{+}} e^{-c(t-s)} dL_{t}^{\pi} \right] - V(s + h, x, w + h) \\
= \mathbb{E}_{s, x, w} \left[ \int_{s}^{(\tau^{\pi} \wedge T)^{+}} e^{-c(t-s)} dL_{t}^{\pi} \mathbb{P}(T_{1}^{s, w} \leq h) \right] \mathbb{P}(T_{1}^{s, w} \leq h) \\
+ \mathbb{E}_{s, x, w} \left[ \int_{s}^{(\tau^{\pi} \wedge T)^{+}} e^{-c(t-s)} dL_{t}^{\pi} \mathbb{P}(T_{1}^{s, w} > h) \right] \mathbb{P}(T_{1}^{s, w} > h) \\
- V(s + h, x, w + h). \tag{15}
\]

For the first term of the right-hand side of (15), because \( \mathbb{P}(T_{1}^{s, w} \leq h) = 1 - \exp[-\int_{w}^{w+h} \lambda(u)du] \), we know that for a sufficiently small \( h \) and any compact set \( \mathcal{D} \), there exists a constant \( Q_{1}(\mathcal{D}) > 0 \) such that
\[\mathbb{E}_{s,w} \left[ \int_{s}^{(\tau^{\pi} \land T) +} e^{-c(t-s)} dL^{\pi}_{t} \mid T_{1}^{s,w} \leq h \right] \mathbb{P}(T_{1}^{s,w} \leq h) \leq Q_{1}(\mathcal{X})h, \quad (16)\]

which means that for all \(\varepsilon > 0\), there exists a constant \(\delta_{1} > 0\), such that for all \(0 < h < \delta_{1}\) and all strategies \(\pi \in U_{ad}^{x,w}[s, T],\)

\[\mathbb{E}_{s,w} \left[ \int_{s}^{(\tau^{\pi} \land T) +} e^{-c(t-s)} dL^{\pi}_{t} \mid T_{1}^{s,w} > h \right] \mathbb{P}(T_{1}^{s,w} > h) \leq V(s + h, x, w + h) - V(s + h, x, w + h). \quad (18)\]

As on \(\{T_{1}^{s,w} > h\}\), there is no claim on \([s, s + h]\), and it is clear that \(\tau^{\pi} > s + h\). Thus, we have that

\[\mathbb{E}_{s,w} \left[ \int_{s}^{(\tau^{\pi} \land T) +} e^{-c(t-s)} dL^{\pi}_{t} \mid T_{1}^{s,w} > h \right] \leq \mathbb{E}_{s,w} \left[ x + p h - X_{s+h}^{\pi} + V(s + h, X_{s+h}^{\pi}, w + h) \right] \cdot (19)\]

The inequality of (19) is based on the fact that \(L_{s+h}^{\pi} = x + p h - X_{s+h}^{\pi} \) on \(\{T_{1}^{s,w} \geq h\}\). Then, we can obtain

\[\mathbb{E}_{s,w} \left[ x + p h - X_{s+h}^{\pi} + V(s + h, X_{s+h}^{\pi}, w + h) \right] \cdot (20)\]

As \(0 \leq X_{s+h}^{\pi} \leq x + p h\) on \(\{T_{1}^{s,w} > h\}\), conditioning on whether \(X_{s+h}^{\pi}\) belongs to \([0, x]\), we can see that on \(\{0 \leq X_{s+h}^{\pi} \leq x, T_{1}^{s,w} > h\}\),

\[V(s + h, X_{s+h}^{\pi}, w + h) - V(s + h, x, w + h) \leq -(x - X_{s+h}^{\pi}),\]

and on \(\{x < X_{s+h}^{\pi} \leq x + p h, T_{1}^{s,w} > h\}\),

\[V(s + h, X_{s+h}^{\pi}, w + h) - V(s + h, x, w + h) \leq V(s + h, x + p h, w + h) - V(s + h, x, w + h).\]
Thus, we can see that
\[
\mathbb{E}_{s,x,w} \left[ x + ph - X_{s+h}^\pi + V(s+h, X_{s+h}^\pi, w+h) - V(s+h, x, w+h)|T_1^{s,w} > h \right] \\
\leq \; ph + V(s+h, x + ph, w+h) - V(s+h, x, w+h).
\]
(21)

Substituting (21) into (20) and combining this with (18), we obtain
\[
\mathbb{E}_{s,x,w} \left[ \int_s^{(\pi \wedge T)+} e^{-c(t-s)} dL_t^\pi |T_1^{s,w} > h \right] \mathbb{P}(T_1^{s,w} > h) - V(s+h, x, w+h) \\
\leq \; ph + V(s+h, x + ph, w+h) - V(s+h, x, w+h).
\]
(22)

By Proposition 3.3, \( \lim_{h \downarrow 0} V(s+h, x + ph, w+h) - V(s+h, x, w+h) = 0. \)
Thus, for all \( \varepsilon > 0 \), there exists a constant \( \delta_2 > 0 \) such that for all \( 0 < h < \delta_2 \),
\( V(s+h, x + ph, w+h) - V(s+h, x, w+h) \leq \frac{\varepsilon}{3} \) holds. Setting \( \delta_3 = \min\{\delta_2, \frac{\varepsilon}{3p}\} \),
for all \( 0 < h < \delta_3 \), substituting \( h \) in (22), we obtain
\[
\mathbb{E}_{s,x,w} \left[ \int_s^{(\pi \wedge T)+} e^{-c(t-s)} dL_t^\pi |T_1^{s,w} > h \right] \mathbb{P}(T_1^{s,w} > h) - V(s+h, x, w+h) \leq \frac{2\varepsilon}{3}.
\]
(23)

From (15), (17) and (23), we know that for all \( \varepsilon > 0 \), there exists a constant \( \delta_4 := \min\{\delta_1, \delta_3\} \) such that for all \( 0 < h < \delta_4 \) and all strategies \( \pi \in U_ad^{x,w}[s, T] \), it holds that
\[
\mathbb{E}_{s,x,w} \left[ \int_s^{(\pi \wedge T)+} e^{-c(t-s)} dL_t^\pi \right] - V(s+h, x, w+h) \leq \varepsilon.
\]
As \( \pi \in U_ad^{x,w}[s, T] \) is arbitrary, we have
\[
V(s, x, w) - V(s+h, x, w+h) \leq \varepsilon.
\]
(24)

From (14) and (24), we obtain \( \lim_{h \downarrow 0}[V(s, x, w) - V(s, x, w+h)] \leq 0. \)

**Corollary 3.1** The value function \( V \) is continuous on \( D \) and uniformly continuous on any compact set \( \mathcal{D} \).

**Proof** The proof is obtained directly from the above propositions. \( \square \)

**4 Dynamic Programming Principle and the HJB equation**

Now, we state the dynamic programming principle (DPP) for our optimization problem. We can refer to Soner [20] and Pham [26] for the derivation from the DPP to the HJB equation. For any strategy \( \pi \in U_ad^{x,w}[s, T] \), denote \( R_t^\pi = R_t^{\pi,x,s,w} = \ldots \)
We call $v$ a "constrained viscosity solution" of (26) on $\mathcal{D}^*$ if it is both a viscosity subsolution on $\mathcal{D}^*$ and a viscosity supersolution on $\mathcal{D}$.

Let $O \subseteq \mathcal{D}^*$ be a subset such that $\partial_T O := \{(T, y, v) \in \partial O \} \neq \emptyset$ and $\text{cl}(O)$ denote the closure of $O$. Denote $\text{USC}(O)$ as the set of all upper semicontinuous functions on $O$ and $\text{LSC}(O)$ as the set of all lower semicontinuous functions on $O$.

(a) We call $v \in \text{USC}(O)$ a viscosity subsolution of (26) on $O$ if $v(T, y, v) \leq y$ for $(T, y, v) \in \partial_T O$; for any $(s, x, w) \in O$ and $\varphi \in \mathcal{C}^{1,1}(\text{cl}(O))$ such that $0 = [v - \varphi](s, x, w) = \max_{(t, y, v) \in O} [v - \varphi](t, y, v)$, it holds that $\max \{1 - v_x, \mathcal{L}[\varphi]\}(s, x, w) \leq 0$.

(b) We call $v \in \text{LSC}(O)$ a viscosity supersolution of (26) on $O$ if $v(T, y, v) \geq y$ for all $(T, y, v) \in \partial_T O$; for any $(s, x, w) \in O$ and $\varphi \in \mathcal{C}^{1,1}(\text{cl}(O))$ such that $0 = [v - \varphi](s, x, w) = \min_{(t, y, v) \in O} [v - \varphi](t, y, v)$, it holds that $\max \{1 - v_x, \mathcal{L}[\varphi]\}(s, x, w) \geq 0$.

In particular, we call $u$ a "constrained viscosity solution" of (26) on $\mathcal{D}^*$ if it is both a viscosity subsolution on $\mathcal{D}^*$ and a viscosity supersolution on $\mathcal{D}$.
There is an equivalent formulation of viscosity solutions; the proof of equivalence of definitions is standard (e.g., see Benth et al. [13] and Awatif [27]). In this paper, we use both definitions interchangeably. Now, we introduce the alternative definition of viscosity solutions. Given a continuously differentiable function $\varphi$ and a continuous function $u$, we define the operator

$$\mathcal{L}[u, \varphi](s, x, w) := [-cu + \varphi_x + p\varphi_x + \varphi_w](s, x, w) + \lambda(w) \left[ \int_0^x u(s, x - \alpha, 0) dG(\alpha) - u(s, x, w) \right].$$

**Definition 4.2** Let $v \in \text{USC}(O)$; we call $v$ a viscosity subsolution of (26) on $O$ if $v(T, y, v) \leq y$ for $(T, y, v) \in \partial T O$; for any $(s, x, w) \in O$, $\varphi \in C^{1,1}(\text{cl}(O))$ such that

$$0 = [v - \varphi](s, x, w) = \max_{(t, y, v) \in O} [v - \varphi](t, y, v),$$

it holds that $\max\{1 - \varphi_x, \mathcal{L}[v, \varphi]\}(s, x, w) \geq 0$.

Let $v \in \text{LSC}(O)$; we call $v$ a viscosity supersolution of (26) on $O$ if $v(T, y, v) \geq y$ for all $(T, y, v) \in \partial T O$; for any $(s, x, w) \in O$ and $\varphi \in C^{1,1}(\text{cl}(O))$ such that

$$0 = [v - \varphi](s, x, w) = \min_{(t, y, v) \in O} [v - \varphi](t, y, v),$$

it holds that $\max\{1 - \varphi_x, \mathcal{L}[v, \varphi]\}(s, x, w) \leq 0$.

**Theorem 4.2** Assume that Assumption 2.1 holds. Then, the value function is a constrained viscosity solution of the HJB equation (26) on $\mathcal{D}^*$. 

**Proof** We omit the proof of $V$ being a viscosity supersolution since the proof is similar to that in Azcue and Muler [4] and Bai et al. [22].

Now, we prove that $V$ is the viscosity subsolution of the HJB equation on $\mathcal{D}^*$. If we assume the contrary, then there exists a point $(s, x, w) \in \mathcal{D}^*$ and $\psi^0 \in C^{1,1}(D)$ such that

$$0 = [V - \psi^0](s, x, w) = \max_{(t, y, v) \in \mathcal{D}^*} [V - \psi^0](t, y, v),$$

but $\max\{1 - \psi^0_x, \mathcal{L}(\psi^0)\}(s, x, w) = -2\eta < 0$, where $\eta > 0$ is a constant and $\mathcal{L}$ is the first-order integro-differential operator. Now, we show that there exists a function $\psi \in C^{1,1}(\mathcal{D}^*)$ and constants $\varepsilon > 0$ and $\rho > 0$ such that

$$\mathcal{L}[\psi](t, y, v) \leq -\varepsilon c, (t, y, v) \in \text{cl}(B_\rho(s, x, w) \cap \mathcal{D}^*) \setminus \{t = T\};$$

$$1 - \psi_x(t, y, v) \leq -\varepsilon c, (t, y, v) \in \text{cl}(B_\rho(s, x, w) \cap \mathcal{D}^*) \setminus \{t = T\};$$

$$V(t, y, v) \leq \psi(t, y, v) - \varepsilon, (t, y, v) \in \partial B_\rho(s, x, w) \cap \mathcal{D}^*,$$

where $B_\rho(R)$ denotes the open sphere centered at $R$ with radius $\rho$ and the notation $\text{cl}$ denotes the closure of sets.
Case I If \( x > 0 \), then we define \( \psi(t, y, v) := \psi^0(t, y, v) + \eta \frac{(t-s)^2 + (y-v)^2 + (v-w)^2}{\lambda(w)(x^2 + w^2)^2} \). By simple calculation, we have \( \max\{1 - \psi_x, \mathcal{L}[\psi]\}(s, x, w) \leq -\eta < 0 \). By the continuity of \( \psi_x \) and \( \mathcal{L}[\psi] \), we know that there exists a constant \( \rho > 0 \), such that

\[
\max\{1 - \psi_x, \mathcal{L}[\psi]\}(t, y, v) \leq -\frac{\eta}{2} < 0,
\]

\( (t, y, v) \in \text{cl}(B_\rho(s, x, w) \cap \mathcal{D}^*) \{t = T\} \).

Note that for all \( (t, y, v) \in \partial B_\rho(s, x, w) \cap \mathcal{D}^* \), the following holds:

\[
V(t, y, v) \leq \psi(t, y, v) - \frac{\eta \rho^4}{\lambda(w)(x^2 + w^2)^2}.
\]

If we choose \( \varepsilon = \min\left\{\frac{\eta \rho^4}{\lambda(w)(x^2 + w^2)^2}, \frac{\eta}{2c}\right\} \), then we obtain (27).

Case II If \( x = 0 \), then

\[
\max\left\{1 - \psi^0_x, \mathcal{L}[\psi^0]\right\}(s, 0, w) = \max\left\{1 - \psi^0_x, -c\psi^0 + \psi^0_x + p\psi^0_y + \psi^0_w - \lambda(w)\psi^0\right\}(s, 0, w) \leq -2\eta < 0.
\]

Define \( \psi(t, y, v) := \psi^0(t, y, v) + \eta [(t-s)^2 + y^2 + (v-w)^2] \). If we denote \( \varepsilon = \min\left\{\frac{\eta}{2c}, \eta \rho^2\right\} \) and perform a calculation similar to that of Case I, then we can show that (27) still holds. In the following, we will argue that (27) leads to a contradiction.

For any strategy \( \pi \in U_{ad}^{x,w}[s, T] \), denote \( R^x,t,w = (t, X^x,t,w, W^x,t,w) \). Additionally, we denote \( R_t = R^x,t,w \) for simplicity. Define \( \tau_\rho := \inf\{t > s : R_t \notin \text{cl}(B_\rho(s, x, w) \cap \mathcal{D}^*)\} \) and \( \theta := \tau_\rho \wedge T_s^x,w \). Applying Itô’s formula to \( e^{-c(t-s)}\psi(R_t) \), we get

\[
\psi(s, x, w) \geq \mathbb{E}_{s,x,w}\left[e^{-c(\theta-s)}\psi(\theta, X_\theta, W_\theta)\right] + \mathbb{E}_{s,x,w}\left[\int_s^\theta e^{-c(t-s)}c\varepsilon dt\right] + \mathbb{E}_{s,x,w}\left[\int_s^\theta e^{-c(t-s)}dL_t\right].
\]

(28)

On \( \{\tau_\rho < T_s^x,w, (\theta, X_\theta, W_\theta) \neq (\theta, X_{\theta+}, W_\theta)\} \), the wealth process jumps out of \( \text{cl}(B_\rho(s, x, w) \cap \mathcal{D}^*) \) due to dividend payments at time \( \theta \). There exists a random variable \( \nu \in [0, 1] \) and a point \( (\theta, X_\nu, W_\nu) \in \partial B_\rho(s, x, w) \cap \mathcal{D}^* \) such that \( X_\nu = X_\theta - \nu[L_{\theta+} - L_\theta] \). From (27), we observe that

\[
\psi(\theta, X_\nu, W_\nu) \geq V(\theta, X_\nu, W_\nu) + \varepsilon.
\]

\[
\psi(\theta, X_\theta, W_\theta) - \psi(\theta, X_\nu, W_\nu) \geq (c\varepsilon + 1)(X_\theta - X_\nu) = (c\varepsilon + 1)\nu(L_{\theta+} - L_\theta).
\]

(29)

From Proposition 3.3, we obtain

\[
V(\theta, X_\nu, W_\nu) \geq V(\theta, X_{\theta+}, W_\theta) + (1 - \nu)(L_{\theta+} - L_\theta).
\]

(30)
Combining (29) with (30), we observe that on \( \{ \tau_\rho < T_1^{s, w}, (\theta, X_\theta, W_\theta) \neq (\theta, X_{\theta+}, W_\theta) \} \),
\[
\psi(\theta, X_\theta, W_\theta) \geq V(\theta, X_{\theta+}, W_\theta) + \varepsilon + (L_{\theta+} - L_\theta).
\] (31)

On the other hand, on \( \{ \tau_\rho < T_1^{s, w}, (\theta, X_\theta, W_\theta) = (\theta, X_{\theta+}, W_\theta) \} \), \( (\theta, X_\theta, W_\theta) \in \partial B_\rho(s, x, w) \cap \mathcal{D}^* \), then we have
\[
\psi(\theta, X_\theta, W_\theta) \geq V(\theta, X_\theta, W_\theta) + \varepsilon = V(\theta, X_{\theta+}, W_\theta) + \varepsilon.
\] (32)

Similarly, in the case of \( \{ \tau_\rho \geq T_1^{s, w}, X_{T_1^{s, w}} < 0 \} \), we have
\[
\psi(\theta, X_\theta, W_\theta) = V(\theta, X_\theta, W_\theta) = 0.
\] (33)

In the case of \( \{ \tau_\rho \geq T_1^{s, w}, X_{T_1^{s, w}} \geq 0 \} \), we have
\[
\psi(\theta, X_\theta, W_\theta) \geq V(\theta, X_\theta, W_\theta) \geq V(\theta+, X_{\theta+}, W_{\theta+}) + L_{\theta+} - L_\theta.
\] (34)

Combining (31), (32), (33) with (34), we obtain
\[
\mathbb{E}_{s, x, w} \left[ e^{-c(\theta-s)} \psi(\theta, X_\theta, W_\theta) + \int_s^{\theta} e^{-c(t-s)} dL_t \right]
\geq \mathbb{E}_{s, x, w} \left[ e^{-c(\theta-s)} V(\theta, X_{\theta+}, W_\theta) + \int_s^{\theta+} e^{-c(t-s)} dL_t : \tau_\rho \geq T_1^{s, w} \right]
+ \mathbb{E}_{s, x, w} \left[ e^{-c(\theta-s)} V(\theta, X_{\theta+}, W_\theta) + \int_s^{\theta+} e^{-c(t-s)} dL_t + \varepsilon e^{-c(\theta-s)} : \tau_\rho < T_1^{s, w} \right].
\] (35)

Combining (28) with (35), we observe that
\[
\psi(s, x, w) \geq \mathbb{E}_{s, x, w} \left[ e^{-c(\theta-s)} V(\theta, X_{\theta+}, W_\theta) + \int_s^{\theta+} e^{-c(t-s)} dL_t \right]
+ \varepsilon - \varepsilon \mathbb{E}_{s, x, w} \left[ e^{-c(\theta-s)} : \tau_\rho \geq T_1^{s, w} \right]
\geq \mathbb{E}_{s, x, w} \left[ e^{-c(\theta-s)} V(\theta, X_{\theta+}, W_\theta) + \int_s^{\theta+} e^{-c(t-s)} dL_t \right]
+ \varepsilon \mathbb{E}_{s, x, w} \left[ 1 - e^{-c(T_1^{s, w}-s)} \right].
\] (36)

For \( h > 0 \), \( \theta + h \) is a stopping time. Recalling the dynamic programming principle (25) and the continuity of \( V \), letting \( h \downarrow 0 \), we obtain
\[ V(s, x, w) = \sup_{\pi \in U^i_{ad}[s, T]} \mathbb{E}_{x, w}^\pi \left[ \int_s^{\theta^+} e^{-c(t-s)} dL_t \right] + e^{-c(\theta-s)} V(\theta, X_{\theta^+}, W_{\theta}). \]  

(37)

As \( \mathbb{P}(T^s, w > s) = 1 \) and \( V(s, x, w) = \psi(s, x, w) \), we observe that (37) contradicts (36).

**Remark 4.1** From the aspect of surplus modeling, the main difference between our paper and [4,28–35] is that they study the dividend optimization of Markov process while we study the non-Markovian process. From the aspect of techniques, the idea of constrained viscosity solution is not only different from the pre-mentioned papers [28–35] in which they explore the smooth solution for the HJB equation but also different from Azcue and Muler [4] in which the notion of viscosity solution is also used.

## 5 The Candidate for the Value Function

In this section, we provide a candidate for the value function. Before we state the theorems, we introduce several notations.

**Definition 5.1** If \( u : D \to \mathbb{R}^3 \), define

\[ u^*(s, x, w) = \limsup_{r \downarrow 0} \sup \{ u(t, y, v) : (t, y, v) \in D \text{ and } \sqrt{|t-s|^2 + |y-x|^2 + |v-w|^2} \leq r \}. \]

We call \( u^* \) the upper semicontinuous envelope of \( u \). \( u^* \) is the smallest upper semicontinuous function satisfying \( u \leq u^* \).

Denote \( d_{\mathcal{D}}(s, x, w) \) the distance between \((s, x, w)\) and the boundary of \( \mathcal{D} \), which means

\[ d_{\mathcal{D}}(s, x, w) = (T - s) \wedge w \wedge \frac{\sqrt{2}}{2} (s - w) \wedge x. \]  

(38)

Recall the definition \( \mathcal{D} = \{(s, x, w) \in D | 0 < s < T, 0 < x, 0 < w < s\} \). We observe that \( d_{\mathcal{D}}(s, x, w) \leq \frac{T}{2 + \sqrt{2}} \).

From now on, denote \( M_1 := \left(\frac{\sqrt{2}}{2} + 1 + 2p\right)T + \frac{T}{2 + \sqrt{2}} \) and \( M_2 := -\left(\frac{\sqrt{2}}{2} + 1 + (c + A)\frac{2T}{2 + \sqrt{2}}\right)T \) for simplicity. In what follows, we construct a viscosity supersolution of (26) on \( \mathcal{D} \).

**Theorem 5.1** Define

\[ \overline{V}(s, x, w) = x + d_{\mathcal{D}}(s, x, w) + N_1(T - s), \]  

(39)

where the constant \( N_1 = \frac{\sqrt{2}}{2} + 1 + 2p \), and \( d_{\mathcal{D}} \) is defined in (38). Then, \( \overline{V} \) is a viscosity supersolution of (26) on \( \mathcal{D} \), and for all \((s, x, w) \in D, x + M_2 \leq \overline{V}(s, x, w) \leq x + M_1 \).
**Proof** Note that for all \((s, x, w) \in \mathcal{D}, d_{\mathcal{D}}(s, x, w) \leq \frac{T}{2+\sqrt{2}}\). Thus, it is obvious that for all \((s, x, w) \in D, x + M_2 \leq \overline{V}(s, x, w) \leq x + M_1\). Denote \(R := (s, x, w)\) for simplicity. Consider a function \(\phi \in C^{1,1}(\mathcal{D})\) such that \(\overline{V} - \phi\) attains its minimum at \(R\). Now, we discuss various cases.

Case I If \(\frac{\sqrt{2}}{2}(s - w) < w \land x \land (T - s)\), we observe that \(d_{\mathcal{D}}(s, x, w) = \frac{\sqrt{2}}{2}(s - w)\) near point \(R\); this leads to \((\phi_x, \phi_s, \phi_w)(R) = (1, \frac{\sqrt{2}}{2} - N_1, -\frac{\sqrt{2}}{2})\). Thus, we obtain

\[
\mathcal{L}[\overline{V}, \phi](R) = \left[-(c + \lambda(w))\overline{V} + \phi_s + p\phi_x + \phi_w\right](R) + \lambda(w) \int_0^x \overline{V}(s, x - u, 0)dG(u) < 0.
\]

Case II If \(w < \frac{\sqrt{2}}{2}(s - w) \land x \land (T - s)\), then \(\phi_w(R) = 1, \phi_s(R) = -N_1, \phi_x(R) = 1\). Similar to Case I, we can deduce that \(\max\{1 - \phi_x, \mathcal{L}[(\overline{V}, \phi)](R) \leq -N_1 + p + 1 < 0\). Similarly, we can verify that \(\overline{V}\) is a viscosity supersolution of (26) in other cases of \(\mathcal{D}\).

In what follows, we provide a subsolution of (26) on \(\mathcal{D}^*\).

**Theorem 5.2** Define

\[
\overline{V}(s, x, w) = x + d_{\mathcal{D}}(s, x, w) - N_2(T - s), \tag{40}
\]

where the constant \(N_2 = \frac{\sqrt{2}}{2} + 1 + (c + \Lambda)\frac{2T}{2+\sqrt{2}}\) and \(d_{\mathcal{D}}(s, x, w)\) is defined in (38). Then, \(\overline{V}\) is a viscosity subsolution of (26) on \(\mathcal{D}^*\), and for all \((s, x, w) \in D, x + M_2 \leq \overline{V}(s, x, w) \leq x + M_1\).

**Proof** It is obvious that for all \((s, x, w) \in D, x + M_2 \leq \overline{V}(s, x, w) \leq x + M_1\). Now, we show that \(\overline{V}\) is a viscosity subsolution of (26) on \(\mathcal{D}^*\). For a fixed point \(R := (s, x, w) \in D^*\), consider the function \(\varphi \in C^{1,1}(D)\) such that \(\overline{V} - \varphi\) reaches its maximum at \(R\).

Case I If \(x > \frac{T}{2+\sqrt{2}}\). First, we note that for all \(x > \frac{T}{2+\sqrt{2}}, x \leq \max\{T - s, w, \frac{\sqrt{2}}{2}(s - w)\}\), which means that for all \(x > \frac{T}{2+\sqrt{2}}, \varphi_x(R) = 1\). Obviously, this leads to \(\max\{1 - \varphi_x, \mathcal{L}[(\overline{V}, \varphi)]\}(R) \geq 0\). Thus, we have shown that \(\overline{V}\) is a viscosity subsolution on \([0, T] \times \frac{T}{2+\sqrt{2}}\), \(\infty \times [0, s]\).

Case II If \(R \in \text{int} \mathcal{D}^*\) and \(x > (T - s) \land w \land \frac{\sqrt{2}}{2}(s - w)\), then it is obvious that \(x\) is less than \(\frac{T}{2+\sqrt{2}}\). By several simple calculations, we obtain \(\varphi_x(R) = 2, \varphi_s(R) = N_2, \varphi_w(R) = 0\).

\[
\mathcal{L}[(\overline{V}, \varphi)](R) = \left[-(c + \lambda(w))\overline{V} + \varphi_s + p\varphi_x + \varphi_w\right](R) + \lambda(w) \int_0^x \overline{V}(s, x - u, 0)dG(u) > 0.
\]

Thus, \(\overline{V}\) is a viscosity subsolution of (26) at \(R\). Now, we prove a special case of the boundary.
Case III If $R = (0, 0, 0) \in \partial \mathcal{D}^*$, then $\varphi_y(R_0) \geq N_2, [\varphi_y + \varphi_z](R_0) \geq N_2, \varphi_x(R_0) \geq 1$. Thus, it is easy to verify that $\mathcal{L} [V, \varphi](R) \geq N_2 + p > 0$. Now, we observe that $V$ is a viscosity subsolution of (26) at $(0, 0, 0)$. We can prove other cases similarly. 

\[ \square \]

Lemma 5.1 Let $\overline{V}$ be the viscosity supersolution constructed in (39) and $V$ be the viscosity subsolution constructed in (40). Define

$$\omega(s, x, w) = \sup_{u \in \mathcal{G}} u(s, x, w),$$

where $\mathcal{G} := \{ u|u$ is a viscosity subsolution of (26) on $\mathcal{D}^*$ and $V \leq u \leq \overline{V}$ in $D\}$. Then, $\omega^*$ is a viscosity subsolution of (26) on $\mathcal{D}^*$ satisfying $x + M_2 \leq \omega^*(s, x, w) \leq M_1 + x$.

We omit the proof, one can refer to Crandall et al. [16] and Mou [36] for similar arguments. At this point, we are ready to provide the representation of the value function $V$.

Theorem 5.3 The value function $V = \omega^*$, where $\omega$ is defined in (41).

Proof From Lemma 5.1, we see that $\omega^*$ is a viscosity subsolution of (26). Note that the value function $V$ is a viscosity subsolution of (26) that satisfies $\overline{V} \leq V \leq V$. By the definition of $\omega^*$, we know that $\omega^* \geq V$. Now, we only need to show that $V \geq \omega^*$. Recall that $c > 0$ is the discount factor. Choose a sufficiently large $\tilde{K}$ such that $c\tilde{K} \geq 1$. Consider the function $V^\theta := \theta V + (\theta - 1)\tilde{K}$, where $\theta > 1$. Since $V$ is a supersolution, it is easy to verify that $V^\theta$ is also a viscosity supersolution of (26).

Indeed, consider any continuously differentiable function $\varphi$ such that $V^\theta - \varphi$ attains its minimum at $R_0$. Then, max \(1 - \left(\frac{\varphi}{\theta} - \frac{\theta - 1}{\theta} \tilde{K}\right)_x, \mathcal{L} [V, \frac{\varphi}{\theta} - \frac{\theta - 1}{\theta} \tilde{K}]\) \leq 0. Thus,

$$\max \left\{ 1 - \varphi_x, \mathcal{L} [V^\theta, \varphi] \right\} \leq -(\theta - 1),$$

which shows that $V^\theta$ is a viscosity supersolution of (26). Instead of comparing $\omega^*$ and $V$, we will compare $\omega^*$ and $V^\theta$. If we can show that $\omega^* \leq V^\theta$, then by simply letting $\theta \downarrow 1$, we obtain the desired comparison result $\omega^* \leq V$ in $D$. Observe that for all $(s, x, w) \in D, x + M_2 \leq V(s, x, w) \leq x + M_1$ and $x + M_2 \leq \omega^*(s, x, w) \leq x + M_1$; thus, we have

$$[\omega^* - V^\theta](s, x, w) \leq x + M_1 - \left[ \theta(x + M_2) + (\theta - 1)\tilde{K} \right] = (1 - \theta)x + (M_1 - \theta M_2 - (\theta - 1)\tilde{K}).$$

In view of (43), we can choose $b := \frac{M_1 - \theta M_2 - (\theta - 1)\tilde{K}}{\theta - 1}$ such that for all $x \geq b, \omega^* \leq V^\theta$. Although $D$ is unbounded, we can restrict our attention to the bounded domain

$$\mathcal{D}_b := \{(t, y, v)|0 \leq t < T, 0 \leq y < b, 0 \leq v \leq t\}$$

\[ \square \]
and prove that $\omega^* \leq V^\theta$ on $\mathcal{D}_b$. Now, we assume on the contrary that

$$M_b := \max_{\mathcal{cl}(\mathcal{D}_b)} [\omega^* - V^\theta] = (\omega^* - V^\theta)(\bar{s}, \bar{x}, \bar{w}) > 0$$

for some $\bar{R} := (\bar{s}, \bar{x}, \bar{w}) \in \mathcal{cl}(\mathcal{D}_b)$. Observe that we have only two cases $\bar{R} \in \mathcal{D}_b^0$ and $\bar{R} \in \mathcal{D}_b^1$ to consider, where $\mathcal{D}_b^1 := \partial \mathcal{D}_b \setminus \{t = T \cup \{y = b\}\}$ is the state constraint boundary restricted by $b$.

Case I Consider $\bar{R} \in \mathcal{D}_b^1$. The construction presented below is a suitable adaption of the construction of Benth et al. [13]. Denote $R := (t, y, v)$ for simplicity. As $\mathcal{D}_b$ is piecewise linear, there exist constants $h_0, k > 0$ and a uniformly continuous map $\eta : \mathcal{cl}(\mathcal{D}_b) \mapsto \mathbb{R}^3$ satisfying

$$B(R + h\eta(R), h\kappa) \subset \mathcal{D}_b^0 \quad \text{for all } R \in \mathcal{cl}(\mathcal{D}_b) \text{ and } h \in [0, h_0],$$

where $B(z, \rho)$ denotes the sphere with radius $\rho$ and center $z$. Note that $\eta(R)$ is a three-dimensional vector, we can write $\eta(R) = (\eta_1(R), \eta_2(R), \eta_3(R))$. For any $\kappa > 1$ and $0 < \varepsilon < 1$, define the function $\Phi(s, x, w, t, y, v)$ on $\mathcal{cl}(\mathcal{D}_b) \times \mathcal{cl}(\mathcal{D}_b)$ by

$$\Phi(s, x, w, t, y, v) = \omega^*(s, x, w) - V^\theta(t, y, v) - (\kappa(s - t) + \varepsilon \eta_1(\bar{R}))^2 - (\kappa(x - y) + \varepsilon \eta_2(\bar{R}))^2 - (\kappa(w - v) + \varepsilon \eta_3(\bar{R}))^2 - \varepsilon(s - \bar{s})^2 - \epsilon(x - \bar{x})^2 - \epsilon(w - \bar{w})^2.$$

Let $M_\kappa = \max_{\mathcal{cl}(\mathcal{D}_b) \times \mathcal{cl}(\mathcal{D}_b)} \Phi(s, x, w, t, y, v)$. Let $(s_k, x_k, w_k, t_k, y_k, v_k) \in \mathcal{cl}(\mathcal{D}_b) \times \mathcal{cl}(\mathcal{D}_b)$ be a maximizer of $\Phi$, i.e., $M_\kappa = \Phi(s_k, x_k, w_k, t_k, y_k, v_k)$. Then, we can get that for a fixed small number $\varepsilon_0$, $M_\kappa \geq \omega^*(\bar{s}, \bar{x}, \bar{w}) - V^\theta(\bar{s}, \bar{x}, \bar{w}) - \varepsilon^2|\eta(\bar{R})|^2 > 0$ for any $\kappa > 1$ and $\varepsilon < \varepsilon_0$. By (44), we assume that $\kappa$ is so large that $\bar{R} + \frac{\varepsilon}{\kappa} \eta(\bar{R}) \in \mathcal{D}_b^0$. From

$$\Phi(s_k, x_k, w_k, t_k, y_k, v_k) \geq \Phi \left(\bar{s}, \bar{x}, \bar{w}, \bar{s} + \frac{\varepsilon}{\kappa} \eta_1(\bar{R}), \bar{x} + \frac{\varepsilon}{\kappa} \eta_2(\bar{R}), \bar{w} + \frac{\varepsilon}{\kappa} \eta_3(\bar{R})\right),$$

we obtain that

$$|\kappa(s_k - t_k) + \varepsilon \eta_1(\bar{R})|^2 + |\kappa(x_k - y_k) + \varepsilon \eta_2(\bar{R})|^2 + |\kappa(w_k - v_k) + \varepsilon \eta_3(\bar{R})|^2 + \varepsilon(s_k - \bar{s})^2 + \epsilon(x_k - \bar{x})^2 + \epsilon(w_k - \bar{w})^2$$

$$\leq \omega^*(s_k, x_k, w_k) - V^\theta(t_k, y_k, v_k) - (\omega^* - V^\theta)(\bar{s}, \bar{x}, \bar{w}) - V^\theta(\bar{s}, \bar{x}, \bar{w}) + V^\theta \left(\bar{s} + \frac{\varepsilon}{\kappa} \eta_1(\bar{R}), \bar{x} + \frac{\varepsilon}{\kappa} \eta_2(\bar{R}), \bar{w} + \frac{\varepsilon}{\kappa} \eta_3(\bar{R})\right).$$

As $\omega^*$ and $-V^\theta$ are bounded on $\mathcal{cl}(\mathcal{D}_b)$, it follows that $|\kappa(s_k - t_k)|, |\kappa(x_k - y_k)|, |\kappa(w_k - v_k)|$ are uniformly bounded in $\kappa$. Hence, we have

$$s_k - t_k \to 0, x_k - y_k \to 0, w_k - v_k \to 0, \quad \text{as } \kappa \to \infty$$

and

$$\lim_{\kappa \to \infty} \left[\omega^*(s_k, x_k, w_k) - V^\theta(t_k, y_k, v_k)\right] \leq M_b.$$
Letting $\kappa \to \infty$ in (45) and using the upper semi-continuity of $\omega^*$, $-V^\theta$ in $\text{cl}(\mathcal{D}_b)$, we conclude that

$$|\kappa(s_k - t_k) + \varepsilon \eta_1(\bar{R})| \to 0, \quad |\kappa(x_k - y_k) + \varepsilon \eta_2(\bar{R})| \to 0,$$

$$|\kappa(w_k - v_k) + \varepsilon \eta_3(\bar{R})| \to 0,$$

$(s_k, x_k, w_k) \to \bar{R}$, $(t_k, y_k, v_k) \to \bar{R}$ and $M_k \to M_b$. Using the uniformly continuity of $\eta$, we can have that $t_k = s_k + \frac{b}{\kappa} \eta_1(s_k, x_k, w_k) + o\left(\frac{1}{\kappa}\right)$. Similarly, we get that $y_k = x_k + \frac{b}{\kappa} \eta_2(s_k, x_k, w_k) + o\left(\frac{1}{\kappa}\right)$, $v_k = w_k + \frac{b}{\kappa} \eta_3(s_k, x_k, w_k) + o\left(\frac{1}{\kappa}\right)$. Thus, combining this with (44) we obtain that $(t_k, y_k, v_k) \in \mathcal{D}_b^0$ for sufficiently large $\kappa$.

Now, we define

$$\phi(s, x, w) := V^\theta(t_k, y_k, v_k) + |\kappa(s - t_k) + \varepsilon \eta_1(\bar{R})|^2 + |\kappa(x - y_k) + \varepsilon \eta_2(\bar{R})|^2$$

$$+ |\kappa(w - v_k) + \varepsilon \eta_3(\bar{R})|^2 + \varepsilon(s - \bar{s})^2 + \varepsilon(x - \bar{x})^2 + \varepsilon(w - \bar{w})^2.$$

$$\varphi(t, y, v) := \omega^*(s_k, x_k, w_k) - |\kappa(s_k - t) + \varepsilon \eta_1(\bar{R})|^2 - |\kappa(x_k - y) + \varepsilon \eta_2(\bar{R})|^2$$

$$- |\kappa(w_k - v) + \varepsilon \eta_3(\bar{R})|^2 - \varepsilon(s_k - \bar{s})^2 - \varepsilon(x_k - \bar{x})^2 - \varepsilon(w_k - \bar{w})^2.$$

Noting that $V^\theta$ is a viscosity supersolution and $\omega^*$ is a viscosity subsolution, we get that

$$\max \left\{ 1 - 2\kappa \left[ |\kappa(x_k - y_k) + \varepsilon \eta_2(\bar{R})| \right],
- (c + \lambda(w_k)) V^\theta(t_k, y_k, v_k) + 2\kappa[|\kappa(s_k - t_k) + \varepsilon \eta_1(\bar{R})|]
+ 2p\kappa[|\kappa(x_k - y_k) + \varepsilon \eta_2(\bar{R})|] + 2\kappa[|\kappa(w_k - v_k) + \varepsilon \eta_3(\bar{R})|]
+ \lambda(v_k) \int_0^{x_k} V^\theta(t_k, y_k - u, 0)dG(u) \right\} \leq -(\theta - 1).$$

(46)

$$\max \left\{ 1 - 2\kappa[|\kappa(x_k - y_k) + \varepsilon \eta_2(\bar{R})| - 2\varepsilon(x_k - \bar{x})] - (c + \lambda(w_k))\omega^*(s_k, x_k, w_k)
+ 2\kappa[|\kappa(s_k - t_k) + \varepsilon \eta_1(\bar{R})|] + 2\varepsilon(s_k - \bar{s})
+ p\left[ 2\kappa[|\kappa(x_k - y_k) + \varepsilon \eta_2(\bar{R})|] + 2\varepsilon(x_k - \bar{x}) \right] + 2\kappa[|\kappa(w_k - v_k) + \varepsilon \eta_3(\bar{R})|]
+ 2\varepsilon(w_k - \bar{w}) + \lambda(w_k) \int_0^{x_k} \omega^*(s_k, x_k - u, 0)dG(u) \right\} \geq 0.$$  

(47)

Combining (46) with (47), we let $\kappa \to \infty$, $\varepsilon \to 0$ to obtain the desired contradiction

$$[c + \lambda(\bar{w})](\omega^* - V^\theta)(\bar{s}, \bar{x}, \bar{w}) < \lambda(\bar{w})(\omega^* - V^\theta)(\bar{s}, \bar{x}, \bar{w}).$$

Case II Let us consider the case $\bar{R} \in \mathcal{D}_b^0$. For any $\kappa > 1$, define the function $\Psi$ on $\text{cl}(\mathcal{D}_b) \times \text{cl}(\mathcal{D}_b)$ by

$$\square$$

Springer
\[ \Psi(s, x, w, t, y, v) = \omega^*(s, x, w) - V^\theta(t, y, v) - \frac{\kappa}{2} (s - t)^2 - \frac{\kappa}{2} (x - y) - \frac{\kappa}{2} (w - v)^2. \]

Let \( M_\kappa = \max_{\text{cl}(\mathcal{D}_b) \times \text{cl}(\mathcal{D}_b)} \Psi(s, x, w, t, y, v). \) Let \((s_\kappa, x_\kappa, w_\kappa, t_\kappa, y_\kappa, v_\kappa)\) be a maximizer. We have \( M_\kappa \geq M_b > 0 \) for all \( \kappa > 1. \) As \( \text{cl}(\mathcal{D}_b) \times \text{cl}(\mathcal{D}_b) \) is compact, we can find a subsequence that may be assumed to be \((s_\kappa, x_\kappa, w_\kappa, t_\kappa, y_\kappa, v_\kappa)\) itself such that \((s_\kappa, x_\kappa, w_\kappa, t_\kappa, y_\kappa, v_\kappa) \to (\hat{s}, \hat{x}, \hat{w}, \hat{t}, \hat{y}, \hat{v}). \) Since \( M_\kappa \geq \Psi(\hat{s}, \hat{x}, \hat{w}, \hat{s}, \hat{x}, \hat{w}), \)

we obtain

\[
\frac{\kappa}{2} (s_\kappa - I_\kappa)^2 + \frac{\kappa}{2} (x_\kappa - y_\kappa)^2 + \frac{\kappa}{2} (w_\kappa - v_\kappa)^2 
\leq \omega^*(s_\kappa, x_\kappa, w_\kappa) - V^\theta(t_\kappa, y_\kappa, v_\kappa) - \omega^*(\hat{s}, \hat{x}, \hat{w}) + V^\theta(\hat{s}, \hat{x}, \hat{w}). \quad (48)
\]

As all upper semicontinuous functions attain their maximum on any compact set, we obtain that \( \frac{\kappa}{2} (s_\kappa - t_\kappa)^2 + \frac{\kappa}{2} (x_\kappa - y_\kappa)^2 + \frac{\kappa}{2} (w_\kappa - v_\kappa)^2 \) is bounded uniformly in \( \kappa. \) Thus, we obtain \( s_\kappa - y_\kappa \to 0, s_\kappa - t_\kappa \to 0, w_\kappa - v_\kappa \to 0 \) as \( \kappa \to \infty, \) which means \( \hat{s} = \hat{t}, \hat{w} = \hat{v}, \hat{x} = \hat{y}. \) Letting \( \kappa \to \infty \) in (48), we get

\[
\lim_{\kappa \to \infty} \left[ \frac{\kappa}{2} (s_\kappa - I_\kappa)^2 + \frac{\kappa}{2} (x_\kappa - y_\kappa)^2 + \frac{\kappa}{2} (w_\kappa - v_\kappa)^2 \right] + \omega^*(\hat{s}, \hat{x}, \hat{w})
- V^\theta(\hat{s}, \hat{x}, \hat{w}) \leq \omega^*(\hat{s}, \hat{x}, \hat{w}) - V^\theta(\hat{s}, \hat{x}, \hat{w}).
\]

By the definition of \((\hat{s}, \hat{x}, \hat{w}), \) we obtain \( \hat{s} = \hat{s}, \hat{x} = \hat{x}, \hat{w} = \hat{w}. \) As \((\hat{s}, \hat{x}, \hat{w}) \in \mathcal{D}_b^0, \) we observe that for sufficiently large \( \kappa, (s_\kappa, x_\kappa, w_\kappa), (t_\kappa, y_\kappa, v_\kappa) \in \mathcal{D}_b^0. \) Note that for any given \( \kappa, \Psi(s, x, w, t, y, v) \) attains its maximum at \((s_\kappa, x_\kappa, w_\kappa, t_\kappa, y_\kappa, v_\kappa). \) Define the functions

\[
\phi(s, x, w) := V^\theta(t_\kappa, y_\kappa, v_\kappa) + \frac{\kappa}{2} (x - y)^2 + \frac{\kappa}{2} (s - t)^2 - \frac{\kappa}{2} (w - v)^2,
\]

\[
\psi(t, y, v) := \omega^*(s_\kappa, x_\kappa, w_\kappa) - \frac{\kappa}{2} (s_\kappa - t)^2 - \frac{\kappa}{2} (x_\kappa - y)^2 - \frac{\kappa}{2} (w_\kappa - v)^2.
\]

We observe that \([\omega^* - \phi](s, x, w) \) attains its maximum at \((s_\kappa, x_\kappa, w_\kappa) \in \mathcal{D}_b^0 \) and \([V^\theta - \psi](t, y, v) \) attains its minimum at \((t_\kappa, y_\kappa, v_\kappa) \in \mathcal{D}_b^0; \) \( \omega^* \) is a viscosity subsolution of (26), and \( V^\theta \) is a viscosity supersolution of (26); noting that \( \max \{1 - \psi, \mathcal{L}[V^\theta, \psi] \} \leq - (\theta - 1) \) because of (42), we obtain

\[
\max \left\{ 1 - \kappa (x_\kappa - y_\kappa), -(c - \lambda(w_\kappa)) \omega^*(s_\kappa, x_\kappa, w_\kappa) + pk(x_\kappa - y_\kappa)
+ \kappa (s_\kappa - t_\kappa) + \kappa (w_\kappa - v_\kappa) + \lambda(w_\kappa) \int_0^{x_\kappa} \omega^*(s_\kappa, x_\kappa - u, 0) dG(u) \right\} \geq 0, \quad (49)
\]

\[
\max \left\{ 1 - \kappa (x_\kappa - y_\kappa), -(c + \lambda(v_\kappa)) V^\theta(t_\kappa, y_\kappa, v_\kappa) + \kappa (s_\kappa - t_\kappa) + \kappa (w_\kappa - v_\kappa)
+ pk(x_\kappa - y_\kappa) + \lambda(v_\kappa) \int_0^{y_\kappa} V^\theta(t_\kappa, y_\kappa - u, 0) dG(u) \right\} \leq -(\theta - 1). \quad (50)
\]
From (50), we observe that \( \lim_{\kappa \to \infty} [1 - \kappa (x_\kappa - y_\kappa)] < 0 \). Combining this fact with (49) and (50) and letting \( \kappa \to \infty \), we obtain

\[-[c + \lambda(\bar{w})]M_b = -[c + \lambda(\bar{w})](\omega^* - V^\theta)(\bar{s}, \bar{x}, \bar{w}) > -\lambda(\bar{w})M_b,\]

which is a contradiction. Until now, we have shown that \( \omega^* \leq V \), which leads to \( \omega^* = V \).

**Remark 5.1** In fact, the proof of \( \omega^* \leq V \) implies the comparison principle. The comparison principle shows that for all supersolutions \( \bar{u} \) and subsolutions \( u \), if \( \bar{u} \) and \( u \) both satisfy the linear growth condition, then \( u \leq \bar{u} \). The comparison principle shows that the value function is the unique constrained viscosity solution of HJB equation satisfying the growth condition.

**Remark 5.2** If we consider the optimal singular dividend and investment problems, then we will add a second-order partial derivative about \( x \) in the HJB equation. All theorems can be verified similarly combining this paper and Bai et al. [22]. Currently, we focus on the analysis of optimal singular dividend problem. Thus, there is no second-order partial derivative in our HJB equation.

In the future, we will construct the auxiliary equations so that we can use their solutions to approximate the value function. Meanwhile, using the smooth solutions of the auxiliary equations we can construct the \( \varepsilon \)-optimal strategy for our problem. In our paper, there is a “max” operator in the HJB equation which makes our optimization problem far more complicated. The closest study to our HJB equation exploring the regularity of solutions of obstacle integro-differential operators is Caffarelli et al. [37]. In [37], the authors consider the obstacle problem in \( \mathbb{R}^n \)

\[
\min\{-Lu, u - \varphi\} = 0 \quad \text{in} \quad \mathbb{R}^n, \quad \lim_{|x| \to \infty} u(x) = 0, \quad (51)
\]

where \( L \) is an infinitesimal generator of a Lévy process, and \( \varphi \) is a given bounded differentiable function on \( \mathbb{R}^n \) that we call an obstacle. The authors showed that a solution of (51) belongs to \( C^{1,s} \) near all regular points, where \( s \in ]0, 1[ \). We can construct the auxiliary equations based on this result, and then we construct the \( \varepsilon \)-optimal strategy. Because the construction proof is too long, we will present it in our next paper.

**6 Conclusions**

In this paper, we consider the optimal dividend problem when the underlying reserve process is modeled by the Sparre Andersen model. We derive the verification theorem in the framework of viscosity solutions. Comparing to the classical verification theorem, the analysis of viscosity solutions is far more complex. We introduce the distance function to investigate the viscosity supersolution and subsolution. The results are then applied to express the optimal value function. Our finding demonstrates again that the
viscosity solution can serve as a powerful tool for dealing with stochastic optimal control problems. However, the strategy analysis is beyond the scope of this paper and it will appear in forthcoming research.

Acknowledgements Here, we want to express our thanks to Jacques Rioux for his dedication to the improvement in this paper. This research is supported by Chinese NSF Grants No. 11471171, No. 11911530091 and No. 11931018.

References

1. De Finetti, B., Su un’impostazione alternativa della teoria collettiva del rischio. In: Transactions of the XVth International Congress of Actuaries, vol. 2, pp. 433–443. New York (1957)
2. Asmussen, S., Taksar, M.: Controlled diffusion models for optimal dividend pay-out. Insur. Math. Econ. 20(1), 1–15 (1997)
3. Guo, X., Liu, J., Zhou, X.Y.: A constrained non-linear regular-singular stochastic control problem, with applications. Stoch. Process. Appl. 109(2), 167–187 (2004)
4. Azcue, P., Muler, N.: Optimal reinsurance and dividend distribution policies in the Cramér–Lundberg model. Math. Finance 15(2), 261–308 (2005)
5. Belhaj, M.: Optimal dividend payments when cash reserves follow a jump-diffusion process. Math. Finance 20(2), 313–325 (2010)
6. Gerber, H.U., Shiu, S.W.: On optimal dividend strategies in the compound Poisson model. North Am. Actuar. J. 10(2), 76–93 (2006)
7. Andersen, E.S.: On the collective theory of risk in case of contagion between claims. Bull. Inst. Math. Appl. 12(2), 275–279 (1957)
8. Li, S.M., Garrido, J.: On a class of renewal risk models with a constant dividend barrier. Insur. Math. Econ. 35(3), 691–701 (2004)
9. Albrecher, H., Márton, M., Claramunt, M.M.: On the distribution of dividend payments in a Sparre Andersen model with generalized Erlang(n) interclaim times. Insur. Math. Econ. 37(2), 324–334 (2005)
10. Albrecher, H., Hartinger, J.: On the non-optimality of horizontal barrier strategies in the Sparre Andersen model. HERMES Int. J. Comput. Math. Appl. 7, 1–14 (2006)
11. Albrecher, H., Thonhauser, S.: Optimality results for dividend problems in insurance. RACSAM-Revista de la Real Academia de Ciencias Exactas, Fisicas y Naturales. Serie A. Matematicas 103(2), 295–320 (2009)
12. Azcue, P., Muler, N.: Optimal investment policy and dividend payment strategy in an insurance company. Ann. Appl. Probab. 20(4), 1253–1302 (2010)
13. Benth, F.E., Karlsen, K.H., Reikvam, K.: Optimal portfolio selection with consumption and nonlinear integro-differential equations with gradient constraint: a viscosity solution approach. Finance Stoch. 5(3), 275–303 (2001)
14. Seydel, R.C.: Existence and uniqueness of viscosity solutions for QVI associated with impulse control of jump-diffusions. Stoch. Process. Appl 119(10), 3719–3748 (2009)
15. Crandall, M.G., Lions, P.L.: Viscosity solutions of Hamilton–Jacobi equations. Trans. Am. Math. Soc. 277(1), 1–42 (1983)
16. Crandall, M.G., Ishii, H., Lions, P.L.: User’s guide to viscosity solutions of second order partial differential equations. Bull. Am. Math. Soc. 27(1), 1–67 (1992)
17. Lions, P.L.: Optimal control of diffusion processes and Hamilton–Jacobi–Bellman equations part 2: viscosity solutions and uniqueness. Commun. Partial Differ. Equ. 8(11), 1229–1276 (1983)
18. Lions, P.L.: Optimal control of diffusion processes and Hamilton–Jacobi–Bellman equations. I. The dynamic programming principle and applications. Commun. Partial Differ. Equ. 8(10), 1101–1174 (1983)
19. Soner, H.M.: Optimal control with state-space constraint. II. SIAM J. Control Optim. 24(6), 1110–1122 (1986)
20. Fleming, W.H., Soner, H.M.: Controlled Markov Processes and Viscosity Solutions. Stochastic Modelling and Applied Probability, vol. 25, 2nd edn. Springer, New York (2006)
21. Yong, J., Zhou, X.Y.: Stochastic Controls: Hamiltonian Systems and HJB Equations, vol. 43. Springer, New York (1999)
22. Bai, L.H., Ma, J., Xing, X.J.: Optimal dividend and investment problems under Sparre Andersen model. Ann. Appl. Probab. 27(6), 3588–3632 (2017)
23. Rolski, T., Schmidli, H., Schmidt, V., Teugels, J.L.: Stochastic Processes for Insurance and Finance, vol. 505. Wiley, New York (2009)
24. Protter, P.: Stochastic integration and differential equations (2nd ed.). In: Stochastic Modelling and Applied Probability, vol. 21. Springer-Verlag, Berlin (2004)
25. Scheer, N., Schmidli, H.: Optimal dividend strategies in a Cramér–Lundberg model with capital injections and administration costs. Eur. Actuar. J. 1(1), 57–92 (2011)
26. Pham, H.: Continuous-Time Stochastic Control and Optimization with Financial Applications. Springer, Berlin (2009)
27. Avatif, S.: Équations d’Hamilton–Jacobi du premier ordre avec termes intégro-différentiels: Partie 1: Unicité Des solutions de viscousité. Commun. Partial Differ. Equ. 16(6–7), 1057–1074 (1991)
28. Czarna, I., Palmowski, Z.: Dividend problem with Parisian delay for a spectrally negative Lévy risk process. J. Optim. Theory Appl. 161(1), 239–256 (2014)
29.Marciniak, E., Palmowski, Z.: On the optimal dividend problem for insurance risk models with surplus-dependent premiums. J. Optim. Theory Appl. 168(2), 723–742 (2016)
30. Pérez, J.L., Yamazaki, K., Yu, X.: On the bail-out optimal dividend problem. J. Optim. Theory Appl. 179(2), 553–568 (2018)
31. Zhu, J., Yang, H.: Optimal capital injection and dividend distribution for growth restricted diffusion models with bankruptcy. Insur. Math. Econ. 70, 259–271 (2016)
32. Zhu, J., Chen, F.: Dividend optimization for regime-switching general diffusions. Insur. Math. Econ. 53(2), 439–456 (2013)
33. Zhao, Y., Wang, R., Yao, D., Chen, P.: Optimal dividends and capital injections in the dual model with a random time horizon. J. Optim. Theory Appl. 167(1), 272–295 (2015)
34. Wei, J., Yang, H., Wang, R.: Classical and impulse control for the optimization of dividend and proportional reinsurance policies with regime switching. J. Optim. Theory Appl. 147(2), 358–377 (2010)
35. Choulli, T., Taksar, M., Zhou, X.Y.: A diffusion model for optimal dividend distribution for a company with constraints on risk control. SIAM J. Control Optim. 41(6), 1946–1979 (2003)
36. Mou, C.: Perron’s method for nonlocal fully nonlinear equations. Anal. PDE 10(5), 1227–1254 (2017)
37. Caffarelli, L.A., Ros-Oton, X., Serra, J.: Obstacle problems for integro-differential operators: regularity of solutions and free boundaries. Invent. Math. 208(3), 1155–1211 (2017)

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