Finite Horizon Stochastic Optimal Switching and Impulse Controls with a Viscosity Solution Approach*

Shanjian Tang and Jiongmin Yong

Department of Mathematics, Fudan University

Shanghai 200433, China

Abstract. A stochastic optimal switching and impulse control problem in a finite horizon is studied. The continuity of the value function, which is by no means trivial, is proved. The Bellman dynamic programming principle is shown to be valid for such a problem. Moreover, the value function is characterized as the unique viscosity solution of the corresponding Hamilton-Jacobi-Bellman equation.

Keywords. optimal stochastic control, dynamic programming, value function, viscosity solution, switching control, impulse control.

AMS(MOS) classifications. 49L20, 49L25, 49N25, 93E20.

§1. Introduction.

In this paper, we consider the following stochastic system in some fixed probability space $(\Omega, \mathcal{F}, P)$ with a filtration $\mathcal{F}_t$:

$$
\begin{aligned}
y_{t,x}(s) =& x + \int_s^t f(r, y_{t,x}(r), u(r), a(r))dr \\
&+ \int_t^s \sigma(r, y_{t,x}(r), u(r), a(r))dw_r + \xi(s), \qquad s \in [0,T],
\end{aligned}
\tag{1.1}
$$

where f and σ are given maps; w_t is a d-dimensional $\mathcal{F}_t$-Brownian motion; $y_{t,x}(s)$, with the initial value $x \in \mathbb{R}^n$ at the initial time t (the subscript in $y_{t,x}(\cdot)$ indicates this dependence), is the state with values in $\mathbb{R}^n$; $u(s)$ is the continuous control with values in a metric space

* This work was partially supported by the NSF of China under Grant 19131050, the Fok Ying Tung Education Foundation and the Chinese State Education Commission Science Foundation under Grant 9024617.

U; The switching and impulse controls $a(s)$ and $\xi(s)$ have values in some finite set A and some cone $K \subseteq \mathbb{R}^n$, respectively, and have the following forms

$$a(\cdot) = \sum_{i\geq 1} a_{i-1}\chi_{[\theta_{i-1},\theta_i)}(\cdot), \qquad \xi(\cdot) = \sum_{j\geq 1} \xi_{j-1}\chi_{[\tau_{i-1},T]}(\cdot), \tag{1.2}$$

are the switching and impulse controls. Here, θ_i and τ_j are $\mathcal{F}_t$-stopping times and a_i and ξ_j are $\mathcal{F}_{\theta_i}$- and $\mathcal{F}_{\tau_j}$-adapted, respectively. The cost functional is defined by

$$\begin{aligned} J^a_{t,x}(u(\cdot), a(\cdot), \xi(\cdot)) = E\{h(y_{t,x}(T)) + \int_t^T f^0(r, y_{t,x}(r), u(r), a(r))dr \\ + \sum_{i\geq 1} k(\theta_i, a_{i-1}, a_i) + \sum_{j\geq 1} \ell(\tau_j, \xi_j)\}, \end{aligned} \tag{1.3}$$

where h, f^0, k and ℓ are some real valued functions. The four terms in the right hand side of (1.3) are usually referred as the final state penalty, the running, switching and impulse costs, respectively. Our optimal control problem is to minimize the cost functional (1.3) over the given class of admissible controls.

The optimal switching and/or impulse coontrol problems were studied by many authors. The readers are referred to [1,5,21,22] for deterministic cases, to [2–4,8,14–17] for stochastic cases and to [18–20] for infinite dimensional deterministic cases. We note that those papers concerning the stochastic cases did not treat the finite horizon impulse control problem. Also, the viscosity solution approach has not been used to treat general stochastic switching and impulse control problems, to our best knowledge. Here, the "general" means general stochastic differential systems allowing degenerate diffusion coefficients. We should notice that in [17], a semigroup approach was used to treat the infinite horizon problem allowing the diffusion to be degenerate. For the theory of viscosity solutions, we refer the readers to [6,7,10–13].

In this paper, we use dynamic programming approach and viscosity solutions to discuss the finite horizon optimal switching and impulse control problem stated above. We allow the diffusion to be degenerate. We should point out that due to the appearance of the impulse control the continuity of the value function in time t is by no means trivial. Such a continuity is obtained by carrying out some careful estimates and by modifying some ideas used in [22]. Then, using a similar argument as [9], we prove the validity of the dynamic programming principle for our problem. As a consequence, we show that the value function is a viscosity solution of the corresponding Hamilton-Jacobi-Bellman equation. In proving the uniqueness of the viscosity solution, we adopt a technique of [16], which improves the one used in [21,22].

The rest of this paper is organized as follows. In Section 2, we give the formulation of our problem and make some assumptions. Section 3 is devoted to the proof of the continuity of the value function. In Section 4, the dynamic programming principle is shown to be valid. Finally, in Section 5, the value function is proved to be the unique viscosity solution of HJB equation in some given function space.

§2. Formulation of the Problem and Assumptions.

Let us start with the framework of our optimal control problem. Some of them are adapted from [9].

For each $t \in [0,T]$, let

$$\Omega_t = \{\omega \in C([t,T];\mathbb{R}^d) : \omega(t) = 0\}. \tag{2.1}$$

Denote by $\mathcal{F}_{t,T}$ the topological σ-field of Ω_t and endow the measurable space $(\Omega_t, \mathcal{F}_{t,T})$ with the Wiener measure P_t. Then, Ω_t becomes a canonical sample space. We set

$$\Omega_{t,s} = \{\omega \in C([t,s];\mathbb{R}^d) : \omega(t) = 0\}, \qquad 0 \le t < s \le T. \tag{2.2}$$

Clearly,

$$\Omega_t = \Omega_{t,T}.$$

Now, if $\tau \in (t,T)$ and $\omega \in \Omega_t$, we let

$$\begin{cases} \omega_1 = \omega\big|_{[t,\tau]}, \\ \omega_2 = (\omega - \omega(\tau))\,\big|_{[\tau,T]}, \\ \Pi\omega = (\omega_1, \omega_2). \end{cases} \tag{2.3}$$

We see that the map $\Pi : \Omega_t \to \Omega_{t,\tau} \times \Omega_\tau$ induces an identification

$$\Omega_t = \Omega_{t,\tau} \times \Omega_\tau. \tag{2.4}$$

Moreover, the inverse of Π is defined in an evident way, $\omega = \Pi^{-1}(\omega_1, \omega_2)$. Finally, if we let $P_{t,\tau}$ and P_τ be the Wiener measures on $\Omega_{t,\tau}$ and Ω_τ, respectively, then, we have

$$P_t = P_{t,\tau} \otimes P_\tau.$$

We refer the interested readers to [9] for relevant details about the above.

Now, we define

$$w_\tau(\omega) = \omega(\tau), \qquad \omega \in \Omega_t. \tag{2.5}$$

Then, $\{w_\tau\}$ is a standard Wiener process. We let $\mathcal{F}_{t,s}$ be the σ-algebra generated by $\{w_\tau : t \le \tau \le s\}$ with $t < s$. It is clear that for $s = T$, $\mathcal{F}_{t,s}$ coincides with the topological σ-field of Ω_t defined at the beginning of this section.

Next, we let $A = \{1, 2, \cdots, m\}$, K be a convex cone in $\mathbb{R}^n$ and U be a metric space. Let $f : [0,T] \times \mathbb{R}^n \times U \times A \to \mathbb{R}^n$, $\sigma : [0,T] \times \mathbb{R}^n \times U \times A \to \mathbb{R}^{n\times d}$, $f^0 : [0,T] \times \mathbb{R}^n \times U \times A \to \mathbb{R}$, $h : \mathbb{R}^n \to \mathbb{R}$, $k : [0,T] \times A \times A \to \mathbb{R}$ and $\ell : [0,T] \times K \to \mathbb{R}$ be continuous and satisfy the following:

(C1) There exist constants $L > 0$, $\delta \in (0,1]$, $\gamma, \nu \in [0,1)$, such that for all $t \in [0,T]$, $x, \widehat{x} \in \mathbb{R}^n$, $u \in U$ and $a \in A$,

$$\begin{aligned}
&|f(t,x,u,a)| \le L(1 + |x|^\nu),\\
&|\sigma(t,x,u,a)| \le L(1 + |x|^{\nu/2}),\\
&|f(t,x,u,a) - f(t,\widehat{x},u,a)| + |\sigma(t,x,u,a) - \sigma(t,\widehat{x},u,a)| \le L|x - \widehat{x}|,\\
&-L \le f^0(t,x,u,a), h(x) \le L(1 + |x|^{\gamma+\delta}),\\
&|f^0(t,x,u,a) - f^0(t,\widehat{x},u,a)| + |h(x) - h(\widehat{x})| \le L(1 + |x|^\gamma + |\widehat{x}|^\gamma)|x - \widehat{x}|^\delta.
\end{aligned} \tag{2.6}$$

(C2) For any $a_1, a_2, a_3 \in A$ with $a_1 \ne a_2 \ne a_3$ and $0 \le t \le \widehat{t} \le T$, it holds

$$\begin{aligned}
&k(t,a_1,a_1) = 0, \qquad 0 < k(\widehat{t},a_1,a_2) \le k(t,a_1,a_2),\\
&k(t,a_1,a_3) \le k(t,a_1,a_2) + k(t,a_2,a_3).
\end{aligned} \tag{2.7}$$

(C3) There exist constants $\ell_0, b_0 > 0$ and $\mu \in (0,1]$, such that for any $0 \le t \le \widehat{t} \le T$ and $\xi, \widehat{\xi} \in K$, it holds

$$\begin{aligned}
&\ell_0 \le \ell(\widehat{t},\xi) \le \ell(t,\xi),\\
&\ell(t,\xi) \ge b_0|\xi|^\mu,\\
&\ell(t,\xi + \widehat{\xi}) \le \ell(t,\xi) + \ell(t,\widehat{\xi}).
\end{aligned} \tag{2.8}$$

(C4) The constants μ, ν, δ, γ appearing in (C1) and (C3) satisfy

$$\delta + \gamma < \mu, \qquad \nu \le \mu. \tag{2.9}$$

Next, we introduce the following notion of control processes.

Definition 2.1. (i) An $\mathcal{F}_{t,s}$-adapted process $u(\cdot)$ is called an admissible (continuous) control process on $[t,T]$ if it takes values in U almost surely.

(ii) An admissible switching process on $[t,T]$ with initial value a_0 is defined to be a pair of sequences $\{a_i, \theta_i\}_{i\ge 0}$, which could be finite or infinite, such that each θ_i is an $\mathcal{F}_{t,\cdot}$-stopping time with

$$t = \theta_0 \le \theta_1 \le \theta_2 \le \cdots \le T, \qquad \text{a.s.} \tag{2.10}$$

and each a_i is $\mathcal{F}_{t,\theta_i}$ measurable with values in A and $a_{i-1} \ne a_i$, and

$$E\{\sum_{i\ge 1} k(\theta_i, a_{i-1}, a_i)\} < \infty. \tag{2.11}$$

(iii) An admissible impulse control process on $[t,T]$ is defined to be

$$\xi(s) = \sum_{j\ge 1} \xi_j \chi_{[\tau_j, T]}(s), \qquad t \le s \le T, \tag{2.12}$$

where each τ_j is an $\mathcal{F}_{t,\cdot}$-stopping time with

$$t \le \tau_1 \le \tau_1 \le \tau_2 \le \cdots \le T, \qquad \text{a.s.} \tag{2.13}$$

each ξ_j is $\mathcal{F}_{t,\tau_j}$ measurable with values in K, and

$$E\{\sum_{j\ge 1} \ell(\tau_j, \xi_j)\} < \infty. \tag{2.14}$$

We let $\mathcal{U}[t,T]$, $\mathcal{A}^a[t,T]$ and $\mathcal{K}[t,T]$ be the set of all admissible continuous control processes, switching control processes with initial value $a \in A$ and impulse control processes on $[t,T]$, respectively.

Hereafter, we will identify $\{a_i, \theta_i\}_{i\ge 0} \in \mathcal{A}^a[t,T]$ with

$$a(s) = \sum_{i\ge 1} a_{i-1} \chi_{[\theta_{i-1}, \theta_i)}(s), \qquad s \in [t,T]. \tag{2.15}$$

Here, we should note that even in the case, say, $\theta_1 = \theta_2$, which makes the term $a_1\chi_{[\theta_1,\theta_2)}(s)$ void in (2.15), we still keep it in the expression (2.15). This is due to the fact that the sequence $\{a_i, \theta_i\}$ with or without (a_1, θ_1) represents two different switching controls and their costs are different. We will see that by introducing such an identification, notations are simplified (compare [5,18]). Also, we should note that an impulse control with *no* impulse at, say, $s = \widehat{\tau}$, and with a *zero* impulse at $s = \widehat{\tau}$ are different due to the positivity of the impulse cost. Hereafter, we refer the impulse control with no impulses at all as the trivial impulse control and the switching control with no switches at all as the trivial switching control.

Now, for any $(x, t, a) \in \mathbb{R}^n \times [0, T) \times A$ and $(u(\cdot), a(\cdot), \xi(\cdot)) \in \mathcal{U}[t, T] \times \mathcal{A}^a[t, T] \times \mathcal{K}[t, T]$, there exists a unique strong solution of the following stochastic system

$$(2.16)\quad y_{t,x}(s) = x + \int_t^s f(r, y_{t,x}(r), u(r), a(r))dr + \xi(s) + \int_t^s \sigma(r, y_{t,x}(r), u(r), s(r))dw_r.$$

Here, the subscripts in $y_{t,x}(\cdot)$ emphasize the dependence of the state on (x, t). Of course, we should keep in mind that it also depends on the control $(u(\cdot), a(\cdot), \xi(\cdot))$. Next, we define the cost functional as in (1.3):

$$(2.17)\quad \begin{aligned} J^a_{t,x}(u(\cdot), a(\cdot), \xi(\cdot)) = E\{h(y_{t,x}(T)) + \int_t^T f^0(r, y_{t,x}(r), u(r), a(r))dr \\ + \sum_{i\ge1} k(\theta_i, a_{i-1}, a_i) + \sum_{j\ge1} \ell(\tau_j, \xi_j)\}, \end{aligned}$$

where $\{a_i, \theta_i\}$ and $\{\xi_j, \tau_j\}$ are associated with the controls $a(\cdot)$ and $\xi(\cdot)$. In the above, the four terms on the right hand side are referred as final state penalty, running cost, switching cost and impulse cost, respectively. We should note that due to the positivity of the nontrivial switching (i.e., $k(t, a, \widehat{a}) > 0$ for $a \neq \widehat{a}$), different initial values of the switching controls yield different values of the cost functional. Thus, the cost functional is labeled by superscript a. Now, we can state our optimal control problem.

Problem CSI. For any $(t, x, a) \in [0, T) \times \mathbb{R}^n \times A$, find a triplet $(\bar{u}(\cdot), \bar{a}(\cdot), \bar{\xi}(\cdot)) \in \mathcal{U}[t, T] \times \mathcal{A}^a[t, T] \times \mathcal{K}[t, T]$, such that

$$(2.18)\qquad V^a(t, x) \equiv J^a_{t,x}(\bar{u}(\cdot), \bar{a}(\cdot), \bar{\xi}(\cdot)) = \inf_{\mathcal{U}[t,T]\times\mathcal{A}^a[t,T]\times\mathcal{K}[t,T]} J^a_{t,x}(u(\cdot), a(\cdot), \xi(\cdot)).$$

The function $V(t,x) \equiv (V^1(t,x),\cdots,V^m(t,x))$ is called the value function of Problem CSI.

To conclude this section, let us state a lemma found in [9]. This result is technically needed in sequel.

Lemma 2.2. *For any bounded continuous function* $\varphi(\cdot)$*, any* $\beta(\cdot) \equiv (u(\cdot),a(\cdot),\xi(\cdot)) \in \mathcal{U}[t,T] \times \mathcal{A}^a[t,T] \times \mathcal{K}[t,T]$ *and any* $\tau \in [s,T]$*,*

$$E_{t,x}[\varphi(y_{t,x}(\tau),\beta(\tau))|\mathcal{F}_{t,s}] = E_{s,y_{t,x}(s)}[\varphi(y_{t,x}(\tau),(\beta\circ\Pi^{-1})(\tau))], \quad \mathcal{P}_{t,s}-\text{a.s.}\ . \tag{2.19}$$

§3. Continuity of the Value Function.

In this section, we prove the continuity of the value function. The main result of this section can be stated as follows.

Theorem 3.1. *Let (C1)–(C4) hold. Then, there exists a constant* $C>0$*, such that for all* $t,\widehat t \in [0,T]$*,* $x,\widehat x \in \mathbb{R}^n$ *and* $a\in A$*,*

$$-L(T+1) \le V^a(t,x) \le C(1+|x|^{\gamma+\delta}). \tag{3.1}$$

$$|V^a(t,x)-V^a(\widehat t,\widehat x)| \le C\big[(1+|x|^\mu+|\widehat x|^\mu)|t-\widehat t|^{\delta/2}+(1+|x|^\gamma+|\widehat x|^\gamma)|x-\widehat x|^\delta\big]. \tag{3.2}$$

The proof is long and technical. We separate it into several lemmas. For $t_0 \le t$, we let $\mathcal{R}^n_{t_0,t}$ be the set of all $\mathbb{R}^n$-valued $\mathcal{F}_{t_0,t}$-measurable random variables. We see that $\mathcal{R}^n_{t,t}=\mathcal{R}^n$. It is clear that under (C1), for any $(t,a)\in[0,T)\times A$, $t_0\in[0,t]$, $x\in\mathcal{R}^n_{t_0,t}$ and $(u(\cdot),a(\cdot),\xi(\cdot))\in\mathcal{U}[t,T]\times\mathcal{A}^a[t,T]\times\mathcal{K}[t,T]$, (2.16) admits a unique solution $y_{t,x}(\cdot)$. We first present some preliminary results on $y_{t,x}(\cdot)$. Our first lemma is concerned with the certain "uniform continuity" of $y_{t,x}(\cdot)$ in x.

Lemma 3.2. *For any* $\beta>0$*, there exists a constant* $C>0$*, such that for all* $(t,a)\in[0,T)\times A$*, all* $x,\widehat x\in\mathcal{R}^n_{t_0,t}$ *(with* $t_0\in[0,t]$*) and* $(u(\cdot),a(\cdot),\xi(\cdot))\in\mathcal{U}[t,T]\times\mathcal{A}^a[t,T]\times\mathcal{K}[t,T]$*,*

$$E|y_{t,x}(s)-y_{t,\widehat x}(s)|^\beta \le CE|x-\widehat x|^\beta, \qquad \forall s\in[t,T]. \tag{3.3}$$

Proof. For any $\varepsilon > 0$, we let $\langle\, x \,\rangle_\varepsilon = \sqrt{\varepsilon^2 + |x|^2}$. Then, by Itô's formula, we have

$$
\begin{aligned}
E\langle\, y_{t,x}(s) - y_{t,\widehat{x}}(s) \,\rangle_\varepsilon^\beta &\le E\langle\, x - \widehat{x} \,\rangle_\varepsilon^\beta + \beta L \int_t^s \langle\, y_{t,x}(r) - y_{t,\widehat{x}}(r) \,\rangle_\varepsilon^\beta dr \\
&+ \frac{\beta}{2}(1 + |\beta - 2|) L^2 \int_t^s E\langle\, y_{t,x}(r) - y_{t,\widehat{x}}(r) \,\rangle_\varepsilon^\beta dr.
\end{aligned} \tag{3.4}
$$

Thus, by Gronwall's inequality,

$$
E\langle\, y_{t,x}(s) - y_{t,\widehat{x}}(s) \,\rangle_\varepsilon^\beta \le C E\langle\, x - \widehat{x} \,\rangle_\varepsilon^\beta. \tag{3.5}
$$

Finally, we let $\varepsilon \to 0$ to obtain (3.3). $\square$

We note that in our problem, the bounds of the state $y_{t,x}(\cdot)$ heavily depends on the impulse control $\xi(\cdot)$ (as well as the initial state x). The following lemma, needed for continuity of the value function, gives bounds of $E|y_{t,x}(s)|^\beta$ in terms of x and $\xi(\cdot)$.

Lemma 3.3. *Let (C1) hold. Then, for any $\beta > 0$, there exists a constant $C > 0$, such that for any $(t, a) \in [0, T) \times A$, $x \in \mathcal{R}^n_{t_0,t}$ (with $t_0 \in [0, t]$) and $(u(\cdot), a(\cdot), \xi(\cdot)) \in \mathcal{U}[t, T] \times \mathcal{A}^a[t, T] \times \mathcal{K}[t, T]$, the solution $y_{t,x}(\cdot)$ of (2.16) satisfies the following:*

$$
\begin{aligned}
E|y_{t,x}(s)|^\beta &\le C\{1 + E|x|^\beta + E|\xi(s)|^\beta + E\Big(\int_t^s |\xi(r)| dr\Big)^\beta\}, \\
&\forall s \in [t, T], \quad \beta \in [\nu, \infty).
\end{aligned} \tag{3.6}
$$

$$
\begin{aligned}
E|y_{t,x}(s)|^\beta &\le C\{1 + (E|x|^\nu)^{\frac{\beta}{\nu}} + \big(E|\xi(s)|^\nu\big)^{\frac{\beta}{\nu}} + \big(E\Big(\int_t^s |\xi(r)| dr\Big)^\nu\big)^{\frac{\beta}{\nu}}\}, \\
&\forall s \in [t, T], \quad \beta \in (0, \nu).
\end{aligned} \tag{3.7}
$$

Proof. For notational simplicity, hereafter, the C represents a generic constant in different places. Also, in what follows, we suppress the subscripts in $y_{t,x}(\cdot)$. From (2.16) and (2.6), we have

$$
\begin{aligned}
1 + |y(s)| &\le 1 + |x| + |\xi(s)| + \Big|\int_t^s \sigma(r, y(r)) dw_r\Big| + L\int_t^s (1 + |y(r)|) dr, \\
&s \in [t, T], \quad \text{a.s. } \omega \in \Omega.
\end{aligned} \tag{3.8}
$$

where $\sigma(r, y) \equiv \sigma(r, y, u(r), a(r))$. In what follows, we suppress "a.s. $\omega \in \Omega$". Then following the standard proof of Gronwall's inequality, we are able to obtain

$$
\begin{aligned}
1 + |y(s)| &\le 1 + |x| + |\xi(s)| + \Big|\int_t^s \sigma(r, y(r)) dw_r\Big|, \\
&+ L\int_t^s e^{L(s-\tau)}[1 + |x| + |\xi(\tau)| + \Big|\int_t^\tau \sigma(r, y(r)) dw_r\Big|] d\tau.
\end{aligned} \tag{3.9}
$$

Thus, it follows that

$$
\begin{aligned}
|y(s)| \le & C\{1+|x|+|\xi(s)|+\int_t^s |\xi(r)|dr \\
& +|\int_t^s \sigma(r,y(r))dw_r|+\int_t^s|\int_t^\tau \sigma(r,y(r))dw_r|d\tau\}.
\end{aligned} \tag{3.10}
$$

Then, for any $\beta \in (0,\infty)$, we have

$$
\begin{aligned}
E|y(s)|^\beta \le & C\{1+E|x|^\beta+E|\xi(s)|^\beta+E\big(\int_t^s |\xi(r)|dr\big)^\beta \\
& +E|\int_t^s \sigma(r,y(r))dw_r|^\beta+E\big(\int_t^s|\int_t^\tau \sigma(r,y(r))dw_r|d\tau\big)^\beta\}.
\end{aligned} \tag{3.11}
$$

Now, we separate three cases. First, we let $\nu \le \beta \le 2$. Then, we have (by (2.6))

$$
\begin{aligned}
& E|\int_t^s \sigma(r,y(r))dw_r|^\beta+E\big(\int_t^s|\int_t^\tau \sigma(r,y(r))dw_r|d\tau\big)^\beta \\
& \le \big(E|\int_t^s \sigma(r,y(r))dw_r|^2\big)^{\beta/2}+(s-t)^{\beta/2}\big(\int_t^s E|\int_t^\tau \sigma(r,y(r))dw_r|^2 d\tau\big)^{\beta/2} \\
& = \big(\int_t^s E|\sigma(r,y(r))|^2 dr\big)^{\beta/2}+(s-t)^{\beta/2}\big(\int_t^s\int_t^\tau E|\sigma(r,y(r))|^2 dr d\tau\big)^{\beta/2} \\
& \le [1+(s-t)^\beta]\big(\int_t^s E|\sigma(r,y(r))|^2 dr\big)^{\beta/2} \le C\{(s-t)^{\beta/2}+\big(\int_t^s E|y(r)|^\nu dr\big)^{\beta/2}\} \\
& \le C\{(s-t)^{\beta/2}+(s-t)^{\frac{\beta-\nu}{2}}\big(\int_t^s E|y(r)|^\beta dr\big)^{\nu/2}\} \le C\{1+\int_t^s E|y(r)|^\beta dr\}.
\end{aligned} \tag{3.12}
$$

Hence, it follows from (3.11)–(3.12) that

$$
E|y(s)|^\beta \le C\{1+E|x|^\beta+E|\xi(s)|^\beta+E\big(\int_t^s |\xi(r)|dr\big)^\beta+\int_t^s E|y(r)|^\beta dr\}. \tag{3.13}
$$

Then, by Gronwall's inequality, we obtain (3.6) for the case $\beta \in [\nu, 2]$. Now, we let $\beta > 2$. Denote

$$
\eta(s)=\int_t^s \sigma(r,y(r))dw_r.
$$

Then, by Itô's formula and Young's inequality, we have

$$
\begin{aligned}
E|\eta(s)|^\beta = & \beta E \mathrm{tr}\,\Big\{\int_t^s \big[|\eta(r)|^{\beta-2}I \\
& +(\beta-2)|\eta(r)|^{\beta-4}\eta(r)\eta(r)^T\big]\sigma(r,y(r))\sigma(r,y(r))^T dr\Big\} \\
\le & C\int_t^s E|\eta(r)|^{\beta-2}(1+|y(r)|^\nu)dr \\
\le & C\int_t^s \{1+E|\eta(r)|^\beta+E|y(r)|^\beta\}dr.
\end{aligned} \tag{3.14}
$$

Then, it follows from the Gronwall's inequality that

$$E|\eta(s)|^\beta \leq C\{1+\int_t^s E|y(r)|^\beta dr\}. \tag{3.15}$$

Hence, we obtain

$$\begin{aligned}
&E|\int_t^s \sigma(r,y(r))dw_r|^\beta + E\big(\int_t^s |\int_t^\tau \sigma(r,y(r))dw_r|d\tau\big)^\beta \\
&\equiv E|\eta(s)|^\beta + E\big(\int_t^s |\eta(\tau)|d\tau\big)^\beta \\
&\leq C\big\{1+\int_t^s E|y(r)|^\beta dr + \int_s^t E|\eta(\tau)|^\beta d\tau\big\} \\
&\leq C\big\{1+\int_t^s E|y(r)|^\beta dr + \int_s^t \big[1+\int_t^\tau E|y(r)|^\beta dr\big]d\tau\big\} \\
&\leq C\{1+\int_t^s E|y(r)|^\beta dr\}.
\end{aligned} \tag{3.16}$$

This is the same form as (3.12). Then, similar to the case $\beta \in [\nu, 2]$, we obtain (3.6) for $\beta > 2$. Finally, for the case $\beta \in (0, \nu)$, by (3.6), we have

$$\begin{aligned}
&E|y(s)|^\beta \leq \{E|y(s)|^\nu\}^{\frac{\beta}{\nu}} \\
&\leq C\{1+E|x|^\nu + E|\xi(s)|^\nu + E\big(\int_t^s |\xi(r)|dr\big)^\nu\}^{\frac{\beta}{\nu}} \\
&\leq C\big\{1+(E|x|^\nu)^{\frac{\beta}{\nu}} + [E|\xi(s)|^\nu]^{\frac{\beta}{\nu}} + \big[E\big(\int_t^s |\xi(r)|dr\big)^\nu\big]^{\frac{\beta}{\nu}}\big\}.
\end{aligned} \tag{3.17}$$

This proves (3.7). □

The following corollary is obvious.

Corollary 3.4. *Let $(t,x,a) \in [0,T) \times \mathbb{R}^n \times A$ and $(u(\cdot), a(\cdot), \xi_0(\cdot)) \in \mathcal{U}[t,T] \times \mathcal{A}^a[t,T] \times \mathcal{K}[t,T]$ with $\xi_0(\cdot) = 0$ being the trivial impulse control. Then for any $\beta \in (0,\infty)$, there exists a constant $C > 0$, such that*

$$E|y_{t,x}(s)|^\beta \leq C(1+|x|^\beta), \qquad s \in [t,T]. \tag{3.18}$$

The following result is crucial in proving the continuity of the value function in time variable t.

Lemma 3.5. *For any $\beta \geq \nu$, there exists a constant $C > 0$, such that for any $(t, x, a) \in [0, T) \times \mathbb{R}^n \times A$ and $(u(\cdot), a(\cdot), \xi(\cdot)) \in \mathcal{U}[t, T] \times \mathcal{A}^a[t, T] \times \mathcal{K}[t, T]$, it holds*

$$
\begin{aligned}
E|y_{t,x}(s) - x - \xi(s)|^\beta \leq C\big\{(1 + |x|^\beta)(s - t)^{\frac{\beta}{2}\wedge 1} + E\Big(\int_t^s |\xi(r)|^{\beta\vee 1} dr\Big)^{\beta\wedge 1}\big\} \\
\forall s \in [t, T],
\end{aligned}
\tag{3.19}
$$

where $a \wedge b = \min\{a, b\}$ and $a \vee b = \max\{a, b\}$. In particular, if $\xi(\cdot)$ is the trivial impulse control, then,

$$
E|y_{t,x}(s) - x|^\beta \leq C(1 + |x|^\beta)(s - t)^{\frac{\beta}{2}\wedge 1}. \tag{3.20}
$$

Proof. We suppress the subscripts t, x in $y_{t,x}(\cdot)$ and set

$$
z(s) \equiv y(s) - x - \xi(s) = \int_t^s f(r, y(r))dr + \int_t^s \sigma(r, y(r))dw_r, \tag{3.21}
$$

with

$$
f(r, y) \equiv f(r, y, u(r), a(r)), \qquad \sigma(r, y) \equiv \sigma(r, y, u(r), a(r)).
$$

This yields that

$$
\begin{aligned}
|z(s)| &\leq L \int_t^s (1 + |y(r)|)dr + \Big|\int_t^s \sigma(r, y(r))dw_r\Big| \\
&\leq L \int_t^s (1 + |x| + |\xi(r)| + |z(r)|)dr + \Big|\int_t^s \sigma(r, y(r))dw_r\Big|.
\end{aligned}
\tag{3.22}
$$

By Gronwall's inequality, we obtain

$$
\begin{aligned}
|z(s)| &\leq L \int_t^s e^{L(s-r)}(1 + |x| + |\xi(r)|)dr + \Big|\int_t^s \sigma(r, y(r))dw_r\Big| \\
&\quad + \int_t^s e^{L(s-\tau)}\Big|\int_t^\tau \sigma(r, y(r))dw_r\Big|d\tau \\
&\leq C\{(1 + |x|)(s - t) + \int_t^s |\xi(r)|dr + \Big|\int_t^s \sigma(r, y(r))dw_r\Big| \\
&\quad + \int_t^s \Big|\int_t^\tau \sigma(r, y(r))dw_r\Big|d\tau.
\end{aligned}
\tag{3.23}
$$

Then, note $\beta \geq \nu$, by (3.12), we obtain

$$
\begin{aligned}
&E|z(s)|^\beta \leq C\big\{(1 + |x|^\beta)(s - t)^\beta + E\Big(\int_t^s |\xi(r)|dr\Big)^\beta + (s - t)^{\beta/2} + \int_t^s E|y(r)|^\beta dr\big\} \\
&\leq C\big\{(1 + |x|^\beta)(s - t)^{\beta/2} + E\Big(\int_t^s |\xi(r)|dr\Big)^\beta \\
&\quad + \int_t^s \big[|x|^\beta + E|z(r)|^\beta + E|\xi(r)|^\beta\big]dr\big\}.
\end{aligned}
\tag{3.24}
$$

By Gronwall's inequality again,

$$
\begin{aligned}
E|z(s)|^\beta &\le C\big\{(1+|x|^\beta)(s-t)^{\beta/2}+|x|^\beta(s-t) \\
&\quad + E\Big(\int_t^s|\xi(r)|dr\Big)^\beta+\int_t^s E|\xi(r)|^\beta dr\big\} \\
&\le C\big\{(1+|x|^\beta)(s-t)^{\frac{\beta}{2}\wedge 1}+E\Big(\int_t^s|\xi(r)|^{\beta\vee 1}dr\Big)^{\beta\wedge 1}\big\}.
\end{aligned}
\tag{3.25}
$$

This proves (3.19). Finally, (3.20) follows easily. □

Let us now start the proof of Theorem 3.1.

Proof of (3.1). Let $(t,x,a)\in[0,T)\times\mathbb{R}^n\times A$ and fix any $u(\cdot)\in\mathcal{U}[t,T]$. Let $a_0(\cdot)\equiv a$ and $\xi_0(\cdot)\equiv 0$ be the trivial switching control and the trivial impulse control, respectively. Then, by Corollary 3.4, we have

$$
\begin{aligned}
V^a(t,x) &\le J^a_{t,x}(u(\cdot),a_0(\cdot),\xi_0(\cdot)) \\
&= E\int_t^T f^0(r,y_{t,x}(r),u(r),a)dr+Eh(y_{t,x}(T)) \\
&\le C\{1+\int_t^T E|y_{t,x}(r)|^{\gamma+\delta}dr+E|y_{t,x}(T)|^{\gamma+\delta}\} \\
&\le C_0\{1+|x|^{\gamma+\delta}\}.
\end{aligned}
\tag{3.26}
$$

On the other hand, for any $(u(\cdot),a(\cdot),\xi(\cdot))\in\mathcal{U}[t,T]\times\mathcal{A}^a[t,T]\times\mathcal{K}[t,T]$, by (C1), we have

$$
J^a_{t,x}(u(\cdot),a(\cdot),\xi(\cdot))\ge -L(T+1).
\tag{3.27}
$$

Thus, the other half of (3.1) follows. □

Next, for any $\rho\in[0,\infty)$, we define

$$
\begin{aligned}
\mathcal{K}_\rho[t,T]=&\{\xi(\cdot)\equiv\sum_{j\ge 1}\xi_j\chi_{[\tau_j,T]}(\cdot)\in\mathcal{K}[t,T]\ \big| \\
&\qquad E\sum_{j\ge 1}\ell(\tau_j,\xi_j)\le C_0(1+\rho^{\gamma+\delta})+2L(T+1)+1\},
\end{aligned}
\tag{3.28}
$$

where C_0 is the constant determined by (3.26). We have the following result which is important in the proof of (3.2).

Lemma 3.6. *For any $(t,x,a)\in[0,T)\times\mathbb{R}^n\times A$ with $|x|\le\rho$,*

$$
V^a(t,x)=\inf_{\mathcal{U}[t,T]\times\mathcal{A}^a[t,T]\times\mathcal{K}_\rho[t,T]}J^a_{t,x}(u(\cdot),a(\cdot),\xi(\cdot)).
\tag{3.29}
$$

Proof. Let $\xi(\cdot) \in \mathcal{K}[t,T] \setminus \mathcal{K}_\rho[t,T]$. Then, for any $(u(\cdot), a(\cdot)) \in \mathcal{U}[t,T] \times \mathcal{A}^a[t,T]$, we have

$$\begin{aligned} J^a_{t,x}(u(\cdot), a(\cdot), \xi(\cdot)) &\geq -L(T+1) + E\sum_{j\geq 1} \ell(\tau_j, \xi_j) \\ &> C_0(1+\rho^{\gamma+\delta}) + L(T+1) + 1 \geq V^a(t,x) + 1. \end{aligned} \tag{3.30}$$

Hence, (3.29) follows. $\square$

One of the important consequence of above Lemma 3.6 is the following.

Corollary 3.7. *Let (C1)–(C4) hold. Then, there exists a constant $C > 0$, such that for any $(t,x,a) \in [0,T) \times \mathbb{R}^n \times A$ with $|x| \leq \rho$ and any $(u(\cdot), a(\cdot), u(\cdot)) \in \mathcal{U}[t,T] \times \mathcal{A}^a[t,T] \times \mathcal{K}_\rho[t,T]$, it holds that*

$$\max_{t\leq s\leq T} E|\xi(s)|^\mu \leq C(1+\rho^{\gamma+\delta}). \tag{3.31}$$

$$E(\int_t^s |\xi(r)|dr)^\mu \leq C(1+\rho^{\gamma+\delta})(s-t)^\mu. \tag{3.32}$$

$$\max_{t\leq s\leq T} E|y_{t,x}(s)|^\mu \leq C(1+\rho^\mu). \tag{3.33}$$

Proof. By (C3) and the above definition, for any $\xi(\cdot) \in \mathcal{K}_\rho[t,T]$, we have (note $0 < \mu \leq 1$)

$$\begin{aligned} C_0(1+\rho^{\gamma+\delta}) + 2L(T+1) + 1 &\geq E\sum_{j\geq 1} \ell(\tau_j, \xi_j) \\ &\geq b_0 E\big(\sum_{j\geq 1} |\xi_j|^\mu\big) \geq b_0 E\big(\sum_{j\geq 1} |\xi_j|\big)^\mu \\ &\geq b_0 E\big(\max_{t\leq s\leq T} |\xi(s)|^\mu\big) \geq b_0 \max_{t\leq s\leq T} E|\xi(s)|^\mu. \end{aligned} \tag{3.34}$$

This gives (3.31). Next, by the above, we have

$$\begin{aligned} E(\int_t^s |\xi(r)|dr)^\mu &\leq E\big(\max_{t\leq r\leq s} |\xi(r)|^\mu\big)(s-t)^\mu \\ &\leq \frac{1}{b_0}[C_0(1+\rho^{\gamma+\delta}) + 2L(T+1) + 1](s-t)^\mu. \end{aligned} \tag{3.35}$$

This proves (3.32). Finally, by Lemma 3.3 (note (C4)), we obtain

$$
\begin{aligned}
E|y_{t,x}(s)|^\mu &\le C\{1+|x|^\mu + E|\xi(s)|^\mu + E(\int_t^s |\xi(r)|dr)^\mu\} \\
&\le C\{1+|x|^\mu + \frac{1}{b_0}[C_0(1+\rho^{\gamma+\delta} + 2L(T+1)+1] \\
&+ \frac{T^\mu}{b_0}[C_0(1+\rho^{\gamma+\delta}) + 2L(T+1)+1]\} \\
&\le C(1+\rho^\mu).
\end{aligned} \tag{3.36}
$$

Hence, (3.33) follows. $\square$

From Lemma 3.5 and Corollary 3.7, we have the following.

Corollary 3.8. *There exists a constant $C>0$, such that for all $(t,x,a)\in[0,T)\times\mathbb{R}^n\times A$ and $(u(\cdot),a(\cdot),\xi(\cdot))\in\mathcal{U}[t,T]\times\mathcal{A}^a[t,T]\times\mathcal{K}_{|x|}[t,T]$, it holds*

$$
E|y_{t,x}(s)-x-\xi(s)|^\mu \le C(1+|x|^\mu)(s-t)^{\mu/2}, \qquad s\in[t,T]. \tag{3.37}
$$

Now, let us come back to the proof of Theorem 3.1.

Proof of (3.2). For any $(t,a)\in[0,T)\times A$, $x,\widehat{x}\in\mathbb{R}^n$ and $(u(\cdot),a(\cdot),\xi(\cdot))\in\mathcal{U}[t,T]\times\mathcal{A}^a[t,T]\times\mathcal{K}_{|x|\vee|\widehat{x}|}[t,T]$, by Lemma 3.2 and Corollary 3.7, we have

$$
\begin{aligned}
&|J^a_{t,x}(u(\cdot),a(\cdot),\xi(\cdot)) - J^a_{t,\widehat{x}}(u(\cdot),a(\cdot),\xi(\cdot))| \\
&\le L\int_t^T E\big\{\big(1+|y_{t,x}(r)|^\gamma + |y_{t,\widehat{x}}(r)|^\gamma\big)|y_{t,x}(r)-y_{t,\widehat{x}}(r)|^\delta\big\}dr \\
&\quad + LE\big\{\big(1+|y_{t,x}(T)|^\gamma + |y_{t,\widehat{x}}(T)|^\gamma\big)|y_{t,x}(T)-y_{t,\widehat{x}}(T)|^\delta\big\} \\
&\le L\{\int_t^T \big[E|y_{t,x}(r)-y_{t,\widehat{x}}(r)|^\delta + \big((E|y_{t,x}(r)|^\mu)^{\gamma/\mu} \\
&\quad + (E|y_{t,\widehat{x}}(r)|^\mu)^{\frac{\gamma}{\mu}}\big)(E|y_{t,x}(r)-y_{t,\widehat{x}}(r)|^\mu)^{\frac{\delta}{\mu}}\big]dr \\
&\quad + \big[E|y_{t,x}(T)-y_{t,\widehat{x}}(T)|^\delta + \big((E|y_{t,x}(T)|^\mu)^{\gamma/\mu} \\
&\quad + (E|y_{t,\widehat{x}}(T)|^\mu)^{\frac{\gamma}{\mu}}\big)(E|y_{t,x}(T)-y_{t,\widehat{x}}(T)|^\mu)^{\frac{\delta}{\mu}}\big]\} \\
&\le C\big\{|x-\widehat{x}|^\delta + (1+|x|^\mu)^{\gamma/\mu}|x-\widehat{x}|^\delta + (1+|\widehat{x}|^\mu)^{\gamma/\mu}|x-\widehat{x}|^\delta\big\} \\
&\le C(1+|x|^\mu+|\widehat{x}|^\mu)|x-\widehat{x}|^\delta.
\end{aligned} \tag{3.38}
$$

This proves (3.2) for the case $t=\widehat{t}$. Now, we prove the continuity of the value function in time t. To this end, we let $0\le t<\widehat{t}\le T$ and let $(x,a)\in\mathbb{R}^n\times A$. Then, for any

$(u(\cdot), a(\cdot), \xi(\cdot)) \in \mathcal{U}[\widehat{t}, T] \times \mathcal{A}^a[\widehat{t}, T] \times \mathcal{K}_{|x|}[\widehat{t}, T]$, we extend $u(s) = u_0$, with $u_0 \in U$, $a(s) = a$ and $\xi(s) = 0$ for $s \in [t, \widehat{t})$ (i.e., there are no switches and impulses on $[t, \widehat{t})$). Then, we see that $(u(\cdot), a(\cdot), \xi(\cdot)) \in \mathcal{U}[t, T] \times \mathcal{A}^a[t, T] \times \mathcal{K}_{|x|}[t, T]$. Thus, by (3.18) and (3.20),

$$
\begin{aligned}
V^a(t,x) &\le J^a_{t,x}(u(\cdot), a(\cdot), \xi(\cdot)) \\
&= J^a_{\widehat{t},x}(u(\cdot), a(\cdot), \xi(\cdot)) + E\int_t^{\widehat{t}} f^0(r, y_{t,x}(r), u_0, a)dr \\
&\quad + E\int_{\widehat{t}}^T [f^0(r, y_{t,x}(r), u(r), a(r)) - f^0(r, y_{\widehat{t},x}(r), u(r), a(r))]dr \\
&\quad + E[h(y_{t,x}(T)) - h(y_{\widehat{t},x}(T))]
\end{aligned}
\tag{3.39}
$$

$$
\begin{aligned}
&\le J^a_{\widehat{t},x}(u(\cdot), a(\cdot), \xi(\cdot)) + L\int_t^{\widehat{t}} (1 + E|y_{t,x}(r)|^{\gamma+\delta})dr \\
&\quad + L\int_{\widehat{t}}^T E(1 + |y_{t,x}(r)|^\gamma + |y_{\widehat{t},x}(r)|^\gamma)|y_{t,x}(r) - y_{\widehat{t},x}(r)|^\delta dr \\
&\quad + LE(1 + |y_{t,x}(T)|^\gamma + |y_{\widehat{t},x}(T)|^\gamma)|y_{t,x}(T) - y_{\widehat{t},x}(T)|^\delta \\
&\le J^a_{\widehat{t},x}(u(\cdot), a(\cdot), \xi(\cdot)) + C(1 + |x|^\mu)(\widehat{t} - t) \\
&\quad + C\int_{\widehat{t}}^T \big[1 + \big(E|y_{t,x}(r)|^\mu\big)^{\gamma/\mu} + \big(E|y_{\widehat{t},x}(r)|^\mu\big)^{\gamma/\mu}\big]\big(E|y_{t,x}(r) - y_{\widehat{t},x}(r)|^\mu\big)^{\delta/\mu} dr \\
&\quad + C\big[1 + \big(E|y_{t,x}(T)|^\mu\big)^{\gamma/\mu} + \big(E|y_{\widehat{t},x}(T)|^\mu\big)^{\gamma/\mu}\big]\big(E|y_{t,x}(T) - y_{\widehat{t},x}(T)|^\mu\big)^{\delta/\mu} \\
&\le J^a_{\widehat{t},x}(u(\cdot), a(\cdot), \xi(\cdot)) + C(1 + |x|^\mu)(\widehat{t} - t) \\
&\quad + C(1 + |x|^\gamma)\big(E|y_{t,x}(\widehat{t} - 0) - x|^\mu\big)^{\delta/\mu} \\
&\le J^a_{\widehat{t},x}(u(\cdot), a(\cdot), \xi(\cdot)) + C(1 + |x|^\mu)(\widehat{t} - t) \\
&\quad + C(1 + |x|^\mu)[(1 + |x|^\mu)(\widehat{t} - t)^{\mu/2}]^{\delta/\mu} \\
&\le J^a_{\widehat{t},x}(u(\cdot), a(\cdot), \xi(\cdot)) + C(1 + |x|^\mu)(\widehat{t} - t)^{\delta/2}.
\end{aligned}
$$

Hence,

$$
V^a(t,x) \le V^a(\widehat{t}, x) + C(1 + |x|^\mu)(\widehat{t} - t)^{\delta/2}. \tag{3.40}
$$

Conversely, for any $\varepsilon > 0$, there exists a $(u(\cdot), a(\cdot), \xi(\cdot)) \in \mathcal{U}[t, T] \times \mathcal{A}^a[t, T] \times \mathcal{K}_{|x|}[t, T]$, such that

$$
\varepsilon + V^a(t,x) \ge J^a_{t,x}(u(\cdot), a(\cdot), \xi(\cdot)). \tag{3.41}
$$

We let

$$
\left\{
\begin{aligned}
& a(\cdot) = \sum_{i\ge 1} a_{i-1}\chi_{[\theta_{i-1},\theta_i)}(\cdot), \\
& \xi(\cdot) = \sum_{j\ge 1} \xi_j \chi_{[\tau_j,T]}(\cdot).
\end{aligned}
\right. \tag{3.42}
$$

Define

$$
\widehat a(\cdot) \equiv \sum_{i\ge 1} a_{i-1}\chi_{[\widehat\theta_{i-1},\widehat\theta_i)}(\cdot), \tag{3.43}
$$

with

$$
\widehat\theta_i = \begin{cases} \widehat t, & \text{if } \theta_i \le \widehat t, \\ \theta_i, & \text{if } \theta_i > \widehat t, \end{cases} \tag{3.44}
$$

and

$$
\left\{
\begin{aligned}
& \bar\xi(\cdot) \equiv \sum_{j\ge 1} \xi_j \chi_{[\widehat\tau_j,T]}(\cdot) = \sum_{\tau_j \le \widehat t} \xi_j \chi_{[\widehat t,T]}(\cdot) + \sum_{\tau_j > \widehat t} \xi_j \chi_{[\tau_j,T]}(\cdot), \\
& \widehat\xi(\cdot) \equiv \sum_{j\ge 1} \xi_j \chi_{[\widehat\tau_j,T]}(\cdot) = \sum_{\tau_j > \widehat t} \xi_j \chi_{[\tau_j,T]}(\cdot).
\end{aligned}
\right. \tag{3.45}
$$

Next, we let

$$
\bar x = x + \xi(\widehat t) = x + \sum_{\tau_j \le \widehat t} \xi_j. \tag{3.46}
$$

Clearly, $\bar x \in \mathcal{R}^n_{t,\widehat t}$. Then, we note that

$$
J^a_{\widehat t,x}(u(\cdot), \widehat a(\cdot), \bar\xi(\cdot)) = J^a_{\widehat t,\bar x}(u(\cdot), \widehat a(\cdot), \widehat\xi(\cdot)) + E \sum_{\tau_j \le \widehat t} \ell(\widehat t, \xi_j). \tag{3.47}
$$

Hence, by the monotonicity of $k(\cdot, a, \widehat a)$ and $\ell(\cdot, \xi)$, using Corollaries 3.7 and 3.8 and Lemma

2.2, we have

(3.48)

$$\begin{aligned}
&\varepsilon + V^a(t,x) \geq Eh(y_{t,x}(T)) + E\int_t^T f^0(r, y_{t,x}(r), u(r), a(r))dr \\
&\qquad + E\sum_{i\geq 1} k(\widehat\theta_i, a_{i-1}, a_i) + E\sum_{j\geq 1} \ell(\widehat\tau_j, \xi_j) \\
&\geq J^a_{\widehat t,x}(u(\cdot), \widehat a(\cdot), \bar\xi(\cdot)) + E\int_t^{\widehat t} f^0(r, y_{t,x}(r), u(r), a(r))dr + E\big[h(y_{t,x}(T)) - h(y_{\widehat t,\bar x}(T))\big] \\
&\qquad + E\int_{\widehat t}^T \big[f^0(r, y_{t,x}(r), u(r), \widehat a(r)) - f^0(r, y_{\widehat t,\bar x}(r), u(r), \widehat a(r))\big]dr \\
&\geq V^a(\widehat t, x) - L\int_t^{\widehat t} (1 + E|y_{t,x}(r)|^{\gamma+\delta})dr \\
&\qquad - LE\big[\big(1 + |y_{t,x}(T)|^\gamma + |y_{\widehat t,\bar x}(T)|^\gamma\big)|y_{t,x}(T) - y_{\widehat t,\bar x}(T)|^\delta \\
&\qquad - L\int_{\widehat t}^T E\big[\big(1 + |y_{t,x}(r)|^\gamma + |y_{\widehat t,\bar x}(r)|^\gamma\big)|y_{t,x}(r) - y_{\widehat t,\bar x}(r)|^\delta dr \\
&\geq V^a(\widehat t, x) - C(1 + |x|^{\gamma+\delta})(\widehat t - t) - L(T+1)E|y_{t,x}(\widehat t) - \bar x|^\delta \\
&\qquad - L\big\{1 + (E|y_{t,x}(T)|^\mu)^{\gamma/\mu} + (E|y_{\widehat t,\bar x}(T)|^\mu)^{\gamma/\mu}\big\}(E|y_{t,x}(T) - y_{\widehat t,\bar x}(T)|^\mu)^{\delta/\mu} \\
&\qquad - L\int_{\widehat t}^T \big\{1 + (E|y_{t,x}(r)|^\mu)^{\gamma/\mu} + (E|y_{\widehat t,\bar x}(r)|^\mu)^{\gamma/\mu}\big\}(E|y_{t,x}(r) - y_{\widehat t,\bar x}(r)|^\mu)^{\delta/\mu}dr.
\end{aligned}$$

Note that (see (3.33) and (3.37))

$$\begin{aligned}
E|y_{\widehat t,\bar x}(s)|^\mu &\leq E|y_{\widehat t,\bar x}(s) - y_{t,x}(s)|^\mu + E|y_{t,x}(s)|^\mu \\
&\leq CE|y_{t,x}(\widehat t) - \bar x|^\mu + C(1 + |x|^\mu) \\
&\leq C(1 + |x|^\mu)(\widehat t - t)^\mu + C(1 + |x|^\mu) \\
&\leq C(1 + |x|^\mu).
\end{aligned} \tag{3.49}$$

Therefore, (3.48) is reduced to the following

$$\begin{aligned}
\varepsilon + V^a(t,x) &\geq V^a(\widehat t, x) - C(1 + |x|^{\gamma+\delta})(\widehat t - t) \\
&\qquad - C\big\{(1 + |x|^\mu)^{\delta/\mu}(\widehat t - t)^{\delta/2}[1 + (1 + |x|^\mu)^{\gamma/\mu}]\big\} \\
&\geq V^a(\widehat t, x) - C(1 + |x|^\mu)(\widehat t - t)^{\delta/2},
\end{aligned} \tag{3.50}$$

where we used (C4). $\square$

§4. Bellman's Dynamic Programming Principle.

In this section, we show the validity, as well as some consequences, of the Bellman dynamic programming principle for our control problem. The first result of this section is the following.

Theorem 4.1. *The value function $V(\cdot,\cdot)$ satisfies the following optimality principle: For all $(t,x,a)\in[0,T)\times\mathbb{R}^n\times A$ and $s\in(t,T]$,*

$$
\begin{aligned}
V^a(t,x) = \inf_{\mathcal{U}[t,T]\times\mathcal{A}^a[t,T]\times\mathcal{K}[t,T]} & E\{V^{a(s)}(s,y_{t,x}(s)) + \int_t^s f^0(r,y_{t,x}(r),u(r),a(r))ds \\
& + \sum_{\theta_i\le s} k(\theta_i,a_{i-1},a_i) + \sum_{\tau_j\le s}\ell(\tau_j,\xi_j)\},
\end{aligned}
\tag{4.1}
$$

where $\{\theta_i,a_a\}$ and $\{\tau_j,\xi_j\}$ are associated with $a(\cdot)$ and $\xi(\cdot)$), respectively, and $a(s)=a(s+0)$, $y_{t,x}(s)=y_{t,x}(s+0)$.

Proof. Set the right hand side of (4.1) to be $W(t,x)$. Then, for any $\varepsilon>0$, there exists a control $\bar\beta(\cdot)\equiv(\bar u(\cdot),\bar a(\cdot),\bar\xi(\cdot))\in\mathcal{U}[t,T]\times\mathcal{A}^a[t,T]\times\mathcal{K}[t,T]$, such that

$$
\begin{aligned}
W(t,x)+\varepsilon \ge & E\{V^{\bar a(s)}(s,y_{t,x}(s)) + \int_t^s f^0(r,y_{t,x}(r),\bar u(r),\bar a(r))ds \\
& + \sum_{\bar\theta_i\le s} k(\bar\theta_i,\bar a_{i-1},\bar a_i) + \sum_{\bar\tau_j\le s}\ell(\bar\tau_j,\bar\xi_j)\},
\end{aligned}
\tag{4.2}
$$

where, $\{\bar\theta_i,\bar a_i\}$ and $\{\bar\tau_j,\bar\xi_j\}$ are associated with $\bar a(\cdot)$ and $\bar\xi(\cdot)$, respectively. On the other hand, for every $(s,z,b)\in[0,T]\times\mathbb{R}^n\times A$, there exists a control $\beta_{s,z,b}(\cdot)\equiv(u_{s,z,b}(\cdot),a_{s,z,b}(\cdot),\xi_{s,z,b}(\cdot))\in\mathcal{U}[s,T]\times\mathcal{A}^b[t,T]\times\mathcal{K}[s,T]$, such that

$$
V^b(s,z)\ge J^b_{s,z}(\beta_{s,z,b}(\cdot))-\varepsilon. \tag{4.3}
$$

Take a partition $\{B_i,i\ge1\}$ of $\mathbb{R}^n$, such that each B_i is a Borel set satisfying

$$
|V^b(s,x_1)-V^b(s,x_2)|\le\varepsilon,\qquad \forall b\in A, x_1,x_2\in B_i,\quad i\ge1. \tag{4.4}
$$

The continuity of the value function (see Theorem 3.1) makes the above is possible. Here, we should note that since the value function is not necessarily uniformly continuous, in general, the size of B_i depends on the position of its center and they may shrink as the center goes to infinity. We fix a $\zeta_i\in B_i$ for each $i\ge1$. Now, let us construct a new control triplet as follows:

$$
\widehat\beta(r)=\begin{cases}\bar\beta(r), & r\in[t,s],\\ \sum_{i\ge1}\sum_{j=1}^m\beta_{s,\zeta_i,j}(r)\chi_{B_i}(y_{t,x}(s;\bar\beta(\cdot)))\chi_{\{a(s)=j\}}(s), & r\in(s,T].\end{cases} \tag{4.5}
$$

Then, we see that $\widehat{\beta}(\cdot) \in \mathcal{U}[t,T] \times \mathcal{A}^a[t,T] \times \mathcal{K}[t,T]$ and one has

$$J^a_{t,x}(\widehat{\beta}(\cdot)) \leq W(t,x) + 2\varepsilon. \tag{4.6}$$

This implies

$$V^a(t,x) \leq W(t,x). \tag{4.7}$$

On the other hand, for any $\varepsilon > 0$, from the definition of $V^a(t,x)$, one concludes that there exists a $\widetilde{\beta}(\cdot) \equiv (\widetilde{u}(\cdot), \widetilde{a}(\cdot), \widetilde{\xi}(\cdot)) \in \mathcal{U}[t,T] \times \mathcal{A}^a[t,T] \times \mathcal{K}[t,T]$, such that

$$V^a(t,x) + \varepsilon \geq J^a_{t,x}(\widetilde{\beta}(\cdot)). \tag{4.8}$$

Making use of Lemma 2.2, one has

$$\begin{aligned}
&E\Big[h(y_{t,x}(T;\widetilde{\beta}(\cdot))) + \int_s^T f^0(r, y_{t,x}(r;\widetilde{\beta}(\cdot)), \widetilde{u}(r), \widetilde{a}(r))dr \\
&\qquad + \sum_{\widetilde{\theta}_i > s} k(\widetilde{\theta}_i, \widetilde{a}_{i-1}, \widetilde{a}_i) + \sum_{\widetilde{\tau}_j > s} \ell(\widetilde{\tau}_j, \widetilde{\xi}_j) \;\Big|\; \mathcal{F}_{t,s}\Big] \\
&= J^{\widetilde{a}(s)}_{s, y_{t,x}(s;\widetilde{\beta}(\cdot))}(\widetilde{\beta}(\cdot; \Pi^{-1}(\omega_1, \omega_2))) \\
&\geq V^{\widetilde{a}(s)}(s, y_{t,x}(s;\widetilde{\beta}(\cdot))), \qquad \mathcal{P}_{t,s} - \text{a.s. } \omega_1.
\end{aligned} \tag{4.9}$$

Hence, one concludes that

$$V^a(t,x) + \varepsilon \geq W(t,x), \tag{4.10}$$

and the desired result follows. $\square$

Next, let us define the following operators: For any $\mathbb{R}^m$-valued function $W(\cdot) \equiv (W^1(\cdot), \cdots, W^m(\cdot))$, and $(t,x) \in [0,T] \times \mathbb{R}^n$,

$$M^a[W](t,x) = \min_{b \in A \setminus \{a\}} \{W^b(t,x) + k(t,a,b)\}, \tag{4.11}$$

$$N[W^a](t,x) = \inf_{\xi \in K} \{W^a(t, x+\xi) + \ell(t,\xi)\}. \tag{4.12}$$

They are referred to as the switching and impulse obstacles, respectively. Next, we give a consequence of Theorem 4.1.

Theorem 4.2. *The value function $V(\cdot)$ has the following properties:*

(i) *For any $(t,x,a)\in[0,T]\times\mathbb{R}^n\times A$,*

$$V^a(t,x)\le\min\{M^a[V](t,x),N[V^a](t,x)\};\tag{4.13}$$

(ii) *If at $(t,x,a)\in[0,T)\times\mathbb{R}^n\times A$, a strict inequality in (4.13) holds, then there exists an $s_0\in(t,T]$, such that*

$$V^a(t,x)=\inf_{\mathcal{U}[t,T]}E\{\int_t^s f^0(r,y_{t,x}(r),u(r),a)dr+V^a(s,y_{t,x}(s))\},\quad s\in[t,s_0].\tag{4.14}$$

The proof of above result is similar to [5]. However, we note that the $s_0\in(t,T]$ is deterministic. The existence of such an $s_0>t$ is by no means obvious. We would like to sketch the proof of (ii) below (The proof of (i) is almost obvious). To this end, let us prove the following lemma first.

Lemma 4.3. *Let $\{a_i;1\le i\le i_0\}\subset A$ and $\{\xi_j;1\le j\le j_0\}\subset K$ with $i_0,j_0\ge 2$. Then, the following hold:*

1) *If for some $i_1=1,\cdots,i_0$, $a_{i_1}\ne a_{i_1+1}$, then*

$$V^{a_{i_0}}(t,x)+\sum_{i=1}^{i_0-1}k(t,a_i,a_{i+1})\ge M^{a_1}[V](t,x),\qquad \forall(t,x)\in[0,T]\times\mathbb{R}^n.\tag{4.15}$$

2) *If for some $j_1=1,\cdots,j_0$, $\xi_{j_1}\ne 0$, then*

$$V^a(t,x+\sum_{j=1}^{j_0}\xi_j)+\sum_{j=1}^{j_0}\ell(t,\xi_j)\ge N[V^a](t,x),\qquad \forall(t,x,a)\in[0,T]\times\mathbb{R}^n\times A.\tag{4.16}$$

3) *If either of the condition of (i) or (ii) holds, then*

$$\begin{aligned}&V^{a_{i_0}}(t,x+\sum_{j=1}^{j_0}\xi_j)+\sum_{i=1}^{i_0-1}k(t,a_i,a_{i+1})+\sum_{j=1}^{j_0}\ell(t,\xi_j)\\&\ge\min\{M^{a_1}[V](t,x),N[V^{a_1}](t,x)\},\qquad\forall(t,x)\in[0,T]\times\mathbb{R}^n.\end{aligned}\tag{4.17}$$

Proof. Let us prove 1). For any $(t,x) \in [0,T] \times \mathbb{R}^n$, by the definition of operator M and (4.13) (the proof of which is easy), we have

$$
\begin{aligned}
& V^{a_{i_0}}(t,x) + \sum_{i=1}^{i_0-1} k(t,a_i,a_{i+1}) \\
& \geq M^{a_{i_0-1}}[V](t,x) + \sum_{i=1}^{i_0-2} k(t,a_i,a_{i+1}) \\
& \geq V^{a_{i_0-1}}(t,x) + \sum_{i=1}^{i_0-2} k(t,a_i,a_{i+1}) \geq \cdots \\
& \geq V^{a_2}(t,x) + k(t,a_1,a_2) \geq M^{a_1}[V](t,x).
\end{aligned}
\tag{4.18}
$$

This gives 1). The proof of 2) is similar and then 3) follows from 1) and 2) easily. $\square$

Sketch of the proof of Theorem 4.2 (ii). From Theorem 4.1, we see that the inequality "$\leq$" holds. Thus, it suffices to prove the other direction. Suppose the contrary. That means there exist sequences $s \to t$ and $\varepsilon \to 0$, such that

$$
V^a(t,x) + \varepsilon < \inf_{u(\cdot)\in\mathcal{U}[t,T]} E\Big[\int_t^s f^0(r, y_{t,x}(r;u(\cdot))dr + V^a(s, y_{t,x}(s;u(\cdot)))\Big]. \tag{4.19}
$$

On the other hand, by definition, one concludes that there exists a triplet $(u^\varepsilon(\cdot), a^\varepsilon(\cdot), \xi^\varepsilon(\cdot)) \in \mathcal{U}[t,T] \times \mathcal{A}^a[t,T] \times \mathcal{K}[t,T]$, such that

$$
\begin{aligned}
V^a(t,x) + \varepsilon \geq E\Big[& V^{a^\varepsilon(s)}(s, y_{t,x}^{u^\varepsilon(\cdot),a^\varepsilon(\cdot),\xi^\varepsilon(\cdot)}(s)) \\
& + \int_t^s f^0(r, y_{t,x}^{u^\varepsilon(\cdot),a^\varepsilon(\cdot),\xi^\varepsilon(\cdot)}(r), u^\varepsilon(r), a^\varepsilon(r))dr \\
& + \sum_{t\leq\theta_i^\varepsilon\leq s} k(\theta_i^\varepsilon, a_{i-1}, a_i) + \sum_{t\leq\tau_j^\varepsilon\leq s} \ell(\tau_i^\varepsilon, \xi_j)\Big],
\end{aligned}
\tag{4.20}
$$

where $y_{t,x}^{u^\varepsilon(\cdot),a^\varepsilon(\cdot),\xi^\varepsilon(\cdot)}(\cdot)$ stands for the trajectory corresponding to $(t,x,u^\varepsilon(\cdot),a^\varepsilon(\cdot),\xi^\varepsilon(\cdot))$. Next, we set

$$
B_1 = \{t \leq \theta_1^\varepsilon \leq s\}, \qquad B_2 = \{t \leq \tau_1^\varepsilon \leq s\}.
$$

Then, (4.19) and (4.20) yield

$$
E\big(\chi_{B_1\cup B_2}\big) > 0, \tag{4.21}
$$

and

$$
0 > (I) + (II) + (III), \tag{4.22}
$$

where (with $\xi_0(\cdot)$ being the trivial impulse control)

$$
\begin{aligned}
(I) =& E\Big\{\int_t^s \big[f^0(r, y_{t,x}^{u^\varepsilon(\cdot),a^\varepsilon(\cdot),\xi^\varepsilon(\cdot)}(r), u^\varepsilon(r), a^\varepsilon(r)) \\
&- f^0(r, y_{t,x}^{u^\varepsilon(\cdot),a,\xi_0(\cdot)}(r), u^\varepsilon(r), a)\big]dr\Big\} = o(1)E\chi_{B_1\cup B_2},
\end{aligned}
\tag{4.23}
$$

$$
\begin{aligned}
(II) &= E\big[\sum_{t\le\theta_i^\varepsilon\le s} k(\theta_i^\varepsilon, a_{i-1}, a_i) + \sum_{t\le\tau_j^\varepsilon\le s} \ell(\tau_j^\varepsilon, \xi_j)\big] \\
&\ge E\big[\sum_{t\le\theta_i^\varepsilon\le s} k(s, a_{i-1}, a_i) + \sum_{t\le\tau_j^\varepsilon\le s} \ell(s, \xi_j)\big],
\end{aligned}
\tag{4.24}
$$

$$
\begin{aligned}
(III) &= E\big[V^{a^\varepsilon(s)}(s, y_{t,x}^{u^\varepsilon(\cdot),a^\varepsilon(\cdot),\xi^\varepsilon(\cdot)}(s)) - V^a(s, y_{t,x}^{u^\varepsilon(\cdot),a,\xi_0(\cdot)}(s))\big] \\
&= E\big[V^{a^\varepsilon(s)}(s, x + \sum_{t\le\tau_j^\varepsilon\le s}\xi_j)\chi_{B_2}\big] + E\big[V^{a^\varepsilon(s)}(s,x)\chi_{B_1\cap B_2^c}\big] \\
&\quad - V^a(s,x)E\chi_{B_1\cup B_2} + o(1)E(\chi_{B_1\cup B_2}).
\end{aligned}
\tag{4.25}
$$

In (4.23)–(4.25), we have used (2.6)–(2.8) and Corollary 3.8, with $(s-t)\to 0$. Hence, from Lemma 4.3, we obtain

$$
\begin{aligned}
0 &> (I) + (II) + (III) \\
&\ge \{o(1) + \big[\min\{M^a[V](s,x), N[V^a](s,x)\} - V^a(s,x)\big]\}E\chi_{B_1\cup B_2} \\
&\ge \delta_0 E\chi_{B_1\cup B_2} + o(1)E(\chi_{B_1\cup B_2}),
\end{aligned}
\tag{4.26}
$$

with

$$
\delta_0 \equiv \min\{M^a[V](t,x), N[V^a](t,x)\} - V^a(t,x) > 0. \tag{4.27}
$$

Thus, (4.21), (4.26) and (4.27) lead a contradiction. $\square$

Now, let us introduce the following Hamiltonian:

(4.28)

$$
\begin{aligned}
H^a(t,x,p,X) = \inf_{u\in U}\{&\frac{1}{2}\mathrm{tr}[\sigma(t,x,u,a)^* X\sigma(t,x,u,a) + \langle p, f(t,x,u,a)\rangle \\
&+ f^0(t,x,u,a)\}, \qquad (t,x,a,p,X)\in[0,T]\times\mathbb{R}^n\times A\times\mathbb{R}^n\times\mathbb{R}^{n\times n}.
\end{aligned}
$$

From the above theorem, using some standard arguments, we can easily obtain the following result.

Proposition 4.4. *Suppose the value function $V(\cdot,\cdot)$ is smooth. Then, it satisfies the following Hamilton-Jacobi-Bellman equation:*

$$\begin{aligned}\max\{-\frac{\partial V^a(t,x)}{\partial t} - H^a(t,x,DV^a(t,x),D^2V^a(t,x)),V^a(t,x)-M^a[V](t,x),\\ V^a(t,x)-N[V^a](t,x)\}=0, \qquad \forall (t,x,a)\in[0,T)\times\mathbb{R}^n\times A,\end{aligned} \tag{4.29}$$

with the terminal condition

$$V^a(T,x)=h(x), \qquad \forall (x,a)\in\mathbb{R}^n\times A, \tag{4.30}$$

where DV^a and D^2V^a stand for the gradient and the Hessian of V^a.

We see that (4.29) is a system of fully nonlinear second order (possibly degenerate) parabolic quasi-variational inequalities. Thus, in general, (4.29)–(4.30) may have no classical solutions. On the other hand, it is well-known that, in general, the value function $V(\cdot,\cdot)$ is *not* smooth enough to satisfy equation (4.29) in a classical sense. Hence, we need the notion of viscosity solution ([6]) to make the above rigorous.

§5. Characterization of the Value Function.

In this section, we shall characterize the value function $V(\cdot,\cdot)$ as the unique viscosity solution of the Hamilton-Jacobi-Bellman equation (4.29)–(4.30). First of all, let us recall the definition of viscosity solutions. To this end, we let $I\subset\mathbb{R}$ be any interval and $C^{1,2}(I\times\mathbb{R}^n)$ be the set of all continuous functions $\varphi(\cdot,\cdot)$ which are C^1 in t and C^2 in x.

Definition 5.1. Function $V(\cdot)\in C(I\times\mathbb{R}^n;\mathbb{R}^m)$ is called a viscosity sub-(super-) solution of (4.29) on $I\times\mathbb{R}^n$ if for any $\varphi(\cdot,\cdot)\in C^{1,2}(I\times\mathbb{R}^m)$ and $a=1,\cdots,m$, whenever $V^a(\cdot,\cdot)-\varphi(\cdot,\cdot)$ attains a local maximum (minimum) at $(t_0,x_0)\in I\times\mathbb{R}^n$,

$$\begin{aligned}\max\{-\frac{\partial\varphi}{\partial t}(t_0,x_0)-H^a(t_0,x_0,D\varphi(t_0,x_0),D^2\varphi(t_0,x_0)),\\ V^a(t_0,x_0)-M^a[V](t_0,x_0),V^a(t_0,x_0)-N[V^a](t_0,x_0)\}\le 0(\ge 0).\end{aligned} \tag{5.1}$$

If $V(\cdot,\cdot)$ is both a viscosity sub- and super-solution of (4.29), $V(\cdot,\cdot)$ is called a viscosity solution of (4.29).

Proposition 5.2. *Let $V(\cdot,\cdot)\in C([0,T]\times\mathbb{R}^n;\mathbb{R}^m)$ be a viscosity sub-(super-) solution of (4.29) on $(0,T)\times\mathbb{R}^n$. Then, $V(\cdot,\cdot)$ is also a viscosity sub-(super-) solution of (4.29) on $[0,T)\times\mathbb{R}^n$.*

Proof. Let $V(\cdot,\cdot)$ be a viscosity subsolution of (4.29). Let $\varphi(\cdot,\cdot)\in C^{1,2}([0,T)\times\mathbb{R}^n)$ such that $V^a(\cdot,\cdot)-\varphi(\cdot,\cdot)$ attains a local maximum at $(0,x_0)\in[0,T)\times\mathbb{R}^n$. Define

$$\psi_\varepsilon(t,x)=\varphi(t,x)+\frac{\varepsilon}{t},\qquad (t,x)\in(0,T)\times\mathbb{R}^n. \tag{5.2}$$

Then, there exists a sequence $\varepsilon_i\to 0$, such that $V^a(\cdot,\cdot)-\psi_\varepsilon(\cdot,\cdot)$ attains a local maximum at (t_i,x_i) which converges to $(0,x_0)$ as $i\to\infty$. Then, by Definition 5.1, we see that

$$\begin{aligned}\max\{&-\frac{\partial\varphi}{\partial t}(t_i,x_i)+\frac{\varepsilon_i}{t_i^2}-H^a(t_i,x_i,D\varphi(t_i,x_i),D^2\varphi(t_i,x_i)),\\ &V^a(t_i,x_i)-M^a[V](t_i,x_i),V^a(t_i,x_i)-N[V^a](t_i,x_i)\}\le 0.\end{aligned} \tag{5.3}$$

Then, dropping the term $\varepsilon_i/(t^2)$ and sending $i\to\infty$, we obtain the conclusion for the sub-solution. Similarly, we can obtain the conclusion for the super-solution. To do this, we replace ε/t by $-\varepsilon/t$ in (5.2). □

Proposition 5.3. *The value function $V(\cdot,\cdot)$ is a viscosity solution of (4.29) on $[0,T)\times\mathbb{R}^n$ satisfying (4.30).*

The proof follows from a standard argument and Theorem 4.2. Now, we come to the uniqueness of the viscosity solution. To this end, we need a further assumption:

$$\gamma=0. \tag{5.4}$$

In the rest of this section, we keep this assumption. Under (5.4), we see that the value function satisfies the following: (see (3.2))

$$|V^a(t,x)-V^a(\widehat t,\widehat x)|\le C\big[|x-\widehat x|^\delta+(1+|x|^\mu+|\widehat x|^\mu)|t-\widehat t|^{\delta/2}\big]. \tag{5.5}$$

Now, let us introduce the following set of functions:

$$\mathcal{V}_\delta=\{v(\cdot)\in C([0,T]\times\mathbb{R}^n;\mathbb{R}^m)\bigm|\sup_{t,x,\widehat x}\frac{|v(t,x)-v(t,\widehat x)|}{1+|x-\widehat x|^\delta}<\infty\}. \tag{5.6}$$

Clearly, under (5.4), the value function $V(\cdot,\cdot)$ of our Problem CSI is in the set $\mathcal{V}_\delta$. On the other hand, we see that for any $v(\cdot,\cdot)\in\mathcal{V}_\delta$, there exists a constant $C>0$, such that

$$|v(t,x)|\le C(1+|x|^\delta). \tag{5.7}$$

Next lemma gives an important property of viscosity solutions of (4.29)–(4.30).

Lemma 5.4. *Suppose $V(\cdot,\cdot)\in\mathcal{V}_\delta$ is a viscosity subsolution of (4.29)–(4.30). Then,*

$$V^a(t,x)\le\min\{M^a[V](t,x),N[V^a](t,x)\},\qquad\forall(t,x,a)\in[0,T]\times\mathbb{R}^n\times A. \tag{5.8}$$

Proof. It is enough to prove (5.8) for all $(t,x,a)\in(0,T)\times\mathbb{R}^n\times A$. We define

$$\Phi^a(s,y)=V^a(s,y)-\frac{1}{\varepsilon}(|t-s|^2+|x-y|^2),\qquad(s,y)\in[0,T]\times\mathbb{R}^n, \tag{5.9}$$

with some $\varepsilon\in(0,1]$. Since $\delta\in(0,1)$, we see that there exists a point $(s_\varepsilon,y_\varepsilon)\in[0,T]\times\mathbb{R}^n$, such that

$$\Phi^a(s_\varepsilon,y_\varepsilon)=\max_{(s,y)\in[0,T]\times\mathbb{R}^n}\Phi^a(s,y)\ge\Phi^a(t,x)=V^a(t,x). \tag{5.10}$$

We see easily that

$$\lim_{\varepsilon\to0}|s_\varepsilon-t|=0,\qquad\lim_{\varepsilon\to0}|y_\varepsilon-x|=0. \tag{5.11}$$

Thus, for $\varepsilon>0$ small enough, we have $(s_\varepsilon,y_\varepsilon)\in(0,T)\times\mathbb{R}^n$. Then, by the definition of the viscosity subsolution, we obtain

$$V^a(s_\varepsilon,y_\varepsilon)\le\min\{M^a[V](s_\varepsilon,y_\varepsilon),N[V^a](s_\varepsilon,y_\varepsilon)\}. \tag{5.12}$$

Sending $\varepsilon\to0$, and using the continuity of $V(\cdot,\cdot)$, we obtain (5.8). $\square$

Next, let us introduce the following sets, which are adopted from [7]. For function $v:[0,T]\times\mathbb{R}^n\to[-\infty,+\infty]$, $(s,z)\in[0,T)\times\mathbb{R}^n$,

$$\begin{aligned}\mathcal{P}^{2,+}v(s,z)=&\{(b,p,X)\in\mathbb{R}\times\mathbb{R}^n\times\mathcal{S}^n\;\big|\;v(t,x)\le v(s,z)+b(t-s)+\langle\,p,x-z\,\rangle\\&\frac{1}{2}\langle\,X(x-z),x-z\,\rangle+o(|t-s|+|x-z|^2),\\&\text{as }[0,T]\times\mathbb{R}^n\ni(t,x)\to(s,z)\},\end{aligned} \tag{5.13}$$

$$\begin{aligned}\overline{\mathcal{P}}^{2,+}v(s,z)=&\{(b,p,X)\in\mathbb{R}\times\mathbb{R}^n\times\mathcal{S}^n\;\big|\;\exists((t_i,x_i)\in[0,T]\times\mathbb{R}^n\\&(b_i,p_i,X_i)\in\mathbb{R}\times\mathbb{R}^n\times\mathcal{S}^n,(b_i,p_i,X_i)\in\mathcal{P}^{2,+}v(t_i,x_i),\\&(t_i,x_i,v(t_i,x_i),b_i,p_i,X_i)\to(s,z,v(s,z),b,p,X)\},\end{aligned} \tag{5.14}$$

where $\mathcal{S}^n$ is the set of all $(n \times n)$ real symmetric matrices. We define

$$\mathcal{P}^{2,-}v(s,z) = -\mathcal{P}^{2,+}(-v)(s,z), \qquad \overline{\mathcal{P}}^{2,-}v(s,z) = -\overline{\mathcal{P}}^{2,+}(-v)(s,z). \tag{5.15}$$

The following result is almost standard.

Proposition 5.5. *Function $V(\cdot,\cdot) \in C([0,t] \times \mathbb{R}^n)$ is a viscosity sub-solution (resp. super-solution) of (4.29) if and only if the following holds: $\forall (t,x) \in [0,T) \times \mathbb{R}^n$,*

$$\begin{aligned} &\max\{-b - H^a(t,x,p,X), V^a(t,x) - M^a[V](t,x),\\ &\qquad V^a(t,x) - N[V^a](t,x)\} \le 0 (\ge 0),\\ &\qquad\qquad \forall (b,p,X) \in \overline{\mathcal{P}}^{2,+}V^a(t,x) \quad (\textit{resp. } \overline{\mathcal{P}}^{2,-}V^a(t,x)). \end{aligned} \tag{5.16}$$

Now, we are ready to prove the main result of this section.

Theorem 5.6. *Let (C1)–(C4) hold and let (5.4) hold. Then, the value function $V(\cdot,\cdot)$ of Problem CSI is the unique viscosity solution of (4.29)–(4.30) in $\mathcal{V}_\delta$.*

Proof. We only need to prove the uniqueness of the viscosity solution of (4.29)–(4.30) in the set $\mathcal{V}_\delta$. We prove this by contradiction. Thus, let $V(\cdot,\cdot), \widehat{V}(\cdot,\cdot) \in \mathcal{V}_\delta$ be two different viscosity solutions of (4.29)–(4.30). Then, there exists a point $(\bar{t},\bar{x},\bar{a}) \in (0,T) \times \mathbb{R}^n \times A$, such that

$$V^{\bar{a}}(\bar{t},\bar{x}) - \widehat{V}^{\bar{a}}(\bar{t},\bar{x}) \equiv \eta > 0. \tag{5.17}$$

Since $V(\cdot,\cdot), \widehat{V}(\cdot,\cdot) \in \mathcal{V}_\delta$, we may let $\widehat{C} \ge b_0$, such that for all $t \in [0,T]$, $x, \widehat{x} \in \mathbb{R}^n$ and $a \in A$,

$$\begin{aligned} &|V^a(t,x)|, |\widehat{V}^a(t,x)| \le \widehat{C}(1 + |x|^\delta),\\ &|V^a(t,x) - V^a(t,\widehat{x})|, |\widehat{V}^a(t,x) - \widehat{V}^a(t,\widehat{x})| \le \widehat{C}(1 + |x - \widehat{x}|^\delta). \end{aligned} \tag{5.18}$$

It is clear that for any $\xi_0 \in K$ with

$$b_0|\xi_0|^\mu \le \widehat{C}(1 + |\xi_0|^\delta), \tag{5.19}$$

one has

$$|\xi_0| \le \Big(\frac{2\widehat{C}}{b_0}\Big)^{\frac{1}{\mu-\delta}} \equiv \bar{C}. \tag{5.20}$$

Next, we take constants $G > 0$ large enough and $\alpha, \beta, \lambda \in (0,1]$ small enough, so that the following hold:

$$
\begin{cases} G\ell_0 > 2\bar{C}, \quad \beta T < \eta/4, \\ \alpha G < 1, \quad 2\alpha \langle \bar{x} \rangle + \dfrac{\lambda}{\bar{t}} < \eta/4, \quad \alpha G(V^{\bar{a}}(\bar{t}, \bar{x}) + L) < \eta/4. \end{cases} \tag{5.21}
$$

Here, $\langle \bar{x} \rangle = \sqrt{1 + |\bar{x}|^2}$. Then, for small $\varepsilon > 0$, we define

$$
\begin{aligned} \Phi^a(t,x,y) =& (1-\alpha G)V^a(t,x) - \widehat{V}^a(t,y) - \alpha\Big(\frac{2T-t}{2T}\Big)(\langle x \rangle + \langle y \rangle) \\ & - \frac{1}{2\varepsilon}|x-y|^2 + \beta t - \frac{\lambda}{t}. \end{aligned} \tag{5.22}
$$

It is clear that there exists a $(t_0, x_0, y_0, a_0) \in (0,T] \times \mathbb{R}^n \times \mathbb{R}^n \times A$, such that

$$
\begin{aligned} \Phi^{a_0}(t_0, x_0, y_0) &= \max_{t,x,y,a} \Phi^a(t,x,y) \geq \Phi^{a_0}(T,0,0) \\ &= -\alpha G h(0) - \alpha T + \beta T - \frac{\lambda}{T} \geq -|h(0)| - T - \frac{1}{T}. \end{aligned} \tag{5.23}
$$

This implies that for some (absolute) constant $C > 0$,

$$
\alpha(\langle x_0 \rangle + \langle y_0 \rangle) + \frac{1}{2\varepsilon}|x_0 - y_0|^2 \leq C(1 + |x_0|^\delta + |y_0|^\delta). \tag{5.24}
$$

Thus, there exists a constant C_α, depending on α, such that

$$
|x_0|, |y_0|, \ \frac{1}{2\varepsilon}|x_0 - y_0|^2 \leq C_\alpha. \tag{5.25}
$$

On the other hand, from $2\Phi^{a_0}(t_0, x_0, y_0) \geq \Phi^{a_0}(t_0, x_0, x_0) + \Phi^{a_0}(t_0, y_0, y_0)$, we end up with

$$
\frac{1}{2\varepsilon}|x_0 - y_0|^2 \leq C|x_0 - y_0|^\delta \to 0, \qquad \text{as } \varepsilon \to 0. \tag{5.26}
$$

We claim that for $\varepsilon > 0$ small, we have $t_0 < T$. In fact, if $t_0 = T$, then, (note (2.6) and (5.4))

$$
\begin{aligned} & (1-\alpha G)V^{\bar{a}}(\bar{t}, \bar{x}) - \widehat{V}^{\bar{a}}(\bar{t}, \bar{x}) - \alpha\Big(\frac{2T-\bar{t}}{T}\Big)\langle \bar{x} \rangle + \beta\bar{t} - \frac{\lambda}{\bar{t}} \equiv \Phi^{\bar{a}}(\bar{t}, \bar{x}, \bar{x}) \\ & \leq \Phi^{a_0}(t_0, x_0, y_0) \\ & = (1-\alpha G)h(x_0) - h(y_0) - \frac{\alpha}{2}(\langle x_0 \rangle + \langle y_0 \rangle) - \frac{1}{2\varepsilon}|x_0 - y_0|^2 + \beta T - \frac{\lambda}{T} \\ & \leq L|x_0 - y_0|^\delta + \alpha G L + \beta T. \end{aligned} \tag{5.27}
$$

Then, we have (see (5.21))

$$
\begin{aligned}
\eta &= V^{\bar a}(\bar t,\bar x) - \widehat V^{\bar a}(\bar t,\bar x)\\
&\le \alpha G(V^{\bar a}(\bar t,\bar x)+L) + 2\alpha\langle\,\bar x\,\rangle + L|x_0-y_0|^\delta + \beta T + \frac{\lambda}{\bar t}\\
&< \frac{3}{4}\eta + L|x_0-y_0|^\delta.
\end{aligned}
\tag{5.28}
$$

Thus, for the given $\alpha>0$, if $\varepsilon>0$ is small enough, the above gives a contradiction. That means for the given constants α,β,λ and G satisfying (5.21), when $\varepsilon>0$ small, we have $t_0\in(0,T)$. Next, we claim that

$$
\widehat V^{a_0}(t_0,y_0) < \min\{M^{a_0}[\widehat V](t_0,y_0), N[\widehat V^{a_0}](t_0,y_0)\}. \tag{5.29}
$$

In fact, if

$$
\widehat V^{a_0}(t_0,y_0) = M^{a_0}[\widehat V](t_0,y_0) = \widehat V^{a_1}(t_0,y_0) + k(t_0,a_0,a_1), \tag{5.30}
$$

then, noticing Lemma 5.4, we have

$$
\begin{aligned}
&\Phi^{a_1}(t_0,x_0,y_0) - \Phi^{a_0}(t_0,x_0,y_0)\\
&= (1-\alpha G)[V^{a_1}(t_0,x_0) - V^{a_0}(t_0,x_0)] - [\widehat V^{a_1}(t_0,y_0) - \widehat V^{a_0}(t_0,y_0)]\\
&\ge \alpha G k(t_0,a_0,a_1) > 0,
\end{aligned}
\tag{5.31}
$$

which contradicts the definition of a_0. Now, if we have

$$
\widehat V^{a_0}(t_0,y_0) = N[\widehat V^{a_0}](t_0,y_0) = \widehat V^{a_0}(t_0,y_0+\xi_0) + \ell(t_0,\xi_0), \tag{5.32}
$$

for some $\xi_0\in K$, then, we have (see (5.18))

$$
b_0|\xi_0|^\mu \le \ell(\tau_0,\xi_0) \le \widehat C(1+|\xi_0|^\delta). \tag{5.33}
$$

Thus, by (5.20) and (5.21), we obtain

$$
\begin{aligned}
&\Phi^{a_0}(t_0,x_0+\xi_0,y_0+\xi_0) - \Phi^{a_0}(t_0,x_0,y_0)\\
&= (1-\alpha G)[V^{a_0}(t_0,x_0+\xi_0) - V^{a_0}(t_0,x_0)] - [\widehat V^{a_0}(t_0,y_0+\xi_0) - \widehat V^{a_0}(t_0,y_0)]\\
&\qquad - \alpha(\langle\,x_0+\xi_0\,\rangle - \langle\,x_0\,\rangle + \langle\,y_0+\xi_0\,\rangle - \langle\,y_0\,\rangle)\\
&\ge \alpha G\ell(t_0,\xi_0) - 2\alpha|\xi_0| \ge \alpha(G\ell_0 - 2\bar C) > 0.
\end{aligned}
\tag{5.34}
$$

This contradicts the definition of (x_0, y_0). Hence, (5.29) holds. Now, let us denote

$$\varphi(t,x,y) = -\beta t + \alpha(\frac{2T-t}{2T})(\langle\, x\,\rangle + \langle\, y\,\rangle) + \frac{1}{2\varepsilon}|x-y|^2 + \frac{\lambda}{t}. \tag{5.35}$$

Then, we have

(5.36)

$$\begin{cases} D_t\varphi(t,x,y) = -\beta - \dfrac{\alpha}{2T}(\langle\, x\,\rangle + \langle\, y\,\rangle) - \dfrac{\lambda}{t^2}, \\ D_x\varphi(t,x,y) = \alpha(\dfrac{2T-t}{2T})\dfrac{x}{\langle\, x\,\rangle} + \dfrac{1}{\varepsilon}(x-y), \\ D_y\varphi(t,x,y) = \alpha(\dfrac{2T-t}{2T})\dfrac{y}{\langle\, y\,\rangle} + \dfrac{1}{\varepsilon}(y-x), \\ B(t,x,y) \equiv D^2_{(x,y)}\varphi(t,x,y) = \dfrac{1}{\varepsilon}\begin{pmatrix} 1 & -1 \\ -1 & 1 \end{pmatrix} + \alpha \begin{pmatrix} \frac{I}{\langle\, x\,\rangle} - \frac{xx^T}{\langle\, x\,\rangle^3} & 0 \\ 0 & \frac{I}{\langle\, y\,\rangle} - \frac{yy^T}{\langle\, y\,\rangle^3} \end{pmatrix}. \end{cases}$$

Then, applying Theorem 9 of [7] to the function

$$(1-\alpha G)V^{a_0}(t,x) + (-\widehat{V}^{a_0})(t,y) - \varphi(t,x,y)$$

at point (t_0, x_0, y_0), we can find $b, c \in \mathbb{R}$ and $X, Y \in \mathcal{S}^n$, such that

$$\begin{cases} (b, (D_x\varphi)(t_0,x_0,y_0), X) \in \overline{\mathcal{P}}^{2,+}\big((1-\alpha G)V^{a_0}\big)(t_0,x_0), \\ (c, (D_y\varphi)(t_0,x_0,y_0), Y) \in \overline{\mathcal{P}}^{2,+}\big(-\widehat{V}^{a_0}\big)(t_0,y_0), \\ b + c = (D_t\varphi)(t_0,x_0,y_0), \\ -(\dfrac{1}{\varepsilon} + \|B(t_0,x_0,y_0)\|)I \le \begin{pmatrix} X & 0 \\ 0 & Y \end{pmatrix} \le B(t_0,x_0,y_0) + \varepsilon B(t_0,x_0,y_0)^2. \end{cases} \tag{5.37}$$

For notational simplicity, in what follows, we suppress the subscript 0 in (t_0, x_0, y_0, a_0). By the definition of viscosity solution, we have

$$-b - (1-\alpha G)H^a(t,x,\frac{1}{1-\alpha G}(\alpha\frac{2T-t}{2T}\frac{x}{\langle\, x\,\rangle} + \frac{x-y}{\varepsilon}), \frac{1}{1-\alpha G}X) \le 0, \tag{5.38}$$

$$c - H^a(t,y,-\frac{\alpha y}{\langle\, y\,\rangle} - \frac{y-x}{\varepsilon}, -Y) \ge 0. \tag{5.39}$$

Thus, we have (see (5.37) and note $\lambda > 0$)

(5.40)

$$\begin{aligned}
&\beta + \frac{\alpha}{2T}(\langle x \rangle + \langle y \rangle) \\
&\le (1-\alpha G)H^a(t,x,\frac{1}{1-\alpha G}(\alpha\frac{2T-t}{2T}\frac{x}{\langle x \rangle} + \frac{x-y}{\varepsilon}),\frac{1}{1-\alpha G}X) \\
&\qquad - H^a(t,y,-\frac{\alpha y}{\langle y \rangle} - \frac{y-x}{\varepsilon},-Y) \\
&= \inf_{u\in U}\{\frac{1}{2}\mathrm{tr}[\sigma(t,x,u,a)^* X\sigma(t,x,u,a)] + \langle\, \alpha\frac{2T-t}{2T}\frac{x}{\langle x \rangle} + \frac{x-y}{\varepsilon}, f(t,x,u,a)\,\rangle \\
&\qquad\quad + (1-\alpha G)f^0(t,x,u,a)\} \\
&- \inf_{u\in U}\{\frac{1}{2}\mathrm{tr}[\sigma(t,y,u,a)^*(-Y)\sigma(t,y,u,a)] + \langle -\alpha\frac{2T-t}{2T}\frac{y}{\langle y \rangle} + \frac{x-y}{\varepsilon}, f(t,y,u,a)\,\rangle \\
&\qquad\quad + f^0(t,x,u,a)\} \\
&\le \sup_{u\in U}\{\frac{1}{2}\mathrm{tr}[\sigma(t,x,u,a)^* X\sigma(t,x,u,a) + \sigma(t,y,u,a)^* Y\sigma(t,y,u,a)] \\
&\quad + [\langle\, \alpha\frac{2T-t}{2T}\frac{x}{\langle x \rangle} + \frac{x-y}{\varepsilon}, f(t,x,u,a)\,\rangle - \langle -\alpha\frac{2T-t}{2T}\frac{y}{\langle y \rangle} + \frac{x-y}{\varepsilon}, f(t,y,u,a)\,\rangle] \\
&\quad + (1-\alpha G)f^0(t,x,u,a) - f^0(t,y,u,a)\} \equiv \sup\{(I)+(II)+(III)\}.
\end{aligned}$$

We now estimate $(I), (II)$ and (III) separately. It is immediate that

$$B \equiv B(t,x,y) \le \frac{1}{\varepsilon}\begin{pmatrix} I & -I \\ -I & I \end{pmatrix} + C\alpha I. \tag{5.41}$$

Thus,

$$B + \varepsilon B^2 \le \frac{3}{\varepsilon}\begin{pmatrix} I & -I \\ -I & I \end{pmatrix} + C\alpha = \frac{3}{\varepsilon}\begin{pmatrix} I \\ -I \end{pmatrix}(\, I \quad -I\,) + C\alpha. \tag{5.42}$$

Then, we have

$$(I) \le \frac{C}{\varepsilon}|x-y|^2 + C\alpha(1+|x|^\nu+|y|^\nu). \tag{5.43}$$

Also,

$$(II) \le C\alpha(1+|x|^\nu+|y|^\nu) + \frac{L}{\varepsilon}|x-y|^2. \tag{5.44}$$

$$(III) \le L|x-y|^\delta + \alpha GL(1+|x|^\delta). \tag{5.45}$$

Hence, we have

$$\begin{aligned}\beta \le C\frac{|x-y|^2}{\varepsilon} + C\alpha(1+|x|^\nu+|y|^\nu) + L|x-y|^\delta \\ + \alpha GL(1+|x|^\delta) - \frac{\alpha}{2T}(\langle\, x\,\rangle + \langle\, y\,\rangle).\end{aligned} \tag{5.46}$$

We let $\varepsilon \to 0$ to get

$$\beta \le C\alpha(1+|x|^\nu+|y|^\nu) + \alpha GL(1+|x|^\delta) - \frac{\alpha}{2T}(\langle\, x\,\rangle + \langle\, y\,\rangle). \tag{5.47}$$

Finally, since $\nu, \delta < 1$, by sending $\alpha \to 0$, we get $\beta \le 0$ which is a contradiction. □

Remark 5.7. If the function $\sigma(t,x,u,a)$ is identically zero, then our Problem SCI is reduced to the deterministic case. Clearly, our result covers such a situation.

Remark 5.8. It is not hard to discuss the time-invariant problem in infinite horizon. In that case, as usual, the discount factor comes in the cost functional and the value function is time independent satisfying a system of fully nonlinear elliptic type quasi-variational inequalities (in viscosity sense). We prefer not get into these details. However, we would like to point out that in [21], to obtain the uniqueness of the viscosity solution for the time-invariant deterministic problem, a technical condition was imposed. Here, we adopt the method used in [16], and that technical condition is removed.

References

[1] G. Barles, *Deterministic impulse control problems, SIAM J. Control & Optim.*, 23 (1985), 419–432.

[2] S. A. Belbas and S. M. Lenhart, *Nonlinear PDE's for stochastic optimal control with switchings and impulses, Appl. Math. Optim.*, 14 (1986), 215–227.

[3] A. Bensoussan and J. L. Lions, *Impulse Control and Quasi-Variational Inequalities,* Bordes, Paris, 1984.

[4] A. Bensoussan, *Stochastic Control by Functional Analysis Methods,* North-Holland, Amsterdam, 1982.

[5] I. Capuzzo-Dolcetta and L. C. Evans, *Optimal switching for ordinary differential equations, SIAM J. Control & Optim.*, 22 (1984), 143–161.

[6] M. G. Crandall, L. C. Evans and P. L. Lions, *Some properties of viscosity solutions of Hamilton-Jacobi equations,* Trans. Amer. Math. Soc., 282 (1984), 487–502.

[7] M. G. Crandall and H. Ishii, *The maximum principle for semicontinuous functions,* Diff. Int. Eqn., 3 (1990), 1001–1014.

[8] L. C. Evans and A. Friedman, *Optimal stochastic switching and the Dirichlet problem for the Bellman equation,* Trans. Amer. Math. Soc., 253 (1979), 365–389.

[9] W. H. Fleming and P. E. Souganidis, *On the existence of value functions of two-player, zero-sum stochastic differential games,* Indiana Univ. Math. J., 38 (1989), 293–314.

[10] H. Ishii, *Uniqueness of unbounded viscosity solution of Hamilton-Jacobi equations,* Indiana Univ. Math. J. 33 (1984), 721–748.

[11] H. Ishii, *On uniqueness and existence of viscosity solutions of fully nonlinear second-order elliptic PDE's,* Comm. Pure & Appl. Math., 42 (1989), 15–45.

[12] H. Ishii and P. L. Lions, *Viscosity solutions of fully nonlinear second-order elliptic partial differential equations,* J. Diff. Eqn., 83 (1990), 26–78.

[13] R. Jensen, *Uniqueness criteria for viscosity solutions of fully nonlinear elliptic partial differential equations,* Indiana Math. J., 38 (1989), 629–667.

[14] S. M. Lenhart and S. A. Belbas, *A system of nonlinear partial differential equations arising in the optimal control of stochastic systems with switching costs,* SIAM J. Appl. Math. 43 (1983), 465–475.

[15] S. M. Lenhart, *Bellman equation for optimal stopping time problems,* Indiana Univ. Math. J., 32 (1983), 363–375.

[16] S. M. Lenhart, *Viscosity solutions associated with impulse control problems for piece-wise-deterministic processes,* Int. J. Math. & Math. Sci., 12 (1989), 145–158.

[17] J. L. Menaldi, *Optimal impulse control problems for degenerate diffusions with jumps,* Acta Appl. Math., 8 (1987), 165–198.

[18] S. Stojanovic and J. Yong, *Optimal switching for systems governed by nonlinear evolution equations,* Numer. Funct. Anal. Optim., 9(9&10) (1987), 995–1030.

[19] S. Stojanovic and J. Yong, *Optimal switching for partial differential equations I,II, J. Math. Anal. Appl.,* 138 (1989), 418–460.

[20] J. Yong, *Optimal switching and impulse controls for distributed parameter systems, System Sci. & Math. Sci.,* 2 (1989), 137–160.

[21] J. Yong, *Systems governed by ordinary differential equations with continuous, switching and impulse controls, Appl. Math. Optim.,* 20 (1989), 223–236.

[22] J. Yong, *Zero-sum differential games involving impulse controls,* submitted.