CONTINUITY OF THE VALUE FUNCTION FOR DETERMINISTIC OPTIMAL IMPULSE CONTROL WITH TERMINAL STATE CONSTRAINT

Yue Zhou[1], Xinwei Feng[2,*,**] and Jiongmin Yong[3,***]

Abstract. Deterministic optimal impulse control problem with terminal state constraint is considered. Due to the appearance of the terminal state constraint, the value function might be discontinuous in general. The main contribution of this paper is the introduction of an intrinsic condition under which the value function is proved to be continuous. Then by a Bellman dynamic programming principle, the corresponding Hamilton-Jacobi-Bellman type quasi-variational inequality (QVI, for short) is derived. The value function is proved to be a viscosity solution to such a QVI. The issue of whether the value function is characterized as the unique viscosity solution to this QVI is carefully addressed and the answer is left open challengingly.

1. Introduction

It is well-known that in general classical continuous-time optimal control theory, there are two major approaches: variational method leading to Pontryagin's maximum principle (MP, for short), and dynamic programming method leading to Hamilton-Jacobi-Bellman (HJB, for short) equation. The former could work for the problems with possible terminal state constraint and it gives necessary conditions for (possibly existed) open-loop optimal controls [27, 39]. The latter usually works for the problems *without* terminal state constraint and it leads to a characterization of the value function as the unique viscosity solution to the HJB equation, then formally, optimal control of state feedback form can be obtained [1, 5, 17, 23, 39]. In the case that there exists a terminal state constraint, the value function might not be everywhere defined (which involves the controllability issue), and even in the set on which the value function is defined, due to the set of admissible controls is depending on the initial pair, the continuity of the value function is not guaranteed. As a result, the corresponding satisfactory general theory of viscosity solution to the HJB equation for the terminal state

[1] School of Mathematics and Statistics, Central South University, Changsha, Hunan, China.
[2] Zhongtai Securities Institute for Financial Studies, Shandong University, Jinan, Shandong 250100, China.
[3] Department of Mathematics, University of Central Florida, Orlando, FL 32816, USA.

* Corresponding author: xwfeng@sdu.edu.cn

constrained is not available as of today. Therefore, people had made some efforts to introduce proper conditions so that the continuity of the value function can still be obtained. One such an effort is for time optimal control problem (with a target set which is a terminal state constraint). To ensure the continuity of the value function, the so-called *small time local controllability* (STLC, for short) was introduced by Sussmann in 1987 [1, 30, 37]. This condition means that when the state gets close to the boundary of the target set (from outside), only a small amount of time is needed to drive the state to the target by a control action. This then will lead to the continuity of the value function.

For optimal impulse control problems, similar to the continuous control case, one also has two major approaches. For variational method, there are works on MP; we mention [9, 12, 16, 18, 28, 33, 38], for a partial list of both stochastic and deterministic cases. On the other hand, since the initiation of optimal impulse control problems by Bensoussan–Lions in the early 1970s [6, 7], the dynamic programming method has been a very popular approach to the problem. It is interesting that the corresponding HJB equation is a *quasi-variational inequality* (QVI, for short) to which the value function is the unique viscosity solution [2, 3] under proper conditions. There are quite a few follow-up works, see [13–15, 23, 35, 36] for deterministic cases, and [4, 10, 20, 22, 24, 25, 31] for stochastic cases. The same as the continuous control case, to our best knowledge, in all the existing literature treating optimal impulse control problems by dynamic programming principle, the terminal state is constraint-free. A natural question arises: What if the terminal state is constrained? Then one expects that, in general, the value function is not continuous, and might even not be defined somewhere. Recall that for continuous control case, there is an STLC condition that ensures the continuity of the value function. The major contribution of this paper is the discovery of an intrinsic condition that can play a similar role as STLC in the optimal impulse control problems. Under such a condition, the continuity of the value function will be proved and, consequently, it will be proved by mean of dynamic programming method that the value function is a viscosity solution to the Hamiton-Jacobi-Bellman QVI.

As far as applications are concerned, it is known that optimal impulse controls can be used in may areas, for examples, [8] for management problems, [26] for SIR epidemic problems, [34] for HIV treatment, [19] for mathematical finance, [21] for some biology systems, to mention a few. For the optimal impulse control problem with a terminal state constraint, our motivation is as follow: Suppose a unit (could be a company, a bank, a state, or even a country) is running its business during a certain time period, say a month, a quarter-year, one year, etc. Besides it keeps normal running, at the end of the period, certain types of goods/assets (such as cash, food, gas, medicine, etc.) have to reach a certain reserve level. If the goal could not be reached by its own production, it is allowed and has to buy from outside with some costs. This means that the state (goods/assets) of the unit can be controlled by some impulses. The problem is to minimize the total cost with the terminal state constraint being satisfied. Clearly, such a framework is very general and could cover many real application problems. This also shows that the problem that we are going to study in the current paper is quite meaningful, both in mathematics and in applications.

Due to the presence of the terminal state constraint, the value function of the optimal impulse control problem is proved to be locally Hölder (or Lipschitz) continuous only, even under our discovered condition, and it could grow at least linearly (no slower than the growth of the impulse cost). These properties essentially prevent us from directly using the current available techniques to prove the value function being the uniqueness of viscosity solution to the corresponding HJB QVI. Actually, we remind ourselves that the study of uniqueness of viscosity solutions to HJB equations for terminal state constrained problems was not successful in some other situations (see [29, 32] for the so-called stochastic target problems).

On the other hand, the obtained HJB QVI for the value function of our optimal impulse control problem with terminal state constraint looks like that for an optimal impulse control problem *without* terminal state constraint. Then a natural question arises: Can we indirectly characterize our value function by via the problem without constraint? In fact, if we can show that our impulse control problem is equivalent to a problem without terminal state constraint, then our goal is achieved. Some discussions will be carried out and the answer is not definite at the moment because we do not yet have the uniqueness of the viscosity solution to HJB QVI in the function class that our value function belongs to. Combining the above, we see that the issue of unique

viscosity solution characterization for the value function of the problem with terminal state constraint remains challengingly open at the moment.

The rest of the paper is organized as follows. Section 2 is devoted to some preliminary results, including the formulation of the problem, listing of the hypotheses, and the description of the value function domain. In Section 3, we will investigate the continuity of the value function and present an interesting example there. In Section 4, we will derive dynamic programming principle and HJB QVI to which the value function is a viscosity solution. A comparison is made in Section 5 between the optimal impulse control problems with and without terminal state constraint, which reveals some interesting facts. Finally, concluding remarks are collected in Section 6.

2. Preliminary results

In this section, we present some preliminary results.

2.1. Formulation of the problem

Let us first formulate our impulse control problem. Consider the following equation:

$$X(s) = x + \int_t^s f(r, X(r)) \mathrm{d}r + \xi(s), \qquad s \in [t, T], \tag{2.1}$$

where $f : [0, T] \times \mathbb{R}^n \to \mathbb{R}^n$ is a given map, $(t, x) \in [0, T) \times \mathbb{R}^n$ is called an *initial pair*, and

$$\xi(s) = \sum_{k \geqslant 1} \xi_k \mathbf{1}_{[\tau_k, T]}(s), \qquad s \in [t, T] \tag{2.2}$$

is called an *impulse control* with $\{\tau_k\}_{k \geqslant 1} \subset [t, T]$ being a non-decreasing finite sequence, and $\xi_k \in K$, $k \geqslant 1$, called *admissible impulses*, for some non-empty closed convex cone $K \subseteq \mathbb{R}^n$ with the vertex at the origin. In the above, we allow $\tau_k = \tau_{k+1}$ for some $k \geqslant 1$. Let $\mathscr{K}[t, T]$ be the set of all impulse controls of form (2.2). Note that $\xi(\cdot)$ defined by (2.2) is completely determined by the sequence $\{(\tau_k, \xi_k) \mid k \geqslant 1\}$. The notation $\xi(\cdot)$ includes the information of both $\{\xi_k\}$ and $\{\tau_k\}$. Under some mild conditions, for any initial pair $(t, x) \in [0, T) \times \mathbb{R}^n$ and impulse control $\xi(\cdot) \in \mathscr{K}[t, T]$, equation (2.1) admits a unique solution $X(\cdot) \equiv X(\cdot; t, x, \xi(\cdot))$. We also note that (2.1) can be equivalently written as the following impulsive way:

$$\begin{cases} \dot{X}(s) = f(s, X(s)), & s \in (\tau_{k-1}, \tau_k), \qquad k \geqslant 2, \\ X(t) = x, \\ X(\tau_k) = X(\tau_k^-) + \xi_k. \end{cases} \tag{2.3}$$

Clearly, both $\xi(\cdot)$ and $X(\cdot)$ are right-continuous. In addition, we require that the terminal state satisfies the following constraint:

$$X(T) \in \bar{D}, \tag{2.4}$$

where D is a non-empty proper domain in $\mathbb{R}^n$ (non-empty open and connected subset $D \neq \mathbb{R}^n$) with $\bar{D}$ being its closure. We may also call $\bar{D}$ a *target*. For any initial pair $(t, x) \in [0, T) \times \mathbb{R}^n$, we introduce the following associated *admissible impulse control* set

$$\mathscr{K}^x[t, T] = \Big\{ \xi(\cdot) \in \mathscr{K}[t, T] \mid X(T; t, x, \xi(\cdot)) \in \bar{D} \Big\}. \tag{2.5}$$

In general, $\mathscr{K}^x[t,T]$ is different from $\mathscr{K}[t,T]$, and $\mathscr{K}^x[t,T]$ could even be empty for some (t,x). In the case $\mathscr{K}^x[t,T] \neq \varnothing$, to measure the performance of the impulse control $\xi(\cdot)$, we introduce the following cost functional

$$J(t,x;\xi(\cdot)) = \int_t^T g(s,X(s))\mathrm{d}s + h(X(T)) + \sum_{k \geqslant 1} \ell\big(\tau_k, \widetilde{X}(\tau_k - 0), \xi_k\big), \tag{2.6}$$

where $\xi(\cdot)$ is of form (2.2) which is identified with $\{\tau_k, \xi_k\}_{k \geqslant 1}$, and

$$g : [0,T] \times \mathbb{R}^n \to [0,\infty), \quad h : \mathbb{R}^n \to [0,\infty), \quad \ell : [0,T] \times \mathbb{R}^n \times K \to (0,\infty) \tag{2.7}$$

are suitable maps. Here, the terms on the right-hand side of (2.6) are called the *running cost*, the *terminal cost* and the *impulse cost*, respectively. The meaning of $\widetilde{X}(\tau_k - 0)$ stands for the following: Suppose

$$\tau_i < \tau_{i+1} = \tau_{i+2} = \cdots = \tau_{i'} < \tau_{i'+1},$$

then

$$\widetilde{X}(\tau_k - 0) = X(\tau_{i+1} - 0) + \sum_{j=i+1}^{k-1} \xi_j, \qquad i+1 \leqslant k \leqslant i', \qquad \sum_{j=i+1}^i \xi_j \triangleq 0,$$

which is the state right before the impulse ξ_k is made. This is needed only if there are more than one separate impulses made at a same time (although such an impulse cannot be optimal). In the above, we may assume that g and h are just bounded uniformly from below. By a translation (if necessary), we can simply assume that they are non-negative, for convenience. This will be assumed throughout of the paper. We emphasize that the impulse cost $\ell(t,x,\xi)$ is strictly positive. Mimicking the classical case, we formulate the following optimal impulse control problem.

Problem (C). For any initial pair $(t,x) \in [0,T) \times \mathbb{R}^n$, find a $\bar{\xi}(\cdot) \in \mathscr{K}^x[t,T]$ such that

$$J(t,x;\bar{\xi}(\cdot)) = \inf_{\xi(\cdot) \in \mathscr{K}^x[t,T]} J(t,x;\xi(\cdot)) = V(t,x). \tag{2.8}$$

We call $\bar{\xi}(\cdot)$ an *optimal impulse control*, the corresponding $\bar{X}(\cdot) \equiv X(\cdot;t,x,\bar{\xi}(\cdot))$ an *optimal state trajectory*, $(\bar{X}(\cdot), \bar{\xi}(\cdot))$ an *optimal pair*, and $V(\cdot,\cdot)$ the *value function* of Problem (C).

Recall a common convention that $\inf \varnothing = \infty$, regarding $\varnothing \subset \mathbb{R}$. Thus, it is convenient to make the following convention:

$$J(t,x;\xi(\cdot)) = \infty, \qquad \forall \xi(\cdot) \in \mathscr{K}[t,T] \setminus \mathscr{K}^x[t,T]. \tag{2.9}$$

We let

$$\mathscr{D}(V) \equiv \Big\{ (t,x) \in [0,T] \times \mathbb{R}^n \mid V(t,x) \text{ is well-defined and finite} \Big\}, \tag{2.10}$$

which is called the *domain* of the value function $V(\cdot,\cdot)$. Since $g(\cdot,\cdot), h(\cdot)$ are non-negative and $\ell(\cdot,\cdot,\cdot)$ is positive (see (2.7)), one automatically has

$$\mathscr{D}(V) = \Big\{ (t,x) \in [0,T] \times \mathbb{R}^n \mid \mathscr{K}^x[t,T] \neq \varnothing \Big\} \equiv \mathscr{D}(K; \bar{D}). \tag{2.11}$$

The notation $\mathscr{D}(K; \bar{D})$, meaning that K-valued impulse control brings the state to the target set $\bar{D}$, emphasizes the compatibility of the set K, D, and the dynamics (2.1).

2.2. Domain of the value function

Before going further, let us introduce the following hypotheses, which will be used in the following. First, we impose conditions on the terminal state constraint set D and the set K in which the impulse control takes values.

(H1) Let $D \subset \mathbb{R}^n$ be a non-empty proper convex domain (open and connected subset, different from $\mathbb{R}^n$) and $K \subseteq \mathbb{R}^n$ be a closed convex cone with the vertex located at the origin.

Note that K being a closed convex cone with the vertex located at the origin implies that if ξ and ξ' are two admissible impulses, so is $\xi + \xi'$. Also, K is allowed to have empty interior.

Sometime, we need a little stronger conditions for D and K.

(H1′) Let $D \subset \mathbb{R}^n$ be a bounded domain and $K = \mathbb{R}^n$; or $D \subset \mathbb{R}^n$ be a conic domain with the vertex being at the origin and $K \subseteq \mathbb{R}^n$ be a closed convex cone with the vertex located at the origin such that

$$D - K = \mathbb{R}^n. \tag{2.12}$$

Concerning the well-posedness of the state equation, we introduce the following assumptions.

(H2) The map $f : [0, T] \times \mathbb{R}^n \to \mathbb{R}^n$ is continuous and there exists a constant $L > 0$ such that

$$|f(t, x) - f(t, x')| \leqslant L|x - x'|, \qquad \forall t \in [0, T], \ x, x' \in \mathbb{R}^n. \tag{2.13}$$

$$|f(t, 0)| \leqslant L, \qquad \forall t \in [0, T]. \tag{2.14}$$

Now, for the cost functional, we introduce the following two assumptions for the running, the terminal and the impulse costs, respectively.

(H3) The maps $g : [0, T] \times \mathbb{R}^n \to [0, \infty)$ and $h : \mathbb{R}^n \to [0, \infty)$ are continuous. There exist constants $L, \mu > 0$ and $0 < \delta \leqslant 1$ such that

$$0 \leqslant g(t, x), h(x) \leqslant L\big(1 + |x|^{\mu+\delta}\big), \qquad \forall (t, x) \in [0, T] \times \mathbb{R}^n, \tag{2.15}$$

$$|g(t, x) - g(t, x')|, |h(x) - h(x')| \leqslant L\Big(1 + |x|^\mu \vee |x'|^\mu\Big)|x - x'|^\delta, \qquad \forall t \in [0, T], x, x' \in \mathbb{R}^n. \tag{2.16}$$

(H4) The map $\ell : [0, T] \times \mathbb{R}^n \times K \to (0, \infty)$ is continuous. There exist constants $\ell_0, \delta_0, \alpha_0, \alpha > 0$, $\beta \in (0, 1]$ and $L, \mu, \delta > 0$ the same as those in (H3) such that

$$\ell_0 + \alpha_0|\xi|^\beta \leqslant \ell(t, x, \xi) \leqslant L + \alpha|\xi|^\beta, \qquad (t, x, \xi) \in [0, T] \times \mathbb{R}^n \times K, \tag{2.17}$$

$$|\ell(t, x, \xi) - \ell(t, x', \xi)| \leqslant L\Big(1 + |x|^\mu \vee |x'|^\mu\Big)|x - x'|^\delta, \qquad \forall (t, \xi) \in [0, T] \times K, \ x, x' \in \mathbb{R}^n, \tag{2.18}$$

$$\min\big\{\ell(t, x, \xi) + \ell(t, x + \xi, \xi'), \ell(t, x, \xi') + \ell(t, x + \xi', \xi)\big\} - \ell(t, x, \xi + \xi') \geqslant \delta_0, \\ (t, x) \in [0, T] \times \mathbb{R}^n, \ \xi, \xi' \in K, \tag{2.19}$$

$$\ell(t, x, \xi) - L|t - t'| \leqslant \ell(t', x, \xi) \leqslant \ell(t, x, \xi), \qquad t \leqslant t', \ x \in \mathbb{R}^n, \ \xi \in K. \tag{2.20}$$

As we have indicated in the introduction section, one can assume that g and h are bounded below uniformly. Here, we directly assume them to be non-negative just for convenience. Condition (2.17) implies that as long as an impulse is made, no matter how small the ξ is, there is a strictly positive fixed cost ℓ_0. Also, roughly speaking, the larger the $|\xi|$, the larger the cost. Condition (2.19) means that if at (t,x) an impulse of size $\xi + \xi'$ needs to be made, then one should make just one impulse of that size instead of making an impulse of size ξ immediately followed by another with size ξ'. Hence, in an optimal impulse control, $\tau_k < \tau_{k+1}$ if both are impulsive moments. In the case that $\ell(t,x,\xi)$ is independent of x, this condition is reduced to

$$\ell(t, \xi + \xi') < \ell(t, \xi) + \ell(t, \xi'),$$

which is a classical condition assumed in the optimal impulse control problems. Because of this condition, $\xi \mapsto \ell(t,x,\xi)$ should be "sublinear". Hence, $\beta \in (0,1]$ and $\xi \mapsto \ell(t,x,\xi)$ grows at most linearly (see (2.17)). Condition (2.20) means that if an impulse is going to be made, then the later the better, which is essentially due to the discount effect.

In what follows, we call the impulse control that contains no impulses the *trivial impulse control*, denote it by $\xi_0(\cdot)$. Clearly, under $\xi_0(\cdot)$, the state equation is a usual ordinary differential equation (ODE, for short) without jumps. Thus, usual estimates for ODEs are valid. Note that due to the presence of the (strictly positive) impulse cost, the trivial impulse control is different from the zero impulse control (which contains impulses with $\xi_k = 0$). Let us first present the following result concerning the state trajectories.

Proposition 2.1. *Let* (H1)–(H2) *hold. Then for any* $(t,x) \in [0,T] \times \mathbb{R}^n$ *and* $\xi(\cdot) \in \mathcal{K}[t,T]$ *of form* (2.2), *state equation* (2.1) *admits a unique solution* $X(\cdot) = X(\cdot\,; t, x, \xi(\cdot))$, *and the following estimates hold:*

$$|X(s)| \leqslant e^{L(s-t)}(1+|x|) + \sum_{k \geqslant 1} e^{L(s-\tau_k)} |\xi_k| \mathbf{1}_{[\tau_k, T]}(s), \qquad s \in [t,T], \tag{2.21}$$

$$|X(s'+0) - X(s+0)| \leqslant L(s'-s) + \Big[e^{-Lt}(1+|x|) + \sum_{\tau_k \leqslant s'} |\xi_k| e^{-L\tau_k} \Big] \big(e^{Ls'} - e^{Ls} \big) + \sum_{s < \tau_k \leqslant s'} |\xi_k|, \quad 0 \leqslant s < s' \leqslant T. \tag{2.22}$$

If $\widehat{X}(\cdot) = X(\cdot\,; t, \hat{x}, \xi(\cdot))$ *with* $\hat{x} \in \mathbb{R}^n$, *then*

$$|X(s) - \widehat{X}(s)| \leqslant e^{L(s-t)} |x - \hat{x}|, \qquad s \in [t,T]. \tag{2.23}$$

Proof. First of all, for any $(t,x) \in [0,T] \times \mathbb{R}^n$ and $\xi(\cdot) \in \mathcal{K}[t,T]$, by a standard argument, (2.1) admits a unique solution $X(\cdot) = X(\cdot\,; t, x, \xi(\cdot))$. If $\widehat{X}(\cdot) = X(\cdot\,; t, \hat{x}, \xi(\cdot))$, then $X(\cdot) - \widehat{X}(\cdot)$ satisfies

$$X(s) - \widehat{X}(s) = x - \hat{x} + \int_t^s [f(s, X(s)) - f(s, \widehat{X}(s))] \mathrm{d}s, \qquad s \in [t,T].$$

Thus, by (H2) and Gronwall's inequality, we can get (2.23). We now prove (2.21). Let $\xi(\cdot)$ be of form (2.2), and $t \leqslant \tau_{k-1} < \tau_k = \tau_{k+1} = \cdots = \tau_{k'} < \tau_{k'+1}$. Then, on $(\tau_k, \tau_{k'+1})$, state $X(\cdot)$ satisfies (2.3). Thus,

$$|X(s)| \leqslant |X(\tau_k^-)| + \sum_{i=k}^{k'} |\xi_i| + \int_{\tau_k}^s L\big(1 + |X(r)|\big) \mathrm{d}r, \qquad s \in (\tau_k, \tau_{k'+1}).$$

Consequently, by Gronwall's inequality, we have

$$X(s)| \leqslant e^{L(s-\tau_k)}\Big(1 + |X(\tau_k^-)| + \sum_{i=k}^{k'} |\xi_i|\Big), \qquad s \in (\tau_k, \tau_{k'+1}).$$

Then, recursively, we obtain (2.21).

Now, let $t \leqslant s < s' \leqslant T$. Then

$$\begin{aligned}
|X(s'+0) - X(s+0)| &= \Big|\int_s^{s'} f(r, X(r))\mathrm{d}r + \sum_{s<\tau_k} \xi_k \mathbf{1}_{[\tau_k, T]}(s')\Big| \\
&\leqslant \int_s^{s'} L\big(1 + |X(r)|\big)\mathrm{d}r + \sum_{s<\tau_k} |\xi_k|\mathbf{1}_{[\tau_k, T]}(s') \\
&\leqslant \int_s^{s'} L\Big(1 + e^{L(r-t)}(1+|x|) + \sum_{k\geqslant 1} e^{L(r-\tau_k)}|\xi_k|\mathbf{1}_{[\tau_k, T]}(r)\Big)\mathrm{d}r + \sum_{s<\tau_k} |\xi_k|\mathbf{1}_{[\tau_k, T]}(s') \\
&\leqslant L(s'-s) + \Big(e^{-Lt}(1+|x|) + \sum_{\tau_k \leqslant s'} |\xi_k|e^{-L\tau_k}\Big)\big(e^{Ls'} - e^{Ls}\big) + \sum_{s<\tau_k \leqslant s'} |\xi_k|.
\end{aligned}$$

This proves (2.22). $\qquad\square$

In the above, (2.21) and (2.23) are standard; whereas, (2.22) seems to be new, from which, we see that although $s \mapsto X(s)$ might have jumps, these jumps can be controlled in some specific way. This is pretty natural and will be useful in the sequel.

To look at the domain $\mathscr{D}(V)$ of the value function, we first note that under (H1), one always has

$$\{T\} \times \bar{D} \subseteq \mathscr{D}(V) = \mathscr{D}(K; \bar{D}) \subseteq \mathscr{D}(\mathbb{R}^n; \bar{D}) = [0, T] \times \mathbb{R}^n. \tag{2.24}$$

Thus, $\mathscr{D}(V)$ is always non-empty, and when $K = \mathbb{R}^n$, the domain $\mathscr{D}(V)$ of $V(\cdot, \cdot)$ is the whole space $[0, T] \times \mathbb{R}^n$. Let us now present the following result.

Proposition 2.2. *Let* (H1)–(H2) *hold. Let*

$$\partial D \subseteq D - K \equiv \{\eta - \xi \mid \eta \in D, \ \xi \in K\}, \tag{2.25}$$

Then $\mathscr{D}(V)$ *is a non-empty open set in* $[0, T] \times \mathbb{R}^n$.

Proof. We already know that $\mathscr{D}(V)$ is non-empty. Let $(t, x) \in \mathscr{D}(V)$, then there exists an impulse control $\xi(\cdot) \in \mathscr{K}^x[t, T]$ such that $X(T; t, x, \xi(\cdot)) \in \bar{D}$. There are two cases.

Case 1. $X(T; t, x, \xi(\cdot)) \in D$. Then there exists an $\varepsilon > 0$ such that

$$B_\varepsilon\big(X(T; t, x, \xi(\cdot))\big) \subseteq D,$$

where $B_\varepsilon(x)$ is the open ball centered at x with radius ε. Consequently, for $0 < \delta < e^{-L(T-t)}\varepsilon$, as long as $|x - \hat{x}| < \delta$, one has

$$|X(T; t, \hat{x}, \xi(\cdot)) - X(T; t, x, \xi(\cdot))| \leqslant e^{L(T-t)}|x - \hat{x}| < \varepsilon.$$

Hence, $\xi(\cdot) \in \mathscr{K}^{\hat{x}}[t, T]$, leading to $(t, \hat{x}) \in \mathscr{D}(V)$. On the other hand, for $\hat{t} > t$, we let

$$\widehat{\xi}(\cdot) = \sum_{\tau_k \leqslant \hat{t}} \xi_k \mathbf{1}_{[\hat{t}, T]}(\cdot) + \sum_{\tau_k > \hat{t}} \xi_k \mathbf{1}_{[\tau_k, T]}(\cdot).$$

This amounts to moving all the impulses no later than $\hat{t}$ to $\hat{t}$. Denote $\widehat{X}(\cdot) = X(\cdot; \hat{t}, x, \widehat{\xi}(\cdot))$. Then for $s \in [\hat{t}, T]$,

$$|X(s) - \widehat{X}(s)| \leqslant e^{L(s - \hat{t})} \Big| X(\hat{t}) - x - \sum_{\tau_k \leqslant \hat{t}} \xi_k \Big|$$

$$\leqslant e^{L(s - \hat{t})} \int_t^{\hat{t}} |f(r, X(r))| \mathrm{d}r \leqslant L e^{L(s - \hat{t})} \int_t^{\hat{t}} (1 + |X(r)|) \mathrm{d}r$$

$$\leqslant L e^{L(s - \hat{t})} \int_t^{\hat{t}} \Big(1 + e^{L(r - t)}(1 + |x|) + \sum_{t \leqslant \tau_k \leqslant r} e^{L(r - \tau_k)} |\xi_k| \Big) \mathrm{d}r \leqslant C \Big(1 + |x| + \sum_{t \leqslant \tau_k \leqslant \hat{t}} |\xi_k| \Big)(\hat{t} - t).$$

Hereafter, $C > 0$ stands for a generic constant which could be different from line to line. Thus, when $\hat{t} - t > 0$ small enough, we have $X(T; \hat{t}, x, \widehat{\xi}(\cdot)) \in D$, leading to $\widehat{\xi}(\cdot) \in \mathscr{K}^x[\hat{t}, T]$. Finally, for $\hat{t} < t$, we take

$$\widehat{\xi}(\cdot) = \sum_{k \geqslant 1} \xi_k \mathbf{1}_{[\tau_k, T]}(\cdot),$$

i.e., we make a trivial extension of $\xi(\cdot)$ from $[t, T]$ to $[\hat{t}, T]$. Denote $\widehat{X}(\cdot) = X(\cdot; \hat{t}, x, \widehat{\xi}(\cdot))$. Then

$$|\widehat{X}(t) - x| = |X(t; \hat{t}, x, \xi_0(\cdot)) - x| \leqslant \int_{\hat{t}}^t |f(r, \widehat{X}(r))| \mathrm{d}r \leqslant L \int_{\hat{t}}^t \Big(1 + |\widehat{X}(r)| \Big) \mathrm{d}r$$

$$\leqslant L \int_{\hat{t}}^t e^{L(r - \hat{t})}(1 + |x|) \mathrm{d}r \leqslant C(1 + |x|)(t - \hat{t}).$$

Hence, for $s \in [t, T]$,

$$|X(s) - \widehat{X}(s)| \leqslant e^{L(s - t)}|x - \widehat{X}(t)| \leqslant C(1 + |x|)(t - \hat{t}).$$

Consequently, when $t - \hat{t} > 0$ small enough, $X(T; \hat{t}, x, \widehat{\xi}(\cdot)) \in D$, leading to $\widehat{\xi}(\cdot) \in \mathscr{K}^x[\hat{t}, T]$. Combining the above, we obtain

$$(\hat{t}, \hat{x}) \in \mathscr{D}(V), \qquad \text{if } |\hat{t} - t| + |\hat{x} - x| \text{ is small enough.}$$

Case 2. $X(T; t, x, \xi(\cdot)) \in \partial D$. In this case, by (2.25), there exists a $\bar{\xi} \in K$ such that by defining

$$\widehat{\xi}(\cdot) = \sum_{k \geqslant 1} \xi_k \mathbf{1}_{[\tau_k, T]}(\cdot) + \bar{\xi} \mathbf{1}_{\{T\}}(\cdot),$$

we have

$$X(T; t, x, \widehat{\xi}(\cdot)) \in D.$$

Then it is reduced to Case 1. □

The following result tells us more about $\mathscr{D}(V)$.

Proposition 2.3. *Let* (H1)–(H2) *hold.*
(i) *It holds that*

$$\bar{D} - K = \mathbb{R}^n, \tag{2.26}$$

if and only if (2.12) *holds. In this case,*

$$\mathscr{D}(V) = \mathscr{D}(K; \bar{D}) = [0, T] \times \mathbb{R}^n. \tag{2.27}$$

(ii) *If D is bounded, then* (2.26) *holds if and only if $K = \mathbb{R}^n$.*

Proof. (i) First of all, it is always true that $D - K \subseteq \bar{D} - K$. Thus, the sufficiency is clear. We now prove the necessity. Under (H1), both $\bar{D} - K$ and $D - K$ are convex. Moreover, $D - K$ is open and

$$\bar{D} - K \subseteq \overline{D - K}. \tag{2.28}$$

In fact, for any $x_0 \in D - K$, we have some $\eta_0 \in D$ and $\xi_0 \in K$ such that

$$x_0 = \eta_0 - \xi_0.$$

Since D is open, there exists a $\delta > 0$ such that

$$\mathscr{O}_\delta(\eta_0) \equiv \{\eta \in \mathbb{R}^n \mid |\eta - \eta_0| < \delta\} \subseteq D.$$

Now, for any $x \in \mathscr{O}_\delta(x_0)$, we have

$$\eta \equiv \eta_0 + x - x_0 \in \mathscr{O}_\delta(\eta_0) \subseteq D,$$

which leads to

$$x = x_0 + x - x_0 = \eta_0 + (x - x_0) - \xi_0 \equiv \eta - \xi_0 \in D - K.$$

Thus, $D - K$ is open. The convexity of $\bar{D} - K$ and $D - K$ is clear. Next, for any

$$x = \eta - \xi \in \bar{D} - K, \qquad \eta \in \bar{D}, \quad \xi \in K,$$

we can find a sequence $\eta_k \in D$ such that $\eta_k \to \eta$. Then

$$x = \lim_{k \to \infty} (\eta_k - \xi) \in \overline{D - K},$$

proving (2.28). Now, if $D - K \neq \bar{D} - K$, by the convexity of $D - K$, we must have $\overline{D - K} \neq \mathbb{R}^n$. This, together with (2.28), contradicts (2.26).

Finally, we prove (2.27). For any $(t, x) \in [0, T] \times \mathbb{R}^n$, under the trivial impulse control $\xi_0(\cdot)$, the state will arrive at $X(T - 0; t, x, \xi_0(\cdot)) \in \mathbb{R}^n$. By (2.26), we have some $\eta \in \bar{D}$ and $\xi \in K$ such that

$$X(T - 0; t, x, \xi_0(\cdot)) = \eta - \xi.$$

Then by defining impulse control

$$\widehat{\xi}(\cdot) = \xi \mathbf{1}_{\{T\}}(\cdot),$$

we have

$$X(T; t, x, \widehat{\xi}(\cdot)) = X(T - 0; t, x, \xi_0(\cdot)) + \xi = \eta \in \bar{D}.$$

Thus, $\mathscr{K}^x[t, T] \neq \varnothing$. This proves our conclusion (see (2.11)).

(ii) First of all, if $K = \mathbb{R}^n$, then of course. (2.26) holds. Now, if $K \neq \mathbb{R}^n$, then there must be a $\zeta \in K$, $|\zeta| = 1$ such that

$$\langle \zeta, \xi \rangle \geqslant 0, \qquad \forall \xi \in K.$$

Now, we claim that $\lambda \zeta \notin \bar{D} - K$ for large enough $\lambda > 0$. In fact, if there exists an $\eta^\lambda \in \bar{D}$ and a $\xi^\lambda \in K$ such that

$$\lambda \zeta = \eta^\lambda - \xi^\lambda.$$

This leads to

$$\eta^\lambda = \lambda \zeta + \xi^\lambda.$$

Hence,

$$\langle \eta^\lambda, \zeta \rangle = \lambda + \langle \xi^\lambda, \zeta \rangle \geqslant \lambda.$$

Since $\{\eta^\lambda\}_{\lambda > 0}$ is bounded, we may assume that $\eta^\lambda \to \bar{\eta}$. But this will lead to a contradiction. $\square$

One of the most interesting examples satisfying (2.26) is the following:

$$\bar{D} = K = \mathbb{R}^n_+ \equiv \{x \in \mathbb{R}^n \mid x_i \geqslant 0\}.$$

Consequently, for such a case, one has (2.27). The above proposition gives two important cases: D is a bounded set with $K = \mathbb{R}^n$ and D is unbounded with $K \neq \mathbb{R}^n$ such that (2.26) holds. They are mutually exclusive. However, we point out that they are not exhausting. Here is a simple example that is neither of the above cases: In $\mathbb{R}^2$, let

$$K = \{(x_1, x_2) \mid x_1, x_2 \geqslant 0\}, \qquad D = \left\{(x_1, x_2) \mid x_2 > \frac{1}{1 - x_1}, \ x_1 < 1\right\}.$$

Then D is unbounded and

$$\bar{D} \subseteq \bar{D} - K = \{(x_1, x_2) \mid x_1 < 1\} \neq \mathbb{R}^2.$$

If we regard (2.26) as the best case, then since $0 \in K$, the worst case should be

$$\bar{D} - K = \bar{D}. \tag{2.29}$$

An example of such is the following:

$$K = \{(x_1, x_2) \mid x_1, x_2 \geqslant 0\}, \qquad D = \{(x_1, x_2) \mid x_1, x_2 < 0\}. \tag{2.30}$$

For the case that (2.26) fails, including the case of (2.29), when $X(T-0) \notin \bar{D}$, there is no way to make an impulse at T so that $X(T) \in \bar{D}$. Therefore,

$$\left[\{T\} \times (\bar{D}-K)^c\right] \cap \mathscr{D}(V) = \varnothing, \quad \text{or} \quad \left[\{T\} \times \mathbb{R}^n\right] \cap \mathscr{D}(V) \neq \{T\} \times \mathbb{R}^n.$$

For such a case, one has to make impulses before T and drive the state to $\bar{D}$ (at T) via the state equation. Thus, it might still be possible that

$$\left[\{t\} \times \mathbb{R}^n\right] \cap \mathscr{D}(V) = \{t\} \times \mathbb{R}^n,$$

for some $t \in [0,T)$. We will see a concrete example below.

Let us now look at the following simple example to get some more feeling.

Example 2.4. Consider

$$X(s) = x + (s-t) + \xi(s), \qquad s \in [t, T]. \tag{2.31}$$

We consider several cases.

(i) $K = [0,\infty)$, $D = (0,1)$. For this case,

$$\bar{D} - K = [0,1] - [0,\infty) = (-\infty, 1],$$

and

$$\mathscr{D}(V) = \left\{ (t,x) \in [0,T] \times \mathbb{R} \mid x + T - t \leqslant 1 \right\}.$$

(ii) $K = (-\infty, 0]$, $D = (0,1)$. For this case,

$$\bar{D} - K = [0,1] - (-\infty, 0] = [0, \infty),$$

and

$$\mathscr{D}(V) = \left\{ (t,x) \in [0,T] \times \mathbb{R} \mid x + T - t \geqslant 0 \right\}.$$

(iii) $K = [0,\infty)$, $D = (0,\infty)$. For this case,

$$\bar{D} - K = [0,\infty) - [0,\infty) = \mathbb{R}, \qquad \mathscr{D}(V) = [0,T] \times \mathbb{R}.$$

(iv) $K = [0,\infty)$, $D = (-\infty, 0)$. For this case,

$$\bar{D} - K = (-\infty, 0] - [0,\infty) = (-\infty, 0] = \bar{D},$$

and

$$\mathscr{D}(V) = \left\{ (t,x) \in [0,T] \times \mathbb{R} \mid x + T - t \leqslant 0 \right\}.$$

Now, we look at the general situation. Under (H1)–(H2), for any $x \in D$, we may let $B_\varepsilon(x) \subseteq D$, with $B_\varepsilon(x)$ being the ball centered at x with radius ε. Let $t \in [0,T]$ such that $T - t > 0$ is small enough so that

$$|X(s; t, x, \xi_0(\cdot)) - x| \leqslant (1+L)e^{LT}(1+|x|)(T-t) < \varepsilon, \qquad s \in [t, T].$$

This means that $\xi_0(\cdot) \in \mathcal{K}^x[t,T]$. Hence, under (H1)–(H2), the following is always true:

$$\mathscr{D}(V) = \mathscr{D}(K; \bar{D}) \neq \varnothing. \tag{2.32}$$

We now would like to get a more precise description of $\mathscr{D}(V)$. For state equation (2.1), we consider the following "backward" system

$$Y(s) = \zeta - \int_s^T f(r, Y(r))\mathrm{d}r, \qquad s \in [0, T],$$

with $\zeta \in \bar{D} - K$. The solution is denoted by $Y(\cdot; T, \zeta)$. Let

$$\mathscr{Y}(t; T, \bar{D} - K) = \{Y(t; T, \zeta) \mid \zeta \in \bar{D} - K\}.$$

For any $x \in \mathscr{Y}(t; T, \bar{D} - K)$, one has some $\zeta \in \bar{D} - K$ such that

$$x = Y(t; T, \zeta).$$

Then, with the trivial impulse control $\xi_0(\cdot)$, we have

$$X(T - 0; t, x, \xi_0(\cdot)) = Y(T; T, \zeta) = \zeta \in \bar{D} - K.$$

One can choose some $\xi \in K$ such that

$$X(T - 0; t, x, \xi_0(\cdot)) + \xi \in \bar{D}.$$

Thus, $\mathscr{Y}(t; T, \bar{D} - K)$ is the set of all possible initial state that if the system starts at (t, x), the state will reach $\bar{D} - K$ at T under $\xi_0(\cdot)$. Then, under a possible impulse at T, the state will hit $\bar{D}$. Now, let $\Pi = \{t_0, t_1, \cdots, t_N\}$ be a partition of $[t, T]$ with $t = t_0 < t_1 < t_2 < \cdots < t_N = T$. Then we may define inductively

$$\begin{aligned}
\mathscr{Y}_\Pi^N &\equiv \bar{D}, \\
\mathscr{Y}_\Pi^{N-1} &= \mathscr{Y}(t_{N-1}; t_N, \bar{D} - K) \equiv \mathscr{Y}(t_{N-1}; t_N, \mathscr{Y}_\Pi^N - K), \\
\mathscr{Y}_\Pi^{N-2} &= \mathscr{Y}(t_{N-2}; t_{N-1}, \mathscr{Y}_\Pi^{N-1} - K), \\
&\cdots\cdots \\
\mathscr{Y}_\Pi^1 &= \mathscr{Y}(t_1; t_2 \mathscr{Y}_\Pi^2 - K), \\
\mathscr{Y}_\Pi^0 &= \mathscr{Y}(t_0; t_1 \mathscr{Y}_\Pi^1 - K).
\end{aligned}$$

We denote

$$\mathscr{Y}(t; \Pi) = \mathscr{Y}_\Pi^0 - K,$$

which is the set of all initial states $x \in \mathbb{R}^n$ such that if the system starts at (t, x), with possible impulses at $t_0, t_1, \cdots, t_N$, the state will reach $\bar{D}$ at T. This can be described by the following:

$$\bar{D} \equiv \mathscr{Y}_\Pi^N \Leftarrow \mathscr{Y}_\Pi^N - K$$
$$\Uparrow$$
$$\mathscr{Y}_\Pi^{N-1} - K \Rightarrow \mathscr{Y}_\Pi^{N-1} \equiv \mathscr{Y}(t_{N-1}; t_N, \mathscr{Y}_\Pi^N - K)$$
$$\Uparrow$$
$$\mathscr{Y}(t_{N-2}; t_{N-1}, \mathscr{Y}_\Pi^{N-1} - K) \equiv \mathscr{Y}_\Pi^{N-2} \Leftarrow \mathscr{Y}_\Pi^{N-2} - K$$
$$\cdots\cdots\cdots\cdots$$
$$\Uparrow$$
$$\mathscr{Y}_\Pi^1 - K \Rightarrow \mathscr{Y}_\Pi^1 \equiv \mathscr{Y}(t_1; t_2, \mathscr{Y}_\Pi^2 - K)$$
$$\Uparrow$$
$$\mathscr{Y}(t_0; t_1, \mathscr{Y}_\Pi^1 - K) \equiv \mathscr{Y}_\Pi^0 \Leftarrow \mathscr{Y}_\Pi^0 - K$$

In the above, horizontal arrows represent making impulses, and upper arrows represent running state equations. Clearly, for any two partitions Π_1 and Π_2 of $[t, T]$ with $\Pi_1 \subseteq \Pi_2$, i.e., Π_2 is a refinement of Π_1, we have

$$\mathscr{Y}(t; \Pi_1) \subseteq \mathscr{Y}(t; \Pi_2).$$

Hence, we may define

$$\mathscr{Y}(t) = \bigcup_\Pi \mathscr{Y}(t; \Pi) = \lim_{\|\Pi\| \to 0} \mathscr{Y}(t; \Pi),$$

where $\|\Pi\|$ is the mesh size of Π defined by

$$\|\Pi\| = \max_{1 \leqslant k \leqslant N} (t_k - t_{k-1}).$$

From the construction, we see that $\mathscr{Y}(t)$ is the set of all initial state that if the system starts from (t, x), then with impulse controls, the state can reach $\bar{D}$ at T, i.e.,

$$\mathscr{K}^x[t, T] \neq \varnothing \qquad \Longleftrightarrow \qquad x \in \mathscr{Y}(t). \tag{2.33}$$

Hence, we have the following characterization of $\mathscr{D}(V)$:

$$\overline{\mathscr{D}(V)} = \overline{\bigcup_{t \in [0, T]} \big[\{t\} \times \mathscr{Y}(t)\big]}. \tag{2.34}$$

The following example gives a concrete construction of $\mathscr{Y}(t)$.

Example 2.5. Consider

$$\begin{cases} X_1(s) = x_1 + \displaystyle\int_t^s X_2(r)\mathrm{d}r + \xi_1(s), \\ X_2(s) = x_2 - \displaystyle\int_t^s X_1(r)\mathrm{d}r + \xi_2(s), \end{cases} \qquad s \in [t, T].$$

Let

$$D = \{(x_1, x_2) \mid x_1^2 + x_2^2 < 1\}, \qquad K = \mathbb{R}_+^2 \equiv \{(x_1, x_2) \mid x_1, x_2 \geqslant 0\}.$$

The backward system reads

$$\begin{cases} Y_1(s) = \eta_1 - \displaystyle\int_s^T Y_2(r)\mathrm{d}r - \xi_1(s), \\ Y_2(s) = \eta_2 + \displaystyle\int_s^T Y_1(r)\mathrm{d}r - \xi_2(s), \end{cases} \qquad s \in [0, T].$$

For any $\eta = (\eta_1, \eta_2) \in \bar{D}$ and $(\zeta_1, \zeta_2) \in K$, let

$$\begin{pmatrix} Y_1(s) \\ Y_2(s) \end{pmatrix} = \begin{pmatrix} \cos(T - s) & -\sin(T - s) \\ \sin(T - s) & \cos(T - s) \end{pmatrix} \begin{pmatrix} \eta_1 - \zeta_1 \\ \eta_2 - \zeta_2 \end{pmatrix}.$$

Note that as s decreases from T, the vector $(Y_1(s), Y_2(s))^\top$ turns counter-clockwise. We may keep making impulses to see that

$$\mathscr{Y}(T) = \operatorname{co}\left(\bar{D} \cup \{(x_1, x_2) \mid x_1, x_2 \leqslant 0\}\right),$$
$$\mathscr{Y}(t) = \operatorname{co}\left(\mathscr{Y}(T) \cup \{(x_1, x_2) \mid x_2 \leqslant x_1 \cot(T - t), \ x_1 > 0\}\right), \qquad 0 \leqslant T - t < \frac{\pi}{4},$$
$$\mathscr{Y}(T - \frac{\pi}{2}) = \{(x_1, x_2) \mid x_2 \leqslant 1\}, \qquad \mathscr{Y}(t) = \mathbb{R}^2, \qquad T - t > \frac{\pi}{2},$$

where $\operatorname{co}(M)$ is the convex hull of the set M, *i.e.*, the smallest convex set containing M. In the illustrative figures, the blue arrow lines give the directions of impulses; the dashed arcs give the directions of the points turning. Thus, in the last figure (of the situation $T - t > \frac{\pi}{2} + \varepsilon$, any initial point (x_1, x_2) with $x_2 > 1$, one could first make a horizontal impulse $\xi = (\xi_1, 0)$ so that $(x_1 + \xi_1, x_2)$ is on the right of the dashed red line. Then by the original system which makes the point turning clockwise, and at $t = T$, the point will be in $\mathscr{Y}(T)$. By making an impulse at $t = T$, the state will get into $\bar{D}$.

From the above, we see that

$$\mathscr{D}(V) \cap \left[\{t\} \times \mathbb{R}^n\right] = \mathscr{D}(K; \bar{D}) \cap \left[\{t\} \times \mathbb{R}^n\right] \neq \{t\} \times \mathbb{R}^n, \qquad 0 \leqslant T - t \leqslant \frac{\pi}{2},$$
$$\mathscr{D}(V) \cap \left[\{t\} \times \mathbb{R}^n\right] = \mathscr{D}(K; \bar{D}) \cap \left[\{t\} \times \mathbb{R}^n\right] = \{t\} \times \mathbb{R}^n, \qquad T - t > \frac{\pi}{2}.$$

However, one has

$$\mathscr{D}(V) = \mathscr{D}(K; \bar{D}) \supseteq [0, T] \times \bar{D}.$$

The above (2.34) gives a characterization for the domain $\mathscr{D}(V)$ of the value function, in some sense. More needs to be done. For example, what will be the boundary of such a domain, how the value function behaves near the boundary of this domain, and so on. We are going to leave these problems open for the time-being, and hope that we will be able to report some relevant results in the near future. Instead, in the current paper, we consider the things more relevant to the continuity of the value function. Let us make some more preparations.
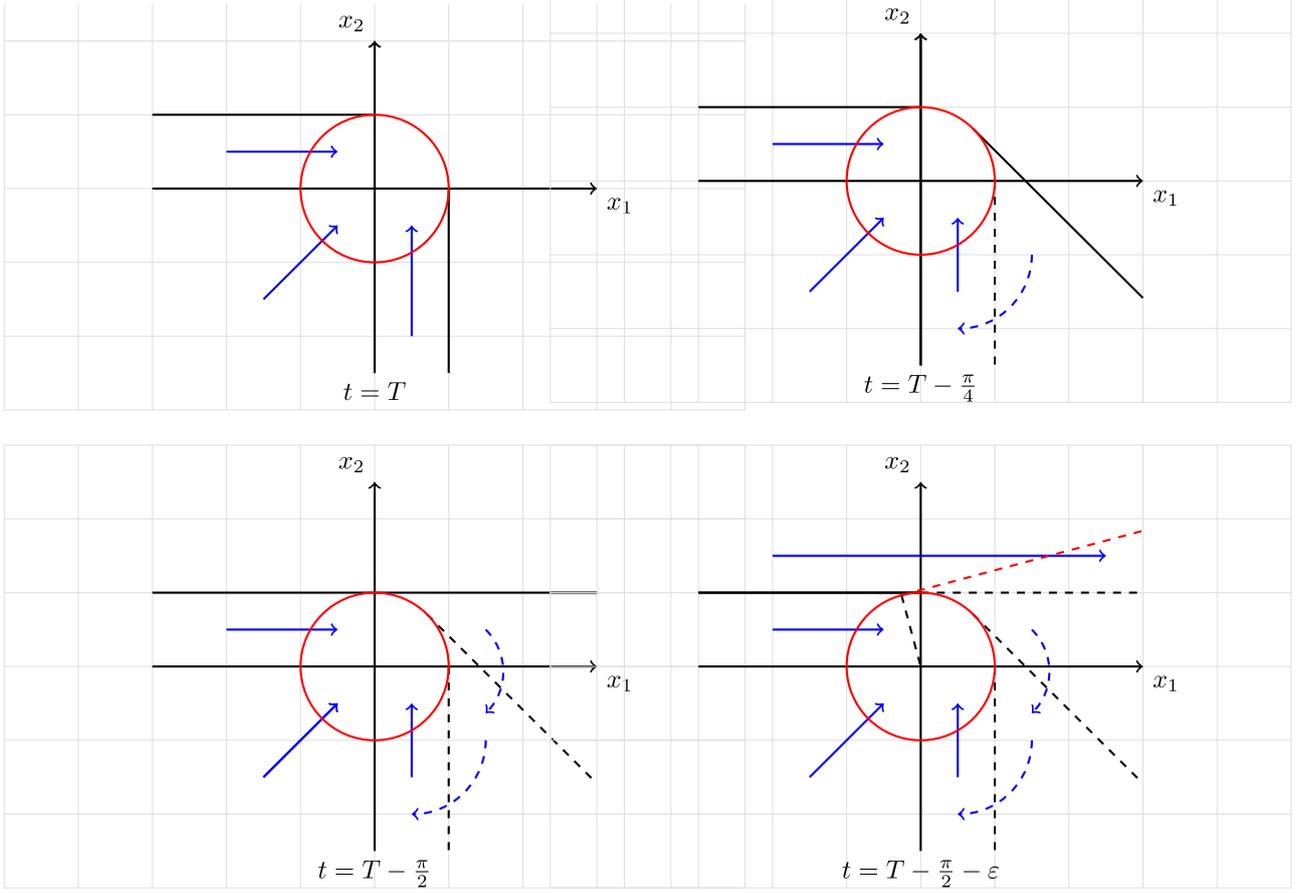
Lemma 2.6. *Let* (H1) *and* (2.12) *hold. Then there exists a nondecreasing continuous function* $\nu : (0, \infty) \to (0, \infty)$ *only depending on* K *and* D *such that for any* $x \in \mathbb{R}^n$, *there exists a* $\xi \in K$ *satisfying*

$$x + \xi \in \bar{D}, \qquad |\xi| \leqslant \nu(|x|). \tag{2.35}$$

Further, if (H1′) *holds, then for some constant* C_0, *only depending on* D *and* K,

$$\nu(r) = C_0(1 + r). \tag{2.36}$$

Proof. First we claim that for any fixed $M > 0$, there exists a $\nu_M > 0$ only depending on K and D such that for any $x \in \mathbb{R}^n$ with $|x| \leqslant M$, there exists a $\xi \in K$ with $|\xi| \leqslant \nu_M$ satisfying $x + \xi \in \bar{D}$, *i.e.*,

$$d(x + \xi, \bar{D}) \equiv \inf_{\eta \in \bar{D}} |x + \xi - \eta| = 0.$$

Suppose this claim fails. Then there exists some $M > 0$ such that one can find a sequence $x_k \in \mathbb{R}^n$ with $|x_k| \leqslant M$, and

$$\inf_{\eta \in \bar{D}} |x_k + \xi - \eta| = d(x_k + \xi, \bar{D}) \geqslant \delta_k, \qquad \forall \xi \in K, \ |\xi| \leqslant k, \ k \geqslant 1,$$

for some $\delta_k > 0$. We may assume $x_k \to x_0$. Now for x_0, by (2.12), there exists a $\xi_0 \in K$ such that

$$\eta_0 = x_0 + \xi_0 \in D.$$

Hence, for $k > |\xi_0|$, large enough, we have (noting that D is open)

$$x_k + \xi_0 = x_k - x_0 + x_0 + \xi_0 = x_k - x_0 + \eta_0 \in D \subseteq \bar{D}.$$

This is a contradiction. Hence, the claim is true. Now, we define

$$\nu(x) = \begin{cases} \nu_{k-1} + 2(\nu_k - \nu_{k-1})(|x| - k + 1), & k - 1 \leqslant |x| \leqslant k - \frac{1}{2}, \\ \nu_k, & k - \frac{1}{2} \leqslant |x| \leqslant k, \end{cases} \qquad k \geqslant 1.$$

which satisfies the our requirement. Hence, (2.35) holds.

Now, let (H$'$) holds. If D is bounded with the bound $M > 0$, then for any $x \in \mathbb{R}^n$, let $\xi \in K = \mathbb{R}^n$ such that $x + \xi = \eta \in D$. Clearly,

$$|\xi| \leqslant |x| + |\eta| \leqslant M + |x|,$$

which is what we want.

Finally, let D be a conic domain with vertex located at the origin. Then under our condition (2.12), we can find a $\xi_0 \in D \cap K$, $\xi_0 \neq 0$. (Note here that $0 \notin D$.) Now, for any $x \in \mathbb{R}^n$, if $x \in \bar{D}$ (which includes the case that $x = \lambda\xi_0$ for some $\lambda \geqslant 0$), we trivially have $x = x - 0 \in \bar{D} - K$. Hence, by taking $\eta = x$ and $\xi = 0$, we have (2.35)–(2.36). Next, let $x \notin \bar{D}$. If

$$x = -\lambda\xi_0,$$

for some $\lambda > 0$, then by taking $\eta = 0 \in \bar{D}$ and $\xi = \lambda\xi_0$, we have (2.35)–(2.36). Hence, we need only to look at the case that $x \notin \bar{D}$, with x and ξ_0 being linearly independent. Consider the two-dimensional space $\mathbb{H}$ spanned by x and ξ_0. After a proper linear transformation, we may assume the following situation in $\mathbb{R}^2$:

$$D \cap K \ni \xi_0 = (1,0)^\top, \quad \{-\lambda\xi_0 \mid \lambda \geqslant 0\} = \{(-\lambda, 0)^\top \mid \lambda \geqslant 0\} = (-\infty, 0] \times \{0\},$$
$$\bar{D} \cap \mathbb{H} = \{(\eta_1, \eta_2)^\top \mid -\gamma_1\eta_1 \leqslant \eta_2 \leqslant \gamma_2\eta_1, \ \eta_1 \geqslant 0\},$$

for some $\gamma_1, \gamma_2 > 0$ (depending on D and K). Then for $(x_1, x_2) \in \mathbb{R}^2 \setminus \bar{D}$, in the case that $x_2 > 0$, we must have $x_2 > \gamma_2 x_1$. Take

$$\eta_1 = \frac{x_2}{\gamma_2} > 0, \quad \eta_2 = x_2, \quad \lambda = \frac{x_2}{\gamma_2} - x_1 > 0. \tag{2.37}$$

One has $\eta \equiv (\eta_1, \eta_2)^\top \in \bar{D} \cap \mathbb{H} \subseteq \bar{D}$, $\xi \equiv \lambda\xi_0 \in K$, and

$$(\eta_1, \eta_2) - \lambda(1,0) = \left(\frac{x_2}{\gamma_2}, x_2\right) - \left(\frac{x_2}{\gamma_2} - x_1\right)(1,0) = (x_1, x_2) = x.$$

Clearly,

$$|\xi| = \lambda \leqslant \frac{\sqrt{1 + \gamma_2^2}}{\gamma_2}|x|, \quad |\eta| = \frac{\sqrt{1 + \gamma_2^2}}{\gamma_2}|x_2| \leqslant \frac{\sqrt{1 + \gamma_2^2}}{\gamma_2}|x|.$$

Thus, (2.35)–(2.36) hold. Likewise, in the case $x_2 < 0$, we must have $x_2 < -\gamma_1 x_1$. Take

$$\eta_1 = \frac{-x_2}{\gamma_1} > 0, \quad \eta_2 = x_2, \quad \lambda = -\frac{x_2}{\gamma_1} - x_1 > 0. \tag{2.38}$$

Then $\eta = (\eta_1, \eta_2)^\top \in \bar{D} \cap \mathbb{H} \subseteq \bar{D}$, $\xi = \lambda \xi_0 \in K$, and

$$(\eta_1, \eta_2) - \lambda(1,0) = \left(\frac{-x_2}{\gamma_1}, x_2 \right) + \left(\frac{x_2}{\gamma_1} + x_1 \right)(1,0) = (x_1, x_2) = x.$$

Also,

$$|\xi| = \lambda \leqslant \frac{\sqrt{1+\gamma_1^2}}{\gamma_1} |x|, \quad |\eta| = \frac{\sqrt{1+\gamma_1^2}}{\gamma_1} |x_2| \leqslant \frac{\sqrt{1+\gamma_1^2}}{\gamma_1} |x|.$$

Hence, (2.35)–(2.36) hold for this case as well. $\qquad\square$

The point of the above result is that under condition (2.12), the size of the impulse ξ that drives the state x to the constraint D is controlled by $|x|$. This is very important below. We also note that (2.12) seems to be a little stronger than (2.26). In the case that $K^\circ \neq \varnothing$ (K° is the interior of K), they are equivalent. It is not clear to us at the moment if they are equivalent in general. Also, we point out that the function $r \to \nu(r)$ can grow arbitrarily fast. Here is a simple example.

Example 2.7. Let $\nu_0 : [0, \infty) \to [0, \infty)$ be a continuous strictly increasing function with $\nu_0(0) = 0$ and $\nu_0(r) \to \infty$ as $r \to \infty$ (for example, $\nu_0(r) = e^r - 1$). Let

$$D = \left\{ (x_1, x_2) \in \mathbb{R}^2 \mid x_1 > \nu_0(|x_2|) \right\}, \qquad K = \left\{ (x_1.x_2) \in \mathbb{R}^2 \mid x_1, x_2 \geqslant 0 \right\}.$$

Then, for any $(x_1, x_2) \in \mathbb{R}^2 \setminus \bar{D}$, one can take $\xi = (\xi_1, \xi_2) \in K$ with

$$\xi_1 \geqslant \nu_0(|x_2|) - x_1, \qquad \xi_2 = 0,$$

which will lead to $x + \xi \in \bar{D}$. This is actually the best choice as far as the norm $|\xi|$ is concerned. Clearly, if $\nu(\cdot)$ is the function appeared in Lemma 2.6, then with $x_1 = 0$, one has

$$\nu_0(|x_2|) \leqslant \xi_1 \leqslant |\xi| \leqslant \nu(|x_2|), \qquad \forall x_2 \in \mathbb{R}.$$

Thus, $\nu(\cdot)$ cannot be growing slower than $\nu_0(\cdot)$.

To conclude this section, let us present one more example.

Example 2.8. Let $K = \mathbb{R}^2$, and

$$D = \left\{ (x_1, x_2) \in \mathbb{R}^2 \mid x_1 \geqslant 0 \right\},$$

then for any $x = (x_1, 0) \in \mathbb{R}^2$ with $x_1 < 0$, we see that the $\xi = (\xi_1, \xi_2) \in K$ that makes $x + K \in \bar{D}$ with smallest possible $|\xi|$ should be $\xi = (-x_1, 0)$ and $|\xi| = |x|$.

The point that we want to make in the above example is that as long as $D \neq \mathbb{R}^n$ and D is convex, the function $\nu(r)$ appears in Lemma 2.6 will be at least of linear growth.

3. Properties of the value functions

In this section, we will present some properties of the value function $V(\cdot, \cdot)$, including its continuity.

3.1. Some bounds

Our goal in this subsection is to obtain, under certain conditions, including (H1)–(H4), the bounds of the value functions, the smaller class of impulse controls on which the value functions are the infimum of the cost functional, and each impulse control in this smaller class has no more than a fixed number of impulses with the sizes of the impulses being bounded. More precisely, we have the following result.

Proposition 3.1. *Let* (H1)–(H4) *and* (2.12) *hold. Then* $\mathscr{D}(V) = [0,T] \times \mathbb{R}^n$ *and*

$$0 \leqslant V(t,x) \leqslant \bar{\nu}(|x|), \qquad \forall (t,x) \in [0,T] \times \mathbb{R}^n, \tag{3.1}$$

for some continuous increasing function $\bar{\nu} : [0,\infty) \to [0,\infty)$. *Moreover, for any* $(t,x) \in [0,T] \times \mathbb{R}^n$, *Problem* (C) *admits an optimal impulse control, and*

$$V(t,x) = \inf_{\xi(\cdot) \in \mathscr{K}_0^x[t,T]} J(t,x;\xi(\cdot)), \tag{3.2}$$

where

$$\mathscr{K}_0^x[t,T] = \left\{ \xi(\cdot) \equiv \sum_{k=1}^N \xi_k \mathbf{1}_{[\tau_k,T]}(\cdot) \in \mathscr{K}^x[t,T] \mid N \leqslant \frac{\bar{\nu}(|x|) + 1}{\ell_0}, \quad \sum_{k=1}^N |\xi_k|^\beta \leqslant \frac{\bar{\nu}(|x|) + 1}{\alpha_0} \right\}. \tag{3.3}$$

Further, if (H1$'$) *is assumed, then the following holds:*

$$0 \leqslant V(t,x) \leqslant C(1 + |x|^{(\mu+\delta)\vee\beta}), \qquad \forall (t,x) \in [0,T] \times \mathbb{R}^n, \tag{3.4}$$

and in the definition of $\mathscr{K}_0^x[t,T]$, $\bar{\nu}(r) \leqslant C(1 + r^{(\mu+\delta)\vee\beta})$.

Proof. For any $(t,x) \in [0,T] \times \mathbb{R}^n$, under the trivial impulse control $\xi_0(\cdot)$, the state $X(\cdot)$ satisfies an ODE without jumps. Thus, we have

$$|X(s;t,x,\xi_0(\cdot))| \leqslant e^{LT}(1 + |x|), \qquad t \leqslant s < T.$$

Under (H1), making use of Lemma 2.6, there is a $\xi_T \in K$ such that

$$X(T-0;t,x,\xi_0(\cdot)) + \xi_T \in \bar{D}, \qquad |\xi_T| \leqslant \nu\big(|X(T-0;t,x,\xi_0(\cdot))|\big),$$

Then we define

$$\widehat{\xi}(\cdot) = \xi_T \mathbf{1}_{\{T\}}(\cdot) \in \mathscr{K}^x[t,T].$$

This is the impulse control that only makes one impulse at T and make the state jump into $\bar{D}$. Clearly,

$$\begin{aligned}
0 \leqslant V(t,x) \leqslant J(t,x;\widehat{\xi}(\cdot)) &= \int_t^T g(s,X(s))\mathrm{d}s + h\big(X(T-0) + \xi_T\big) + \ell\big(T, X(T-0), \xi_T\big) \\
&\leqslant \int_t^T L\big(1 + |X(s)|^{\mu+\delta}\big)\mathrm{d}s + L(1 + |X(T-0) + \xi_T|^{\mu+\delta}) + L + \alpha|\xi_T|^\beta\big) \leqslant \bar{\nu}(|x|),
\end{aligned} \tag{3.5}$$

for some continuous increasing function $\bar{\nu} : [0,\infty) \to [0,\infty)$. This proves (3.1), which also leads to $\mathscr{D}(V) = [0,T] \times \mathbb{R}^n$.

Next, let $(t, x) \in \mathscr{D}(V) = [0, T] \times \mathbb{R}^n$, *i.e.*, $V(t, x) < \infty$. Let $\xi(\cdot) \in \mathscr{K}^x[t, T]$ satisfy

$$V(t, x) + 1 \geq J(t, x; \xi(\cdot)),$$

then we have that the optimal impulse control must belong to this smaller class. Moreover,

$$J(t, x; \xi(\cdot)) \geqslant N\ell_0 + \sum_{k=1}^{N} \alpha_0 |\xi_k|^\beta.$$

Then

$$N \leqslant \frac{V(t, x) + 1}{\ell_0} \leqslant \frac{\bar{\nu}(|x|) + 1}{\ell_0},$$

and

$$\sum_{k=1}^{N} |\xi_k|^\beta \leqslant \frac{V(t, x) + 1}{\alpha_0} \leqslant \frac{\bar{\nu}(|x|) + 1}{\alpha_0}.$$

Hence, (3.2)–(3.3) hold.

Now, let $\xi^\varepsilon(\cdot) \in \mathscr{K}_0^x[t, T]$ be a minimizing sequence for the cost functional $\xi(\cdot) \mapsto J(t, x; \xi(\cdot))$. Then, we may assume that

$$\lim_{\varepsilon \to 0} \xi^\varepsilon(s) = \bar{\xi}(s) \equiv \sum_{k \geqslant 1} \bar{\xi}_k \mathbf{1}_{[\bar{\tau}_k, T]}(\cdot), \qquad s \in [t, T],$$

with $t \leqslant \bar{\tau}_1 < \bar{\tau}_2 < \cdots$ and $\bar{\xi}_k \neq 0$. This can be done as follows: Let us begin with the sequence $(\tau_1^\varepsilon, \xi_1^\varepsilon)$. We may assume that

$$\lim_{\varepsilon \to 0} (\tau_1^\varepsilon, \xi_1^\varepsilon) = (\tau_1^0, \xi_1^0).$$

If $\xi_1^0 \neq 0$, we define

$$(\bar{\tau}_1, \bar{\xi}_1) = (\tau_1^0, \xi_1^0).$$

If $\xi_1^0 = 0$, we skip (τ_1^0, ξ_1^0). By taking sub-subsequence, we may assume that

$$\lim_{\varepsilon \to 0} (\tau_2^\varepsilon, \xi_2^\varepsilon) = (\tau_2^0, \xi_2^0).$$

If $\tau_2^0 > \tau_1^0$ and $\xi_2^0 \neq 0$, we define

$$(\bar{\tau}_2, \bar{\xi}_2) = (\tau_2^0, \xi_2^0).$$

Otherwise, if $\xi_2^0 = 0$, we skip $(\tau_2^0, \bar{\xi}_2^0)$; and if $\tau_2^0 = \tau_1^0$, $\xi_2^0 \neq 0$, and $\xi_1^0 + \xi_2^0 = 0$, we skip both (τ_1^0, ξ_1^0) and (τ_2^0, ξ_2^0); if $\tau_2^0 = \tau_1^0$ and $\xi_1^0 + \xi_2^0 \neq 0$, we redefine

$$(\bar{\tau}_1, \bar{\xi}_1) = (\tau_1^0, \xi_1^0 + \xi_2^0).$$

Clearly, with such a procedure, we can complete the construction of $\bar{\xi}(\cdot)$. By the convergence $\xi^{\varepsilon}(\cdot) \to \bar{\xi}(\cdot)$ pointwise, we see that actually the convergence is uniform. Then one also has the uniform convergence of $X^{\varepsilon}(\cdot) \equiv X(\cdot\,; t, x, \xi^{\varepsilon}(\cdot))$ to $\bar{X}(\cdot) \equiv X(\cdot\,; t, x, \bar{\xi}(\cdot))$. It is ready to see that $(\bar{X}(\cdot), \bar{\xi}(\cdot))$ is an optimal pair.

Finally, if (H1$'$) holds, by Lemma 2.6, we have from (3.5) that

$$0 \leqslant V(t,x) \leqslant C(1 + |x|^{\mu+\delta}) + \alpha C_0(1 + |X(T-0)|^{\beta}) \leqslant C(1 + |x|^{\mu+\delta}) + C_0 e^{\beta L(T-t)} \alpha |x|^{\beta},$$

proving (3.4). $\qquad\square$

We see that the appearance of the term $C_0 e^{\beta L(T-t)} \alpha |x|^{\beta}$ is due to the terminal state constraint. It is possible that $\beta > \mu + \delta$. In this case, the value function might grow with the same order as the impulse cost.

3.2. Continuity of the value function

In this subsection, we will establish the continuity of the value functions $V(\cdot, \cdot)$. Note that unlike the classical situation, when the terminal state constraint is presented, the value functions could be discontinuous. Also, some proper conditions will ensure the continuity of the value functions. To be convincing, let us first look at a simple example.

Example 3.2. Consider state equation (which is the same as that in Exam. 2.4)

$$X(s) = x + (s - t) + \xi(s), \qquad s \in [t, T]. \tag{3.6}$$

The cost functional is defined by

$$J(t, x; \xi(\cdot)) = \sum_{k \geqslant 1} \ell(\tau_k, \widetilde{X}(\tau_k - 0), \xi_k), \tag{3.7}$$

with

$$\ell(t, x, \xi) = 1 + |\xi|. \tag{3.8}$$

Suppose $K = \mathbb{R}$, $D = (0, 1)$. Let us consider Problem (C). For any $(t, x) \in [0, T] \times \mathbb{R}$ with $x + T - t \in [0, 1]$, we take $\xi(\cdot) = \xi_0(\cdot)$, the trivial impulse control. Hence,

$$V(t, x) = 0, \qquad \text{if} \ \ x + T - t \in [0, 1].$$

Now, if

$$x + T - t > 1,$$

then, under $\xi_0(\cdot)$, we have

$$X(T; t, x, \xi_0(\cdot)) = x + T - t > 1.$$

Hence, during $[t, T]$ an impulse has to be made. The most economical impulse will be

$$\xi(\cdot) = -(x + T - t - 1)\mathbf{1}_{[\tau_1, T]}(\cdot),$$

where the choice $\tau_1 \in [t, T]$ is irrelevant. Under such an impulse control, we have

$$J(t, x; \xi(\cdot)) = 1 + |x + T - t - 1| = x + T - t.$$

Apparently, such an impulse control is optimal. Finally, if

$$x + T - t < 0,$$

then we take

$$\xi(\cdot) = -(x + T - t)\mathbf{1}_{[\tau_1, T]}(\cdot),$$

with an arbitrary $\tau_1 \in [t, T]$. Again, this impulse control is optimal. With such a control, one has

$$J(t, x; \xi(\cdot)) = 1 + |x + T - t| = 1 - x - T + t.$$

Consequently,

$$V(t, x) = \begin{cases} 1 - x - T + t, & x + T - t < 0, \\ 0, & x + T - t \in [0, 1], \\ x + T - t, & x + T - t > 1. \end{cases} \tag{3.9}$$

Clearly, this value function $V(\cdot, \cdot)$ is discontinuous (along the lines $x + T - t = 0$ and $x + T - t = 1$).

Now, we modify the cost functional as follows:

$$J(t, x; \xi(\cdot)) = h(X(T)) + \sum_{k \geqslant 1} \ell(\tau_k, \widetilde{X}(\tau_k - 0), \xi_k), \tag{3.10}$$

with

$$h(x) = 9\left(x - \frac{2}{5}\right)^2, \qquad x \in \mathbb{R}. \tag{3.11}$$

For any $X \in \mathbb{R}$ (a possible terminal state location), take $\xi \in K \equiv \mathbb{R}$ and look at the following:

$$h(X + \xi) + \ell(T, X, \xi) = 9\left(X + \xi - \frac{2}{5}\right)^2 + 1 + |\xi|,$$

requiring $b \equiv X + \xi \in [0, 1]$. This is the cost at the terminal time T if the terminal state is X and an impulse ξ is made at T. Hence, let us consider the following function

$$F(b, X) = h(b) + \ell(T, X, b - X) = 9\left(b - \frac{2}{5}\right)^2 + 1 + |b - X|, \qquad b \in [0, 1],$$

which will help us to decide whether we should make an impulse at T. For any given $X \in \mathbb{R}$, we want to find the minimum of $b \mapsto F(b, X)$ over $b \in [0, 1]$. To this end, we first observe that

$$F_b(b, X) = \begin{cases} 18\left(b - \frac{2}{5}\right) + 1 = 18b - \frac{31}{5}, & \text{if } b > X, \\ 18\left(b - \frac{2}{5}\right) - 1 = 18b - \frac{41}{5}, & \text{if } b < X, \end{cases}$$

Hence,

$$
\begin{cases}
F_b\left(\dfrac{31}{90}, X\right) = 0, & X < \dfrac{31}{90} \equiv b_0, \\[3mm]
F_b\left(\dfrac{41}{90}, X\right) = 0, & X > \dfrac{41}{90} \equiv b_1.
\end{cases}
$$

Clearly,

$$
0 < b_0 \equiv \frac{31}{90} < \frac{2}{5} < b_1 \equiv \frac{41}{90} < 1.
$$

Further, for $X \in (b_0, b_1)$, we have

$$
\begin{cases}
F_b(b, X) < 0, & b \in (b_0, X), \\
F_b(b, X) > 0, & b \in (X, b_1).
\end{cases}
$$

Hence, for $X \in (b_0, b_1)$,

$$
\min_{b \in [0,1]} F(b, X) = F(X, X) = 9\left(X - \frac{2}{5}\right)^2 + 1.
$$

To summarize, we have

$$
\min_{b \in [0,1]} F(b, X) =
\begin{cases}
F(b_0, X) = 9\left(\dfrac{31}{90} - \dfrac{2}{5}\right)^2 + 1 + \dfrac{31}{90} - X = \dfrac{247}{180} - X, & X < b_0, \\[4mm]
F(X, X) = 9\left(X - \dfrac{2}{5}\right)^2 + 1, & X \in [b_0, b_1], \\[4mm]
F(b_1, X) = 9\left(\dfrac{41}{90} - \dfrac{2}{5}\right)^2 + 1 + X - \dfrac{41}{90} = \dfrac{103}{180} + X, & X > b_1,
\end{cases}
$$

and $X \mapsto \min\limits_{b \in [0,1]} F(b, X)$ is continuous. Note that

$$
h(0) = \frac{36}{25} > \frac{247}{180} = \min_{b \in [0,1]} F(b, 0) = \min_{\xi \in [0,1]} \left[h(\xi) + \ell(T, 0, \xi)\right], \tag{3.12}
$$

and

$$
h(1) = \frac{81}{25} > \frac{283}{180} = \min_{b \in [0,1]} F(b, 1) = \min_{\xi \in [0,1]} \left[h(1 + \xi) + \ell(T, 1, \xi)\right], \tag{3.13}
$$

Now, we look at the equation

$$
h(X) = \min_{b \in [0,1]} F(b, X),
$$

which give the point X at which there is no difference if the best impulse is made or no impulse is made. A direct check shows that the above does not have solutions in $[b_0, b_1]$. Now, on $(0, b_0)$, we solve

$$
9\left(X - \frac{2}{5}\right)^2 = \frac{247}{180} - X,
$$

whose unique solution is

$$X = \frac{1}{90} \equiv a_0.$$

On $(b_1, 1)$, we solve

$$9\left(X - \frac{2}{5}\right)^2 = \frac{103}{180} + X,$$

whose unique solution is

$$X = \frac{71}{90} \equiv a_1.$$

The above tells us that (recalling $b = X + \xi$)

$$\begin{cases} h(X) < \min\limits_{X+\xi \in [0,1]} \left[h(X+\xi) + \ell(T, X, \xi) \right], & X \in (a_0, a_1), \\ h(X) \geqslant \min\limits_{X+\xi \in [0,1]} \left[h(X+\xi) + \ell(T, X, \xi) \right], & X \in \mathbb{R} \setminus (a_0, a_1). \end{cases}$$

We see that

$$0 < a_0 \equiv \frac{1}{90} < b_0 \equiv \frac{31}{90} < b_1 \equiv \frac{41}{90} < a_1 \equiv \frac{71}{90} < 1.$$

The above means that if the terminal state $X(T-0) \in (a_0, a_1)$, we should not make an impulse at T, and if $X(T-0) \in \mathbb{R} \setminus (a_0, a_1)$, we should make an impulse as follows:

$$\xi_1 = \begin{cases} \dfrac{31}{90} - X(T-0), & X(T-0) < a_0 = \dfrac{1}{90}, \text{ or } x + T - t < a_0, \\ \dfrac{41}{90} - X(T-0), & X(T-0) > a_1 = \dfrac{71}{90}, \text{ or } x + T - t > a_1, \end{cases} \tag{3.14}$$

so that either $X(T) = b_0 \equiv \frac{31}{90}$ or $X(T) = b_1 \equiv \frac{41}{90}$. Combining the above analysis, we obtain the value function

$$V(t,x) = \begin{cases} 9\left(x + T - t - \dfrac{2}{5}\right)^2, & x + T - t \in \left[\dfrac{1}{90}, \dfrac{71}{90}\right] \equiv [a_0, a_1], \\ \dfrac{247}{180} - (x + T - t), & x + T - t < \dfrac{1}{90} \equiv a_0, \\ \dfrac{103}{180} + x + T - t, & x + T - t > \dfrac{71}{90} \equiv a_1, \end{cases} \tag{3.15}$$

which is continuous.

Now, let $K = [0, \infty)$ and $D = (0, 1)$. Then from Example 2.4, we see that

$$\mathscr{D}(V) = \left\{ (t, x) \in [0, T] \times \mathbb{R} \mid x + T - t \leqslant 1 \right\},$$

and only positive impulses can be made. Hence, by looking above computation, we see that if $X(T - 0; t, x, \xi_0(\cdot)) < a_0$, we could make an impulse; for all other cases, we could not/should not make impulses.

Therefore,

$$V(t,x) = \begin{cases} \dfrac{247}{180} - (x + T - t), & x + T - t < \dfrac{1}{90}, \\[2mm] 9\Big(x + T - t - \dfrac{2}{5}\Big)^2, & x + T - t \in \Big[\dfrac{1}{90}, 1\Big], \\[2mm] +\infty, & x + T - t > 1. \end{cases} \tag{3.16}$$

This value function is continuous over $\mathscr{D}(V) = \overline{\mathscr{D}(V)}$ which is a closed set.

Further, let $K = [0, \infty)$ and $D = (0, \infty)$. The feature is that the state $X(T-0)$ will be either in $\bar{D}$, or, it can always be pulled back to $\bar{D}$ by an admissible impulse. Therefore, $\mathscr{D}(V) = [0, T] \times \mathbb{R}$. When $X(T-0) < a_0 = \frac{1}{90}$, an impulse is necessary to meet the terminal state constraint, or to make the total cost smaller. Hence,

$$V(t,x) = \begin{cases} \dfrac{247}{180} - (x + T - t), & x + T - t < \dfrac{1}{90}, \\[2mm] 9\Big(x + T - t - \dfrac{2}{5}\Big)^2, & x + T - t \geqslant \dfrac{1}{90}. \end{cases} \tag{3.17}$$

This function is continuous as well.

The above example shows that when the terminal cost function $h(\cdot)$ and the impulse cost are compatible, one could get the continuity of the value function $V(\cdot, \cdot)$. In the above example, a careful observation shows that when the terminal state gets close to the boundary ∂D of the constraint set D from inside, an impulse should be made to reduce the cost. This essentially eliminates the possible jumps of the best costs between the terminal state $X(T-0)$ being close to the boundary ∂D from outside and from inside of D. On the other hand, due to the terminal constraint, the value $h(x)$ of $h(\cdot)$ for $x \in \mathbb{R}^n \setminus \bar{D}$ is irrelevant to our problem. We now would like to present general results.

Theorem 3.3. *Let* (H1)–(H4) *hold and* (2.12) *be satisfied. Suppose*

$$\inf_{\xi \in K, x + \xi \in D} \Big[h(x + \xi) + \ell(T, x, \xi) \Big] < h(x), \qquad \forall x \in \mathbb{R}^n \setminus D. \tag{3.18}$$

Then there exists a continuous increasing function $\widehat{\nu} : [0, \infty) \to [0, \infty)$ *such that*

$$\begin{cases} |V(t,x) - V(t,\widehat{x})| \leqslant C(1 + |x|^\mu \vee |\widehat{x}|^\mu)|x - \widehat{x}|^\delta, & \forall x, \widehat{x} \in \mathbb{R}^n, \ |x - \widehat{x}| \ small, \\[2mm] |V(t,x) - V(\widehat{t},x)| \leqslant \widehat{\nu}(|x| \vee |\widehat{x}|)|t - \widehat{t}| & \forall t, \widehat{t} \in [0, T]. \end{cases} \tag{3.19}$$

In the case that (H1′) *holds,*

$$|V(t,x) - V(\widehat{t},\widehat{x})| \leqslant C\big[1 + (|x| \vee |\widehat{x}|)^\mu\big]|x - \widehat{x}|^\delta + C\big[1 + (|x| \vee |\widehat{x}|)^{(\mu+\delta)\vee\beta}\big]|t - \widehat{t}|,$$
$$\forall (t,x), (\widehat{t},\widehat{x}) \in [0, T] \times \mathbb{R}^n \ |x - \widehat{x}| \ small. \tag{3.20}$$

Proof. Let $(t, x) \in [0, T] \times \mathbb{R}^n$. From Proposition 3.1, there exists an optimal impulse control $\bar{\xi}(\cdot) \in \mathscr{K}_0^x[t, T]$. Due to condition (3.18), we claim that $\bar{X}(T) \in D$. In fact, if $\bar{X}(T) \in \partial D$, then there exists a $\zeta \in K$ such that

$$\bar{X}(T) + \zeta \in D, \qquad h(\bar{X}(T) + \zeta) + \ell(T, \bar{X}(T), \zeta) < h(\bar{X}(T)).$$

Thus, by letting

$$\widehat{\zeta}(\cdot) = \bar{\xi}(\cdot) + \zeta \mathbf{1}_{\{T\}},$$

we have

$$J(t,x;\widehat{\zeta}(\cdot)) = J(t,x;\bar{\xi}(\cdot)) + h(\bar{X}(T) + \zeta) + \ell(T, \bar{X}(T), \zeta) - h(\bar{X}(T)) < J(t,x;\bar{\xi}(\cdot)) = V(t,x),$$

contradicting the optimality of $\bar{\xi}(\cdot)$. Hence, we may assume that $\bar{X}(T) \in D$. Now, for any $\widehat{x} \in \mathbb{R}^n$, let $\widehat{X}(\cdot) = X(\cdot\,; t, \widehat{x}, \bar{\xi}(\cdot))$, we have

$$|\bar{X}(s) - \widehat{X}(s)| \leqslant e^{L(T-t)}|x - \widehat{x}|, \qquad t \leqslant s \leqslant T.$$

Recalling that D is open, for $|x - \widehat{x}|$ small, one sees that $\bar{\xi}(\cdot) \in \mathscr{K}^{\widehat{x}}[t,T]$. Consequently, making use of Propositions 2.1 and 3.1, together with the Lipschitz continuity of $x \mapsto \ell(t,x,\xi)$, we have (noting $\bar{\xi}(\cdot) \in \mathscr{K}_0^x[t,T]$)

$$
\begin{aligned}
V(t,\widehat{x}) \leqslant J(t,\widehat{x};\bar{\xi}(\cdot)) &\leqslant J(t,x;\bar{\xi}(\cdot)) + L\int_t^T \Big(1 + |\bar{X}(r)|^\mu \vee |\widehat{X}(r)|^\mu\Big)|\bar{X}(r) - \widehat{X}(r)|^\delta \mathrm{d}r \\
&\quad + L\Big(1 + |\bar{X}(T)|^\mu \vee |\widehat{X}(T)|^\mu\Big)|\bar{X}(T) - \widehat{X}(T)|^\delta \\
&\quad + \sum_{k \geqslant 1}|\ell(\bar{\tau}_k, \bar{X}(\bar{\tau}_k - 0), \bar{\xi}_k) - \ell(\bar{\tau}_k, \widehat{X}(\bar{\tau}_k - 0), \bar{\xi}_k)| \\
&\leqslant V(t,x) + C(1 + |x|^\mu \vee |\widehat{x}|^\mu)|x - \widehat{x}|^\delta.
\end{aligned}
$$

for some constant $C > 0$. By symmetry, we obtain the first estimate in (3.19).

Next, let $0 \leqslant t < \widehat{t} \leqslant T$. Let $\bar{\xi}(\cdot) \in \mathscr{K}_0^x[t,T]$ be optimal for the initial pair (t,x). Let

$$t \leqslant \bar{\tau}_1 < \bar{\tau}_2 < \cdots < \bar{\tau}_{k_0} \leqslant \widehat{t} < \bar{\tau}_{k_0+1}.$$

Define

$$\widehat{\xi}(\cdot) = \sum_{k=1}^{k_0} \bar{\xi}_k \mathbf{1}_{[\widehat{t},T]}(\cdot) + \sum_{k \geqslant k_0+1} \bar{\xi}_k \mathbf{1}_{[\bar{\tau}_k,T]}(\cdot).$$

Denote $\widehat{X}(\cdot) = X(\cdot\,; \widehat{t}, x, \widehat{\xi}(\cdot))$. Then

$$
\begin{aligned}
J(\widehat{t},x;\widehat{\xi}(\cdot)) &= \int_{\widehat{t}}^T g(r,\widehat{X}(r))\mathrm{d}r + h(\widehat{X}(T)) + \ell\Big(\widehat{t},x;\sum_{\bar{\tau}_k \leqslant \widehat{t}} \bar{\xi}_k\Big) + \sum_{\tau_k > \widehat{t}} \ell(\bar{\tau}_k, \widehat{X}(\bar{\tau}_k - 0), \bar{\xi}_k) \\
&\leqslant J(t,x;\bar{\xi}(\cdot)) + \int_t^{\widehat{t}} |g(r,X(r))|\mathrm{d}r + \int_{\widehat{t}}^T |g(r,\widehat{X}(r)) - g(r,X(r))|\mathrm{d}r + |h(\widehat{X}(T)) - h(X(T))| \\
&\quad + \Big[\ell\Big(\widehat{t},x;\sum_{k=1}^{k_0} \bar{\xi}_k\Big) - \sum_{k=1}^{k_0} \ell(\bar{\tau}_k, X(\bar{\tau}_k - 0), \bar{\xi}_k)\Big] + \sum_{k > k_0}\Big[\ell(\bar{\tau}_k, \widehat{X}(\tau_k - 0), \bar{\xi}_k) - \ell(\bar{\tau}_k, X(\bar{\tau}_k - 0), \bar{\xi}_k)\Big].
\end{aligned}
$$

Note that

$$\ell\Big(\widehat{t}, x; \sum_{k=1}^{k_0} \bar{\xi}_k\Big) \leqslant \sum_{k=1}^{k_0} \ell\Big(\widehat{t}, x + \sum_{i=1}^{k-1} \bar{\xi}_i, \bar{\xi}_k\Big) \leqslant \sum_{k=1}^{k_0} \ell\Big(\bar{\tau}_k, x + \sum_{i=1}^{k-1} \bar{\xi}_i, \bar{\xi}_k\Big).$$

On the other hand,

$$|X(\tau_1 - 0) - x| \leqslant \int_t^{\tau_1} |f(r, X(r))| \mathrm{d}r \leqslant L \int_t^{\tau_1} (1 + |X(r)|) \mathrm{d}r$$
$$\leqslant L \int_t^{\tau_1} \Big(1 + e^{L(r-t)}(1 + |x|)\Big) \mathrm{d}r \leqslant C(1 + |x|)(\tau_1 - t).$$

Next,

$$|X(\bar{\tau}_2 - 0) - x - \xi_1| \leqslant \int_{\bar{\tau}_1}^{\bar{\tau}_2} |f(r, X(r))| \mathrm{d}r \leqslant L \int_{\bar{\tau}_1}^{\bar{\tau}_2} \big(1 + |X(r)|\big) \mathrm{d}r$$
$$\leqslant L \int_{\bar{\tau}_1}^{\bar{\tau}_2} \Big(1 + e^{L(r-t)}(1 + |x|) + e^{L(r-\bar{\tau}_1)}|\bar{\xi}_1|\Big) \mathrm{d}r \leqslant C(1 + |x| + |\bar{\xi}_1|)(\bar{\tau}_2 - \bar{\tau}_1).$$

By induction, we see that

$$\Big|X(\bar{\tau}_k - 0) - x - \sum_{i=1}^{k-1} \bar{\xi}_i\Big| \leqslant C\Big(1 + |x| + \sum_{i=1}^{k-1} |\bar{\xi}_i|\Big)(\bar{\tau}_k - t), \qquad 1 \leqslant k \leqslant k_0.$$

Also, for any $s > \widehat{t}$,

$$|\widehat{X}(s) - X(s)| \leqslant e^{L(s-\widehat{t})}\Big|x + \sum_{i=1}^{k_0} \bar{\xi}_k - X(\widehat{t})\Big| \leqslant C\Big(1 + |x| + \sum_{i=1}^{k_0} |\bar{\xi}_i|\Big)(\widehat{t} - t).$$

Consequently, noting that $\bar{\xi}(\cdot) \in \mathscr{K}_0^x[t, T]$,

$$J(\widehat{t}, x; \widehat{\bar{\xi}}(\cdot)) \leqslant J(t, x; \bar{\xi}(\cdot)) + \int_t^{\widehat{t}} L(1 + |X(r)|^\delta) \mathrm{d}r + L \int_{\widehat{t}}^T |X(r) - \widehat{X}(r)| \mathrm{d}r + L|X(T) - \widehat{X}(T)|$$
$$+ L \sum_{k=1}^{k_0} \Big|x + \sum_{i=1}^{k-1} \bar{\xi}_i - X(\bar{\tau}_k - 0)\Big| + L \sum_{k > k_0} |\widehat{X}(\bar{\tau}_k - 0) - X(\bar{\tau}_k - 0)|$$
$$\leqslant V(t, x) + C\Big(1 + |x| + \sum_{k \geqslant 1} |\bar{\xi}_k|\Big)(\widehat{t} - t) \leqslant V(t, x) + \widetilde{\nu}(|x|)(\widehat{t} - t),$$

for some $\widetilde{\nu}(|x|)$. Finally, let $\widehat{t} < t$. Then we extend $\bar{\xi}(\cdot)$ on $[t, T]$ to $\widehat{\xi}(\cdot)$ on $[\widehat{t}, T]$ trivially. One has

$$
\begin{aligned}
V(\widehat{t}, x) &\leqslant J(\widehat{t}, x; \widehat{\xi}(\cdot)) \leqslant J(t, x; \bar{\xi}(\cdot)) + \int_{\widehat{t}}^{t} |g(r, \widehat{X}(r))| \mathrm{d}r + \int_{t}^{T} |g(r, \widehat{X}(r)) - g(r, X(r))| \mathrm{d}r \\
&\quad + |h(\widehat{X}(T)) - h(X(r))| + \sum_{k \geqslant 1} |\ell(\bar{\tau}_k, \widehat{X}(\bar{\tau}_k - 0), \xi_k^{\varepsilon}) - \ell(\bar{\tau}_k, X(\bar{\tau}_k - 0), \bar{\xi}_k^{\varepsilon})| \\
&\leqslant V(t, x) + \int_{\widehat{t}}^{t} L(1 + |\widehat{X}(r)|) \mathrm{d}r + L \int_{t}^{T} |X(r) - \widehat{X}(r)| \mathrm{d}r \\
&\quad + L|X(T) - \widehat{X}(T)| + \sum_{k \geqslant 1} |X(\bar{\tau}_k - 0) - \widehat{X}(\bar{\tau}_k - 0)| \\
&\leqslant V(t, x) + C\Big(1 + |x| + \sum_{k \geqslant 1} |\bar{\xi}_k|\Big)(t - \widehat{t}) \leqslant \widetilde{\nu}(|x|)(t - \widehat{t}).
\end{aligned}
$$

This completes the proof of (3.19). Finally, in the case that (H1$'$) holds, our conclusion follows from the above arguments, together with Proposition 3.1. $\qquad\square$

We see that due to the appearance of the terminal constraint, the Hölder continuity of the map $x \mapsto V(t, x)$ is only locally.

4. Dynamic programming principle and HJB quasi-variational inequality

In this section, we are going to establish Bellman's principle of optimality for our Problem (C). Then the value function $V(\cdot, \cdot)$ is showed to be a viscosity solution of the corresponding HJB equation, which is a quasi-variational inequality. For convenience, in what follows, we will keep assumptions (H1$'$), (H2)–(H4) and (3.18).

Theorem 4.1. *Let $V(\cdot, \cdot)$ be the value function of Problem (C). Then for any $(t, x) \in [0, T) \times \mathbb{R}^n$, the following principle of optimality holds:*

$$
V(t, x) \leqslant \inf_{\xi \in K} \big\{ V(t, x + \xi) + \ell(t, x, \xi) \big\} \equiv \mathbf{N}[V](t, x), \qquad \forall (t, x) \in [0, T) \times \mathbb{R}^n, \tag{4.1}
$$

$$
V(t, x) \leqslant \int_{t}^{\widehat{t}} g(r, X(r; t, x, \xi_0(\cdot))) \mathrm{d}r + V(\widehat{t}, X(\widehat{t}; t, x, \xi_0(\cdot))), \quad \forall 0 \leqslant t \leqslant \widehat{t} \leqslant T, \ x \in \mathbb{R}^n. \tag{4.2}
$$

Furthermore, if the strict inequality holds in (4.1), then there exists a $\bar{t} \in (t, T]$ such that

$$
V(t, x) = \int_{t}^{\widehat{t}} g(r, X(r; t, x, \xi_0(\cdot))) \mathrm{d}r + V(\widehat{t}, X(\widehat{t}; t, x, \xi_0(\cdot))), \qquad 0 \leqslant t \leqslant \widehat{t} < \bar{t} \leqslant T, \ x \in \mathbb{R}^n.
$$

For $x \in \mathbb{R}^n$, it holds that

$$
V(T, x) = h^D(x) \equiv \begin{cases} \min \big\{ h(x), \mathbf{N}^D[h](x) \big\}, & x \in \bar{D}, \\ \mathbf{N}^D[h](x), & x \in \mathbb{R}^n \setminus \bar{D}, \end{cases} \tag{4.3}
$$

where

$$\mathbf{N}^D[h](x) = \inf_{\xi \in K, x+\xi \in \bar{D}} \big\{ h(x+\xi) + \ell(T, x, \xi) \big\}, \qquad x \in \mathbb{R}^n. \tag{4.4}$$

Proof. First of all, for $x \in \mathbb{R}^n$, we clearly have (4.3). Next, let $(t, x) \in [0, T) \times \mathbb{R}^n$ and $\zeta \in K$. For any $\xi(\cdot) \in \mathscr{K}^{x+\zeta}[t, T]$, we see that

$$\widehat{\xi}(\cdot) = \zeta \mathbf{1}_{[t,T]}(\cdot) + \xi(\cdot) \in \mathscr{K}^x[t, T].$$

Thus,

$$V(t, x) \leqslant J(t, x; \widehat{\xi}(\cdot)) = \ell(t, x, \zeta) + J(t, x+\zeta; \xi(\cdot)).$$

Consequently,

$$V(t, x) \leqslant \ell(t, x, \zeta) + V(t, x+\zeta), \qquad \forall \zeta \in K.$$

Therefore,

$$V(t, x) \leqslant \inf_{\zeta \in K} \{ V(t, x+\zeta) + \ell(t, x, \zeta) \} = \mathbf{N}[V](t, x).$$

On the other hand, for any $0 \leqslant t < \widehat{t} \leqslant T$, we take any $\xi(\cdot) \in \mathscr{K}^{X(\widehat{t}; t, x, \xi_0(\cdot))}[\widehat{t}, T]$. Extend it to $\widehat{\xi}(\cdot) \in \mathscr{K}^x[t, T]$ in the way that no impulses are made on $[t, \widehat{t})$. Then

$$V(t, x) \leqslant J(t, x; \widehat{\xi}(\cdot)) = \int_t^{\widehat{t}} g(r, X_{t,x}(r))\mathrm{d}r + J(\widehat{t}, X(\widehat{t}; t, x, \xi_0(\cdot)); \xi(\cdot)).$$

Consequently,

$$V(t, x) \leqslant \int_t^{\widehat{t}} g(r, X(r; t, x, \xi_0(\cdot)))\mathrm{d}r + V(\widehat{t}, X(\widehat{t}; t, x, \xi_0(\cdot))).$$

Finally, we assume that

$$V(t, x) < \mathbf{N}[V](t, x) \equiv \inf_{\zeta \in K} \big\{ V(t, x+\zeta) + \ell(t, x, \zeta) \big\}. \tag{4.5}$$

For any $\varepsilon > 0$, there exists an impulse control $\xi^\varepsilon(\cdot) = \sum_{i \geqslant 1} \xi_i^\varepsilon \mathbf{1}_{[\tau_i^\varepsilon, T]}(s) \in \mathscr{K}^x[t, T]$ such that

$$J(t, x; \xi^\varepsilon(\cdot)) \leqslant V(t, x) + \varepsilon.$$

If $\tau_1^\varepsilon = t$, then

$$V(t, x) + \varepsilon \geqslant J(t, x+\xi_1^\varepsilon; \widehat{\xi}^\varepsilon(\cdot)) + \ell(t, x, \xi_1^\varepsilon) \geqslant V(t, x+\xi_1^\varepsilon) + \ell(t, x, \xi_1^\varepsilon), \tag{4.6}$$

where

$$\widehat{\xi}^{\varepsilon}(\cdot) = \sum_{i \geqslant 1} \widehat{\xi}_i^{\varepsilon} \mathbf{1}_{[\widehat{\tau}_i^{\varepsilon}, T]}(s); \qquad \widehat{\xi}_i^{\varepsilon} = \xi_{i+1}^{\varepsilon}, \quad \widehat{\tau}_i^{\varepsilon} = \tau_{i+1}^{\varepsilon}, \quad i \geqslant 1.$$

This is contradicting (4.5). Hence, $\tau_1^{\varepsilon} > t$. We further claim that there exists a $\bar{t} > t$ such that $\tau_1^{\varepsilon} \geqslant \bar{t}$ for all $\varepsilon > 0$ small. If this is not the case, then for some $\varepsilon \downarrow 0$, $\tau_i^{\varepsilon} \downarrow t$. Thus,

$$\begin{aligned}
V(t,x) + \varepsilon &\geqslant J(t,x;\xi^{\varepsilon}(\cdot)) \\
&= \int_t^{\tau_1^{\varepsilon}} g(r, X(s;t,x,\xi_0(\cdot)))\mathrm{d}s + \ell(\tau_1^{\varepsilon}, X(\tau_1^{\varepsilon}-0;t,x,\xi_0(\cdot)), \xi_1^{\varepsilon}) + J(\tau_1^{\varepsilon}, X(\tau_1^{\varepsilon}-0;t,x,\xi_0(\cdot)) + \xi_1^{\varepsilon}; \widehat{\xi}^{\varepsilon}(\cdot)) \\
&\geqslant \int_t^{\tau_1^{\varepsilon}} g(r, X(s;t,x,\xi_0(\cdot)))\mathrm{d}s + \ell(\tau_1^{\varepsilon}, X(\tau_1^{\varepsilon}-0;t,x\xi_0(\cdot)), \xi_1^{\varepsilon}) + V(\tau_1^{\varepsilon}, X(\tau_1^{\varepsilon}-0;t,x,\xi_0(\cdot))).
\end{aligned} \tag{4.7}$$

Letting $\varepsilon \downarrow 0$ and (we may assume that) $\xi_1^{\varepsilon} \to \xi_1$, we have

$$V(t,x) \geqslant V(t, x + \xi_1) + \ell(t, x, \xi_1).$$

which is a contradiction again. Therefore, we get the existence of $\bar{t} > t$. Then for any $\widehat{t} < \bar{t}$,

$$\begin{aligned}
V(t,x) + \varepsilon &\geqslant \int_t^{\widehat{t}} g(r, X(r;t,x,\xi_0(\cdot)))\mathrm{d}r + J(\widehat{t}, X(\widehat{t};t,x,\widehat{\xi}^{\varepsilon}) \\
&\geqslant \int_t^{\widehat{t}} g(r, X(r;t,x,\xi_0(\cdot)))\mathrm{d}s + V(\widehat{t}, X(\widehat{t};t,x,\xi_0(\cdot))).
\end{aligned} \tag{4.8}$$

Combing this with (4.2), the proof is complete. $\qquad\square$

The following proposition is about the properties of $\mathbf{N}[\cdot]$ and $\mathbf{N}^D[\cdot]$.

Proposition 4.2. *The maps $(t,x) \mapsto \mathbf{N}[V](t,x)$ and $x \mapsto \mathbf{N}^D[h](x)$ are continuous. Moreover, if $(t_0, x_0) \in [0,T) \times \mathbb{R}^n$ such that*

$$V(t_0, x_0) = \mathbf{N}[V](t_0, x_0) = V(t_0, x_0 + \xi_0) + \ell(t_0, x_0, \xi_0), \tag{4.9}$$

for some $\xi_0 \in K$, then

$$\mathbf{N}[V](t_0, x_0 + \xi_0) - V(t_0, x_0 + \xi_0) \geqslant \delta_0. \tag{4.10}$$

Similarly, if $x_0 \in \mathbb{R}^n$ such that

$$h(x_0) = \mathbf{N}^D[h](x_0) = h(x_0 + \xi_0) + \ell(T, x_0, \xi_0), \tag{4.11}$$

for some $\xi_0 \in K$, then

$$\mathbf{N}^D[h](x_0 + \xi_0) - h(x_0 + \xi_0) \geqslant \delta_0. \tag{4.12}$$

Proof. First of all, suggested by (3.2)–(3.3), for any $x \in \mathbb{R}^n$, we may introduce

$$K_0^{|x|} = \left\{ \xi \in K \mid |\xi| \leqslant \left(\frac{\bar{\nu}(|x|) + 1}{\alpha_0} \right)^{\frac{1}{\beta}} \right\}. \tag{4.13}$$

Then (3.2)–(3.3) tells us that for given initial pair $(t, x) \in [0, T] \times \mathbb{R}^n$, we may restrict ourselves to the impulse controls with the impulse vector taken from $K_0^{|x|}$. Now, for any $(t, x), (\widehat{t}, \widehat{x}) \in [0, T) \times \mathbb{R}^n$, and $\xi \in K_0^{|x| \vee |\widehat{x}|}$,

$$\left| V(t, x + \xi) + \ell(t, x, \xi) - [V(\widehat{t}, \widehat{x} + \xi) + \ell(\widehat{t}, \widehat{x}, \xi)] \right| \leqslant \left[\widehat{\nu}(|x| \vee |\widehat{x}|) + L \right] \left(|t - \widehat{t}| + |x - \widehat{x}| \right).$$

Note that in the definition of $\mathbf{N}[\,\cdot\,]$, $\xi \in K$ is uniform in (t, x). Thus, we have

$$|\mathbf{N}[V](t, x) - \mathbf{N}[V](\widehat{t}, \widehat{x})| \leqslant \left[\widehat{\nu}(|x| \vee |\widehat{x}|) + L \right] \left(|t - \widehat{t}| + |x - \widehat{x}| \right). \tag{4.14}$$

Next, if for some $(t_0, x_0) \in [0, T) \times \mathbb{R}^n$, (4.9) holds for some $\xi_0 \in K_0^{|x_0|}$, then we claim that (4.10) holds. In fact, for any $\xi_1 \in K_0^{|x_0 + \xi_0|}$

$$
\begin{aligned}
& V(t_0, x_0 + \xi_0 + \xi_1) + \ell(t_0, x_0 + \xi_0, \xi_1) - V(t_0, x_0 + \xi_0) \\
&= V(t_0, x_0 + \xi_0 + \xi_1) + \ell(t_0, x_0 + \xi_0, \xi_1) + \ell(t_0, x_0, \xi_0) - \Big[V(t_0, x_0 + \xi_0) + \ell(t_0, x_0, \xi_0) \Big] \\
&= V(t_0, x_0 + \xi_0 + \xi_1) + \ell(t_0, x_0, \xi_0 + \xi_1) - V(t_0, x_0) + \Big[\ell(t_0, x_0, \xi_0) + \ell(t_0, x_0 + \xi_0, \xi_1) - \ell(t_0, x_0, \xi_0 + \xi_1) \Big] \\
&\geqslant \ell(t_0, x_0, \xi_0) + \ell(t_0, x_0 + \xi_0, \xi_1) - \ell(t_0, x_0, \xi_0 + \xi_1) \geqslant \delta_0.
\end{aligned}
$$

This proves (4.10).

Now, we look at $\mathbf{N}^D[\,\cdot\,]$. By (4.13), we can redefine (compare with (4.4))

$$\mathbf{N}^D[h](x) = \inf_{\xi \in K_0^x, x + \xi \in D} \left\{ h(x + \xi) + \ell(T, x, \xi) \right\}, \qquad x \in \mathbb{R}^n.$$

We now show that $x \mapsto \mathbf{N}^D[h](x)$ is continuous. For any given $x \in \mathbb{R}^n$, there exists a $\xi \in K_0^x \subset K$ such that

$$\mathbf{N}^D[h](x) = h(x + \xi) + \ell(T, x, \xi), \qquad x + \xi \in \bar{D}.$$

We claim that $x + \xi \in D$ (not on the boundary ∂D of $\bar{D}$). In fact, if $x + \xi \in \partial D$, then by (3.18), there exists a $\xi_1 \in K$ such that

$$
\begin{aligned}
\mathbf{N}^D[h](x) = h(x + \xi) + \ell(T, x, \xi) &> h(x + \xi + \xi_1) + \ell(T, x + \xi, \xi_1) + \ell(T, x, \xi) \\
&> h(x + \xi + \xi_1) + \ell(T, x, \xi + \xi_1) \geqslant \mathbf{N}^D[h](x),
\end{aligned}
$$

a contradiction. Hence, by the openness of D, there exists a $\delta > 0$ such that $\mathscr{O}_\delta(x + \xi) \subseteq D$. Then for $y \in \mathscr{O}_\delta(x)$, $\xi \in D - \{y\}$ and thus

$$
\begin{aligned}
\mathbf{N}^D[h](x) = h(x + \xi) + \ell(T, x, \xi) &\geqslant h(y + \xi) + \ell(T, y, \xi) - \nu(|x| \vee |y|, |x - y|) \\
&\geqslant \mathbf{N}^D[h](y) - \nu(|x| \vee |y|, |x - y|),
\end{aligned}
$$

for some continuous function $\nu : [0, \infty) \times [0, \infty) \to [0, \infty)$ with $\nu(r, 0) = 0$, for any $r \geqslant 0$. Switch the positions of x and y, we obtain the continuity of $x \mapsto \mathbf{N}^D[h](x)$. The proof of last conclusion is similar to the case of $\mathbf{N}[V](\cdot, \cdot)$, with the restriction that $x_0 + \xi_0, x_0 + \xi_0 + \xi_1 \in \bar{D}$ and $t_0 = T$. $\square$

The above result leads to the following Hamilton-Jacobi-Bellman equation for the value function $V(\cdot, \cdot)$. The proof is standard.

Theorem 4.3. *Suppose that the value function $V(\cdot,\cdot) \in C^1([0,T] \times \mathbb{R}^n)$. Then $V(\cdot,\cdot)$ satisfies the following HJB quasi-variational inequality:*

$$\begin{cases} V_t(t,x) + \langle V_x(t,x), f(t,x) \rangle + g(t,x) \geqslant 0, \qquad \mathbf{N}[V](t,x) - V(t,x) \geqslant 0, \quad (t,x) \in [0,T] \times \mathbb{R}^n, \\ \Big(V_t(t,x) + \langle V_x(t,x), f(t,x) \rangle + g(t,x) \Big) \Big(\mathbf{N}[V](t,x) - V(t,x) \Big) = 0, \quad (t,x) \in [0,T] \times \mathbb{R}^n, \\ V(T,x) = h^D(x), \qquad x \in \mathbb{R}^n, \end{cases} \tag{4.15}$$

which can also be written as

$$\begin{cases} \min \Big\{ V_t(t,x) + \langle V_x(t,x), f(t,x) \rangle + g(t,x), \mathbf{N}[V](t,x) - V(t,x) \Big\} = 0, \quad (t,x) \in [0,T) \times \mathbb{R}^n, \\ V(T,x) = h^D(x), \qquad x \in \mathbb{R}^n. \end{cases} \tag{4.16}$$

From the previous sections, we see that under (H1′), (H2)–(H4) and (3.18), the value function $V(\cdot,\cdot) \in C([0,T] \times \mathbb{R}^n)$. However, it is known that the value function might not be $C^1([0,T] \times \mathbb{R}^n)$ in general. Therefore, the above is a formal result. For (4.16), inspired by the viscosity solution notion introduced by Crandall–Lions [11], Barles introduced the following corresponding notion [2, 3], which has been modified here for our impulse control problem in finite time horizon (see [31]).

Definition 4.4. A continuous function $V(\cdot,\cdot)$ is called a *viscosity sub-solution* of HJB quasi-variational inequality (4.16) if

$$V(T,x) \leqslant h^D(x), \qquad x \in \mathbb{R}^n, \tag{4.17}$$

and for any function $\varphi \in C^1([0,T] \times \mathbb{R}^n)$ such that $V(\cdot,\cdot) - \varphi(\cdot,\cdot)$ attains a local maximum at $(t,x) \in [0,T) \times \mathbb{R}^n$, it holds

$$\min \Big\{ \varphi_t(t,x) + \langle \varphi_x(t,x), f(t,x) \rangle + g(t,x), \mathbf{N}[V](t,x) - V(t,x) \Big\} \geqslant 0. \tag{4.18}$$

A continuous function $V(\cdot,\cdot)$ is called a *viscosity super-solution* of HJB quasi-variational inequality (4.16) if

$$V(T,x) \geqslant h^D(x), \qquad x \in \mathbb{R}^n, \tag{4.19}$$

and for any function $\varphi \in C^1([0,T] \times \mathbb{R}^n)$ such that $V(\cdot,\cdot) - \varphi(\cdot,\cdot)$ attains a local minimum at $(t,x) \in [0,T) \times \mathbb{R}^n$, it holds

$$\min \Big\{ \varphi_t(t,x) + \langle \varphi_x(t,x), f(t,x) \rangle + g(t,x), \mathbf{N}[V](t,x) - V(t,x) \Big\} \leqslant 0. \tag{4.20}$$

A continuous function $V(\cdot,\cdot)$ is called a *viscosity solution* of the HJB quasi-variational inequality (4.16) if it is both viscosity super-solution and viscosity sub-solution.

We now state the following result of existence of viscosity solution.

Theorem 4.5. *Let* (H1′), (H2)–(H4) *and* (3.18) *hold, then the value function $V(\cdot,\cdot)$ is a viscosity solution to the HJB quasi-variational inequality* (4.16).

Proof. Obviously, $V(t,x)$ satisfies the boundary condition.

Assume that $\Psi \in C^1([0,T] \times R^n)$ such that $V - \Psi$ takes a local maximum at some point (t_0, x_0). For sufficiently small $|t - t_0|$,

$$V(t_0, x_0) - \Psi(t_0, x_0) \geqslant V(t, x_{t_0, x_0}(t)) - \Psi(t, x_{t_0, x_0}(t)),$$

where $x_{t_0,x_0}(t)$ is the trajectory corresponding to no-impulse case. Hence,

$$0 \leqslant \int_{t_0}^t g(r, x_{t_0,x_0}(r))\mathrm{d}r + V(t, x_{t_0,x_0}(t)) - V(t_0, x_0)$$
$$\leqslant \int_{t_0}^t g(r, x_{t_0,x_0}(r))\mathrm{d}r + \Psi(t, x_{t_0,x_0}(t)) - \Psi(t_0, x_0).$$

Dividing both side by $t - t_0$ and letting $t \to t_0$,

$$0 \leqslant g(t_0, x_0) + \Psi_t(t_0, x_0) + \Psi_x(t_0, x_0)f(t_0, x_0),$$

i.e., $V(t, x)$ is a viscosity subsolution of the HJB quasi-variational inequality.

Next assume that $\Psi \in C^1([0, T] \times R^n)$ such that $V - \Psi$ takes a local minimum at some point (t_0, x_0). For sufficiently small $|t - t_0|$,

$$V(t_0, x_0) - \Psi(t_0, x_0) \leqslant V(t, x_{t_0,x_0}(t)) - \Psi(t, x_{t_0,x_0}(t)),$$

where $x_{t_0,x_0}(t)$ is the trajectory corresponding to no-impulse case. In this case, if $V(t, x) = \inf_{\xi \in K}\{V(t, x + \xi) + \ell(t, x, \xi)\}$. It is obvious that

$$\min\{\Psi_t(t_0, x_0) + \Psi_x(t_0, x_0)f(t_0, x_0) + g(t_0, x_0), \inf_{\xi \in K}\{V(t_0, x_0 + \xi) + \ell(t_0, \xi)\} - V(t_0, x_0)\} \leqslant 0.$$

If $V(t, x) < \inf_{\xi \in K}\{V(t, x + \xi) + \ell(t, \xi)\}$, there exists $\bar{t} \in (t_0, T]$ such that

$$V(t_0, x_0) = \int_{t_0}^t g(r, x_{t_0,x_0}(r))\mathrm{d}r + V(t, x_{t_0,x_0}(t)), \qquad \forall t \in [t_0, \bar{t}].$$

Therefore,

$$0 = \int_{t_0}^t g(r, x_{t_0,x_0}(r))\mathrm{d}r + V(t, x_{t_0,x_0}(t)) - V(t_0, x_0)$$
$$\geqslant \int_{t_0}^t g(r, x_{t_0,x_0}(r))\mathrm{d}r + \Psi(t, x_{t_0,x_0}(t)) - \Psi(t_0, x_0),$$

i.e., $V(t, x)$ is a viscosity supersolution of the HJB quasi-variational inequality. $\square$

We have seen that Problem (C) admits an optimal impulse control (see Prop. 3.1). It is almost standard that, via the value function, an optimal impulse control can also be constructed. We omit the details here.

5. AN OPTIMAL IMPULSE CONTROL PROBLEM WITHOUT TERMINAL STATE CONSTRAINT

HJB QVI (4.16) suggests us to introduce the following modified cost functional:

$$\widetilde{J}(t, x; \xi(\cdot)) = \int_t^T g(s, X(s))\mathrm{d}s + h^D(X(T)) + \sum_{k \geqslant 1} \ell(\tau_k, \widetilde{X}(\tau_k - 0), \xi_k), \tag{5.1}$$

and consider the problem with the above cost functional without terminal state constraint, call it Problem $(\widetilde{\mathrm{C}})$. Let the value function be $\widetilde{V}(\cdot\,,\cdot)$. The question is whether $\widetilde{V}(\cdot\,,\cdot) = V(\cdot\,,\cdot)$? If this is true, then since $\widetilde{V}(\cdot\,,\cdot)$ can be characterized as the unique viscosity solution to (4.16), we indirectly obtain a characterization of the value function $V(\cdot\,,\cdot)$. In this section, we discuss this issue.

We recall that

$$\mathbf{N}^D[h](x) = \min_{\xi \in K, x+\xi \in \bar{D}} \Big(h(x+\xi) + \ell(T, x, \xi) \Big), \qquad x \in \mathbb{R}^n, \tag{5.2}$$

and (for convenience, we denote the value function by $V^D(t, x)$)

$$V^D(T, x) = h^D(x) \equiv \begin{cases} \min\Big\{ h(x), \mathbf{N}^D[h](x) \Big\}, & x \in \bar{D}, \\ \mathbf{N}^D[h](x), & x \in \mathbb{R}^n \setminus \bar{D}. \end{cases} \tag{5.3}$$

Taking $D = \mathbb{R}^n$, the above is reduced to the case of no terminal state constraint. In other words, if there is no terminal state constraint, then the terminal value of the value function is given by the following:

$$V^{\mathbb{R}^n}(T, x) = h^{\mathbb{R}^n}(x) \equiv \min\Big\{ h(x), \mathbf{N}^{\mathbb{R}^n}[h](x) \Big\}, \qquad x \in \mathbb{R}^n, \tag{5.4}$$

with

$$\mathbf{N}^{\mathbb{R}^n}[h](x) = \min_{\xi \in K} \Big(h(x+\xi) + \ell(T, x, \xi) \Big), \qquad x \in \mathbb{R}^n. \tag{5.5}$$

According to the above, we have

$$\widetilde{V}(T, x) = (h^D)^{\mathbb{R}^n}(x) \equiv \min\Big\{ h^D(x), \mathbf{N}^{\mathbb{R}^n}[h^D](x) \Big\}, \qquad x \in \mathbb{R}^n.$$

Let us calculate the following: (remember $D \neq \mathbb{R}^n$ and (2.19))

$$\begin{aligned} \mathbf{N}^{\mathbb{R}^n}\big[\mathbf{N}^D[h]\big](x) &= \min_{\xi \in K} \Big(\mathbf{N}^D[h](x+\xi) + \ell(T, x, \xi) \Big) \\ &= \min_{\xi \in K} \Big[\min_{\zeta \in K, x+\xi+\zeta \in \bar{D}} \Big(h(x+\xi+\zeta) + \ell(T, x+\xi, \zeta) + \ell(T, x, \xi) \Big) \Big] \\ &\geqslant \min_{\xi \in K} \Big[\min_{\zeta \in K, x+\xi+\zeta \in \bar{D}} \Big(h(x+\xi+\zeta) + \ell(T, x, \xi+\zeta) \Big) \Big] + \delta_0 = \mathbf{N}^D[h](x) + \delta_0. \end{aligned}$$

Also,

$$\mathbf{N}^{\mathbb{R}^n}[h](x) = \min_{\xi \in K} \Big(h(x+\xi) + \ell(T, x, \xi) \Big) \leqslant \min_{\xi \in K, x+\xi \in \bar{D}} \Big(h(x+\xi) + \ell(T, x, \xi) \Big) = \mathbf{N}^D[h](x), \qquad x \in \mathbb{R}^n.$$

We claim that under (3.18), the above equality holds. In fact, let

$$\min_{\xi \in K} \Big(h(x+\xi) + \ell(T, x, \xi) \Big) = h(x+\xi_0) + \ell(T, x, \xi_0),$$

for some $\xi_0 \in K$. If $x + \xi_0 \notin \bar{D}$, then by (3.18), the above leads to

$$\min_{\xi \in K} \left(h(x + \xi) + \ell(T, x, \xi) \right) = h(x + \xi_0) + \ell(T, x, \xi_0)$$
$$> h(x + \xi_0) > \min_{\xi \in K, x + \xi \in \bar{D}} \left(h(x + \xi) + \ell(T, x, \xi) \right),$$

which is a contradiction. Hence, our claim holds and for $x \in \mathbb{R}^n \setminus \bar{D}$,

$$\min \left\{ h^D(x), \mathbf{N}^{\mathbb{R}^n}[h^D](x) \right\} = \min \left\{ \mathbf{N}^D[h](x), \mathbf{N}^{\mathbb{R}^n} \left[\mathbf{N}^D[h] \right](x) \right\} = \mathbf{N}^D[h](x); \tag{5.6}$$

For $x \in \bar{D}$,

$$\min \left\{ \min \left\{ h(x), \mathbf{N}^D[h](x) \right\}, \mathbf{N}^{\mathbb{R}^n} \left[\min \left\{ h, \mathbf{N}^D[h] \right\} \right](x) \right\}$$
$$= \min \left\{ h(x), \mathbf{N}^D[h](x), \mathbf{N}^{\mathbb{R}^n}[h](x), \mathbf{N}^{\mathbb{R}^n} \left[\mathbf{N}^D[h] \right](x) \right\}$$
$$= \min \left\{ h(x), \mathbf{N}^{\mathbb{R}^n}[h](x) \right\} = \min \left\{ h(x), \mathbf{N}^D[h](x) \right\}.$$

These imply

$$\left(h^D \right)^{\mathbb{R}^n}(x) = h^D(x), \qquad x \in \mathbb{R}^n. \tag{5.7}$$

From the above discussion, by a standard argument, we see that the value function $\widetilde{V}(\cdot, \cdot)$ of Problem $(\widetilde{C})$, under some mild conditions, is the unique viscosity solution to the HJB QVI (4.16). We obtained

$$V(T, x) = \widetilde{V}(T, x), \qquad \forall x \in \mathbb{R}^n.$$

But this does not mean that

$$V(t, x) = \widetilde{V}(t, x), \qquad \forall (t, x) \in [0, T] \times \mathbb{R}^n. \tag{5.8}$$

The reason is that it is unknown whether HJB QVI (4.16) admits a unique viscosity solution in the function class that the value function $V(\cdot, \cdot)$ belongs to. On the other hand, if this were the case, then (5.8) would hold. Consequently, under proper conditions that make $\widetilde{V}(\cdot, \cdot)$ to be globally Lipschitz (or Hölder) continuous, one should have the same continuity for the value function $V(\cdot, \cdot)$. However, from our discussion in Section 3, this seems to be unlikely. Hence, we end up with a challenging open questions:

(i) *Is the viscosity solution to HJB QVI (4.16) unique within the class of locally Lipschitz (or Hölder) continuous functions?*

(ii) *Should there be some additional conditions for the HJB QVI so that it could characterize the value function?*

6. Concluding remarks

In this paper, we have introduced an intrinsic condition under which, together with other routine conditions, the value function of the optimal impulse control with terminal state constraint is continuous. This makes a big step towards the characterization of the value function. Due to the presence of the terminal state constraint, the value function is only locally Lipschitz (or Hölder) continuous and its growth is not slower than the impulse cost.

Therefore, the available techniques are not enough to characterizing the value function as the unique viscosity solution of the HJB QVI. Moreover, efforts are made to the discussion on a seemingly equivalent optimal impulse control problem without terminal state constraint. It leads to a challenging open question about the uniqueness of viscosity solutions to the HJB QVI.

REFERENCES

[1] M. Bardi and I. Capuzzo-Dolcetta, Optimal Control and Viscosity Solutions of Hamilton-Jacobi-Bellman Equations. Birkhäuser, Boston (1997).

[2] G. Barles, Quasi-variational inequalities and first-order Hamilton-Jacobi equations. *Nonlinear Anal.: Theory, Methods Appl.* **9** (1985) 131–148.

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

[4] C. Belak, S. Christensen and F.T. Seifried, A general verification result for stochastic impulse control problems. *SIAM J. Control Optim.* **55** (2019) 627–649.

[5] R. Bellma, Dynamic Programming. Princeton Univ. Press, Princeton (1957).

[6] A. Bensoussan and J.L. Lions, Nouvells formulation de problèmes de contrôle impulsonnel et applications. *C. R. Acad. Sci. Paris* **276** (1973) 1182–1192.

[7] A. Bensoussan and J.L. Lions, Impulse Control and Quasi-Variational Inequalities. Bordes, Paris (1984).

[8] A. Bensoussan and C.S. Tapiero, Impulsive control in management: prospects and applications. *J. Optim. Theory Appl.* **17** (1982) 419–442.

[9] M. Chahim, R.F. Hartl and P.M. Kort, A tutorial on the deterministic impulse control maximum principle: necessary and sufficient optimality conditions. *Eur. J. Oper. Res.* **219** (2012) 18–26.

[10] L. Chu, T. Kompas and Q. Grafton, Impulse controls and uncertainty in economics: method and application. *Environ. Model. Softw.* **65** (2015) 50–57.

[11] M.G. Crandall and P.L. Lions, Viscosity solutions of Hamilton-Jacobi equations. *Trans. AMS* **277** (1983) 1–42.

[12] V. Dykhta and O. Samsonuk, Maximum principle for nonsmooth optimal impulse problems. *Proc. IFAC Nonlinear Control Systems* (2001) 1303–1307.

[13] B. El Asri, Deterministic minimax impulse control in finite horizon: the viscosity solution approach. *ESAIM COCV* **19** (2013) 63–77.

[14] N. El Farouq, Degenerate first-order quasi-variational inequalities: an approach to approximate the value function. *SIAM J. Control Optim.* **55** (2017) 2714–2733.

[15] N. El Farouq, G. Barles and P. Bernhard, Deterministic minimax impulse control. *Appl. Math. Optim.* **61** (2010) 353–378.

[16] X. Feng, Maximum principle for optimal control problems involving impulse controls with nonsmooth data. *Stochastics* **88** (2016) 1188–1206.

[17] W.H. Fleming and H.M. Soner, Controlled Markov Processes and Viscosity Solutions. Springer-Verlag, New York (1992).

[18] Y. Hu and J. Yong, Maximum principle for stochastic optimal impulse controls. *Chin. Ann. Math. Ser. A, Suppl.* **12** (1991) 109–114 (in Chinese).

[19] R. Korn, Some applications of impulse control in mathematical finance. *Math. Methods Oper. Res.* **50** (1999) 189–218.

[20] R. Korn, Y. Melnyk and F.T. Seifried, Stochastic impulse control with regime-switching dynamics. *Eur. J. Oper. Res.* **260** (2017) 1024–1042.

[21] R. Leander, S. Lenhart and V. Protopopescu, Optimal control of continuous systems with impusle controls. *Optim. Control Appl. Meth.* **36** (2015) 535–549.

[22] S.M. Lenhart, Viscosity solutions associated with impulse control problems for piecewise-deterministic process. *Internat. J. Math. Math. Sci.* **12** (1989) 145–157.

[23] X. Li and J. Yong, Optimal Control Theory for Infinite Dimensional Systems. Birkhäuser, Boston (1995).

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

[25] J.L. Mendaldi and M. Robin, On some impulse control problems with constraint. *SIAM J. Control Optim.* **55** (2017) 3204–3225.

[26] A. Piunovskiy, A. Plakhov and M. Tumanov, Optimal impulse control os a SIR epidemic. *Optim. Control Appl. Meth.* **41** (2020) 448–468.

[27] L.S. Pontryagin, V.G. Boltyanskii, R.V. Gamkrelidze and E.F. Misjcjenko, The Mathematical Theory of Optimal Processes. Interscience, New York (1962).

[28] R. Rempala and J. Zabczyk, On the maximum principle for Deterministic impulse control problems. *J. Optim. Theory Appl.* **59** (1988) 281–288.

[29] H.M. Soner and N. Touzi, Stochastic target problems, dynamic programming and viscosity solutions. *SIAM J. Control Optim.* **41** (2002) 404–424.

[30] H.J. Sussmann, Small-time local controllability and continuity of the optimal time function for linear systems. *J. Optim. Theory Appl.* **53** (1987) 281–296.

[31] S. Tang and J. Yong, Finite horizon stochastic optimal switching and impulse controls with a viscosity solution approach. *Stochastics Stochastics Reports* **45** (1993) 145–176.

[32] N. Touzi, Optimal Stochastic Control, Stochastic Target Problems, and Backward SDE. Fields Institute Monographs, 29, Springer, New York (2013).

[33] Z. Wu and F. Zhang, Maximum principle for stochastic recursive optimal control problems involving impusle controls. *Abstr. Appl. Anal.* (2012) Art. ID 709682.

[34] V. Yadav and S.N. Balakrishnan, Optimal impulse control of systems with control constraints and application to HIV treatment. *Proc. Am. Control Conf.* (2006) 4824–4829.

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

[36] J. Yong, Zero-sum differential games involving impulse controls. *Appl. Math. Optim.* **29** (1994) 243–261.

[37] J. Yong, Differential Games: A Concise Introduction. World Scientific, Singapore (2015).

[38] J. Yong and P. Zhang, Necessary conditions of optimal impulse controls for distributed parameter systems. *Bull. Aust. Math. Soc.* **45** (1992) 305–326.

[39] J. Yong and X.Y. Zhou, Stochastic Controls: Hamiltonian Systems and HJB Equations. Springer-Verlag, New York (1999).