2019 18th European Control Conference (ECC)
Napoli, Italy, June 25-28, 2019

Sparse optimal feedback control for continuous-time systems

Takuya Ikeda[1] and Kenji Kashima[1]

Abstract— **In this paper, we investigate a sparse optimal control in a feedback framework. We adopt the dynamic programming approach and analyze the optimal control via the value function. Due to the non-smoothness of the L^0 cost functional, in general, the value function is not differentiable in the domain. Then, we characterize the value function as a viscosity solution to the associated Hamilton-Jacobi-Bellman (HJB) equation. Based on the result, we derive a sufficient and necessary condition for the L^0 optimality, which immediately gives the optimal feedback map. In addition, we consider the relationship with L^1 optimal control problem and show an equivalence theorem.**

I. INTRODUCTION

This work investigates an optimal control problem for nonlinear systems with the L^0 control cost. This cost functional penalizes the length of the support of control variables, and the optimization based on the criteria tends to make the control input identically zero on a set with positive measures. The optimal control is switched off completely on parts of the time domain. Hence, this type of control is also referred to as *sparse optimal control*. For example, this optimal control framework is applied to actuator placements [1], [2], [3], networked control systems [4], [5], and multi-period investments [6], to name a few.

The sparse optimal control involves the discontinuous and non-convex cost functional. Then, in order to deal with the difficulty of analysis, some relaxed problems with the L^p cost functional have been often investigated, akin to methods used in compressed sensing applications [7]. In [8], the L^1 cost functional is analyzed with an aim to show the relationship between the L^0 optimality and the L^1 optimality, and an equivalence theorem is derived. In [9], the result is extended to general linear systems including infinite dimensional systems. The L^1 control cost is also considered in [10], [11], [12]. In [9], the sparsity properties of optimal controls for the L^p cost with $p \in (0, 1)$ is discussed. While the literature aforementioned have investigated open loop optimal controls, more recently, optimal controls in the feedback framework have been also considered in [13], [14], [15]. The work [13] studies an infinite horizon optimal control problem with the L^p cost functional, where $p \in (0, 1]$, and derives the existence and the discreteness results for the time-discretized problem. In [14], an infinite horizon optimal control problem involving mixed quasi-norms of L^p-type cost functionals is considered, and the switching structure is studied and the optimal feedback control is computed in the numerical simulation. In [15], a finite horizon optimal control problem with the L^1 cost functional is discussed for stochastic control-affine dynamical systems, and proposes a sampling-based algorithm utilizing forward and backward stochastic differential equations.

In the previous works, the exact L^0 optimal feedback control was not investigated. To the best of our knowledge, while there are works that directly address the L^0 cost in the context of the open loop design, e.g. [16], [17], the counterparts in the context of the feedback control do not exist. Then, the purpose of our paper is to directly deal with the underlying non-smooth and non-convex L^0 optimal control problem without the aid of any L^p relaxations. For this purpose, we adopt dynamic programming approach and investigate the sparse optimal control via the value function. Due to the non-smoothness of the L^0 cost functional, our value function is not differentiable, and hence it does not satisfy the associated HJB equation in the classical sense. Then, we first characterize the value function as a viscosity solution of the HJB equation. Based on the result, we show a sufficient and necessary condition for the L^0 optimality, which immediately gives an optimal feedback map. In addition, motivated by the discussion given in previous works, we also consider the relationship between the L^0 optimality and the L^1 optimality, and show an equivalence theorem by utilizing the uniqueness theorem of the viscosity solution. Moreover, the discreteness property of the sparse control is often reported in some frameworks. Then, we further show this property of our sparse optimal control for control-affine systems with a box constraint. This paper thus investigates a non-differentiable value function. We lastly point out [18] to readers as a recent relevant study, which investigates an optimal control problem with a discontinuous value function.

The remainder of this paper is organized as follows: In Section II, we give mathematical preliminaries for our subsequent discussion. In section III, we formulate our optimal control problem. Section IV is devoted to the theoretical analysis. We first characterize the value function as a viscosity solution to the associated HJB equation, and next show a sufficient and necessary condition for the L^0 optimality. We also mention the relationship with the L^1 optimization problem and some basic properties of the sparse optimal control for control-affine systems with box constraints. In Section V we offer concluding remarks.

[1]T. Ikeda and K. Kashima are with Graduate School of Informatics, Kyoto University, Kyoto, 606-8501, Japan. ikeda.t@bode.amp.i.kyoto-u.ac.jp, kk@i.kyoto-u.ac.jp

II. MATHEMATICAL PRELIMINARIES

This section reviews notation that will be used throughout the paper.

Let N, N_1, and N_2 be positive integers. For a matrix $M \in \mathbb{R}^{N_1 \times N_2}$, $M^\top$ denotes the transpose of M. For a vector $a = [a_1, a_2, \ldots, a_N]^\top \in \mathbb{R}^N$, we denote the Euclidean norm by $\|a\| \triangleq (\sum_{i=1}^N a_i^2)^{1/2}$, the open ball with center at a and radius $r > 0$ by $B(a,r)$, i.e., $B(a,r) \triangleq \{x \in \mathbb{R}^N : \|x - a\| < r\}$, and the closed ball with center at a and radius $r > 0$ by $\bar{B}(a,r)$, i.e., $B(a,r) \triangleq \{x \in \mathbb{R}^N : \|x-a\| \leq r\}$. We denote the inner product of $a \in \mathbb{R}^N$ and $b \in \mathbb{R}^N$ by $a \cdot b$. For scalars $a \in \mathbb{R}$ and $b \in \mathbb{R}$, $\min\{a,b\}$ (resp. $\max\{a,b\}$) returns the smaller one (resp. larger one) of a and b.

Let $T > 0$. For $p \in \{0, 1, \infty\}$, L^p denotes the set of all continuous-time signals $u(t) = [u_1(t), u_2(t), \ldots, u_N(t)]^\top \in \mathbb{R}^N$ over a time interval $[0,T]$ such that $\|u\|_p < \infty$, where $\|\cdot\|_p$, referred to as L^p *norm*, is defined by

$$
\begin{aligned}
\|u\|_0 &\triangleq \sum_{j=1}^N \mu_L(\{t \in [0,T] : u_j(t) \neq 0\}), \\
\|u\|_1 &\triangleq \sum_{j=1}^N \int_0^T |u_j(t)| dt, \\
\|u\|_\infty &\triangleq \max_{j=1,2,\ldots,N} \operatorname*{ess\,sup}_{0 \leq t \leq T} |u_j(t)|,
\end{aligned}
$$

with the Lebesgue measure μ_L on $\mathbb{R}$. The L^0 norm is also expressed by $\|u\|_0 = \int_0^T \psi_0(u(t))dt$, where $\psi_0 : \mathbb{R}^N \to \mathbb{R}$ is a function that returns the number of nonzero components, i.e.,

$$
\psi_0(a) \triangleq \sum_{j=1}^m |a_j|^0
$$

with $0^0 = 0$.

For a given set $\Omega \subset \mathbb{R}^N$, $C(\Omega)$ denotes the set of all continuous functions on Ω, and $C^1(\Omega)$ denotes the set of all differentiable functions with continuous partial derivatives in Ω.

For a locally Lipschitz continuous function $v : \Omega \to \mathbb{R}$ with an open set $\Omega \subset \mathbb{R}^N$, the *lower Dini derivative* and the *upper Dini derivative* of v at $x \in \Omega$ in the direction $q \in \mathbb{R}^N$ are defined by

$$
\begin{aligned}
\partial^- v(x;q) &\triangleq \liminf_{h \to 0+} \frac{v(x+hq) - v(x)}{h}, \\
\partial^+ v(x;q) &\triangleq \limsup_{h \to 0+} \frac{v(x+hq) - v(x)}{h}.
\end{aligned}
$$

If these two values coincide, then we denote the derivative by $\partial v(x;q)$. We also denote the gradient of v at x by $Dv(x)$.

Given a partial differential equation

$$
F(x, v(x), Dv(x)) = 0, \quad x \in \Omega, \tag{1}
$$

where Ω is an open set of $\mathbb{R}^N$ and $Dv(x)$ is the gradient of v at x, a function $v \in C(\Omega)$ is said to be a *viscosity subsolution* of (1) if, for any $\phi \in C^1(\Omega)$,

$$
F(x_0, v(x_0), D\phi(x_0)) \leq 0
$$

at any local maximum point $x_0 \in \Omega$ of $v - \phi$. Similarly, a function $v \in C(\Omega)$ is said to be a *viscosity supersolution* of (1) if, for any $\phi \in C^1(\Omega)$,

$$
F(x_0, v(x_0), D\phi(x_0)) \geq 0
$$

at any local minimum point $x_0 \in \Omega$ of $v - \phi$. Finally, v is said to be a *viscosity solution* of (1), if it is simultaneously a viscosity subsolution and supersolution.

III. PROBLEM FORMULATION

We consider the following control system:

$$
\begin{aligned}
&\dot{y}(t) = f(y(t), u(t)), \quad t > 0, \\
&y(0) = x,
\end{aligned} \tag{2}
$$

where $y(t) \in \mathbb{R}^n$ is the state variable, $x \in \mathbb{R}^n$ is the initial state, and $u(t) \in \mathbb{R}^m$ is the control variable. For given initial state $x \in \mathbb{R}^n$ and final time of control $t \in (0, \infty)$, we consider the cost functional

$$
J(x,t,u) \triangleq \int_0^t \psi_0(u(s))ds + g(y(t)),
$$

where y and u satisfy the equation (2), g is the terminal cost, and the function $\psi_0 : \mathbb{R}^m \to \mathbb{R}$ returns the number of nonzero components, which is defined in the Section II. Note that the first term expresses the L^0 cost of the control input, and hence the minimization of J enhances the sparsity. We assume the range of the control u is constrained in a compact set $\mathbb{U} \subset \mathbb{R}^m$ that contains $0 \in \mathbb{R}^m$, i.e., $u(t) \in \mathbb{U}$ for all t, and we denote the set of all such functions by $\mathcal{U}$, i.e.,

$$
\mathcal{U} \triangleq \{u \in L^\infty : u(t) \in \mathbb{U} \text{ for all } t\}.
$$

In other words, the main problem is formulated as follows:

Problem 1: Given $x \in \mathbb{R}^n$ and $t \geq 0$, find a control input u on $[0,t]$ that solves

$$
\begin{aligned}
\underset{u}{\text{minimize}} \quad & J(x,t,u) \\
\text{subject to} \quad & \dot{y}(s) = f(y(s), u(s)), \\
& y(0) = x, \\
& u \in \mathcal{U}.
\end{aligned}
$$

In this paper, we assume the following conditions for the dynamic $f(y,u)$ and the terminal cost $g(y)$:

(A_1) $f : \mathbb{R}^n \times \mathbb{U} \to \mathbb{R}^n$ is continuous;

(A_2) f is Lipschitz continuous in the state variable, uniformly in the control variable, i.e., there exists a constant L such that

$$
\|f(y,u) - f(z,u)\| \leq L\|y - z\| \tag{3}
$$

for all $y, z \in \mathbb{R}^n$ and $u \in \mathbb{U}$;

(A_3) $g : \mathbb{R}^n \to \mathbb{R}$ is continuous.

Assumptions (A_1) and (A_2) guarantee the existence and the uniqueness of a solution to the differential equation (2); see [19]. For given $x \in \mathbb{R}^n$ and $u \in \mathcal{U}$, we denote the state at time τ by $y(x;\tau;u)$, or briefly by $y(\tau)$ if no confusion may arise. As will be seen in Theorem 1, Assumption (A_3) is used to get the continuity of the *value function*, which is defined by

$$
V(x,t) \triangleq \inf_{u \in \mathcal{U}} J(x,t,u), \quad x \in \mathbb{R}^n, \quad t \in [0, \infty).
$$

IV. ANALYSIS

In this section, we derive the associated HJB equation and a sufficient and necessary condition for the L^0 optimality, which gives an optimal feedback map for Problem 1. Our analysis is based on the dynamic programming approach, and hence we first investigate the value function V.

A. Characterization of the Value Function

We here characterize the value function as a viscosity solution to HJB equation. For the purpose, we first show the continuity of the value function, which is a fundamental property since we investigate the solutions to the HJB equation in the class $C(\mathbb{R}^n \times [0,T])$ for some $T > 0$. Here, the basic estimates of the solution y to (2) are the following [19]:

Lemma 1: Fix any $x, z \in \mathbb{R}^n$. If (A_1) and (A_2) are satisfied, then

$$\|y(x;t;u)\| \le (\|x\| + \sqrt{2Kt})e^{Kt} \quad \forall u \in \mathcal{U}, \forall t > 0, \tag{4}$$

$$\|y(x;t;u) - y(z;t;u)\| \le e^{Lt}\|x - z\| \quad \forall u \in \mathcal{U}, \forall t > 0, \tag{5}$$

$$\|y(x;t;u) - x\| \le F_x t \quad \forall u \in \mathcal{U}, \forall t \in [0, 1/F_x], \tag{6}$$

where $K \triangleq L + \sup\{f(0,u) : u \in \mathbb{U}\}$, L is a constant that satisfies (3), and $F_x \triangleq \sup\{\|f(w,u)\| : \|x - w\| \le 1, u \in \mathbb{U}\}$.

Theorem 1: Fix $T > 0$. Under assumptions (A_1), (A_2), and (A_3), the value function V is continuous on $\mathbb{R}^n \times [0,T]$. If in addition the terminal cost g is Lipschitz continuous, then V is locally Lipschitz continuous.

Proof: Fix any $(x,t) \in \mathbb{R}^n \times [0,T]$ and $\varepsilon > 0$. Then, the set $\tilde{D}_{x,t} \triangleq \{y(x;t;u) : u \in \mathcal{U}\} \subset \mathbb{R}^n$ is bounded from (4) in Lemma 1. Hence, we can take a compact set $D_{x,t} \subset \mathbb{R}^n$ that contains $\tilde{D}_{x,t}$ in its interior. Since g is continuous on $\mathbb{R}^n$, g is uniformly continuous on $D_{x,t}$. It follows that there exists $\delta > 0$ depending on x, t, ε such that, for any $c, d \in D_{x,t}$ with $\|c - d\| < \delta$, $|g(c) - g(d)| < \varepsilon$ holds. In addition, from (5), there exists $r > 0$ depending on x, t, ε such that, for any $(z,\tau) \in B((x,t), r)$, $\|y(x;t;u) - y(z;\tau;u)\| < \delta$ and $y(z;\tau;u) \in D_{x,t}$ hold for all $u \in \mathcal{U}$. This can be observed from

$$\begin{aligned}
&\|y(x;t;u) - y(z;\tau;u)\| \\
&\quad \le \|y(x;t;u) - y(x;\tau;u)\| + \|y(x;\tau;u) - y(z;\tau;u)\| \\
&\quad \le \left\| \int_\tau^t f(y(x;s;u), u(s))ds \right\| + e^{L\tau}\|x - z\| \\
&\quad \le C_x|t - \tau| + e^{LT}\|x - z\|
\end{aligned}$$

for all $u \in \mathcal{U}$, where

$$C_x \triangleq \max\{\|f(y,u)\| : (y,u) \in \bar{B}(0,\bar{x}) \times \mathbb{U}\} \tag{7}$$

with $\bar{x} \triangleq (\|x\| + \sqrt{2KT})e^{KT}$ and $C_x < \infty$ from the assumption (A_1). Thus, for $(z,\tau) \in B((x,t), r)$,

$$|g(y(x;t;u)) - g(y(z;\tau;u))| < \varepsilon$$

for all $u \in \mathcal{U}$.

Here, fix any $(z,\tau) \in B((x,t), \min\{r, \varepsilon\})$. Then, by definition of V, there exists a control $\bar{u} \in \mathcal{U}$ such that

$$V(z,\tau) + \varepsilon \ge J(z,\tau,\bar{u}). \tag{8}$$

Then,

$$\begin{aligned}
&V(x,t) - V(z,\tau) \\
&\quad \le J(x,t,\bar{u}) - J(z,\tau,\bar{u}) + \varepsilon \\
&\quad = \int_{\min\{t,\tau\}}^{\max\{t,\tau\}} \psi_0(\bar{u}(s))ds + g(y(x;t;\bar{u})) - g(y(z;\tau;\bar{u})) \\
&\qquad + \varepsilon \\
&\quad < m|t - \tau| + \varepsilon + \varepsilon < (m+2)\varepsilon,
\end{aligned}$$

where we used the boundedness of ψ_0 in the third estimation, i.e., $|\psi_0(a)| \le m$ for all $a \in \mathbb{R}^m$. Similarly, we can also show $V(z,\tau) - V(x,t) < (m+2)\varepsilon$ by taking a control $\tilde{u} \in \mathcal{U}$ such that $V(x,t) + \varepsilon \ge J(x,t,\tilde{u})$. This shows the continuity of V.

If in addition g is Lipschitz continuous, then the result follows from the local boundedness of f in the state variable. More precisely, take any bounded neighborhood $D_{x,t}$ that contains (x,t). Define $C_x^* \triangleq \sup\{C_w : (w,s) \in D_{x,t}$ for some $s\}$, where C_x is defined in (7) and $C_x^* < \infty$. Take any $(w,s) \in D_{x,t}$ and $(z,\tau) \in D_{x,t}$. Fix any $\varepsilon > 0$ and take $\bar{u}$ that satisfies (8) for (z,τ). Let G be the Lipschitz constant of g. Then, from the estimation above, we have

$$\begin{aligned}
&V(w,s) - V(z,\tau) \\
&\quad \le m|s - \tau| + |g(y(w;s;\bar{u})) - g(y(z;\tau;\bar{u}))| + \varepsilon \\
&\quad \le m|s - \tau| + G\|y(w;s;\bar{u}) - y(z;\tau;\bar{u})\| + \varepsilon \\
&\quad \le (m + GC_x^*)|s - \tau| + Ge^{LT}\|w - z\| + \varepsilon.
\end{aligned}$$

The arbitrarity of ε and the similar discussion above complete the proof. ■

For the value function, we have another fundamental lemma.

Lemma 2: Assume (A_1), (A_2), and (A_3). Fix any $t \in (0,\infty)$. Then

$$V(x,t) = \inf_{u \in \mathcal{U}} \left\{ \int_0^\tau \psi_0(u(s))ds + V(y(x;\tau;u), t - \tau) \right\}$$

for all $x \in \mathbb{R}^n$ and $\tau \in (0,t]$.

Proof: See Appendix A. ■

The next theorem is one of the main results of this paper.

Theorem 2: Assume (A_1), (A_2), and (A_3). Fix $T > 0$. Then, the value function V is a viscosity solution of the following Hamilton-Jacobi-Bellman equation with an initial condition:

$$\begin{cases} v_t(x,t) + H(x, D_x v(x,t)) = 0 & \text{in } \mathbb{R}^n \times (0,T), \quad (9) \\ v(x,0) = g(x) & \text{in } \mathbb{R}^n, \quad (10) \end{cases}$$

where $H : \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}$ is defined by

$$H(x,p) \triangleq \sup_{u \in \mathbb{U}}\{-f(x,u) \cdot p - \psi_0(u)\},$$

v_t denotes the partial derivative with respect to the last variable and $D_x v$ denotes the gradient with respect to the first n variables.

Proof: We first show that V is a viscosity subsolution of the equation (9). Fix any $\phi \in C^1(\mathbb{R}^n \times (0,T))$, and let (x,t) be a local maximum point of $V-\phi$. Then, there exists $r>0$ such that, for all $z \in \mathbb{R}^n$ and $s \in (0,T)$ with $\|x-z\| < r$ and $|t-s|<r$,

$$\phi(x,t) - \phi(z,s) \le V(x,t) - V(z,s). \tag{11}$$

Take any $u \in \mathbb{U}$ and define $\bar{u}(\tau) = u$ for all τ. Then, by continuity of y, for sufficiently small τ, we have $\|y(x;\tau;\bar{u}) - x\| < r$ and $\tau < r$. It follows from (11) that

$$\phi(x,t) - \phi(y(x;\tau;\bar{u}), t-\tau) \le V(x,t) - V(y(x;\tau;\bar{u}), t-\tau).$$

Here, we have

$$\begin{aligned} V(x,t) &\le \int_0^\tau \psi_0(\bar{u}(s))ds + V(y(x;\tau;\bar{u}), t-\tau) \\ &= \tau\psi_0(u) + V(y(x;\tau;\bar{u}), t-\tau) \end{aligned}$$

from Lemma 2. Hence,

$$\phi(x,t) - \phi(y(x;\tau;\bar{u}), t-\tau) \le \tau\psi_0(u).$$

Devide τ and let $\tau \to 0$, then

$$-D_x\phi(x,t)\cdot f(x,u) + \phi_t(x,t) \le \psi_0(u).$$

This inequality holds for all $u \in \mathbb{U}$. This means

$$\phi_t(x,t) + \sup_{u\in\mathbb{U}}\{-D_x\phi(x,t)\cdot f(x,u) - \psi_0(u)\} \le 0.$$

We next show that V is the viscosity supersolution of (9). Fix any $\phi \in C^1(\mathbb{R}^n \times (0,T))$, and let (x,t) be a local minimum point of $V-\phi$. Then, there exists $r>0$ such that, for all $z \in \mathbb{R}^n$ and $s \in (0,T)$ with $\|x-z\|<r$ and $|t-s|<r$,

$$V(x,t) - V(z,s) \le \phi(x,t) - \phi(z,s). \tag{12}$$

Here, fix any $\varepsilon > 0$ and $\tau \in (0,t]$. Then, from Lemma 2, there exists $\tilde{u} \in \mathcal{U}$, which depends on ε and τ, such that

$$V(x,t) + \tau\varepsilon \ge \int_0^\tau \psi_0(\tilde{u}(s))ds + V(y(x;\tau;\tilde{u}), t-\tau).$$

Denote simply $y(x;s;\tilde{u})$ by $\tilde{y}(s)$ for all s. For sufficiently small $\tau>0$, we have $\|\tilde{y}(\tau) - x\| < r$ and $\tau < r$ by (6). It follows from (12) that

$$\begin{aligned} \phi(x,t) - \phi(\tilde{y}(\tau), t-\tau) &\ge V(x,t) - V(\tilde{y}(\tau), t-\tau) \\ &\ge \int_0^\tau \psi_0(\tilde{u}(s))ds - \tau\varepsilon. \end{aligned} \tag{13}$$

Here,

$$\begin{aligned} &\phi(x,t) - \phi(\tilde{y}(\tau), t-\tau) \\ &\quad = -\int_0^\tau \frac{d}{ds}\phi(\tilde{y}(s), t-s)ds \\ &\quad = -\int_0^\tau D_x\phi(\tilde{y}(s), t-s)\cdot f(\tilde{y}(s), \tilde{u}(s))ds \\ &\qquad\quad + \int_0^\tau \phi_t(\tilde{y}(s), t-s)ds. \end{aligned} \tag{14}$$

Note that

$$\begin{aligned} &\int_0^\tau D_x\phi(\tilde{y}(s), t-s)\cdot f(\tilde{y}(s), \tilde{u}(s))ds \\ &\qquad = \int_0^\tau D_x\phi(x,t)\cdot f(x, \tilde{u}(s))ds + o(\tau), \end{aligned}$$

where $o(\tau)$ denotes a function $h(\tau)$ such that $\lim_{\tau\to 0+} h(\tau)/\tau = 0$, and we used $\phi \in C^1$, (4), (A_1), and (6) with a calculation

$$\begin{aligned} &\int_0^\tau \big\{D_x\phi(\tilde{y}(s), t-s)\cdot f(\tilde{y}(s), \tilde{u}(s)) \\ &\qquad - D_x\phi(x,t)\cdot f(x,\tilde{u}(s))\big\}ds \\ &= \int_0^\tau \big\{D_x\phi(\tilde{y}(s), t-s) - D_x\phi(x,t)\big\}\cdot f(\tilde{y}(s), \tilde{u}(s))ds \\ &\qquad + \int_0^\tau D_x\phi(x,t)\cdot \big\{f(\tilde{y}(s), \tilde{u}(s)) - f(x,\tilde{u}(s))\big\}ds. \end{aligned}$$

Similarly,

$$\int_0^\tau \phi_t(\tilde{y}(s), t-s)ds = \phi_t(x,t) + o(\tau).$$

It follows from (13) and (14) that

$$\begin{aligned} -\tau\varepsilon &\le \int_0^\tau \big\{-D_x\phi(x,t)\cdot f(x,\tilde{u}(s)) - \psi_0(\tilde{u}(s))\big\}ds \\ &\qquad + \phi_t(x,t) + o(\tau) \\ &\le \tau \sup_{u\in\mathbb{U}}\big\{-D_x\phi(x,t)\cdot f(x,u) - \psi_0(u)\big\} \\ &\qquad + \phi_t(x,t) + o(\tau). \end{aligned}$$

Devide $\tau > 0$ and let $\tau \to 0$, then

$$-\varepsilon \le \sup_{u\in\mathbb{U}}\{-D_x\phi(x,t)\cdot f(x,u) - \psi_0(u)\} + \phi_t(x,t).$$

The arbitrarity of ε shows that V is the viscosity supersolution of (9). Thus, V is a viscosity solution of (9). Moreover, V satisfies $V(x,0) = g(x)$ in $\mathbb{R}^n$ by definition, and hence V is a viscosity solution of (9) with an initial condition (10). ∎

We next mention the connection to the L^1 optimization problem, which is based on the uniqueness theorem of the viscosity solution.

Theorem 3: Assume (A_1), (A_2), (A_3), and fix $T>0$. If the dynamics f is affine in u, i.e.,

$$f(y,u) = f_0(y) + \sum_{j=1}^m f_j(y)u_j \tag{15}$$

for some $f_j : \mathbb{R}^n \to \mathbb{R}^n$, $j = 0,1,2,\ldots,m$, and $\mathbb{U} = \{u \in \mathbb{R}^m : |u_j| \le 1, \forall j\}$, then the value function V coincides with the value function of the corresponding L^1 *optimal control problem*. More precisely, the L^1 optimal control problem is defined by the optimization problem in which the cost functional of Problem 1 is replaced with

$$J_1(x,t,u) \triangleq \sum_{j=1}^m \int_0^t |u_j(s)|ds + g(y(t)),$$

and the value function V_1 of this problem is defined by

$$V_1(x,t) \triangleq \inf_{u\in\mathcal{U}} J_1(x,t,u).$$

Then $V = V_1$ holds on $\mathbb{R}^n \times [0,T)$.

Proof: In this case, for any $x, p \in \mathbb{R}^n$,

$$\begin{aligned} H(x,p) &= \sup_{u\in\mathbb{U}}\Big\{ -\sum_{j=1}^m f_j(x)u_j\cdot p - \sum_{j=1}^m |u_j|^0 \Big\} - f_0(x)\cdot p \\ &= \sum_{j=1}^m \sup_{u_j\in\mathbb{U}_j} \Big\{ -\big(f_j(x)\cdot p\big)u_j - |u_j|^0 \Big\} - f_0(x)\cdot p, \end{aligned}$$

where $\mathbb{U}_j$ is the set of all j-th components of $\mathbb{U}$, i.e., $\mathbb{U}_j \triangleq \{a \in \mathbb{R} : |a| \le 1\}$. Here, it follows from an elementary calculation that

$$\sup_{u_j\in\mathbb{U}_j}\Big\{ -a^j_{x,p}u_j - |u_j|^0 \Big\} = \sup_{u_j\in\mathbb{U}_j}\Big\{ -a^j_{x,p}u_j - |u_j| \Big\}$$

for all $x, p \in \mathbb{R}^n$ and $j = 1, 2, \ldots, m$, where $a^j_{x,p} \triangleq f_j(x)\cdot p$. Indeed, the supremum of both sides is given by

$$\begin{cases} a^j_{x,p} - 1, & \text{if} \quad -a^j_{x,p} + 1 < 0, \\ 0, & \text{if} \quad -a^j_{x,p} + 1 = 0, \\ 0, & \text{if} \quad -a^j_{x,p} + 1 > 0, \\ & \quad\ \text{and } a^j_{x,p} + 1 > 0, \\ 0, & \text{if} \quad a^j_{x,p} + 1 = 0, \\ -a^j_{x,p} - 1, & \text{if} \quad a^j_{x,p} + 1 < 0. \end{cases}$$

Hence, the equation (9) is equivalent to

$$v_t(x,t) + H_1(x, D_x v(x,t)) = 0, \tag{16}$$

where

$$\begin{aligned} H_1(x,p) &\triangleq \sup_{u\in\mathbb{U}}\{-f(x,u)\cdot p - \psi_1(u)\}, \quad \forall x, \forall p \in \mathbb{R}^n, \\ \psi_1(a) &\triangleq \sum_{j=1}^m |a_j|, \quad \forall a \in \mathbb{R}^m. \end{aligned}$$

Note that the equation (16) is the HJB equation for the L^1 optimal control problem, and it is known that the value function V_1 is the unique viscosity solution of the equation with initial condition (10) in the class $C(\mathbb{R}^n \times (0,T))$ [19]. This means $V = V_1$. ■

B. Optimality of a Control

We here derive a sufficient and necessary condition for the L^0 optimality, which is the second main result.

Theorem 4: Assume (A_1), (A_2), and g is Lipschitz continuous. Fix any initial state $x \in \mathbb{R}^n$ and final time of control $T > 0$. Then, a control u is an optimal solution to Problem 1 if and only if

$$\partial V(y(t), T-t; f(y(t),u(t)), -1) + \psi_0(u(t)) = 0 \tag{17}$$

almost everywhere on $(0,T)$.

Proof: We first show the sufficiency. We assume (17) and take any u that satisfies (17). Put

$$h(t) \triangleq \int_0^t \psi_0(u(s))ds + V(y(t), T-t) \tag{18}$$

for $t \in (0,T)$. Since ψ_0 is integrable and the value function V is locally Lipschitz continuous by Theorem 1, $h(t)$ is differentiable almost everywhere, and

$$\begin{aligned} \frac{d}{dt}h(t) &= \psi_0(u(t)) + \frac{d}{dt}V(y(t), T-t) \\ &= \psi_0(u(t)) + \partial V(y(t), T-t; f(y(t),u(t)), -1). \end{aligned} \tag{19}$$

It follows from (17) that $\frac{d}{dt}h(t) = 0$ almost everywhere on $(0,T)$. Then, the Lipschitz continuity of h shows $h(0) = h(T) = J(x,T,u)$. On the other hand, $h(0) = V(x,T)$ by definition of h. This means $J(x,T,u) = V(x,T)$, and hence u is optimal.

We next show the necessity. We assume u is optimal. Then, from Lemma 2, $h(t)$ defined by (18) is constant. Hence, $\frac{d}{dt}h(t) = 0$. This with (19) completes the proof. ■

Remark 1: Theorem 4 mentions the relationship with the optimal control value and the state value at the current time, and hence the optimal control is immediately characterized in terms of the feedback control.

Remark 2: In addition, Theorem 4 reveals the discreteness of the sparse optimal control for control-affine systems with a box constraint, i.e., f is given by (15) and $\mathbb{U} = \{u \in \mathbb{R}^m : U_j^- \le u_j \le U_j^+, \forall j\}$ for some $U_j^- < 0$ and $U_j^+ > 0$. Here we take any (z,t) such that the value function V is differentiable at $(z, T-t)$. Note that such points exist almost everywhere, since V is locally Lipschitz continuous from Theorem 1. Denote the optimal input value for the current point (z,t) by $u(z,t)$. Then, the differentiability of V and (17) imply that

$$\begin{aligned} -\psi_0(u(z,t)) &= \partial V(z, T-t; f(z,u(z,t)), -1) \\ &= DV(z, T-t)\cdot(f(z,u(z,t)), -1) \\ &= D_xV(z,T-t)\cdot f(z,u(z,t)) - V_t(z,T-t). \end{aligned}$$

Here, V is a viscosity solution of (9), and hence V satisfies the equation at any point where V is differentiable. In other words,

$$V_t(z,T-t) - D_xV(z,T-t)\cdot f(z,u) - \psi_0(u) \le 0$$

for all $u \in \mathbb{U}$. Hence,

$$u(z,t) \in \arg\min_{u\in\mathbb{U}}\{D_xV(z,T-t)\cdot f(z,u) + \psi_0(u)\}.$$

Here, if the system is control-affine, such as (15), then the j-th component $u_j(z,t)$ of $u(z,t)$ is given by

$$u_j(z,t) \in \arg\min_{u_j\in\mathbb{U}_j}\{D_xV(z,T-t)\cdot f_j(z)u_j + |u_j|^0\}, \tag{20}$$

where $\mathbb{U}_j$ is the set of all j-th components of $\mathbb{U}$, i.e., $\mathbb{U}_j = \{a \in \mathbb{R} : U_j^- \le a \le U_j^+\}$, and define $b^j_{z,t} \triangleq D_xV(z,T-$

$t) \cdot f_j(z)$. Then, it follows from (20) that

$$u_j(z,t) \in \begin{cases} \{U_j^-\}, & \text{if} \quad b_{z,t}^j U_j^- + 1 < 0, \\ \{U_j^-, 0\}, & \text{if} \quad b_{z,t}^j U_j^- + 1 = 0, \\ \{0\}, & \text{if} \quad b_{z,t}^j U_j^- + 1 > 0, \\ & \quad \text{and } b_{z,t}^j U_j^+ + 1 > 0, \\ \{0, U_j^+\}, & \text{if} \quad b_{z,t}^j U_j^+ + 1 = 0, \\ \{U_j^+\}, & \text{if} \quad b_{z,t}^j U_j^+ + 1 < 0. \end{cases}$$

Thus, the optimal control takes only three values of $\{U_j^-, 0, U_j^+\}$.

V. CONCLUSIONS

We have investigated a finite horizon optimal control problem with the L^0 control cost functional. We have first characterized the value function as a viscosity solution to the associated HJB equation, and shown an equivalence theorem between the L^0 optimality and the L^1 optimality via the uniqueness theorem of the viscosity solution. In addition, we have derived a sufficient and necessary condition for the L^0 optimality that connects the current state and the current optimal control value. We have finally revealed the discreteness property of the sparse optimal control for control-affine systems. Future work includes an extension to stochastic control systems.

APPENDIX

A. Proof of Lemma 2

The result is obvious for $\tau = t$, since we have $V(x,0) = g(x)$ for all $x \in \mathbb{R}^n$. Then, we take any $\tau \in (0,t)$. Fix any $u \in \mathcal{U}$, and denote $\bar{u}(s) \triangleq u(s+\tau)$, then we have

$$\begin{aligned} J(x,t,u) &= \int_0^\tau \psi_0(u(s))ds + \int_0^{t-\tau} \psi_0(\bar{u}(s))ds + g(y(t)) \\ &= \int_0^\tau \psi_0(u(s))ds + J(y(x;\tau;u), t-\tau, \bar{u}) \\ &\geq \int_0^\tau \psi_0(u(s))ds + V(y(x;\tau;u), t-\tau). \end{aligned}$$

Taking the infimum over $u \in \mathcal{U}$, we get the inequality

$$V(x,t) \geq \inf_{u \in \mathcal{U}} \left\{ \int_0^\tau \psi_0(u(s))ds + V(y(x;\tau;u), t-\tau) \right\}.$$

We next show the inverse inequality. Fix any $u \in \mathcal{U}$ and $\varepsilon > 0$. Denote $z \triangleq y(x;\tau;u)$. Then, there exists $\tilde{u} \in \mathcal{U}$ such that

$$V(z, t-\tau) + \varepsilon \geq J(z, t-\tau, \tilde{u}).$$

Put $\hat{u}(s) \triangleq u(s)$ for $0 \leq s \leq \tau$ and $\hat{u}(s) \triangleq \tilde{u}(s-\tau)$ for $\tau < s \leq t$. Then,

$$\begin{aligned} V(x,t) &\leq J(x,t,\hat{u}) \\ &= \int_0^\tau \psi_0(u(s))ds + \int_0^{t-\tau} \psi_0(\tilde{u}(s))ds + g(y(x;t;\hat{u})). \end{aligned}$$

Here, $y(x;t;\hat{u}) = y(z;t-\tau;\tilde{u})$, and hence

$$\begin{aligned} V(x,t) &\leq \int_0^\tau \psi_0(u(s))ds + J(z, t-\tau, \tilde{u}) \\ &\leq \int_0^\tau \psi_0(u(s))ds + V(z, t-\tau) + \varepsilon. \end{aligned}$$

The arbitrarity of ε and $u \in \mathcal{U}$ show the desired inequality.

ACKNOWLEDGMENT

This work was supported by JSPS KAKENHI Grant Number 18H01461.

REFERENCES

[1] G. Stadler, "Elliptic optimal control problems with L^1-control cost and applications for the placement of control devices," *Computational Optimization and Applications*, vol. 44, no. 2, pp. 159–181, 2009.

[2] R. Herzog, G. Stadler, and G. Wachsmuth, "Directional sparsity in optimal control of partial differential equations," *SIAM Journal on Control and Optimization*, vol. 50, no. 2, pp. 943–963, 2012.

[3] K. Kunisch, K. Pieper, and B. Vexler, "Measure valued directional sparsity for parabolic optimal control problems," *SIAM Journal on Control and Optimization*, vol. 52, no. 5, pp. 3078–3108, 2014.

[4] M. Nagahara, D. Quevedo, and J. Østergaard, "Sparse packetized predictive control for networked control over erasure channels," *IEEE Transactions on Automatic Control*, vol. 59, no. 7, pp. 1899–1905, 2014.

[5] T. Ikeda and K. Kashima, "Sparsity-constrained controllability maximization with application to time-varying control node selection," *IEEE Control Systems Letters*, vol. 2, no. 3, pp. 321–326, 2018.

[6] S. Boyd, M. T. Mueller, B. O' Donoghue, Y. Wang *et al.*, "Performance bounds and suboptimal policies for multi–period investment," *Foundations and Trends® in Optimization*, vol. 1, no. 1, pp. 1–72, 2014.

[7] D. L. Donoho, "Compressed sensing," *IEEE Transactions on information theory*, vol. 52, no. 4, pp. 1289–1306, 2006.

[8] M. Nagahara, D. E. Quevedo, and D. Nešić, "Maximum hands-off control: a paradigm of control effort minimization," *IEEE Transactions on Automatic Control*, vol. 61, no. 3, pp. 735–747, 2016.

[9] T. Ikeda and K. Kashima, "On sparse optimal control for general linear systems," *IEEE Transactions on Automatic Control*, 2018.

[10] W. Alt and C. Schneider, "Linear-quadratic control problems with L^1-control cost," *Optimal control applications and methods*, vol. 36, no. 4, pp. 512–534, 2015.

[11] G. Vossen and H. Maurer, "On L^1-minimization in optimal control and applications to robotics," *Optimal Control Applications and Methods*, vol. 27, no. 6, pp. 301–321, 2006.

[12] M. Athans and P. L. Falb, *Optimal Control: An Introduction to the Theory and Its Applications*. Dover Publications, 1966.

[13] D. Kalise, K. Kunisch, and Z. Rao, "Infinite horizon sparse optimal control," *Journal of Optimization Theory and Applications*, vol. 172, no. 2, pp. 481–517, 2017.

[14] ——, "Sparse and switching infinite horizon optimal control with nonconvex penalizations," *arXiv preprint arXiv:1808.10659*, 2018.

[15] I. Exarchos, E. A. Theodorou, and P. Tsiotras, "Stochastic l^1-optimal control via forward and backward sampling," *Systems & Control Letters*, vol. 118, pp. 101–108, 2018.

[16] S. Sukumar and D. Chatterjee, "A jammer's perspective of reachability and LQ optimal control," *Automatica*, vol. 70, pp. 295–302, 2016.

[17] D. Chatterjee, M. Nagahara, D. E. Quevedo, and K. M. Rao, "Characterization of maximum hands-off control," *Systems & Control Letters*, vol. 94, pp. 31–36, 2016.

[18] G. Reissig and M. Rungger, "Symbolic optimal control," *IEEE Transactions on Automatic Control*, 2019. [Online]. Available: https://doi.org/10.1109/TAC.2018.2863178

[19] M. Bardi and I. Capuzzo-Dolcetta, *Optimal control and viscosity solutions of Hamilton-Jacobi-Bellman equations*. Springer Science & Business Media, 2008.