



# On the Existence of Lipschitz Continuous Optimal Feedback Control

A. L. Dontchev<sup>1,2</sup> · M. I. Krastanov<sup>3,4</sup> · V. M. Veliov<sup>5</sup>

Received: 19 June 2018 / Accepted: 11 January 2019 / Published online: 29 April 2019  
© The Author(s) 2019

## Abstract

We consider an optimal control problem involving a nonlinear ODE with control, an integral cost functional, and a control constraint. Our main assumptions include a coercivity condition and the condition that the optimal control is an isolated solution of the variational inequality appearing in the first-order optimality condition. We show that the optimal open-loop control is Lipschitz continuous in time; moreover, we identify the dependence of the Lipschitz constant of the optimal control on the data of the problem. Then, we establish the existence of a Lipschitz continuous optimal feedback control. As an application, we study regularity properties of the optimal value function. A main tool for obtaining these results is the property of uniform strong metric regularity.

**Keywords** Optimal control · Optimal feedback control · Lipschitz continuity · Value function

**Mathematics Subject Classification (2010)** 49N60 · 49K40 · 49J30

## 1 Introduction

In this paper, we consider an optimal control problem for a time-dependent nonlinear control system over a fixed time interval  $[0, T]$  with an integral cost functional. The set of feasible

---

Dedicated to Alexander Ioffe.

✉ V. M. Veliov  
vladimir.veliov@tuwien.ac.at

A. L. Dontchev  
dontchev@umich.edu

M. I. Krastanov  
krastanov@fmi.uni-sofia.bg

<sup>1</sup> American Mathematical Society, Providence, RI, USA

<sup>2</sup> The University of Michigan, Ann Arbor, MI 48109, USA

<sup>3</sup> Faculty of Mathematics and Informatics, University of Sofia, Sofia, Bulgaria

<sup>4</sup> Institute of Mathematics and Informatics, Bulgarian Academy of Sciences, Sofia, Bulgaria

<sup>5</sup> Institute of Statistics and Mathematical Methods in Economics, Vienna University of Technology, Vienna, Austria

controls consists of all functions in  $L^\infty$  (the space of measurable and essentially bounded functions over  $[0, T]$ ) with values in a given convex and closed set in  $\mathbb{R}^m$ . We assume twice differentiability with respect to the state and the control of the functions involved in the problem and local Lipschitz continuity of these functions together with all their derivatives with respect to all arguments. We also assume the existence of a reference optimal solution. Since the reference optimal control is a function in  $L^\infty$ , its values can be changed in a subset of  $[0, T]$  with Lebesgue measure zero without violating the optimality. In fact, the optimal control is a class of functions that differ from each other on a set of measure zero.

Our first task is to prove that, under an integral coercivity condition at the reference solution, we can select from the class of optimal controls a function which satisfies the first-order optimality condition for all  $t \in [0, T]$ , instead of for almost every (a.e.)  $t \in [0, T]$ . Then, we show that under a coercivity condition this representative of the optimal controls is Lipschitz continuous with respect to time  $t \in [0, T]$  provided that it is an *isolated* solution of the Hamiltonian variational inequality in the first-order optimality condition. Moreover, we establish that the Lipschitz constant of the optimal control depends only on two constants: the coercivity constant and the Lipschitz constant of all functions defining the problem and their first and second derivatives over a bounded set in the space of variables (time, state, control).

The integral coercivity condition is a rather standard assumption in optimal control; the specific condition we use here goes back to the work of Hager [7]. In contrast, the isolatedness condition was introduced only recently in [2, Definition 3.6] in the context of the so-called differential variational inequalities, with the aim to prevent different solution curves from crossing each other. The isolatedness assumption is automatically satisfied when the Hamiltonian has a unique minimizer for each  $t \in [0, T]$ , e.g., when the Hamiltonian is strictly convex. In [2, Theorem 4.1], it was established that if an optimal control  $\bar{u}$  is an isolated solution of the Hamiltonian variational inequality and for each  $t \in [0, T]$  the mapping defining this variational inequality is strongly metrically regular at  $\bar{u}$  for 0; then, the optimal control  $\bar{u}$  is Lipschitz continuous on  $[0, T]$ . We also mention the earlier work [4] in that direction for an optimal control problem with linear dynamics and a strongly convex cost for which strong regularity holds automatically; in fact, only continuity of the optimal control is claimed there but the Lipschitz continuity can be gleaned from the proof. We note that the coercivity condition implies strong metric regularity in the respective function spaces, see [2, Theorem 4.2].

Our next task is to prove the existence of a Lipschitz continuous optimal feedback control. We show that under the coercivity and isolatedness conditions for the optimal control, there exists an optimal feedback control  $(\tau, \xi) \mapsto u^*(\tau, \xi)$  which is a Lipschitz continuous function; here  $(\tau, \xi)$  represents the parametrizing pair initial time–initial condition.

Our third and last task is to show that the existence of a Lipschitz continuous optimal feedback control implies that the optimal value function  $(\tau, \xi) \mapsto V(\tau, \xi)$  is differentiable with respect to  $\xi$  and its derivative is Lipschitz continuous.

An outline of the paper follows. In Section 2, we introduce the optimal control problem considered and set the stage for the further developments. Section 3 contains preliminary material showing in particular that the optimal control can be redefined on a set of measure zero so that the first-order optimality system holds for all  $t \in [0, T]$ . Section 4 gives conditions for Lipschitz continuity in time of the optimal open-loop control while Section 5 is devoted to the existence of a Lipschitz continuous optimal feedback control. The last Section 6 applies the latter result to show Lipschitz differentiability of the value function.

## 2 The Optimal Control Problem

We consider the following optimal control problem:

$$\min \left\{ J(u) := g(x(T)) + \int_0^T h(t, x(t), u(t))dt \right\}, \tag{1}$$

subject to

$$\begin{aligned} \dot{x}(t) &= f(t, x(t), u(t)), & x(0) &= x_0, \\ x &\in W^{1,\infty}, & u \in \mathcal{U} &:= \{u \in L^\infty : u(t) \in U \text{ for a.e. } t \in [0, T]\}, \end{aligned} \tag{2}$$

where the state  $x(t) \in \mathbb{R}^n$ , the set  $U$  of feasible control values is a closed and convex subset of  $\mathbb{R}^m$ , and the functions  $g : \mathbb{R}^n \rightarrow \mathbb{R}$ ,  $h : [0, T] \times \mathbb{R}^n \times \mathbb{R}^m \rightarrow \mathbb{R}$  and  $f : [0, T] \times \mathbb{R}^n \times \mathbb{R}^m \rightarrow \mathbb{R}^n$ . The final time  $T$  and the initial state  $x_0$  are fixed.

Throughout we assume that the function  $g$  is twice differentiable and its second derivative is locally Lipschitz continuous, the functions  $h(t, \cdot, \cdot)$  and  $f(t, \cdot, \cdot)$  are two times continuously differentiable (with respect to  $(x, u)$ ), and these functions, together with all their derivatives, are locally Lipschitz continuous (with respect to  $(t, x, u)$ ).

We also assume that problem (1)–(2) has a locally optimal solution  $(\bar{x}, \bar{u})$ . The local optimality is understood in the following way: there exists a number  $e_0 > 0$  such that for every  $u \in \mathcal{U}$  with  $\|u - \bar{u}\|_\infty \leq e_0$  either there is no solution of (2) over  $[0, T]$  or such a solution exists and  $J(u) \geq J(\bar{u})$ .

In this paper, we employ the standard function spaces  $L^\infty, L^2, W^{1,\infty}, W^{1,2}$ , all over  $[0, T]$ . Specifically, the space of controls  $u$  is  $L^\infty$ , the space of measurable and essentially bounded functions. The state trajectory  $x$  is in  $W^{1,\infty}$ , the space of Lipschitz continuous functions. For the controls we also use the space  $L^2$  of measurable square integrable functions, and for the state trajectory  $x$  the space  $W^{1,2}$  such that both  $x$  and its derivative  $\dot{x}$  are in  $L^2$ . Furthermore, for an element  $x$  of a metric space we denote by  $\mathbb{B}_a(x)$  (respectively  $\overset{\circ}{\mathbb{B}}_a(y)$ ) the closed (respectively open) ball centered at  $x$  with radius  $a$ .

Clearly, any feasible control  $u$  is actually a class of functions which differ from each other on a set of Lebesgue measure zero. We call any particular function from this class a *representative* and denote it in the same way, by  $u$ .

Introducing the Hamiltonian  $H(t, x, u, \lambda) = h(t, x, u) + \lambda^\top f(t, x, u)$ , where  $^\top$  means transposition, we employ the standard first-order necessary optimality condition (a consequence of the Pontryagin maximum principle) in the form used, e.g., in [7], according to which there exists a Lipschitz continuous function  $\bar{\lambda} : [0, T] \rightarrow \mathbb{R}^n$  such that the triple  $(\bar{x}, \bar{u}, \bar{\lambda})$  satisfies for a.e.  $t \in [0, T]$  the following optimality system:

$$\begin{aligned} -\dot{x}(t) + f(t, x(t), u(t)) &= 0, & x(0) - x_0 &= 0, \\ \dot{\lambda}(t) + H_x(t, x(t), u(t), \lambda(t)) &= 0, & \lambda(T) - g_x(x(T)) &= 0, \\ H_u(t, x(t), u(t), \lambda(t)) + N_U(u(t)) &\ni 0, \end{aligned} \tag{3}$$

where  $H_x$  denotes the derivative of  $H$  with respect to  $x$ , etc., and  $N_U$  is the normal cone mapping to the set  $U$  defined as

$$u \mapsto N_U(u) = \begin{cases} \{y \in \mathbb{R}^n \mid \langle y, v - u \rangle \leq 0 \text{ for all } v \in U\} & \text{if } u \in U, \\ \emptyset & \text{otherwise.} \end{cases}$$

In further lines, we give the following long but important remark, which summarizes various observations that will be used later on.

*Remark 1* It is a standard fact that under our assumptions there exist positive reals  $d_0$  and  $d$  such that for every  $\tilde{u} \in \mathcal{U}$  with  $\|\tilde{u} - \bar{u}\|_\infty \leq d$  and for every  $\xi \in B_{d_0}(x_0)$  there exists a unique solution  $\tilde{x}$  of the differential equation

$$\dot{\tilde{x}}(t) = f(t, x(t), \tilde{u}(t)) \quad \text{for a.e. } t \in [0, T], \quad x(0) = \xi, \tag{4}$$

which satisfies  $\|\tilde{x} - \bar{x}\|_{W^{1,\infty}} \leq 1$ . Moreover, making  $d_0$  and  $d$  smaller if necessary, we obtain that the (unique) solution  $\tilde{\lambda}$  of the linear adjoint equation

$$\dot{\tilde{\lambda}}(t) + H_x(t, \tilde{x}(t), \tilde{u}(t), \tilde{\lambda}(t)) = 0 \quad \text{for a.e. } t \in [0, T], \quad \tilde{\lambda}(T) = g_x(\tilde{x}(T)) \tag{5}$$

satisfies  $\|\tilde{\lambda} - \bar{\lambda}\|_{W^{1,\infty}} \leq 1$ . Without loss of generality, we assume that  $d \leq 1$  and  $d \leq e_0$ , where  $e_0$  appears in the definition of local optimality given in the beginning of this section.

Since  $\bar{u} \in L^\infty$ , there exists a compact set  $\bar{U}$  such that  $\bar{u}(t) \in \bar{U}$  for a.e.  $t \in [0, T]$ . Define the set

$$\Omega = \{(t, x, u, \lambda) : t \in [0, T], \text{dist}(u, \bar{U}) \leq 1, |x - \bar{x}(t)| \leq 1, |\lambda - \bar{\lambda}(t)| \leq 1\}.$$

Denote by  $L$  the Lipschitz constant on  $\Omega$  of each of the functions  $f, g, h, f_x, f_u, h_x, h_u, f_{xx}, f_{xu}, f_{uu}, h_{xx}, h_{xu}, h_{uu}$ , as well as of the functions  $H, H_x, H_u, H_{xx}, H_{xu}, H_{uu}$ . Since  $f$  and  $H_x$  are bounded in  $\Omega$ , then  $\dot{\tilde{x}}$  and  $\dot{\tilde{\lambda}}$  are also bounded. Make  $L$  larger if needed so that for every  $\tilde{x}$  and  $\tilde{\lambda}$  that satisfy (4) and (5), respectively, the functions  $(t, v) \mapsto H_u(t, \tilde{x}(t), v, \tilde{\lambda}(t))$  and  $(t, v) \mapsto H_{uu}(t, \tilde{x}(t), v, \tilde{\lambda}(t))$  are Lipschitz continuous with constant  $L$  in the set  $\{(t, v) : t \in [0, T], \text{dist}(v, \bar{U}) \leq 1\}$ . This concludes Remark 1.

To shorten the notations, we skip arguments with ‘‘bar’’, shifting the ‘‘bar’’ to the functions, e.g.,  $\bar{H}(t) := H(t, \bar{x}(t), \bar{u}(t), \bar{\lambda}(t))$ ,  $\bar{H}(t, u) := H(t, \bar{x}(t), u, \bar{\lambda}(t))$ ,  $\bar{f}(t) := f(t, \bar{x}(t), \bar{u}(t))$ ,  $\bar{g}_{xx} := g_{xx}(\bar{x}(T))$ , etc. Define the matrices

$$A(t) = \bar{f}_x(t), \quad B(t) = \bar{f}_u(t), \quad Q(t) = \bar{H}_{xx}(t), \quad S(t) = \bar{H}_{xu}(t), \quad R(t) = \bar{H}_{uu}(t).$$

Our first main assumption is the following:

- COERCIVITY: there exists a constant  $\rho > 0$  such that

$$\begin{aligned} & y(T)^\top \bar{g}_{xx} y(T) + \int_0^T \left( y(t)^\top Q(t) y(t) + w(t)^\top R(t) w(t) + 2y(t)^\top S(t) w(t) \right) dt \\ & \geq \rho \int_0^T |w(t)|^2 dt \end{aligned} \tag{6}$$

for all  $w \in L^2, y \in W^{1,2}$  such that  $w(t) \in U - U, \dot{y}(t) = A(t)y(t) + B(t)w(t)$  for a.e.  $t \in [0, T]$ , and  $y(0) = 0$ .

The coercivity condition was first used in [7] to show convergence of the multiplier method and later in [5] to establish Lipschitz stability as well as convergence of discrete approximations in optimal control. It can be viewed as a strong second-order sufficient condition in optimal control. Checking this condition would very much depend on the specific problem at hand; sometimes it is enforced numerically by adding penalty terms to the cost. The coercivity condition has also been used for a posteriori numerical verification of optimality after an approximate solution is found.

In the following section, we present some preparatory material. In particular, we show that the coercivity condition implies a pointwise in time coercivity property which plays an important role in further analysis.

### 3 Preliminaries

Denote by  $\text{meas}(E)$  the Lebesgue measure of a set  $E$ . Let  $\Delta \subset [0, T]$  be a measurable set with  $\text{meas}(\Delta) > 0$ , and let  $v : \Delta \rightarrow \mathbb{R}^m$  be a measurable and bounded function. For  $t \in \Delta$  denote by  $V_\Delta(v; t)$  the set of points  $w \in \mathbb{R}^m$  with the following property: there is a sequence of measurable sets  $E_k \subset \Delta$  such that

$$\text{meas}(E_k) > 0, \quad E_k \subset [t - 1/k, t + 1/k], \quad \lim_{k \rightarrow \infty} \sup_{s \in E_k} |v(s) - w| = 0.$$

We denote  $V_{[0, T]}(v; t)$  simply by  $V(v; t)$ .

A point  $t \in \Delta$  is said to be *essentially non-isolated* if for every  $\varepsilon > 0$  the set  $[t - \varepsilon, t + \varepsilon] \cap \Delta$  is of positive measure.

**Lemma 1** *Let  $\Delta \subset [0, T]$  be a measurable set and let  $v : \Delta \rightarrow \mathbb{R}^m$  be a measurable and bounded function. Then, for any  $t \in \Delta$ , the following statements are equivalent:*

- (i)  $V_\Delta(v; t) \neq \emptyset$ ;
- (ii)  $t$  is essentially non-isolated point of  $\Delta$ .

*Proof* If (i) holds, then the very definition of  $V_\Delta(v; t)$  implies that  $t$  is essentially non-isolated.

Let us pick an essentially non-isolated point  $t$  of  $\Delta$ . Let  $K \subset \mathbb{R}^m$  be a compact set such that  $v(s) \in K$  for every  $s \in \Delta$ . Take an arbitrary  $w \in K$ . If for every  $\varepsilon > 0$  and every natural number  $k$  there exists  $E_k \subset [t - 1/k, t + 1/k] \cap \Delta$  such that  $\text{meas}(E_k) > 0$  and  $\sup_{s \in E_k} |v(s) - w| < \varepsilon$ , then  $w \in V_\Delta(v; t)$ . If this is not the case, then there exist  $\varepsilon(w) > 0$  and a natural number  $k(w)$  such that  $|v(s) - w| \geq \varepsilon(w)$  for a.e.  $s \in [t - 1/k(w), t + 1/k(w)] \cap \Delta$ ; that is,  $v(s) \notin \overset{\circ}{B}_{\varepsilon(w)}(w)$  for a.e.  $s \in [t - 1/k(w), t + 1/k(w)] \cap \Delta$ . If  $w \notin V_\Delta(v; t)$  for every  $w \in K$ , then, due to the compactness of  $K$ , there exist  $w_1, \dots, w_r \in K$  such that  $K \subset \bigcup_{i=1}^r \overset{\circ}{B}_{\varepsilon(w_i)}(w_i)$ . Denote  $\bar{k} := \max\{k(w_1), \dots, k(w_r)\}$ ; then,  $v(s) \notin \bigcup_{i=1}^r \overset{\circ}{B}_{\varepsilon(w_i)}(w_i)$  for a.e.  $s \in [t - 1/\bar{k}, t + 1/\bar{k}] \cap \Delta$ . This contradicts the essential non-isolatedness of  $t$ , since  $K \subset \bigcup_{i=1}^r \overset{\circ}{B}_{\varepsilon(w_i)}(w_i)$  and  $\text{meas}([t - 1/\bar{k}, t + 1/\bar{k}] \cap \Delta) > 0$ . Hence  $V_\Delta(v; t) \neq \emptyset$  and the proof is complete.  $\square$

Taking  $\Delta = [0, T]$ , we obtain that  $V(v; t)$  is non-empty for every  $t \in [0, T]$ .

**Lemma 2** *Let  $u$  and  $\tilde{u}$  be two measurable and bounded functions acting from  $[0, T]$  to  $\mathbb{R}^m$ , and let  $u(t) \in V(u; t)$  for every  $t \in [0, T]$ . Then, the function  $\tilde{u}$  can be redefined on a set of measure zero in such a way that  $\tilde{u}(t) \in V(\tilde{u}; t)$  and  $|\tilde{u}(t) - u(t)| \leq \|\tilde{u} - u\|_\infty$  for every  $t \in [0, T]$ .*

*Proof* Take an arbitrary  $t \in [0, T]$ . Consider first the case where both functions  $u$  and  $\tilde{u}$  are approximately continuous at  $t$ . We recall that  $u$  is approximately continuous at  $t \in (0, T)$  if there exists a measurable set  $E \subset [0, T]$  containing  $t$  such that

$$\lim_{k \rightarrow \infty} 2k \text{meas}(E \cap [t - 1/k, t + 1/k]) = 1$$

and the restriction of  $u$  to  $E$  is continuous. Let  $\tilde{E}$  be the set in the definition of approximate continuity of  $\tilde{u}$  at  $t \in (0, T)$ . Then, the set  $E'_k := E \cap \tilde{E} \cap [t - 1/k, t + 1/k]$  satisfies

$\lim_{k \rightarrow \infty} 2k \operatorname{meas}(E'_k) = 1$ . In particular,  $\operatorname{meas}(E'_k) > 0$  for all sufficiently large  $k$ . Due to the continuity of  $u$  and  $\tilde{u}$  on  $E \cap \tilde{E}$ , we have

$$|\tilde{u}(t) - u(t)| \leq \lim_k \left| \frac{1}{\operatorname{meas}(E'_k)} \int_{E'_k} (\tilde{u}(s) - u(s)) \, ds \right| \leq \|\tilde{u} - u\|_\infty.$$

Moreover, since the sets  $E_k$  in the definition of  $V$  can be replaced by  $E'_k$ , we conclude that  $\tilde{u}(t) \in V(\tilde{u}; t)$ .

Now, let  $t \in [0, T]$  be such that  $u$  or  $\tilde{u}$  is not approximately continuous at  $t$ , or  $t$  equals 0 or  $T$ .

We will now redefine  $\tilde{u}(t)$  to fit the claim. It is well known (see, e.g., [8, Theorem 7.54]) that almost all  $t \in [0, T]$  are points of approximate continuity of both  $u$  and  $\tilde{u}$ ; therefore we need to redefine  $\tilde{u}$  only on a set of measure zero. Note that the sets  $V(\tilde{u}; t)$  are invariant with respect to changes of  $\tilde{u}$  on a set of measure zero.

Denote  $w := u(t) \in V(u; t)$ . Let  $E_k$  be the sets in the definition of  $V$ . In particular,  $\varepsilon_k := \sup_{s \in E_k} |u(s) - w| \xrightarrow{k} 0$ . Since  $E_k$  is of positive measure, it contains an essentially non-isolated point  $t_k \in E_k$ . According to Lemma 1, there exists  $\tilde{w}_k \in V_{E_k}(\tilde{u}; t_k)$ ; hence, there exists a sequence  $\{E_k^i\}$ ,  $E_k^i \subset E_k$ , such that

$$\operatorname{meas}(E_k^i) > 0, \quad E_k^i \subset [t_k - 1/i, t_k + 1/i], \quad \varepsilon_k^i := \sup_{s \in E_k^i} |\tilde{u}(s) - \tilde{w}_k| \xrightarrow{i} 0.$$

Let  $\tilde{w}$  be a cluster point of the sequence  $\{\tilde{w}_k\}$ . To show that  $\tilde{w} \in V(\tilde{u}; t)$ , we employ the following argument involving choosing a diagonal sequence. For an arbitrary natural number  $j$ , choose  $k = k_j$  so large that

$$|t_{k_j} - t| \leq \frac{1}{2j} \quad \text{and} \quad |\tilde{w}_{k_j} - \tilde{w}| \leq \frac{1}{j}.$$

Then, choose  $i = i_j$  such that

$$\frac{1}{i_j} \leq \frac{1}{2j} \quad \text{and} \quad \varepsilon_{k_j}^{i_j} \leq \frac{1}{j}.$$

We have

$$\begin{aligned} \tilde{E}_j := E_{k_j}^{i_j} &\subset [t_{k_j} - 1/i_j, t_{k_j} + 1/i_j] \subset [t - |t - t_{k_j}| - 1/i_j, t + |t - t_{k_j}| + 1/i_j] \\ &\subset [t - 1/j, t + 1/j] \end{aligned}$$

and

$$\sup_{s \in \tilde{E}_j} |\tilde{u}(s) - \tilde{w}| \leq |\tilde{w}_{k_j} - \tilde{w}| + \sup_{s \in \tilde{E}_j} |\tilde{u}(s) - \tilde{w}_{k_j}| \leq \frac{1}{j} + \varepsilon_{k_j}^{i_j} \leq \frac{2}{j}.$$

Taking also into account that  $\operatorname{meas}(\tilde{E}_j) > 0$ , the last two relations imply that  $\tilde{w} \in V(\tilde{u}; t)$ .

For every  $k$  and  $i$ , we have  $E_k^i \subset E_k$ ,

$$\left| \frac{1}{\operatorname{meas}(E_k^i)} \int_{E_k^i} u(s) \, ds - w \right| = \left| \frac{1}{\operatorname{meas}(E_k^i)} \int_{E_k^i} (u(s) - w) \, ds \right| \leq \varepsilon_k,$$

and

$$\left| \frac{1}{\operatorname{meas}(E_k^i)} \int_{E_k^i} \tilde{u}(s) \, ds - \tilde{w}_k \right| = \left| \frac{1}{\operatorname{meas}(E_k^i)} \int_{E_k^i} (\tilde{u}(s) - \tilde{w}_k) \, ds \right| \leq \varepsilon_k^i.$$

Hence,

$$|\tilde{w}_k - w| \leq \left| \frac{1}{\text{meas}(E_k^i)} \int_{E_k^i} (\tilde{u}(s) - u(s)) ds \right| + \varepsilon_k + \varepsilon_k^i \leq \|\tilde{u} - u\|_\infty + \varepsilon_k + \varepsilon_k^i.$$

Passing to the limit with  $i$  and then with  $k$ , we obtain  $|\tilde{w} - w| \leq \|\tilde{u} - u\|_\infty$ . Then, we redefine  $\tilde{u}(t)$  as  $\tilde{u}(t) = \tilde{w}$ . This completes the proof.  $\square$

**Corollary 1** *Every  $v \in U$  can be redefined on a set of measure zero in such a way that  $v(t) \in V(v; t)$  for every  $t \in [0, T]$ .*

For a proof, apply Lemma 2 with  $\tilde{u} = v$  and the constant function  $u(t) = u$  for all  $t \in [0, T]$ .

*Remark 2* From now on, the element  $\bar{u} \in L^\infty$  will be identified with a function (denoted again by  $\bar{u}$ ) satisfying  $\bar{u}(t) \in V(\bar{u}; t)$  for every  $t \in [0, T]$ .

Observe that the coercivity condition (6) does not depend on the particular representative of  $\bar{u}$ .

**Lemma 3** *Let the coercivity condition (6) hold, where  $\bar{u}$  is identified as in Remark 2. Then,*

$$w^\top R(t)w \geq \rho|w|^2 \quad \text{for every } t \in [0, T] \text{ and } w \in U - U. \tag{7}$$

*Proof* Fix an arbitrary  $t \in [0, T]$ . Since  $\bar{u}(t) \in V(\bar{u}; t)$ , there exists a sequence  $E_k \subset [0, T]$  such that

$$\text{meas}(E_k) > 0, \quad E_k \subset [t - 1/k, t + 1/k], \quad \varepsilon_k := \sup_{s \in E_k} |\bar{u}(s) - \bar{u}(t)| \rightarrow 0. \tag{8}$$

For an arbitrary  $w \in U - U$ , we define a function  $w_k$  as

$$w_k(s) = \begin{cases} w & \text{if } s \in E_k, \\ 0 & \text{if } s \notin E_k. \end{cases}$$

Using the Cauchy formula for the equation

$$\dot{y}_k(s) = A(s)y_k(s) + B(s)w_k(s) \quad \text{for a.e. } s \in [0, T], \quad y(0) = 0,$$

we obtain that  $y_k(s) = 0$  for  $s \in [0, t - 1/k]$  and  $|y_k(s)| \leq c_1 \text{meas}(E_k)$  for  $s \in (t - 1/k, T]$ , where here and further  $c_1, c_2, \dots$  are positive reals independent of  $k$ . Then, for the terms involved in (6), we have

$$\begin{aligned} & \left| y_k(T)^\top g_{xx} y_k(T) \right| + \left| \int_0^T y_k(s)^\top Q(s) y_k(s) ds \right| \leq c_2 (\text{meas}(E_k))^2, \\ & \left| \int_0^T y_k(s)^\top S(s) w_k(s) ds \right| = \left| \int_{E_k} y_k(s)^\top S(s) ds w \right| \leq c_3 (\text{meas}(E_k))^2, \\ & \int_0^T w_k(s)^\top R(s) w_k(s) ds = w^\top \int_{E_k} R(s) ds w, \\ & \int_0^T |w_k(s)|^2 ds = \text{meas}(E_k) |w|^2. \end{aligned}$$

Since  $R(s) = \bar{H}_{uu}(s, \bar{u}(s))$ , using (8), we obtain (see Remark 1) that for  $s \in E_k$  one has

$$|R(s) - R(t)| \leq L(|s - t| + \varepsilon_k) \leq L(1/k + \varepsilon_k) =: \tilde{\varepsilon}_k \rightarrow 0.$$

Using the above estimated in (6) and the above five displayed formulas, we obtain

$$\text{meas}(E_k) w^\top R(t) w \geq \rho \text{meas}(E_k) |w|^2 - c_4 (\text{meas}(E_k))^2 - c_5 \text{meas}(E_k) \tilde{\varepsilon}_k.$$

Dividing by  $\text{meas}(E_k)$  (here we use the first inequality in (8)) and passing to the limit with  $k$ , we obtain (7). □

### 4 Lipschitz Continuity of the Optimal Control

Let us recall the optimality system (3):

$$\begin{aligned} -\dot{x}(t) + f(t, x(t), u(t)) &= 0, \\ x(0) - x_0 &= 0, \\ \dot{\lambda}(t) + H_x(t, x(t), u(t), \lambda(t)) &= 0, \\ \lambda(T) - g_x(x(T)) &= 0, \\ H_u(t, x(t), u(t), \lambda(t)) + N_U(u(t)) &\ni 0. \end{aligned} \tag{9}$$

**Lemma 4** *Let the coercivity condition hold. Then, the optimal control  $\bar{u} \in L^\infty$  has a representative  $\tilde{u}$  such that the matrix  $R(t) = \bar{H}_{uu}(t, \tilde{u}(t))$  satisfies (7) and  $(\bar{x}(t), \tilde{u}(t), \bar{\lambda}(t))$  satisfies (9) for all  $t \in [0, T]$ . In fact, any representative of the optimal control that satisfies  $\tilde{u}(t) \in V(\bar{u}; t)$  for all  $t \in [0, T]$  has this property.*

*Proof* Let us redefine  $\bar{u}$  so that  $\tilde{u}(t) \in V(\bar{u}; t)$  for all  $t \in [0, T]$  (see Corollary 1 and Remark 2). Then, according to Lemma 3, the pointwise coercivity condition (7) holds for every  $t \in [0, T]$ .

Fix an arbitrary  $t \in [0, T]$ . Since  $\tilde{u}(t) \in V(\bar{u}; t)$ , there exists a sequence  $\{E_k\}$  of measurable subsets of  $[0, T]$  such that (8) holds. Since  $\text{meas}(E_k) > 0$  and (9) is satisfied by  $(\bar{x}(t), \tilde{u}(t), \bar{\lambda}(t))$  almost everywhere, there exists  $t_k \in E_k$  such that (9) holds for  $t_k$ . From (8), we obtain that  $t_k \rightarrow t$  and  $\tilde{u}(t_k) \rightarrow \tilde{u}(t)$ . Then, due to the continuity of the function  $(t, u) \mapsto H_u(t, \bar{x}(t), u, \bar{\lambda}(t))$  and the upper semi-continuity of the mapping  $u \mapsto N_U(u)$ , (9) holds for  $t$  as well. □

We recall next the property of *strong metric regularity* of a general set-valued mapping  $\mathcal{F} : \mathcal{Y} \rightrightarrows \mathcal{Z}$ , where  $\mathcal{Y}$  and  $\mathcal{Z}$  are Banach spaces (for more on that, see, e.g., [6, Section 3.7]). A mapping  $\mathcal{F}$  is said to be *strongly metrically regular* at  $\hat{y}$  for  $\hat{z}$  if there exist constants  $\kappa \geq 0$ ,  $a > 0$  and  $b > 0$  such that the truncated inverse mapping

$$IB_b(\hat{z}) \ni z \mapsto \mathcal{F}^{-1}(z) \cap IB_a(\hat{y})$$

is single-valued (a function) and Lipschitz continuous on  $IB_b(\hat{z})$ . Here  $\mathcal{F}^{-1}(z) = \{y \mid z \in \mathcal{F}(y)\}$ .

Our further analysis is based on the following version of Robinson’s implicit function theorem. It was first stated as [6, Theorem 5G.3]<sup>1</sup> and then in corrected form as Theorem 3.2 in [2] (see also [3, Theorem 2.3] for a slight extension):

<sup>1</sup>See Errata and Addenda at <https://sites.google.com/site/adontchev/>



**Theorem 1** *Let  $a, b,$  and  $\kappa$  be positive scalars and let a mapping  $\mathcal{F} : \mathcal{Y} \rightrightarrows \mathcal{Z}$  be strongly metrically regular at  $\hat{y}$  for  $\hat{z}$  with neighborhoods  $\mathcal{B}_a(\hat{y})$  and  $\mathcal{B}_b(\hat{z})$  and constant  $\kappa$ . Let  $\mu > 0$  be such that  $\kappa\mu < 1$  and let  $\kappa' > \kappa/(1 - \kappa\mu)$ . Then, for every positive  $\alpha$  and  $\beta$  such that*

$$\alpha \leq a/2, \quad 2\mu\alpha + 2\beta \leq b \quad \text{and} \quad 2\kappa'\beta \leq \alpha$$

and for every function  $g : \mathcal{Y} \rightarrow \mathcal{Z}$  satisfying

$$\|g(\hat{y})\| \leq \beta \quad \text{and} \quad \|g(y) - g(y')\| \leq \mu\|y - y'\| \quad \text{for every } y, y' \in \mathcal{B}_{2\alpha}(\hat{y}),$$

the mapping  $z \mapsto (g + \mathcal{F})^{-1}(z) \cap \mathcal{B}_a(\hat{y})$  is a Lipschitz continuous function on  $\mathcal{B}_\beta(\hat{z})$  with Lipschitz constant  $\kappa'$ .

Compared with the standard Robinson’s implicit function theorem, see [6, Theorem 2B.1], Theorem 1 exhibits the fact that everything hinges on the constants involved; that is, the constants of metric regularity of the perturbed mapping  $g + \mathcal{F}$  do not depend on the actual perturbations but only on  $\|g(\hat{y})\|$ , the Lipschitz constant of  $g$  and the constants of the strong regularity of  $\mathcal{F}$ . In that sense, Theorem 1 shows strong metric regularity which is *uniform* with respect to perturbations.

Let us get back to the optimal control problem at hand. If  $(t, u) \in \text{cl gph}(\bar{u})$ , then there exists a sequence  $t_k \rightarrow t$  such that  $\bar{u}(t_k) \rightarrow u$ . According to (7), we have

$$w^\top \bar{H}_{uu}(t_k, \bar{u}(t_k)) w \geq \rho|w|^2 \quad \text{for every } w \in U - U.$$

Passing to the limit, we obtain that

$$w^\top \bar{H}_{uu}(t, u) w \geq \rho|w|^2 \tag{10}$$

for every  $(t, u) \in \text{cl gph}(\bar{u})$  and every  $w \in U - U$ . It is well known that the property (10) implies that for every  $(t, u) \in \text{cl gph}(\bar{u})$  the mapping

$$v \mapsto \bar{H}_u(t, u) + \bar{H}_{uu}(t, u)(v - u) + N_U(v) \tag{11}$$

is strongly metrically regular at  $u$  for  $0$  with constants  $\kappa' = 1/\rho, a' = b' = +\infty$  (that is, with any positive  $a'$  and  $b'$ ), see, e.g., [7, Lemma 1]. Note that these constants are independent of  $t$ .

Next, we reformulate, adapted to our notations and needs, a simplified version of Theorem 3.5 in [2], which in turn is a corollary of Theorem 1.

**Theorem 2** *Assume that for every  $(t, u) \in \text{cl gph}(\bar{u})$  the mapping in (11) is strongly metrically regular at  $u$  for  $0$  with constants  $\kappa', a', b'$  that are independent of  $(t, u)$ . Then, for every  $t \in [0, T]$ , the mapping  $u \mapsto \bar{H}_u(t, u) + N_U(u)$  is strongly metrically regular at  $\bar{u}(t)$  for  $0$  with any constants  $\kappa, a, b$  satisfying the inequalities*

$$a \leq \frac{a'}{2}, \quad 2L'ak' < 1, \quad 4L'a^2 < b', \quad \frac{\kappa'}{1 - 2L'ak'} < \kappa, \quad 4L'a^2 + 2b < b', \quad 2\kappa b < a, \tag{12}$$

where  $L'$  is a Lipschitz constant of the mapping  $u \mapsto \bar{H}_{uu}(t, u)$  on  $\mathcal{B}_{a'}(\bar{u}(t))$ , for every  $t \in [0, T]$ .

The conditions (12) are not stated in Theorem 3.5 in [2], but are explicitly written in the beginning of its proof there.

Continuing the analysis of (11), we apply Theorem 2 with  $a' = 1, b' = +\infty$  and  $\kappa' = 1/\rho$ , having the inequalities (12) reduced to

$$a \leq \frac{1}{2}, \quad 2La < \rho, \quad \kappa > \frac{1}{\rho - 2La}, \quad 2\kappa b < a, \tag{13}$$

where now  $L$  is the constant from Remark 1.

*Remark 3* The important consequence of (13) is that the constants  $\kappa, a, b$  of strong regularity of  $u \mapsto \bar{H}_u(t, u) + N_U(u)$  at  $\bar{u}(t)$  for 0 can be chosen to depend only on the constant  $\rho$  in the coercivity condition (6) and the constant  $L$  in Remark 1.

We introduce next our second main assumption:

- ISOLATEDNESS: The function  $\bar{u}$  (represented as in Lemma 4) is an isolated solution of the inclusion  $\bar{H}_u(t, u) + N_U(u) \ni 0$  for all  $t \in [0, T]$ , meaning that there exists a (relatively) open set  $\mathcal{O} \subset [0, T] \times \mathbb{R}^m$  such that

$$\{(t, u) \in [0, T] \times \mathbb{R}^m : \bar{H}_u(t, u) + N_U(u) \ni 0\} \cap \mathcal{O} = \text{gph}(\bar{u}). \tag{14}$$

For example, the isolatedness assumption holds if for every  $t \in [0, T]$  the inclusion  $\bar{H}_u(t, u) + N_U(u) \ni 0$  has a unique solution (which has to be  $\bar{u}(t)$ ). In this case, one can verify the isolatedness condition taking any (relatively) open set  $\mathcal{O} \subset [0, T] \times \mathbb{R}^m$  containing  $\text{gph}(\bar{u})$ .

**Theorem 3** *Suppose that the isolatedness assumption (14) and condition (7) hold. Then, the optimal control  $\bar{u}$  is Lipschitz continuous on  $[0, T]$ . Moreover, the Lipschitz constant of  $\bar{u}$  depends only on the number  $\rho$  in (7) and the constant  $L$  in Remark 1.*

*Proof* The proof is somewhat parallel to the proof of Theorem 3.7 in [2]. Here we use Theorem 2 and (13) instead of the more general Theorem 3.5 in [2] (used in the proof of Theorem 3.7 in [2]), which does not imply the second claim of Theorem 3.

As mentioned around (10), condition (7) implies that for every  $(t, u) \in \text{cl gph}(\bar{u})$  the mapping in (11) is strongly metrically regular at  $u$  for 0. Then, we can apply Theorem 2. Let the numbers  $a, b, \kappa$  be chosen to satisfy conditions (13), so that for every  $t \in [0, T]$  the mapping  $u \mapsto \bar{H}_u(t, u) + N_U(u)$  is strongly metrically regular at  $\bar{u}(t)$  for 0 (see Theorem 2). Let  $L$  be the constant in Remark 1; then, the mappings  $(t, u) \mapsto \bar{H}_u(t, u)$  and  $(t, u) \mapsto \bar{H}_{uu}(t, u)$  are Lipschitz continuous with constant  $L$  on the set  $\{(t, u) : t \in [0, T], u \in \mathcal{B}_a(\bar{u}(t))\}$ . Without loss of generality we consider  $\bar{u}$  as taking values in the set  $\bar{U}$  in Remark 1; we also recall that  $a \leq 1$ .

Take an arbitrary  $t \in [0, T]$ . Then, pick  $\alpha_t < a/2$  and then  $\gamma_t \in (0, 1)$  such that  $(\tau, v) \in \mathcal{O}$  for every  $\tau \in [t - \gamma_t, t + \gamma_t] \cap [0, T]$  and  $v \in \mathcal{B}_{\alpha_t}(\bar{u}(t))$ , and also

$$\kappa L \gamma_t < 1/2, \quad L(a + 2)\gamma_t \leq b, \quad 4\kappa L \gamma_t \leq \alpha_t(1 - \kappa L \gamma_t). \tag{15}$$

For an arbitrary  $\tau \in [t - \gamma_t, t + \gamma_t] \cap [0, T]$  define the mapping  $g_{\tau,t} : U \rightarrow \mathbb{R}^m$  as

$$g_{\tau,t}(u) = \bar{H}_u(\tau, u) - \bar{H}_u(t, u), \quad u \in U.$$

Then, we have that

$$|g_{\tau,t}(\bar{u}(t))| \leq L|\tau - t| \leq L\gamma_t, \tag{16}$$

and, for any  $u, u' \in \mathcal{B}_a(\bar{u}(t))$ ,

$$\begin{aligned} |g_{\tau,t}(u) - g_{\tau,t}(u')| &= |\bar{H}_u(\tau, u) - \bar{H}_u(\tau, u') - \bar{H}_u(t, u) + \bar{H}_u(t, u')| \\ &\leq \int_0^1 |\bar{H}_{uu}(\tau, u' + s(u - u')) - \bar{H}_{uu}(t, u' + s(u - u'))| ds |u - u'| \\ &\leq L\gamma_t |u - u'|. \end{aligned}$$

Set

$$\kappa'_t = 2\kappa/(1 - \kappa L\gamma_t), \quad \beta_t = \mu_t = L\gamma_t.$$

According to the inequality  $\alpha_t < a/2$ , the inequalities in (15), and the definitions of  $\kappa'_t, \alpha_t, \beta_t$ , the following inequalities are fulfilled (for convenience we skip the subscripts  $t$  for a moment):

$$\mu > 0, \quad \kappa\mu < 1, \quad \kappa' > \kappa/(1 - \kappa\mu), \quad \alpha \leq a/2, \quad 2\mu\alpha + 2\beta \leq b, \quad 2\kappa'\beta \leq \alpha. \tag{17}$$

Now, we apply Theorem 1. For short, denote  $G_t(u) := \bar{H}_u(t, u) + N_U(u)$ . In our context all assumptions of the last theorem are satisfied with  $g = g_{\tau,t}$ . Thus we obtain that the mapping

$$\mathcal{B}_{\beta_t}(0) \ni z \mapsto (g_{\tau,t} + G_t)^{-1}(z) \cap \mathcal{B}_{\alpha_t}(\bar{u}(t)) = (G_\tau)^{-1}(z) \cap \mathcal{B}_{\alpha_t}(\bar{u}(t))$$

is Lipschitz continuous with Lipschitz constant  $\kappa'_t = 2\kappa/(1 - \kappa L\gamma_t) \leq 4\kappa$  (see the first inequality in (15)). In particular, there exists a unique  $v \in \mathcal{B}_{\alpha_t}(\bar{u}(t))$  such that  $0 \in G_\tau(v)$ . Since  $\tau \in [t - \gamma_t, t + \gamma_t] \cap [0, T]$  and  $v \in \mathcal{B}_{\alpha_t}(\bar{u}(t))$ , we also have that  $(\tau, v) \in \mathcal{C}$ . Due to isolatedness condition, we obtain that  $v = \bar{u}(\tau)$ . From (16), we obtain that  $g_{\tau,t}(\bar{u}(t)) \in \mathcal{B}_{\beta_t}(0)$ . Thus

$$\bar{u}(t) = (g_{\tau,t} + G_t)^{-1}(g_{\tau,t}(\bar{u}(t))) \cap \mathcal{B}_{\alpha_t}(\bar{u}(t)).$$

Since  $\bar{u}(\tau) = (g_{\tau,t} + G_t)^{-1}(0) \cap \mathcal{B}_{\alpha_t}(\bar{u}(t))$ , using (16), we get that

$$|\bar{u}(t) - \bar{u}(\tau)| \leq \kappa' |g_{\tau,t}(\bar{u}(t))| \leq 4\kappa L |t - \tau|.$$

Summarizing, we obtain that for every  $t \in [0, T]$  there exists a neighborhood  $(t - \gamma_t, t + \gamma_t) \cap [0, T]$  in  $[0, T]$  in which  $\bar{u}$  is Lipschitz continuous with the same constant  $4\kappa L$ . This implies that  $\bar{u}$  is Lipschitz continuous with the same constant in the whole interval  $[0, T]$ .

The second claim of the theorem follows from Remark 3 concerning  $\kappa$ . □

The example displayed in Remark 9 in [5] demonstrates that the isolatedness assumption (14) is essential for the Lipschitz continuity of the optimal control shown in Theorem 3. In this example  $h = (u^2 - 1)^2, g = 0, f = 0, U = \mathbb{R}, T = 1$ . Here, for each measurable set  $\Omega \subset [0, 1]$  the function defined as  $u(t) = -1$  for  $t \in \Omega$  and  $u(t) = 1$  for  $t \in [0, 1] \setminus \Omega$  is an optimal control, and the coercivity condition is satisfied. However, the isolatedness condition is satisfied only if the measure of  $\Omega$  is either zero or 1. In these two cases the optimal control is Lipschitz continuous.

## 5 Lipschitz Continuous Optimal Feedback Control

In this section, we prove the existence of a Lipschitz continuous locally optimal feedback control for problem (1)–(2). For this purpose we embed the problem into a family of problems by replacing the initial time 0 with any  $\tau \in [0, T]$  and the initial condition  $x(0) = x_0$

with  $x(\tau) = \xi \in \mathbb{R}^n$ . Denote this new family of problems by  $P(\tau, \xi)$ , so that  $P(0, x_0)$  is (1)–(2). Also, denote by  $J(\tau, \xi; u)$  the value of the objective function of  $P(\tau, \xi)$  for a control  $u \in \mathcal{U}$  being defined as

$$J(\tau, \xi; u) := g(x(T)) + \int_{\tau}^T h(t, x(t), u(t))dt,$$

where  $x$  is the solution of the initial-value problem

$$\dot{x}(t) = f(t, x(t), u(t)) \quad \text{for a.e. } t \in [\tau, T], \quad x(\tau) = \xi. \tag{18}$$

To set the stage, we give first the following definition which recasts the usual way a locally optimal feedback control is understood. Recall that  $(\bar{x}, \bar{u})$  is a locally unique solution of problem (1)–(2).

**Definition 1** The function  $u^* : [0, T] \times \mathbb{R}^n \rightarrow U$  is said to be a locally optimal feedback control around the reference solution pair  $(\bar{x}, \bar{u})$  if there exist positive numbers  $\varepsilon_0$  and  $\bar{a}$ , and a set  $\Gamma \subset [0, T] \times \mathbb{R}^n$  such that

- (i)  $\text{gph}(\bar{x}) + \{0\} \times B_{\varepsilon_0}(0) \subset \Gamma$ ;
- (ii) for every  $(\tau, \xi) \in \Gamma$  the equation

$$\dot{x}(t) = f(t, x(t), u^*(t, x(t))), \quad x(\tau) = \xi, \tag{19}$$

has a unique absolutely continuous solution  $\hat{x}[\tau, \xi]$  on  $[\tau, T]$  which satisfies  $\text{gph}(\hat{x}[\tau, \xi]) \subset \Gamma$ ;

- (iii) the function  $\hat{u}[\tau, \xi](\cdot) := u^*(\cdot, \hat{x}[\tau, \xi](\cdot))$  is measurable, bounded, and satisfies  $\|\hat{u}[\tau, \xi] - \bar{u}\|_{\infty} \leq \bar{a}$ , and  $J(\tau, \xi; \hat{u}[\tau, \xi]) \leq J(\tau, \xi; u)$ , where  $u$  is any admissible control on  $[\tau, T]$  with  $\|u - \bar{u}\|_{\infty} \leq \bar{a}$  for which the corresponding solution  $x$  of (19) exists on  $[\tau, T]$  and satisfies  $\text{gph}(x) \subset \Gamma$ ;
- (iv)  $u^*(\cdot, \bar{x}(\cdot)) = \bar{u}(\cdot)$ .

The main result of this section follows.

**Theorem 4** *Let the coercivity condition (6) and the isolatedness condition (14) hold. Then, there exists a locally optimal feedback control  $u^* : [0, T] \times \mathbb{R}^n \rightarrow U$  around  $(\bar{x}, \bar{u})$  which is Lipschitz continuous on the set  $\Gamma$  appearing (together with the positive numbers  $\varepsilon_0$  and  $\bar{a}$ ) in Definition 1.*

Let us first sketch the idea of the proof. First, we prove that for  $\xi$  close to  $\bar{x}(\tau)$  a unique solution  $(\bar{x}[\tau, \xi], \bar{u}[\tau, \xi])$  exists and it is close to the restriction of  $(\bar{x}, \bar{u})$  to  $[\tau, T]$ ; moreover,  $\bar{u}[\tau, \xi]$  depends in a Lipschitz way on  $\xi$  (in the space  $L^{\infty}$ ). Then, we show that  $\bar{u}[\tau, \xi]$  is Lipschitz continuous.

For any  $\tau \in [0, T)$ , we define the spaces

$$Y_{\tau} = W^{1,\infty} \times L^{\infty} \times W^{1,\infty}, \quad Z_{\tau} = L^{\infty} \times \mathbb{R}^n \times L^{\infty} \times \mathbb{R}^n \times L^{\infty},$$

where the time interval for these functional spaces is  $[\tau, T]$ . It is convenient to define the norm in  $Y_{\tau}$  as  $\|(x, u, \lambda)\| := \max\{\|x\|_{1,\infty}, \|u\|_{\infty}, \|\lambda\|_{1,\infty}\}$ . For any fixed  $\tau \in [0, T)$ , any (locally) optimal solution-multiplier triple  $y := (x, u, \lambda) \in Y_{\tau}$  for  $P(\tau, \bar{x}(\tau))$  satisfies the inclusion

$$F_{\tau}(y) + G_{\tau}(y) \ni 0, \tag{20}$$

where  $F_\tau : Y_\tau \rightarrow Z_\tau$  and  $G_\tau : Y_\tau \rightrightarrows Z_\tau$  are defined as

$$F_\tau(y) = \begin{pmatrix} -\dot{x} + f(\cdot, x, u) \\ x(\tau) - \bar{x}(\tau) \\ \dot{\lambda} + H_x(\cdot, x, u, \lambda) \\ \lambda(T) - g_x(x(T)) \\ H_u(\cdot, x, u, \lambda) \end{pmatrix}, \quad G_\tau(y) = \begin{pmatrix} 0 \\ 0 \\ 0 \\ 0 \\ N_{\mathcal{U}}^\infty(u) \end{pmatrix}.$$

Here

$$N_{\mathcal{U}}^\infty(u) := \{v \in L^\infty : v(t) \in N_U(u(t)) \text{ for a.e. } t \in [\tau, T]\}.$$

By using the superscript  $\infty$  in the notation of the latter set we emphasize that the cone  $N_{\mathcal{U}}^\infty(u)$  includes only a part of the normal cone  $N_{\mathcal{U}}(u)$  which is a subset of the dual space of  $L^\infty$ ; note that the dependence on  $\tau$  is not indicated.

**Proposition 1** *Let the coercivity condition (6) hold. Then, the mapping  $F_\tau + G_\tau$  is strongly metrically regular at the restriction of  $\bar{y} := (\bar{x}, \bar{u}, \bar{\lambda})$  to  $[\tau, T]$  (denoted in the same way) for 0. Moreover, the constants of strong regularity, call them  $\bar{\kappa}, \bar{\alpha}, \bar{\beta}$ , can be chosen independent of  $\tau$ .*

*Proof* The strong metric regularity of the mapping  $F_\tau$  follows from [5, Theorem 5], with the only difference that in [5] there is no terminal term in the cost functional and the functions  $h$  and  $f$  do not depend on time  $t$ . As is well known, under the smoothness conditions imposed the problem with a terminal cost can be transformed into an equivalent problem without a terminal cost. In addition, the time-dependent problem is handled in exactly the same way as the time-invariant; thus, the difference is basically formal. For reader’s convenience, below we outline the proof by highlighting the main steps and utilizing Theorem 1 as a shortcut.

First, observe that the coercivity condition (6) is fulfilled for problem  $P(\tau, \bar{x}(\tau))$  with the same constant  $\rho$  for all  $\tau$ . To show this, it is enough to take  $w(t) = 0$  on  $[0, \tau]$  in (6). The next step is to linearize the generalized (20) at  $\bar{y} = (\bar{x}, \bar{u}, \bar{\lambda})$ , obtaining

$$F_\tau(\bar{y}) + \mathcal{A}_\tau(y - \bar{y}) + G_\tau(y) \ni 0, \tag{21}$$

where

$$\mathcal{A}_\tau(y) = \begin{pmatrix} -\dot{x} + Ax + Bu \\ x(\tau) - \bar{x}(\tau) \\ \dot{\lambda} + Qx + Su \\ \lambda(T) - \bar{g}_{xx}x(T) \\ S^T x + Ru \end{pmatrix}.$$

The strong regularity of the mapping appearing in the linearization (21), say with constants  $\kappa, a, b$  independent of  $\tau$  is established in [7, Lemma 3] (with the caveat concerning the terminal cost and the dependence on  $t$ ). Consider the function

$$g_\tau(y) = -F_\tau(y) + F_\tau(\bar{y}) + \mathcal{A}_\tau(y - \bar{y}).$$

Then,  $g_\tau(\bar{y}) = 0$ . Since  $\mathcal{A}_\tau$  is the strict derivative (in  $L^\infty$ ) of  $F_\tau$  at  $\bar{y}$ , the Lipschitz modulus of  $g_\tau$  at  $\bar{y}$  is zero. Thus, in the notation of Theorem 1, taking  $\alpha$  sufficiently small one can make  $\mu$  arbitrarily close to zero; furthermore,  $\kappa'$  and  $\beta$  could be chosen accordingly to satisfy (17). It remains to put  $\bar{\kappa} = \kappa', \bar{\alpha} = \alpha, \bar{\beta} = \beta$  and to observe that these constants are independent of  $\tau$ . □

As a consequence of the last proposition, for any  $\xi \in IB_{\bar{y}}(\bar{x}(\tau))$  the inclusion  $F_\tau(y) + G_\tau(y) + \zeta \ni 0$  with  $\zeta = (0, -\xi, 0, 0, 0)$  has a unique solution  $(\bar{x}[\tau, \xi], \bar{u}[\tau, \xi], \bar{\lambda}[\tau, \xi])$  in

$IB_{\bar{a}}((\bar{x}, \bar{u}, \bar{\lambda}))$  and it is Lipschitz continuous with respect to  $\xi \in IB_{\bar{b}}(\bar{x}(\tau))$  with Lipschitz constant  $\bar{\kappa}$  in the norms of  $\mathbb{R}^n$  and  $Y_\tau$ .

Clearly, the constant  $\bar{b}$  can be decreased, if necessary, without affecting the strong regularity property. Then, we may assume that  $\bar{b} > 0$  is chosen so small that the coercivity assumption (6) adapted to problem  $P(\tau, \xi)$  with  $\xi \in IB_{\bar{b}}(\bar{x}(\tau))$  holds with a constant  $\rho/2$  (instead of  $\rho$ ). Here and further “adapted” means that the matrices  $A, B, Q, R, S$  are calculated along  $(\bar{x}[\tau, \xi], \bar{u}[\tau, \xi], \bar{\lambda}[\tau, \xi])$  instead of  $(\bar{x}, \bar{u}, \bar{\lambda})$  and the integration in (6) is on  $[\tau, T]$ . Since under the coercivity assumption, the necessary optimality condition (3) is also sufficient (for local optimality), we obtain the following proposition.

**Proposition 2** *Let the coercivity condition and isolatedness condition hold. Then, for any  $\xi \in IB_{\bar{b}}(\bar{x}(\tau))$ , the pair  $(\bar{x}[\tau, \xi], \bar{u}[\tau, \xi])$  defined in the second to last paragraph is the unique locally optimal solution of problem  $P(\tau, \xi)$ , in the set  $IB_{\bar{a}}((\bar{x}, \bar{u}))$ . Moreover, the function  $IB_{\bar{b}}(\bar{x}(\tau)) \ni \xi \mapsto \bar{u}[\tau, \xi]$  is Lipschitz continuous in the norm of  $L^\infty$  with Lipschitz constant  $\bar{\kappa}$ .*

It is important to note that assuming  $\bar{b}$  small enough we may guarantee that Remark 1 is still valid with  $e_0, d_0$  and  $d$  replaced with  $e_0/2, d_0/2$  and  $d/2$ , respectively, and for the interval  $[\tau, T]$  and the function  $\bar{u}[\tau, \xi], \xi \in IB_{\bar{b}}(\bar{x}(\tau))$ , instead of  $[0, T]$  and  $\bar{u}$ . The constant  $L$  remains the same.

As already mentioned, the coercivity assumption, adapted to problem  $P(\tau, \xi)$ , would hold (with  $\rho/2$  instead of  $\rho$ ) provided that  $\xi \in IB_{\bar{b}}(\bar{x}(\tau))$ . According to Lemma 4, an arbitrary redefinition of  $\bar{u}[\tau, \xi]$  on a set of measure zero which satisfies  $\bar{u}[\tau, \xi](t) \in V_{[\tau, T]}(\bar{u}[\tau, \xi]; t)$  for every  $t \in [\tau, T]$  (and such exists due to Corollary 1) fulfills the conditions that the matrix  $R(t)$  satisfies (7), and  $\bar{y}[\tau, \xi]$  satisfies (9) for every  $t \in [\tau, T]$ , all adapted to problem  $P(\tau, \xi)$ . Moreover, according to Lemma 2,  $\bar{u}[\tau, \xi]$  can be assumed to satisfy  $|\bar{u}[\tau, \xi](t) - \bar{u}(t)| \leq \|\bar{u}[\tau, \xi] - \bar{u}\|_\infty$  for every  $t \in [\tau, T]$ .

**Lemma 5** *Let the coercivity condition and the isolatedness condition hold. Then, there exists a number  $\varepsilon \in (0, \bar{b}]$  such that for every  $\tau \in [0, T)$  and  $\xi \in IB_\varepsilon(\bar{x}(\tau))$  the control  $\bar{u}[\tau, \xi]$  satisfies the isolatedness condition (on  $[\tau, T]$ ), namely there exists a (relatively) open set  $\mathcal{O} \subset [\tau, T] \times \mathbb{R}^m$  such that*

$$\{(t, u) \in [\tau, T] \times \mathbb{R}^m : H_u(t, \bar{x}[\tau, \xi](t), u, \bar{\lambda}[\tau, \xi](t)) + N_U(u) \ni 0\} \cap \mathcal{O} = \text{gph}(\bar{u}[\tau, \xi]). \tag{22}$$

*Proof* Let us take  $\varepsilon > 0$  so small that

$$\varepsilon \leq \bar{b}, \quad L(2\bar{\kappa} + \bar{L} + 1)\varepsilon < \rho, \tag{23}$$

where  $\bar{L}$  is the Lipschitz constant of  $\bar{u}$  (see Theorem 3). For arbitrarily fixed  $\tau \in [0, T)$  and  $\xi \in IB_\varepsilon(\bar{x}(\tau))$  denote (for short)  $\tilde{x}(t) = \bar{x}[\tau, \xi](t), \tilde{u} = \bar{u}[\tau, \xi]$  redefined as described before the statement of the lemma,  $\tilde{\lambda}(t) = \bar{\lambda}[\tau, \xi](t)$  and  $\tilde{y} = (\tilde{x}, \tilde{u}, \tilde{\lambda})$ . Also denote  $\tilde{R}(t) = H_{uu}(t, \tilde{y}(t)), \tilde{H}_u(t, u) = H_u(t, \tilde{x}(t), u, \tilde{\lambda}(t))$ . Then, due to the first inequality in (23) and the redefinition of  $\tilde{u}$ , we know that for every  $t \in [\tau, T]$

$$|\tilde{u}(t) - \bar{u}(t)| \leq \|\tilde{u} - \bar{u}\|_\infty, \quad \tilde{H}_u(t, \tilde{u}(t)) + N_U(\tilde{u}(t)) \ni 0, \tag{24}$$

$$w^\top \tilde{R}(t)w \geq \frac{\rho}{2}|w|^2 \quad \forall w \in U - U. \tag{25}$$

Let us define

$$\mathcal{O} = \left( \text{gph}(\tilde{u}) + (-\varepsilon, \varepsilon) \times \overset{\circ}{B}_\varepsilon(0) \right) \cap ([\tau, T] \times \mathbb{R}^m),$$

which is relatively open in  $[\tau, T] \times \mathbb{R}^m$ . We shall prove that the claim of the lemma holds with this set  $\mathcal{O}$ .

Note that the right side of (22) is contained in the left side; thus, it is sufficient to prove the opposite inclusion. Targeting a contradiction, let us assume that there exists a point

$$(t_0, u_0) \in \{(t, u) \in [\tau, T] \times \mathbb{R}^m : \tilde{H}_u(t, u) + N_U(u) \ni 0\} \cap \mathcal{O} \tag{26}$$

which is not in  $\text{gph}(\tilde{u})$ . Then,  $\tilde{u}(t_0) \neq u_0$ . From (26) and the second relation in (24), we have

$$\tilde{H}_u(t_0, u_0) + N_U(u_0) \ni 0, \quad \tilde{H}_u(t_0, \tilde{u}(t_0)) + N_U(\tilde{u}(t_0)) \ni 0.$$

From here

$$\tilde{H}_u(t_0, u_0)(\tilde{u}(t_0) - u_0) \geq 0, \quad \tilde{H}_u(t_0, \tilde{u}(t_0))(u_0 - \tilde{u}(t_0)) \geq 0,$$

which implies

$$(\tilde{H}_u(t_0, u_0) - \tilde{H}_u(t_0, \tilde{u}(t_0)))(u_0 - \tilde{u}(t_0)) \leq 0.$$

Then, using (25) (notice that  $u_0 \in U$ , since otherwise  $N_U(u_0) = \emptyset$ ), we obtain

$$\begin{aligned} 0 &\geq \left( \tilde{H}_u(t_0, u_0) - \tilde{H}_u(t_0, \tilde{u}(t_0)) \right) (u_0 - \tilde{u}(t_0)) \\ &= \int_0^1 \frac{d}{ds} \tilde{H}_u(t_0, \tilde{u}(t_0) + s(u_0 - \tilde{u}(t_0))) ds (u_0 - \tilde{u}(t_0)) \\ &= \int_0^1 (u_0 - \tilde{u}(t_0))^\top \tilde{H}_{uu}(t_0, \tilde{u}(t_0) + s(u_0 - \tilde{u}(t_0))) (u_0 - \tilde{u}(t_0)) ds \\ &\geq \int_0^1 (u_0 - \tilde{u}(t_0))^\top \tilde{R}(t_0) (u_0 - \tilde{u}(t_0))^\top ds - \int_0^1 sL|u_0 - \tilde{u}(t_0)|^3 ds \\ &\geq \frac{\rho}{2}|u_0 - \tilde{u}(t_0)|^2 - \frac{L}{2}|u_0 - \tilde{u}(t_0)|^3. \end{aligned}$$

Hence,

$$\rho \leq L|\tilde{u}(t_0) - u_0|.$$

Due to the inclusion  $(t_0, u_0) \in \mathcal{O}$ , there exists  $(t_1, u_1) \in \text{gph}(\tilde{u})$  such that  $|t_1 - t_0| \leq \varepsilon$ ,  $|u_1 - u_0| \leq \varepsilon$ . Then, continuing the inequality (26), we obtain

$$\begin{aligned} \rho &\leq L(|\tilde{u}(t_0) - \tilde{u}(t_1)| + |\tilde{u}(t_1) - \tilde{u}(t_1)| + |\tilde{u}(t_1) - u_1| + |u_1 - u_0|) \\ &\leq L(\|\tilde{u} - \tilde{u}\|_\infty + \bar{L}|t_0 - t_1| + \|\tilde{u} - \tilde{u}\|_\infty + \varepsilon) \leq L(2\bar{k} + \bar{L} + 1)\varepsilon. \end{aligned}$$

This last inequality contradicts (23). Hence, (22) holds. □

Having proved that the isolatedness condition is also fulfilled for problem  $P(\tau, \xi)$ , we can apply Theorem 3 to this problem and obtain that the (locally) optimal control  $\tilde{u}[\tau, \xi]$  is Lipschitz continuous. The Lipschitz constant,  $\bar{L}$ , depends on the problem only through the constant  $\rho$  (now  $\rho/2$ ) and the constant  $L$ , therefore can be chosen independent of  $\tau$  and  $\xi$ , provided that  $|\xi - \bar{x}(\tau)| \leq \varepsilon$ , where  $\varepsilon > 0$  is sufficiently small (independent of  $\tau$ ).

*Proof of Theorem 4* Let the number  $\varepsilon \in (0, \bar{b}]$  be the one from Lemma 5. There exists a number  $\varepsilon_0 \in (0, \varepsilon]$  such that for every  $\tau \in [0, T]$  and  $\xi \in B_{\varepsilon_0}(\bar{x}(\tau))$  the corresponding

$(\bar{x}[\tau, \xi], \bar{u}[\tau, \xi])$  exists on  $[\tau, T]$  and satisfies,  $|\bar{x}[\tau, \xi](t) - \bar{x}(t)| \leq \varepsilon$  for every  $t \in [\tau, T]$ , and  $\|\bar{u}[\tau, \xi] - \bar{u}\|_\infty < \bar{a}$ . Such an  $\varepsilon_0$  exists because the mapping  $B_{\bar{b}} \ni \xi \mapsto (\bar{x}[\tau, \xi], \bar{u}[\tau, \xi])$  is Lipschitz continuous in  $W^{1,\infty} \times L^\infty$  with constant  $\bar{k}$ .

Define the set

$$\Gamma = \{\bar{x}[\tau, \xi](t) : \tau \in [0, T], \xi \in B_{\varepsilon_0}(\bar{x}(\tau)), t \in [\tau, T]\}. \tag{27}$$

Clearly,  $\text{gph}(\bar{x}) + \{0\} \times B_{\varepsilon_0}(0) \subset \Gamma$ . For  $\tau \in [0, T]$  denote

$$\Gamma_\tau := \{\xi : (\tau, \xi) \in \Gamma\}.$$

Since for every  $\tau \in [0, T]$  we have  $\Gamma_\tau \subset B_\varepsilon(\bar{x}(\tau))$ , the function  $\Gamma_\tau \ni \xi \mapsto \bar{u}[\tau, \xi]$  is Lipschitz continuous with Lipschitz constant  $\bar{k}$ . Moreover, each of these functions  $\bar{u}[\tau, \xi]$  is Lipschitz continuous on  $[\tau, T]$  with Lipschitz constant  $\bar{L}$ . In addition, according to Proposition 2, for every  $\tau \in [0, T]$  and  $\xi \in \Gamma_\tau$ , the function  $\bar{u}[\tau, \xi]$  is the unique locally optimal control for the corresponding problem  $P(\tau, \xi)$  in the set  $B_{\bar{a}}(\bar{u})$ .

Now, define feedback control mapping  $x \mapsto u^*(\cdot, x)$  as

$$u^*(t, x) := \bar{u}[t, x](t) \quad \text{for } (t, x) \in \Gamma. \tag{28}$$

The values  $\bar{u}[t, x](t)$  are well defined since  $\bar{u}[t, x]$  is a (Lipschitz) continuous function. Clearly, all requirements of Definition 1 are satisfied; the last one following from the identity  $\bar{u}[t, \bar{x}(t)] = \bar{u}(t)$ .

Let us consider two arbitrary pairs  $(\tau, \xi), (s, \eta) \in \Gamma$ . Due to the Dynamic Programming Principle, for every  $s \geq \tau, \tau, s \in [0, T)$  and every  $t \in [s, T]$ , we have

$$\bar{u}[\tau, \xi](t) = \bar{u}[s, \bar{x}[\tau, \xi](s)](t).$$

Then,

$$\begin{aligned} |u^*(\tau, \xi) - u^*(s, \eta)| &= |\bar{u}[\tau, \xi](\tau) - \bar{u}[s, \eta](s)| \\ &\leq |\bar{u}[\tau, \xi](\tau) - \bar{u}[\tau, \xi](s)| + |\bar{u}[\tau, \xi](s) - \bar{u}[s, \eta](s)| \\ &\leq \bar{L}|s - \tau| + |\bar{u}[s, \bar{x}[\tau, \xi](s)](s) - \bar{u}[s, \eta](s)| \\ &\leq \bar{L}|s - \tau| + \bar{k}|\bar{x}[\tau, \xi](s) - \eta| \\ &\leq \bar{L}|s - \tau| + \bar{k}(|\bar{x}[\tau, \xi](s) - \xi| + |\xi - \eta|) \\ &\leq \bar{L}|s - \tau| + \bar{k}(M|s - \tau| + |\xi - \eta|), \end{aligned}$$

where  $M$  is an upper bound of  $|f(t, x, u)|$  in the set  $\Omega$  defined in Remark 1. This completes the proof of Theorem 4. □

*Remark 4* The last part of the proof and the uniqueness claim in Proposition 2 imply that the function  $\hat{u}[\tau, \xi]$  appearing in Definition 1 is the unique locally optimal control for in problem  $P(\tau, \xi)$  in the set  $B_{\bar{a}}(\bar{u})$ .

### 6 Regularity of the Value Function

In this section, we show that the existence of a Lipschitz continuous optimal feedback established in Theorem 4 implies certain smoothness properties of the value function. In the preceding sections we assume only local optimality at the reference point, see Definition 1. In line with that assumption, we introduce the following definition:



**Definition 2** The function  $V : [0, T] \times \mathbb{R}^n \rightarrow \mathbb{R}$  is said to be a local value function of problem (1)–(2) around a reference admissible pair  $(\bar{x}, \bar{u})$  if there exist positive numbers  $\varepsilon_0$  and  $\bar{a}$ , and a set  $\Gamma \subset [0, T] \times \mathbb{R}^n$  such that  $\text{gph}(\bar{x}) + \{0\} \times B_{\varepsilon_0}(0) \subset \Gamma$  and for every  $(\tau, \xi) \in \Gamma$  one has

$$V(\tau, \xi) = \inf_u J(\tau, \xi; u),$$

where the infimum of the objective function  $J(\tau, \xi; u)$  in problem P( $\tau, \xi$ ) is taken over all admissible pairs  $(x, u)$  for which  $\|u - \bar{u}\|_\infty \leq \bar{a}$ , (18) has a unique solution  $x$  on  $[\tau, T]$ , and  $\text{gph}(x) \subset \Gamma$ .

By this definition, the local value function, with a set  $\Gamma$  and a neighborhood  $B_{\bar{a}}(\bar{u})$ , is finite if for every  $(\tau, \xi) \in \Gamma$  there exists at least one admissible pair  $(x, u)$  satisfying  $\|u - \bar{u}\|_\infty \leq \bar{a}$  and  $\text{gph}(x) \in \Gamma$ . Clearly, in that case  $(\bar{x}, \bar{u})$  is a locally optimal solution.

As in Section 5, we denote  $\Gamma_\tau := \{\xi : (\tau, \xi) \in \Gamma\}$ . Thus, the condition  $\text{gph}(\bar{x}) + \{0\} \times B_{\varepsilon_0}(0) \subset \Gamma$  in Definition 2 means that  $B_{\varepsilon_0}(\bar{x}(t)) \subset \Gamma_t$  for every  $t \in [0, T]$ . We also denote  $\overset{\circ}{\Gamma} = \{(\tau, \xi) : \tau \in [0, T], \xi \in \text{int}(\Gamma_\tau)\}$ .

**Theorem 5** *Let the coercivity condition and the isolatedness condition hold. Then, problem (1)–(2) has a (finite) local value function  $V$  around  $(\bar{x}, \bar{u})$  (with a set  $\Gamma$  and parameters  $\varepsilon_0$  and  $\bar{a}$ ); moreover  $V(\tau, \cdot)$  is differentiable with respect to  $\xi$  whenever  $(\tau, \xi) \in \overset{\circ}{\Gamma}$  and the derivative  $V_\xi$  is Lipschitz continuous on  $\overset{\circ}{\Gamma}$ .*

*Proof* The proof is routine, in principle, but we present it in full, because we deal here with a local value function, which requires some attention to detail. We will prove the theorem with  $\Gamma, \varepsilon_0$  and  $\bar{a}$  as in Theorem 4. Then, there is a locally optimal Lipschitz continuous feedback control  $u^*$  in the sense of Definition 1, with the corresponding pairs  $(\hat{x}[\tau, \xi], \hat{u}[\tau, \xi])$ . According to this definition, we have

$$V(\tau, \xi) = g(\hat{x}[\tau, \xi](T)) + \int_\tau^T h(s, \hat{x}[\tau, \xi](s), \hat{u}[\tau, \xi](s)) ds. \tag{29}$$

First, we prove the following claim.

*Claim A:* for every  $(\tau, \xi_0) \in \overset{\circ}{\Gamma}$  there exists a number  $\delta > 0$  such that for every  $\xi, \xi' \in B_\delta(\xi_0)$  one has  $\|\hat{u}[\tau, \xi] - \bar{u}\|_\infty \leq \bar{a}$  and the initial value problem

$$\dot{x}(t) = f(t, x(t), \hat{u}[\tau, \xi](t)), \quad x(\tau) = \xi', \tag{30}$$

has a unique solution  $x^{\xi, \xi'}$  on  $[\tau, T]$  and  $\text{gph}(x^{\xi, \xi'}) \subset \Gamma$ . Note that  $x^{\xi, \xi} = \hat{x}[\tau, \xi]$ .

We recall (see Definition 1) that  $\hat{x}[\tau, \xi]$  is the solution of (18) with control  $u^*(t, \hat{x}[\tau, \xi](t))$ , where  $u^*$  is defined in (28). Since  $x \mapsto u^*(\cdot, x)$  is Lipschitz continuous, we obtain by a standard argument that for any  $(\tau, \xi_0) \in \overset{\circ}{\Gamma}$  there is a constant  $\delta_0 \in (0, \bar{a}/\bar{L}]$  such that  $B_{\delta_0}(\hat{x}[\tau, \xi_0](t)) \subset \Gamma_t$  for every  $t \in [\tau, T]$ . Indeed, due to the Lipschitz continuity of the right-hand side of (18) every solution starting backwards from a point in a sufficiently small neighborhood of  $\hat{x}[\tau, \xi_0](t)$  at time  $t > \tau$  takes values only in  $B_{\varepsilon_0}(\xi_0)$  at time  $\tau$ . From the definition of  $\Gamma$  in (27), the graph of each such trajectory is contained in  $\Gamma$  and  $\|\hat{u}[\tau, \xi] - \bar{u}\|_\infty \leq \bar{a}$ . Then, Claim A follows from (30) thanks to the Lipschitz continuity of  $\hat{u}[\tau, \xi]$  in  $\xi$  (Proposition 2). The proof of that last assertion, e.g., by contradiction, is straightforward

In addition, we have the representation

$$x^{\xi, \xi'}(t) - \hat{x}[\tau, \xi](t) = \Phi[\tau, \xi](t, \tau)(\xi' - \xi) + o(|\xi' - \xi|), \quad t \in [\tau, T],$$

where  $\Phi[\tau, \xi](t, s)$  is the fundamental matrix solution of the linearization of the differential equation in (30) normalized at  $t = s$ , that is,

$$\frac{\partial}{\partial t} \Phi(t, s) = f_x(t, \bar{x}[\tau, \xi](t), \bar{u}[\tau, \xi](t))\Phi(t, s), \quad \Phi(s, s) = \text{the identity.}$$

In particular, there exists a constant  $C$  such that for every  $\xi, \xi' \in \mathcal{B}_\delta(\xi_0)$  one has

$$\|x^{\xi, \xi'} - x^{\xi, \xi}\|_\infty = \|x^{\xi, \xi'} - \hat{x}[\tau, \xi]\|_\infty \leq C|\xi' - \xi|. \tag{31}$$

From Definition 2 and Claim A, we have that

$$V(\tau, \xi') \leq g(x^{\xi, \xi'}(T)) + \int_{\bar{\tau}}^T h(s, \hat{x}^{\xi, \xi'}(s), \hat{u}[\tau, \xi](s)) ds,$$

Utilizing this last inequality in (29), we obtain

$$\begin{aligned} V(\tau, \xi') - V(\tau, \xi) &\leq g(x^{\xi, \xi'}(T)) - g(\hat{x}[\tau, \xi](T)) \\ &\quad + \int_{\tau}^T [h(s, x^{\xi, \xi'}(s), \hat{u}[\tau, \xi](s)) - h(s, \hat{x}[\tau, \xi](s), \hat{u}[\tau, \xi](s))] ds \\ &= g_x(\tilde{x}^T) \Phi[\tau, \xi](T, \tau)(\xi' - \xi) \\ &\quad + \int_{\tau}^T h_x(s, \tilde{x}(s), \hat{u}[\tau, \xi](s))\Phi[\tau, \xi](s, \tau)(\xi' - \xi) ds + o(|\xi' - \xi|), \end{aligned}$$

where  $\tilde{x}^T \in \text{co}\{x^{\xi, \xi'}(T), \hat{x}[\tau, \xi](T)\}$  and  $\tilde{x}(\cdot)$  is a measurable selection of the set-valued map  $s \mapsto \text{co}\{x^{\xi, \xi'}(s), \hat{x}[\tau, \xi](s)\}$ . Due to the Lipschitz continuity of  $g_x$  and  $h_x$ , and (31), we obtain

$$\begin{aligned} V(\tau, \xi') - V(\tau, \xi) &\leq [g_x(\hat{x}[\tau, \xi](T))\Phi[\tau, \xi](T, \tau) \\ &\quad + \int_{\tau}^T h_x(s, \hat{x}[\tau, \xi](s), \hat{u}[\tau, \xi](s))\Phi[\tau, \xi](s, \tau) ds](\xi' - \xi) + o(|\xi' - \xi|). \end{aligned}$$

It is well-known that the expression in the brackets equals  $(\hat{\lambda}[\tau, \xi](\tau))^T$ , where, as before,  $\hat{\lambda}[\tau, \xi]$  is the solution of the adjoint equation (9) for the reference pair  $(\hat{x}[\tau, \xi], \hat{u}[\tau, \xi])$  and end-point condition  $\hat{\lambda}(T) = g_x(\hat{x}[\tau, \xi](T))$ . Hence,

$$V(\tau, \xi') - V(\tau, \xi) \leq \langle \hat{\lambda}[\tau, \xi](\tau), (\xi' - \xi) \rangle + o(|\xi' - \xi|). \tag{32}$$

Using this inequality with  $\xi' = \xi_0$  and taking into account the Lipschitz continuity of  $\hat{\lambda}[\tau, \xi]$  with respect to  $\xi$ , we obtain

$$\begin{aligned} V(\tau, \xi_0) - V(\tau, \xi) &\leq \langle \hat{\lambda}[\tau, \xi](\tau), \xi_0 - \xi \rangle + o(|\xi_0 - \xi|) \\ &= \langle \hat{\lambda}[\tau, \xi_0](\tau), \xi_0 - \xi \rangle + o(|\xi_0 - \xi|). \end{aligned}$$

Using (32) with  $\xi' := \xi$  and  $\xi = \xi_0$ , we obtain

$$V(\tau, \xi) - V(\tau, \xi_0) \leq \langle \hat{\lambda}[\tau, \xi_0](\tau), \xi - \xi_0 \rangle + o(|\xi - \xi_0|).$$

Combining the last two inequalities, we obtain that  $V$  is differentiable with respect to  $\xi$  at  $\xi_0$ ; furthermore,  $V_\xi(\tau, \xi_0) = \hat{\lambda}[\tau, \xi_0](\tau)$ . The Lipschitz continuity of  $V_\xi$  follows from the last expression and the Lipschitz continuity of the function  $(\tau, \xi) \mapsto \hat{\lambda}[\tau, \xi_0](\tau)$ .  $\square$

Observe that  $\Gamma$ ,  $\varepsilon_0$  and  $\bar{a}$  in this theorem can be taken to be those in the proof of Theorem 4. Also, observe that at the end of the last proof we obtained the equality  $V_\xi(\tau, \xi_0) = \hat{\lambda}[\tau, \xi_0](\tau)$ , which, as is well known, holds under various sets of assumptions. Moreover, based on Theorem 5, one can verify that if  $\bar{u}$  is a globally optimal solution, then the value function  $V$  is a classical solution of the corresponding Hamilton-Jacobi-Bellman equation (see, e.g., [1, Chapter III.3]).

**Acknowledgements** Open access funding provided by TU Wien (TUW). A.L. Dontchev is supported by the National Science Foundation Award CMMI 1562209 and by the Australian Research Council (ARC) Project DPI60100854. M.I. Krastanov is supported by the Sofia University “St. Kliment Ohridski” under contract no. 80-10-133/25.04.2018. V.M. Veliov is supported by the Austrian Science Foundation (FWF) under grant no. P31400-N32.

**Open Access** This article is distributed under the terms of the Creative Commons Attribution 4.0 International License (<http://creativecommons.org/licenses/by/4.0/>), which permits unrestricted use, distribution, and reproduction in any medium, provided you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made.

## References

1. Bardi, M., Capuzzo-Dolcetta, I.: Optimal Control and Viscosity Solutions of Hamilton-Jacobi-Bellman Equations. Birkhäuser, Boston (1997)
2. Cibulka, R., Dontchev, A.L., Krastanov, M., Veliov, V.M.: Metrically regular differential generalized equations. *SIAM J. Control Optim.* **56**, 316–342 (2018)
3. Cibulka, R., Preninger, J., Roubal, T.: On Uniform Regularity and Strong Regularity. Research Report, 2018–01, ORCOS, TU Wien (2018)
4. Dontchev, A.L.: Efficient estimates of the solutions of perturbed control problems. *J. Optim. Theory Appl.* **35**, 85–109 (1981)
5. Dontchev, A.L., Hager, W.W.: Lipschitzian stability in nonlinear control and optimization. *SIAM J. Control Optim.* **31**, 569–603 (1993)
6. Dontchev, A.L., Rockafellar, R.T.: *Implicit Functions and Solution Mappings*, 2nd edn. Springer, New York (2014)
7. Hager, W.W.: Multiplier methods for nonlinear optimal control. *SIAM J. Numer. Anal.* **27**, 1061–1080 (1990)
8. Ziemer, W.P.: *Modern Real Analysis*, 2nd edn. Springer, Berlin (2017)

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