Topics on optimal control and PDEs

by

Khai T. Nguyen
Mathematics department
University of Padova, Italy

Section 1. Introduction

Section 2. A model problem

Section 3. Time optimal control for linear systems

Section 4. Regularity of minimum time for nonlinear control systems

Section 5. Compactness properties of solutions to Hamilton-Jacobi equations

1 INTRODUCTION

Consider the control systems

$$\begin{cases} \dot{x}(t) &= f(x(t), u(t)), \quad t \in [0, +\infty[\ a.e., \\ \\ x(0) &= x_0, \end{cases} \tag{1.1}$$

where $x_0 \in \mathbb{R}^n$ and

+ $f : \mathbb{R}^n \times U \to \mathbb{R}^n$ is the dynamics of the control system
+ $U \subset \mathbb{R}^m$ is the control set
+ $u : [0, +\infty[\to U$ is a control function.

NOTATIONS: we will write

$$f(x,u) = \begin{pmatrix} f_1(x,u) \\ \vdots \\ f_n(x,u) \end{pmatrix} \quad \text{and} \quad x(t) = \begin{pmatrix} x^1(t) \\ \vdots \\ x^n(t) \end{pmatrix}.$$

The set of admissible controls is denoted by

$$\mathcal{U}_{ad} := \{ u : [0, +\infty[\to U \mid u \text{ is measurable } \}, \tag{1.2}$$

we will also write that

$$u(t) = \begin{pmatrix} u_1(t) \\ \vdots \\ u_m(t) \end{pmatrix}.$$

Note that a solution of (1.1) depends on initial state x_0 and the choice of admissible control u.

Definition 1.1 *Given any $x_0 \in \mathbb{R}^n$ and $u \in \mathcal{U}_{ad}$, a solution of (1.1) denoted by $y^{x_0,u}(\cdot)$ is called a trajectory of* (1.1) *starting from x_0 associated with the control u.*

In this note, we will assume that our control system satisfies the following standard hypotheses:

STANDARD HYPOTHESES

(SH1) The control set U is closed.

(SH2) The function f is continuous. Moreover, there exists a constant $K_1 > 0$ such that

$$|f(y,u) - f(x,u)| \leq K_1 \cdot |y - x|, \quad \forall x, y \in \mathbb{R}^n, u \in U.$$

Under assumption (SH2), (1.1) admits the unique global trajectory $y^{x_0,u}$.

PAYOFFS: Let's define the *payoff functional*

$$P[u(\cdot)] := \int_0^T r(y^{x_0,u}(t), u(t))dt + g(y^{x_0,u}(T)) \tag{1.3}$$

where the terminal time $T > 0$ and

+ the function $r : \mathbb{R}^n \times U \to \mathbb{R}$ is the running cost,
+ the function $g : \mathbb{R}^n \to \mathbb{R}$ is the terminal cost.

THE BASIC PROBLEM: The goal is to find an admissible control $u^*(\cdot)$ which maximizes the payoff function.

$$P[u^*(\cdot)] \geq P[u(\cdot)], \quad \forall u \in \mathcal{U}_{ad}. \tag{1.4}$$

Definition 1.2 *The admissible control u^* such that (1.4) is called an optimal control.*

THE BASIC QUESTIONS:

(i) Does an optimal control exist ?

(ii) How can we characterize an optimal control ?

(iii) How can we construct an optimal control ?

EXAMPLES:

Example 1: (Rocket railroad car)
Imagine a railroad car powered by rocket engines on each side. We introduce the variables

$x(t)$ is the position of the rocket railroad car on the train track at time t

$v(t)$ is the velocity of the rocket rail road car at time t

$F(t)$ is the force from the rocket engines at time t

where $-1 \leq F(t) \leq 1$ and the sign of $F(t)$ depends on which engine is firing.

Our goal: *is to construct $F(\cdot)$ in order to drive the rocket railroad car to the origin 0 with zero velocity in a minimum amount of time.*

Mathematical model: Assuming that the rocket railroad car has mass m, the motion of law is

$$\ddot{x}(t) = \frac{F(t)}{m} := u(t) \tag{1.5}$$

where $u(\cdot)$ is understood as a control function. For simplicity, we will also assume that $m = 1$. The motion equation of the rocket car is

$$\begin{cases} \ddot{x}(t) = u(t), \\ \\ x(0) = x_0 \quad \text{and} \quad v(0) = v_0 \end{cases} \tag{1.6}$$

where $u(\cdot) \in \mathcal{U} = [-1, 1]$, x_0 is the position of the rocket railroad car at time 0 and v_0 is the velocity of the rocket railroad car at x_0. By setting

$$z(t) = \begin{pmatrix} x(t) \\ \\ v(t) \end{pmatrix}, \quad A = \begin{pmatrix} 0 & 1 \\ \\ 0 & 0 \end{pmatrix} \quad \text{and} \quad b = \begin{pmatrix} 0 \\ \\ 1 \end{pmatrix} \tag{1.7}$$

we can rewrite (1.6) as the first order control system:

$$\begin{cases} \dot{z}(t) = A \cdot z(t) + u(t) \cdot b \\ \\ z(0) = (x_0, v_0)^T. \end{cases} \tag{1.8}$$

The payoff function is

$$P[u(\cdot)] = -\int_0^{\theta} 1 \, ds = \theta$$

where θ is the first time such that $z(\theta) = (0,0)^T$.
The goal is to find $u^* \in \mathcal{U}_{ad}$ such that

$$P[u^*(\cdot)] \geq P[u(\cdot)], \quad \forall u \in \mathcal{U}_{ad}.$$

Example 2: (Minimal surfaces of revolution)
Consider in the space $\mathbb{R}^3$ the two circles

$$\begin{cases} z^2 + y^2 = R_1 \\ \\ x = a_1 \end{cases} \quad \text{and} \quad \begin{cases} z^2 + y^2 = R_2 \\ \\ x = a_2 \end{cases} \tag{1.9}$$

where $a_1 < a_2$.

Let $\mathcal{A}_{ad}$ be the set of functions $\xi : [a_1, a_2] \to \mathbb{R}^3$ such that $\xi(x) = \begin{pmatrix} x \\ 0 \\ r(x) \end{pmatrix}$ where

$r(\cdot) : [a_1, a_2] \to \mathbb{R}^+$ is smooth and satisfies that $r(a_1) = R_1$ and $r(a_2) = R_2$. For each $\xi \in \mathcal{A}_{ad}$, we denote by

$$S_\xi = \Big\{(x, y, z)^T \;\Big|\; a_1 \le x \le a_2, z^2 + y^2 = r(x)\Big\} \tag{1.10}$$

the surface of revolution generated by ξ. The area of S_ξ is

$$Area(S_\xi) = 2\pi \int_{a_1}^{a_2} r(x)\sqrt{r'(x)^2 + 1}\; dx. \tag{1.11}$$

Our goal: Finding $\xi^*(\cdot) \in \mathcal{A}_{ad}$ such that

$$Area(S_{\xi^*}) \le Area(S_\xi) \quad \forall \xi(\cdot) \in \mathcal{A}_{ad}. \tag{1.12}$$

We can reformulate the problem into a control problem. Indeed, we consider the constant control system

$$\begin{cases} \dot{r}(t) &= u(t), \\ r(a_1) &= R_1, \end{cases} \tag{1.13}$$

where $u(\cdot) \in \mathcal{U}_{ad}$ which is denoted by

$$\mathcal{U}_{ad} = \Big\{u \in C^1([a_1, a_2], \mathbb{R}^+) \;\Big|\; \int_{a_1}^{a_2} u(s)ds = R_2 - R_1\Big\}. \tag{1.14}$$

The payoff functional is

$$P[u(\cdot)] = 2\pi \int_{R_1}^{R_2} r(s)\sqrt{1 + u^2(s)}\; ds. \tag{1.15}$$

The goal is to find $u^*(\cdot) \in \mathcal{U}_{ad}$ such that

$$P[u^*(\cdot)] \le P[u(\cdot)] \quad \forall u \in \mathcal{U}_{ad}. \tag{1.16}$$

We will answer the basic questions for some special cases which are contained in section 2 and section 3. Our main goal is to bring you to some more advanced problems related to control problem. More precisely, we shall present some recent results concerning with the regularity and the compactness of viscosity solutions to time optimal control and Hamilton-Jacobi-Bellmann Equations.

2 A model problem

2.1 A problem in calculus of variations

We now start the analysis of our model problem. Given $0 < T \le +\infty$, we suppose that

$$L : \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R} \quad \text{and} \quad g : \mathbb{R}^n \to \mathbb{R} \tag{2.1}$$

are respectively the continuous *running cost (Lagrangian)* and the continuous *terminal cost*. For any $(\bar{t}, \bar{x}) \in]0, T[\times \mathbb{R}^n$, we define the set of admissible arcs as

$$\mathcal{A}_{\bar{t},\bar{x}} = \{y(\cdot) \in AC([0,t], \mathbb{R}^n) \mid y(0) = \bar{x}\}$$

where $AC([0,\bar{t}], \mathbb{R}^n)$ is the set of absolutely continuous functions from $[0,t]$ to $\mathbb{R}^n$. The cost functional

$$J_{\bar{t},\bar{x}}[y(\cdot)] = \int_0^{\bar{t}} L(y(s), \dot{y}(s))ds + g(y(\bar{t})).$$

Then we consider the following problem:

$$\text{minimize } J_{\bar{t},\bar{x}}[y(\cdot)] \text{ over all } y \in \mathcal{A}_{\bar{t},\bar{x}}. \tag{2.2}$$

This problem is classical in the calculus of variations.

For our purpose, we would prefer to reset the above problem in a control sense. In fact, we consider the control system

$$\begin{cases} \dot{x}(t) = u(t) \quad a.e.\ t \in [0,T] \\ x(0) = \bar{x}, \end{cases} \tag{2.3}$$

where $x : [0, +\infty[\to \mathbb{R}^n$ and $u(\cdot) \in \mathcal{U}_{ad}$. Here, the admissible control $\mathcal{U}_{ad}^T$ is denoted by

$$\mathcal{U}_{ad}^T = \{u : [0,T] \to \mathbb{R}^n \mid u \in L^1_{loc}([0,T], \mathbb{R}^n)\}. \tag{2.4}$$

We suppose that

$$L : \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R} \quad \text{and} \quad g : \mathbb{R}^n \to \mathbb{R} \tag{2.5}$$

are respectively the continuous *running cost (Lagrangian)* and the continuous *terminal cost*. For every $\bar{t} \in [0, T[$ and $\bar{x} \in \mathbb{R}^n$, the payoff function is

$$P_{\bar{t},\bar{x}}[u(\cdot)] = \int_0^{\bar{t}} L(y^{\bar{x},u}(s), u(s))ds + g(y^{\bar{x},u}(\bar{t})). \tag{2.6}$$

Our problem is

$$\text{minimize } P_{\bar{t},\bar{x}}[u(\cdot)] \text{ over all } u \in \mathcal{U}_{ad}^{\bar{t}}. \tag{2.7}$$

Before going to study (2.7), we remark that (2.7) is a special case of the Bolza problem. For a general Bolza problem, the control system (2.3) will be nonlinear. Moreover, if the running cost $L = 0$, (2.7) is called a Mayer problem.

Let's consider the value function V which is associated with (2.7).

Definition 2.1 *The value $V : [0, T[\times \mathbb{R}^n \to \mathbb{R}$ of (2.7) is denoted by*

$$V(\bar{t}, \bar{x}) = \inf_{u \in \mathcal{U}_{ad}} P_{\bar{t},\bar{x}}[u(\cdot)], \quad \forall (\bar{t}, \bar{x}) \in [0, T[\times \mathbb{R}^n. \tag{2.8}$$

Noting that $V(0,\cdot) = g(\cdot)$.

We are now going introduce the result which is called *Bellman's optimality principle* or *dynamic programming principle*. This one is to show that V is an alternative characterization of the solution of a suitable partial differential equation.

Theorem 2.1 *Let $(\bar{t}, \bar{x}) \in]0, T[\times \mathbb{R}^n$, $u \in \mathcal{U}_{ad}^{\bar{t}}$. Then, for all $t \in [0, \bar{t}[$,*

$$V(\bar{t}, \bar{x}) \leq \int_0^t L(y^{\bar{x},u}(s), u(s))ds + V(\bar{t} - t, y^{\bar{x},u}(t)) \tag{2.9}$$

In addition, the control u is a minimizer for (2.8) if and only if the above equality holds for all $t \in [0, \bar{t}]$.

Proof. Let $t \in [0, \bar{t})$ and $u \in \mathcal{U}_{ad}$. For any control $v \in \mathcal{U}_{ad}$, we consider the control $u'(\cdot)$ such that

$$u'(s) = \begin{cases} u(s) & \forall s \in [0, t] \\ v(s) & \forall s \in [t, \bar{t}]. \end{cases}$$

By the definition of the value function, one has that

$$\begin{aligned} V(\bar{t}, \bar{x}) &\leq P_{\bar{t},\bar{x}}[u'(\cdot)] = \int_0^{\bar{t}} L(y^{\bar{x},u'}(s), u'(s))ds + g(y^{\bar{x},u'}(\bar{t})) \\ &= \int_0^t L(y^{\bar{x},u}(s), u(s))ds + \int_t^{\bar{t}} L(y^{\bar{x},u'}(s), u'(s))ds + g(y^{\bar{x},u'}(\bar{t})). \end{aligned}$$

Set $\bar{x}_t = y^{\bar{x},u}(t)$ and $\bar{v}_t(s) = v(t+s)$ for all $s \in [0, \bar{t} - t]$, we have that

$$y^{\bar{x},u'}(t+s) = y^{\bar{x}_t,\bar{v}_t}(s), \quad \forall s \in [0, \bar{t} - t].$$

Therefore,

$$V(\bar{t}, \bar{x}) \leq \int_0^t L(y^{\bar{x},u}(s), u(s))ds + P_{\bar{t}-t,\bar{x}_t}[\bar{v}_t(\cdot)].$$

Taking the infimum over all $v \in \mathcal{U}_{ad}$, we obtain that

$$V(\bar{t}, \bar{x}) \leq \int_0^t L(y^{\bar{x},u}(s), u(s))ds + V(\bar{t} - t, \bar{x}_t).$$

(2.9) is proved. We will leave the rest part of the proof for the readers.(**exercise 1**)

We can give a shaper formulation of the dynamic programming principle.

Theorem 2.2 *Let $(\bar{t}, \bar{x}) \in [0, T] \times \mathbb{R}^n$. Then, for all $t \in [0, \bar{t}]$, it holds*

$$V(\bar{t}, \bar{x}) = \inf_{u \in \mathcal{U}_{ad}} \Big\{ \int_0^t L(y^{\bar{x},u}(s), u(s))ds + V(\bar{t} - t, y^{\bar{x},u}(t)) \Big\}. \tag{2.10}$$

Proof. Given $\varepsilon > 0$, let $u \in \mathcal{U}_{ad}^{\bar{t}}$ be such that

$$V(\bar{t}, \bar{x}) \geq \int_0^{\bar{t}} L(y^{\bar{x},u}(s), u(s))ds + g(y^{\bar{x},u}(\bar{t})) - \varepsilon.$$

Then,

$$\begin{aligned} V(\bar{t}, \bar{x}) &\geq \int_0^t L(y^{\bar{x},u}(s), u(s))ds + \int_t^{\bar{t}} L(y^{\bar{x},u}(s), u(s))ds + g(y^{\bar{x},u}(\bar{t})) - \varepsilon \\ &\geq \int_0^t L(y^{\bar{x},u}(s), u(s))ds + V(\bar{t} - t, y^{\bar{x},u}(t)) - \varepsilon. \end{aligned}$$

By taking ε to zero, we obtain the proof. □

Corollary 2.1 *Let $(\bar{t}, \bar{x}) \in [0, T[\times \mathbb{R}^n$. Assume that V is differentiable at $(\bar{t}, \bar{x})$. Then, for any $v \in \mathbb{R}^n$, it holds*

$$V_t(\bar{t}, \bar{x}) + [-\nabla_x V(\bar{t}, \bar{x}) \cdot v - L(\bar{x}, v)] \leq 0 \tag{2.11}$$

Proof. For any $v \in \mathbb{R}^n$, we consider the constant control $u(s) = v$ for all $s \in [0, T[$. Recalling theorem 2.1, we have that

$$V(\bar{t}, \bar{x}) \leq \int_0^t L(y^{\bar{x},u}(s), u(s))ds + V(\bar{t} - t, y^{\bar{x},u}(t))$$

for all $t \in (0, \bar{t}[$. Since $u(s) = v$ and $y^{\bar{x},u}(s) = \bar{x} + sv$ for all $s \in [0, T[$, we obtain from the inequality that

$$V(\bar{t}, \bar{x}) \leq \int_0^t L(\bar{x} + sv, v)ds \ + V(\bar{t} - t, \bar{x} + tv), \quad \forall t \in]0, \bar{t}[.$$

Hence,

$$0 \leq \frac{\int_0^t L(\bar{x} + sv, v)ds}{t} + \frac{V(\bar{t} - t, \bar{x} + tv) - V(\bar{t}, \bar{x})}{t}.$$

Recalling that V is differentiable at $(\bar{t}, \bar{x})$, by taking $t \to 0^+$, we obtain that

$$0 \leq L(\bar{x}, v) - V_t(\bar{t}, \bar{x}) + \nabla_x V(\bar{t}, \bar{x}) \cdot v.$$

The proof is complete. □

We are now giving the standard hypotheses which ensure that (2.8) admits a minimizer. More precisely, we will assume that

(L1) For any $R > 0$, there exists γ_R such that

$$|L(y, u) - L(x, u)| \leq L_R \cdot |y - x|, \quad \forall u \in \mathbb{R}^n, x, y \in B(0, R).$$

(L2) There exists $l_0 > 0$ and a function $l : [0, \infty[\to [0, \infty[$ with $\lim_{r\to\infty} \frac{l(r)}{r} = +\infty$ and such that

$$L(x, u) \geq l(|u|) - l_0, \quad \forall x \in \mathbb{R}^n, u \in \mathbb{R}^n.$$

(L3) For every $\bar{x}$, the function $L(\bar{x}, \cdot)$ is convex.

We will state the existence result of an optimal control.

Theorem 2.3 *Assume that L satisfies (L1)-(L3) and g is locally Lipschitz and bounded below. Then, for every $(\bar{t}, \bar{x}) \in]0, T[\times\mathbb{R}^n$, there exists an optimal control $u^*(\cdot) \in \bar{\mathcal{U}_{ad}}$ of the problem (2.8), i.e.,*

$$V(\bar{t}, \bar{x}) = \min_{u\in\mathcal{U}_{ad}} P_{\bar{t},\bar{x}}[u(\cdot)] = P_{\bar{t},\bar{x}}[u^*(\cdot)].$$

In addition, we have that $\|u^*\|_{L^\infty([0,\bar{t}],\mathbb{R}^n)} < +\infty$.

Proof. Given $(\bar{t}, \bar{x}) \in]0, T[\times\mathbb{R}^n$, we set $\lambda = \inf_{u(\cdot)\in\mathcal{U}_{ad}} P_{\bar{t},\bar{x}}[u(\cdot)]$. There exists a sequence of control functions $u_k(\cdot) \in L^1([0, \bar{t}], \mathbb{R}^n)$ such that $\lim_{k\to\infty} P_{\bar{t},\bar{x}}[u_k(\cdot)] = \lambda$. Thus, there is $k_0 > 0$ such that $P_{\bar{t},\bar{x}}[u_k(\cdot)] \leq \lambda + 1$ for all $k > k_0$. Therefore, by recalling (L2) we have that

$$\begin{aligned}\int_0^{\bar{t}} l(|u_k(s)|)ds &\leq P_{\bar{t},\bar{x}}[u_k(\cdot)] + l_0\bar{t} - \inf_{x\in\mathbb{R}^n} g(x)\\ &\leq \lambda + 1 + l_0\bar{t} - \inf_{x\in\mathbb{R}^n} g(x)\end{aligned}$$

for all $k > k_0$. By $\lim_{r\to\infty} \frac{l(r)}{r} = +\infty$, there is $M_0 > 0$ such that $l(r) > r$ for all $r > M_0$. Hence, for all $k \geq k_0$, it holds that

$$\begin{aligned}\|u_k\|_{L^1([0,\bar{t}],\mathbb{R}^n)} &\leq M_0\bar{t} + \int_{|u_k(s)|>M_0} l(|u_k(s)|)ds\\ &\leq M_0\bar{t} + \lambda + 1 + l_0\bar{t} - \inf_{x\in\mathbb{R}^n} g(x).\end{aligned}$$

Therefore, there exists $M > 0$ such that

$$\|u_k\|_{L^1([0,\bar{t}],\mathbb{R}^n)} \leq M, \quad \forall k \in \mathbb{Z}^+. \tag{2.12}$$

Now, let $u \in L^1([0, \bar{t}], \mathbb{R}^n)$ be such that $\|u\|_{L^1} \leq M$. For $\alpha > 0$, we define

$$u_\alpha(s) = \begin{cases} u(s) & \text{if} \quad |u(s)| \leq \alpha,\\ 0 & \text{if} \quad |u(s)| > \alpha.\end{cases}$$

Observe that $|y^{\bar{x},u_\alpha}(s)|, |y^{\bar{x},u}(s)| \leq R_1$ for all $s \in [0, \bar{t}]$ where $R_1 = |\bar{x}| + M$. Hence, by setting $I_\alpha = \{s \in [0, \bar{t}] \mid |u(s)| > \alpha\}$, we can estimate that

$$
\begin{aligned}
&P_{\bar{t},\bar{x}}[u_\alpha(\cdot)] - P_{\bar{t},\bar{x}}[u(\cdot)] \\
&\qquad = \int_0^{\bar{t}} L(y^{\bar{x},u_\alpha}(s), u_\alpha(s)) - L(y^{\bar{x},u}(s), u(s))ds + g(y^{\bar{x},u_\alpha}(\bar{t})) - g(y^{\bar{x},u}(\bar{t})) \\
&\qquad \leq \left(L_{R_1}\bar{t} + g_{R_1}\right) \cdot \int_0^{\bar{t}} |u_\alpha(s) - u(s)|ds + \int_{I_\alpha} L(y^{\bar{x},u}(s), 0) - L(y^{\bar{x},u}(s), u(s))ds \\
&\qquad\qquad \leq \left(L_{R_1}\bar{t} + g_{R_1}\right) \cdot \int_{I_\alpha} |u(s)|ds + (K_M + l_0) \cdot |I_\alpha| - \int_{I_\alpha} l(|u(s)|)ds
\end{aligned}
$$

where g_{R_1} is a Lipschitz constant of g in $B(0, R_1)$ and $K_M = \sup_{|y| \leq R_1} L(y, 0)$. Let $\alpha > 1$, we then have that

$$
P_{\bar{t},\bar{x}}[u_\alpha(\cdot)] - P_{\bar{t},\bar{x}}[u(\cdot)] \leq \Gamma \cdot \int_{I_\alpha} |u(s)|ds - \int_{I_\alpha} l(|u(s)|)ds \tag{2.13}
$$

where $\Gamma = L_{R_1}\bar{t} + g_{R_1} + K_M + l_0$. Now, by $\lim_{r \to +\infty} \frac{l(r)}{r} = +\infty$ there exists $\alpha_\Gamma > 1$ such that

$$
l(r) > (\Gamma + 1) \cdot r \quad \forall r \geq \alpha_\Gamma.
$$

Hence, we have that

$$
P_{\bar{t},\bar{x}}[u_{\alpha_\Gamma}(\cdot)] - P_{\bar{t},\bar{x}}[u(\cdot)] \leq -\int_{I_{\alpha_\Gamma}} |u(s)|ds \leq -\alpha_\Gamma \cdot |I_{\alpha_\Gamma}|. \tag{2.14}
$$

It implies that

$$
P_{\bar{t},\bar{x}}[u_{\alpha_\Gamma}(\cdot)] \leq P_{\bar{t},\bar{x}}[u(\cdot)]
$$

Thus, by recalling (2.12), if we define

$$
v_k(s) = \begin{cases} u_k(s) & \text{if} \quad |u_k(s)| \leq \alpha_\Gamma, \\ \\ 0 \quad \text{if} & |u_k(s)| > \alpha_\Gamma. \end{cases}
$$

Then, we have that $\|v_k\|_{L^\infty([0,\bar{t}],\mathbb{R}^n)} \leq \alpha_\Gamma$ for $k > k_0$ and $\lim_{k\to\infty} P_{\bar{t},\bar{x}}[v_k(\cdot)] = \lambda$. Since $\|v_k\|_{L^\infty([0,\bar{t}],\mathbb{R}^n)}$ is bounded, there exists $v \in L^\infty([0, \bar{t}], \mathbb{R}^n)$ such that $v_k \rightharpoonup v$, i.e.,

$$
\lim_{k\to\infty} \int_0^{\bar{t}} w(s)v_k(s)ds = \int_0^{\bar{t}} w(s)v(s)ds \quad \forall w \in L^1([0, \bar{t}], \mathbb{R}^n).
$$

In particular, $y^{\bar{x},v_k}$ converges uniformly to $y^{\bar{x},v}$. Thus, we have that $\lim_{k\to\infty} g(y^{\bar{x},v_k}(\bar{t})) = g(y^{\bar{x},v}(\bar{t}))$. Hence,

$$
\liminf_{k\to\infty} P_{\bar{t},\bar{x}}[v_k(\cdot)] - P_{\bar{t},\bar{x}}[v(\cdot)] = \liminf_{k\to\infty} \int_0^{\bar{t}} L(y^{\bar{x},v_k}(s), u_k(s)) - L(y^{\bar{x},v}(s), v(s))ds.
$$

By (L1) we have that

$$\liminf_{k\to\infty}\Big|\int_0^{\bar t} L(y^{\bar x,v_k}(s),v_k(s))-L(y^{\bar x,v}(s),v_k(s))ds\Big| \le C\liminf_{k\to\infty}\int_0^{\bar t}|y^{\bar x,v_k}(s)-y^{\bar x,v}(s)|ds=0.$$

Thus, we obtain that

$$\liminf_{k\to\infty} P_{\bar t,\bar x}[v_k(\cdot)]-P_{\bar t,\bar x}[v(\cdot)]=\liminf_{k\to\infty}\int_0^{\bar t} L(y^{\bar x,v}(s),v_k(s))-L(y^{\bar x,v}(s),v(s))ds$$

Since $L(\bar x,\cdot)$ is convex and $v_k \rightharpoonup v$, we get that

$$\liminf_{k\to\infty}\int_0^{\bar t} L(y^{\bar x,v}(s),v_k(s))-L(y^{\bar x,v}(s),v(s))ds\ge 0.$$

Therefore,

$$\lambda=\liminf_{k\to\infty} P_{\bar t,\bar x}[v_k(\cdot)]\ge P_{\bar t,\bar x}[v(\cdot)].$$

This implies that $P_{\bar t,\bar x}[v(\cdot)]=\lambda$. Hence, $u^*=v$ is an optimal control of (2.8). The proof is complete. □

Now, we define the Hamilton function $H:\mathbb{R}^n\times\mathbb{R}^n\to\mathbb{R}^n$ which is associated with L as the following:

$$H(\bar x,p)=\sup_{q\in\mathbb{R}^n}\big\{-p\cdot q-L(\bar x,q)\big\}\quad \forall(\bar x,p)\in\mathbb{R}^n\times\mathbb{R}^n. \tag{2.15}$$

From (L2), $H(\bar x,p)<+\infty$ for all $(\bar x,p)\in\mathbb{R}^n\times\mathbb{R}^n$ (**exercise 2**). Moreover, by the continuity of L, for each p, there exists at least q_p where the supremum in (2.15) is attained. The function H is called the *Legendre transform* of the running cost L. We can now state the following lemma which is a consequence of corollary 2.1

Lemma 2.1 *Let $(\bar t,\bar x)\in]0,T[\times\mathbb{R}^n$ be a differentiability point of V. Assuming that L satisfies (L2). Then,*

$$V_t(\bar t,\bar x)+H(\bar x,\nabla_x V(\bar t,\bar x))\le 0. \tag{2.16}$$

We are proving a reversed inequality

Lemma 2.2 *Let $(\bar t,\bar x)\in]0,T[\times\mathbb{R}^n$ be a differentiability point of V. Assuming that L satisfies (L1)–(L3), g is locally lipschitz and bounded below. Then,*

$$V_t(\bar t,\bar x)+H(\bar x,\nabla_x V(\bar t,\bar x))\ge 0. \tag{2.17}$$

Proof. From theorem 3.8, let $u^*(\cdot)$ be the optimal control of (2.8). Then by recalling theorem 2.1, we have that

$$V(\bar t,\bar x)=\int_0^t L(y^{\bar x,u^*}(s),u^*(s))ds+V(\bar t-t,y^{\bar x,u^*}(t)),\quad \forall t\in]0,\bar t[.$$

Thus, for $t \in]0, \bar{t}[$

$$\frac{V(\bar{t}-t, y^{\bar{x},u^*}(t)) - V(\bar{t},\bar{x})}{-t} = \frac{\int_0^t L(y^{\bar{x},u^*}(s), u^*(s))ds}{t}. \tag{2.18}$$

Since $\|u^*\|_{L^\infty([0,\bar{t}],\mathbb{R}^n)}$ is bounded, there exists a sequence $t_k \to 0^+$ such that $\lim_{k\to 0^\infty} \frac{\int_0^{t_k} u^*(s)ds}{t_k} = \bar{u}$. Since V is differentiable at $(\bar{t},\bar{x})$, one can compute that

$$\lim_{k\to\infty} \frac{V(\bar{t}-t_k, y^{\bar{x},u^*}(t_k)) - V(\bar{t},\bar{x})}{-t_k} = V_t(\bar{t},\bar{x}) - \nabla_x V(\bar{t},\bar{x})\cdot \bar{u}. \tag{2.19}$$

On the other hand, recalling (L1), we have that

$$|L(y^{\bar{x},u^*}(s), u^*(s)) - L(\bar{x}, u^*(s))| \leq C \cdot t_k$$

for some positive constant C. Thus,

$$\lim_{k\to\infty} \frac{\int_0^{t_k} L(y^{\bar{x},u^*}(s), u^*(s))ds}{t_k} = \lim_{k\to\infty} \frac{\int_0^{t_k} L(\bar{x}, u^*(s))ds}{t_k}.$$

Recalling (L3) that $L(\bar{x},\cdot)$ is convex, by using Jensen's inequality, we then have

$$\lim_{k\to\infty} \frac{\int_0^{t_k} L(\bar{x}, u^*(s))ds}{t_k} \geq \lim_{k\to\infty} L\Big(\bar{x}, \frac{\int_0^{t_k} u^*(s)ds}{t_k}\Big) = L(\bar{x},\bar{u}).$$

Therefore,

$$\lim_{k\to\infty} \frac{\int_0^{t_k} L(y^{\bar{x},u^*}(s), u^*(s))ds}{t_k} \geq L(\bar{x},\bar{u}). \tag{2.20}$$

Recalling (2.18), (2.19) and (2.20), we obtain that

$$V_t(\bar{t},\bar{x}) - \nabla_x V(\bar{t},\bar{x})\cdot\bar{u} - L(\bar{x},\bar{u}) \geq 0.$$

The proof is complete. □

Combining the two above lemmas, we obtain that

Proposition 2.1 *Let $(\bar{t},\bar{x}) \in]0,T[\times\mathbb{R}^n$ be a differentiability point of V. Assuming that L satisfies (L1)–(L3), g is locally lipschitz and bounded below. Then,*

$$V_t(\bar{t},\bar{x}) + H(\bar{x}, \nabla_x V(\bar{t},\bar{x})) = 0. \tag{2.21}$$

Therefore, if the value function V is smooth in $]0,T[\times\mathbb{R}^n$ then V is a smooth solution of the Hamiton-Jacobi equation

$$\begin{cases} u_t(t,x) + H(x, \nabla_x u) = 0, \\ \\ u(0,\cdot) = g(\cdot). \end{cases} \tag{2.22}$$

2.2 The Hopf formula

We consider here the special case of $L(x,q) = L(q)$ and $T = +\infty$. We will assume that

(i) L is convex and $\lim_{|q|\to+\infty} \frac{L(q)}{|q|} = +\infty$

(ii) g is Lipschitz in $\mathbb{R}^n$.

The Hamiltion function associated with L is denoted by

$$H(p) = \sup_{q\in\mathbb{R}^n} \big\{-p\cdot q - L(q)\big\}. \tag{2.23}$$

To start with this subsection, we first introduce the formula of Hopf.

Theorem 2.4 *The value function V satisfies*

$$V(\bar{t},\bar{x}) = \min_{z\in\mathbb{R}^n} \left\{\bar{t}\cdot L\Big(\frac{z-\bar{x}}{\bar{t}}\Big) + g(z)\right\} \tag{2.24}$$

for all $(\bar{t},\bar{x}) \in]0,T[\times\mathbb{R}^n$.

Proof. Let $u(\cdot) \in \mathcal{U}^{\bar{t}}_{ad}$, we set $z_u = y^{\bar{x},u}(\bar{t})$. Since L is convex, we obtain from Jensen's inequality that

$$\int_0^{\bar{t}} L(u(s))ds \geq \bar{t}\cdot L\Big(\frac{1}{\bar{t}}\cdot\int_0^{\bar{t}} u(s)ds\Big) = \bar{t}\cdot L\Big(\frac{z_u-\bar{x}}{\bar{t}}\Big).$$

Hence,

$$P_{\bar{t},\bar{x}}[u(\cdot)] \geq \bar{t}\cdot L\Big(\frac{z_u-\bar{x}}{\bar{t}}\Big) + g(z_u). \tag{2.25}$$

This implies that

$$V(\bar{t},\bar{x}) \geq \inf_{u\in\mathcal{U}_{ad}} \left\{\bar{t}\cdot L\Big(\frac{z_u-\bar{x}}{\bar{t}}\Big) + g(z_u)\right\}.$$

Since $\{z_u \mid u\in\mathcal{U}^{\bar{t}}_{ad}\} = \mathbb{R}^n$, we get that

$$V(\bar{t},\bar{x}) \geq \inf_{z\in\mathbb{R}^n} \left\{\bar{t}\cdot L\Big(\frac{z-\bar{x}}{\bar{t}}\Big) + g(z)\right\}. \tag{2.26}$$

On the other hand, for every $z\in\mathbb{R}^n$ by choosing $u_z(\cdot)\in\mathcal{U}_{ad}$ such that $u_z(s) = \frac{z-\bar{x}}{\bar{t}}$ for all $s\in[0,\bar{t}]$, we have that

$$P_{\bar{t},\bar{x}}[u_z(\cdot)] = \bar{t}\cdot L\Big(\frac{z-\bar{x}}{\bar{t}}\Big) + g(z).$$

Therefore, we obtain that

$$V(\bar{t},\bar{x}) = \inf_{z\in\mathbb{R}^n} \left\{\bar{t}\cdot L\Big(\frac{z-\bar{x}}{\bar{t}}\Big) + g(z)\right\}. \tag{2.27}$$

Now, by using the Lipschit continuity of g and the coercivity of L, one can complete the proof. (**exercise 3**) □

We introduce here two main properties of the value function.

Proposition 2.2 *Let* $(\bar{t},\bar{x}) \in]0,T[\times\mathbb{R}^n$. *Then, the followings hold*

(i) ***A functional identity****: for every* $s \in [0,\bar{t}[$,

$$V(\bar{t},\bar{x}) = \min_{z\in\mathbb{R}^n}\Big\{(\bar{t}-s)\cdot L\Big(\frac{z-\bar{x}}{\bar{t}-s}\Big) + V(s,z)\Big\}. \tag{2.28}$$

(ii) ***The linear programming principle:*** *Let* $0 < s < \bar{t}$ *and assume that* x_0 *a minimum for (2.24). Let* $z = \frac{s}{t}x_0 + (1-\frac{s}{t})\bar{x}$. *Then, we have that* x_0 *is a minimum for* $s\cdot L(\frac{w-z}{s}) + g(w)$.

Exercise 4: Prove the proposition 2.2.

Now, we will use Hopf's fomula to prove the first regularity property of V.

Theorem 2.5 *The value function* V *is Lipschitz continuous in* $[0,T]\times\mathbb{R}^n$. *More precisely, we have that*

$$|V(\bar{t},\bar{x}) - V(t,x)| \le L_g|\bar{x}-x| + L_1|\bar{t}-t|, \tag{2.29}$$

where L_g *is the lipschitz constant of* g *and* $L_1 \ge 0$ *is a suitable constant.*

Proof. We first show that

$$|V(\bar{t},\bar{x}) - V(\bar{t},x)| \le L_g\cdot|\bar{x}-x|, \quad \forall x,\bar{x}\in\mathbb{R}^n, t\in[0,T[. \tag{2.30}$$

Indeed, by Hopf's formula, there exists z_x such that

$$V(\bar{t},x) = \bar{t}\cdot L\Big(\frac{z_x - x}{\bar{t}}\Big) + g(z_x).$$

One the on the hand, we also have that

$$V(\bar{t},\bar{x}) \le \bar{t}\cdot L\Big(\frac{(z_x+\bar{x}-x)-\bar{x}}{\bar{t}}\Big) + g(z_x+\bar{x}-x).$$

Therefore,

$$V(\bar{t},\bar{x}) - V(\bar{t},x) \le g(z_x+\bar{x}-x) - g(z_x) \le L_g|\bar{x}-\bar{x}|.$$

Similarly, we get that

$$V(\bar{t},x) - V(\bar{t},\bar{x}) \le g(z_x+\bar{x}-x) - g(z_x) \le L_g|\bar{x}-\bar{x}|.$$

(2.30) is proved.

Assume that $\bar{t} > t$, from proposition 2.2, there exists $z_t = \frac{t}{\bar{t}}\bar{x} + \frac{\bar{t}-t}{\bar{t}}z_{\bar{x}}$ such that

$$V(\bar{t},\bar{x}) = (\bar{t}-t)\cdot L\Big(\frac{z_t-\bar{x}}{\bar{t}-t}\Big) + V(t,z_t),$$

where z_x is a minimum of $\bar{t} \cdot L(\frac{w-\bar{x}}{\bar{t}}) + g(w)$.
Thus, by (2.30), we get that

$$|V(\bar{t}, \bar{x}) - V(t, x)| \leq (\bar{t} - t) \cdot L\Big(\frac{z_t - \bar{x}}{\bar{t} - t}\Big) + L_g|z_t - x|. \tag{2.31}$$

On the other hand, we have that

$$\bar{t} \cdot L(0) + g(\bar{x}) \geq V(\bar{t}, \bar{x}) = \bar{t} \cdot L\Big(\frac{z_{\bar{x}} - \bar{x}}{\bar{t}}\Big) + g(z_{\bar{x}}).$$

Hence,

$$L\Big(\frac{z_{\bar{x}} - \bar{x}}{\bar{t}}\Big) \leq L(0) + L_g \cdot \Big|\frac{z_{\bar{x}} - \bar{x}}{\bar{t}}\Big|.$$

Thus, by the coercivity of L, there exists a constant C_1 such that $|\frac{z_{\bar{x}} - \bar{x}}{\bar{t}}| \leq C_1$. This implies that

$$\Big|\frac{z_t - \bar{x}}{\bar{t} - t}\Big| = \Big|\frac{z_{\bar{x}} - \bar{x}}{\bar{t}}\Big| \leq C_1.$$

By recalling (2.31), we finally obtain that

$$|V(\bar{t}, \bar{x}) - V(t, x)| \leq L_g \cdot |\bar{x} - x| + \Big(L_g C_1 + \sup_{|q| \leq C_1} L(q)\Big) \cdot |\bar{t} - t|.$$

The proof is complete by setting $L_1 = L_g C_1 + \sup_{|q| \leq C_1} L(q)$. □

It is well-known that a Lipschitz contiuous function is differentiable at most everywhere by Rademacher's theorem. By recalling proposition 2.1, we obtain that

Corollary 2.2 *The value function V satisfies the H-J equation*

$$\begin{cases} u_t(t, x) + H(\nabla_x u) = 0, \\ u(0, \cdot) = g(\cdot) \end{cases} \tag{2.32}$$

almost everywhere in $]0, T[\times \mathbb{R}^n$.

One may expect that V is smooth such that V is a classical solution of (2.32). However, V usually fails to be everywhere differentiable even if g and L are differentiable. Indeed, let's consider the following example:

Example 5: Let's consider the problem (2.7) with $n = 1$, $L(x, q) = L(q) = \frac{q^2}{2}$ and the terminal cost g given by

$$g(z) = \begin{cases} z \quad \text{if } z \leq 0, \\ \frac{1}{2}(1 - (z - 1)^2) \quad \text{if } 0 < z \leq 1, \\ \frac{1}{2} \quad \text{if } z > 1. \end{cases} \tag{2.33}$$

By using Hopf's formula, we have that

$$V(1,x) = \min_{z\mathbb{R}} \Big\{ \frac{|z-x|^2}{2} + g(z) \Big\}.$$

We set

$$f(z) = \frac{|z-x|^2}{2} + g(z) \quad \forall z \in \mathbb{R}.$$

One can compute that

$$f'(z) = \begin{cases} 1 + z - x & \text{if } z \leq 0, \\ \\ 1 - x & \text{if } 0 < z \leq 1, \\ \\ z - x & \text{if } z > 1. \end{cases} \tag{2.34}$$

+ If $x > 1$ we have that $f'(z) = 0 \Longleftrightarrow z = x$. Thus, $V(1,x) = f(x) = \frac{1}{2}$.
+ If $x \leq 1$ we have that $f'(z) = 0 \Longleftrightarrow z = x - 1$. Thus, $V(1,x) = f(x-1) = x - \frac{1}{2}$.
Therefore, one can easily see that $V(1,\cdot)$ is not differentiable at $x = 1$. □

Exercise 5: Computing the value function V in the example.

On the other hand, the property of solving the equation (2.32) almost everywhere is not enough to characterize the value function V. Indeed, the problem (2.32) can have more then one solution which is Lipschitz continuous.

Example 6: The problem

$$\begin{cases} u_t(t,x) + \frac{1}{2}u_x^2 = 0, \\ \\ u(0,\cdot) = 0(\cdot) \end{cases} \tag{2.35}$$

admits the solution $u = 0$. However, for any $a > 0$, the function u_a defined as

$$u_a(t,x) = \begin{cases} 0 & \text{if } |x| \geq \frac{at}{2}, \\ \\ a|x| - \frac{1}{2}a^2 t & \text{if } |x| < \frac{at}{2} \end{cases} \tag{2.36}$$

is a Lipschitz function satisfying the equation almost everywhere together with its initial condition.

The above example show that the property of solving the equation almost everywhere is too weak and does not suffice to provide a satisfactory notion of generalized solution. It is therefore desirable to find additional conditions to ensure uniqueness and characterize the value function among the Lipschitz continuous solutions of the equations. A way of doing this relies on the semiconcavity property which is closely related to the entropy condition.

However, in this note, we prefer to use the concept of viscosity solutions to show that the viscosity solution of (5.49) is coincided with the value function.

2.3 Viscosity solutions

Consider the HJ equation

$$\begin{cases} u_t(t,x) + H(\nabla_x u) = 0, \\ \\ u(0,\cdot) = g(\cdot) \end{cases} \tag{2.37}$$

where $u : [0,+\infty[\times\mathbb{R}^n \to \mathbb{R}$ is the value function, $H : \mathbb{R}^n \to \mathbb{R}$ is the Hamilton function and the notation

$$u_t = \frac{\partial u}{\partial t}, \quad \nabla u = \Big(\frac{\partial u}{\partial x_1}, ..., \frac{\partial u}{\partial x_n}\Big).$$

Definition 2.2 *Let $u : [0,+\infty[\times\mathbb{R}^n$ be uniformly continuous. We say that u is a viscosity solution of (2.22) if $u(0,\cdot) = g(\cdot)$ and*

(1) *u is a viscosity subsolution of (2.22), i.e., for every $v \in C^1([0,+\infty[\times\mathbb{R}^n)$ such that $u - v$ has a local maximal at (t_0, x_0)*

$$v_t(t_0,x_0) + H(\nabla v(t_0,x_0)) \leq 0.$$

(2) *u is a viscosity supersolution of (2.22), i.e., for every $v \in C^1([0,+\infty[\times\mathbb{R}^n)$ such that $u - v$ has a local minimal at (t_0, x_0)*

$$v_t(t_0,x_0) + H(\nabla v(t_0,x_0)) \geq 0.$$

For this subsection, we will assume that

(i) H is convex and $\lim_{|q|\to+\infty} \frac{H(p)}{|p|} = +\infty$

(ii) g is Lipschitz in $\mathbb{R}^n$.

The *running cost (Lagrangian)* associated with H is the Legrende transform of H

$$L(q) = H^*(-q) = \sup_{p\in\mathbb{R}^n} \{-q\cdot p - H(p)\}.$$

The initial data g plays the role of the terminal cost.

Let V be the value function given by (2.24). Our goal is to show that the value function V is the unique viscosity solution of (5.49)

Theorem 2.6 *The value function V is the a viscosity solution of (5.49).*

Proof. Fix any $(t_0, x_0) \in]0,+\infty[\times\mathbb{R}^n$ and consider a function $v \in C^1([0,+\infty[\times\mathbb{R}^n)$ such that $V - v$ has a local maximal at (t_0, x_0), i.e.,

$$V(t,x) - v(t,x) \leq V(t_0,x_0) - v(t_0,x_0) \quad \forall (t,x) \in B((t_0,x_0),\delta)$$

for some $\delta > 0$. Thus, fixing any $w \in \mathbb{R}^n$, we have that for all $s > 0$ sufficiently small

$$\frac{v(t_0 - s, x_0 + sw) - v(t_0,x_0)}{s} \quad \geq \quad \frac{V(t_0 - s, x_0 + sw) - V(t_0,x_0)}{s}. \tag{2.38}$$

by recalling (2.9), we get

$$V(t_0, x_0) \leq sL(w) + V(t_0 - s, x_0 + sw).$$

Combining with (2.38), we have

$$\frac{v(t_0 - s, x_0 + sw) - v(t_0, x_0)}{s} \geq -L(w).$$

By letting s tend to 0^+, we obtain that

$$v_t(t_0, x_0) + [-\nabla v(t_0, x_0) \cdot w - L(w)] \leq 0.$$

Thus,

$$v_t(t_0, x_0) + H(\nabla v(t_0, x_0)) \leq 0.$$

It says that V is a viscosity subsolution of (5.49).

We are now going to show that V is a viscosity supersolution of (5.49). Fix any $(t_0, x_0) \in]0, +\infty[\times\mathbb{R}^n$ and consider a function $v \in C^1([0, +\infty[\times\mathbb{R}^n)$ such that $V - v$ has a local minimal at (t_0, x_0). Thus, one gets that

$$V(t, x) - V(t_0, x_0) \geq v(t, x) - v(t_0, x_0) \quad \forall (t, x) \in B((t_0, x_0), \delta) \tag{2.39}$$

for some $\delta > 0$. By Hopf's formula, there exists z_0 be such that

$$V(t_0, x_0) = t_0 \cdot L(w_0) + g(z_0)$$

where $w_0 = \frac{z_0 - x_0}{t_0}$. Recalling proposition 2.2, we have that

$$V(t_0, x_0) = sL(w_0) + V(t_0 - s, x_0 + sw_0) \quad \forall s \in]0, t_0[.$$

By choosing $(t, x) = (t_0 - s, x_0 + sw_0)$ in (2.39), we obtain for $s > 0$ sufficiently small that

$$-L(w_0) \geq \frac{v(t_0 - s, x_0 + sw_0) - v(t_0, x_0)}{s}.$$

Letting s tend to 0^+, we finally get

$$v_t(t_0, x_0) + H(\nabla v(t_0, x_0)) \geq v_t(t_0, x_0) + [-\nabla v(t_0, x_0) \cdot w_0 - L(w_0)] \geq 0$$

Thus, V is a viscosity solution of (5.49). For the uniqueness, one can show by using the result of comparison principle (see in [7]). □

Exercise 6: Showing that the value function V in theorem 3.8 is a viscosity solution of the Hamilton-Jacobi equation (2.22).

3 Time optimal control for linear systems

Consider the linear control system

$$\begin{cases} \dot{x}(t) &= Ax(t) + Bu(t), t > 0 \\ x(0) &= x_0, \end{cases} \tag{3.1}$$

where $A \in \mathbb{M}^{n\times n}$, $B \in \mathbb{M}^{n\times m}$, $1 \leq m \leq n$, and $U = [-1, 1]^m$. The set of admissible control is

$$\mathcal{U}_{ad} = \{u : [0, +\infty[\to U \mid u \text{ is measurable}\}.$$

Given $u \in \mathcal{U}_{ad}$, the trajectory starting from x_0 with control u can be presented as

$$y^{x_0,u}(t) = e^{At}x_0 + \int_0^t e^{A(t-s)}Bu(s)ds \tag{3.2}$$

where

$$e^{At} = \sum_{k=0}^{\infty} t^k \cdot \frac{A^k}{k!}$$

is the fundamental solution of

$$\begin{cases} \dot{X}(t) &= AX(t), \\ X(0) &= \mathbb{I}_n. \end{cases} \tag{3.3}$$

3.1 Controllability

In this subsection, we will answer the basic question on controllability for the linear control which mentioned particularly in the example of rocket railroad car. More precisely, we will consider the problem driving the system (3.1) to the origin.

Definition 3.1 *The reachable set for time $t > 0$ is*

$$\mathcal{R}(t) = \text{the set of initial points } x_0 \text{ for which there exists an admissible control } u \text{ such that } y^{x_0,u}(t) = 0.$$

The overall reachable set is the set of initial points x_0 for which there exists a control u such that $y^{x_0,u}(t) = 0$ for some $t > 0$, i.e.,

$$\mathcal{R} = \cup_{t\geq 0}\mathcal{R}(t).$$

We now introduce the definitions of controllability.

Definition 3.2 *The linear control system (3.1) is small time controllable on 0 if 0 is in the interior of $\mathcal{R}$. Moreover, if $\mathcal{R} = \mathbb{R}^n$, we say that (3.1) is fully controllable.*

Our goal is to study the controllability properties of (3.1). Firstly, from the definition, $x_0 \in \mathcal{R}(t)$ if and only if there exist a control $u(\cdot) \in \mathcal{U}_{ad}$ such that

$$0 = e^{At}x_0 + \int_0^t e^{A(t-s)}Bu(s)ds.$$

It is equivalent that

$$x_0 = -\int_0^t e^{-As}Bu(s)ds.$$

Therefore,

$$\mathcal{R}(t) = \Big\{x_0 = -\int_0^t e^{-As}Bu(s)ds \mid u \in \mathcal{U}_{ad}\Big\}. \tag{3.4}$$

By using this expression, one can show that

Theorem 3.1 *For a fixed $t > 0$, the followings hold*

(i) $\mathcal{R}(t)$ *is convex, symmetric and compact.*

(ii) $\mathcal{R}(t) \subseteq \mathcal{R}(\bar{t})$ *for all* $\bar{t} \geq t \geq 0$.

Exercise 7: Prove the above theorem.

A simple example: Let $n = 2$, $m = 1$ and $U = [-1, 1]$. Consider the linear control system

$$\begin{cases} \dot{x}(t) &= bu(t), t > 0 \\ \\ x(0) &= x_0, \end{cases} \tag{3.5}$$

where $b = (0, 1)^T$. One can easily compute that $\mathcal{R} = \{(x_1, x_2) \mid x_1 = 0\}$. Thus, $\mathcal{R}$ does not cover a neighborhood of the origin. Therefore, the control system is not fully controllable.

We are going to look for a condition which ensure that the reachable set $\mathcal{R}$ contains a small ball with the center being the origin. Let's start with a single column $B = b \in \mathbb{R}^n$, $U = [-1, 1]$.

Proposition 3.1 *Assuming that $B = b \in \mathbb{R}^n$ and $U = [-1, 1]$. Then,*

$$\mathcal{R} \subseteq H(A, b) = \text{span}\{b, Ab, ..., A^{n-1}b\}.$$

Proof. Fixing any $t > 0$, we have that

$$\mathcal{R}(t) = \Big\{x = -\int_0^t e^{-As}bu(s)ds \;\Big|\; u \in \mathcal{U}_{ad}\Big\}. \tag{3.6}$$

One has that

$$e^{-As}b = \sum_{k=0}^{\infty}(-1)^k \frac{s^k A^k b}{k!}. \tag{3.7}$$

Recall that

$$P(\lambda) := \det(\lambda \cdot \mathbb{I}_n - A)$$

is the characteristic polynomial of A. The Cayley-Hamiltonian states that

$$P(A) = 0.$$

It means that if we write

$$P(\lambda) = \lambda^n + \beta_{n-1}\lambda^{n-1} + ... + \beta_1\lambda + \beta_0,$$

then

$$A^n + \beta_{n-1}A^{n-1} + ... + \beta_1 A + \beta_0\mathbb{I}_n = 0.$$

Equivalently,

$$A^n = -\beta_{n-1}A^{n-1} - ... - \beta_1 A - \beta_0\mathbb{I}_n$$

Thus, one can show that for all $k \in \mathbb{Z}^+$,

$$A^{n+k} = -\beta_{n-1,k}A^{n-1} - ... - \beta_{1,k}A - \beta_{0,k}^k\mathbb{I}_n.$$

Therefore, from (3.7) it holds for every $s \in \mathbb{R}$

$$e^{-As}b \in \text{span}\{b, Ab, ..., A^{n-1}b\}.$$

It implies that

$$\mathcal{R}(t) \subseteq \text{span}\{b, Ab, ..., A^{n-1}b\}.$$

The proof is complete. □

Proposition 3.2 *For every $t > 0$, 0 is an interior point of $\mathcal{R}(t)$ in $H(A, b)$,i.e. there exists $r(t) > 0$ such that*

$$B(0, r(t)) \cap H(A, b) \subseteq \mathcal{R}(t). \tag{3.8}$$

In particular, there exist $\delta > 0$ such that

$$B(0, \delta) \cap H(A, b) \subseteq \mathcal{R}.$$

Proof. Recalling theorem 3.1 and proposition 3.1, $\mathcal{R}(t)$ is a compact, convex set in the subspace $H(A, b)$. Assuming that 0 is not an an interior point of $\mathcal{R}(t)$ in $H(A, b)$, by the convexity of $\mathcal{R}(t)$ in $H(A, b)$, there exists a unit vector $\xi \in H(A, b)$ such that

$$\langle \xi \ , \ y\rangle \leq 0, \quad \forall y \in \mathcal{R}(t).$$

Now, let's consider

$$g_\xi(s) = \langle \xi \ , \ e^{-sA}b\rangle.$$

According to $g_\xi(\cdot)$, we choose the control $u_\xi(\cdot) \in \mathcal{U}_{ad}$ such that

$$u_\xi(s) = -\text{sign}(g_\xi(s)), \quad \forall x \in [0, +\infty).$$

We set

$$y_\xi = -\int_0^t e^{-As} b u_\xi(s) ds.$$

One has that $y_\xi \in \mathcal{R}(t)$. Thus,

$$0 \geq \langle \xi \ , \ y_\xi \rangle = \int_0^t |g_\xi(s)| ds.$$

It implies that $g_\xi(s) = 0$ for all $s \in [0, t]$. Therefore,

$$g_\xi^{(k)}(s) = 0, \quad \forall s \in [0, t), k \in \mathbb{N}.$$

In particular,

$$\langle \xi \ , \ A^k b \rangle = 0, \quad \forall k \in \{0, ..., n-1\}.$$

Hence, ξ is orthogonal to $H(A, b)$. This is a contradiction. □

Theorem 3.2 *Assuming that* $\mathrm{Re}\lambda \leq 0$ *for each eigenvalue of* A. *Then*

$$\mathcal{R} = H(A, b).$$

Proof. Assume by a contradiction, from theorem 3.1 there exists $\bar{x} \in H(A, b)$ such that $\bar{x} \in \partial \mathcal{R}$. Since $\mathcal{R}$ is convex, there exist a unit $\xi \in H(A, b)$ such that

$$\langle \xi \ , y - \bar{x} \rangle \leq 0, \quad \forall y \in \mathcal{R}.$$

Equivalently,

$$\langle \xi \ , \ y \rangle \leq \nu, \quad \forall y \in \mathcal{R} \tag{3.9}$$

where $\nu = \langle \xi \ , \ \bar{x} \rangle$. Let's consider

$$g_\xi(s) = \langle \xi \ , \ e^{-sA} b \rangle.$$

According to $g_\xi(\cdot)$, we choose the control $u_\xi(\cdot) \in \mathcal{U}_{ad}$ such that

$$u_\xi(s) = -\mathrm{sign}(g_\xi(s)), \quad \forall x \in [0, +\infty).$$

We set

$$y_\xi(t) = -\int_0^t e^{-As} b u_\xi(s) ds \in \mathcal{R}(t).$$

One has that

$$\langle \xi \ , y_\xi(t) \rangle = \int_0^t |g_\xi(s)| ds \tag{3.10}$$

On the other hand, one can compute that

$$g_\xi^{(k)}(s) = \big\langle \xi \ , \ (-1)^k A^k b \big\rangle, \quad \forall k \in \mathbb{N}.$$

Therefore, by Cayley-Hamiltonian theorem, we have that

$$P\Big(-\frac{d}{ds}\Big)g_\xi(s) = 0, \quad \forall s \in \mathbb{R}$$

where $P(\lambda) = \det(\lambda \cdot \mathbb{I}_n - A)$. It implies that $g_\xi(\cdot)$ solves the $n-th$ order ODE. Let $\lambda_1, ..., \lambda_n$ be the n solutions of $P(\lambda) = 0$, or the n engenvalues of A. Therefore, according to ODE theory, we can write

$$g_\xi(s) = \sum_{i=1}^{n} p_i(s) \cdot e^{-\lambda_i s} \tag{3.11}$$

where $p_i(\cdot)$ are appropriate polynomials.
Now, assume that

$$\lim_{t\to+\infty} \langle \xi \ , y_\xi(t)\rangle = \int_0^{+\infty} |g_\xi(s)|ds \leq +\infty.$$

Then, $\lim_{s\to+\infty} g_\xi(s) = 0$. Since, $\text{Re}\lambda_i \leq 0$ for all $i \in \{1, 2, ..., n\}$, from the expression of $g_\xi(s)$ in (3.11), one can show that $g_\xi(s) = 0$ for all $s \in \mathbb{R}$. This is a contradiction. Hence,

$$\lim_{s\to\infty} \langle \xi \ , y_\xi(t)\rangle = +\infty.$$

This contradict to (3.9). Therefore,

$$\mathcal{R} = H(A, b).$$

The proof is complete. □

From proposition 3.1 and proposition 3.2, we will introduction a algebraic condition which ensures that the linear control system (3.1) is small time controllable.

Definition 3.3 *Controllability matrix is*

$$G(A, B) := [B, AB, ..., A^{n-1}B]. \tag{3.12}$$

Theorem 3.3 *Let $B = b \in \mathbb{R}^n$ and $U = [-1, 1]$. Then, the linear system (3.1) is small time controllable if and only if*

$$\text{Rank } G(A, b) \ = \ n. \tag{3.13}$$

Proof. Assume that the linear system (3.1) is small time controllable, i.e., there exists $\delta > 0$ such that

$$B(0, \delta) \subset \mathcal{R}.$$

Recalling proposition 3.1, we obtain that

$$B(0, \delta) \subset H(A, b) = \text{span}\{b, Ab, ..., A^{n-1}b\}.$$

It implies that

$$\text{Rank } G(A, b) \ = \ n.$$

On the other hand, assuming that

$$\text{Rank } G(A, b) \ = \ n,$$

we have $H(A, b) = \mathbb{R}^n$. From proposition 3.2, $0 \in \text{int}\mathcal{R}$, i.e., the linear system (3.1) is small time controllable. □

Together with theorem 3.2, we obtain that.

Theorem 3.4 *Let* $B = b \in \mathbb{R}^n$ *and* $U = [-1, 1]$. *Assume that* $\text{Re}\lambda \leq 0$ *for each eigenvalue of* A *and* rank $G(A, b) = n$. *Then, the linear control system (3.1) is fully controllable.*

We are now going to state the general results for linear control system.

Theorem 3.5 *The linear control system (3.1) is small time controllable if and only if*

$$\text{Rank } G(A, B) \ = \ n. \tag{3.14}$$

Proof. We write that $B = [b_1, ..., b_m]$ where $b_i \in \mathbb{R}^n$ for all $i \in \{1, ..., m\}$ and the control $u = (u_1, ..., u_m)^T$. One has that

$$-\int_0^t e^{-As} Bu(s)ds = -\sum_{i=1}^{m} \int_0^t e^{-As} b_i u_i(s)ds.$$

Hence,

$$\mathcal{R} = \mathcal{R}_1 + ... + \mathcal{R}_m, \tag{3.15}$$

where $\mathcal{R}_i$ is the reachable set of the linear system (3.1) with $B = b_i$ for all $i \in \{1, 2, ..., m\}$.

Now, assume that

$$\text{Rank } G(A, B) \ = \ n,$$

we have

$$H(A, b_1) + ... + H(A, b_m) = \mathbb{R}^n.$$

From proposition 3.2, for every $i \in \{1, ..., m\}$, there exists $\delta_i > 0$ such that

$$B(0, \delta_i) \cap H(A, b_i) \subseteq \mathcal{R}_i.$$

Therefore, from (3.15), one has that 0 is in the interior of $\mathcal{R}$, i.e., the linear control system (3.1) is small time controllable.

On the other hand, assume that the linear control system (3.1) is small time controllable, we will prove that rank $G(A, B) = n$. Assume by a contradiction, one has that

$$\dim(H(A, b_1) + ... + H(A, b_m)) < n.$$

In particular, $0 \notin \text{int}(H(A, b_1) + ... + H(A, b_m))$. By recalling proposition 3.1, we finally obtain that $0 \notin \text{int}\mathcal{R}$. This is a contradiction. □

We conclude this subsection with the following result.

Theorem 3.6 *Assume that* rank $G(A, B) = n$ *and* $\text{Re}\lambda \leq 0$ *for each eigenvalue of* A. *Then, the linear control system 3.1 is fully controllable.*

Exercise 8. Prove the above theorem.
Exercise 9. Showing that for every $t > 0$, there exists $r(t) > 0$ such that

$$B(0, r(t)) \cap H(A, B) \subset \mathcal{R}(t).$$

3.2 Bang-bang principle

Definition 3.4 *The control* $u \in \mathcal{U}_{ad}$ *is called bang-bang control if for each time* $t \geq 0$ *and for each index* $i \in \{1, ...m\}$, *we have* $|u_i(t)| = 1$ *where*

$$u(t) = \begin{pmatrix} u_1(t) \\ \vdots \\ u_m(t) \end{pmatrix}. \tag{3.16}$$

Our main result in this subsection is the following:

Theorem 3.7 (Bang-bang principle) *Given any* $\bar{x} \in \mathcal{R}$, *then there exists a bang-bang control* u *which steers* $\bar{x}$ *to the origin.*

Recalling (3.15) that

$$\mathcal{R} = \mathcal{R}_1 + ... + \mathcal{R}_m,$$

where $\mathcal{R}_i$ is the reachable set of the linear system (3.1) with $B = b_i$ for all $i \in \{1, 2, ..., m\}$. To prove theorem 3.7, we only need to show it in the case of $B = b \in \mathbb{R}^n$ and $U = [-1, 1]$. In this case, from the previous subsection, we have that for every $t > 0$

$$\mathcal{R}(t) \subseteq H(A, b) = \text{span}\{b, Ab, ..., A^{n-1}b\}$$

is convex and compact.

Proposition 3.3 *For every* $t > 0$, *consider the reachable set* $\mathcal{R}(t)$. *Let* x *be on the boundary of* $\mathcal{R}(t)$ *in the space* $H(A, b)$. *Then, there exists an bang-bang control which steers* x *to the origin in time* t.

Proof. Since $\mathcal{R}(t)$ is convex in $\mathbb{R}^n$, we have that $\mathcal{R}(t)$ is also convex in $H(A, b)$. Thus, for every $x \in \partial\mathcal{R}(t)$, there exists a unit vector $\xi_x \in H(A, b)$ such that $\xi_x \in N_{\mathcal{R}(t)}(x)$, i.e.,

$$\langle \xi_x \ , \ y - x \rangle \leq 0, \quad \forall y \in \mathcal{R}(t). \tag{3.17}$$

From (3.4), we have

$$x = -\int_0^t e^{-As} b u_x(s) ds,$$

where $u_x : [0, t[\to U$ is measurable. Let's consider

$$g(s) = \langle \xi_x \ , \ e^{-sA} b \rangle, \quad \forall s \in \mathbb{R}. \tag{3.18}$$

We define

$$\bar{u}(s) = -\text{sign}(g(s)), \quad \forall s \in [0, +\infty). \tag{3.19}$$

One can see that $\bar{u} \in \mathcal{U}_{ad}$. Hence,

$$\bar{y} = -\int_0^t e^{-As} b\bar{u}(s)ds$$

is in $\mathcal{R}(t)$. By recalling (3.17), we have that

$$\langle \xi_x \ , \ \bar{y} - x \rangle \leq 0.$$

That is

$$\int_0^t \Big\langle \xi_x \ , \ -e^{-As} b\bar{u}(s) + e^{-As} b u_x(s) \Big\rangle ds \leq 0.$$

Recalling (3.18) and (3.19), we obtain that

$$\int_0^t |g(s)| \cdot [1 + \text{sign}(g(s)) u_x(s)] ds \leq 0. \tag{3.20}$$

Since $|g(s)| \cdot [1 + \text{sign}(g(s)) u_x(s)] \geq 0$ for all $s \in [0, t]$, one has that

$$|g(s)| \cdot [1 + \text{sign}(g(s)) u_x(s)] = 0 \text{ for a.e. } s \in [0, t].$$

On the other hand, since $\xi_x \in H(A, b)$, one can show that the set

$$Z_g^t = \{s \in [0, t] \mid g(s) = 0\}$$

is finite. Hence, $1 + \text{sign}(g(s)) u_x(s) = 0$ for a.e. $s \in [0, t]$. It implies that

$$u_x(s) = \bar{u}(s) \text{ for a.e. } s \in [0, t].$$

Therefore, x is also steered to the origin by control $\bar{u}(s)$. The proof is complete. □

Exercise 10. Proving that the set $Z_g^t = \{s \in [0, t] \mid g(s) = 0\}$ is finite.

Proposition 3.4 *For every $t > 0$, the set $\mathcal{R}(t)$ is strictly convex in the subspace $H(A, b)$.*

Exercise 11. Proving the above proposition.

3.3 The continuity of the minimum time function

In this subsection, we will study the continuity of the minimum time function function for the linear control system with the target 0. Indeed, let's recall the linear control system

$$\begin{cases} \dot{x}(t) &= Ax(t) + Bu(t), \quad t > 0 \\ \\ x(0) &= x_0, \end{cases} \tag{3.21}$$

where $A \in \mathbb{M}^{n\times n}$, $B \in \mathbb{M}^{n\times m}$, $1 \le m \le n$, and $U = [-1,1]^m$. The set of admissible control is

$$\mathcal{U}_{ad} = \{u : [0, +\infty[\to U \mid u \text{ is measurable}\}.$$

Given $u \in \mathcal{U}_{ad}$, the trajectory starting from x_0 with control u can be presented as

$$y^{x_0,u}(t) = e^{At}x_0 + \int_0^t e^{A(t-s)}Bu(s)ds. \tag{3.22}$$

Now, assume that the target $S = \{0\}$.

Definition 3.5 *For fixed $x_0 \in \mathbb{R}^n\backslash 0$, let $T(x)$ be the minimum amount of time to reach to the target S from x_0, i.e.,*

$$T(x_0) := \inf\{t > 0 \mid y^{x_0,u}(t) = 0, u \in \mathcal{U}_{ad}\}. \tag{3.23}$$

Setting $T(0) = 0$ and $T(x_0) = +\infty$ for all $x_0 \in \mathbb{R}^n\backslash\mathcal{R}$, we say that $T : \mathbb{R}^n \to [0, +\infty]$ is the minimum time function of the control system (3.21) with target S.

From definition 3.5, one can see that

+ $T(x_0)$ is finite for all $x_0 \in \mathcal{R}$,

+ $T(x_0) = +\infty$ for all $x_0 \in \mathbb{R}^n\backslash\mathcal{R}$.

Theorem 3.8 (Existence of time-optimal control) *For every $x_0 \in \mathcal{R}$, we have that*

$$T(x_0) = \min\{t > 0 \mid y^{x_0,u}(t) = 0, u \in \mathcal{U}_{ad}\}.$$

It means that there exists an optimal control $u^* \in \mathcal{U}_{ad}$ such that $y^{x_0,u^*}(T(x_0)) = 0$.

Proof. Since $x_0 \in \mathcal{R}$, we have that $T(x_0) < \infty$. Thus, there exist a sequence of admissible controls $\{u_k\} \subset \mathcal{U}_{ad}$ and an increasing sequence $\{t_k\}$ converging to $T(x_0)$ such that

$$y^{x_0,u_k}(t_k) = 0, \quad \forall k \in \mathbb{N}.$$

Since $\|u_k\|_{L^\infty[0,T(x_0)]} \le 1$, there exists a subsequence $\{u_{k_l}\}$ such that $\{u_{k_l}\}$ converges weakly to $\bar{u}$ in $L^1([0,T])$. Therefore, $y^{x_1,u_{k_l}}(\cdot)$ converges uniformly to $y^{x_0,u}(\cdot)$. In particular, we obtain that $y^{x_0,\bar{u}}(t) = \lim_{k_l\to\infty} y^{x_0,u_{k_l}}(t_{k_l}) = 0$. The proof is complete. □

Before going to study the continuity property of T, we state the result of dynamics programming principle.

Theorem 3.9 (Dynamic programming principle) *Let x be in $\mathcal{R}$. Then, it holds*

$$\min\Big\{t + T(y^{x,u}(t)) \mid u \in \mathcal{U}_{ad}\Big\} = T(x)$$

for every $0 < t < T(x)$.

Exercise 12. Proving the above theorem.

Theorem 3.10 *Let $T_{\mathcal{R}} : \mathcal{R} \to [0, \infty)$ be denoted by $T_{\mathcal{R}}(x) = T(x)$ for all $x \in \mathcal{R}$. Then, $T_{\mathcal{R}}$ is continuous.*

Proof. We first show that $T_{\mathcal{R}}(\cdot)$ is continuous at 0. Let $\{x_k\} \subset \mathcal{R}$ be a sequence converging to 0. Since for every $t > 0$, there is $r(t) > 0$ such that

$$B(0, r(t)) \cap \mathcal{R}(t) \subset H(A, B),$$

there exists $t_k > 0$ converging to 0 such that $x_k \in \mathcal{R}(t_k)$. It implies that $0 \leq T_{\mathcal{R}}(x_k) \leq t_k$. Thus, $\lim_{k\to 0} T_{\mathcal{R}}(x_k) = 0$. Hence, $T_{\mathcal{R}}$ is continuous at 0.

Now, let $x \in \mathcal{R}\backslash\{0\}$. From theorem 3.8, let $u^* \in \mathcal{U}_{ad}$ be an optimal control steering x to 0, i.e., $y^{x,u^*}(T_{\mathcal{R}}(x)) = 0$. Given $z \in \mathcal{R}\backslash\{0\}$ be such that $T_{\mathcal{R}}(z) \geq T_{\mathcal{R}}(x)$, we consider trajectory $y^{z,u^*}(\cdot)$ staring from z with control u^*. By setting $\bar{z} = y^{z,u^*}(T_{\mathcal{R}}(x))$, we have that

$$\|\bar{z}\| = \|e^{T_{\mathcal{R}}(x)A}(z - x)\| \leq e^{T_{\mathcal{R}}(x)\cdot\|A\|} \cdot \|z - x\|.$$

By the dynamics programming principle, we have that

$$T_{\mathcal{R}}(z) - T_{\mathcal{R}}(x) \leq T_{\mathcal{R}}(\bar{z}).$$

Hence,

$$T_{\mathcal{R}}(z) - T_{\mathcal{R}}(x) \leq \sup_{\bar{z}\in B(0,\delta_z)} T_{\mathcal{R}}(\bar{z})$$

where $\delta_z = e^{T_{\mathcal{R}}(x)\cdot\|A\|} \cdot \|z - x\|$.

On the other hand, let $z \in \mathcal{R}\backslash\{0\}$ be such that $T_{\mathcal{R}}(z) \leq T_{\mathcal{R}}(x)$. Let $u_z^*(\cdot)$ be an optimal control steering from z to 0 in time $T_{\mathcal{R}}(z)$. By setting $\bar{x} = y^{x,u_z^*}(T_{\mathcal{R}}(z))$, we have that

$$\|\bar{x}\| = \|e^{T_{\mathcal{R}}(z)A}(z - x)\| \leq e^{T_{\mathcal{R}}(z)\cdot\|A\|} \cdot \|z - x\| \leq e^{T_{\mathcal{R}}(x)\cdot\|A\|} \cdot \|z - x\|.$$

By the dynamics programming principle, we have that

$$T_{\mathcal{R}}(x) - T_{\mathcal{R}}(z) \leq T_{\mathcal{R}}(\bar{x}).$$

Hence,

$$T_{\mathcal{R}}(x) - T_{\mathcal{R}}(z) \leq \sup_{\bar{z}\in B(0,\delta_z)} T_{\mathcal{R}}(\bar{z})$$

where $\delta_z = e^{T_{\mathcal{R}}(z)\cdot\|A\|} \cdot \|z - x\|$.

Therefore, for all $x \in \mathcal{R}\backslash\{0\}$, we have that

$$|T_{\mathcal{R}}(z) - T_{\mathcal{R}}(x)| \leq \sup_{\bar{z}\in B(0,\delta_z)} T_{\mathcal{R}}(\bar{z}) \tag{3.24}$$

where $\delta_z = e^{T_{\mathcal{R}}(x)\cdot\|A\|} \cdot \|z - x\|$. Thus, since $T_{\mathcal{R}}(\cdot)$ is continuous at 0, $T_{\mathcal{R}}(\cdot)$ is continuous on $\mathcal{R}$. The proof is complete. □

Together with theorem 3.6, we obtain immediately the following.

Theorem 3.11 *Assume that* rank $G(A,B) = n$ *and* $\mathrm{Re}\lambda \leq 0$ *for each eigenvalue of* A. *Then,* T *is continuous on* $\mathbb{R}^n$.

4 Regularity of minimum time function for nonlinear control system

4.1 Seminconcave functions with linear modulus

In this subsection, we collect some main properties of a semiconcave function with linear modulus. Some proofs will be postponed to the following subsection. Standard reference is in [14].

Definition 4.1 *Let $\Omega \subset \mathbb{R}^n$ be an open set. We say that a function $f : \Omega \to \mathbb{R}$ is semiconcave with linear modulus if there exists $C \geq 0$ such that*

$$\lambda f(x) + (1-\lambda)f(y) - f(\lambda x + (1-\lambda)y) \leq \lambda(1-\lambda) \cdot C \cdot |y-x|^2, \tag{4.1}$$

for all $\lambda \in [0,1]$ such that $[x,y] \subset \Omega$. The constant C is called a semiconcavity constant for f in Ω.

A function f is called semiconvex in Ω if $-f$ is semiconcave in Ω.

Noting that if the constant $C = 0$ in (4.1), we say that f is concave in Ω. Hence, a semiconcave function is a concave function up to a quadratic term. More precisely,

Remark 4.1 *The function $f : \Omega \to \mathbb{R}$ is semiconcave with the semiconcavity constant C in Ω if and only if $f(\cdot) - C \cdot \frac{|\cdot|^2}{2}$ is concave in Ω or $D^2 f \leq C$ in the sense of distributions.*

We now introduce a standard criterion of semiconcave functions.

Proposition 4.1 *Let $f : \Omega \to \mathbb{R}$. Assume that f continuous and*

$$f(x+h) + f(x-h) - 2f(x) \leq C \cdot |h|^2 \tag{4.2}$$

for all $[x-h, x+h] \subset \Omega$. Then, f is semiconcave with a semiconcavity constant C in Ω.

Proof. We set

$$g(x) = f(x) - C \cdot |x|^2, \quad \forall x \in \Omega.$$

From (4.7), we have that

$$g(x+h) + g(x-h) - 2g(x) \leq 0, \tag{4.3}$$

for all $[x-h, x+h] \subset \Omega$. Moreover, (4.1) follows that

$$\lambda g(y) + (1-\lambda)g(x) - g(\lambda x + (1-\lambda)y) \leq 0, \tag{4.4}$$

for all $\lambda \in [0,1]$ such that $[x,y] \subset \Omega$. Now, one can show that (4.3) implies (4.4) for all $\lambda \in \mathbb{Q} \cap [0,1]$. Then, by using the continuity, we obtain (4.4) for all $\lambda \in [0,1]$. □

Proposition 4.2 *A semiconcave function $f : \Omega \to \mathbb{R}$ is locally Lipschitz continuous in Ω.*

Proof. Setting

$$g(x) = f(x) - C \cdot |x|^2, \quad \forall x \in \Omega,$$

we have that

$$\lambda g(y) + (1-\lambda) g(x) \le g(\lambda x + (1-\lambda) y). \tag{4.5}$$

Given any $x_0 \in \Omega$, we consider a closed cube Q with center x_0 such that $Q \subset \Omega$. Let $x_1, ..., x_{2^n}$ be the vertices of Q and

$$m = \min\{f(x_i) \mid i = 1, ..., 2^n\}.$$

For every $x \in Q$, there exists $0 \ge \lambda_1, ..., \lambda_{2^n} \le 1$ such that $\sum_{i=1}^{2^n} \lambda_i = 1$ and

$$y = \sum_{i=1}^{2^n} \lambda_i \cdot x_i.$$

Therefore, from (4.5), one proves that

$$m \le \sum_{i=1}^{2^n} \lambda_i g(x_i) \le g(y). \tag{4.6}$$

It implies that for all $y \in Q$

$$f(y) \ge m - C \cdot \|y\|^2 \ge m_0 := m - C \cdot \max_{z \in Q} \|z\|^2.$$

Therefore, f is bounded below in Q.

On the other hand, for every $x \in Q$, we have that

$$g(x) + g(2x_0 - x) \le 2 \cdot g(x_0).$$

Hence,

$$g(x) \le 2g(x_0) - g(2x_0 - x) \le 2g(x_0) - m.$$

It implies that

$$f(x) \le 2f(x_0) - C\|x_0\|^2 - m + C\|x\|^2 \le M_0 = 2f(x_0) - C\|x_0\|^2 - m + C \cdot \max_{z \in Q} \|z\|^2.$$

Therefore, f is bounded above in Q.

We are now going to show that f is Lipschitz in $Q_1 = x_0 + \frac{1}{2}(Q - x_0)$. Given any $x, y \in Q_1$, there exists $x_1 \in \partial Q$ such that $x \in [y, x_1]$.

$$x = \frac{|y - x|}{|x_1 - y|} \cdot x_1 + \frac{|x - x_1|}{|x_1 - y|} \cdot y.$$

Hence, from (4.5), we get

$$\frac{g(x) - g(y)}{|x - y|} \le \frac{g(y) - g(x_1)}{|x_1 - y|}.$$

It implies that

$$\frac{f(x)-f(y)}{|y-x|} \le \frac{f(y)-f(x_1)}{|x_1-y|} - C\cdot|x_1+y| + C\cdot|x+y|$$

Since $f(\cdot)$ is bounded in Q and $|x_1-y| \le \frac{\mathrm{diam}(Q)}{4}$, we have that

$$\frac{f(x)-f(y)}{|y-x|} \le L_Q$$

for a suitable constant $L_Q > 0$. Similarly, one gets that

$$\frac{f(y)-f(x)}{|y-x|} \le L_Q.$$

Therefore,

$$|f(y)-f(x)| \le L_Q \cdot |y-x|, \quad \forall x,y \in Q.$$

The proof is complete. □

From the above two propositions, one has that

Corollary 4.1 *Let $f:\Omega\to\mathbb{R}$. f is semiconcave with a semiconcavity constant C in Ω if and only if f continuous and*

$$f(x+h)+f(x-h)-2f(x) \le C\cdot|h|^2 \tag{4.7}$$

for all $[x-h,x+h]\subset\Omega$.

Let us now recall the result of H. Rademacher.

Theorem 4.1 (H. Rademacher) *A locally Lipschitz function $f:\Omega\to\mathbb{R}$ is a.e. differentiable in Ω.*

Hence, we obtains the first result on the differentiability of semiconcave function.

Corollary 4.2 *A semiconcave function $f:\Omega\to\mathbb{R}$ is a.e. differentiable in Ω.*

Moreover, a semiconcave function with linear modulus is a concave function up to a quadratic term. This allows to extend immediately some well-know properties of concave functions.

Theorem 4.2 *Let $f:\Omega\to\mathbb{R}$ be semiconcave. Then the following holds:*

*(i) (**Alexandroff's Theorem**) f is a.e. twice differentiable in Ω, i.e., for a.e. $x\in\Omega$, there exists a vector $p\in\mathbb{R}^n$ and a symmetric matrix B such that*

$$\lim_{y\to x}\frac{f(y)-f(x)-\langle p,y-x\rangle+\langle B(y-x),y-x\rangle}{|y-x|^2}=0.$$

(ii) The gradient of f, defined almost everywhere in Ω, belongs to the class $BV_{loc}(\Omega, \mathbb{R}^n)$.

Example 7. (Distance function) Let $S \subset \mathbb{R}^n$ be closed. The distance function from a point to S is defined by

$$d_S(x) = \min_{y \in S} |y - x|, \quad (x \in \mathbb{R}^n)$$

is locally semiconcave in $\mathbb{R}^n \backslash S$.

Exercise 13 Proving that

(1) $d_S(\cdot)$ is locally semiconcave in $\mathbb{R}^n \backslash S$.

(2) $d_S(\cdot)$ is not locally semiconcave in $\mathbb{R}^n$.

(3) $d_S^2(\cdot)$ is seminconcave with semiconcavity constant 2.

4.2 Generalize differentials

We are now introducing a concept in nonsmooth analysis which are used to studying deeper in the regularity of semiconcave functions.

Definition 4.2 *Let f be a real valued function defined on the open set $\Omega \subset \mathbb{R}^n$. For any $x \in \Omega$, the sets*

$$D^- f(x) = \Big\{ p \in \mathbb{R}^n \ \Big| \ \liminf_{y \to x} \frac{f(y) - f(x) - \langle p, y - x \rangle}{|y - x|} \geq 0 \Big\} \tag{4.8}$$

$$D^+ f(x) = \Big\{ p \in \mathbb{R}^n \ \Big| \ \limsup_{y \to x} \frac{f(y) - f(x) - \langle p, y - x \rangle}{|y - x|} \leq 0 \Big\} \tag{4.9}$$

are called, respectively, the (Fréchet) superdifferential and subdifferential of f at x.

We recall the basic properties of superdifferential and subdifferential of f.

Proposition 4.3 *Let $f : \Omega \to \mathbb{R}^n$ and $x \in \Omega$. Then, the following properties hold:*

(i) $D^+ f(x) = -D^-(-f)(x)$.

(ii) $D^+ f(x)$ *and* $D^- f(x)$ *are convex (possibly empty).*

(iii) $D^+ f(x)$ *and* $D^- f(x)$ *are both nonempty if and only if f is differentiable at x. In this case, we have that*

$$D^+ f(x) = D^- f(x) = Df(x).$$

We are now stating the properties of superdifferential of a semiconcave function which are not valid for a general Lipschitz continuous function.

Proposition 4.4 *Let $f : \Omega \to \mathbb{R}$ be semiconcave with semiconcavity constant C. Then, a vector $p \in \mathbb{R}^n$ belongs to $D^+ f(x)$ if and only if*

$$f(y) - f(x) - \langle p, y - x \rangle \leq \frac{C}{2} \cdot |y - x|^2$$

for every $y \in \Omega$ such that $[x, y] \subset \Omega$.

Exercise 14 Proving the above proposition.

Corollary 4.3 *Let $f : \Omega \to \mathbb{R}$ be semiconcave with semiconcavity constant C and let $[x, y] \subset \Omega$. Then, for every $p \in D^+ f(x)$, $q \in D^+ f(y)$, it holds*

$$\langle q - p, y - x \rangle \leq 2C \cdot |y - x|^2.$$

Before going to give a presentation of superdifferential of a semiconcave function. We introduce the concept of reachable gradient.

Definition 4.3 *Let $f : \Omega \to \mathbb{R}^n$ be locally Lipschitz. For every $x \in \Omega$, we denote by*

$$D^* f(x) = \Big\{ p = \lim_{k\to\infty} Df(x_k) \mid f \text{ is differentiable at } x_k \text{ and } x_k \to x \Big\}.$$

From Rademacher's Thereom, one can see that $D^* f(x)$ is nonempty. In the case of seminconcave function, we also have that

Proposition 4.5 *Let $f : \Omega \to \mathbb{R}$ be semiconcave with semiconcavity constant C and let $x \in \Omega$. Then,*

(i) $D^+ f(x) = \operatorname{co}(D^* f(x))$ *where* $\operatorname{co}(D^* f(x))$ *is the convex hull of* $D^* f(x)$.

(ii) $D^+ f(x)$ *is singleton if and only if* f *is differentiable at* x.

(iii) $D^+ f(\cdot)$ *is upper semicontinuous.*

(iv) if $D^+ f(y)$ *is singleton in the neighborhood* $\mathcal{O}_x$ *of* x *then* $f(\cdot)$ *is* C^1 *in* $\mathcal{O}_x$.

To conclude this subsection, we are now discussing on the singular set of f. We denote by

$$\Sigma_f = \{x \in \Omega \mid f \text{ is not differentiable at } x\}.$$

From proposition 4.5, if f be semiconcave then

$$\Sigma_f = \{x \in \Omega \mid \dim_{\mathcal{H}} D^+ f(x) \geq 1\}. \tag{4.10}$$

Theorem 4.3 *Let $f : \Omega \to \mathbb{R}$ be semiconcave. Then, Σ_f is countable $\mathcal{H}^{(n-1)}$-rectifiable. More generally, if we denote by*

$$\Sigma_f^k = \{x \in \Omega \mid \dim_{\mathcal{H}} D^+ f(x) \geq k\}$$

then Σ_f^k is countable $\mathcal{H}^{(n-k)}$-rectifiable.

Proposition (4.5) and the above theorem will be proved in the following subsection.

4.3 Semiconcavity and time optimal control

Consider the control systems

$$\begin{cases} \dot{x}(t) = f(x(t), u(t)), & t \in [0, +\infty[\ a.e., \\ x(0) = x_0, \end{cases} \tag{4.11}$$

where $x_0 \in \mathbb{R}^n$ and

+ $f : \mathbb{R}^n \times U \to \mathbb{R}^n$ is the dynamics of the control system
+ $U \subset \mathbb{R}^m$ is the control set
+ $u : [0, +\infty[\to U$ is a control function.

Standard hypotheses

(H1) $f : \mathbb{R}^n \times U \to \mathbb{R}^n$ is Lipschitz

$$|f(y, u) - f(x, u)| \leq L_1 \cdot |y - x|, \quad \forall x, y \in \mathbb{R}^n, u \in U. \tag{4.12}$$

Moreover, the gradient of f with respect to x exists everywhere and is locally Lipschitz in x, uniformly in u.

(H2) U is compact.

The set of admissible control is

$$\mathcal{U}_{ad} = \big\{u : [0, \infty) \to U \mid u \text{ is measurable}\big\}.$$

For every $u \in \mathcal{U}_{ad}$, we recall that $y^{x_0,u}(\cdot)$ is the trajectory staring from x with control u which is the unique solution of (4.11). The minimum time needed to steer x to the closed target $\mathcal{S}$, regarded as a function of x, is called the minimum time function and is denoted by

$$T_{\mathcal{S}}(x) := \inf \ \{t \geq 0 \mid y^{x,u}(t) \in \mathcal{S}, u \in \mathcal{U}_{ad}\}. \tag{4.13}$$

Now, we define

$$H(x, p) = \sup_{u \in U} \langle p, f(x, u)\rangle. \tag{4.14}$$

By the dynamic programming principle, one can show that $T_S(\cdot)$ is a vicosity solution of *Hamilton-Jacobi-Bellman* equation

$$H(x, \nabla T_S(x)) - 1 = 0, \quad \forall x \in \mathcal{R} \backslash S, \tag{4.15}$$

i.e., for all $x \in \mathcal{R}\backslash$,

$$H(x, p) - 1 \geq 0, \quad \forall p \in D^- T_S(x),$$

$$H(x, p) - 1 \leq 0, \quad \forall p \in D^+ T_S(x),$$

where $\mathcal{R}$ is the reachable set denoted by

$$\mathcal{R} = \{x \in \mathbb{R}^n \mid T_S(x) < \infty\}.$$

In particular, the equation (4.15) hold at all differentiability points of $T_S(x)$. Thus,

$$H(x,p) - 1 = 0, \quad \forall x \in \mathcal{R}\backslash S, p \in D^*T(x).$$

It is proved in [6] that $T_\mathcal{S}$ is the unique viscosity solution of (4.15) in $\mathcal{R}\backslash S$ satisfying suitable boundary condition.

We want to study the properties of $T_\mathcal{S}$ under the following controllability assumption:

(H3) For very $R > 0$, there exist $\mu_R > 0$ such that for all $x \in (B(0,R) \cap \mathcal{R})\backslash S$, there is $u_x \in U$:

$$f(x,u_x) \cdot \frac{x - \pi_S(x)}{|x - \pi_S(x)|} \le -\mu_R. \tag{4.16}$$

It is well studied in [23, 24, 8] that

Proposition 4.6 *Assume that system (4.11) satisfies (H1)-(H3). Then, T_S is locally Lipschitz in $\mathbb{R}^d$. Moreover, for every $R > 0$, it holds*

$$T_S(x) \le C_R \cdot d_S(x), \quad \forall x \in (B(0,R) \cap \mathcal{R})\backslash S$$

for some constant C_R.

Therefore, $T_S(x)$ is differentiable almost everywhere in $\mathcal{R}\backslash S$ and

$$H(x, \nabla T_S(x)) - 1 = 0, \qquad a.e.\ x \in \mathcal{R}\backslash S.$$

We now state the main result of this subsection (see in [13]).

Theorem 4.4 *Assume that system (4.11) satisfies (H1)-(H3) and the target S satisfies a ρ_0-internal sphere condition, i.e., for every $x \in \partial S$, there exists x_0 such that $x \in B'(x_0, \rho_0) \subset S$. Then, $T_S(\cdot)$ is locally semiconcave in $\mathcal{R}\backslash S$.*

Sketch of proof. (**The method of middle point**)
Fixing any $x \in \mathcal{R}\backslash S$, let $h \in \mathbb{R}^n$ be such that $[x-h, x+h] \subset \mathcal{R}\backslash S$, one needs to show that

$$T_S(x+h) + T_S(x-h) - 2T_S(x) \le C_x \cdot |h|^2. \tag{4.17}$$

Let $u^*(\cdot)$ be an optimal control steering x to S in time $T_S(x)$. We define

$$y_h^+(t) = y^{x+h,u^*}(t), \quad y(t) = y^{x,u^*}(t) \quad \text{and} \quad y_h^-(t) = y^{x-h,u^*}(t).$$

By the dynamics programming principle, we have that

$$T_S(x+h) + T_S(x-h) - 2T_S(x) \le T(y_h^+(t)) + T(y_h^-(t)) - 2T(y(t)). \tag{4.18}$$

Moreover, observing that

$$|y_h^+(t) + y_h^-(t) - 2y(t)| \leq C \cdot |h|^2. \tag{4.19}$$

From (4.18), (4.19) and the locally Lipschitz continuity of $T_{\mathcal{S}}$, we finally obtain that

$$T_S(x+h)+T_S(x-h)-2T_S(x) \leq T(y_h^+(t))+T(y_h^-(t))-2T\Big(\frac{y_h^+(t))+T(y_h^-(t)}{2}\Big)+C_1|h|^2.$$

Therefore, we only need to study the semiconcavity property of $T_S(\cdot)$ near to the target S. We leave the rest part for the reader. □

4.4 Sets with finite perimeter

In this subsection, we will recall some basic concepts from geometric measure theory. The major references are [25], [26] and [2].

Definition 4.4 *Let $A \subseteq \mathbb{R}^d$ and $0 \le p \le d$. The p-*dimensional Hausdorff measure *$\mathcal{H}^p(A)$ is defined by $\mathcal{H}^p(A) = \lim_{\delta \to 0^+} \mathcal{H}^p_\delta(A)$, where*

$$\mathcal{H}^p_\delta(A) = \omega_p \inf \left\{ \sum_{i=1}^{\infty} (\mathrm{diam}(U_i))^p : A \subseteq \bigcup_i U_i,\, \mathrm{diam}(U_i) < \delta \right\},$$

and

$$\omega_p := \frac{2^p \Gamma(\frac{p}{2}+1)}{\pi^{p/2}}, \qquad \Gamma(p) := \int_0^\infty t^{p-1}\, e^{-t}\, dt.$$

The constant ω_p is chosen so that $\mathcal{H}^p(A)$ equals the Lebesgue measure $\mathcal{L}^p(A)$ if $p \in \mathbb{N}$ and A is a subset of a p-dimensional subspace of $\mathbb{R}^d$. We define the Hausdorff dimension *$\dim_{\mathcal{H}}(A)$ of A by setting:*

$$\dim_{\mathcal{H}}(A) := \inf\{d \ge 0 : \mathcal{H}^d(A) = 0\}.$$

Let $k \in \mathbb{N}$, we say that $A \subset \mathbb{R}^d$ is countably k-rectifiable *if*

$$A \subset \mathcal{N} \cup \bigcup_{i=1}^{\infty} S_i$$

where S_i are suitable Lipschitz k-dimensional surfaces and $\mathcal{N}$ is a $\mathcal{H}^k$-negligible set. We say A is k-rectifiable if it is countably k-rectifiable and $\mathcal{H}^k(A) < \infty$, while A is locally k-rectifiable if $A \cap K$ is k-rectifiable for any compact set $K \subset \mathbb{R}^d$. Given an open subset Ω of $\mathbb{R}^d$ and a Lipschitz continuous function $f : \Omega \to \mathbb{R}^m$, with Lipschitz rank $L \ge 0$, for every $0 \le k \le d$, the estimate $\mathcal{H}^k(f(S)) \le L^k \mathcal{H}^k(S)$ holds for all $S \subseteq \Omega$. (see Proposition 2.49(iv) in [2]).

The concepts of functions of bounded variation and of sets with finite perimeter will also be used (see p. 117 and p. 143 in [2]):

Definition 4.5 *Let $\Omega \subset \mathbb{R}^d$ be open, and $u \in L^1(\Omega)$. We say that u is a* function of bounded variation in Ω *(denoted by $u \in BV(\Omega)$) if the distributional derivative of u is representable by a finite Radon measure in Ω, i.e., if*

$$\int_\Omega u \frac{\partial \varphi}{\partial x_i}\, dx = -\int_\Omega \varphi\, dD_i u \text{ for all } \varphi \in C_c^\infty(\Omega),\, i = 1, \ldots, d$$

for some Radon measure $Du = (D_1 u, \ldots, D_d u)$. We denote by $\|Du\|$ the total variation of the vector measure Du, i.e.

$$\|Du\|(\Omega) := \sup \left\{ \int_\Omega u(x) \mathrm{div}\phi(x)\, dx : \phi \in C_c^1(\Omega, \mathbb{R}^d),\, \|\phi\|_{L^\infty(\Omega)} \le 1 \right\}.$$

Accordingly, $u \in L^1_{\mathrm{loc}}(\Omega)$ is a function of locally bounded variation in Ω *(denoted by $u \in BV_{\mathrm{loc}}(\Omega)$) if $u \in BV(U)$ for every $U \subseteq \Omega$ open and bounded with $\overline{U} \subset \Omega$.*

Lemma 4.1 *Let $f \in BV(a,b)$; then there exists a measurable set $I \subseteq (a,b)$ such that $\mathcal{L}^1(I) = b - a$ and*

$$\|Df\|(a,b) \geq |f(t) - f(s)| \qquad \text{for any } t, s \in I\,.$$

Definition 4.6 *Let $E \subset \mathbb{R}^d$ be $\mathcal{L}^d$-measurable, and let $\Omega \subseteq \mathbb{R}^d$ be open. E has* finite perimeter in Ω *if its characteristic function*

$$\chi_E(x) := \begin{cases} 1, & \text{if } x \in E, \\ 0, & \text{otherwise,} \end{cases}$$

has bounded variation in Ω, and we say that the perimeter *of E in Ω is $P(E,\Omega) = \|D\chi_E\|(\Omega)$. We say that E has perimeter locally finite in Ω if $P(E,U) < +\infty$ for every open bounded subset U of Ω with $\overline{U} \subset \Omega$.*

Definition 4.7 *Let μ be a Radon measure on $\mathbb{R}^d$, and let M be the union of all open sets $U \subset \mathbb{R}^d$ such that $\mu(U) = 0$; the complement of M is called the support of μ and it is denoted by* $\mathrm{supp}(\mu)$.

The following concept of normal vector was introduced by De Giorgi.

Definition 4.8 *Let Ω be a nonempty open subset of $\mathbb{R}^d$ and $E \subset \mathbb{R}^d$ be a set of finite perimeter in Ω; we call* reduced boundary of E in Ω *the set $\partial^* E$ of all points $x \in \mathrm{supp}(\|D\chi_E\|) \cap \Omega$ such that*

$$\nu_E(x) := \lim_{\rho \to 0^+} \frac{D\chi_E(B(x,\rho))}{\|D\chi_E(B(x,\rho))\|} = \frac{dD\chi_E}{d\|D\chi_E\|}(x)$$

exists in $\mathbb{R}^d$ and satisfies $\|\nu_E(x)\| = 1$. The function $-\nu_E : \partial^* E \to \mathbb{R}^d$ is called the De Giorgi outer normal *to E in x.*

Finally, the following measure-theoretic concepts will be used in our analysis.

Definition 4.9 *Let $E \subset \mathbb{R}^d$ be a Borel set. We set, for $x \in \mathbb{R}^d$ and $0 \leq k \leq d$,*

$$\delta_E^k(x) = \liminf_{\rho \to 0^+} \frac{\mathcal{H}^k(E \cap B(x,\rho))}{\omega_k \rho^k},$$

where ω_k is the k-dimensional Lebesgue measure of the unit ball in $\mathbb{R}^k$. It is well known that for $k = d$ the limit actually exists and is equal to 1 for $\mathcal{L}^d$-a.e. $x \in E$.

Definition 4.10 *Let $E \subseteq \mathbb{R}^d$ be $\mathcal{L}^d$-measurable. We define (see p. 158 in [2]):*

$$\begin{aligned} E^0 &:= \{x \in \mathbb{R}^d : \delta_E^d(x) = 0\}, && \text{the measure theoretic exterior of } E; \\ E^1 &:= \{x \in \mathbb{R}^d : \delta_E^d(x) = 1\}, && \text{the measure theoretic interior of } E; \\ \partial_M E &:= \mathbb{R}^d \setminus (E^0 \cup E^1), && \text{the measure theoretic boundary of } E. \end{aligned}$$

Concerning the relations among the above introduced concepts of boundary, we recall the following (see Theorem 3.61, p. 158, in [2]).

Theorem 4.5 (De Giorgi-Federer) *Let* Ω *be a nonempty open subset of* $\mathbb{R}^d$ *and* $E \subseteq \mathbb{R}^d$ *be a set of finite perimeter in* Ω*. Then*

$$\partial^* E \cap \Omega \subseteq \Big\{x \in \mathbb{R}^d : \delta_E^d(x) = 1/2\Big\} \subseteq \partial_M E \subseteq \partial E,$$

and

$$\mathcal{H}^{d-1}\Big(\Omega \setminus (E^0 \cup \partial^* E \cup E^1)\Big) = 0.$$

In particular, E *has density either* 0*, or* $\frac{1}{2}$*, or* 1 *at* $\mathcal{H}^{d-1} -$ *a.e.* $x \in \Omega$*, and* $\mathcal{H}^{d-1}(\partial_M E \setminus \partial^* E) = 0$*.*

We conclude this subsection with the following criterion for sets of finite perimeter.

Theorem 4.6 (Federer) *Let* Ω *be a nonempty open subset of* $\mathbb{R}^d$ *and* $E \subseteq \mathbb{R}^d$ *be measurable. If* $\mathcal{H}^{d-1}(\partial(\Omega \cap E)) < +\infty$ *then* $P(E, \Omega) < +\infty$*.*

4.5 External sphere condition and semiconcavity

We now introduce new concepts for sets which is associated with semiconcavity concepts. Basing on these ones, we can extend to study the regularity of a class of continuous functions which is applied to time optimal control.

Definition 4.11 *Let $Q \subset \mathbb{R}^d$ be closed and $v \in \mathbb{R}^d$. We say that v is a proximal normal vector to Q at $x \in \partial Q$, denoted by $v \in N_Q^P(x)$, if there exists a constant $\sigma > 0$ such that*

$$\langle v \ , \ y - x \rangle \leq \sigma \cdot |y - x|^2, \quad \forall y \in Q. \tag{4.20}$$

Equivalently $v \in N_Q^P(x)$ if and only if there exists $\lambda > 0$ such that $\pi_Q(x + \lambda v) = \{x\}$.

Definition 4.12 *Let $Q \subset \mathbb{R}^d$ be closed and $x \in \partial Q$. The vector $v \in N_Q^P(x)$ is realized by a ball of radius ρ if and only if (4.20) satisfies for $\sigma = \frac{|v|}{2\rho}$.*

We are ready to give the main concept for this subsection.

Definition 4.13 *Let $Q \subset \mathbb{R}^d$ be closed and let $\theta(\cdot) : \partial Q \to (0, \infty)$ be continuous. We say that Q satisfies the $\theta(\cdot)$-external sphere condition if and only if for every $x \in \partial Q$, there exists a vector $v_x \neq 0$ such that $v_x \in N_Q^P(x)$ is realized by a ball of radius $\theta(x)$, i.e.,*

$$\left\langle \frac{v_x}{|v_x|}, y - x \right\rangle \leq \frac{1}{2\theta(x)} |y - x|^2.$$

for all $y \in Q$.

We will say that Q satisfies the ρ_0-external sphere condition for a constant $\rho_0 > 0$ if $\rho(\cdot) = \rho_0$. We are now going to study the main properties of sets which satisfies an external sphere condition.

Theorem 4.7 (Locally finite perimeter) *Let $Q \subset \mathbb{R}^d$ be closed. Assuming that Q satisfies the $\theta(\cdot)$-external sphere condition. Then, $\partial Q \cap \overline{\mathcal{O}}$ is finitely $\mathcal{H}^{d-1}$-rectifiable for any bounded, open set $\mathcal{O}$. In particular, Q has locally finite perimeter.*

Proof. Since $\mathcal{O}$ is bounded, we have that $\overline{\mathcal{O}}$ is compact. Therefore, there is a constant $\rho_0 > 0$ such that for every $x \in \partial Q \cap \overline{\mathcal{O}}$, there exists a unit vector $v_x \in N_Q^P(x)$ is realized by a ball of radius ρ_0, i.e.,

$$\langle v_x, y - x \rangle \leq \frac{1}{2\rho_0} |y - x|^2.$$

for all $y \in Q$.

Step 1: By the compactness of $\mathbb{S}^{d-1}$, we can find $M_1 \in \mathbb{N}$ and a finite set $\{v_1, ..., v_{M_1}\} \subset \mathbb{R}^{d-1}$ such that

$$\mathbb{S}^{d-1} \subset \bigcup_{i=1}^{M_1} v_i + \frac{1}{3} B'(0, 1)$$

where $\mathbb{S}^{d-1} = \{v \in \mathbb{R}^d \mid |v| = 1\}$ is the unit sphere with center 0. We partition ∂Q as

$$\partial Q = \bigcup_{i=1}^{M_1} \partial Q_i \tag{4.21}$$

where

$$\partial Q_i \ := \ \Big\{ x \in \partial Q \ \Big| \ |v_x - v_i| \leq \frac{1}{3} \Big\}.$$

One has first that

$$\partial Q \cap \overline{\mathcal{O}} = \bigcup_{i=1}^{M} \partial Q_i \cap \overline{\mathcal{O}}. \tag{4.22}$$

Moreover, for every $x \in \partial Q_i \cap \overline{\mathcal{O}}$, it holds

$$\begin{aligned} \langle v_i, y - x \rangle \ &\leq \ \langle v_x, y - x \rangle + |v_i - v_x| \cdot |y - x|, \quad \forall y \in Q \\ &\leq \ \Big(\frac{1}{2\rho_0} |y - x| + \frac{1}{3} \Big) \cdot |y - x| \end{aligned}$$

for all $y \in Q$. Therefore, for every $x, y \in \partial Q_i \cap \overline{\mathcal{O}}$, we have that

$$|\langle v_i, y - x \rangle| \leq \Big(\frac{1}{2\rho_0} |y - x| + \frac{1}{3} \Big) \cdot |y - x|. \tag{4.23}$$

Step 2: We are now going to show that $\partial Q_i \cap \overline{\mathcal{O}}$ is finitely $\mathcal{H}^{d-1}$-rectifiable for all $i \in \{1, .., M\}$. Fixing any $i \in \{1, .., M_1\}$, since $\partial Q_i \cap \overline{\mathcal{O}}$ is compact, there exists $M_2 \in \mathbb{N}$ and $x_1, ..., x_{M_2}$ such that

$$\partial Q_i \cap \overline{\mathcal{O}} \subset \bigcup_{k=1}^{M_2} B'(x_k, \delta),$$

where $\delta = \frac{\rho_0}{6}$. Setting $\partial Q_i^k = \partial Q_i \cap \overline{\mathcal{O}} \cap B'(x_k, \delta)$, we have that

$$\partial Q_i \cap \overline{\mathcal{O}} = \bigcup_{k=1}^{M_2} \partial Q_i^k. \tag{4.24}$$

Moreover, by (4.23) and the choice of δ, we have that for every $x, y \in \partial Q_i^k$, it holds

$$|\langle v_i, y - x \rangle| \leq \frac{1}{2} \cdot |y - x|.$$

Now, let $v_i^{\perp}$ be the subspace of $\mathbb{R}^d$ which is orthogonal to v_i. Let $\pi_i(\cdot)$ be the projection on $v_i^{\perp}$. From (4.24), one shows that

$$|\pi_i(y) - \pi_i(x)| \geq \frac{1}{\sqrt{2}} \cdot |y - x|, \quad \forall x, y \in \partial Q_i^k.$$

Thus, $\pi_i : \partial Q_i^k \to v_i^{\perp}$ is injective. Hence, if we set $A_i^k = \pi_i(\partial Q_i^k) \subset v_i^{\perp}$, the map $\pi_i^{-1} : A_i^k \to Q_i^k$ is Lipschitz with constant $\sqrt{2}$. Therefore, $\partial Q_i \cap \overline{\mathcal{O}}$ is finitely $\mathcal{H}^{d-1}$-rectifiable. By recalling (4.22), $\partial Q \cap \overline{\mathcal{O}}$ is finitely $\mathcal{H}^{d-1}$-rectifiable.

Finally, noting that $\mathcal{H}^{d-1}(A_i^k) < +\infty$, it implies that $\mathcal{H}^{d-1}(\partial Q_i^k) \leq 2^{\frac{d-1}{2}} \mathcal{H}^{d-1}(A_i^k) \leq +\infty$. Recalling (4.22), we obtain that $Q \cap \mathcal{O}$ has finite perimeter. The proof is complete.

Theorem 4.8 *Let $Q \subset \mathbb{R}^d$ be closed. Assuming that Q satisfies the $\theta(\cdot)$-external sphere condition. For every $k \in \{1, ..., d-1\}$, we denote by*

$$\partial Q^k = \big\{x \in \partial Q \mid \dim_{\mathcal{H}} N_Q^P(x) \geq k\big\}.$$

Then, ∂Q^k is countably $\mathcal{H}^{d-k}$-rectifiable.

The proof is based of the same technique of the previous theorem. A more general result is studied in [19].

Exercise 15. Proving the about theorem for $k = 2$.

We now recall the definition of Fréchet normal vector of a set.

Definition 4.14 *Let $Q \subset \mathbb{R}^d$ be closed and $v \in \mathbb{R}^d$. We say that v is a Fréchet normal vector to Q at x, denoted by $v \in N_Q^F(x)$, if*

$$\limsup_{y \in Q \to x} \left\langle v \ , \ \frac{y-x}{|y-x|} \right\rangle \leq 0. \tag{4.25}$$

Lemma 4.2 *Let $Q \subset \mathbb{R}^d$ be closed. Assuming that Q satisfies a $\theta(\cdot)$-external sphere condition. Then, the map $N_Q^F(\cdot) : \partial Q \rightrightarrows \mathbb{R}^d$ is upper-semicontinuous, i.e.,*

$$\lim_{y \to x} N_Q^F(y) \subseteq N_Q^F(x).$$

Proposition 4.7 *Let $Q \subset \mathbb{R}^d$ be closed. Assuming that Q satisfies a $\theta(\cdot)$-external sphere condition. Then, the set Q is smooth in ∂Q^1, i.e., for every $x \in \partial Q^1$, it holds*

$$\lim_{y \in \partial Q \to x} \left\langle v_x \ , \ \frac{y-x}{|y-x|} \right\rangle = 0,$$

where v_x is the unique unit proximal normal vector to Q at x.

Proof. Assume by a contradiction, there exists a sequence $\{y_n\} \subset \partial Q$ converging to x such that

$$\left\langle -v_x \ , \ \frac{y_n-x}{|y_n-x|} \right\rangle \geq \delta \tag{4.26}$$

for a constant $\delta > 0$ and for all $n \in \mathbb{N}$. Let v_n be the unit proximal normal vector to Q at y_n realized by a ball of radius $\theta(y_n)$. Since y_n converges to x, there exists a constant ρ_0 such that for every n, it holds

$$\langle v_n, z - y_n \rangle \leq \rho_0 \cdot |z - y_n|^2, \quad \forall z \in Q. \tag{4.27}$$

Therefore, v_n must converge to v_x. On the other hand, from the above inequality, we get in particularly that

$$\langle v_n, x - y_n \rangle \leq \rho_0 \cdot |x - y_n|^2.$$

It implies that

$$\left\langle v_n \ , \ \frac{x - y_n}{|x - y_n|} \right\rangle \leq \rho_0 \cdot |x - y_n|.$$

Since $\lim_{n\to\infty} v_n = v_x$ and $\lim_{n\to\infty} y_n = x$, we obtain that

$$\left\langle v_x \ , \ \frac{x - y_n}{|x - y_n|} \right\rangle \leq 0.$$

This contradicts to (4.26). The proof is complete.

We are now discussing the connection between external sphere condition and semiconcavity. Let $f : \Omega \to \mathbb{R}$ be upper semicontinuous where $\Omega \subset \mathbb{R}^n$ is open. We denote by

$$\text{hypo}(f) := \{(x, \beta) \mid x \in \Omega, \beta \leq f(x)\}$$

the hypograph of f.

Theorem 4.9 *The function f is locally semiconcave in Ω if and only if f is locally Lipschitz and* hypo(f) *satisfies a $\theta(\cdot)$ external sphere condition.*

Proof. Assume that f is locally semiconcave. From proposition (4.2), we have that f is locally Lipschitz in Ω. We now prove that hypo(f) satisfies a $\theta(\cdot)$ external sphere condition. For every $x \in \Omega$, there exists $v_x \in Df^-(x)$ such that

$$f(y) - f(x) - \langle v_x, y - x \rangle \leq \frac{C_x}{2} \cdot |y - x|^2, \quad \forall y \in B(x, \delta_x) \tag{4.28}$$

where C_x is a suitable constant and δ_x is a suitable constant such that $B(x, \delta_x) \subset \Omega$. It implies that

$$\Big\langle (-v_x, 1), (y - x, f(y) - f(x)) \Big\rangle \leq \frac{C_x}{2} \cdot |y - x|^2, \quad \forall y \in B(x, \delta_x).$$

Therefore, there exists $\rho_x > 0$ be such that

$$\left\langle \frac{(-v_x, 1)}{|(-v_x, 1)|}, (y - x, \beta - f(x)) \right\rangle \leq \rho_x \cdot \Big(|y - x|^2 + |\beta - f(x)|^2 \Big)$$

Thus, $(-v_x, 1) \in N^P_{\text{hypo}(f)}(x, f(x))$ is realized by a ball of radius $\frac{1}{2\rho_x}$. From here, one can show that hypo(f) satisfies a $\theta(\cdot)$ external sphere condition.

For the reversed side, we prefer to leave as an **exercise 16**.

Remark 4.2 *Thereom 4.3 is a particular case of theorem 4.8. Moreover, (ii)-(iv) of proposition 4.5 are consequences of proposition 4.7.*

4.6 A class of continuous functions

In this subsection, we will study the regularity properties a class of continuous functions whose hypograph satisfies an external sphere condition. From theorem 4.9, such class is a generalization of the class of semiconcave functions and is applied to study the regularity of the minimum time function under a weak controllability condition.

Let $\Omega \subset \mathbb{R}^n$ be open. For $\rho > 0$, we denote by

$$\mathcal{F}_\rho(\Omega, \mathbb{R}) = \{f \in \mathcal{C}(\Omega, \mathbb{R}) \mid \text{hypo}(f) \text{ satifies the } \rho - \text{extermal sphere condition}\}$$

where $\mathcal{C}(\Omega, \mathbb{R})$ is the class of continuous function from Ω to $\mathbb{R}$.

Theorem 4.10 *For every $\rho > 0$, it holds that*

$$\mathcal{F}_\rho(\Omega, \mathbb{R}) \subset BV_{loc}(\Omega, \mathbb{R}).$$

Proof. Let f be in $\mathcal{F}_\rho(\Omega, \mathbb{R})$. Given any $U \subset \Omega$ open and bounded, we need to show that $f \in BV(U, \mathbb{R})$. Indeed, since $\overline{U}$ is compact and f is continuous, there exist $M > 0$ such that $\|f\|_{\mathbb{L}^\infty(\overline{U}, \mathbb{R})} < M$. Thus, recalling theorem 4.7, we obtain that

$$P(\text{hypo}(f), U \times \mathbb{R}) = P(\text{hypo}(f), U \times [-M, M]) < \infty.$$

Hence, for every $\varphi \in C_c^\infty(U \times]-M, M[)$, we have:

$$\begin{aligned}
\int_{U\times]-M,M[} \chi_{\text{hypo}\, f}(x,t) \,\text{div}\, \varphi(x,t)\, dx\, dt &= \int_{]-M,M[} \left(\int_U \chi_{\text{hypo}\, f}(x,t) \,\text{div}\, \varphi(x,t)\, dx \right) dt \\
&= \int_{]-M,M[} \left(\int_U \chi_{]-\infty, f(x)]}(t) \,\text{div}\, \varphi(x,t)\, dx \right) dt \\
&= \int_{\mathbb{R}} \left(\int_U \chi_{]-\infty, f(x)]}(t) \,\text{div}\, \varphi(x,t)\, dx \right) dt \le c.
\end{aligned}$$

For every $t \in]-M, M[$, let $\psi_t \in C_c^\infty(]-M, M[)$ be such that $\psi_t(]-M, M[) \subseteq [0,1]$ and $\psi_t(s) = 1$ if s belongs to a neighborhood of t. We have for every $t \in]-M, M[$:

$$\begin{aligned}
P(\{x \in U : f(x) \ge t\}, U) &= \sup \left\{ \int_U \chi_{]-\infty, f(x)]}(t) \text{div}\, \sigma(x)\, dx : \sigma \in C_c^\infty(U),\ \|\sigma\|_\infty \le 1 \right\} \\
&= \sup \left\{ \int_U \chi_{]-\infty, f(x)]}(t) \text{div}\, (\sigma(x)\psi_t(t))\, dx : \sigma \in C_c^\infty(U),\ \|\sigma\|_\infty \le 1 \right\}
\end{aligned}$$

Setting $\varphi(x,s) = \sigma(x)\psi_t(s)$, we have that $\varphi \in C_c^\infty(U \times]-M, M[)$ and $\|\varphi\|_\infty \le 1$, whence

$$\int_{\mathbb{R}} P(\{x \in U : f(x) \ge t\}, U)\, dt \le c.$$

According to Theorem 1, p. 185 in [25], we have that $f \in BV(U, \mathbb{R})$. $\square$

We now introduce the definition of horizontal superdifferential.

Definition 4.15 *Let $f : \Omega \to \mathbb{R}$ be continuous. For every $x \in \Omega$, the unit vector $v \in \mathbb{S}^{n-1}$ is a horizontal superdifferential of f at x, denoted by $v \in \partial^\infty f(x)$, if*

$$(-v, 0) \in N^P_{\text{hypo}(f)}(x, f(x)).$$

We also set

$$\mathcal{S}_f = \{x \in \Omega \mid \partial^\infty f(x) \neq \varnothing\}.$$

Lemma 4.3 *Let $f : \Omega \to \mathbb{R}$ be continuous. If f is Lipschitz in a neighborhood of $x \in \Omega$ then the set $\partial^\infty f(x)$ is empty.*

Exercise 17: Proving the above Lemma.

From the above Lemma, if f is locally semiconcave in Ω, the set $\partial^\infty f(x)$ is empty for every $x \in \Omega$. However, we consider here $f \in \mathcal{F}_\rho(\Omega, \mathbb{R})$. The horizontal superdifferential may appeal at the non-lipschitz points. Our main goal is now to study the properties of $\mathcal{S}_f$.

Proposition 4.8 *Assuming that $f \in \mathcal{F}_\rho(\Omega, \mathbb{R})$. Then, the set $\mathcal{S}_f$ is closed in Ω.*

Skecth of proof.
Main step: For every $x \in \mathcal{S}_f$, there exists $v \in \mathbb{S}^{n-1}$ such that $(-v, 0) \in N^P_{\text{hypo}(f)}(x, f(x))$ is realized by a ball of radius ρ, i.e.,

$$\langle -v, y - x\rangle \leq \rho \cdot (|y - x|^2 + |\beta - f(x)|^2), \quad \forall y \in \Omega, \beta \leq f(y).$$

Indeed, let $(-w, 0) \in N^P_{\text{hypo}(f)}(x, f(x))$, along the ray $x(t) = x - t \cdot w$ $(t > 0)$, by using Clarke's density theorem, one can find a sequence x_n converge to x such that f is differentiable at x_n and $\lim_{n\to\infty} |Df(x_n)| = +\infty$. Moreover, since f is differentiable at x, we have that $(-Df(x_n), 1) \in N^P_{\text{hypo}(f)}(x_n, f(x_n))$ realized by a ball of radius ρ. Therefore, there exists a subsequence $\{x_{n_k}\}$ converge to x such that

$$\lim_{n_k\to\infty} \frac{(-Df(x_{n_k}, 1)}{|(-Df(x_{n_k}, 1)|} = (-v, 0).$$

Thus, $(-v, 0) \in N^P_{\text{hypo}(f)}(x, f(x))$ is realized by a ball of radius ρ.

Last step: Taking any $x_n \in \mathcal{S}_f$ converging to $x \in \Omega$. From the main step, there exists $v_n \in \mathbb{S}^{n-1}$ such that $(-v_n, 0) \in N^P_{\text{hypo}(f)}(x_n, f(x_n))$ is realized by a ball of radius ρ. Therefore, there exists $(-v_x, 0) \in N^P_{\text{hypo}(f)}(x, f(x))$ is realized by a ball of radius ρ. Thus, the proof is complete. □

Before going to study the size of $\mathcal{S}_f$, we need the following key Lemma.

Lemma 4.4 *Let f be in $\mathcal{F}_\rho(\Omega, \mathbb{R})$. Let $x \in \mathcal{S}_f$ be such that $N^P_{\text{hypo}(f)}(x, f(x)) = \mathbb{R}^+(v, 0)$ for some $v \in \mathbb{S}^{n-1}$. Then, there exists $\delta_0 = \delta_0(x) > 0$ such that*

$$\|Df\|_{\text{Sq}(x,\delta)} \geq 2^{n-2} \cdot \delta^{n-\frac{1}{2}}, \quad \forall 0 < \delta < \delta_0 \tag{4.29}$$

where $\mathrm{Sq}(x,\delta) := \big\{(y_1,\dots,y_n)\in\mathbb{R}^n : \max_{i=1,\dots,n}|y_i-x_i|<\delta\big\}$. *Thus,*

$$\|Df\|_{B(x,\delta)} \geq 2^{-\frac{3}{2}}\cdot\delta^{n-\frac{1}{2}},\quad \forall 0<\delta<\delta_0. \tag{4.30}$$

Proof. Without loss of generality we may assume that

$$x=0\in\Omega,\quad f(x)=0\quad\text{and}\quad N^P_{\mathrm{hypo}(f)}(0,0)=\mathbb{R}^+(e_1,0).$$

For any $\delta>0$ define

$$R_\delta := \{y=(y_1,\dots,y_n)\in\mathrm{Sq}(0,\delta) : \frac{3}{4}\delta<y_1<\delta\},$$
$$S_\delta := \{y=(y_1,\dots,y_d)\in\mathrm{Sq}(0,\delta) : -\delta<y_1<-\delta/2\}\,.$$

Claim: There exist $\delta_1,\delta_2>0$ such that

$$f(y)\leq -\frac{1}{2}\,\delta^{\frac{1}{2}},\quad \forall y\in R_\delta, \delta<\delta_1 \tag{4.31}$$

and

$$f(y)>0\quad \forall y\in S_\delta,\delta<\delta_2. \tag{4.32}$$

Proof of Claim: Let us prove (4.31). For $y\in R_\delta$ we have

$$\tfrac{3}{4}\delta < \langle(e_1,0),(y,\beta)\rangle \leq \rho\cdot\big(\|y\|^2+|\beta|^2\big)\quad \forall\beta\leq f(y)$$

whence

$$\tfrac{3}{4}\delta \leq \rho\cdot\big(n\delta^2+|\beta|^2\big),\quad \forall\beta\leq f(y). \tag{4.33}$$

Notice that, for δ small enough, it holds

$$f(y)<0\qquad\text{for any } y\in R_\delta;$$

indeed, by contradiction, if $f(y)\geq 0$ one could choose $\beta=0$, thus violating (4.33) for δ sufficiently small. Formula (4.31) easily follows for a small enough δ_1 on taking $\beta=f(y)<0$ in (4.33).

Let us prove (4.32). Assume by contradiction that there exist sequences $\{\delta_k\}_k$ and $\{y_k\}_k$ such that

$$\delta_k\to 0^+,\quad y_k\in S_{\delta_k}\quad\text{and}\quad f(y_k)\leq 0.$$

Since $y_k\to 0$ and $f(0)=0$ we get $\lim_{k\to\infty}f(y_k)=f(0)=0$. Let $(v_k,\alpha_k)\in N^P_{\mathrm{hypo}(f)}(x_k,f(x_k))$ be realized by a ball of radius ρ and $(v_k,\alpha_k$ converges to $(e_1,0)$. Moreover, for all $\beta\leq 0=f(0)$ it holds

$$\Big\langle\frac{(v_k,\alpha_k)}{|(v_k,\alpha_k)|},(0,\beta)-(y_k,f(y_k))\Big\rangle \leq \rho\cdot\big(\|y_k\|^2+|\beta-f(y_k)|^2\big).$$

Since $f(y_k) \leq f(0) = 0$ we can choose $\beta = f(y_k)$ in the above inequality and get

$$\langle v_k, -y_k \rangle \leq C \cdot \|y_k\|^2.$$

Thus,

$$\tfrac{\delta_k}{2} - \|v_k - e_1\| \sqrt{n}\, \delta_n \leq \langle e_1, -y_k \rangle + \langle v_k - e_1, -y_k \rangle = \langle v_k, -y_k \rangle \leq Cn\, \delta_k^2.$$

Dividing both sides by δ_k and passing to the limit as $k \to \infty$ we obtain a contradiction. This concludes the proof of the Claim. $\diamond$

Last step: The Claim allows us to conclude: indeed, for any $\delta < \delta_0 := \min\{\delta_1, \delta_2\}$ and any $z \in (-\delta, \delta)^{n-1}$ we get

$$|f(y_a, z) - f(y_b, z)| \geq \frac{1}{2}\delta^{\frac{1}{2}} \quad \forall y_a \in]\tfrac{3}{4}\delta, \delta[, y_b \in]-\delta, -\delta/2[.$$

By virtue of Lemma 4.1, for any $z \in (-\delta, \delta)^{d-1}$ there exist $y_a(z) \in]\frac{3}{4}\delta, \delta[$ and $y_b(z) \in]-\delta, -\delta/2[$ such that

$$\|Df_z\|(-\delta, \delta) \geq |f(y_a(z), z) - f(y_b(z), z)| \geq \frac{1}{2}\delta^{\frac{1}{2}}$$

where $f_z := f(\cdot, z)$. We obtain

$$\begin{aligned} \|Df\|(\mathrm{Sq}(0,\delta)) &\geq \int_{]-\delta,\delta[^{n-1}} \|D_{e_1} f\|(z+]-\delta,\delta[e_1)dz \\ &= \int_{]-\delta,\delta[^{n-1}} \|Df_z\|(-\delta,\delta)dz \\ &\geq (2\delta)^{n-1} \cdot \tfrac{1}{2}\delta^{\frac{1}{2}} = 2^{n-2}\delta^{n-\frac{1}{2}}, \end{aligned}$$

where we have denoted by $D_{e_1} f$ the distributional derivative of f along e_1 and by $z+]-\delta, \delta[\cdot e_1$ the line segment joining $(-\delta, z)$ and (δ, z). This concludes the proof of the Lemma. $\square$

We are ready to prove the main result.

Theorem 4.11 *Let f be in $\mathcal{F}_\rho(\Omega, \mathbb{R})$. Then, $\mathcal{H}^{n-\frac{1}{2}}(\mathcal{S}_f) \cap U$ is finite for every $U \subset \Omega$ open and bounded.*

Proof. We divide the set $\mathcal{S}_f$ into two sets:

$$\mathcal{S}_f = \mathcal{S}_f^1 + \mathcal{S}_f^2$$

where

$$\begin{aligned} \mathcal{S}_f^1 &:= \Big\{ x \in \mathcal{S}_f \mid N^P_{\mathrm{hypo}(f)}(x, f(x)) = \mathbb{R}^+(v, 0) \Big\}, \\ S_f^2 &:= \Big\{ x \in \mathcal{S}_f \mid \dim_{\mathcal{H}} N^P_{\mathrm{hypo}(f)}(x, f(x)) \geq 2 \Big\}. \end{aligned}$$

Recalling 4.8, one can show that S_f^2 is $\mathcal{H}^{n-1}$-countably rectifiable. In particular, $\mathcal{H}^{n-\frac{1}{2}}(\mathcal{S}_f^2) = 0$. Hence, we only need to prove that $\mathcal{H}^{n-\frac{1}{2}}(\mathcal{S}_f^1) \cap U$ is finite.

We can construct a covering of $S_f^1 \cap U$ by setting:

$$\mathcal{B} := \left\{ (x + r\overline{\mathbb{B}}^n) : x \in S_f^1 \cap U, \, r < \frac{\min\{\delta_0(x), \mathrm{dist}(U, \partial\Omega)/2\}}{10} \right\}.$$

Since $\mathcal{B}$ is a fine covering of $S_f^1 \cap U$, by using Vitali's covering Theorem, there exists a countable subset of pairwise disjoint balls $\mathcal{B}' := \{x_i + r_i\overline{\mathbb{B}^n} : i \in \mathbb{N}\} \subset \mathcal{B}$ such that

$$\bigcup_{B \in \mathcal{B}} B \subseteq \bigcup_{i=1}^{\infty} (x_i + 5r_i\mathbb{B}^n),$$

which implies that $\{x_i + 5r_i\mathbb{B}^n : i \in \mathbb{N}\}$ is a covering of $S_f^1 \cap U$ and

$$\bigcup_{i \in \mathbb{N}} (x_i + 5r_i\mathbb{B}^n) \subseteq (U + c\mathbb{B}^n) =: W$$

for a suitable constant $c > 0$ and thus W is an open bounded subset of Ω.

$$\begin{aligned} |Df|(W) &\geq |Df| \left(\bigcup_{i=1}^{\infty} (x_i + 5r_i\mathbb{B}^n) \right) \\ &\geq \sum_{i=1}^{\infty} |Df|(x_i + r_i\mathbb{B}^n) \geq \sum_{i=1}^{\infty} 2^{-\frac{3}{2}} \cdot r_i^{n-\frac{1}{2}} \geq C \cdot \mathcal{H}^{n-\frac{1}{2}}(\mathcal{S}_f^1). \end{aligned}$$

On the other hand, from theorem 4.10, we have that $|Df|(W) < +\infty$. Therefore, $\mathcal{H}^{n-\frac{1}{2}}(\mathcal{S}_f^1) < +\infty$. The proof is complete. □

Corollary 4.4 *Let f be in $\mathcal{F}_\rho(\Omega, \mathbb{R})$. Then, $\mathcal{L}^n(\mathcal{S}_f) = 0$.*

We conclude this subsection with the following theorems.

Theorem 4.12 *Let f be in $\mathcal{F}_\rho(\Omega, \mathbb{R})$. Then, f is locally semiconcave in the open set $\Omega \backslash \mathcal{S}_f$. In particular, f is a.e. twice diffrentiable in Ω.*

Excercise 18. Prove the above theorem.

Theorem 4.13 *Let $f : \Omega \to \mathbb{R}$ be continuous. Assuming that* hypo(f) *satisfies a $\theta(\cdot)$-external sphere condition. Then, f is locally semiconcave in the open set $\Omega \backslash \mathcal{S}_f$.*

4.7 Application to time optimal control

There are several cases showing that the controllability condition (H3) in subsection 4.3 does not hold, or the minimum time function $T_S(\cdot)$ is not locally Lipschitz (e.g., the rocket car case). Therefore, $T_S(\cdot)$ may not be locally semiconcave and differentiable almost everywhere. A natural problem is to study T_S under a weaker controllability assumption. More precisely, in this subsection we will study the regularity of T_S when T_S is just continuous.

Consider the control systems

$$\begin{cases} \dot{x}(t) = f(x(t), u(t)), \quad t \in [0, +\infty[\ a.e., \\ x(0) = x_0, \end{cases} \tag{4.34}$$

where $x_0 \in \mathbb{R}^n$ and

+ $f : \mathbb{R}^n \times U \to \mathbb{R}^n$ is the dynamics of the control system
+ $U \subset \mathbb{R}^m$ is the control set
+ $u : [0, +\infty[\to U$ is a control function.

Standard hypotheses

(H1) $f : \mathbb{R}^n \times U \to \mathbb{R}^n$ is Lipschitz

$$|f(y, u) - f(x, u)| \leq L_1 \cdot |y - x|, \quad \forall x, y \in \mathbb{R}^n, u \in U. \tag{4.35}$$

Moreover, the gradient of f with respect to x exists everywhere and is locally Lipschitz in x, uniformly in u.

(H2) U is compact.

The set of admissible control is

$$\mathcal{U}_{ad} = \{u : [0, \infty) \to U \mid u \text{ is measurable}\}.$$

For every $u \in \mathcal{U}_{ad}$, we recall that $y^{x_0,u}(\cdot)$ is the trajectory staring from x with control u which is the unique solution of (4.11). The minimum time needed to steer x to the closed target $\mathcal{S}$, regarded as a function of x, is called the minimum time function and is denoted by

$$T_{\mathcal{S}}(x) := \inf \{t \geq 0 \mid y^{x,u}(t) \in \mathcal{S}, u \in \mathcal{U}_{ad}\}. \tag{4.36}$$

A condition on $\mathcal{S}$

(S) The target $\mathcal{S}$ satisfies the ρ_0-internal sphere condition, i.e., for every $x \in \partial\mathcal{S}$, there exists a unit vector $v \in \mathbb{R}$ such that

$$\langle v, y - x \rangle \leq \frac{1}{2\rho_0} \cdot |y - x|^2, \quad \forall y \in \mathbb{R}^n \backslash \mathcal{S}.$$

Since the internal sphere property is closed with respect to the union operator, one can see intuitively that the sublevel set,

$$\mathcal{S}(r) = \{x \in \mathbb{R}^n \mid T_{\mathcal{S}}(x) \le r\},$$

the set of points reachable from $\mathcal{S}$ in time less than equal to r, inherits such property from $\mathcal{S}$.

Proposition 4.9 *Let (H1),(H2) and (S) hold. Assume that $T_{\mathcal{S}}(\cdot)$ is continuous in $\mathbb{R}^n$. Then, for every $r > 0$, the sublevel set $\mathcal{S}(r)$ satisfies the $\rho(r)$-internal sphere condition where $\rho(r) > 0$ can be computed.*

Proof. Given any $\bar{x} \in \partial\mathcal{S}(r)$, let $\bar{u} \in \mathcal{U}_{ad}$ be an optimal control steering x to $\mathcal{S}$ in time $T(x)$. We set $x(t) = y^{\bar{x},\bar{u}}(t)$ and $x_r = x(r) \in \partial\mathcal{S}$. Since $\mathcal{S}$ satisfies the ρ_0-internal sphere condition, there exists a unit vector $p_r \in N^P_{\overline{\mathcal{S}^c}}(x_0)$ such that

$$\langle p_r, z_r - x_r\rangle \le \frac{1}{2\rho_0} \cdot |z_r - x_r|^2, \quad \forall z_r \in \overline{\mathcal{S}^c}. \tag{4.37}$$

Let $p(\cdot)$ be the adjoint arc which solves the following ODE:

$$\begin{cases} \dot{p}(t) \;=\; -p(t) \cdot D_x f(x(t), \bar{u}(t)), \quad a.e.\ t \in [0, r], \\ \\ p(T(\bar{x})) \;=\; p_r. \end{cases} \tag{4.38}$$

We will prove now that $p(0)$ is a proximal normal vector to $\overline{\mathcal{S}^c(r)}$ at $\bar{x}$ realized by a ball of radius $\rho(r) > 0$, i.e.,

$$\left\langle \frac{p(0)}{|p(0)|}\ ,\ \bar{z} - \bar{x} \right\rangle \le \frac{1}{\rho(r)} \cdot |\bar{z} - \bar{x}|^2, \quad \forall \bar{z} \in \overline{\mathcal{S}^c(r)}. \tag{4.39}$$

Indeed, let $\bar{z}$ be in $\overline{\mathcal{S}^c(r)}$. One first has that $T_{\mathcal{S}}(\bar{z}) \ge T_{\mathcal{S}}(\bar{x})$. We set $z(t) = y^{\bar{z},\bar{u}}(t)$ and $z_r = z(r)$. By the dynamics programming principle, one can show that $z_0 \in \overline{\mathcal{S}^c}$, this implies that

$$\langle p_0, z_0 - x_0\rangle \le \frac{1}{2\rho_0} \cdot |z_0 - x_0|^2. \tag{4.40}$$

On the other hand, one can estimate that

$$-\frac{d}{dt}\langle p(t), z(t) - x(t)\rangle \le C_1 \cdot |z(t) - x(t)|^2. \tag{4.41}$$

Hence,

$$\langle p(0), \bar{z} - \bar{x}\rangle \le C_1 \cdot |z(t) - x(t)|^2 + \langle p_r, \bar{z_r} - \bar{x_r}\rangle.$$

By Gronwall's inequality, we have that $|z(t) - x(t)| \le \beta(t) \cdot |\bar{z} - \bar{x}|$ and $|p(0)| \le \alpha(t)$. Therefore, from the above estimate and (5.36), one concludes the proof. □

Our main result is the following.

Theorem 4.14 *Let (H1),(H2) and (S) hold. Assume that $T_S(\cdot)$ is continuous in $\mathbb{R}^n$. Then,* hypo(f) *satisfies the $\theta(\cdot)$-external sphere condition.*

Sketch of proof. Let $x(\cdot)$, $p(\cdot)$ be in the above proposition.
Step 1: Let

$$\lambda_{\bar{x}} = H(\bar{x}, p(0)) = \sup_{u \in U} \langle -p(0), f(\bar{x}, u) \rangle.$$

By using the dynamics programming principle, one shows that $\lambda_{\bar{x}} \geq 0$.

Step 2: Showing that

$$(p(0), \lambda_x) \in N^P_{\text{hypo}(T_{\mathcal{S}})}(\bar{x}, T_{\mathcal{S}}(\bar{x}))$$

is realized by a ball of radius $\theta_x > 0$, i.e., for all $\bar{z} \in \overline{\mathcal{S}^c}$, $\beta \leq T_{\mathcal{S}}(\bar{z})$, it holds

$$\langle p(0), \bar{z} - \bar{x} \rangle + \lambda_x \cdot (\beta - T_S(x)) \leq \frac{1}{2\theta_x} \cdot |(p(0), \lambda_x)| \cdot (|\bar{z} - \bar{x}|^2 + |\beta - T_{\mathcal{S}}(\bar{x})|^2). \quad (4.42)$$

Hint: Using the proposition 4.9 and step 1. □

Corollary 4.5 *Under the assumptions in theorem 4.14, $T_{\mathcal{S}}(\cdot)$ is locally semiconcave outside a closed set Γ where $\mathcal{H}^{n-\frac{1}{2}}(\Gamma \cap U) < \infty$, U is compact. In particular, $T_{\mathcal{S}}$ is a.e. twice differentiable.*

5 Compactness properties of solutions to Hamilton-Jacobi equations

Consider the first order Hamilton-Jacobi equation

$$\begin{cases} u_t(t,x) + H(\nabla_x u(t,x)) = 0, \\ \\ u(0,\cdot) = g(\cdot) \end{cases} \tag{5.1}$$

where $u : [0,+\infty[\times\mathbb{R}^n \to \mathbb{R}$ is the value function, $H : \mathbb{R}^n \to \mathbb{R}$ is the Hamilton function and the notation

$$u_t = \frac{\partial u}{\partial t}, \quad \nabla u = \Big(\frac{\partial u}{\partial x_1}, ..., \frac{\partial u}{\partial x_n}\Big).$$

The function H satisfies the standard assumptions:

(H1) H is coercive, i.e., $\lim_{|p|\to\infty} \frac{H(p)}{|p|} = +\infty$,

(H2) $H \in C^2(\mathbb{R}^n, \mathbb{R})$ and $D^2H(p) \geq \alpha \cdot \mathbb{I}_n$ for all $p \in \mathbb{R}^n$ where α is a positive constant.

Since H is convex and coercive, the Legendre transform L of H is denoted by

$$L(q) = \sup_{p\in\mathbb{R}^n} \{q\cdot p - H(p)\} \tag{5.2}$$

is convex and coercive. Moreover, (H2) implies that

$$L \in C^2(\mathbb{R}^n, \mathbb{R}) \quad \text{and} \quad D^2L(q) \leq \frac{1}{\alpha}\cdot \mathbb{I}_n \quad \forall q \in \mathbb{R}^n. \tag{5.3}$$

Hence, for every $g \in \mathrm{Lip}(\mathbb{R}^n, \mathbb{R})$, the HJ equation (5.1) admits a unique viscosity solution given by Hopf's formula:

$$V(t,x) = \min_{y\in\mathbb{R}^n} \Big\{ t\cdot L\Big(\frac{x-y}{t}\Big) + g(y)\Big\} \tag{5.4}$$

and $V(0,\cdot) = g(\cdot)$. We recall some main properties of the viscosity solution defined by Hopf's formulate:

Proposition 5.1 *Let V be the viscosity solution of (5.1). Then the followings hold*

(i) ***A functional identity:*** *for each $x \in \mathbb{R}^n$ and $0 \leq s < t$,*

$$V(t,x) = \min_{y\in\mathbb{R}^n} \Big\{ V(s,y) + (t-s)\cdot L\Big(\frac{x-y}{t-s}\Big)\Big\}.$$

(ii) ***The linear programming principle:*** *Let $0 < s < t$, $x \in \mathbb{R}^n$ and assume that y is a minimizer for (5.4). Let $z = \frac{s}{t}\cdot x + (1-\frac{s}{t})\cdot y$. Then y is the unique minimum for $s\cdot L(\frac{z-w}{s}) + u_0(w)$.*

(iii) ***Characteristics***: *(5.1) admits a unique minimum y_x if and only if $\nabla V(t,\cdot)$ is differentiable at x. Moreover, in this case we have that $y_x = x - t\cdot\nabla H(\nabla V(t,x))$.*

From (i) of the above proposition, if we define $S_t : \mathrm{Lip}(\mathbb{R}^n,\mathbb{R}) \to \mathrm{Lip}(\mathbb{R}^n,\mathbb{R})$ that

$$S_t(g)(\cdot) := \min_{y\in\mathbb{R}^n}\Big\{t\cdot L\Big(\frac{x-y}{t}\Big) + g(y)\Big\} \quad \text{and} \quad S_0(g) = g$$

for $t > 0$ and $g \in \mathrm{Lip}(\mathbb{R}^n,\mathbb{R})$, we have that

$$S_{t+s}(g) = S_t(S_s(g)).$$

Thus $S_t(\cdot)$ is a semigroup.

Exercise 19: Proving (iii) in the above proposition.

Lemma 5.1 *Let $T > 0$ and $u_0 \in \mathrm{Lip}(\mathbb{R}^n,\mathbb{R})$. Then $S_T(u_0)$ is semiconcave with semiconcave constant $\frac{2}{\alpha T}$.*

Proof. Fix any $x, h \in \mathbb{R}^n$, we need to show that

$$S_T(u_0)(x+h) + S_T(u_0)(x-h) - 2S_T(u_0)(x) \le \frac{2}{\alpha T}\cdot|h|^2. \tag{5.5}$$

Let y_x be such that

$$S_T(u_0)(x) = T\cdot L\Big(\frac{x-y_x}{T}\Big) + u_0(y_x).$$

By Hopf's formula, we have

$$S_T(u_0)(x\pm h) \le T\cdot L\Big(\frac{x\pm h-y_x}{T}\Big) + u_0(y_x).$$

It implies that

$$\begin{aligned} &S_T(u_0)(x+h) + S_T(u_0)(x-h) - 2S_T(u_0)(x) \\ &\qquad \le T\cdot\Big[L\Big(\frac{x+h-y_x}{T}\Big) + L\Big(\frac{x-h-y_x}{T}\Big) - 2L\Big(\frac{x-y_x}{T}\Big)\Big]. \end{aligned} \tag{5.6}$$

Recalling that L is semiconcave with semiconcave constant $\frac{2}{\alpha}$, we obtain (5.5). □

Therefore, by using Helly's theorem, one can show that for any $T > 0$ the map $S_T : \big(\mathrm{Lip}(\mathbb{R}^n,\mathbb{R}), \|\cdot\|_{W^{1,1}}\big) \to W^{1,1}_{loc}(\mathbb{R}^n,\mathbb{R})$ is compact. A natural question is to study a quantitative compactness of S_T.

Definition 5.1 (Komogorove ε-entropy) *Let (X,d) be a metric space and K a totally bounded subset of X. For $\varepsilon > 0$, let $N_\varepsilon(K \mid X)$ be the minimal number of sets in a cover of K by subsets of X having diameter no larger than 2ε. Then the ε-entropy of K is defined as*

$$\mathcal{H}_\varepsilon(K \mid X) \doteq \log_2 N_\varepsilon(K \mid X).$$

Given $L, M > 0$, we define

$$C_{[L,M]} = \{u_0 \in \text{Lip}(\mathbb{R}^n, \mathbb{R}) \mid \text{Supp(u}_0) \subset [-L, L]^n, \|\nabla u_0\|_{L^\infty(\mathbb{R}^n,\mathbb{R})} \leq M\}.$$

Our goal: is to show that for every $T > 0$ and for every $\epsilon > 0$ sufficiently small, it holds

$$\mathcal{H}_\epsilon\Big(S_T(C_{[L,M]}) \mid W^{1,1}_{loc}(\mathbb{R}^n, \mathbb{R})\Big) \simeq \frac{1}{\epsilon^n}. \tag{5.7}$$

5.1 An upper bound of $\mathcal{H}_\epsilon\Big(S_T(C_{[L,M]}) \mid W^{1,1}_{loc}(\mathbb{R}^n, \mathbb{R})\Big)$

To start this subsection, we recall some basic properties of the elements of $S_T(C_{[L,M]})$.

Lemma 5.2 *Given any $T > 0$ and $u_0 \in C_{[L,M]}$, we have that*

(i) $S_T^{u_0}(\cdot)$ is semiconcave with semiconcavity constant $\frac{2}{\alpha T}$.

(ii) $\|\nabla S_T(u_0)\|_{L^\infty(\mathbb{R}^n,\mathbb{R})} \leq M$ and $\text{Supp}(S_T(u_0) + \beta_T) \subset [-L_T, L_T]^n$ *where*

$$L_T = L + T \cdot \sup_{\|p\| \leq M} DH(p) \text{ and } \beta_T = T \cdot H(0). \tag{5.8}$$

Proof. (i) is prove in Lemma 5.1. Moreover, by recalling Theorem 2.5 we have that $\|\nabla S_T(u_0)\|_{L^\infty(\mathbb{R}^n,\mathbb{R})} \leq \|\nabla u_0\|_{L^\infty(\mathbb{R}^n,\mathbb{R})} \leq M$.

Now, let $x \notin \mathbb{R}^n \backslash [-L_T, L_T]^n$ be a differentiable point of $S_T(u_0)$. We will show that

$$S_T(u_0)(x) = -\beta_T.$$

Recalling (iii) of proposition 5.1, we have that

$$S_T(u_0)(x) = T \cdot L(DH(\nabla S_T(u_0)(x))) + u_0(y_x)$$

where $y_x = x - T \cdot DH(\nabla S_T(u_0)(x))$. Since $|\nabla S_T(u_0)(x)| \leq M$, $x \in \mathbb{R}^n \backslash [-L_T, L_T]^n$, by recalling (5.8), we get that $y_x \in \mathbb{R}^n \backslash [-L, L]^n$. It implies that $\nabla u_0(y_x) = 0$. Therefore, since y_x is a minimum of $T \cdot L\big(\frac{x-y}{T}\big) + u_0(y)$, over $\mathbb{R}^n$, one has that $-DL\big(\frac{x-y_x}{T}\big) = \nabla u_0(y_x) = 0$. Hence, $\nabla S_T(u_0)(x) = DL(DH(\nabla S_T(u_0)(x))) = 0$. Thus,

$$S_T(u_0)(x) = T \cdot L(DH(0)) = -T \cdot H(0) = -\beta_T.$$

The proof in complete. □

For any $K, L, M > 0$, we denote by

$$\mathcal{SC}_{[K,L,M]} = \{f \in \mathcal{SC}_{[K]} \mid \|Df\|_{\mathbb{L}^\infty(\mathbb{R}^n)} \leq M, \ \text{Supp}(f) \subseteq [-L, L]^N\}$$

where

$$\mathcal{SC}_{[K]} = \{f : \mathbb{R}^n \to \mathbb{R} \mid f \text{ is semiconcave with seminconcavity constant } K\}.$$

From Lemma 5.2, we immediately obtain that

$$S_T\big(C_{[L,M]}\big) + \beta_T \subset SC_{[\frac{2}{\alpha T}, L_T, M]} \tag{5.9}$$

where

$$L_T = L + T \cdot \sup_{\|p\| \le M} |DH(p)| \text{ and } \beta_T = T \cdot H(0).$$

This implies that

Corollary 5.1 *Given any $L, M > 0$ and $T > 0$. For any $\varepsilon > 0$, it holds*

$$\mathcal{H}_\varepsilon\Big(S_T(C_{[L,M]}) + \beta_T \mid W^{1,1}(\mathbb{R}^n, \mathbb{R})\Big) \le \mathcal{H}_\varepsilon\Big(SC_{[\frac{2}{\alpha T}, L_T, M]} \mid W^{1,1}(\mathbb{R}^n, \mathbb{R})\Big) \tag{5.10}$$

where

$$L_T = L + T \cdot \sup_{\|p\| \le M} |DH(p)| \text{ and } \beta_T = T \cdot H(0).$$

We are now going to study on upper estimate for a class of semiconcave functions in $W^{1,1}(\mathbb{R}^n, \mathbb{R})$. More precisely, we will show that

$$\mathcal{H}_\varepsilon\Big(SC_{[K,L,M]} \,\Big|\, W^{1,1}(\mathbb{R}^n, \mathbb{R})\Big) \le \Gamma_{[K,L,M]} \cdot \frac{1}{\varepsilon^n}$$

ONE DIMENSIONAL CASE (n=1): For every $f \in SC_{[K,L,M]}$, we first have that $g(\cdot) = f(\cdot) - \frac{K}{2} \cdot |\cdot|^2$ is concave. Therefore, $D^+g : \mathbb{R} \rightrightarrows \mathbb{R}$ is a monotone multifunction, i.e.,

$$p_y \le p_x, \quad \forall x < y, p_x \in D^+g(x) \text{ and } p_y \in D^+g(y). \tag{5.11}$$

Since D^+g is a.e. differentiable in $\mathbb{R}$, $D^+g(\cdot)$ is a.e. univalued in $\mathbb{R}$. Therefore, we can consider that $D^+g(\cdot)$ is a decreasing function. Moreover,

$$\|D^+g(\cdot)\|_{\mathbb{L}^\infty([-L,L],\mathbb{R})} \le M + KL.$$

Let $f, \tilde{f}$ be in $SC_{[K,L,M]}$. We have that

$$\begin{aligned}
\|f - \tilde{f}\|_{\mathbb{W}^{1,1}(\mathbb{R},\mathbb{R})} &= \|f - \tilde{f}\|_{\mathbb{L}^1(\mathbb{R},\mathbb{R})} + \|D^+f - D^+\tilde{f}\|_{\mathbb{L}^1(\mathbb{R},\mathbb{R})} \\
&= \|f - \tilde{f}\|_{\mathbb{L}^1([-L,L],\mathbb{R})} + \|D^+f - D^+\tilde{f}\|_{\mathbb{L}^1([-L,L],\mathbb{R})} \\
&\le (L+1) \cdot \|D^+f - D^+\tilde{f}\|_{\mathbb{L}^1([-L,L],\mathbb{R})} \\
&= (L+1) \cdot \|D^+g - D^+\tilde{g}\|_{\mathbb{L}^1([-L,L],\mathbb{R})}.
\end{aligned}$$

Therefore, one has that for a given $\varepsilon > 0$, it holds

$$\mathcal{H}_\varepsilon\Big(SC_{[K,L,M]} \,\Big|\, W^{1,1}(\mathbb{R}, \mathbb{R})\Big) \le \mathcal{H}_{\varepsilon_1}\big(\mathcal{DC}_{[L,M+KL]} \mid \mathbb{L}^1([-L,L],\mathbb{R})\big) \tag{5.12}$$

where $\varepsilon_1 = \frac{\varepsilon}{2(L+1)}$ and

$$\mathcal{DC}_{[L,M]} = \{F : [-L, L[\to [-M, M[\mid F \text{ is not increasing}\}. \tag{5.13}$$

Therefore, our goal is now to study an upper bound on $\mathcal{H}_\varepsilon\big(\mathcal{DC}_{[L,M]} \mid \mathbb{L}^1([-L,L],\mathbb{R})\big)$.

Proposition 5.2 *Given $L, M > 0$, for $\varepsilon > 0$ sufficiently small, it holds*

$$\mathcal{H}_\varepsilon\big(\mathcal{DC}_{[L,M]} \mid \mathbb{L}^1([-L,L],\mathbb{R})\big) \leq 2LM \cdot \frac{1}{\varepsilon} \tag{5.14}$$

Proof. Given any $N \in \mathbb{Z}^+$, we divide $[-L, L[$ and $[-M, M[$ into $2N$ intervals, i.e.,

$$[-L, L[= \bigcup_{i=0}^{2N-1} \Big[-L + \frac{i}{2N}, -L + \frac{i+1}{2N}\Big[$$

and

$$]-M, M] = \bigcup_{i=1}^{2N} \Big]-M + \frac{i}{2N}, -M + \frac{i+1}{2N}\Big].$$

Given any function F be in $\mathcal{DC}_{[L,M]}$, we approximate F by a piecewise constant function $\overline{F}$ as follows:

$$\overline{F}(x) = -M + \frac{k}{2N} \quad \forall x \in \Big[-L + \frac{i}{2N}, -L + \frac{i+1}{2N}\Big[$$

if $-M + \frac{k-1}{2N} < F\big(-L + \frac{i}{2N}\big) \leq -M + \frac{k}{2N}$. One can compute that

$$\|F - \overline{F}\|_{\mathbb{L}^1([-L,L],\mathbb{R})} \leq \frac{LM}{2} \cdot \frac{1}{N}. \tag{5.15}$$

On the other hand, $\overline{F} \in \mathcal{F}$ where

$$\mathcal{F}_N = \Big\{\overline{F} : [-L, L[\to \big\{-M + \frac{1}{2N}M, -M + \frac{2}{2N}M..., \frac{2N}{N}M\big\} \;\Big|\; \overline{F} \text{ is not increasing}\Big\}.$$

We have that

$$|\mathcal{F}| = \binom{2N}{N} < 2^{2N}.$$

Now, given $\epsilon > 0$ sufficiently small, we choose $\bar{N} = \left\lfloor \frac{LM}{2\varepsilon} \right\rfloor + 1$ to obtain that for every $F \in \mathcal{DC}_{[L,M]}$, there exists $\overline{F} \in \mathcal{F}_N$ such that $\|F - \overline{F}\|_{\mathbb{L}^1([-L,L],\mathbb{R})} \leq \varepsilon$. Therefore,

$$DC_{[L,M]} \subset \bigcup_{\overline{F} \in \mathcal{F}_{\bar{N}}} B_{\mathbb{L}^1}\big(\overline{F}, \varepsilon\big).$$

On the other hand,

$$|\mathcal{F}| \leq 2^{\frac{2LM}{\varepsilon}}$$

Hence,

$$\mathcal{N}_\varepsilon\big(DC_{[L,M]} \mid \mathbb{L}^1([-L,L],\mathbb{R})\big) \leq 2^{\frac{2LM}{\varepsilon}}.$$

The proof is complete. □

Therefore, we obtain that

Proposition 5.3 *Given $K, L, M > 0$, for $\varepsilon > 0$ sufficiently small, it holds*

$$\mathcal{H}_\varepsilon\Big(SC_{[K,L,M]} \;\Big|\; W^{1,1}(\mathbb{R},\mathbb{R})\Big) \leq \Gamma_{[K,L,M]} \cdot \frac{1}{\varepsilon}$$

where $\Gamma_{[K,L,M]} = 4L(L+1)(M+KL)$.

Therefore, by recalling corollary 5.1, we obtain our first result in one dimensional case.

Theorem 5.1 *Given any $L, M > 0$ and $T > 0$. For $\varepsilon > 0$ sufficiently small, one has*

$$\mathcal{H}_\varepsilon\Big(S_T(C_{[L,M]}) + \beta_T \mid W^{1,1}(\mathbb{R},\mathbb{R})\Big) \leq \Gamma_{[K_T,L_T,M]} \cdot \frac{1}{\varepsilon} \tag{5.16}$$

where $L_T = L + T \cdot \sup_{\|p\| \leq M} |DH(p)|$, $K_T = \frac{2}{\alpha T}$.

GENERAL CASE ($n \geq 1$): We first introduce some results of monotone functions which will play a role in this part. A standard reference is in [1].

Definition 5.2 *Let $F : \mathbb{R}^n \rightrightarrows \mathbb{R}^n$ be a multifunction. We say that F is monotone decreasing if*

$$\langle v_2 - v_1, x_2 - x_1 \rangle \leq 0, \quad \forall x_i \in \mathbb{R}^n, v_i \in F(x_i), i = 1, 2. \tag{5.17}$$

Moreover, the monotone decreasing function F is maximal if it is maximal with respect to the inclusion in the class of monotone decreasing functions, i.e.,

$$F_1(x) \subseteq F(x) \text{ for all } x \in \mathbb{R}^n \Longrightarrow F_1 = F.$$

Let $Dm(F) = \{x \in \mathbb{R}^n \mid F(x) \neq \varnothing\}$ be the domain of F. Given an open set $\Omega \subset \mathbb{R}^n$ relatively compact in the interior of $Dm(F)$. From [1] we have that if F is a monotone decreasing function then F is bounded and almost everywhere univalued in Ω. Therefore, we can consider F as an element of $\mathbb{L}^\infty(\Omega, \mathbb{R}^n)$. We recall now a result in [1]:

Proposition 5.4 *The monotone function F, viewed as an element in $\mathbb{L}^\infty(\Omega, \mathbb{R}^n)$, is in $BV(\Omega, \mathbb{R}^n)$. Moreover,*

$$\int_\Omega |DF| \leq 2^{\frac{n}{2}} \omega_n \cdot \big[\mathrm{diam}(\Omega) + \mathrm{osc}(F, \Omega)\big]^n \tag{5.18}$$

where ω_n is the Lebesgue measure on a unit ball in $\mathbb{R}^n$,

$$\mathrm{diam}(\Omega) = \sup\{|y - x| \mid x, y \in \Omega\} \quad \text{and} \quad \mathrm{osc}(F, \Omega) = \sup\big\{|F(y) - F(x)| \mid x, y \in \Omega\big\}.$$

For any $f \in SC_{[K,L,M]}$, $g(\cdot) = f(\cdot) - \frac{K}{2}|\cdot|^2$ is a concave function. Therefore,

$$D^+ g : \mathbb{R}^n \rightrightarrows \mathbb{R}$$

is monotone decreasing. Moreover,

$$\|D^+g(\cdot)\|_{\mathbb{L}^\infty([-L,L]^n,\mathbb{R}^n)} \leq M + \sqrt{n}KL =: M_1 \tag{5.19}$$

and

$$\|D(D^+g)\|_{BV([-L,L]^n)} \leq 2^{\frac{3n}{2}} n^{\frac{n}{2}} \omega_n \cdot [L(\sqrt{n}K+1)+M]^n =: C_1. \tag{5.20}$$

On the other hand, let $f, \tilde{f}$ be in $SC_{[K,L,M]}$. We have that

$$\begin{aligned} \|f-\tilde{f}\|_{\mathbb{W}^{1,1}(\mathbb{R}^n,\mathbb{R}^n)} &= \|f-\tilde{f}\|_{\mathbb{L}^1(\mathbb{R}^n,\mathbb{R}^n)} + \|D^+f - D^+\tilde{f}\|_{\mathbb{L}^1(\mathbb{R}^n,\mathbb{R}^n)} \\ &= \|f-\tilde{f}\|_{\mathbb{L}^1([-L,L]^n,\mathbb{R}^n)} + \|D^+f - D^+\tilde{f}\|_{\mathbb{L}^1([-L,L]^n,\mathbb{R}^n)} \\ &\leq (L+1)\cdot\|D^+f - D^+\tilde{f}\|_{\mathbb{L}^1([-L,L]^n,\mathbb{R}^n)} \\ &= (L+1)\cdot\|D^+g - D^+\tilde{g}\|_{\mathbb{L}^1([-L,L]^n,\mathbb{R}^n)}. \end{aligned}$$

Therefore, one has that for a given $\varepsilon > 0$, it holds

$$\mathcal{H}_\varepsilon\Big(SC_{[K,L,M]} \;\Big|\; W^{1,1}(\mathbb{R}^n,\mathbb{R})\Big) \leq \mathcal{H}_{\varepsilon_1}\big(\mathcal{DC}_{[L,M_1,C_1]} \mid \mathbb{L}^1([-L,L]^n,\mathbb{R}^n)\big) \tag{5.21}$$

where $\varepsilon_1 = \frac{\varepsilon}{2(L+1)}$ and

$$\mathcal{DC}_{[L,M_1,C_1]} = \{F : [-L,L[^n \to [-M_1,M_1[^n \mid F \text{ is not increasing and } \|DF\|_{BV} \leq C_1\}. \tag{5.22}$$

Kolmogorov-ε entropy of class of monotone function in $\mathbb{L}^1$: Let $\mathcal{L} = [0,L]^n$ and $\mathcal{M} = [0,M]^n$. We denote $\mathcal{F}_{[\mathcal{L},\mathcal{M},C]}$ by a set of multifunctions $F : \mathcal{L} \to \mathcal{M}$ such that $\|F\|_{BV(\mathcal{L})} \leq C$. Our main result is as the following:

Proposition 5.5 *For any $\varepsilon > 0$ sufficiently small, it holds*

$$\mathcal{H}_\epsilon\Big(\mathcal{F}_{[\mathcal{L},\mathcal{M},C]} \mid \mathbb{L}^1(\mathcal{L},\mathbb{R}^n)\Big) \leq \Gamma_{[\mathcal{L},\mathcal{M},C,n]} \cdot \frac{1}{\varepsilon^n},$$

where $\Gamma_{[\mathcal{L},\mathcal{M},C,n]} = 2^n \cdot (LC + \sqrt{n}M \cdot L^n)^n$.

Proof. For a fixed $N \in \mathbb{N}$, we divide $\mathcal{L}$ into N^n cubes which have a side $\frac{L}{N}$. More precisely,

$$\mathcal{L} = \bigcup_{\iota \in \{0,\dots,N-1\}^n} \Box_\iota$$

where $\Box_\iota = [\frac{\iota_1}{N}L, \frac{\iota_1+1}{N}L] \times \dots \times [\frac{\iota_n}{N}L, \frac{\iota_n+1}{N}L]$ for every $\iota = (\iota_1,\dots,\iota_n) \in \{0,\dots,N-1\}^n$.

We first approximate $F \in \mathcal{F}_{[\mathcal{L},\mathcal{M},C]}$ by $\overline{F}$ which is constant in $\Box_\iota$. More precisely, given any $F \in \mathcal{F}_{[\mathcal{L},\mathcal{M},C]}$, for every $\iota \in \{0,\dots,N-1\}^n$, let

$$\overline{F}_\iota = \frac{1}{\text{Vol}(\Box_\iota)} \cdot \int_{\Box_\iota} F(x)dx$$

be an average of F over $\Box_\iota$. One observes that $\overline{F}_\iota = (\overline{F}_\iota^1, ..., \overline{F}_\iota^n) \in \mathcal{M}$. By using Poincare inequality for a convex domain, we have

$$\|F - \overline{F}_\iota\|_{\mathbb{L}^1(\mathrm{int}(\Box_\iota))} \leq \frac{L}{2N} \cdot \|DF\|_{\mathrm{BV}(\mathrm{int}(\Box_\iota))}. \tag{5.23}$$

Let $\overline{F} : \mathcal{L} \to \mathcal{M}$ be such that

$$\overline{F}(x) = \begin{cases} \overline{F}_\iota & \text{if} \quad x \in \mathrm{int}(\Box_\iota), \\ \\ 0 & \text{if} \quad x \in \mathcal{L} \backslash \bigcup_{\iota \in \{0,...,N-1\}^n} \mathrm{int}(\Box_\iota). \end{cases} \tag{5.24}$$

Recalling (5.23), we estimate

$$\begin{aligned} \|F - \overline{F}\|_{\mathbb{L}^1(\mathcal{L})} &= \sum_{\iota \in \{0,...,N-1\}^n} \|F - \overline{F}\|_{\mathbb{L}^1(\mathrm{int}(\Box_\iota))} \\ &\leq \frac{L}{2N} \cdot \sum_{\iota \in \{0,...,N-1\}^n} \|DF\|_{\mathrm{BV}(\mathrm{int}(\Box_\iota))} \\ &\leq \frac{L}{2N} \cdot \|DF\|_{\mathrm{BV}(\mathcal{L})}. \end{aligned}$$

Since $\|DF\|_{\mathrm{BV}(\mathcal{L})} \leq C$, we then have

$$\|F - \overline{F}\|_{\mathbb{L}^1(\mathcal{L})} \leq \frac{L}{2N} \cdot C. \tag{5.25}$$

Moreover, fixing $i \in \{1, ..., n\}$, we denote by $e_i = (\delta_1^i, ..., \delta_n^i)$ where $\delta_j^i = 0$ if $j \neq i$ and $\delta_i^i = 1$. For any $\iota = (\iota_1, ..., \iota_n) \in \{0, ..., N-1\}^n$ be such that $\iota_i \in \{0, ..., N-2\}$, we set $\bar{\iota} = \iota + e_i$ where e_i is the $i - th$ coordinate vector of $\mathbb{R}^n$. We compute

$$\overline{F}_{\bar{\iota}} - \overline{F}_\iota = \frac{L^n}{N^n} \cdot \int_{\Box_\iota} F\Big(x + \frac{L}{N} e_i\Big) - F(x) dx.$$

Since F is monotone decreasing, we have

$$\begin{aligned} \langle \overline{F}_{\bar{\iota}} - \overline{F}_\iota, e_i \rangle &= \frac{L^{n-1}}{N^{n-1}} \cdot \int_{\Box_\iota} \Big\langle F\Big(x + \frac{L}{N} e_i\Big) - F(x), x + \frac{L}{N} e_i - x \Big\rangle dx \\ &\leq 0. \end{aligned}$$

This implies that $\overline{F}_{\bar{\iota}}^i \leq \overline{F}_\iota^i$. Therefore, by the same argument, one can show that

$$\overline{F}_{\iota + e_j}^j \leq \overline{F}_\iota^j \quad \forall j \in \{1, ..., n\}, \iota \in \{0, ..., N-1\}^n. \tag{5.26}$$

Now, dividing $[0, M]$ into N intervals

$$[0, M[= \Big[0, \frac{M}{N}\Big[\bigcup ... \bigcup \Big[\frac{N-1}{N} M, M\Big[,$$

we construct a new function $\tilde{F} : \mathcal{L} \to \mathcal{M}$ which corresponds to $\overline{F}$. For any $\iota \in \{0, ..., N-1\}^n$, for each $i \in \{1, ..., n\}$, we set $\tilde{F}^i_\iota = (k + \frac{1}{2}) \cdot \frac{M}{N}$ if $\overline{F}^i_\iota \in \left[k \cdot \frac{M}{N}, (k+1) \cdot \frac{M}{N}\right[$. We define $\tilde{F} : \mathcal{L} \to \mathcal{M}$ as

$$\tilde{F}(x) = \begin{cases} \tilde{F}_\iota = (\tilde{F}^1_\iota, ..., \tilde{F}^n_\iota) & \text{if} \quad x \in \mathrm{int}(\square_\iota), \\ \\ 0 \quad \text{if} \quad x \in \mathcal{L} \setminus \bigcup_{\iota \in \{0,...,N-1\}^n} \mathrm{int}(\square_\iota). \end{cases} \tag{5.27}$$

For every $\iota \in \{0, ..., N-1\}^n$, it holds

$$|\tilde{F}_\iota - \overline{F}_\iota| \leq \frac{\sqrt{n}}{2} \cdot \frac{M}{N}.$$

Thus,

$$\|\tilde{F} - \overline{F}\|_{\mathbb{L}^1(\mathcal{L})} \leq \frac{\sqrt{n} M \cdot L^n}{2} \cdot \frac{1}{N}.$$

Combining the above estimate with (5.25), we obtain that

$$\|\tilde{F} - F\|_{\mathbb{L}^1(\mathcal{L})} \leq \Gamma^1_{[\mathcal{L}, \mathcal{M}, C, n]} \cdot \frac{1}{N} \tag{5.28}$$

where $\Gamma^1_{[\mathcal{L}, \mathcal{M}, C, n]} = \frac{LC}{2} + \frac{\sqrt{n} M \cdot L^n}{2}$.

On the other hand, (5.26) implies that

$$\tilde{F}^j_{\iota + e_j} \leq \tilde{F}^j_\iota \qquad \forall j \in \{1, ..., n\}, \iota \in \{0, ..., N-1\}^n. \tag{5.29}$$

Let $\tilde{\mathcal{F}}$ be the set of all functions $\tilde{F} : \mathcal{L} \to \left\{0, (0 + \frac{1}{2}) \cdot \frac{M}{N}, ..., ((N-1) + \frac{1}{2}) \cdot \frac{M}{N}\right\}^n$ that are constant in each $\mathrm{int}(\square_\iota)$ for all $\iota \in \{0, ..., N-1\}^n$, equal to 0 in $\mathcal{L} \setminus \bigcup_{\iota \in \{0,...,N-1\}^n} \mathrm{int}(\square_\iota)$ and satisfy (5.29). We are now giving an upper bound of $|\tilde{\mathcal{F}}|$. Fixing $i \in \{1, ..., n\}$, we define

$$\mathcal{J}_i = \{\iota \in \{0, ..., N-1\}^n \mid \iota_i = 0\}.$$

We have that $|\mathcal{J}_i| = N^{n-1}$. For any $\iota \in \mathcal{J}_i$, we denote by $\mathcal{K}^i_\iota$ the number of sets $\{\tilde{F}^i_\iota, \tilde{F}^i_{\iota+e_i}, ..., \tilde{F}^i_{\iota+(N-1)\cdot e_i}\}$ such that $\tilde{F}^i_{\iota+k\cdot e_i} \in \left\{(0 + \frac{1}{2}) \cdot \frac{M}{N}, ..., ((N-1) + \frac{1}{2}) \cdot \frac{M}{N}\right\}$ and $\tilde{F}^i_{\iota+(k+1)\cdot e_i} \leq \tilde{F}^i_{\iota+k\cdot e_i}$. One has that $\mathcal{K}^i_\iota \leq 2^{2N}$ (see in [27]). Hence by recalling that $|\mathcal{J}_i| = N^{n-1}$, if we define $\mathcal{K}^i$ as the number of sets

$$\left\{\tilde{F}^i_\iota \in \left\{(0 + \frac{1}{2}) \cdot \frac{M}{N}, ..., ((N-1) + \frac{1}{2}) \cdot \frac{M}{N}\right\} \;\middle|\; \iota \in \{0, ..., N-1\}^n\right\}$$

such that $\tilde{F}^i_{\iota+e_i} \leq \tilde{F}^i_\iota$ for all $\iota \in \{0, ..., N-1\}^n$, then we have that $\mathcal{K}^i \leq 2^{2N^n}$. Therefore, by letting i be from 1 to N, we obtain that

$$|\tilde{\mathcal{F}}| \leq 2^{2nN^n} \leq 2^{(2N)^n}. \tag{5.30}$$

Now, given any $0 < \varepsilon < \Gamma^1_{[\mathcal{L},\mathcal{M},C,n]}$, we choose $N = \left\lfloor \frac{\Gamma^1_{[\mathcal{L},\mathcal{M},C,n]}}{\varepsilon} \right\rfloor + 1$. From (5.28) and (5.30), we finally obtain that

$$\mathcal{H}_\epsilon\Big(\mathcal{F}_{[\mathcal{L},\mathcal{M},C]} \mid \mathbb{L}^1(\mathcal{L})\Big) \le \Gamma_{[\mathcal{L},\mathcal{M},C,n]} \cdot \frac{1}{\varepsilon^n}$$

where $\Gamma_{[\mathcal{L},\mathcal{M},C,n]} = \Big(2 \cdot \Gamma^1_{[\mathcal{L},\mathcal{M},C,n]}\Big)^n$. The proof is complete. □

Remark 5.1 *Given any $L_1 < L_2$, $M_1 < M_2$ and $C > 0$, we denote by $\mathcal{L}_1^2 = [L_1, L_2]$ and $\mathcal{M}_1^2 =]M_1, M_2[$. Let $\mathcal{F}_{[\mathcal{L}_1^2,\mathcal{M}_1^2,C]}$ be a set of functions $F : \mathcal{L}_1^2 \to \mathcal{M}_1^2$ which are monotone decreasing in $\mathcal{L}_1^2$ and $\|F\|_{\mathrm{BV}(\mathcal{L}_1^2)} \le C$. Then, for every $\varepsilon > 0$, one has that*

$$\mathcal{H}_\epsilon\Big(\mathcal{F}_{[\mathcal{L}_1^2,\mathcal{M}_1^2,C]} \mid \mathbb{L}^1\big(\mathcal{L}_1^2\big)\Big) \le \Gamma_{[\mathcal{L}_1^2,\mathcal{M}_1^2,C,n]} \cdot \frac{1}{\varepsilon^n}$$

where $\Gamma_{[\mathcal{L}_1^2,\mathcal{M}_1^2,C,n]} = 2^n \cdot (C \cdot (L_2 - L_1) + \sqrt{n} \cdot (M_2 - M_1) \cdot (L_2 - L_1)^n)^n$.

Therefore, by recalling (5.21), we obtain that

Theorem 5.2 *Given any $L, M, K > 0$, for $\varepsilon > 0$ sufficiently small, it holds*

$$\mathcal{H}_\varepsilon\Big(\mathcal{SC}_{[K,L,M]} \mid \mathbb{W}^{1,1}\big(\mathbb{R}^n, \mathbb{R}\big)\Big) \le \Gamma^{\mathcal{SC}}_{[K,L,M,n]} \cdot \frac{1}{\varepsilon^n} \tag{5.31}$$

where

$$\Gamma^{\mathcal{SC}}_{[K,L,M,n]} = 2^{n^2+3n}(L+1)^n\Big(2^{n/2}\omega_n n^{n/2} L(L + M + K\sqrt{n}L)^n + \sqrt{n}L^n(M + K\sqrt{n}L)\Big)^n. \tag{5.32}$$

Finally, by recalling Corollary 5.1, we have that

Theorem 5.3 *Let H satisfy **(H)**. For every $\varepsilon > 0$ and for every $T > 0$, we have that*

$$\mathcal{H}_\epsilon(S_T(C_{[L,M]}) + \beta_T \mid W^{1,1}(\mathbb{R}^n)) \le \Gamma^{\mathcal{SC}}_{[K_T,L_T,M,n]} \cdot \frac{1}{\varepsilon^n} \tag{5.33}$$

where $L_T = L + T \cdot \sup_{\|p\| \le M} DH(p)$, $K_T = \frac{2}{\alpha T}$, $\beta_T = T \cdot H(0)$ and $\Gamma^{\mathcal{SC}}_{[K_T,L_T,M,n]}$ is computed by (5.32).

5.2 An lower bound of $\mathcal{H}_\epsilon\Big(S_T(C_{[L,M]}) \mid W^{1,1}_{loc}(\mathbb{R}^n,\mathbb{R})\Big)$

We start this subsection with a controllability result.

Theorem 5.4 *Given $K, L, M > 0$ and time $T > 0$ such that*

$$KT \leq \frac{1}{2\alpha_M} \text{ where } \alpha_M = \|DH^2(p)\|_{\mathbb{L}^\infty([-M,M]^n)}. \tag{5.34}$$

Then, it holds

$$\mathcal{SC}_{[K,L,M]} \subset \mathcal{S}_T(C_{[L_T,M]}) + \beta_T \tag{5.35}$$

provided that $L_T = L + T \cdot \sup_{\|p\| \leq M} DH(p)$ and $\beta_T = T \cdot H(0)$.

Proof. Let u_T be any function in $\mathcal{SC}_{[K,L,M]} - \beta_T$. We need to find a function $u_0 \in C_{[L,M]}$ such that $\mathcal{S}_T(u_0) = u_T$. Set

$$w_0(x) = -u_T(-x), \quad \forall x \in \mathbb{R}^n.$$

Since $u_T \in \mathcal{SC}_{[K,L,M]} - \beta_T$, we have that

(i) $w_0(x) = \beta_T$ for all $x \in \mathbb{R}^n \backslash [-L, L]^n$ and $\|Dw_0\|_{\mathbb{L}^\infty(\mathbb{R}^n)} < M$,

(ii) w_0 is semiconvex with semiconvexity constant K.

We define $w : [0,T] \times \mathbb{R}^n \to \mathbb{R}$ be as the following:

$$w(t,x) = S_t(w_0)(x), \quad \forall (t,x) \in [0,T] \times \mathbb{R}^n. \tag{5.36}$$

From lemma 5.2, one can get that $w(T,\cdot)$ is lipschitz and $\|\nabla w(T,\cdot)\|_{\mathbb{L}^\infty(\mathbb{R}^n)} < M$. Moreover, since $\beta_T = T \cdot H(0)$

$$w(T,x) = 0, \quad \forall x \in \mathbb{R}^n \backslash [-L_T, L_T]^n. \tag{5.37}$$

Now, we set

$$u(t,x) = -w(T-t,-x), \quad \forall (t,x) \in [0,T] \times \mathbb{R}^n.$$

One first has that

$$u(0,\cdot) \in C_{[L_T,M]}.$$

On the other hand, if u is differentiable at $(t_0,x_0) \in [0,T] \times \mathbb{R}^n$ then we have

$$u_t(t_0,x_0) = w_T(T-t_0,-x_0) \text{ and } \nabla u(t_0,x_0) = \nabla w(T-t_0,-x_0).$$

Hence, since w is the viscosity solution of (2.22), we have

$$u_t(t_0,x_0) + H(\nabla u(t_0,x_0)) = 0.$$

It means that u satisfies (2.22) at (t_0,x_0). It is well-known from the Hopf's formula that w is lipschitz in $[0,T] \times \mathbb{R}^n$. It implies that u is lipschitz in $[0,T] \times \mathbb{R}^n$. Therefore, u solves (2.22) almost everywhere in $[0,T] \times \mathbb{R}^n$. Our goal is now to show that u is a viscosity solution of (2.22) in $[0,T]\times\mathbb{R}^n$. Indeed, by the uniqueness of viscosity solutions of (2.22), from Corollary 1.5.7 in [14], we need to check that for every $t\in(0,T]$, it holds

$$u(t,x+h)+u(t,x-h)-2u(t,x)\leq C\cdot\left(1+\frac{1}{t}\right)\cdot|h|^2,\quad\forall x,h\in\mathbb{R}^n. \tag{5.38}$$

Recalling that $u(t,x)=-w(T-t,-x)$, (5.38) holds if we show there exist a constant $C_1>0$ such that

$$w(t,x+h)+w(t,x-h)-2w(x)\geq -C_1\cdot|h|^2,\quad\forall t\in[0,T],x,h\in\mathbb{R}^n. \tag{5.39}$$

Fixing $t\in(0,T]$, for every $x,h\in\mathbb{R}^n$, let y_h^+ be a minimum of $\min_{y\in\mathbb{R}^n}\left\{t\cdot L\Big(\frac{x+h-y}{t}\Big)+w_0(y)\right\}$. This implies that

$$DL\Big(\frac{x+h-y_h^+}{t}\Big)=\nu_{y_h^+}\in D^-w_0(y_h^+) \tag{5.40}$$

By Hopf's formula, it holds

$$w(t,x+h)=t\cdot L\Big(\frac{x+h-y_h^+}{t}\Big)+w_0(y_h^+).$$

Similarly, let y_h^- be a minimum of $\min_{y\in\mathbb{R}^N}\left\{t\cdot L\Big(\frac{x-h-y}{t}\Big)+w_0(y)\right\}$, we also have that

$$DL\Big(\frac{x-h-y_h^-}{t}\Big)=\nu_{y_h^-}\in D^-w_0(y_h^-). \tag{5.41}$$

and

$$w(t,x-h)=t\cdot L\Big(\frac{x-h-y_h^-}{t}\Big)+w_0(y_h^-).$$

On the other hand, by Hopf's formula

$$w(t,x)\leq t\cdot L\Big(\frac{x-\frac{y_h^++y_h^-}{2}}{t}\Big)+w_0\Big(\frac{y_h^++y_h^-}{2}\Big).$$

Hence, we get that

$$\begin{aligned}w(t,x+h)+w(t,x-h)-2w(t,x)\geq{}& w_0(y_h^+)+w_0(y_h^-)-2w_0\Big(\frac{y_h^++y_h^-}{2}\Big)\\&+t\cdot\Big[L\Big(\frac{x+h-y_h^+}{t}\Big)+L\Big(\frac{x-h-y_h^-}{t}\Big)-2\cdot L\Big(\frac{x-\frac{y_h^++y_h^-}{2}}{t}\Big)\Big].\end{aligned}$$

Since L is convex and w_0 is semiconvex with semiconvexity constant K, we obtain from the above inequality that

$$w(t,x+h)+w(t,x-h)-2w(t,x)\geq-\frac{K}{4}\cdot|y_h^+-y_h^-|^2. \tag{5.42}$$

From (5.40) and (5.41), one gets that

$$\Big\langle DL\Big(\frac{x+h-y_h^+}{t}\Big) - DL\Big(\frac{x-h-y_h^-}{t}\Big), -\frac{y_h^+-y_h^-}{t}\Big\rangle = \big\langle \nu_{y_h^+} - \nu_{y_h^-}, -\frac{y_h^+-y_h^-}{t}\big\rangle. \quad (5.43)$$

By the semiconvex property of w_0, we can estimate the right hand side of (5.43) as

$$\big\langle \nu_{y_h^+} - \nu_{y_h^-}, -\frac{y_h^+-y_h^-}{t}\big\rangle \leq \frac{K}{t}\cdot |y_h^+ - y_h^-|^2. \quad (5.44)$$

On the other hand,

$$\Big\langle DL\Big(\frac{x+h-y_h^+}{t}\Big) - DL\Big(\frac{x-h-y_h^-}{t}\Big), -\frac{y_h^+-y_h^-}{t}\Big\rangle = \int_0^1 \Big\langle D^2L(z_\tau)\Big(\frac{2h-(y_h^+-y_h^-)}{t}\Big), -\frac{y_h^+-y_h^-}{t}\Big\rangle d\tau \quad (5.45)$$

where $z_\tau = \tau\cdot\frac{x+h-y_h^+}{t} + (1-\tau)\cdot\frac{x-h-y_h^-}{t}$. From (5.40), (5.41) and $\|w_0\|_{\mathbb{L}^\infty(\mathbb{R}^n)} \leq M$, we have that $|z_\tau| \leq M_1 = \|DL^{-1}\|_{\mathbb{L}^\infty([-M,M]^n)}$ for all $\tau\in[0,1]$. Hence, $\|D^2L(z_\tau)\| \leq K_M$ where $K_M = \|D^2L\|_{\mathbb{L}^\infty([-M_1,M_1]^n)}$, Thus,

$$\int_0^1 \Big\langle D^2L(z_\tau)\Big(\frac{2h}{t}\Big), -\frac{y_h^+-y_h^-}{t}\Big\rangle d\tau \geq -2K_M\cdot\frac{|h|\cdot|y_h^+-y_h^-|}{t^2}.$$

On the other hand, since $\alpha_M = \|DH^2(p)\|_{\mathbb{L}^\infty([-M,M]^n)}$, we have that $DH^2(p)\leq \alpha_M\cdot I$ for all $p\in[-M,M]^n$. Hence, for every $q = DL^{-1}(p)$ where $p\in[-M,M]^n$, one has that

$$DL^2(q) = [D^2H(p)]^{-1} \geq \frac{1}{\alpha_M}\cdot I.$$

Thus,

$$\int_0^1 \Big\langle D^2L(z_\tau)\Big(-\frac{y_h^+-y_h^-}{t}\Big), -\frac{y_h^+-y_h^-}{t}\Big\rangle d\tau \geq \frac{1}{\alpha_M}\cdot\frac{|y_h^+-y_h^-|^2}{t^2}.$$

Therefore, one gets that

$$\Big\langle DL\Big(\frac{x+h-y_h^+}{t}\Big) - DL\Big(\frac{x-h-y_h^-}{t}\Big), -\frac{y_h^+-y_h^-}{t}\Big\rangle \geq \frac{1}{\alpha_M}\cdot\frac{|y_h^+-y_h^-|^2}{t^2} - 2K_M\cdot\frac{|h|\cdot|y_h^+-y_h^-|}{t^2}. \quad (5.46)$$

Combining (5.43), (5.44) and (5.46), we obtain that

$$2K_M\cdot|h|\cdot|y_h^+-y_h^-| \geq \Big(\frac{1}{\alpha_M} - Kt\Big)\cdot|y_h^+-y_h^-|^2.$$

Since $t\in(0,T]$, by recalling (5.34), we get from the above inequality that

$$|y_h^+-y_h^-| \leq 4K_M\alpha_M\cdot|h|.$$

Hence, from (5.42), we obtain that for every $t \in (0,T]$, it holds

$$w(t,x+h)+w(t,x-h)-2w(t,x) \geq -C_1 \cdot |h|^2. \tag{5.47}$$

where $C_1 = 4KK_M^2\alpha_M^2$. Therefore, u is a viscosity of (2.22) in $[0,T]\times\mathbb{R}^n$. By the uniqueness of viscosity solutions of (2.22) if we set $u_0(\cdot) = u(0,\cdot) \in C_{[L_T,M]}$, we have that $S_t(u_0)(\cdot) = u(t,\cdot)$. Moreover, one can see that $S_T(u_0) = u_T$. The proof is complete. □

Remark 5.2 *Given $K, M > 0$ and time $T > 0$ such that*

$$KT \leq \frac{1}{2\alpha_M} \text{ where } \alpha_M = \|DH^2(p)\|_{\mathbb{L}^\infty([-M,M]^n)}. \tag{5.48}$$

Let $w_0 : \mathbb{R}^n \to \mathbb{R}$ be semiconvex with the semiconvex constant K and $\|Dw_0\|_{\mathbb{L}^\infty(\mathbb{R}^n)} \leq M$. Then,

$$\begin{cases} w_t(t,x) + H(\nabla_x w(t,x)) = 0, \\ \\ w(0,\cdot) = w_0(\cdot) \end{cases} \tag{5.49}$$

admits the uniquess viscosity solution which is in $C^1(\mathbb{R}^n \times (0,T))$.

Proof. Let w be the uniqueness viscosity solution of (5.49). From the above proof, one has that $w(\cdot,\cdot)$ is locally semiconcave in $(0,+\infty)\times\mathbb{R}^n$ and $w(t,\cdot)$ is smooth for every $t \in (0,T)$. Given any $(t,x) \in \mathbb{R}^n \times (0,T)$, let $(\lambda, v) \in \mathbb{R}\times\mathbb{R}^n$ be in $D^*w(t,x)$. There exists a sequence $\{(t_k,x_k)\}_k$ converging to (t,x) such that w is differentiable at (t_k,x_k) and

$$\lim_{k\to\infty} (w_t(t_k,x_k), \nabla_x w(t_k,x_k)) = (\lambda, v). \tag{5.50}$$

Since $w(t,\cdot)$ is differentiable at x, one first has that $v = \nabla_x w(t,x)$. Thus,

$$\lim_{k\to\infty} \nabla_x w(t_k,x_k) = \nabla_x w(t,x).$$

On the other hand, we have that

$$w_t(t_k,x_k) = -H(\nabla_x w(t_k,x_k)).$$

Hence,

$$\lambda = \lim_{k\to\infty} w_t(t_k = -\lim_{n\to\infty} H(\nabla_x w(t_k,x_k)) = -H(\nabla_x w(t,x)).$$

Therefore, $D^*w(t,x) = \{(-H(\nabla_x w(t,x)), \nabla_x w(t,x))\}$. Hence, from proposition 4.5 we have that $D^+w(t,x)$ is singleton for all $(t,x) \in \mathbb{R}^n \times (0,T)$. The proof is complete by recalling proposition 4.5. □

Our main goal is now to give a lower bound on $\mathcal{H}_\varepsilon\Big(\mathcal{SC}_{[K,L,M]} \mid \mathbb{W}^{1,1}(\mathbb{R}^n,\mathbb{R})\Big)$. We will first introduce a bump function which is a sample to construct a subset of $\mathcal{SC}_{[K,L,M]}$ such that we can give an estimate on the Kolmogorov-ε of such subset in $\mathbb{W}^{1,1}(\mathbb{R}^n,\mathbb{R})$.

A bump function: consider the function $c : [0,1] \to \mathbb{R}$ such that

$$\begin{cases} c(r) = (\frac{1}{4})^2 - \int_0^r (\frac{1}{4} - |\frac{1}{4} - s|)ds \quad \forall r \in [0, \frac{1}{2}], \\ \\ c(r) = 0 \quad \forall r \in [\frac{1}{2}, 1]. \end{cases} \tag{5.51}$$

One see that

$$c'(r) = \left|\frac{1}{4} - r\right| - \frac{1}{4} \quad \forall r \in \left[0, \frac{1}{2}\right] \quad \text{and} \quad c'(r) = 0 \quad \forall r \in \left[\frac{1}{2}, 1\right].$$

Thus, $c'(\cdot)$ is Lipschitz with Lipschitz constant 1 in $[0,1]$ and $\|c'\|_{\mathbb{L}^\infty([0,1])} \leq \frac{1}{4}$.
Given $L, K > 0$, by setting $\bar{c}(r) = KL^2 \cdot c(\frac{r}{L})$ for every $r \in [0, L]$, we have that

(i) $\bar{c}(r) = \bar{c}'(r) = 0 \quad \forall r \in [\frac{L}{2}, L]$ and $0 \leq \bar{c}(r) \leq \frac{KL^2}{16} \quad \forall r \in [0, \frac{L}{2}]$,

(ii) $\|\bar{c}'\|_{\mathbb{L}^\infty([0,L])} \leq \frac{KL}{4}$ and $\bar{c}'(\cdot)$ is Lipschitz with Lipschitz constant K.

We construct now the bump function $b : [-L, L]^n \to \mathbb{R}$ as the followings:

$$b(x) \ := \ \frac{1}{3} \cdot \bar{c}(|x|), \quad \forall x \in B\Big(0, \frac{L}{2}\Big) \quad \text{and} \quad b(x) = 0 \quad \forall x \in [-L, L]^n \backslash B\Big(0, \frac{L}{2}\Big). \tag{5.52}$$

One can check that

$$\nabla b(x) = \frac{1}{3} \cdot \bar{c}'(|x|) \cdot \frac{x}{|x|}, \ \forall x \in [-L, L]^n \backslash 0 \quad \text{and} \quad \nabla b(0) = 0. \tag{5.53}$$

Thus,

$$\|\nabla b\|_{\mathbb{L}^\infty([-L,L]^n)} \leq \frac{KL}{12} \tag{5.54}$$

Moreover, one can also compute that

$$\begin{aligned} \|\nabla b\|_{\mathbb{L}^1([-L,L]^n)} &= \frac{1}{3} \cdot \int_{B(0,L)} |\bar{c}'(|x|)| dx = \frac{KL}{3} \cdot \int_{B(0,L)} \left|c'\Big(\frac{|x|}{L}\Big)\right| dx \\ &= \frac{KL^{n+1}}{3} \cdot \int_{B(0,1)} |c'(|x|)| dx = \frac{2KL^{n+1}\omega_{n-1}}{3} \int_0^1 |c'(r)| r^{n-1} dr \\ &= \frac{2KL^{n+1}\omega_{n-1}}{3} \cdot \frac{2^n - 1}{2n(n+1)4^n}. \end{aligned}$$

Thus, by setting

$$C_{[K,L,n]} = \frac{2KL^{n+1}\omega_{n-1}}{3} \cdot \frac{2^n - 1}{2n(n+1)4^n}, \tag{5.55}$$

we have that

$$\|\nabla b\|_{\mathbb{L}^1([-L,L]^n)} = C_{[K,L,n]}. \tag{5.56}$$

Since $\vec{c}'(\cdot)$ is lipschitz with lipschitz constant K, we have that for $x \in [-L,L]^n \backslash 0$, it holds

$$|\nabla b(x) - \nabla b(0)| = |\vec{c}'(|x|) - \vec{c}'(0)| \leq K \cdot |x|.$$

For $x, y \in [-L,L]^n \backslash 0$, we compute that

$$\begin{aligned} |\nabla b(y) - \nabla b(x)| &= \frac{1}{3} \cdot \left| \vec{c}'(|y|) \cdot \frac{y}{|y|} - \vec{c}'(|x|) \cdot \frac{x}{|x|} \right| \\ &\leq \frac{1}{3} \cdot |\vec{c}'(|y|) - \vec{c}'(|x|)| + \frac{1}{3} \cdot \vec{c}'(|x|) \cdot \left| \frac{y}{|y|} - \frac{x}{|x|} \right| \\ &\leq \frac{K}{3} \cdot \big| |y| - |x| \big| + \frac{K}{3} \cdot |x| \cdot \left(\left| \frac{y}{|y|} - \frac{y}{|x|} \right| + \left| \frac{y-x}{|x|} \right| \right) \\ &\leq K \cdot |y - x|. \end{aligned}$$

It implies that ∇b is lipschitz with lipschitz constant K in $[-L,L]^n$.

Given $N \in \mathbb{Z}^+$, we define $b_N : [-\frac{L}{N}, \frac{L}{N}]^n \to \mathbb{R}$ as

$$b_N(x) = \frac{1}{N^2} \cdot b(Nx), \quad \forall x \in \Big[-\frac{L}{N}, \frac{L}{N} \Big]^n.$$

One computes that $\nabla b_N(x) = \frac{1}{N} \cdot \nabla b(Nx)$ for all $x \in [-\frac{L}{N}, \frac{L}{N}]^n$. Thus, $b_N(\cdot)$ satisfies the following properties:

(i) $b_N(x) = 0$ for all $x \in [-\frac{L}{N}, \frac{L}{N}]^n \backslash B(x, \frac{L}{2N})$ and $\|\nabla b_N\|_{\mathbb{L}^\infty([-\frac{L}{N}, \frac{L}{N}]^n)} \leq \frac{KL}{12N}$,

(ii) $\|\nabla b_n\|_{\mathbb{L}^1([-\frac{L}{N}, \frac{L}{N}]^n)} = \frac{C_{[K,L,n]}}{N^{n+1}}$ where $C_{[K,L,n]} = \frac{2KL^{n+1}\omega_{n-1}}{3} \cdot \frac{2^n - 1}{2n(n+1)4^n}$,

(iii) b_N and $-b_N$ are semiconcave with semiconcavity constant K.

Our main result is as follow:

Theorem 5.5 *Given $K, L, M > 0$. For $\varepsilon > 0$ sufficiently small, it holds for every fixed $\beta \in \mathbb{R}^n$*

$$\mathcal{H}_\varepsilon\Big(\mathcal{SC}_{[K,L,M]} \mid \mathbb{W}^{1,1}(\mathbb{R}^n, \mathbb{R})\Big) \geq \Gamma^-_{[K,L,n]} \cdot \frac{1}{\varepsilon^n} \tag{5.57}$$

where $\Gamma^-_{[K,L,n]} = \frac{C^n_{[K,L,n]}}{2^{n+3}}$ and $C_{[K,L,n]} = \frac{2KL^{n+1}\omega_{n-1}}{3} \cdot \frac{2^n-1}{2n(n+1)4^n}$.

Proof. Given $N \in \mathbb{Z}^+$, we divide $[-L,L]^n$ into N^n cubes which have the side $\frac{2L}{N}$. More precisely,

$$[-L,L]^n = \bigcup_{\iota \in \{1,\dots,N\}^n} \square_\iota \tag{5.58}$$

where $\square_\iota = (-L, ..., -L) + \frac{L}{N} \cdot \iota + [-\frac{L}{N}, \frac{L}{N}]^n$, $\iota \in \{1, ..., N\}^n$. For each $\iota \in \{1, .., N\}^n$, we denote by

$$O_\iota = (-L, ..., -L) + \left[\iota - \Big(\frac{1}{2}, ..., \frac{1}{2} \Big) \right] \cdot \frac{L}{N}$$

the center of $\square_\iota$. The bump function which we are going to define on $\square_\iota$ is

$$b^\iota_N(x) = b_N(x - C_\iota), \quad \forall x \in \square_\iota \quad \text{and} \quad b^\iota_N(x) = 0, \quad \forall x \in \mathbb{R}^n \backslash \square_\iota.$$

One sees that,

(i) $b_N^\iota(x) = 0$ for all $x \in \mathbb{R}^n \backslash B(C_\iota, \frac{L}{2N})$,

(ii) $\|\nabla b_N^\iota\|_{\mathbb{L}^\infty(\mathbb{R}^n)} \leq \frac{KL}{12N}$ and $\|\nabla b_N^\iota\|_{\mathbb{L}^1(\mathbb{R}^n)} = \|\nabla b_N\|_{\mathbb{L}^1([-\frac{L}{N},\frac{L}{N}]^n)}$,

(iii) b_N^ι and $-b_N^\iota$ are semiconcave with semiconcave constant K.

Let

$$\Delta_N := \left\{\delta = (\delta_\iota)_{\iota \in \{1,\dots,N\}^n} \;\middle|\; \delta_\iota \in \{-1,1\}\right\}$$

we construct our class of semiconcave functions with semiconcavity constant K which have a support contained in $[-L,L]^n$ and whose gradients are bounded by M as follow:

$$\mathcal{U}_N = \left\{u_\delta = \sum_{\iota \in \{1,\dots,N\}^n} \delta_\iota \cdot b_N^\iota \;\middle|\; \delta \in \Delta_N\right\}. \tag{5.59}$$

For every $u \in \mathcal{U}_N$, we have that for $N \geq \frac{KL}{12M}$

$$\mathrm{Supp}(u) \subset [-L,L]^n, \quad \|\nabla u\|_{\mathbb{L}^\infty(\mathbb{R}^n)} \leq M$$

and u is semiconcave with semiconcavity constant K. Therefore,

$$\mathcal{U}_N \subset \mathcal{SC}_{[K,L,M]}. \tag{5.60}$$

Moreover, $|\mathcal{U}_N| = 2^{N^n}$. Fixing $\bar{\delta} \in \Delta_N$, for every $\delta \in \Delta_N$, we have that

$$\|u_{\bar{\delta}} - u_\delta\|_{\mathbb{W}^{1,1}(\mathbb{R}^n)} = 2d(\bar{\delta},\delta) \cdot \|b_N\|_{\mathbb{W}^{1,1}([-\frac{L}{N},\frac{L}{N}]^n)}. \tag{5.61}$$

where $d(\bar{\iota},\iota) = \mathrm{Card}\{\iota \in \{1,...,N\}^n \mid \bar{\delta}_\iota \neq \delta_\iota\}$. By Poincare inequality, we obtain that

$$\|b_N\|_{\mathbb{L}^1([-\frac{L}{N},\frac{L}{N}]^n)} \leq \frac{L}{N} \cdot \|\nabla b_N\|_{\mathbb{L}^1([-\frac{L}{N},\frac{L}{N}]^n)}.$$

Thus, for $N \geq L$, it holds

$$\|b_N\|_{\mathbb{W}^{1,1}([-\frac{L}{N},\frac{L}{N}]^n)} \leq 2 \cdot \|\nabla b_N\|_{\mathbb{L}^1([-\frac{L}{N},\frac{L}{N}]^n)} = \frac{2C_{[K,L,n]}}{N^{n+1}}.$$

Therefore, from (5.61), we have that

$$\|u_{\bar{\delta}} - u_\delta\|_{\mathbb{W}^{1,1}(\mathbb{R}^n)} \leq \varepsilon \quad \text{if} \quad d(\bar{\delta},\delta) \leq \frac{N^{n+1}\varepsilon}{4C_{[K,L,n]}}. \tag{5.62}$$

We define now

$$\mathcal{I}_{\bar{\delta},N}(\epsilon) = \left\{\delta \in \Delta_N \;\middle|\; \|u_{\bar{\delta}} - u_\delta\|_{\mathbb{W}^{1,1}(\mathbb{R}^n)} \leq \varepsilon\right\}.$$

Noting that $C_N(\varepsilon) = |\mathcal{I}_{\bar{\delta},N}(\epsilon)|$ is independent the choice of $\bar{\delta}$. From (5.62), one has that

$$C_N(\varepsilon) \leq \sum_{l=0}^{\left\lfloor \frac{N^{n+1}\varepsilon}{4C_{[K,L,n]}} \right\rfloor} \binom{N^n}{l} \tag{5.63}$$

where $\lfloor \alpha \rfloor = \max\{z \in \mathbb{Z} \mid z \leq \alpha\}$. Let $X_1, ..., X_{N^n}$ be independent random variables with Bernoulli distribution $\mathbb{P}(X_i = 1) = \mathbb{P}(X_i = 0) = \frac{1}{2}$, then for any $k \leq N^n$, one has

$$\sum_{l=0}^{k} \binom{N^n}{l} = 2^{N^n} \cdot \mathbb{P}(X_1 + ... + X_{N^n} \leq k). \tag{5.64}$$

Setting $S_{N^n} = X_1 + ... + X_{N^n}$, and using Hoeffding's inequality ([22, Theorem 2]) that, for any fixed $\mu > 0$, gives

$$\mathbb{P}(S_{N^n} - \mathbb{E}(S_{N^n}) \leq -\mu) \leq \exp\left(-\frac{2\mu^2}{N^n}\right), \tag{5.65}$$

where $\mathbb{E}(S_{N^n})$ denotes the expectation of S_{N^n}. Since, by the above assumptions on $X_1, \ldots, X_{N^n}$, we have $\mathbb{E}(S_{N^n}) = \frac{N^n}{2}$, taking $\mu = \frac{N^n}{2} - \left\lfloor \frac{N^{n+1}\varepsilon}{4C_{[K,L,n]}} \right\rfloor$ and assuming

$$\frac{N^{n+1}\varepsilon}{4C_{[K,L,n]}} \leq \frac{N^n}{2}$$

we deduce from (5.63), (5.64) and (5.65) that

$$\begin{aligned} C_N(\varepsilon) &\leq 2^{N^n} \cdot \exp\Bigg(- \frac{2\left(\frac{N^n}{2} - \left\lfloor \frac{N^{n+1}\varepsilon}{4C_{[K,L,n]}} \right\rfloor\right)^2}{N^n} \Bigg) \\ &\leq 2^{N^n} \cdot \exp\Bigg(- \frac{N^n}{2} \cdot \Big(1 - \frac{N\varepsilon}{2C_{[K,L,n]}}\Big)^2 \Bigg). \end{aligned}$$

By choosing

$$\bar{N} = \left\lfloor \frac{C_{[K,L,n]}}{\varepsilon} \right\rfloor + 1. \tag{5.66}$$

We obtain that

$$C_{\bar{N}}(\varepsilon) \leq 2^{\bar{N}^n} \cdot \exp\Big(- \frac{\bar{N}^n}{8} \Big).$$

Thus, by recalling that $|\mathcal{U}_{\bar{N}}| = 2^{\bar{N}^n}$, we have that the number of sets in an $\frac{\varepsilon}{2}$-cover of $\mathcal{U}_{\bar{N}}$ is at least

$$\mathcal{N}_{\frac{\varepsilon}{2}}(\mathcal{U}_{\bar{N}} \mid W^{1,1}(\mathbb{R}^n)) \geq \exp\Big(\frac{\bar{N}^n}{8}\Big) \geq \exp\Big(\frac{C^n_{[K,L,n]}}{2^3 \cdot \varepsilon^n}\Big). \tag{5.67}$$

Since $\mathcal{U}_{\bar{N}} \subset \mathcal{SC}_{[K,L,M]}$, we get that

$$\mathcal{N}_{\frac{\varepsilon}{2}}(\mathcal{SC}_{[K,L,M]} \mid W^{1,1}(\mathbb{R}^n)) \geq \exp\Big(\frac{C^n_{[K,L,n]}}{2^3 \cdot \varepsilon^n}\Big).$$

Therefore, for $\varepsilon > 0$ sufficiently small, we have that

$$\mathcal{N}_{\varepsilon}(\mathcal{SC}_{[K,L,M]} \mid W^{1,1}(\mathbb{R}^n)) \geq \exp\Big(\frac{C^n_{[K,L,n]}}{2^{n+3} \cdot \varepsilon^n}\Big).$$

The proof is complete. □

We now state our main result of the lower estimate of compactness.

Theorem 5.6 *Given $L, M \geq 0$. For $T > 0$ and $\varepsilon > 0$ sufficiently small, it holds*

$$\mathcal{H}_\epsilon(S_T(C_{[L,M]}) + \beta_T \mid W^{1,1}(\mathbb{R}^n)) \geq \frac{\Gamma^-_{[K_T,\frac{L}{2},n]}}{\ln(2)} \cdot \frac{1}{\varepsilon^n} \tag{5.68}$$

where $K_T = \frac{1}{T|D^2H(0)|}$ and $\beta_T = T \cdot H(0)$.

Proof. Let $0 < h < M$ be such that

$$\alpha_h = \|DH^2(p)\|_{\mathbb{L}^\infty([-h,h]^n)} \geq \frac{1}{2} \cdot |DH^2(0)| \quad \text{and} \quad T \cdot \sup_{\|p\| \leq h} |DH(p)| \leq \frac{L}{2}. \tag{5.69}$$

We have first that

$$S_T(C_{[L,h]}) \subset S_T(C_{[L,M]}). \tag{5.70}$$

From theorem 5.4, we get

$$\mathcal{SC}_{\left[K_T,\frac{L}{2},h\right]} \subset S_T(C_{[L,h]}) + \beta_T$$

where $\beta_T = T \cdot H(0)$ and $K_T = \frac{1}{T|D^2H(0)|}$. Therefore, (5.70) implies that

$$\mathcal{SC}_{\left[K_T,\frac{L}{2},h\right]} \subset S_T(C_{[L,M]}) + \beta_T. \tag{5.71}$$

From theorem 5.5, we have that

$$\mathcal{H}_\epsilon\Big(\mathcal{SC}_{\left[K_T,\frac{L}{2},h\right]} \mid \mathbb{W}^{1,1}(\mathbb{R}^n)\Big) \geq \frac{\Gamma^-_{[K_T,\frac{L}{2},n]}}{\ln(2)} \cdot \frac{1}{\varepsilon^n}. \tag{5.72}$$

By combining (5.71) and (5.72), we complete the proof. □

References

[1] G. Alberti and L. Ambrosio, A geometrical approach to monotone functions in $\mathbb{R}^N$. Math. Z. 230 (1999), 259–316.

[2] L. Ambrosio, N. Fusco, D. Pallara, *Functions of bounded variation and free discontinuity problems*, Oxford Mathematical Monographs, The Clarendon Press Oxford University Press 2000.

[3] Fabio Ancona, Olivier Glass, Khai T. Nguyen, *Lower compactness estimates for scalar balance laws*, Comm. Pure Appl. Math. 65 (2012), no. 9, 1303-1329.

[4] Fabio Ancona, Olivier Glass, Khai T. Nguyen, *The compactness estimates for systems of conservation laws*, in preparation.

[5] Fabio Ancona, Piermarco Cannarsa, Khai T. Nguyen, *The compactness estimates for Hamilton-Jacobi equations*, in preparation.

[6] M. Bardi, *A boundary value problem for the minimum time function*, Siam J. Control Optim. 27 (1989), 776–785.

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

[8] M. Bardi, Falcone. M, *An approximation scheme for the minimum time function*, Siam J. Control Optim. 28 (1990), 950–965.

[9] A. Bressan and B. Piccoli, *Introduction to the Mathematical Theory of Control*, Appl. Math. 2, Amer. Inst. Math. Sci., Springfield, MO, 2007.

[10] P. Cannarsa, H. Frankowska & C. Sinestrari, *Optimality conditions and synthesis for the minimum time problem*, Set-Valued Anal. 8 (2000), 127–148.

[11] P. Cannarsa, F. Marino & P. R. Wolenski, *Semiconcavity of the minimum time function for differential inclusions*, Discrete Contin. Dyn. Syst. Ser. B, vol. 19 (2012),187–206.

[12] P. Cannarsa & Khai T. Nguyen, *Exterior sphere condition and time optimal control for differential inclusions*, SIAM J. Control Optim. 46 (2011), 2558–2576.

[13] P. Cannarsa & C. Sinestrari, *Convexity properties of the minimum time function*, Calc. Var. Partial Differential Equations 3 (1995), 273-298.

[14] P. Cannarsa, C. Sinestrari, *Semiconcave Functions, Hamilton-Jacobi Equations, and optimal Control*, Birkhäuser, Boston (2004).

[15] G. Colombo, A. Marigonda, *Differentiability properties for a class of non-convex functions*, Calc. Var. Partial Differential Equations 25 (2006), 1–31.

[16] G. Colombo, Khai T. Nguyen, *On the structure of the Minimum Time Function*, Siam J. Control Optim. 48 (2010), 4776–4814.

[17] G. Colombo, Khai T. Nguyen, *On the minimum time function around the origin*, Math. Control Relat. Fields, to appear.

[18] G. Colombo, Khai T. Nguyen, Luong V. Nguyen, *Non-Lipschitz points and the SBV regularity of the minimum time function*, submitted.

[19] A. Marigonda, Khai T. Nguyen, D. Vittone, *Some regularity results for a class of upper semicontinuous functions*, Indiana Univ. Math. J., in print.

[20] Khai T. Nguyen, *Hypographs satisfying an external sphere condition and the regularity of the minimum time function*, J. Math. Anal. Appl. 372 (2010), 611–628.

[21] Khai T. Nguyen & Davide Vittone *Rectifiability of siguralities of non-lipschitz functions*, Journal of Convex Analysis, 19 (2012).

[22] Hoeffding W., Probability inequalities for sums of bounded random variables. J. Amer. Statist. Assoc. 58 (1963), 13–30.

[23] Petrov N. M., *Controllability of autonomous systems*, Differential equations 4 (1968), 311–317.

[24] Petrov N. M., *On the Bellman function for the time optimal process problem*, J. Appl. Math. Mech 34 (1970), 785-791.

[25] Evans L.C. , Gariepy R. F., *Measure theory and fine properties of functions*, Studies in Advanced Mathematics, CRC Press, Boca Raton, 1992.

[26] Federer Herbert, *Curvature measures*, Trans. Amer. Math. Soc. 93 (1959), 418–491.

[27] Bartlett P. L., Kulkarni S. R., Posner S. E., Covering numbers for real-valued function classes. IEEE Trans. Inform. Theory 43 (1997), no. 5, 1721–1724.