Metric entropy for Hamilton-Jacobi equation with uniformly directionally convex Hamiltonian

Stefano Bianchini[(1)], Prerona Dutta[(2)] and Khai T. Nguyen[(2)]

[(1)] SISSA, via Beirut 2, IT-34014 Trieste, ITALY
[(2)] Department of Mathematics, NC State University, Raleigh, NC 27695, USA

E-mails: bianchin@sissa.it, pdutta2@ncsu.edu, khai@math.ncsu.edu

Preprint SISSA 29/2020/MATE

Abstract

The present paper first aims to study the BV-type regularity for viscosity solutions of the Hamilton-Jacobi equation

$$u_t(t,x)+H\big(D_xu(t,x)\big) \ = \ 0 \qquad \text{for all } (t,x)\in]0,\infty[\times\mathbb{R}^d$$

with coercive and uniformly directionally convex Hamiltonian $H\in\mathcal{C}^1(\mathbb{R}^d)$. More precisely, we establish a BV bound on the slope of backward characteristics $DH(u(t,\cdot))$ starting at a positive time $t>0$. Relying on the BV bound, we quantify the metric entropy in $\mathbf{W}^{1,1}_{\mathrm{loc}}(\mathbb{R}^d)$ for the map S_t that associates to every given initial data $u_0\in\mathbf{Lip}\big(\mathbb{R}^d\big)$ the corresponding solution S_tu_0. Finally, a counter example is constructed to show that both $D_xu(t,\cdot)$ and $DH(D_xu(t,\cdot))$ fail to be in BV_{loc} for general strictly convex and coercive $H\in\mathcal{C}^2(\mathbb{R}^d)$.

Keywords: Hamilton-Jacobi equations, Hopf-Lax semigroup, Kolmogorov entropy, semiconcave functions, bounded variation

1 Introduction

Consider a first-order Hamilton-Jacobi equation

$$u_t(t,x)+H\big(D_xu(t,x)\big) \ = \ 0 \qquad \text{for all } (t,x)\in]0,\infty[\times\mathbb{R}^d \tag{1.1}$$

where $u:[0,+\infty[\times\mathbb{R}^d\to\mathbb{R}^d$, $D_xu=(u_{x_1},\dots,u_{x_d})$, and $H:\mathbb{R}^d\to\mathbb{R}$ is a Hamiltonian. Due to the nonlinear dependence of the characteristic speeds on the gradient of the solution, in general a classical solution u will develop singularities, and the gradient Du will become discontinuous in finite time. To cope with this difficulty, the concept of viscosity solution was introduced by Crandall and Lions in [10] to guarantee global existence, uniqueness and stability of the Cauchy problem. Under standard assumptions of the convexity and the coercivity on Hamiltonian H, (1.1) generates a Hopf-Lax semigroup of viscosity solutions $(S_t)_{t\geq 0}: \mathbf{Lip}(\mathbb{R}^d) \to \mathbf{Lip}(\mathbb{R}^d)$. More precisely, for every Lipschitz initial data $u_0 \in \mathbf{Lip}(\mathbb{R}^d)$, the corresponding unique viscosity solution of equation (1.1) with $u(0,x) = u_(x)$ is computed by the Hopf-Lax representation formula

$$u(t,x) = S_t(u_0)(x) \;\; = \;\; \min_{y\in\mathbb{R}^d} \left\{ u_0(y) + t \cdot L\left(\frac{x-y}{t}\right)\right\} \tag{1.2}$$

where L is the Legendre transform of H. In addition, if H is strongly convex, i.e., there exists a constant $\lambda > 0$ such that

$$D^2H(p) \;\geqslant\; \lambda \cdot \mathbf{I}_d \qquad \text{for all } p \in \mathbb{R}^d,$$

then the map $x \to u(t,x) - \dfrac{1}{2\lambda t} \cdot \|x\|^2$ is concave for every $t > 0$. In particular, $u(t,\cdot)$ is almost everywhere twice differentiable and $D_xu(t,\cdot)$ has locally bounded total variation. Moreover, thanks to Helly's compactness theorem, the map $S_t : \mathbf{Lip}(\mathbb{R}^d) \to \mathbf{Lip}(\mathbb{R}^d)$ is compact in $\mathbf{W}^{1,1}_{\text{loc}}(\mathbb{R}^d)$. A natural question arises on how to measure the degree of compactness of S_t. This involves using the concept of metric entropy (or ε-entropy), introduced by Kolmogorov and Tikhomirov in [14]:

Definition 1.1 *Let (E,ρ) be a metric space and K a totally bounded subset of E. For $\varepsilon > 0$, let $\mathcal{N}_\varepsilon(K)$ be the minimal number of sets in a covering of K by subsets of E having diameter no larger than 2ε. Then the ε-entropy of K is defined as*

$$\mathcal{H}_\varepsilon(K|E) \doteq \log_2 \mathcal{N}_\varepsilon(K|E).$$

In other words, it is the minimum number of bits needed to represent a point in a given set K in the space E, up to an accuracy ε with respect to the metric ρ.

Such a question stems from a conjecture of Lax in [15] for scalar conservation laws with uniformly convex fluxes. A complete answer to Lax's conjecture was provided in [5, 13]. This study was also extended to hyperbolic systems of conservation laws in [7, 8]. Recently, the first results on the ε-entropy for sets of viscosity solutions of (1.1) were obtained in [3]. The authors proved that the minimal number of bits needed to represent a viscosity solution of (1.1) up to accuracy ε with respect to the $\mathbf{W^{1,1}}$-distance is of order ε^{-d} under the strongly convex condition on Hamiltonian H. A similar result was also proved in [4] by the same authors, when the Hamiltonian depends on the state variable x. The main idea is to provide controllability results for Hamilton-Jacobi equations and a compactness result for a class of semiconcave functions. However, such a gain of BV regularity does not hold for (1.1) with general strictly convex Hamiltonian functions and the previous approach to finding the ε-entropy of the solution set in [3, 4] cannot be applied. In this case, a study of the fine regularity properties of viscosity solutions is still lacking, and quantitative estimates on the ε-entropy of viscosity solution sets are not yet available.

In the present paper, we extend the analysis of the metric entropy for sets of viscosity solutions to (1.1) when the Hamiltonian $H \in \mathcal{C}^1(\mathbb{R}^d)$ is strictly convex, coercive, and in the form of a uniformly directionally convex function, i.e., for every constant $R > 0$ it holds that

$$\inf_{p\neq q\in\overline{B}(0,R)} \left\langle \frac{DH(p)-DH(q)}{|DH(p)-DH(q)|}, \frac{p-q}{|p-q|} \right\rangle \ := \ \lambda_R \ > \ 0. \tag{1.3}$$

By the Hopf-Lax representation formula (1.2), it is well-known from [9] that the set of slopes of backward optimal rays through (t, x), denoted by

$$\mathbf{b}(t,x) = \left\{ \frac{x-y}{t} : u(t,x) = \bar{u}(y) + t \cdot L\left(\frac{x-y}{T}\right) \right\}, \tag{1.4}$$

reduces to a singleton $\mathbf{b}(t, x) = DH(D_x u(t, x))$ for almost every $(t, x) \in \mathbb{R}_+ \times \mathbb{R}^d$. Moreover, if M is a Lipschitz constant of u_0 then for all $t > 0$, $\mathbf{b}(t, \cdot)$ can be viewed as an element in $\mathbf{L}^\infty(\mathbb{R}^d)$ with

$$\|\mathbf{b}(t,\cdot)\|_{\mathbf{L}^\infty(\mathbb{R}^d)} \ \leq \ \Lambda_M \ := \ \max\{|q| : L(q) \leq M|q|\}. \tag{1.5}$$

Towards the sharp estimate on the ε-entropy of the semigroup S_t for all $t > 0$, we first establish a BV bound on $\mathbf{b}(t, \cdot)$.

Theorem 1.2 *Assume that $H \in \mathcal{C}^1(\mathbb{R}^d)$ is strictly convex, coercive, and satisfies (1.3). For every $t > 0$ and $u_0 \in \mathbf{Lip}(\mathbb{R}^d)$ with a Lipschitz constant M, the function $\mathbf{b}(t, \cdot)$ has locally bounded variation, and for every open and bounded set $\Omega \subset \mathbb{R}^d$ of finite perimeter*

$$|D\mathbf{b}(t,\cdot)|(\Omega) \ \leq \ \frac{1}{\gamma_M} \cdot \left(\Lambda_M + \frac{\mathrm{diam}(\Omega)}{t}\right) \cdot \mathcal{H}^{d-1}(\partial\Omega) + \frac{\sqrt{d}}{t} \cdot |\Omega| \tag{1.6}$$

with $\gamma_M := \lambda_{\left(\max_{|q|\leq\Lambda_M} |DL(q)|\right)}$.

Intuitively, the uniform directional convexity of H yields a bound on the directional derivatives of $\mathbf{b}(t, \cdot)$ in terms of $\mathrm{div}_x(\mathbf{b}(t, \cdot))$. Indeed, the main idea in the proof of Theorem 1.2 is to achieve an upper bound on the quotient $|D_x\mathbf{b}(t, \cdot)|/|\mathrm{div}_x\mathbf{b}(t, \cdot)|$ for a suitable sequence of approximate solutions, which converges uniformly and monotonically to the given solution. In turn, the approximations of $\mathbf{b}$ will converge (in the sense that their graphs converge with respect to the Hausdorff distance). As a consequence of Theorem 1.2, the map $S_t : \mathbf{Lip}(\mathbb{R}^d) \to \mathbf{Lip}(\mathbb{R}^d)$ is compact in $\mathbf{W}^{1,1}_{\mathrm{loc}}(\mathbb{R}^d)$ for every $t > 0$. Indeed, for every $m, M > 0$, denote by

$$\mathcal{U}_{[m,M]} \ := \ \left\{\bar{u} \in \mathbf{Lip}(\mathbb{R}^d) : |\bar{u}(0)| \leq m, \mathrm{Lip}[\bar{u}] \leq M\right\},$$

the set $S_t\left(\mathcal{U}_{[m,M]}\right)$ is compact in $\mathbf{W}^{1,1}_{\mathrm{loc}}(\mathbb{R}^d)$. From the BV bound in (1.6) and a result in [12], the ε-entropy of the sets of slopes of optimal rays starting at time t in $\mathbf{L}^1_{\mathrm{loc}}(\mathbb{R}^d)$ is of the order ε^{-d}. By establishing a quantitative relation (depending on the nonlinearity of H) between the $\mathbf{W}^{1,1}$-distance of two solutions and the $\mathbf{L}^1$-distance of slopes of two corresponding optimal rays, in Theorem 4.1 we provide upper and lower estimates of $\mathcal{H}_\varepsilon\left(S_t\left(\mathcal{U}_{[m,M]}\right) \middle| \mathbf{W}^{1,1}(]-R,R[^d)\right)$ for all $t, R > 0$ and $\varepsilon > 0$ sufficiently small. In particular, if $H(p) = |p|^{2k}$ for $k \geq 1$ then

$$\mathcal{H}_\varepsilon\left(S_T\left(\mathcal{U}_{[m,M]}\right) \middle| \mathbf{W}^{1,1}(]-R,R[^d)\right) \ \approx \ \varepsilon^{-(2k-1)d}.$$

In the last section, a counter-example is constructed to show that if $H \in \mathcal{C}^2(\mathbb{R}^d)$ is strictly convex and coercive but does not satisfy (1.3) then both $Du(t,\cdot)$ and $\mathbf{b}(t,\cdot)$ fail to be in BV_{loc} in general.

2 Notations and preliminaries

Given a positive integer d and a measurable set $\Omega \subseteq \mathbb{R}^d$, throughout the paper we shall denote by

- $|\cdot|$, the Euclidean norm in $\mathbb{R}^d$ and for any $R > 0$

$$B_d(x,R) \;=\; \{y \in \mathbb{R}^d : |x-y| < R\}, \qquad \square_R \;=\;]-R,R[^d;$$

- $\langle\cdot,\cdot\rangle$, the Euclidean inner product in $\mathbb{R}^d$;
- $\partial\Omega$, the boundary of Ω;
- $[x,y]$, the segment joining two points $x, y \in \mathbb{R}^d$;
- $\#S$, the number of elements in any finite set S;
- $\mathrm{Vol}(D)$, the Lebesgue measure of a measurable set $D \subset \mathbb{R}^d$,
- $\omega_d := \mathrm{Vol}(B(0,1))$, the Lebesgue measure of the unit ball in $\mathbb{R}^d$,
- $\mathbf{L}^1(\Omega)$, the Lebesgue space of all (equivalence classes of) summable real valued functions on Ω, equipped with the usual norm $\|\cdot\|_{\mathbf{L}^1(\Omega)}$ (we shall use the same symbol in case u is vector-valued);
- $\mathbf{L}^\infty(\Omega)$, the space of all essentially bounded real valued functions functions on Ω, and by $\|u\|_{\mathbf{L}^\infty(\Omega)}$ the essential supremum of a function $u \in \mathbf{L}^\infty(\Omega)$ (we shall use the same symbol in case u is vector-valued);
- $\mathbf{W}^{1,1}(\Omega)$, the Sobolev space of functions with summable first order distributional derivatives, and by $\|\cdot\|_{\mathbf{W}^{1,1}(\Omega)}$ its norm;
- $\mathcal{C}^1(\Omega)$, the space of continuously differentiable functions on Ω;
- $BV(\Omega,\mathbb{R}^m)$, the space of all vector-valued functions $F : \Omega \to \mathbb{R}^m$ of bounded variation (i.e., all $F \in \mathbf{L}^1(\Omega,\mathbb{R}^m)$ such that the first partial derivatives of F in the sense of distributions are measures with finite total variation in Ω);
- $\mathbf{Lip}(\Omega)$, the space of all Lipschitz functions $f : \Omega \to \mathbb{R}$, and $\mathrm{Lip}[f]$ the Lipschitz seminorm of f;
- $\mathcal{H}^k(E)$, the k-dimensional Hausdorff measure of $E \subset \mathbb{R}^d$;
- For any function f, the function $f_{\llcorner\Omega}$ is the restriction of f on Ω;
- $\mathbf{I}_d$, the identity matrix of size d;
- $\lfloor a \rfloor := \max\{z \in \mathbb{Z} : z \le a\}$, the integer part a;

2.1 Semiconcave and BV functions in $\mathbb{R}^d$

2.1.1 Semiconcave functions

Let us recall here some basic definitions and properties of semiconcave (semiconvex) functions in $\mathbb{R}^d$. We refer to [9] for a general introduction to the respective theories.

Definition 2.1 *A continuous function $u : \Omega \to \mathbb{R}$ is semiconcave if there exists a nondecreasing continuous function $\omega : [0, \infty[\to [0, \infty[$ with $\lim_{s\to 0+} \omega(s) = 0$ such that*

$$u(x + h) + u(x - h) - 2u(x) \ \leq \ \omega(|h|) \cdot |h|$$

for all $x, h \in \mathbb{R}^d$ such that $[x - h, x + h] \subset \Omega$. We will say that

- *u is semiconcave in Ω with a semiconcavity constant K if $\omega(s) = Ks$ for all $s \in [0, +\infty[$;*
- *u is semiconvex (with constant -K) if $-u$ is semiconcave (with constant K);*
- *u is locally semiconcave (semiconvex) if u is semiconcave (semiconvex) in every compact set $A \subset \Omega$.*

For every $x \in \Omega$ with $\Omega \subseteq \mathbb{R}^d$ open, the sets

$$D^+u(x) \ := \ \left\{ p \in \mathbb{R}^d : \limsup_{y\to x} \frac{u(y) - u(x) - \langle p, y - x\rangle}{|y - x|} \leq 0 \right\}$$

and

$$D^-u(x) \ := \ \left\{ p \in \mathbb{R}^d : \liminf_{y\to x} \frac{u(y) - u(x) - \langle p, y - x\rangle}{|y - x|} \geq 0 \right\}$$

are called the superdifferential and the subdifferential of u at x respectively. It is clear that $D^{\pm}u(x)$ is convex and

$$D^-u(x) \ = \ -D^+(-u)(x) \qquad \text{for all } x \in \Omega.$$

From [9, Proposition 3.3.4, Proposition 3.3.10], the superdifferential of a semiconcave function enjoys the following properties:

Proposition 2.1 *Let $u : \Omega \to \mathbb{R}$ be locally semiconcave with $\Omega \subseteq \mathbb{R}^d$ open convex. Then:*

(i) The superdifferential $D^+u(x)$ is a compact, convex, nonempty set for all $x \in \Omega$. Moreover, the set-valued map $x \mapsto D^+u(x)$ is upper semicontinuous;

(ii) $D^+u(x)$ is a singleton if and only if u is differentiable at x;

(iii) If $D^+u(x)$ is a singleton for all $x \in \Omega$, then $u \in C^1(\Omega)$;

(iv) For every $x, y \in \Omega$, it holds that

$$\langle p_y - p_x, y - x\rangle \ \leq \ \omega(|x - y|) \cdot |y - x|$$

for all $p_x \in D^+u(x)$ and $p_y \in D^+u(y)$.

As a consequence of (ii)-(iii) if u is both locally semiconcave and locally semiconvex then u is in $C^1(\Omega)$. This is crucial to prove a further regularity for viscosity solutions in Proposition 2.2 which allows us to construct a backward smooth solutions of (1.1). From (iv), one immediately gets:

Corollary 2.2 *If $u : \Omega \to \mathbb{R}$ is semiconvex with constant $-K$ then for every $x, y \in \Omega$, it holds*

$$\langle p_y - p_x, y - x\rangle \ \geq \ -K \cdot |y-x|^2$$

for all $p_x \in D^- u(x)$ and $p_y \in D^- u(y)$.

For any given constants $r, K > 0$, we denote by

$$\mathcal{SC}_{[r,K]} \ := \ \Big\{ v \in \mathbf{Lip}(\mathbb{R}^d) : \mathrm{Lip}[v] \leq r \text{ and } v \text{ is semiconcave with constant } K \Big\} \qquad (2.1)$$

From the proof of [3, Proposition 10], one obtains an lower bound on the ε-entropy for the set $\left\{ Dv_{\llcorner \Box_R} : \mathcal{SC}_{[r,K]} \right\}$ in $\mathbf{L}^1(\Box_R)$ which will be used to establish a lower estimate on the ε-entropy of a set of viscosity solutions in subsection 4.2.

Corollary 2.3 *Given any $r, R, K > 0$, for every*

$$0 \ < \ \varepsilon \ \leq \ \min\{r, K\} \cdot \frac{\omega_d \cdot R^d}{(d+1)2^{d+8}},$$

there exists a subset $\mathcal{G}^R_{[r,K]}$ of $\mathcal{SC}_{[r,K]}$ such that

$$\#\mathcal{G}^R_{[r,K]} \ \geq \ 2^{\beta_{[R,K]} \cdot \varepsilon^{-d}} \qquad \text{with} \qquad \beta_{[R,K]} \ = \ \frac{1}{3^d 2^{d^2+4d+3} \ln 2} \cdot \left(\frac{K \omega_d R^{d+1}}{(d+1)} \right)^d$$

and

$$\left\| Dv_{\llcorner \Box_R} - Dw_{\llcorner \Box_R} \right\|_{\mathbf{L}^1(\Box_R)} \ \geq \ 2\varepsilon \qquad \textit{for all } v \neq w \in \mathcal{G}^R_{[r,K]}.$$

2.1.2 Functions of bounded total variations

Let us now introduce the concept of functions of bounded variation. We refer to [2] for a comprehensive analysis.

Definition 2.4 *The function $u \in \mathbf{L}^1(\Omega)$ is a function of bounded variation on $\Omega \subseteq \mathbb{R}^d$, denoted by $BV(\Omega, \mathbb{R}^m)$, if the distributional derivative of u, denote by Du, is an $m \times d$ matrix of finite measures $D_i u^\alpha$ in Ω satisfying*

$$\sum_{\alpha=1}^m \int_\Omega u^\alpha div \varphi^\alpha \, dx \ = \ -\sum_{\alpha=1}^m \sum_{i=1}^d \int_\Omega \varphi_i^\alpha d D_i u^\alpha \quad \textit{for all } \varphi \in \left[\mathcal{C}_c^1(\Omega, \mathbb{R}^d)\right]^m, i \in \{1, \dots, d\}.$$

We denote by $|Du|$ the total variation of Du, i.e.,

$$|Du|(\Omega) \ = \ \sup\left\{ \sum_{\alpha=1}^m \int_\Omega u^\alpha div^\alpha \varphi : \varphi \in \left[\mathcal{C}_c^1(\Omega, \mathbb{R}^d)\right]^m, \|\varphi\|_\infty \leq 1 \right\}.$$

We recall a Poincaré-type inequality for bounded total variation functions on convex domains that will be used in the paper. This result is based on [1, Theorem 3.2].

Theorem 2.5 *(Poincaré inequality) Let $\Omega \subset \mathbb{R}^n$ be an open, bounded, convex set with Lipschitz boundary. For any $u \in BV(\Omega, \mathbb{R})$, it holds*

$$\int_\Omega \left|u(x) - u_\Omega\right| \, dx \ \leq \ \frac{\mathrm{diam}(\Omega)}{2} \cdot |Du|(\Omega)$$

where

$$u_\Omega \ = \ \frac{1}{\mathrm{Vol}(\Omega)} \cdot \int_\Omega u(x) \, dx$$

is the mean value of u over Ω.

To complete this subsection, let us recall a result on the metric entropy for a class of functions with bounded total variation which will be used in subsection 4.1. For any given constants $R, M, V > 0$, we consider a of uniformly bounded total variation functions on $\Box_R$

$$\mathcal{F}_{[R,M,V]} \ = \ \left\{ f : \Box_R \to \mathbb{R}^d \ \middle| \ \|f\|_{\mathbf{L}^\infty(\Box_R)} \leq M, |Df|(\Box_R) \leq V \right\}. \tag{2.2}$$

By a slightly modification in the proof of [12, Theorem 1], one can obtain the following upper bound of the ε-entropy of $\mathcal{F}_{[R,M,V]}$ in $\mathbf{L}^1(\Box_R)$.

Corollary 2.6 *For every $0 < \varepsilon < \min\left\{1, 4RV, 4VM^{-1/d}\right\}$, it holds*

$$\mathcal{H}_\varepsilon\left(\mathcal{F}_{[R,M,V]} \ \middle| \ \mathbf{L}^1(\Box_R)\right) \ \leq \ 48d \cdot \left(\frac{12\sqrt{d}RV}{\varepsilon}\right)^d. \tag{2.3}$$

Proof. By the definition of the ε-entropy, we have

$$\mathcal{H}_\varepsilon\left(\mathcal{F}_{[R,M,V]} \ \middle| \ \mathbf{L}^1(\Box_R)\right) \ \leq \ d \cdot \mathcal{H}_\varepsilon\left(\mathcal{F}^1_{[R,M,V]} \ \middle| \ \mathbf{L}^1(\Box_R)\right) \tag{2.4}$$

with

$$\mathcal{F}^1_{[R,M,V]} \ = \ \left\{ f : \Box_R \to \mathbb{R} \ \middle| \ \|f\|_{\mathbf{L}^\infty(\Box_R)} \leq M, |Df|(\Box_R) \leq V \right\}.$$

Consider a class of bounded total variation real valued functions

$$\mathcal{B}_{[R,M,V]} \ = \ \left\{ g : [0,R] \to [0,M] \ \middle| \ |Dg|([0,R]) \leq V \right\}.$$

From [12, Lemma 2.3], for every $0 < \varepsilon < \frac{RV}{3}$, one has

$$\mathcal{N}_\varepsilon\left(\mathcal{B}_{[R,9/8V,V]} \middle| \mathbf{L}^1([0,R])\right) \ \leq \ 2^{\frac{17RV}{\varepsilon}}$$

and this implies that

$$\mathcal{N}_\varepsilon\left(\mathcal{B}_{[R,M,V]} \middle| \mathbf{L}^1([0,R])\right) \ \leq \ \frac{8M}{V} \cdot \mathcal{N}_\varepsilon\left(\mathcal{B}_{[R,9/8V,V]} \middle| \mathbf{L}^1([0,R])\right) \ \leq \ \frac{8M}{V} \cdot 2^{\frac{17RV}{\varepsilon}}.$$

In particular, for every $0 < \varepsilon < \dfrac{RV^2}{3V+M}$ such that $\dfrac{8M}{V} \leq 2^{\frac{RV}{3}}$, it holds

$$\mathcal{H}_\varepsilon\left(\mathcal{B}_{[R,M,V]}\Big|\mathbf{L}^1([0,R])\right) \;=\; \log_2\left(\mathcal{N}_\varepsilon\left(\mathcal{B}_{[R,M,V]}\Big|\mathbf{L}^1([0,R])\right)\right) \;\leq\; \frac{18RV}{\varepsilon}.$$

Using the above estimate, one can following the same argument in the proof of [12, Theorem 1] to obtain that for every $0 < \varepsilon < \min\left\{1, 4RV, 4VM^{-1/d}\right\}$, it holds

$$\mathcal{H}_\varepsilon\left(\mathcal{F}^1_{[R,M,V]}\;\Big|\;\mathbf{L}^1(\Box_R)\right) \;\leq\; 48\cdot\left(\frac{12\sqrt{d}RV}{\varepsilon}\right)^d$$

and (2.4) yields (2.3). □

2.2 Semigroup of Hamilton-Jacobi equation

Consider the Hamilton-Jacobi equation (1.1) with coercive and strictly convex Hamiltonian $H \in \mathcal{C}^1(\mathbb{R}^d)$, i.e., $\displaystyle\lim_{|p|\to\infty}\frac{H(p)}{|p|} \;=\; +\infty$ and

$$H(tp+(1-t)q) \;<\; t\cdot H(p)+(1-t)H(q) \quad \text{for all } t\in]0,1[, p,q\in\mathbb{R}^d,$$

and satisfying the additional condition (1.3). Moreover, without loss of generality, we shall assume that the Hamiltonian satisfies further conditions

$$H(0) \;=\; 0 \qquad \text{and} \qquad DH(0) \;=\; 0, \tag{2.5}$$

otherwise the transformations $x \mapsto x + tDH(0)$, $u(t,\cdot) \mapsto u(t,x) + t\cdot H(0)$ and $H(p) \mapsto H(p) - \langle DH(0), p\rangle$ reduce the general case to this one.

Remark 2.7 *If $H \in \mathcal{C}^2(\mathbb{R}^d)$ satisfies*

$$D^2H(p) \;=\; |D^2H(p)|\cdot A(p) \qquad \text{with} \qquad A(p) \;\geq\; \lambda\cdot\mathbf{I}_d \tag{2.6}$$

for some $\lambda > 0$ then H satisfies (1.3).

Proof. For any $p \neq q \in \mathbb{R}^d$, by mean value theorem, it holds that

$$\begin{aligned} DH(p) - DH(q) = &\int_0^1 DH^2(tp+(1-t)q)\cdot(p-q)dt \\ = &\left[\int_0^1 A(tp+(1-t)q)\big|D^2H(tp+(1-t)q)\big|dt\right]\cdot(p-q) \end{aligned}$$

and

$$|DH(p)-DH(q)| \;\leq\; |p-q|\cdot\int_0^1 |DH^2(tp+(1-t)q)|dt.$$

Thus, using (2.6), we estimate

$$\begin{aligned}\langle DH(p)-DH(q),p-q\rangle = & \int_0^1 \Big[(p-q)^T A(tp+(1-t)q)(p-q)\Big]\cdot\big|D^2H(tp+(1-t)q)\big|dt \\ \geq & \ \lambda\cdot|p-q|^2\int_0^1 \big|D^2H(tp+(1-t)q)\big|dt \\ = & \ \lambda\cdot|DH(p)-DH(q)|\cdot|p-q|\end{aligned}$$

and this implies (1.3). □

It is well-known that classical smooth solutions of (1.1) in general break down and Lipschitz continuous functions that satisfy (1.1) almost everywhere together with a given initial condition are not unique. To handle this problem, the following concept of a generalized solution was introduced in [10] to guarantee global existence and uniqueness results.

Definition 2.8 *(Viscosity solution) We say that a continuous function* $u:[0,T]\times\mathbb{R}^d$ *is a viscosity solution of* (1.1) *if:*

(1) *u is a viscosity subsolution of* (1.1)*, i.e., for every point* $(t_0,x_0)\in]0,T[\times\mathbb{R}^d$ *and test function* $v\in C^1\big(]0,+\infty[\times\mathbb{R}^d\big)$ *such that* $u-v$ *has a local maximum at* (t_0,x_0)*, it holds that*
$$v_t(t_0,x_0)+H\big(D_xv(t_0,x_0)\big)\leqslant 0\,,$$

(2) *u is a viscosity supersolution of (1.1), i.e., for every point* $(t_0,x_0)\in]0,T[\times\mathbb{R}^d$ *and test function* $v\in C^1\big(]0,+\infty[\times\mathbb{R}^d\big)$ *such that* $u-v$ *has a local minimum at* (t_0,x_0)*, it holds that*
$$v_t(t_0,x_0)+H\big(D_xv(t_0,x_0)\big)\geqslant 0\,.$$

By the alternative equivalent definition of viscosity solution expressed in terms of the sub- and superdifferential of the function (see [10]), and because of Proposition 2.1 one immediately sees that every C^1 solution of (1.1) is also a viscosity solution of (1.1). On the other hand, if u is a viscosity solution of (1.1) then u satisfies the equation at every point of differentiability. Let us state a result on further regularity for viscosity solutions (see in [3, Proposition 3]) which says that smoothness in the pair (t,x) follows from smoothness in the second variable.

Proposition 2.2 *Let u be a viscosity solution of (1.1) in* $[0,T]\times\mathbb{R}^d$*. If* $u(t,\cdot)$ *is both locally semiconcave and semiconvex in* $\mathbb{R}^d$ *for all* $t\in]0,T]$ *then u is a* C^1 *solution of (1.1) in* $]0,T]\times\mathbb{R}$*.*

The viscosity solution of the Hamilton-Jacobi equation (1.1) with initial data $u(0,\cdot)=u_0\in$ Lip$(\mathbb{R}^n)$ can be represented as the value function of a classical problem in calculus of variation, which admits the Hopf-Lax representation formula

$$u(t,x)=\min_{y\in\mathbb{R}^n}\left\{t\cdot L\Big(\frac{x-y}{t}\Big)+u_0(y)\right\}.\qquad t>0,\quad x\in\mathbb{R}^d\,, \tag{2.7}$$

where $L \in \mathcal{C}^1(\mathbb{R}^d)$ denotes the Legendre transform of H, defined by

$$L(q) := \max_{p\in\mathbb{R}^d} \{p \cdot q - H(p)\} \qquad q \in \mathbb{R}^d. \tag{2.8}$$

The main properties of viscosity solutions defined by the Hopf-Lax formula, which are of interest to this paper are recalled below (cfr. [9, Section 1.1, Section 6.4]).

Proposition 2.3 *Let u be the viscosity solution of* (1.1) *on $[0, +\infty[\, \times \mathbb{R}^d$, with continuous initial data u_0, defined by* (2.7)*. Then the followings hold true:*

(*i*) ***Functional identity**: For all $x \in \mathbb{R}^d$ and $0 \leqslant s < t$, it holds that*

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

(*ii*) ***Differentiability of u and uniqueness**:* (2.7) *admits a unique minimizer y_x if and only if $u(t, \cdot)$ is differentiable at x. In this case we have*

$$y_x \;=\; x - t \cdot DH\big(D_x u(t, x)\big) \qquad \text{and} \qquad D_x u(t, x) \;\in\; D^- u_0(y_x).$$

(*iii*) ***Dynamic programming principle**: Let $t > s > 0$, $x \in \mathbb{R}^d$, assume that y is a minimizer for* (2.7)*, and define $z = \frac{s}{t}x + \Big(1 - \frac{s}{t}\Big)\,y$. Then y is the unique minimizer over $\mathbb{R}^d$ of*

$$w \mapsto s \cdot L\Big(\frac{z - w}{s}\Big) + u_0(w) \quad \textit{for all } w \in \mathbb{R}^d.$$

As a consequence, the family of nonlinear operators

$$S_t : \mathbf{Lip}(\mathbb{R}^d) \to \mathbf{Lip}(\mathbb{R}^d), \qquad u_0 \mapsto S_t u_0, \quad t \geqslant 0,$$

defined by

$$\begin{cases} S_t u_0(x) \;:=\; \min_{y\in\mathbb{R}^d} \left\{t \cdot L\Big(\dfrac{x - y}{t}\Big) + u_0(y)\right\} & t > 0\,,\ x \in \mathbb{R}^d, \\[2ex] S_0 u_0(x) \;:=\; u(x) & x \in \mathbb{R}^d, \end{cases} \tag{2.9}$$

enjoys the following properties:

(i) For every $u_0 \in \mathbf{Lip}(\mathbb{R}^d)$, $u(t, x) := S_t u_0(x)$ provides the unique viscosity solution of the Cauchy problem (1.1) with initial data $u(0, \cdot) = u_0$.

(ii) (Semigroup property)

$$S_{t+s} u_0 = S_t\, S_s u_0\,, \quad \text{for all } t, s \geq 0\,, \text{ for all } u_0 \in \mathbf{Lip}(\mathbb{R}^d).$$

(iii) (Translation) For every constant $c \in \mathbb{R}$ we have that

$$S_t(u_0 + c) = S_t u_0 + c\,, \qquad \text{for all } u_0 \in \mathbf{Lip}(\mathbb{R}^d)\,, \text{ for all } t \geqslant 0\,. \tag{2.10}$$

(iv) The map S_t is continuous on sets of functions with uniform Lipschitz constant w.r.t $\mathbf{W}^{1,1}_{\mathrm{loc}}$-topology, i.e., for every $u_n \in \mathbf{Lip}(\mathbb{R}^d)$ with a Lipschitz constant M such that

$$u_n \;\longrightarrow\; u \quad \text{in } \mathbf{W}^{1,1}_{\mathrm{loc}}(\mathbb{R}^d),$$

we have that $S_t(u_n)$ also converges to $S_t(u)$ in $\mathbf{W}^{1,1}_{\mathrm{loc}}(\mathbb{R}^d)$.

3 BV bound on $\mathbf{b}(t,\cdot)$

In this section, we shall establish a BV bound on $\mathbf{b}(t,\cdot)$ stated in Theorem 1.2 for a given $t>0$. Here, we recall that

$$\mathbf{b}(t,x)=\left\{\frac{x-y}{t}: u(t,x)=u_0(y)+t\cdot L\left(\frac{x-y}{T}\right)\right\}\qquad \text{for all } x\in\mathbb{R}^d,$$

and u_0 is a Lipschitz function with a Lipschitz constant M. For any $n\geq 1$, set

$$\mathcal{Z}_n \;=\; \frac{2^{-n+1}}{\sqrt{d}}\cdot\mathbb{Z}^d \;=\; \{y_1,y_2,\dots,y_k,\dots\},$$

we approximate the solution u by a monotone decreasing sequence of continuous function $u_n:]0,+\infty[\times\mathbb{R}^d\to\mathbb{R}$ defined by

$$u_n(t,x) \;:=\; \min_{y\in\mathcal{Z}_n}\left\{(1-\varepsilon_n)\cdot u_0(y)+t\cdot L\left(\frac{x-y}{t}\right)\right\} \tag{3.1}$$

with

$$\varepsilon_n \;=\; \frac{1}{M\cdot\Lambda_M}\cdot\max_{|q|\leq\frac{2^{-n}}{t}}\left[M|q|+L(q)\right]\qquad\text{and}\qquad \Lambda_M \;=\; \max\{|q|: L(q)\leq M|q|\}. \tag{3.2}$$

Let $C^n_{t,x}$ be the set of optimal points $y\in\mathcal{Z}_n$ such that (3.1) holds, i.e.,

$$C^n_{t,x} \;:=\; \arg\min_{y\in\mathcal{Z}_n}\left\{(1-\varepsilon_n)\cdot u_0(y)+t\cdot L\left(\frac{x-y}{t}\right)\right\},$$

we define the multivalued-function $\mathbf{b}_n:]0,\infty[\times\mathbb{R}^d\to\mathbb{R}^d$ by

$$\mathbf{b}_n(t,x) \;=\; \left\{\frac{x-y}{t}: y\in C^n_{t,x}\right\}\qquad \text{for all } (t,x)\in]0,\infty[\times\mathbb{R}^d. \tag{3.3}$$

1. For any $x\in\mathbb{R}^d$ and $y^n_x\in C^n_{t,x}$, it holds

$$\begin{aligned} u_n(t,x')-u_n(t,x) &\leq t\cdot\left[L\left(\frac{x'-y^n_x}{t}\right)-L\left(\frac{x-y^n_x}{t}\right)\right]\\ &\leq DL\left(\frac{x-y^n_x}{t}\right)\cdot(x'-x)+O(|x'-x|)\end{aligned}$$

for all $x'\in\mathbb{R}^d$. In particular, we have

$$\mathbf{b}_n(t,x)\;\subseteq\; DH(D^+u_n(t,x))$$

and the set of points where $\mathbf{b}_n(t,\cdot)$ is not singleton,

$$\Sigma^n_t \;=\; \{x \mid \#\mathbf{b}_n(t,x)\geq 2\},$$

is $\mathcal{H}^{n-1}$-rectifiable. Moreover, $\{\mathbf{b}_n(t,\cdot)\}_{n\geq1}$ is a bounded sequence in $[\mathbf{L}^\infty(\mathbb{R}^d)]^d$. Indeed, for any $(x,y^n_x)\in\mathbb{R}^d\times C^n_{t,x}$, one has that $\{\mathbf{b}_n(t,\cdot)\}_{n\geq1}$ is a bounded sequence in

$[\mathbf{L}^\infty(\mathbb{R}^d)]^d$. Indeed, for any given $(x, y_x^n) \in \mathbb{R}^d \times C_{t,x}^n$, let $\bar{x} \in \mathcal{Z}_n$ be the closet point to x such that $|x - \bar{x}| \leq 2^{-n}$. Using the Lipschitz estimate on u_0, we obtain

$$\begin{aligned}
L\left(\frac{x-y_x^n}{t}\right) &\leq \frac{1-\varepsilon_n}{t}\cdot\left(u_0(\bar{x}) - u_0(y_x^n)\right) + L\left(\frac{x-\bar{x}}{t}\right)\\
&\leq \frac{(1-\varepsilon_n)\cdot M}{t}\cdot\left(|x-y_x^n| + 2^{-n}\right) + L\left(\frac{x-\bar{x}}{t}\right)\\
&\leq (1-\varepsilon_n)\cdot M\cdot\left|\frac{x-y_x^n}{t}\right| + \max_{|q|\leq\frac{2^{-n}}{t}}\left[M|q| + L(q)\right].
\end{aligned}$$

Thus, (3.2) implies that

$$L\left(\mathbf{b}_n(t,x)\right) - M\cdot|\mathbf{b}_n(t,x)| \;\leq\; \varepsilon_n\cdot M\cdot\left(\Lambda_M - |\mathbf{b}_n(t,x)|\right)$$

and this yields

$$\|\mathbf{b}_n\|_{\mathbf{L}^\infty(\mathbb{R}^d)} \;\leq\; \Lambda_M \qquad \text{for all } n \geq 1. \tag{3.4}$$

On the other hand, let $z_x \in \mathcal{Z}_n$ be the closet point to $y_x \in C_{t,x}$ such that $|z_x - y_x| \leq 2^{-n}$. Since $|x - y_x|/t \leq \Lambda_M$, we estimate

$$\begin{aligned}
|u_n(t,x) - u(t,x)| &= \min_{y\in\mathcal{Z}_n}\left\{(1-\varepsilon_n)\cdot u_0(y) + t\cdot L\left(\frac{x-y}{t}\right)\right\} - u_0(y_x) - t\cdot L\left(\frac{x-y_x}{t}\right)\\
&\leq (1-\varepsilon_n)\cdot u_0(z_x) - u_0(y_x) + t\cdot\left[L\left(\frac{x-z_x}{t}\right) - L\left(\frac{x-y_x}{t}\right)\right]\\
&\leq \left((1-\varepsilon_n)\cdot M + \sup_{|q|\leq\Lambda_M+\frac{2^{-n}}{t}}|DL(q)|\right)\cdot 2^{-n} + \varepsilon_n\cdot|u_0(y_x)|\\
&\leq \left(M + \sup_{|q|\leq\Lambda_M+\frac{2^{-n}}{t}}|DL(q)|\right)\cdot 2^{-n} + \left(|u_0(0)| + M|x| + M\Lambda_M t\right)\cdot\varepsilon_n.
\end{aligned}$$

In particular, $u_n(t,\cdot)$ converges uniformly to $u(t,\cdot)$ in any compact subset of $\mathbb{R}^d$.

2. Fixing $n \in \mathbb{Z}^+$, for any $i \neq j \geq 1$, the set

$$\mathcal{O}_{i,j}^t \;=\; \left\{x\in\mathbb{R}^d : (1-\varepsilon_n)u_0(y_i) + t\cdot L\left(\frac{x-y_i}{t}\right) < (1-\varepsilon_n)u_0(y_j) + t\cdot L\left(\frac{x-y_j}{t}\right)\right\}$$

is an open subset of $\mathbb{R}^d$ with C^1-boundary

$$\Gamma_{i,j}^t : \left\{x\in\mathbb{R}^d : (1-\varepsilon_n)u_0(y_i) + t\cdot L\left(\frac{x-y_i}{t}\right) = (1-\varepsilon_n)u_0(y_j) + t\cdot L\left(\frac{x-y_j}{t}\right)\right\}.$$

Set $\mathcal{V}_i^t := \bigcup_{j\neq i}\mathcal{O}_{i,j}^t$. From (3.1) and (3.3), it holds

$$\mathbf{b}_n(t,x) \;=\; \frac{x-y_i}{t} \qquad \text{for all } x \in \mathcal{V}_i^t. \tag{3.5}$$

In particular, $\mathbf{b}_n(t,\cdot)$ is in $BV_{loc}(\mathbb{R}^d)$ with

$$D\mathbf{b}_n(t,\cdot) \;=\; \frac{\mathbf{I}_d}{t}\cdot\mathcal{L}^d+\sum_{i=1}^{p(\Omega)}\mathbf{b}_n(t,\cdot)\otimes\nu_i\mathcal{H}^{d-1}_{\llcorner\Omega\bigcap\partial\mathcal{V}_i^t}$$

and

$$\operatorname{div}\mathbf{b}_n(t,\cdot) \;=\; \frac{d}{t}\cdot\mathcal{L}^d+\sum_{i=1}^{p(\Omega)}\langle\mathbf{b}_n(t,\cdot),\nu_i\rangle\mathcal{H}^{d-1}_{\llcorner\Omega\bigcap\partial\mathcal{V}_i^t}$$

where ν_i is the inner normal vector to $\mathcal{V}_i^t$.

Proposition 3.1 *For every $t>0$, it holds*

$$|D\mathbf{b}_n(t,\cdot)|(\Omega) \;\leq\; \frac{1}{\gamma_M}\cdot\left(\|\mathbf{b}_n(t,\cdot)\|_{\mathbf{L}^\infty(\mathbb{R}^d)}+\frac{\operatorname{diam}(\Omega)}{t}\right)\cdot\mathcal{H}^{d-1}(\partial\Omega)+\frac{\sqrt{d}}{t}\cdot|\Omega| \tag{3.6}$$

for every open and bounded set $\Omega\subset\mathbb{R}^d$ of finite perimeter.

Proof. 1. We first rewrite

$$D\mathbf{b}_n(t,\cdot) \;=\; \frac{\mathbf{I}_d}{t}\cdot\mathcal{H}^d+\frac{1}{2t}\cdot\sum_{i\neq j\in\{1,\cdots,p(\Omega)\}}(y_j-y_i)\otimes\nu_i\mathcal{H}^{d-1}_{\llcorner\Omega\bigcap\partial\mathcal{V}_i^t\bigcap\partial\mathcal{V}_i^t}$$

and

$$\operatorname{div}\mathbf{b}_n(t,\cdot) \;=\; \frac{d}{t}\cdot\mathcal{H}^d+\frac{1}{2t}\cdot\sum_{i\neq j\in\{1,\cdots,p(\Omega)\}}\langle y_j-y_i,\nu_i\rangle\,\mathcal{H}^{d-1}_{\llcorner\Omega\bigcap\partial\mathcal{V}_i^t\bigcap\partial\mathcal{V}_i^t} \tag{3.7}$$

with

$$\nu_i(x) \;=\; \frac{DL\left(\frac{x-y_j}{t}\right)-DL\left(\frac{x-y_i}{t}\right)}{\left|DL\left(\frac{x-y_j}{t}\right)-DL\left(\frac{x-y_i}{t}\right)\right|}\qquad \text{for } \mathcal{H}^{d-1}\ a.e.\ x\in\Omega\bigcap\partial\mathcal{V}_i^t\bigcap\partial\mathcal{V}_j^t.$$

In particular, this implies

$$\left|D\mathbf{b}_n(t,\cdot)-\frac{\mathbf{I}_d}{t}\right|(\Omega) \;\leq\; \frac{1}{2t}\sum_{i\neq j\in\{1,\cdots,p(\Omega)\}}|y_j-y_i|\cdot\mathcal{H}^{d-1}(\Omega\bigcap\partial\mathcal{V}_i^t\bigcap\partial\mathcal{V}_j^t). \tag{3.8}$$

For a fixed $x\in\Omega\bigcap\partial\mathcal{V}_i^t\bigcap\partial\mathcal{V}_j^t$, set $p_i:=DL\left(\frac{x-y_i}{t}\right)$ and $p_j:=DL\left(\frac{x-y_j}{t}\right)$, we have

$$\nu_i(x) \;=\; \frac{p_j-p_i}{|p_j-p_i|}\qquad\text{and}\qquad y_j-y_i \;=\; DH(p_i)-DH(p_j).$$

From (3.4), it holds that $|p_i|\leq\max_{|q|\leq\Lambda_M}|DL(q)|$. Thus, the assumption (1.3) and (1.5) yield

$$|y_i-y_j| \;\leq\; -\frac{1}{\gamma_M}\cdot\langle y_j-y_i,\nu_i(x)\rangle\,. \tag{3.9}$$

Recalling (3.7) and (3.8), we get

$$\left|D\mathbf{b}_n(t,\cdot)-\frac{\mathbf{I}_d}{t}\right|(\Omega) \ \leq \ \frac{1}{\gamma_M}\cdot\left|\operatorname{div}\mathbf{b}_n(t,\cdot)-\frac{d}{t}\right|(\Omega). \tag{3.10}$$

2. Let us now provide a bound on $\left|\operatorname{div}\mathbf{b}_n(t,\cdot)-\frac{d}{t}\right|(\Omega)$. Pick a point $x_0\in\Omega$, from (3.5), (3.7) and (3.9), the function $\mathbf{d}_n(t,x):=\frac{x-x_0}{t}-\mathbf{b}_n(t,x)$ is in $BV_{\text{loc}}(\mathbb{R}^d)$ and

$$\operatorname{div}\mathbf{d}_n(t,\cdot) \ = \ \frac{1}{2t}\cdot\sum_{i\neq j\in\{1,\cdots,p(\Omega)\}}\langle y_i-y_j,\nu_i\rangle\,\mathcal{H}^{d-1}_{\llcorner\Omega\bigcap\partial\nu_i^t\bigcap\partial\nu_i^t}$$

is a positive Radon measure. In particular, this implies that

$$|\operatorname{div}\mathbf{d}_n(t,\cdot)|(\Omega) \ = \ \int_\Omega\operatorname{div}\mathbf{d}_n(t,\cdot).$$

Let $\rho_\varepsilon\in\mathcal{C}_c^\infty(\mathbb{R}^d)$ be a family of modifiers, i.e., $\rho_\varepsilon(x)=\varepsilon^{-d}\rho\left(\frac{x}{\varepsilon}\right)$ for $\rho\in\mathcal{C}_c^\infty(\mathbb{R}^d)$ satisfying $\rho(x)\geq 0$, $\rho(x)=\rho(-x)$, $\operatorname{supp}(\rho)\subset B_d(0,1)$ and $\int_{\mathbb{R}^d}\rho(x)dx=1$. For every test function $\varphi_\varepsilon=\chi_\Omega*\rho_\varepsilon$, it holds that

$$\int_{\mathbb{R}^d}\varphi_\varepsilon\operatorname{div}\mathbf{d}_n(t,\cdot) \ = \ -\int_{\mathbb{R}^d}\mathbf{d}_n(t,x)\cdot\nabla\varphi_\varepsilon(x)dx \ \leq \ \|\mathbf{d}_n(t,\cdot)\|_{\mathbf{L}^\infty(\mathbb{R}^d)}\cdot\int_{\mathbb{R}^d}|\nabla\varphi_\varepsilon(x)|dx.$$

Thus, taking $\varepsilon\to 0+$, we get

$$\int_\Omega\operatorname{div}\mathbf{d}_n(t,\cdot) \ \leq \ \|\mathbf{d}_n(t,\cdot)\|_{\mathbf{L}^\infty(\mathbb{R}^d)}\cdot\mathcal{H}^{d-1}(\partial\Omega),$$

and (3.10) yields

$$|D\mathbf{b}_n(t,\cdot)|(\Omega) \ \leq \ \frac{1}{\gamma_M}\cdot\left(\|\mathbf{b}_n(t,\cdot)\|_{\mathbf{L}^\infty(\mathbb{R}^d)}+\frac{\operatorname{diam}(\Omega)}{t}\right)\cdot\mathcal{H}^{d-1}(\partial\Omega)+\frac{\sqrt{d}}{t}\cdot|\Omega|.$$

The proof is complete. □

Using Proposition 3.1, one can easily prove Theorem 1.2.

Proof of Theorem 1.2. We first claim that $\mathbf{b}_n(t,\cdot)$ converges to $\mathbf{b}(t,\cdot)$ in $\mathbf{L}^1_{\text{loc}}$. Since the sequence $\mathbf{b}_n(t,\cdot)$ is bounded in $[\mathbf{L}^\infty(\mathbb{R}^d)]^d$ and the set $\left(\bigcup_{n\geq 1}\Sigma_t^n\bigcup\Sigma_t\right)$ has zero Lebesgue measure, it is sufficient to show that

$$\lim_{n\to\infty}\mathbf{b}_n(t,x) \ = \ \mathbf{b}(t,x) \qquad \text{for all } x\in\mathbb{R}^d\backslash\left(\bigcup_{n\geq 1}\Sigma_t^n\bigcup\Sigma_t\right).$$

Assume by a contradiction that there exists a subsequence $\mathbf{b}_{n_k}(t,x)$ converging to some $w\neq\mathbf{b}(t,x)$. Since $u_n(t,\cdot)$ converges uniformly to $u(t,\cdot)$ in any compact subset of $\mathbb{R}^d$, we have

$$u(t,x) = \lim_{n_k\to\infty} u_{n_k}(t,x) = \lim_{n_k\to\infty} u_0(x - t\mathbf{b}_{n_k}(t,x)) + t\cdot L(\mathbf{b}_{n_k}(t,x))$$
$$= u_0(x-tw) + t\cdot L(w) = u_0(x-tw) + t\cdot L\left(\frac{x-(x-tw)}{t}\right).$$

Thus, $\mathbf{b}(t,x)$ is not singleton and this yields a contradiction. Finally, from (3.6) and [2, Proposition 3.13], the function $\mathbf{b}_n(t,\cdot)$ converges weakly to $\mathbf{b}(t,\cdot)$ in $BV(\Omega,\mathbb{R}^d)$. In particular, $\mathbf{b}(t,\cdot)$ has locally bounded variation and

$$|D\mathbf{b}(t,\cdot)|(\Omega) \leq \liminf_{n\to\infty} |D\mathbf{b}_n(t,\cdot)|(\Omega)$$
$$\leq \frac{1}{\gamma_M}\cdot\left(\liminf_{n\to+\infty}\|\mathbf{b}_n(t,\cdot)\|_{\mathbf{L}^\infty(\mathbb{R}^d)} + \frac{\mathrm{diam}(\Omega)}{t}\right)\cdot\mathcal{H}^{d-1}(\partial\Omega) + \frac{\sqrt{d}}{t}\cdot|\Omega|.$$

To conclude the proof, we recall (3.4), which yields (1.6). □

As a consequence of Theorem 1.2, the following holds:

Corollary 3.1 *Under the same assumption in Theorem 1.2, the set*

$$S_T\left(\mathcal{U}_{[m,M]}\right) = \left\{S_T(\bar{u}) : \bar{u}\in\mathcal{U}_{[m,M]}\right\}$$

with

$$\mathcal{U}_{[m,M]} := \left\{\bar{u}\in\mathbf{Lip}\left(\mathbb{R}^d\right) : |\bar{u}(0)|\leq m, \mathrm{Lip}[\bar{u}]\leq M\right\} \tag{3.11}$$

is compact in $\mathbf{W}^{1,1}_{\mathrm{loc}}(\mathbb{R}^d)$ *for every* $T>0$.

Proof. Let $(\bar{u}_n)_{n\geq 1}\subseteq\mathcal{U}_{[m,M]}$ be a sequence of initial data. Set

$$v_n(x) := S_T(\bar{u}_n)(x) \qquad \text{for all } x\in\mathbb{R}^d, n\geq 1,$$

we have

$$\|DH(Dv_n)\|_{\mathbf{L}^\infty} \leq \Lambda_M \qquad \text{and} \qquad |v_n(0)| \leq m + MT\Lambda_M.$$

Moreover, for any given $R>0$, Theorem 1.2 implies that

$$|D(DH(Dv_n))|(B(0,R)) \leq C_R \quad \text{for some constant } C_R>0.$$

Thanks to Helly's theorem, there exist a subsequence $(v_{n_k})_{k\geq 1}$ and $w\in BV_{\mathrm{loc}}(\mathbb{R}^d,\mathbb{R}^d)$ such that

- $v_{n_k}(0)$ converges to some $\bar{v}_0\in\mathbb{R}$;
- $DH(Dv_{n_k})$ converges to w point-wise and

$$\lim_{k\to\infty}\|DH(Dv_{n_k}) - w\|_{\mathbf{L}^1(B_d(0,R))} = 0.$$

This implies that $Dv_{n_k} = DL(DH(Dv_{n_k}))$ converges to $DL(w)$ point-wise and

$$\lim_{n_k\to+\infty} \|Dv_{n_k} - DL(w)\|_{\mathbf{L}^1(B_d(0,R))} \;=\; 0\,.$$

Thus, denoting by

$$\bar{v}^R_{n_k} \;:=\; \frac{1}{|B_d(0,R)|}\cdot\int_{B_d(0,R)} v_{n_k}(x)dx$$

the average value of v_{n_k} in $B_d(0,R)$, we have

$$\begin{aligned}\lim_{n_k\to\infty}\left(\bar{v}^R_{n_k} - v_{n_k}(0)\right) \;&=\; \lim_{n_k\to\infty}\frac{1}{|B_d(0,R)|}\cdot\int_0^1\int_{B_d(0,R)} Dv_{n_k}(sx)(x)dxds\\ &=\; \frac{1}{|B_d(0,R)|}\cdot\int_0^1\int_{B_d(0,R)} DL(w(sx))(x)dxds \;:=\; \bar{v}^R\end{aligned}$$

and this yields $\lim_{n_k\to\infty}\bar{v}^R_{n_k} = \bar{v}_0 + \bar{v}^R$. On the other hand, by the Poincaré inequality, it holds that

$$\left(\left\|\left(v_{n_k} - \bar{v}^R_{n_k}\right) - \left(v_{n'_k} - \bar{v}^R_{n'_k}\right)\right\|_{\mathbf{L}^1(B_d(0,R))}\right) \;\leq\; R\cdot\left\|Dv_{n_k} - Dv_{n'_k}\right\|_{\mathbf{L}^1(B_d(0,R))}.$$

Therefore, the sequence $(v_{n_k})_{k\geq1}$ is a Cauchy sequence in $\mathbf{W}^{1,1}(B_d(0,R))$ for every $R>0$ and thus converges to $\bar{v}$ in $\mathbf{W}^{1,1}_{\mathrm{loc}}(\mathbb{R}^d)$. □

4 Metric entropy in $\mathbf{W}^{1,1}$ for S_T

Assuming that H is in $\mathcal{C}^2(\mathbb{R}^d)$, we shall establish upper and lower estimates for the metric entropy of

$$S^R_T(\mathcal{U}_{[m,M]}) \;:=\; \left\{v_{\llcorner\square_R} \;:\; v\in S_T\left(\mathcal{U}_{[m,M]}\right)\right\}$$

in $\mathbf{W}^{1,1}(\square_R)$ for given constants $T,R,m,M>0$. In order to do so, let us introduce the following continuous functions:

- $\Psi_M:[0,M]\to[0,\infty[$ with $\Psi(0)=0$ and

$$\Psi_M(s) \;=\; s\cdot\min_{|p-q|\geq s,\,p,q\in\overline{B}_d(0,M)}\frac{|DH(p)-DH(q)|}{|p-q|}\quad\text{for all } s\in]0,M]; \tag{4.1}$$

- $\Phi_M:[0,M]\to[0,\infty[$ with $\Phi(0)=0$ and

$$\Phi_M(s) \;=\; s\cdot\min_{p\in\overline{B}_d\left(0,M-\frac{s}{2}\right)}\left(\max_{q\in\overline{B}_d(p,\frac{s}{2})}\left\|D^2H(q)\right\|_\infty\right)\quad\text{for all } 0<s\leq M \tag{4.2}$$

  with $\left\|D^2H(q)\right\|_\infty := \max_{|v|\leq1}\left|D^2H(q)(v)\right|$.

Notice that both maps $s \mapsto \Psi_M(s)$ and $s \mapsto \Phi_M(s)$ are strictly increasing, and the strictly convex property of H implies that

$$0 \;<\; \Psi_M(s) \;\leq\; \Phi_M(s) \;<\; M \cdot \max_{p \in \overline{B}_d(0,M)} \left\| D^2 H(p) \right\|_\infty \qquad \text{for all } s \in]0, M].$$

From Theorem 1.2, for any $v \in S_T^R(\mathcal{U}_{[m,M]})$, it holds that

$$|DH(Dv)|(\Box_R) \;\leq\; V_T \qquad \text{and} \qquad \|v\|_{\mathbf{L}^\infty(\mathbb{R}^d)} \;\leq\; m_T \tag{4.3}$$

with

$$\begin{cases} V_T & := \; \dfrac{d 2^d R^{d-1}}{\gamma_M} \cdot \left(\Lambda_M + \dfrac{2\sqrt{d}R}{T} \right) + \dfrac{\sqrt{d} 2^d R^d}{T} \\ m_T & := \; m + \sqrt{d} M R + T \cdot \sup\limits_{|q| \leq \Lambda_M} L(q). \end{cases}$$

Our main result is stated as follows.

Theorem 4.1 *Assume that $H \in \mathcal{C}^2(\mathbb{R}^d)$ is strictly convex, coercive, and satisfies (1.3). Then, for every $0 < \varepsilon < \min\left\{ R^+ \Psi_M^{-1}\left(\min\left\{ 2, 8V_T R, 8V_T \Lambda_M^{-1/d} \right\} \right), R^- \cdot \Phi_M^{-1}\left(\dfrac{\lambda_M}{2T} \right) \right\}$, it holds*

$$\begin{aligned} \log_2\left(\left\lfloor \frac{\beta^-}{\varepsilon} \right\rfloor \right) + \Gamma^- \cdot \left(\Phi_M\left(\frac{\varepsilon}{R^-} \right) \right)^{-d} \;&\leq\; \mathcal{H}_\varepsilon\left(S_T^R(\mathcal{U}_{[m,M]}) \;\middle|\; \mathbf{W}^{1,1}(\Box_R) \right) \\ &\leq\; \log_2\left(\frac{\beta^+}{\varepsilon} \right) + \Gamma^+ \cdot \left(\Psi_M\left(\frac{\varepsilon}{R^+} \right) \right)^{-d} \end{aligned} \tag{4.4}$$

where the constants $\beta^\pm, R^\pm$ and $\Gamma^\pm$ are explicitly computed to be

$$\beta^- \;=\; 2^d R^d m, \qquad \beta^+ \;=\; \left(2^{d+1} R^d + 2 \right)(3 + \sqrt{d}R) m_T, \qquad R^- \;=\; \frac{\omega_d \cdot R^d}{(d+1) 2^{d+9}}$$

$$R^+ \;=\; \left(2^d R^d + 1 \right)(3 + \sqrt{d}R), \quad \Gamma^- \;=\; \frac{1}{8 \ln 2} \cdot \left(\frac{8 R \lambda_M}{3T} \right)^d, \quad \Gamma^+ \;=\; 48d \cdot \left(24\sqrt{d} R V_T \right)^d.$$

Before proving Theorem 4.1 in the next two subsections, we present some cases where the estimates in (4.4) are sharp.

Remark 4.2 The following Hamiltonian

$$H(p) = |p|^{2k} \quad \text{for some postive integer } k \geq 2$$

is not uniformly convex but satisfies all assumptions in Theorem 1.2. Moreover, a direct computation yields

$$\alpha_1 s^{2k-1} \;\leq\; \Psi_M(s) \;\leq\; \Phi_M(s) \;\leq\; \alpha_2 s^{2k-1} \qquad \text{for all } s \in [0, M]$$

for some constants $\alpha_1, \alpha_2 > 0$ depending on k, and

$$\mathcal{H}_\varepsilon\left(S_T^R(\mathcal{U}_{[m,M]}) \;\middle|\; \mathbf{W}^{1,1}(\Box_R) \right) \;\approx\; \varepsilon^{-(2k-1)d}.$$

Remark 4.3 If $H \in \mathcal{C}^2(\mathbb{R}^d)$ is uniformly convex then

$$\alpha_1 s \ < \ \Psi_M(s) \ \leq \ \Phi_M(s) \ < \ \alpha_2 s \quad \text{for some } 0 < \alpha_1 < \alpha_2,$$

and (4.4) yields the same result in [3] that

$$\mathcal{H}_\varepsilon\left(S_T^R(\mathcal{U}_{[m,M]}) \mid \mathbf{W}^{1,1}(\Box_R)\right) \ \approx \ \varepsilon^{-d}.$$

Remark 4.4 For the one dimensional case ($d = 1$), every strictly convex $H \in \mathcal{C}^2(\mathbb{R})$ satisfies (1.3). In addition, assume that H has polynomial degeneracy, i.e., the set $I_H = \{\omega \in \mathbb{R} : H''(\omega) = 0\} \neq \varnothing$ is finite and for each $w \in I_H$, there exists a natural number $p_\omega \geq 2$ such that

$$H^{(p_\omega+1)}(\omega) \ \neq \ 0 \qquad \text{and} \qquad H^{(j)}(\omega) \ = \ 0 \quad \text{for all } j \in \overline{2,p}.$$

The polynomial degeneracy of H is defined by

$$\mathbf{p}_H \ := \ p_{\omega_H} \ = \ \max_{\omega \in I_H} p_\omega \quad \text{for some } \omega_H \in I_H.$$

For every $M > \omega_H$, there exist $0 < \alpha_1 < \alpha_2$ such that

$$\alpha_1 \cdot s^{\mathbf{p}_H} \ < \ \Psi_M(s) \ \leq \ \Phi_M(s) \ < \ \alpha_2 \cdot s^{\mathbf{p}_H}$$

and (4.4) implies that

$$\mathcal{H}_\varepsilon\left(S_T^R(\mathcal{U}_{[m,M]}) \mid \mathbf{W}^{1,1}(\Box_R)\right) \ \approx \ \varepsilon^{-\mathbf{p}_H}$$

4.1 Upper estimate

Towards the upper estimate of $\mathcal{H}_\varepsilon\left(S_T^R(\mathcal{U}_{[m,M]}) \mid \mathbf{W}^{1,1}(\Box_R)\right)$ in (4.4), we first provide a bound on the $\mathbf{L}^1$-distance between elements Du_1 and Du_2 in term of the $\mathbf{L}^1$-distance between $DH(Du_1)$ and $DH(Du_2)$ for every $u_1, u_2 \in S_T^R(\mathcal{U}_{[m,M]})$ by using the function Ψ defined in (4.1). Observing that the map $s \mapsto \dfrac{\Psi_M(s)}{s}$ is monotone increasing, and

$$\Psi_M(|p-q|) \ \leq \ |DH(p) - DH(q)| \quad \text{for all } p, q \in \overline{B}_d(0, M), \tag{4.5}$$

we prove the following lemma.

Lemma 4.5 *For any $u_1, u_2 \in S_T^R(\mathcal{U}_{[m,M]}$, it holds that*

$$\|Du_1 - Du_2\|_{\mathbf{L}^1(\Box_R)} \ \leq \ \left(2^d R^d + 1\right) \cdot \Psi_M^{-1}\left(\|\mathbf{b}_1 - \mathbf{b}_2\|_{\mathbf{L}^1(\Box_R)}\right) \tag{4.6}$$

with $\mathbf{b}_1 := DH(Du_1)$ and $\mathbf{b}_2 := DH(Du_2)$.

Proof. For simplicity, setting

$$\beta \ := \ \Psi_M^{-1}\left(\|\mathbf{b}_1 - \mathbf{b}_2\|_{\mathbf{L}^1(\Box_R)}\right),$$

we claim that

$$|Du_1(x) - Du_2(x)| \ \leq \ \beta \cdot \max\left\{1, \frac{|\mathbf{b}_1(x) - \mathbf{b}_2(x)|}{\|\mathbf{b}_1 - \mathbf{b}_2\|_{\mathbf{L}^1(\square_R)}}\right\} \qquad \text{for } a.e.\ x \in \square_R. \tag{4.7}$$

Indeed, assume that $|Du_1(x) - Du_2(x)| > \beta$. From (4.5), it holds

$$\begin{aligned}|Du_1(x) - Du_2(x)| \ &= \ \frac{|Du_1(x) - Du_2(x)|}{|DH(Du_1(x)) - DH(Du_2(x))|} \cdot |\mathbf{b}_1(x) - \mathbf{b}_2(x)| \\ &\leq \ \frac{|Du_1(x) - Du_2(x)|}{\Psi_M(|Du_1(x) - Du_2(x)|)} \cdot |\mathbf{b}_1(x) - \mathbf{b}_2(x)|.\end{aligned}$$

By the monotone increasing property of the map $s \longmapsto \dfrac{\Psi_M(s)}{s}$, one has

$$|Du_1(x) - Du_2(x)| \ \leq \ \frac{\beta}{|\Psi_M(\beta)|} \cdot |\mathbf{b}_1(x) - \mathbf{b}_2(x)| \ = \ \beta \cdot \frac{|\mathbf{b}_1(x) - \mathbf{b}_2(x)|}{\|\mathbf{b}_1 - \mathbf{b}_2\|_{\mathbf{L}^1(\square_R)}}.$$

and this implies (4.7). Therefore, the $\mathbf{L}^1$-distance between Du_1 and Du_2 is bounded by

$$\begin{aligned}\|Du_1 - Du_2\|_{\mathbf{L}^1(\square_R)} \ &= \ \int_{\square_d} |Du_1 - Du_2(x)| dx \\ &\leq \ \beta \cdot \int_{\square_d} 1 + \frac{|\mathbf{b}_1(x) - \mathbf{b}_2(x)|}{\|\mathbf{b}_1 - \mathbf{b}_2\|_{\mathbf{L}^1(\square_R)}} dx \ = \ (2^d R^d + 1)\beta \\ &= \ (2^d R^d + 1) \cdot \Psi_M^{-1}\left(\|\mathbf{b}_1 - \mathbf{b}_2\|_{\mathbf{L}^1(\square_R)}\right)\end{aligned} \tag{4.8}$$

and this yields (4.6). □

Proof of the upper estimate of $\mathcal{H}_\varepsilon\left(S_T^R(\mathcal{U}_{[m,M]}) \mid \mathbf{W}^{1,1}(\square_R)\right)$ in (4.4)

1. Recalling the second inequality in (4.3), one has

$$\bar{v}^R \ := \ \frac{1}{\mathrm{Vol}(\square_R)} \cdot \int_{\square_R} v(x)\ dx \ \in \ [-m_T, m_T] \qquad \text{for all } v \in S_T^R\left(\mathcal{U}_{[m,M]}\right).$$

For any $\varepsilon' > 0$, we cover $[-m_T, m_T]$ by $K_{\varepsilon'} = \left\lfloor \dfrac{m_T}{\Psi_M^{-1}(\varepsilon')} \right\rfloor + 1$ small intervals with length $2\Psi_M^{-1}(\varepsilon')$, i.e.,

$$[-m_T, m_T] \ \subseteq \ \bigcup_{i=1}^{K_{\varepsilon'}} B\big(a_i, \Psi_M^{-1}(\varepsilon')\big) \quad \text{for some } a_i \in [-m_T, m_T],$$

and then decompose the set $S_T^R(\mathcal{U}_{[m,M]})$ into $K_{\varepsilon'}$ subsets as follows:

$$S_T^R\left(\mathcal{U}_{[m,M]}\right) \ \subseteq \ \bigcup_{i=1}^{K_{\varepsilon'}} S_T^{R,i}(\mathcal{U}_{[m,M]})$$

with

$$S_T^{R,i}(\mathcal{U}_{[m,M]}) \ := \ \left\{v \in S_T^{R,i}\left(\mathcal{U}_{[m,M]}\right) : \bar{v}^R \in B\left(a_i, \Psi_M^{-1}(\varepsilon')\right)\right\}.$$

Thus, for all $\varepsilon > 0$, it holds

$$\mathcal{N}_\varepsilon\left(S_T^R\left(\mathcal{U}_{[m,M]}\right)\Big|\mathbf{W}^{1,1}(\Box_R)\right) \ \leq \ \sum_{i=1}^{K_{\varepsilon'}}\mathcal{N}_\varepsilon\left(S_T^{R,i}\left(\mathcal{U}_{[m,M]}\right)\Big|\mathbf{W}^{1,1}(\Box_R)\right). \tag{4.9}$$

2. Given $i \in \{1, 2, \dots, K_{\varepsilon'}\}$, we shall provide an upper bound on the covering number $\mathcal{N}_\varepsilon\left(S_T^{R,i}\left(\mathcal{U}_{[m,M]}\right)\Big|\mathbf{W}^{1,1}(\Box_R)\right)$ by introducing the following set

$$\mathcal{B}_T^{R,i}\left(\mathcal{U}_{[m,M]}\right) \ = \ \left\{DH(Dv) : v \in S_T^{R,i}(\mathcal{U}_{[m,M]})\right\}.$$

From (4.3) and (2.2), one has

$$\mathcal{B}_T^{R,i}\left(\mathcal{U}_{[m,M]}\right) \ \subseteq \ \mathcal{F}_{[R,\Lambda_M,V_T]},$$

and Corollary 2.6 yields

$$\mathcal{H}_{\varepsilon'/2}\left(\mathcal{B}_T^{R,i}\left(\mathcal{U}_{[m,M]}\right)\Big|\mathbf{L}^1\left(\Box_R\right)\right) \ \leq \ 48d\cdot\left(\frac{24\sqrt{d}RV_T}{\varepsilon'}\right)^d \ = \ \Gamma^+\cdot(\varepsilon')^{-d} \tag{4.10}$$

for all $0 < \varepsilon' < \min\left\{2, 8V_TR, 8V_T\Lambda_M^{-1/d}\right\}$ sufficiently small. Hence, there exists a set of initial data $\mathcal{U}_{[m,M]}^{\varepsilon'}$ with its image under the map S_T defined by

$$\mathbf{S}_{\varepsilon'}^i \ := \ \left\{\mathbf{v}_1, \dots, \mathbf{v}_{\beta_1^{\varepsilon'}}\right\} \ \subset \ S_T^{R,i}\left(\mathcal{U}_{[m,M]}\right) \qquad \text{with} \qquad \beta_1^{\varepsilon'} \ := \ \#\mathcal{U}_{[m,M]}^{\varepsilon'} \ \leq \ 2^{\Gamma^+\cdot(\varepsilon')^{-d}}$$

such that the following inclusion holds

$$\mathcal{B}_T^{R,i}\left(\mathcal{U}_{[m,M]}\right) \ \subseteq \ \bigcup_{j=1}^{\beta_1^{\varepsilon'}} B_{\mathbf{L}^1}(\mathbf{b}_j, \varepsilon') \qquad \text{with} \qquad \mathbf{b}_j \ := \ DH(D\mathbf{v}_j).$$

In particular, for any given $v \in S_T^R(\mathcal{U}_{[m,M]})$, it holds

$$\|DH(Dv) - \mathbf{b}_{j_0}\|_{\mathbf{L}^1(\Box_R)} \ < \ \varepsilon' \quad \text{for some } j_0 \in \overline{1, \beta_1^{\varepsilon'}}.$$

Recalling Lemma 4.5, we obtain

$$\begin{aligned}\|Dv - D\mathbf{v}_{j_0}\|_{\mathbf{L}^1(\Box_R)} \ &\leq \ \left(2^dR^d + 1\right)\cdot\Psi_M^{-1}\left(\|DH(Dv) - \mathbf{b}_{j_0}\|_{\mathbf{L}^1(\Box_R)}\right) \\ &\leq \ \left(2^dR^d + 1\right)\cdot\Psi_M^{-1}(\varepsilon'),\end{aligned}$$

and the Poincaré's inequality yields

$$\left\|\left(v - \bar{v}^R\right) - \left(\mathbf{v}_{j_0} - \bar{\mathbf{v}}_{j_0}^R\right)\right\|_{\mathbf{L}^1(\Box_R)} \ \leq \ \sqrt{d}R\cdot\|Dv - D\mathbf{v}_{j_0}\|_{\mathbf{L}^1(\Box_R)} \ \leq \ \sqrt{d}R\left(2^dR^d + 1\right)\cdot\Psi_M^{-1}(\varepsilon').$$

On the other hand, since $v, \mathbf{v}_{i_0} \in S_T^{R,i}(\mathcal{U}_{[m,M]})$, one has

$$\left|\bar{v}^R - \bar{\mathbf{v}}_{j_0}^R\right| \ \leq \ \left|\bar{v}^R - a_i\right| + \left|\bar{\mathbf{v}}_{j_0}^R - a_i\right| \ \leq \ 2\Psi_M^{-1}(\varepsilon').$$

The $\mathbf{W}^{1,1}$-distance between v and $\mathbf{v}_{j_0}$ can be estimated by

$$\begin{aligned}\|v - \mathbf{v}_{j_0}\|_{\mathbf{W}^{1,1}(\Box_R)} \ &\leq \ \|Dv - D\mathbf{v}_{j_0}\|_{\mathbf{L}^1(\Box_R)} + \left\|\left(v - \bar{v}^R\right) - \left(\mathbf{v}_{j_0} - \bar{\mathbf{v}}_{j_0}^R\right)\right\|_{\mathbf{L}^1(\Box_R)} \\ &+ \left|\bar{v}^R - \bar{\mathbf{v}}_{j_0}^R\right| \cdot \left|\Box_R\right| \ \leq \ \left(2^d R^d + 1\right)(3 + \sqrt{d}R) \cdot \Psi_M^{-1}(\varepsilon') \ = \ R^+ \cdot \Psi_M^{-1}(\varepsilon').\end{aligned}$$

Thus, for any given $0 < \varepsilon < R^+\Psi_M^{-1}\left(\min\left\{2, 8V_TR, 8V_T\Lambda_M^{-1/d}\right\}\right)$, if we choose $\varepsilon' = \Psi_M\left(\frac{\varepsilon}{R^+}\right)$ such that $0 < \varepsilon' < \min\left\{2, 8V_TR, 8V_T\Lambda_M^{-1/d}\right\}$ then the set $S_T^{R,i}\left(\mathcal{U}_{[m,M]}\right)$ is covered by $\beta_1^{\varepsilon'}$ open balls in $\mathbf{W}^{1,1}(\Box_R)$ centered at $\mathbf{v}_i$ of radius ε, i.e.

$$S_T^{R,i}(\mathcal{U}_{[m,M]}) \ \subseteq \ \bigcup_{i=1}^{\beta_1^{\varepsilon'}} B_{\mathbf{W}^{1,1}}(\mathbf{v}_i, \varepsilon),$$

and thus

$$\mathcal{N}_\varepsilon\left(S_T^{R,i}(\mathcal{U}_{[m,M]}\Big|\mathbf{W}^{1,1}(\Box_R)\right) \ \leq \ \beta_1^{\varepsilon'} \ = \ 2^{\Gamma^+ \cdot \left(\Psi_M\left(\frac{\varepsilon}{R^+}\right)\right)^{-d}}.$$

Finally, recalling (4.9), we get

$$\begin{aligned}\mathcal{N}_\varepsilon\left(S_T^R(\mathcal{U}_{[m,M]}\Big|\mathbf{W}^{1,1}(\Box_R)\right) \ &\leq \ \frac{\left(2^{d+1}R^d + 2\right)(3 + \sqrt{d}R)m_T}{\varepsilon} \cdot 2^{\Gamma^+ \cdot \left(\Psi_M\left(\frac{\varepsilon}{R^+}\right)\right)^{-d}} \\ &= \ \frac{\beta^+}{\varepsilon} \cdot 2^{\Gamma^+ \cdot \left(\Psi_M\left(\frac{\varepsilon}{R^+}\right)\right)^{-d}}\end{aligned}$$

and this yields the second inequality in (4.4). □

Remark 4.6 *To obtain the upper bound of* $\mathcal{H}_\varepsilon\left(S_T^R(\mathcal{U}_{[m,M]}) \;\middle|\; \mathbf{W}^{1,1}(\Box_R)\right)$ *in (4.4), we only require that* H *belongs to* $\mathcal{C}^1(\mathbb{R}^d)$.

4.2 Lower estimate

In this subsection, we will prove the first inequality in (4.4). In order to do so, for any given $p \in \mathbb{R}^d$, let $\Phi(\cdot, p) : [0, \infty[\to [0, \infty[$ be the strictly increasing continuous function defined by $\Phi(0, p) = 0$ and

$$\Phi(s, p) \ = \ s \cdot \left(\max_{p' \in \overline{B}_d(p, \frac{s}{2})} \left\|D^2H(p')\right\|_\infty\right) \qquad \text{for all } s > 0.$$

From the definition of Φ_M in (4.2), it holds that

$$\Phi_M(s) \ = \ \min_{p \in \overline{B}_d\left(0, M - \frac{s}{2}\right)} \Phi(s, p) \qquad \text{for all } s \in [0, M]. \tag{4.11}$$

The following proposition shows that a solution to (1.1) with a semiconvex initial condition preserves the semiconvexity on a given time interval, provided the semiconvexity constant of the initial data is sufficiently small in absolute value.

Proposition 4.1 *Given $T, M, r > 0$ and $\bar{p} \in \overline{B}_d\left(0, M - \frac{r}{2}\right)$, let $\bar{u}$ be a semiconvex function with semiconvexity constant $-K$ such that*

$$D^-\bar{u}(\mathbb{R}^d) \subseteq \overline{B}_d\left(\bar{p}, \frac{r}{2}\right) \qquad \text{and} \qquad K \leq \frac{\lambda_M}{4T} \cdot \frac{r}{\Phi(r,\bar{p})} \tag{4.12}$$

with λ_M defined in (1.3). Then, the map $(t,x) \mapsto S_t(\bar{u})(x)$ is a classical solution for $0 < t \leq T$ and

$$DS_t(\bar{u})(x) \in \overline{B}_d\left(\bar{p}, \frac{r}{2}\right) \qquad \textit{for all } (t,x) \in]0,T] \times \mathbb{R}^d.$$

Proof. For simplicity, we set

$$u(t,x) := S_t(\bar{u})(x) \qquad \text{for all } (t,x) \in [0,\infty[\times\mathbb{R}^d.$$

It is well-known from [9, Theorem 5.3.8] that $u(t,\cdot)$ is locally semiconcave for every $t > 0$. Thus, by Proposition 2.2, it is sufficient to show that $u(t,\cdot)$ is semiconvex with some semiconvexity constant $-C < 0$ for all $t \in [0,T]$, i.e., for any fixed $(t,x) \in [0,T[\times\mathbb{R}^d$, it holds that

$$u(t,x+h) + u(t,x-h) - 2u(t,x) \geq -C \cdot |h|^2 \quad \text{for all } h \in \mathbb{R}^d. \tag{4.13}$$

By the Lipschitz continuity of $u(t,\cdot)$, we can assume that $u(t,\cdot)$ is differentiable at $x \pm h$. In this case, $\mathbf{b}(t, x \pm h)$ reduce to a single value denoted by $\mathbf{b}^\pm = DH(\mathbf{p}^\pm)$ with $\mathbf{p}^\pm = Du(t, x \pm h)$ and satisfy the following relations

$$\begin{cases} \mathbf{p}^\pm \in D^-\bar{u}(x \pm h - t\mathbf{b}^\pm) \subseteq \overline{B}_d\left(\bar{p}, \frac{r}{2}\right) \subseteq \overline{B}_d(0,M), \\ \\ u(t, x \pm h) = \bar{u}(x \pm h - t\mathbf{b}^\pm) + t \cdot L(\mathbf{b}^\pm). \end{cases} \tag{4.14}$$

Since $\bar{u}$ is semiconvex with semiconvexity constant $-K$, denoting $x^\pm := x \pm h$, one has from Corollary 2.2 that

$$\left\langle \mathbf{p}^+ - \mathbf{p}^-, x^+ - x^- - t\left(\mathbf{b}^+ - \mathbf{b}^-\right) \right\rangle \geq -K \cdot \left|2h - t\left(\mathbf{b}^+ - \mathbf{b}^-\right)\right|^2$$

and

$$\begin{aligned} \left\langle \mathbf{p}^+ - \mathbf{p}^-, \mathbf{b}^+ - \mathbf{b}^- \right\rangle &\leq \frac{K}{t} \cdot \left|2h - t\left(\mathbf{b}^+ - \mathbf{b}^-\right)\right|^2 + \frac{2|h|}{t} \cdot |\mathbf{p}^+ - \mathbf{p}^-| \\ &\leq 2Kt|\mathbf{b}^+ - \mathbf{b}^-|^2 + \frac{8K|h|^2}{t} + \frac{2|h|}{t} \cdot |\mathbf{p}^+ - \mathbf{p}^-| \\ &\leq 2KT|DH(\mathbf{p}^+) - DH(\mathbf{p}^-)|^2 + \frac{8K|h|^2}{t} + \frac{2|h|}{t} \cdot |\mathbf{p}^+ - \mathbf{p}^-|. \end{aligned}$$

Since $\mathbf{p}^\pm \in \overline{B}_d(\bar{p}, \frac{r}{2})$, it holds

$$|DH(\mathbf{p}^+) - DH(\mathbf{p}^-)| \leq \frac{\Phi(r,\bar{p})}{r} \cdot |\mathbf{p}^+ - \mathbf{p}^-|.$$

Thus, recalling (4.12) and (4.14), we estimate

$$
\begin{aligned}
2KT|DH(\mathbf{p}^+) - DH(\mathbf{p}^-)|^2 \ &\leq\ 2KT \cdot \frac{\Phi(r,\bar{p})}{r} \cdot |DH(\mathbf{p}^+) - DH(\mathbf{p}^-)| \cdot |\mathbf{p}^+ - \mathbf{p}^-| \\
&\leq\ \frac{\lambda_M}{2} \cdot |DH(\mathbf{p}^+) - DH(\mathbf{p}^-)| \cdot |\mathbf{p}^+ - \mathbf{p}^-| \ =\ \frac{\lambda_M}{2} \cdot |\mathbf{b}^+ - \mathbf{b}^-| \cdot |\mathbf{p}^+ - \mathbf{p}^-|
\end{aligned}
$$

and

$$
\left\langle \mathbf{p}^+ - \mathbf{p}^-, \mathbf{b}^+ - \mathbf{b}^- \right\rangle \ \leq\ \frac{\lambda_M}{2} \cdot |\mathbf{b}^+ - \mathbf{b}^-| \cdot |\mathbf{p}^+ - \mathbf{p}^-| + \frac{8K|h|^2}{t} + \frac{2|h|}{t} \cdot |\mathbf{p}^+ - \mathbf{p}^-|. \tag{4.15}
$$

On the other hand, one deduces from (1.3) that

$$
\begin{aligned}
\left\langle \mathbf{p}^+ - \mathbf{p}^-, \mathbf{b}^+ - \mathbf{b}^- \right\rangle \ &=\ \left\langle \mathbf{p}^+ - \mathbf{p}^-, DH(\mathbf{p}^+) - DH(\mathbf{p}^-) \right\rangle \\
&\geq\ \lambda_M \cdot |DH(\mathbf{p}^+) - DH(\mathbf{p}^-)| \cdot |\mathbf{p}^+ - \mathbf{p}^-| \\
&=\ \lambda_M \cdot |\mathbf{b}^+ - \mathbf{b}^-| \cdot |\mathbf{p}^+ - \mathbf{p}^-|,
\end{aligned}
$$

and (4.15) yields

$$
\frac{\lambda_M}{2} \cdot |t(\mathbf{b}^+ - \mathbf{b}^-)| \cdot |\mathbf{p}^+ - \mathbf{p}^-| \ \leq\ 8K|h|^2 + 2|h| \cdot |\mathbf{p}^+ - \mathbf{p}^-|. \tag{4.16}
$$

Two cases are considered:

- If $|\mathbf{p}^+ - \mathbf{p}^-| \leq K|h|$ then

$$
\begin{aligned}
|\mathbf{b}^+ - \mathbf{b}^-| \ &=\ |DH(\mathbf{p}^+) - DH(\mathbf{p}^-)| \ \leq\ \frac{\Phi(r,\bar{p})}{r} \cdot |\mathbf{p}^+ - \mathbf{p}^-| \\
&\leq\ \frac{K\Phi(r,\bar{p})}{r} \cdot |h|.
\end{aligned}
$$

- Otherwise, (4.16) implies that

$$
\frac{\lambda_M}{2} \cdot |t(\mathbf{b}^+ - \mathbf{b}^-)| \ \leq\ 10|h|.
$$

Hence,

$$
|t(\mathbf{b}^+ - \mathbf{b}^-)| \ \leq\ \left(\frac{KT\Phi(r,\bar{p})}{r} + \frac{20}{\lambda_M} \right) \cdot |h|. \tag{4.17}
$$

By the Hopf-Lax representation formula, we have

$$
u(t, x \pm h) \ =\ \bar{u}(x \pm h - t\mathbf{b}^{\pm}) + t \cdot L(\mathbf{b}^{\pm})
$$

and

$$
u(t,x) \ \leq\ 2\bar{u}\left(x - t \cdot \frac{\mathbf{b}^+ + \mathbf{b}^-}{2} \right) + t \cdot L\left(\frac{\mathbf{b}^+ + \mathbf{b}^-}{2} \right).
$$

Using the convexity of L and semiconvexity of $\bar{u}$, we estimate

$$\begin{aligned}
u(t,x+h)+u(t,x-h)-2u(t,x) \ &\geq \ t\cdot\left[L(\mathbf{b}^+)+L(\mathbf{b}^-)-2L\left(\frac{\mathbf{b}^++\mathbf{b}^-}{2}\right)\right] \\
&+\bar{u}(x+h-t\mathbf{b}^+)+\bar{u}(x-h-t\mathbf{b}^-)-2\bar{u}\left(x-t\cdot\frac{\mathbf{b}^++\mathbf{b}^-}{2}\right) \\
&\geq \ -K\cdot\left|2h-t\left(\mathbf{b}^+-\mathbf{b}^-\right)\right|^2 \ \geq \ -8K|h|^2-2K|t(\mathbf{b}^+-\mathbf{b}^-)|^2 \\
&\geq \ -2K\cdot\left[4+\left(\frac{KT\Phi(r,\bar{p})}{r}+\frac{20}{\lambda_M}\right)^2\right]\cdot|h|^2
\end{aligned}$$

and this yields (4.13). □

Relying the above Proposition and Lemma 2.3, we prove the first inequality in (4.4).

Proof of the lower estimate of $\mathcal{H}_\varepsilon\left(S_T^R(\mathcal{U}_{[m,M]})\ \middle|\ \mathbf{W}^{1,1}(\square_R)\right)$ in (4.4)

1. For any given $r>0$ and $p\in\mathbb{R}^d$, we denote by

$$\mathcal{SC}_r^p \ := \ \left\{\varphi=v+\langle p,\cdot\rangle : v\in SC_{[\frac{r}{2},K_r]}\right\} \qquad \text{with} \qquad K_r \ = \ \frac{\lambda_M}{4T}\cdot\frac{r}{\Phi_M(r)}$$

where $SC_{[\frac{r}{2},K_r]}$ is defined in (2.1). From (4.11), there exists $p_r\in\overline{B}_d(0,M-\frac{r}{2})$ satisfying

$$\Phi(r,p_r) \ = \ \Phi_M(r) \ = \ \min_{p\in\overline{B}\left(0,M-\frac{r}{2}\right)}\Phi(r,p).$$

Consider the operator $\mathcal{T}:\mathcal{SC}_r^{p_r}\to\mathbf{Lip}\big(\mathbb{R}^d\big)$ defined by

$$\mathcal{T}(\varphi) \ = \ \varphi+S_T(\varphi_-)(0) \quad \text{with} \quad \varphi_-(x)=-\varphi(-x) \quad \text{for all } \varphi\in SC_r^{p_r}. \tag{4.18}$$

We show that

$$\mathcal{T}\left(\mathcal{SC}_r^{p_r}\right) \ \subseteq \ S_T\left(\mathcal{U}_{[0,M]}\right) \quad \text{with} \quad \mathcal{U}_{[0,M]} \text{ defined in (3.11).} \tag{4.19}$$

In order to find an initial data for a given function $\varphi\in\mathcal{SC}_r^{p_r}$, we only need to reverse the equation. Since

$$\varnothing \ \neq \ D^+\varphi(x) \ \subset \ \overline{B}_d\left(0,\frac{r}{2}\right) \qquad \text{for all } x\in\mathbb{R}^d,$$

the following function

$$w_0(\cdot) \ := \ -\mathcal{T}(\varphi)(-\cdot) \ = \ \varphi_-(\cdot)-S_T(\varphi_-)(0)$$

is semiconvex with a semiconvexity constant $-K_r$ and

$$D^-w_0(x) \ = \ p_r+D^+\varphi(-x) \ \subseteq \ \overline{B}_d\left(p_r,\frac{r}{2}\right) \qquad \text{for all } x\in\mathbb{R}^d.$$

Let $w(t,x) = S_t(w_0)(x)$ be the unique viscosity solution of (1.1) with initial datum w_0. Recalling Proposition 4.1 and property (ii) in Proposition 2.3, we have that w is a C^1 classical solution of (1.1) in $]0,T] \times \mathbb{R}^d$ and

$$D_x w(T,x) \subseteq \overline{B}_d\left(p_r, \frac{r}{2}\right) \subseteq \overline{B}_d(0,M) \qquad \text{for all } x \in \mathbb{R}^d.$$

Moreover, the translation property (iii) of the semigroup S_t (as defined by 2.10) implies that

$$w(T,0) = S_T(w_0)(0) = S_T\left(\varphi_- - S_T(\varphi_-)(0)\right)(0) = 0.$$

Thus, the continuous function $u : [0,T] \times \mathbb{R}^d \to \mathbb{R}$, defined by

$$u(t,x) = -w(T-t,-x) \qquad \text{for all } (t,x) \in [0,T] \times \mathbb{R}^d,$$

is also a C^1 classical solution of (1.1) in $]0,T[\times\mathbb{R}^d$ with

$$u(T,\cdot) \equiv \mathcal{T}(\varphi)(\cdot) \quad \text{and} \quad u(0,\cdot) = -w(T,-\cdot) \in \mathcal{U}_{[0,M]}.$$

In particular, $u(t,x)$ is a viscosity solution of (1.1) in $[0,T] \times \mathbb{R}^d$, so that, by the uniqueness property of the semigroup map S_t, one has

$$S_T(u_0)(\cdot) \equiv \mathcal{T}(\varphi)(\cdot), \qquad u_0(\cdot) \equiv -w(T,-\cdot)$$

and this implies $\mathcal{T}(\varphi)(\cdot) \in S_T\left(\mathcal{U}_{[0,M]}\right)$.

2. For every $\varepsilon > 0$, we select a finite subset $A_\varepsilon \subseteq [-m,m]$ such that

$$\#A_\varepsilon = \left\lfloor \frac{2^d R^d m}{\varepsilon} \right\rfloor \quad \text{and} \quad |a_i - a_j| \geq \frac{2\varepsilon}{2^d R^d} \quad \text{for all } a_i \neq a_j \in A_\varepsilon. \tag{4.20}$$

From the transition property (iii) of the semigroup S_t, it holds that

$$S_T\big(\mathcal{U}_{[m,M]}\big) \supseteq \bigcup_{a \in A_\varepsilon} S_T\big(a + \mathcal{U}_{[0,M]}\big) = A_\varepsilon + S_T\big(\mathcal{U}_{[0,M]}\big),$$

and (4.19) implies that

$$S_T\big(\mathcal{U}_{[m,M]}\big) \supseteq A_\varepsilon + \mathcal{T}\left(\mathcal{SC}_r^{p_r}\right). \tag{4.21}$$

By Lemma 2.3, there exists a subset $\mathcal{G}_r^R$ of $SC_r^{p_r}$ such that

$$\#\mathcal{G}_r^R \geq 2^{\beta_{[R,r]} \cdot \varepsilon^{-d}} \qquad \text{with} \qquad \beta_{[R,r]} = \frac{1}{3^d 2^{d^2+4d+3} \ln 2} \cdot \left(\frac{\omega_d R^{d+1} K_r}{(d+1)}\right)^d$$

and

$$\left\| D\varphi_{\llcorner \Box_R} - D\phi_{\llcorner \Box_R} \right\|_{\mathbf{L}^1(\Box_R)} \geq 2\varepsilon \qquad \text{for all } \varphi \neq \phi \in \mathcal{G}_r^R.$$

provided that

$$0 < \varepsilon \leq \min\left\{\frac{r}{2}, K_r\right\} \cdot \frac{\omega_d \cdot R^d}{(d+1)2^{d+8}}. \tag{4.22}$$

Since $D\mathcal{T}(\varphi)(x) = D\varphi(x)$ for all $x \in \mathbb{R}^d$, one gets

$$\left\| D\mathcal{T}(\varphi)_{\llcorner \Box_R} - D\mathcal{T}(\phi)_{\llcorner \Box_R} \right\|_{\mathbf{W}^{1,1}(\Box_R)} \geq 2\varepsilon \qquad \text{for all } \varphi \neq \phi \in \mathcal{G}_r^R.$$

Recalling (4.20), we have

$$\left\| f_{\llcorner \Box_R} - g_{\llcorner \Box_R} \right\|_{\mathbf{W}^{1,1}(\Box_R)} \geq 2\varepsilon \quad \text{for all } f \neq g \in A_\varepsilon + \mathcal{T}\left(\mathcal{SC}_r^{p_r}\right)$$

and

$$\mathcal{H}_\varepsilon\left(S_T^R(\mathcal{U}_{[m,M]}) \Big| \mathbf{W}^{1,1}(\Box_R)\right) \geq \log_2\left(\#A_\varepsilon \cdot \#\mathcal{G}_r^R\right) = \log_2\left(\left\lfloor \frac{2^d R^d m}{\varepsilon} \right\rfloor\right) + \frac{\beta_{[R,r]}}{\varepsilon^d}.$$

Finally, by choosing $r = \dfrac{\varepsilon}{R^-}$ with $R^- = \dfrac{\omega_d \cdot R^d}{(d+1)2^{d+9}}$, we compute

$$K_r = \frac{\lambda_M}{4TR^-} \cdot \frac{\varepsilon}{\Phi_M\left(\frac{\varepsilon}{R^-}\right)} \qquad \text{and} \qquad \beta_{[R,r]} = \frac{2^{4d-3}R^d}{3^d \ln 2} \cdot \left(\frac{\varepsilon}{\Phi_M\left(\frac{\varepsilon}{R^-}\right)}\right)^d.$$

Thus, for every $0 < \varepsilon \leq R^- \cdot \Phi_M^{-1}\left(\frac{\lambda_M}{2T}\right)$ such that (4.22) holds, we get

$$\mathcal{H}_\varepsilon\left(S_T^R(\mathcal{U}_{[0,M]}) \Big| \mathbf{W}^{1,1}(\Box_R)\right) \geq \frac{1}{8\ln 2} \cdot \left(\frac{8R\lambda_M}{3T}\right)^d \cdot \left(\Phi_M\left(\frac{\varepsilon}{R^-}\right)\right)^d + \log_2\left(\left\lfloor \frac{2^d R^d m}{\varepsilon} \right\rfloor\right)$$

and this yields the first inequality in (4.4). □

Remark 4.7 *With the same argument, the lower bound of $\mathcal{H}_\varepsilon\left(S_T^R(\mathcal{U}_{[m,M]}) \mid \mathbf{W}^{1,1}(\Box_R)\right)$ in (4.4) can be obtained for $H \in \mathcal{C}^{1,1}(\mathbb{R}^d)$ by defining*

$$\Phi_M(s) = s \cdot \inf_{p \in \overline{B}_d\left(0, M - \frac{s}{2}\right)} \left(\sup_{p_1 \neq p_2 \in \overline{B}_d(p, \frac{s}{2})} \frac{|DH(p_1) - DH(p_2)|}{|p_1 - p_2|} \right)$$

for all $s > 0$.

5 A counter-example

In this section, we show that if the strictly convex and coercive Hamiltonian $H \in \mathcal{C}^2(\mathbb{R}^2)$ does not satisfy the uniform directional convexity condition (1.3) then Theorem 1.2 fails in general. Indeed, let us consider the following Hamiltonian

$$H(p) = \frac{3^3}{4^4} \cdot p_1^4 + p_2^2 \qquad \text{for all } p = (p_1, p_2) \in \mathbb{R}^2.$$

The associated Lagrangian L of H is computed by

$$L(q) = |q_1|^{\frac{4}{3}} + q_2^2 \qquad \text{for all } q = (q_1, q_2) \in \mathbb{R}^2.$$

For any given $\bar{q} = (\bar{q}_1, \bar{q}_2) \in \mathbb{R}^2$, one has

$$\begin{aligned} L(q) = L(q - \bar{q}) \qquad &\Longleftrightarrow \qquad |q_1|^{4/3} + q_2^2 = |q_1 - \bar{q}_1|^{4/3} + (q_2 - \bar{q}_2)^2 \\ &\Longleftrightarrow \qquad q_2 = \frac{|q_1 - \bar{q}_1|^{4/3} - |q_1|^{4/3}}{2\bar{q}_2} + \frac{\bar{q}_2}{2}. \end{aligned}$$

Let $\gamma_{\bar{q}} : \mathbb{R} \to \mathbb{R}$ be such that

$$\gamma_{\bar{q}}(s) = \frac{|s - \bar{q}_1|^{4/3} - |s|^{4/3}}{2\bar{q}_2} + \frac{\bar{q}_2}{2} \qquad \text{for all } s \in \mathbb{R}.$$

In particular, assuming that $\bar{q}_2 = |\bar{q}_1|^{2/3}$ with $|\bar{q}_1| = \delta$ for some $\delta > 0$, it holds

$$\gamma_{\bar{q}}(0) = \bar{q}_2, \qquad \gamma_{\bar{q}}(\bar{q}_1) = 0,$$

and the following curve which connects two points $(0, \bar{q}_2)$ and $(\bar{q}_1, 0)$

$$\Gamma_{\bar{q}} = \begin{cases} \{(s, \gamma_{\bar{q}}(s)) : s \in [0, \delta]\} \subset [0, \delta] \times [0, \delta^{2/3}] & \text{if } \quad \bar{q}_1 = \delta > 0 \\ \\ \{(s, \gamma_{\bar{q}}(s)) : s \in [-\delta, 0]\} \subset [-\delta, 0] \times [0, \delta^{2/3}] & \text{if } \quad \bar{q}_1 = -\delta < 0 \end{cases} \tag{5.1}$$

has a length $> \delta^{2/3}$. From this observation, we shall construct a uniformly Lipschitz initial datum $\bar{u}$ such that both $Du(1, \cdot)$ and $\mathbf{b}(1, \cdot) = DH(Du(1, \cdot))$ does not have a locally bounded variation where $u = S_t(\bar{u})$ is the solution of (1.1) with $u(0, \cdot) \equiv \bar{u}$. Our construction is divided into two main steps.

1. Given $0 < \ell < 1$, we first construct a uniformly Lipschitz function $\bar{u} : \mathbb{R}^2 \to [0, \infty)$ with a Lipschitz constant which does not depend on ℓ such that

$$\text{supp}(\bar{u}) \subset [-2\ell, 2\ell] \quad \text{and} \quad |\mathbf{b}(1, \cdot)|([-\ell, \ell]^2), |Du(1, \cdot)|([-\ell, \ell]^2) \geq 1 \tag{5.2}$$

where

$$u(t, \cdot) = S_t(\bar{u}) \qquad \text{for all } t \geq 0.$$

For every $0 < \delta < \ell$, we consider the periodic lattice

$$y_\iota = \left(\iota_1 \delta, \iota_2 \delta^{2/3}\right) \quad \text{with} \quad \iota \in \mathcal{Z}_2 := \left\{(\iota_1', \iota_2') \in \mathbb{Z}^2 : \iota_1' + \iota_2' \in 2\mathbb{Z}\right\},$$

and the corresponding regions

$$\begin{aligned} \Omega_\iota &= \{x \in \mathbb{R}^2 : L(x - y_\iota) < L(x - y_{\iota'}) \quad \text{for all } \iota' \neq \iota\} \\ &= y_\iota + \left\{q \in \mathbb{R}^2 : L(q) < L(q + y_\iota - y_{\iota'}) \quad \text{for all } \iota' \neq \iota\right\} \\ &\subseteq y_\iota + [-\delta, \delta] \times [-\delta^{2/3}, \delta^{-2/3}] \end{aligned}$$

with

$$\partial\Omega_\iota \;=\; \left[y_\iota + (\Gamma_{\bar q^+}\cup\Gamma_{\bar q^-})\right]\bigcup\left[y_{\iota-(1,1)}+\Gamma_{\bar q^+}\right]\bigcup\left[y_{\iota-(-1,1)}+\Gamma_{\bar q^-}\right], \quad \bar q^{\pm} = (\pm\delta,\delta^{2/3}).$$

This correspond to the function

$$g_1(x) \;=\; L(x-y_\iota) \quad \text{for all } x\in\Omega_\iota, \iota\in\mathcal{Z}_2.$$

The dual solution is

$$g_0(y) \;=\; \max_{x\in\mathbb{R}^2}\{g_1(x)-L(x-y)\} \quad \text{for all } y\in\mathbb{R}^2$$

By the definition of Ω_ι, both g_0 and g_1 are Lipschitz with a Lipschitz constant

$$M_\delta \;=\; \sup_{q\in[-\delta,\delta]\times[-\delta^{2/3},\delta^{-2/3}]}|DL(q)| \;=\; O(\delta^{1/3}).$$

For $\delta>0$ sufficiently small, one can construct a Lipschitz initial data $\bar u$ with a Lipschitz constant M_δ and

$$\mathrm{supp}(\bar u)\;\subset\;[-2\ell,2\ell], \qquad \bar u(y)\;=\;g_0(y) \;\text{ for all } y\in\left[-\frac{3\ell}{2},\frac{3\ell}{2}\right]^2.$$

Let $u(t,\cdot)=S_t(\bar u)$ be the solution. At time $t=1$, we have

$$u(1,x) \;=\; \min_{y\in\mathbb{R}^2}\{\bar u(y)+L(x-y)\} \;=\; \bar u(y_x)+L(x-y_x)$$

for some $y_x\in\overline{B}(x,\Lambda_{M_\delta})$ with $\Lambda_{M_\delta}=\max\{|q|: L(q)\le M_\delta\cdot|q|\}=O(\delta^{1/3})$ being the maximal characteristic speed. Thus, if $M_\delta\le\frac{\ell}{2}$ then for all $x\in[-\ell,\ell]^2\cap\Omega_\iota$, $\iota\in\mathcal{Z}_2$,

$$\begin{aligned} u(1,x) &= \min_{y\in[-\frac{3\ell}{2},\frac{3\ell}{2}]^2}\{\bar u(y)+L(x-y)\} \;=\; \min_{y\in[-\frac{3\ell}{2},\frac{3\ell}{2}]^2}\{g_0(y)+L(x-y)\} \\ &= \min_{y\in\mathbb{R}^2}\{g_0(y)+L(x-y)\} \;=\; g_1(x) \;=\; L(x-y_\iota) \end{aligned}$$

and the slope of backward optimal rays through $(1,x)$ is

$$\mathbf{b}(1,x) \;=\; DH(Du(1,x)) \;=\; x-y_\iota.$$

For any two adjacent $y_\iota, y_{\iota'}$ with $\Omega_\iota,\Omega_{\iota'}\subset[-\ell,\ell]^2$ and $x\in\partial\Omega_\iota\cap\partial\Omega_{\iota'}$, we compute

$$\begin{cases} Du(1,x) &= [DL(x-y_\iota)-DL(x-y_{\iota'})]\otimes\mathbf{n}(x)\mathcal{H}^1_{\llcorner\partial\Omega_\iota\cap\partial\Omega_{\iota'}} \\ D\mathbf{b}(1,x)) &= (y_\iota-y_{\iota'})\otimes\mathbf{n}(x)\mathcal{H}^1_{\llcorner\partial\Omega_\iota\cap\partial\Omega_{\iota'}}. \end{cases}$$

and this implies

$$|D\mathbf{b}(1,\cdot)|(\Omega_\iota\cup\Omega_{\iota'}),|Du(1,\cdot)|(\Omega_\iota\cup\Omega_{\iota'}) \;\ge\; \delta^{2/3}\cdot\mathcal{H}^1(\partial\Omega_\iota\cap\partial\Omega_{\iota'}) \;\ge\; \delta^{4/3}.$$

Since the number of open regions $\Omega_\iota \subset [-\ell, \ell]^2$ is the order of $\frac{\ell^2}{\delta^{5/2}}$, we have

$$|D\mathbf{b}(1,\cdot)|([-\ell,\ell]^2), |D(u(1,\cdot))|([-\ell,\ell]^2) \;\geq\; C\cdot\frac{\ell^2}{\delta^{5/2}}\cdot\delta^{4/3} \;=\; C\cdot\frac{\ell^2}{\delta^{1/3}}$$

for some constant $C > 0$. Thus, choosing $\delta > 0$ sufficiently small, we obtain (5.2).

2. Let's consider a sequence of disjoint squares $\Box_n = c_n + [0, 2^{-n}] \times [0, 2^{-n}]$ such that

$$\bigcup_{n\geq 1} \Box_n \;\subset\; [0,1]^2 .$$

From the previous step, for any $n \geq 1$ one can construct a sequence of Lipschitz functions $\bar{u}_n : \mathbb{R}^2 \to [0, \infty)$ and $\mathrm{supp}(\bar{u}_n) \subset \Box_n$ such that the solution u_n of (1.1) with initial data $\bar{u}_n$ satisfies

$$|Du_n(1,\cdot)|\left(c_n + \frac{1}{2}\cdot(\Box_n - c_n)\right), |DH(Du_n(1,\cdot))|\left(c_n + \frac{1}{2}\cdot(\Box_n - c_n)\right) \;\geq\; 1,$$

and

$$L(x-z) \;\geq\; \min_{y\in\Box_n}\{\bar{u}_n(y) + L(x-y)\} \quad \text{for all } x \in \left(c_n + \frac{1}{2}\cdot(\Box_n - c_n)\right), z \in \mathbb{R}^2\backslash\Box_n .$$

Set $\bar{u} = \sum_{n=1}^{\infty} u_{0,n}$. The solution u of (1.1) with initial data $\bar{u}$ satisfies

$$u(1,x) \;=\; u_n(1,x) \quad \text{for all } x \in \left(c_n + \frac{1}{2}\cdot(\Box_n - c_n)\right),$$

and this implies

$$|D\mathbf{b}(1,\cdot)|([0,1]^2) \;\geq\; \sum_{n=1}^{\infty} |DH(Du_n(1,\cdot))|(\Box_n) \;\geq\; \sum_{n=1}^{\infty} 1 \;=\; +\infty.$$

Similarly, one has that $|Du(1,\cdot)|([0,1]^2) = +\infty$. □

Acknowledgments. This research by K.T.Nguyen was partially supported by a grant from the Simons Foundation/SFARI (521811, NTK).

References

[1] G. Acosta and R. C. Dúran, An optimal Poincaré inequality in $\mathbf{L}^1$ for convex domains, *Proc. Amer. Math. Soc. Vol* **132** (2003), no.1, p. 195-202.

[2] L. Ambrosio, N. Fusco and D. Pallara, Functions of Bounded Variation and Free Discontinuity Problems, *Oxford Science Publications*, Clarendon Press, Oxford, UK, (2000).

[3] F. Ancona, P. Cannarsa and Khai T. Nguyen, Quantitative compactness estimates for Hamilton-Jacobi equations, *Arch. Rat. Mech. Anal.*, **219** (2016), no. 2, 793–828.

[4] F. Ancona, P. Cannarsa and Khai T. Nguyen, The compactness estimates for Hamilton Jacobi Equations depending on space, *Bulletin of the Institute of Mathematics, Academia Sinica* **11** (2016), no. 1, 63–113.

[5] F. Ancona, O. Glass and K. T. Nguyen, Lower compactness estimates for scalar balance laws, *Comm. Pure Appl. Math* **65** (2012), no. 9, 1303–1329.

[6] F. Ancona, O. Glass and Khai T. Nguyen, On Kolmogorov entropy compactness estimates for scalar conservation laws without uniform convexity, *SIAM Journal on Mathematical Analysis*, **51** (2019), no. 4, 3020–3051.

[7] F. Ancona, O. Glass and Khai T. Nguyen, On lower compactness estimates for general nonlinear hyperbolic systems, *Ann. Inst. H. Poincare Anal. Non Lineaire*, **32** (2015), no. 6, 1229–1257.

[8] F. Ancona, O. Glass and K. T. Nguyen, On quantitative compactness estimates for hyperbolic conservation laws, *Proceedings of the 14th International Conference on Hyperbolic Problems (HYP2012)*, AIMS, Springfield, MO, 2014.

[9] P. Cannarsa and C. Sinestrari, Semiconcave functions, Hamilton-Jacobi equations, and optimal control, *Progress in Nonlinear Differential Equations and their Applications*, 58. Birkhäuser Boston, 2004.

[10] M.G. Crandall and P.-L. Lions, Viscosity solutions of Hamilton-Jacobi equations, *Trans. Amer. Math. Soc.* **277** (1983), no. 1, 1–42.

[11] P. Dutta, R. Capuani, and Khai T. Nguyen, Metric entropy for functions of bounded total generalized variation, *SIAM Journal on Mathematical Analysis*, to apprear. (https://arxiv.org/abs/1912.00219)

[12] P. Dutta and K. T. Nguyen, Covering numbers for bounded variation functions, *Journal of Mathematical Analysis and Applications* **468** (2018), no. 2, 1131–1143.

[13] C. De Lellis and F. Golse, A Quantitative Compactness Estimate for Scalar Conservation Laws, *Comm. Pure Appl. Math.* **58** (2005), no. 7, 989–998.

[14] A.N. Kolmogorov and V.M Tikhomirov, ε-Entropy and ε-capacity of sets in functional spaces, *Uspekhi Mat. Nauk* **14** (1959), 3-86.

[15] P.D. Lax, Course on hyperbolic systems of conservation laws, *XXVII Scuola Estiva di Fis. Mat.*, Ravello, 2002.