Uniqueness of the viscosity solution of a constrained Hamilton-Jacobi equation

Vincent Calvez

Institut Camille Jordan, UMR 5208 CNRS & Université Claude Bernard Lyon 1, France
vincent.calvez@math.cnrs.fr

King-Yeung Lam

Department of Mathematics, Columbus, OH, United States
lam.184@math.ohio-state.edu

Abstract

In quantitative genetics, viscosity solutions of Hamilton-Jacobi equations appear naturally in the asymptotic limit of selection-mutation models when the population variance vanishes. They have to be solved together with an unknown function $I(t)$ that arises as the counterpart of a non-negativity constraint on the solution at each time. Although the uniqueness of viscosity solutions is known for many variants of Hamilton-Jacobi equations, the uniqueness for this particular type of constrained problem was not resolved, except in a few particular cases. Here, we provide a general answer to the uniqueness problem, based on three main assumptions: convexity of the Hamiltonian function $H(I, x, p)$ with respect to p, monotonicity of H with respect to I, and BV regularity of $I(t)$.

1 Introduction

This note is intended to address uniqueness of the viscosity solution of the following Hamilton-Jacobi equation, under a non-negativity constraint:

$$\begin{cases} \partial_t u(t,x) + H(I(t), x, d_x u(t,x)) = 0, & t \in (0,T), \ x \in \mathbb{R}^d, \\ \min_{x \in \mathbb{R}^d} u(t,x) = 0, & t \in (0,T), \\ u(0,x) = g(x), & x \in \mathbb{R}^d, \end{cases} \qquad (1.1)$$

where $H : \mathbb{R} \times \mathbb{R}^d \times \mathbb{R}^d \to \mathbb{R}$ is of class $\mathcal{C}^2$, and the initial data $g \in W^{1,\infty}_{loc}(\mathbb{R}^d)$ satisfies $\min g = 0$. This problem arises naturally in the analysis of quantitative genetics model in the asymptotic regime of small variance [10, 19, 5, 20, 6, 17].

The main difficulty, beyond the classical issue of weak solutions in the viscosity sense, stems from the role played by the scalar quantity $I(t)$ which is subject to no

equation, but is attached to the non-negativity constraint $\min u(t, \cdot) = 0$. Moreover, its regularity in context is usually low, typically of bounded variation [19, 20, 6]. Interestingly, we shall see that BV seems to be the natural regularity ensuring uniqueness of the constrained problem. In order to give a sense to problem (1.1) in such a setting, we use the theory of viscosity solutions for equations with a measurable dependence in time which was first studied by H. Ishii [12] and then by P.L. Lions and B. Perthame [15].

1.1 Motivation and previous works

A special case of constrained Hamilton-Jacobi equation (1.1) arises in the asymptotic limit of the following quantitative genetics model proposed in [5, 20] (see also [19, 6, 17]):

$$\epsilon \partial_t n_\epsilon = n_\epsilon R(I_\epsilon(t), x) + \epsilon^2 \Delta n_\epsilon, \quad t > 0, \ x \in \mathbb{R}^d, \quad I_\epsilon(t) = \int_{\mathbb{R}^d} \psi(x) n_\epsilon(t, x) \, dx, \tag{1.2}$$

where $n_\epsilon(t, x)$ is the density of a population structured by a d-dimensional phenotypical trait x. The reproduction rate R of a given individual depends both on its trait x, and the environmental impact of the population I_ϵ. Individuals may burden differently, and that burden is weighted by the function ψ which is bounded below and above by positive constants. The key point is that I_ϵ is a scalar quantity, so that individuals compete for a single resource. As is natural for biological populations, the density-dependent feedback is negative, meaning that $R(I, x)$ is decreasing with respect to I.

The Hopf-Cole transformation $u_\epsilon = -\epsilon \log n_\epsilon$ transforms (1.2) into the following equation:

$$\partial_t u_\epsilon + R(I_\epsilon(t), x) + |d_x u_\epsilon|^2 = \epsilon \Delta u_\epsilon,$$

which yields formally (1.1) in the vanishing viscosity limit $\epsilon \to 0$, with $H(I, x, p) = R(I, x) + |p|^2$:

$$\begin{cases} \partial_t u + R(I(t), x) + |d_x u|^2 = 0, & t \in (0, T), \ x \in \mathbb{R}^d, \\ \min_{x \in \mathbb{R}^d} u(t, x) = 0, & t \in (0, T), \\ u(0, x) = g(x), & x \in \mathbb{R}^d. \end{cases} \tag{1.3}$$

In fact, locally uniform convergence to a viscosity solution was established under suitable assumptions on R and the initial data, but along subsequences $\epsilon_n \to 0$ [20, 6, 17]. Therein, compactness of $\{I_\epsilon\}$ usually follows from a uniform BV estimate. The constraint $\min u(t, \cdot) = 0$ can then be derived from natural properties of the integral $I_\epsilon = \int \psi \exp(-u_\epsilon/\epsilon) \, dx$ being uniformly positive and bounded in ϵ, as a consequence of the negative feedback of I_ϵ on growth. However, it is not possible to characterize the limit function I directly from the convergence of $u_{\epsilon_n} \to u$ in the limit $\epsilon_n \to 0$, because the relationship between u_ϵ and I_ϵ is too singular. Nonetheless, the constraint $\min u(t, \cdot) = 0$ must be satisfied at any time. Hence, $I(t)$ can be viewed as a Lagrange multiplier in problem (1.1), see [10] for a discussion about this analogy.

The uniqueness of the limiting problem (1.3), if available, guarantees the convergence of the whole family of solutions $\{u_\epsilon\}$ as $\epsilon \to 0$. It is interesting in the

mathematical as well as biological perspectives, since the limiting problem (1.3) determines much of the Darwinian evolutionary dynamics of the population model (1.2). When ϵ is small, the population n_ϵ concentrates at the point(s) where the limit function u reaches its minimum value 0.

The uniqueness was first treated in [20], for the particular case when $R(I, x)$ is separable in the following sense:

$$R(I, x) = B(x) - D(x)Q(I), \quad \text{or} \quad R(I, x) = B(x)Q(I) - D(x),$$

with positive functions B, D, and a monotonic function Q such that R is decreasing with respect to I.

Later on, the uniqueness for (1.3) was treated in [18] under convexity assumptions on $R(I, x)$ and the initial condition g, essentially: R decreasing with respect to I, and concave with respect to x, plus g convex. It was proved that convexity is propagated forward in time so that the solution $u(t, x)$ to (1.3) is a solution in the classical $\mathcal{C}^1$ sense. Hence, it has always a unique minimum point $\overline{x}(t)$, which is a smooth function of t. As a result, $I(t)$ is necessarily smooth in that setting, since it can be determined by the implicit relation $R(I(t), \overline{x}(t)) = 0$. Biologically interpreted, their results describe very well the Darwinian dynamics of a monomorphic population as it evolves smoothly towards a (global) evolutionary attractor. Recently, the preprint [14] tackled the uniqueness of (1.3) when the trait space is one-dimensional, with a mixture of separable and non-separable growth rate $R(I, x)$, without convexity assumptions. However, the uniqueness result is also restricted to the case of continuous functions $I(t)$, which is not guaranteed in the absence of convexity.

1.2 Assumptions and main result

In this paper, we establish uniqueness of solutions to problem (1.1) under mild conditions. In distinction with previous works, we assume neither (i) separability of the Hamiltonian $H(I, x, p)$ in any of its variables; nor (ii) convexity in the trait variable $x \in \mathbb{R}^d$. In particular, we can handle solutions (u, I), allowing possibly:

- $x \mapsto u(t, x)$ to possess multiple minimum points $\{\overline{x}_i(t)\}$; and

- the Lagrange multiplier $I(t)$ to be discontinuous.

Both of them are natural and attractive features of the solutions of population genetics models.

We restrict to $\mathcal{C}^2$ Hamiltonian functions $H(I, x, p)$ which are convex and super-linear with respect to the third variable p:

(H1): $\quad (\forall I, x, p) \quad d_{p,p}^2 H(I, x, p) > 0, \quad (\forall I, x) \quad \lim_{|p| \to +\infty} \frac{H(I, x, p)}{|p|} = +\infty.$

Our uniqueness result strongly relies on the following monotonicity assumption:

(H2): $\quad (\forall I, x, p) \quad \dfrac{\partial H}{\partial I}(I, x, p) < 0.$

As we are dealing with convex Hamiltonians, it is appropriate to reformulate the problem using suitable representation formulas: For a given function $I(t)$, we define the *variational solution* $V(t,x)$ of (1.1) as follows:

$$V(t,x) = \inf_{\left\{\begin{subarray}{l} \gamma \in AC(0,t) \\ \gamma(t)=x \end{subarray}\right\}} \left\{ \int_0^t L(I(s), \gamma(s), \dot\gamma(s))\, ds + g(\gamma(0)) \right\}, \qquad (1.4)$$

for $(t,x) \in (0,T) \times \mathbb{R}^d$. Here $AC(0,t)$ is the space of absolutely continuous functions on $(0,t)$, and $L : \mathbb{R} \times \mathbb{R}^d \times \mathbb{R}^d \to \mathbb{R}$ is the Lagrangian function, *i.e.* the Legendre transform (or convex conjugate) of H defined as:

$$L(I,x,v) = \sup_{p \in \mathbb{R}^d} \left\{ p \cdot v - H(I,x,p) \right\}. \qquad (1.5)$$

It is such that $d_p H$ and $d_v L$ are reciprocal functions. In this formulation, the problem (1.1) becomes the determination of $I(t)$ so that the value function $V(t,x)$ satisfies the constraint: $\min V(t,\cdot) = 0$. In the formulation (1.4), the role played by the scalar quantity $I(t)$ is perhaps more apparent: it is adjusted progressively so that the following constraint is satisfied at each time, among all $\gamma \in AC(0,t)$ irrespective of the endpoint of γ:

$$(\forall t) \quad \inf_{\gamma \in AC(0,t)} \left\{ \int_0^t L(I(s), \gamma(s), \dot\gamma(s))\, ds + g(\gamma(0)) \right\} = 0. \qquad (1.6)$$

Our methodology relies on the Lagrangian formulation of the constraint (1.6). Due to the above variational reformulation, it is more appropriate to write the assumptions on the Lagrangian function: Assumptions (H1)-(H2) can be recast as:

(L1): $\quad (\forall I,x,v) \quad d^2_{v,v} L(I,x,v) > 0$.

(L2): $\quad (\forall I,x,v) \quad \dfrac{\partial L}{\partial I}(I,x,v) > 0$.

We need two supplementary conditions, to be satisfied locally in $I \in (-J,J)$ for any constant $J > 0$:

(L3): There exists a positive constant C_Θ and a super-linear function $\Theta : \mathbb{R}_+ \to \mathbb{R}_+$ such that

$$(\forall I,x,v) \in (-J,J) \times \mathbb{R}^d \times \mathbb{R}^d \quad L(I,x,v) \geq \Theta(|v|) - C_\Theta, \quad \lim_{r \to +\infty} \frac{\Theta(r)}{r} = +\infty.$$

(L4): For each $K > 0$, there exists positive constants α_K, β_K such that

$$(\forall I,x,v) \in (-J,J) \times B(0,K) \times \mathbb{R}^d \quad |d_x L(I,x,v)| \leq \alpha_K + \beta_K L(I,x,v),$$

where $B(0,K) = \{x \in \mathbb{R}^d : |x| < K\}$.

Finally, we assume the initial data g to be locally Lipschitz continuous, minimized by 0 and coercive:

(G): $\quad g \in W^{1,\infty}_{loc}(\mathbb{R}^d), \quad \min_{x \in \mathbb{R}^d} g(x) = 0, \quad$ and $\quad \lim_{|x| \to \infty} g(x) = +\infty$.

Remark 1. *The super-linearity in* (L3) *holds true pointwise in* (I, x) *by the very definition of the Legendre transform* (1.5). *The main point here is the uniformity with respect to* x.

Theorem 1. *Assume* (L1)-(L4), *and let* I_1 *and* I_2 *be two BV functions. If the two associated variational solutions* V_1 *and* V_2 *of* (1.4), *equipped with the same initial data* g *satisfying* (G), *satisfy the constraints*

$$(\forall t, i) \quad \min_{x \in \mathbb{R}^d} V_i(t, x) = 0 \tag{1.7}$$

then I_1 *and* I_2 *coincide almost everywhere, and so do* V_1 *and* V_2.

To make the connection with viscosity solutions, we state the following auxiliary result:

Theorem 2. *Assume* (L1)-(L4). *For each* $I(t)$ *of bounded variation and g satisfying* (G), *the variational solution* $V(t, x)$ *with initial data g, given by* (1.4), *is the unique locally Lipschitz viscosity solution of* (1.1) *over* $[0, T] \times \mathbb{R}^d$.

Our result encompasses the framework of the convergence results in *e.g.* [20, Theorem 4.1] and [6, Theorem 1.2] and [17, Theorem 3.1]. The latter established convergence of subsequences (u_{ϵ_n}) towards locally Lipschitz viscosity solutions via uniform BV estimates on the sequences (I_ϵ). With our uniqueness theorem, it can be concluded that the whole sequence (u_ϵ) converges to the unique solution in each problem.

The reason for separating the two statements in Theorem 1 and Theorem 2 is to emphasize the use of the variational formulation in our proof. It would be of considerable interest to by-pass the variational formulation and derive uniqueness from PDE arguments only. Uniqueness of unbounded solutions generally requires stringent conditions on the growth of the solution and the Hamiltonian [4, 7], but here this issue is mediated by the fact that the Hamiltonian function is convex, and the solution is non-negative by definition. We could not find a reference containing precisely Theorem 2, but [8] is close, and we adapt their proof to our context in the Appendix.

1.3 Examples

First, we apply our result to the special case presented in Section 1.1.

Corollary 3. *Consider the problem* (1.3) *with the initial data* g *as in* (G). *Assume that* $R \in \mathcal{C}^2(\mathbb{R} \times \mathbb{R}^d)$ *satisfies*

$$(\forall I, x) \quad \frac{\partial R}{\partial I}(I, x) < 0 \quad and \quad (\forall I) \quad \sup_{x \in \mathbb{R}^d} R(I, x) < +\infty, \tag{1.8}$$

then the solution pair (u, I) *to the constrained Hamilton-Jacobi equation* (1.3) *is unique, in the class of locally Lipschitz viscosity solutions* u, *and BV functions* I.

The condition (1.8) is natural from the biological viewpoint as the net growth rate is (i) depends negatively on the population burden I, and (ii) is presumably bounded from above. Since the three other conditions (L1), (L2) and (L4) are

5

straightforward, it is sufficient to verify (L3). Indeed, the Lagrangian L is given by $\frac{1}{4}|v|^2 - R(I, x)$. Moreover, since the range of the BV function $I(t)$ lies in some bounded interval $[-J, J]$, it is sufficient to restrict R to $[-J, J] \times \mathbb{R}^d$ to address uniqueness over a bounded time interval $(0, T)$. Since it is immediate from (1.8) that R is uniformly bounded over $[-J, J] \times \mathbb{R}^d$, (L3) follows.

Our result also includes relevant examples that were not covered by the previous contributions, particularly non-separable Hamiltonian functions $H(I, x, p)$. For instance, consider the following quantitative genetics model:

$$\epsilon \partial_t n_\epsilon = \int \frac{1}{\epsilon^d} K\left(\frac{x - x'}{\epsilon}\right) B(I_\epsilon(t), x') n_\epsilon(t, x')\, dx' - n_\epsilon D(I_\epsilon(t), x),$$

where $I_\epsilon(t)$ is the same as in (1.2), and K is a probability distribution function that encodes the mutational effects after reproduction: if the parent has trait x', and gives birth at rate $B(I_\epsilon, x')$, the trait x of the offspring is distributed following $K_\epsilon(x - x') = \frac{1}{\epsilon^d} K\left(\frac{x - x'}{\epsilon}\right)$. Assume that K is symmetric, and has finite exponential moments, and denote by $\mathcal{K}$ its Laplace transform: $\mathcal{K}(p) = \int K(z) \exp(p \cdot z)\, dz$. Then, the limiting problem as $\epsilon \to 0$ is (1.1) with the following Hamiltonian function [6]:

$$H(I, x, p) = B(I, x)\mathcal{K}(p) - D(I, x)\,. \tag{1.9}$$

Corollary 4. *Consider the problem* (1.1) *with the Hamiltonian* (1.9), *and the initial data g as in* (G). *Assume that $\mathcal{K}$ is the Laplace transform of a symmetric p.d.f. with finite exponential moments, and that $B, D \in \mathcal{C}^2(\mathbb{R} \times \mathbb{R}^d)$ are non-negative functions which satisfy the following conditions:*

$$(\forall I, x) \quad B > 0 \quad and \quad (\forall J > 0) \quad \sup_{[-J \times J] \times \mathbb{R}^d} B(I, x) < \infty,$$

accompanied with the following monotonicity conditions:

$$(\forall I, x) \quad \frac{\partial B}{\partial I}(I, x) < 0, \ and \ \frac{\partial D}{\partial I}(I, x) > 0\,,$$

then the solution pair (u, I) to the constrained Hamilton-Jacobi equation (1.1) *is unique, in the class of locally Lipschitz viscosity solutions u, and BV functions I.*

Proof. There are a few items to check in order to apply Theorem 1. Firstly, the Hamiltonian function H (1.9) clearly verifies (H1) and (H2), hence (L1) and (L2) follows. Secondly, the Lagrangian function L associated with the Hamiltonian H (1.9) is

$$L(I, x, v) = B(I, x)\mathcal{L}\left(\frac{v}{B(I, x)}\right) + D(I, x)\,,$$

where $\mathcal{L}(v)$ is the Legendre transform of $\mathcal{K}(p)$. Let I_1, I_2 be the two functions that are involved in the uniqueness test. Define $B_\infty := \sup_{[-J, J] \times \mathbb{R}^d} B$, where $[-J, J]$ is a bounded interval in which both I_1 and I_2 take values. Condition (L3) is clearly verified with $\Theta(\cdot) = B_\infty + B_\infty \mathcal{L}\left(\frac{\cdot}{B_\infty}\right) \geq 0$, because L is decreasing with respect to the value B, $B_\infty < +\infty$ and $\mathcal{L}$ satifies (L3) automatically (see Remark 1). The justification of (L4) requires more work. We begin with the following inequality:

$$d_v \mathcal{L}(v) \cdot v \leq 2\left(1 + \mathcal{L}(v)\right)\,. \tag{1.10}$$

6

To derive it, consider the following pointwise inequality: for all $X \in \mathbb{R}$, $\cosh X \leq 1 + \frac{1}{2} X \sinh X$, which in turn implies the following one by symmetry of K:

$$\mathcal{K}(p) = \int K(z) \cosh(p \cdot z)\, dz$$

$$\leq 1 + \frac{1}{2} \int K(z)(p \cdot z) \sinh(p \cdot z)\, dz = 1 + \frac{1}{2} \int K(z)(p \cdot z) \exp(p \cdot z)\, dz \,.$$

By applying this estimate to $p = d_v \mathcal{L}(v) \Leftrightarrow v = d_p \mathcal{K}(p)$ in the dual Legendre transformation $\mathcal{L} \leftrightarrow \mathcal{K}$, we deduce, by the fact $\mathcal{K}(p) + \mathcal{L}(v) \geq p \cdot v$, that

$$d_v \mathcal{L}(v) \cdot v - \mathcal{L}(v) \leq \mathcal{K}(p) \leq 1 + \frac{1}{2} d_v \mathcal{L}(v) \cdot v \,.$$

This yields the simple estimate announced in (1.10).

The technical condition (L4) is reformulated in this context as follows, after division by $B > 0$:

$$\left| d_x(\log B) \mathcal{L}\left(\frac{v}{B}\right) - d_x(\log B)\left(d_v \mathcal{L}\left(\frac{v}{B}\right) \cdot \frac{v}{B}\right) + \frac{d_x D}{B} \right|$$

$$\leq \alpha_K \frac{1}{B} + \beta_K \left(\mathcal{L}\left(\frac{v}{B}\right) + \frac{D}{B} \right) \,.$$

It is indeed guaranteed for a suitable choices of α_K, β_K. The main arguments besides (1.10) are: both $|d_x(\log B)|$ and $|d_x D/B|$ are locally uniformly bounded from above, $1/B$ is locally uniformly bounded from below, and D is non-negative. $\qquad \square$

1.4 The Pessimization Principle: $I(t)$ is non-decreasing

The pessimization principle [9] is a concept in adaptive dynamics, which says that if the environmental feedback is encoded by a scalar quantity $I(t) \in \mathbb{R}$ at any time, mutations and natural selection inevitably lead to deterioration/Verelendung. In the setting of this paper, it can be formulated by claiming that the population burden $I(t)$ is a non-decreasing function.

In this section, we give an additional assumption that guarantees this claim.

Theorem 5. *Under the assumptions* (L1) - (L4), *let* (u, I) *be the unique solution pair to* (1.1) *such that* u *is a locally Lipschitz viscosity solution, and* I *is* BV. *Assume, in addition, that*

(L5): $(\forall I, x, v) \quad L(I, x, v) \geq L(I, x, 0)$.

Then I *is non-decreasing with respect to time.*

Note that (L5) is equivalent to $d_v L(I, x, 0) = 0$ by convexity, and thus to $d_p H(I, x, 0) = 0$ by duality. It is clearly verified for the two examples in Section 1.3 by symmetry of the Hamiltonian function with respect to $p = 0$.

Corollary 6. *Under the assumption* (1.8), *let* (u, I) *be the unique solution pair to* (1.3), *with the initial data* g *as in* (G). *Then* I *is non-decreasing with respect to time.*

Corollary 7. *Under the assumptions of Corollary 4, let* (u, I) *be the unique solution pair to the problem* (1.1) *with the Hamiltonian* (1.9), *and the initial data* g *as in* (G). *Then* I *is non-decreasing with respect to time.*

1.5 Organization

In Section 2, we quantify the BV regularity of minimizing curves. The proof of the main result (Theorem 1) is given in Section 3. The pessimization principle (Theorem 5) is proved in Section 4. Finally, the version of variational principle we need (Theorem 2) is proved in Appendix A.

2 Regularity of the minimizing curves

Let $(t, x) \in (0, T) \times \mathbb{R}^d$, and let $\gamma^{t,x}$ be a minimizing curve in (1.4) for some given $I(t)$. We take the existence of minimizing curves for granted. It follows from standard methods in Lagrangian dynamics, see *e.g.* [11], and the *a priori* compactness estimates obtained in Lemma 9 below. The purpose of this section is to establish BV regularity of the derivative $\dot{\gamma}^{t,x}$, provided that $I(t)$ is BV itself. Such regularity is crucial in our proof of uniqueness.

First, we establish the following consequence of the convexity and the superlinearity of the Lagrangian:

Lemma 8. *Assume* (L1) *and* (L3), *then*

$$(\forall I, x) \quad \lim_{|v| \to +\infty} \frac{d_v L(I, x, v) \cdot v}{|v|} = +\infty \,,$$

and the limit is uniform over (I, x) *lying in compact subsets of* $\mathbb{R}^{d+1}$.

Proof. Fix $K > 0$ and let $(I, x) \in [-K, K] \times B(0, K)$. Let $M > 0$ be given, and choose $r_0 = r_0(K) > 0$ large enough so that for all $r \geq r_0$,

$$\Theta(r) - C_\Theta - L(I, x, 0) \geq Mr \,.$$

Then we have, for $r \geq r_0$ and $e \in \mathbb{S}^{d-1}$,

$$L(I, x, re) - L(I, x, 0) \geq Mr \,.$$

By convexity of L in v, we have, for all $e \in \mathbb{S}^{d-1}$ and $r \geq r_0$,

$$d_v L(I, x, re) \cdot e = \frac{d}{dr} \big(L(I, x, re) \big) \geq \frac{L(I, x, re) - L(I, x, 0)}{r} \geq M \,.$$

Therefore, for each M there exists r_0 such that $d_v L(I, x, v) \cdot v \geq M|v|$, provided that $|v| \geq r_0$ and $(I, x) \in [-K, K] \times B(0, K)$. $\qquad\square$

We will now establish the BV estimate of $\dot{\gamma}^{t,x}$ for $(t, x) \in (0, T) \times \mathbb{R}^d$. For the remainder of this section, we fix $T > 0$ and J so that $\sup_{(0,T)} |I| \leq J$. For ease of notation, dependence of various constants on T and J will be suppressed.

Lemma 9. *Assume* (L1), (L3) *and* (L4), *then* *for each K, there exists a constant C_K such that, uniformly for $(t,x) \in (0,T) \times B(0,K)$, any minimizing curve $\gamma^{t,x} \in AC(0,t)$ associated with* (1.4) *satisfies*

$$\|\gamma^{t,x}\|_{W^{1,\infty}(0,t)} = \|\gamma^{t,x}\|_{L^\infty(0,t)} + \|\dot{\gamma}^{t,x}\|_{L^\infty(0,t)} \leq C_K, \tag{2.1}$$

and

$$[\dot{\gamma}^{t,x}]_{BV(0,t)} \leq C_K \left(t + [I]_{BV(0,t)} \right) . \tag{2.2}$$

Proof. The proof is divided into three steps, wherein classical arguments are recalled for the sake of completeness. For the sake of notation, we drop the superscript of $\gamma^{t,x}$, assuming that the pair $(t,x) \in (0,T) \times B(0,K)$ is fixed throughout the proof.
Step #1: L^∞ bound on γ. We deduce immediately the following bound from (L3):

$$\int_0^t \Theta(|\dot{\gamma}(s)|) \, ds - Ct + g(\gamma(0)) \leq \int_0^t L(I(s), \gamma(s), \dot{\gamma}(s)) \, ds + g(\gamma(0))$$
$$\leq \int_0^t L(I(s), x, 0) \, ds + g(x) , \tag{2.3}$$

from which we deduce a non-optimal $W^{1,1}$ estimate

$$\int_0^t |\dot{\gamma}(s)| \, ds \leq Ct + \int_0^t L(I(s), x, 0) \, ds + g(x) - g(\gamma(0)) .$$

by using the super-linearity of Θ in a crude way, namely, $\Theta(|v|) \geq |v| - C$. Furthermore, we deduce from $g \geq 0$ that $\dot{\gamma}$ belongs to $L^1(0,t)$. Consequently, there exists a constant C_K' such that, uniformly for $(t,x) \in (0,T) \times B(0,K)$, we have

$$\|\gamma(s)\|_{L^\infty(0,t)} \leq C_K' .$$

Let $A = A(g, K)$ be the (local) Lipschitz bound on g in the ball with radius C_K'. By updating the constant C, we can assume that $L(I, x, v) \geq (A+1)|v| - C$. Back to (2.3), we deduce that

$$(A+1) \int_0^t |\dot{\gamma}(s)| \, ds \leq Ct + \int_0^t L(I(s), x, 0) \, ds + A \int_0^t |\dot{\gamma}(s)| \, ds ,$$

We obtain as a consequence the following updated estimate:

$$\frac{1}{t} \int_0^t |\dot{\gamma}(s)| \, ds \leq C + \frac{1}{t} \int_0^t L(I(s), x, 0) \, ds \leq C' , \tag{2.4}$$

where the bound C' is uniform for $(t,x) \in (0,T) \times B(0,K)$, and $I(s)$ taking values in $[-J, J]$.
Step #2: L^∞ bound on $\dot{\gamma}$. We deduce from (2.4) that there exists a subset $\mathcal{S}_{t,x} \subset (0,t)$ of positive measure, such that for all $\hat{s} \in \mathcal{S}_{t,x}$, $|\dot{\gamma}(\hat{s})| \leq 2C'$.
Since γ is a minimizing curve, it satisfies the following Euler-Lagrange condition in the distributional sense:

$$-\frac{d}{ds} \left(d_v L(I(s), \gamma(s), \dot{\gamma}(s)) \right) + d_x L(I(s), \gamma(s), \dot{\gamma}(s)) = 0 \quad \text{in } (0,t). \tag{2.5}$$

9

Let Leb(f) denote the set of Lebesgue points of the function f. Let s_1, $s_2 \in (0, t)$ belong to Leb(I) $\cap$ Leb($\dot{\gamma}$). Let $\{\rho_n\}_{n \in \mathbb{N}}$ be a family of mollifiers so that $\rho_n(\cdot - s_i)$ is compactly supported in $(0, t)$ for $i = 1, 2$ and $n \in \mathbb{N}$. We can test (2.5) against $\theta(s) := \int_0^s (\rho_n(s' - s_2) - \rho_n(s' - s_1)) \, ds'$. After integration by parts, we find that:

$$\left| \int_0^t d_v L(I(s), \gamma(s), \dot{\gamma}(s)) \left(\rho_n(s - s_2) - \rho_n(s - s_1) \right) \, ds \right| \leq \int_0^t |d_x L(I(s), \gamma(s), \dot{\gamma}(s))| \, ds, \tag{2.6}$$

where we have used that the test function θ satisfies $|\theta(s)| \leq 1$ for $s \in (0, t)$, and that $\theta(0) = \theta(t) = 0$ holds. Using the definition of the Lebesgue points, we find that

$$\lim_{n \to +\infty} \int_0^t d_v L(I(s), \gamma(s), \dot{\gamma}(s)) \rho_n(s - s_i) \, ds = d_v L(I(s_i), \gamma(s_i), \dot{\gamma}(s_i)). \tag{2.7}$$

Then, we can specialize $s_1 \in$ Leb(I) $\cap$ Leb($\dot{\gamma}$) $\cap \mathcal{S}_{t,x}$ because the latter has positive measure. We deduce from (2.6)–(2.7), and the definition of $\mathcal{S}_{t,x}$ that

$$|d_v L(I(s_2), \gamma(s_2), \dot{\gamma}(s_2))| \leq C + \int_0^t \left| d_x L(I(s'), \gamma(s'), \dot{\gamma}(s')) \right| \, ds',$$

for all $s_2 \in$ Leb(I) $\cap$ Leb($\dot{\gamma}$). Multiplying by the unit vector $\frac{\dot{\gamma}(s_2)}{|\dot{\gamma}(s_2)|}$, and using (L4), we find:

$$\frac{d_v L(I(s_2), \gamma(s_2), \dot{\gamma}(s_2)) \cdot \dot{\gamma}(s_2)}{|\dot{\gamma}(s_2)|} \leq C + \int_0^t \left| d_x L(I(s'), \gamma(s'), \dot{\gamma}(s')) \right| \, ds'$$

$$\leq C + \int_0^t \left[\alpha_{K'} + \beta_{K'} L(I(s'), \gamma(s'), \dot{\gamma}(s')) \right] \, ds',$$

where $K' = \|\gamma\|_{L^\infty(0,t)}$. We deduce from the uniform bound of $\|\gamma\|_{L^\infty(0,t)}$, and the minimizing property of γ that the right-hand-side is uniformly bounded for $(t, x, I) \in (0, T) \times B(0, K) \times [-J, J]$. Hence,

$$\frac{d_v L(I(s), \gamma(s), \dot{\gamma}(s)) \cdot \dot{\gamma}(s)}{|\dot{\gamma}(s)|} \leq C \quad \text{a.e. } s \in (0, t),$$

and the boundedness of $\dot{\gamma}$ is a consequence of Lemma 8. This proves (2.1).

Step #3: *BV bound on $\dot{\gamma}$.* Back to (2.5), we see that

$$p(s) = d_v L(I(s), \gamma(s), \dot{\gamma}(s))$$

is Lipschitz continuous as $I, \gamma, \dot{\gamma} \in L^\infty$, and L is $\mathcal{C}^2$. By the Fenchel-Legendre duality, we have

$$\dot{\gamma}(s) = d_p H(I(s), \gamma(s), p(s)).$$

Therefore, $\dot{\gamma}$ is BV.

From the chain rule involving BV functions (see [1] and references therein, as well as [2, Theorem 3.96]), we deduce as a by-product that

$$[\dot{\gamma}]_{BV(0,t)} \leq [d_p H]_{\text{Lip}([-J,J] \times B(0, \|\gamma\|_\infty) \times B(0, \|\dot{\gamma}\|_\infty))} \left([p]_{BV(0,t)} + [\gamma]_{BV(0,t)} + [I]_{BV(0,t)} \right).$$

Both p and γ are Lipschitz continuous, so there exists a positive constant C such that $\max([p]_{BV(0,t)}, [\gamma]_{BV(0,t)}) \leq Ct$. Hence, we have obtained (2.2). $\qquad \square$

3 Uniqueness of the variational solution (proof of Theorem 1)

Let $I_1, I_2 \in BV(0,T)$. There exists a constant J such that I_1 and I_2 take value in $[-J, J]$. Recall the definition of the variational solutions V_1, V_2 (1.4):

$$
\begin{cases}
V_1(t,x) = \displaystyle\inf_{\gamma_1 \in AC(0,t): \gamma_1(t)=x} \left\{ \int_0^t L(I_1(s), \gamma_1(s), \dot\gamma_1(s))\, ds + g(\gamma_1(0)) \right\}, \\[2mm]
V_2(t,x) = \displaystyle\inf_{\gamma_2 \in AC(0,t): \gamma_2(t)=x} \left\{ \int_0^t L(I_2(s), \gamma_2(s), \dot\gamma_2(s))\, ds + g(\gamma_2(0)) \right\}.
\end{cases}
$$

Lemma 10. *There exists $C > 0$ such that for $j = 1, 2$,*

$$
V_j(t, x) \geq \min\left\{ \frac{|x|}{2}, \min_{|x'| \geq |x|/2} g(x') \right\} - Ct.
$$

Proof. By definition of the variational solution and (L3), we have

$$
\begin{aligned}
V_j(t, x) &= \int_0^t L\left(I_j(s), \gamma_j^{t,x}(s), \dot\gamma_j^{t,x}(s) \right) ds + g\left(\gamma_j^{t,x}(0) \right) \\
&\geq \int_0^t \left(|\dot\gamma_j^{t,x}(s)| - C \right) ds + g\left(\gamma_j^{t,x}(0) \right) \\
&\geq \left| x - \gamma_j^{t,x}(0) \right| - Ct + g\left(\gamma_j^{t,x}(0) \right) \\
&\geq \min\left\{ |x|/2, \min_{|x'| \geq |x|/2} g(x') \right\} - Ct
\end{aligned}
$$

where we used the fact that $\gamma^{t,x}(t) = x$ and that $|x| \leq |x - \gamma^{t,x}(0)| + |\gamma^{t,x}(0)|$, so that: either $|x - \gamma^{t,x}(0)| \geq |x|/2$ or $|\gamma^{t,x}(0)| \geq |x|/2$. $\qquad\square$

On the way to proving Theorem 1, we seek a pair of uniformly positive weights ϕ_j, $j = 1, 2$, such that $\mu = I_2 - I_1$ verifies both $\int_0^t \phi_2 \mu\, ds \geq 0$ and $\int_0^t \phi_1 \mu\, ds \leq 0$.

By Lemma 10 and (G), we deduce that $x \mapsto V_j(t, x)$ attains minimum in some bounded set, say $B(0, K)$, uniformly for $t \in (0, T)$.

Let x_1^t (resp. x_2^t) be some minimum point for $V_1(t, \cdot)$ (resp. $V_2(t, \cdot)$) – this might not be unique – and let $\gamma_1^t(s)$ (resp. $\gamma_2^t(s)$) be an optimal trajectory ending up at x_1^t (resp. x_2^t). We deduce from Lemma 9 that $\dot\gamma_j^t$ lies in BV, uniformly with respect to $t \in (0, T)$:

$$
\max_{j=1,2} \left[\dot\gamma_j^t \right]_{BV(0,t)} \leq C. \tag{3.1}
$$

Using the formulation (1.6) for I_1 with the curve γ_2^t, then the optimality of γ_2^t for I_2, we find the following set of inequalities:

$$
\begin{aligned}
0 &\leq \int_0^t L(I_1(s), \gamma_2^t(s), \dot\gamma_2^t(s))\, ds + g(\gamma_2^t(0)) \\
&= \int_0^t L(I_2(s), \gamma_2^t(s), \dot\gamma_2^t(s))\, ds + g(\gamma_2^t(0)) + \int_0^t \phi_1(t, s)\, (I_1(s) - I_2(s))\, ds, \\
&= -\int_0^t \phi_1(t, s)\, (I_2(s) - I_1(s))\, ds.
\end{aligned}
$$

i.e.

$$\int_0^t \phi_1(t,s)\left(I_2(s) - I_1(s)\right) ds \le 0\,, \tag{3.2}$$

where the positive weight ϕ_1 is given by

$$\phi_1(t,s) = \int_0^1 \frac{\partial L}{\partial I}\left((1-\theta)I_1(s) + \theta I_2(s), \gamma_2^t(s), \dot\gamma_2^t(s)\right) d\theta\,. \tag{3.3}$$

Similarly, by exchanging the roles of the two solutions, we obtain

$$\int_0^t \phi_2(t,s)\left(I_2(s) - I_1(s)\right) ds \ge 0\,, \tag{3.4}$$

where the positive weight ϕ_2 is given by

$$\phi_2(t,s) = \int_0^1 \frac{\partial L}{\partial I}\left((1-\theta)I_2(s) + \theta I_1(s), \gamma_1^t(s), \dot\gamma_1^t(s)\right) d\theta\,. \tag{3.5}$$

By Assumption (L2) and the uniform boundedness of $(I_j(s), \gamma_j^t(s), \dot\gamma_j^t(s))$ (by (2.1) and (3.1)), there exists $\lambda > 0$ such that:

$$(\forall t, s) \quad \min_{j=1,2} \phi_j(t,s) \ge \lambda\,. \tag{3.6}$$

Functions of bounded variations have left- and right-limits everywhere. Here, we focus on the value of the right-limit at the origin. This is expressed in the following statement.

Lemma 11. *Let ϕ_1, ϕ_2 be defined as in* (3.3) *and* (3.5)*. Then*

$$\lim_{t\to 0+} \left\{ [\phi_1(t,\cdot)]_{BV(0,t)} + [\phi_2(t,\cdot)]_{BV(0,t)} \right\} = 0.$$

Proof. Our first observation is that BV regularity of $\{I_j\}_{j=1,2}$ implies the following smallness estimate:

$$\lim_{t\to 0+} \left\{ [I_1]_{BV(0,t)} + [I_2]_{BV(0,t)} \right\} = 0. \tag{3.7}$$

The important point here is that the left point of the interval is fixed to 0. The same conclusion would not be true if the interval $(0,t)$ would be replaced with $(-t,t)$ due to possible jump discontinuity at the origin. To prove (3.7), let us decompose I_1, say, into a difference of non-decreasing functions $I_1 = I_1^+ - I_1^-$. Then,

$$[I_1]_{BV(0,t)} \le I_1^+\big|_{0+}^t + I_1^-\big|_{0+}^t \underset{t\to 0+}{\longrightarrow} 0\,, \tag{3.8}$$

simply because I_1^+ and I_1^- have right limits at the origin.

By (2.2), we get that this vanishing limit can be extended to $\dot\gamma_j^t$ as well:

$$[\dot\gamma_j^t]_{BV(0,t)} \le C\left(t + [I_j]_{BV(0,t)}\right) \underset{t\to 0+}{\longrightarrow} 0. \tag{3.9}$$

Consequently, we are able to estimate $[\phi_2(t,\cdot)]_{BV(0,t)}$ as follows. To keep the idea concise, we will compute the derivative $\frac{\partial \phi_2}{\partial s}$ of the BV function $\phi_2(t,\cdot)$ in the sense

12

of a finite measure on $(0, t)$, so that $[\phi_2(t, \cdot)]_{BV(0,t)} = \int_{(0,t)} |\frac{\partial \phi_2}{\partial s}(t, s)| \, ds$. We shall adopt this convention for the remainder of the paper.

$$\left| \frac{\partial \phi_2}{\partial s}(t, s) \right| \leq \int_0^1 \left| d_{I,I}^2 L(\Gamma(s)) \left((1 - \theta)\dot{I}_2(s) + \theta \dot{I}_1(s) \right) \right| \, d\theta$$

$$+ \int_0^1 \left| d_{I,x}^2 L(\Gamma(s)) \dot{\gamma}_1^t(s) \right| \, d\theta + \int_0^1 \left| d_{I,v}^2 L(\Gamma(s)) \ddot{\gamma}_1^t(s) \right| \, d\theta$$

$$\leq C \left(|\dot{I}_1(s)| + |\dot{I}_2(s)| + |\dot{\gamma}_1^t(s)| + |\ddot{\gamma}_1^t(s)| \right), \tag{3.10}$$

where we have used the shortcut notation $\Gamma(s) = ((1 - \theta)I_2(s) + \theta I_1(s), \gamma_1^t(s), \dot{\gamma}_1^t(s))$. We may integrate the latter over the open interval $(0, t)$ to obtain

$$[\phi_2(t, \cdot)]_{BV(0,t)} \leq C \left([I_1]_{BV(0,t)} + [I_2]_{BV(0,t)} + t + \left[\dot{\gamma}_1^t \right]_{BV(0,t)} \right)$$

$$\leq C \left([I_1]_{BV(0,t)} + [I_2]_{BV(0,t)} + t \right)$$

where we used (3.9). By (3.8), we deduce that $\lim_{t \to 0+} [\phi_2(t, \cdot)]_{BV(0,t)} = 0$. The proof for ϕ_1 is analogous. □

We are now in position to prove Theorem 1.

Proof of Theorem 1. Let $\mu = I_2 - I_1$ and suppose to the contrary that $\mu \neq 0$ on a set of positive measure in $(0, T)$.

We claim that we may assume, without loss of generality, that $\mu \neq 0$ in a set of positive measure in $(0, t)$, for each $t \in (0, T)$. To see this claim, let

$$t_0 := \sup\{t \geq 0 : \mu(s) = 0 \text{ a.e. in } [0, t]\}.$$

If $t_0 = 0$, we are done. If $t_0 > 0$, then the variational solutions $V_1(t, x) \equiv V_2(t, x)$ are identical for $(t, x) \in [0, t_0] \times \mathbb{R}^d$. Now, $V_j(t_0, \cdot)$ has minimum value zero (by (1.7)) and is coercive (by Lemma 10), so the condition (G) is verified. Therefore, we may re-label the initial time to be t_0. In any case, it suffices to derive a contradiction assuming $\mu \neq 0$ in a set of positive measure in $(0, t)$, for each $t \in (0, T)$.

Using $\int_0^t \phi_2(t, s)\mu(s) \, ds \geq 0$ (by (3.4)), we may integrate by parts to obtain

$$\phi_2(t, t-) \int_0^t \mu(\tau) \, d\tau = \int_{(0,t)} \frac{\partial \phi_2}{\partial s}(t, s) \left(\int_0^s \mu(\tau) \, d\tau \right) ds + \int_0^t \phi_2(t, s)\mu(s) \, ds$$

$$\geq \int_{(0,t)} \frac{\partial \phi_2}{\partial s}(t, s) \left(\int_0^s \mu(\tau) \, d\tau \right) ds$$

Taking the negative part, we deduce the following partial estimate,

$$\phi_2(t, t-) \left(\int_0^t \mu(s) \, ds \right)_- \leq \left(\sup_{s \in (0,t)} \left| \int_0^s \mu(\tau) \, d\tau \right| \right) \int_{(0,t)} \left| \frac{\partial \phi_2}{\partial s}(t, s) \right| ds \tag{3.11}$$

Similarly we deduce from (3.2) that

$$\phi_1(t, t-) \int_0^t \mu(\tau) \, d\tau = \int_0^t \frac{\partial \phi_1}{\partial s}(t, s) \left(\int_0^s \mu(\tau) \, d\tau \right) ds + \int_0^t \phi_1(t, s)\mu(s) \, ds$$

$$\leq \int_0^t \frac{\partial \phi_1}{\partial s}(t, s) \left(\int_0^s \mu(\tau) \, d\tau \right) ds$$

13

Taking the positive part, we deduce the following complementary estimate,

$$\phi_1(t,t-)\left(\int_0^t \mu(s)\,ds\right)_+ \leq \left(\sup_{s\in(0,t)}\left|\int_0^s \mu(\tau)\,d\tau\right|\right)\int_0^t \left|\frac{\partial\phi_1}{\partial s}(t,s)\right|ds \qquad (3.12)$$

Combining (3.11) and (3.12), together with (3.6), we obtain

$$\lambda\left|\int_0^t \mu(s)\,ds\right| \leq \max_{i=1,2}\left\{\int_0^t \left|\frac{\partial\phi_i}{\partial s}(t,s)\right|ds\right\}\left(\sup_{s\in(0,t)}\left|\int_0^s \mu(\tau)\,d\tau\right|\right) \qquad (3.13)$$

Next, Lemma 11 ensures that there exists $t_1 > 0$ so that

$$\sup_{t\in(0,t_1)}\max_{i=1,2}\left\{\int_0^t \left|\frac{\partial\phi_i}{\partial s}(t,s)\right|ds\right\} \leq \frac{\lambda}{2}.$$

Then, taking supremum in (3.13) for $0 < t < t_1$, we have

$$\lambda\left(\sup_{t\in(0,t_1)}\left|\int_0^t \mu(s)\,ds\right|\right) \leq \frac{\lambda}{2}\left(\sup_{s\in(0,t_1)}\left|\int_0^s \mu(\tau)\,d\tau\right|\right).$$

This implies $\int_0^t \mu(s)\,ds = 0$ for all $t \in [0,t_1]$. Hence, $\mu(t) = 0$ almost everywhere on $(0,t_1)$. This is in contradiction with the assumption that $\mu \neq 0$ on a set of positive measure in $(0,t_1)$, and we conclude that $I_2 - I_1 = \mu = 0$ a.e. Finally $V_1 \equiv V_2$ by virtue of the variational formulation. $\qquad\square$

Remark 2. *We discuss a situation when the montonicity assumption (H2) or (L2) may not hold globally, but our arguments in Section 4 can still be applied. Suppose $(V_i(t,x), I_i(t)) \in W_{loc}^{1,\infty}(\mathbb{R}\times[0,T])\times BV([0,T])$ $(i=1,2)$ are two variational solutions such that $V_i(0,x) = g(x)$, and suppose we have reduced to the situation that, for each t small, the set $\{t' \in (0,t) : I_1(t') \neq I_2(t')\}$ has positive measure. By inspection of the proof of Theorem 1, we need the following conditions:*

(U1) *$L(I,t,x,v) = \mathcal{L}(v) - R(I,t,x)$ is of class $\mathcal{C}^2$.*

(U2) *For $i = 1,2$, $V_i(t,x)$ admit the variational characterization*

$$V_i(t,x) = \inf_{\gamma\in AC(0,t):\gamma(t)=x}\left\{\int_0^t L(I_i(s),s,\gamma(s),\dot\gamma(s))\,ds + g(\gamma(0))\right\}$$

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

(U3) *For $i = 1,2$ and $t > 0$ small, there exists $x_i^t \in \{x : V(t,x) = 0\}$ such that*

    (i) *$\lim_{t\to 0+} x_i^t$ exists;*

    (ii) *$\sup_{0<\theta<1}\partial_I R(\lim_{t\to 0+}[(1-\theta)I_1(t)+\theta I_2(t)], 0, \lim_{t\to 0+} x_i^t) < 0$;*

    (iii) *$\limsup_{t\to 0+}\|\dot\gamma_i^t\|_{L^\infty(0,t)} < +\infty$, where γ_i^t is the minimizing path corresponding to value of V_i at the point (t,x_i^t).*

The key is to replace inequality (3.6), which may no longer be true, by a weaker inequality. Indeed, we can follow (3.5) to define

$$\phi_i(t, s) = \int_0^1 \partial_I L((1-\theta)I_1(s) + \theta I_2(s), s, \gamma_{3-i}^t(s), \dot{\gamma}_{3-i}^t(s)) \, d\theta$$

$$= -\int_0^1 \partial_I R((1-\theta)I_1(s) + \theta I_2(s), s, \gamma_{3-i}^t(s)) \, d\theta. \tag{3.14}$$

By (U3)(iii), *we have*

$$\sup_{(0,t)} |\gamma_i^t(s) - x_i^t| = \sup_{(0,t)} |\gamma_i^t(s) - \gamma_i^t(t)| \to 0 \quad as \ t \to 0+. \tag{3.15}$$

Hence, we have the following estimate to replace (3.6):

$$\liminf_{t \to 0+} \inf_{s \in (0,t)} \phi_i(t, s)$$

$$= \liminf_{t \to 0+} \inf_{s \in (0,t)} \int_0^1 -\partial_I R((1-\theta)I_1(s) + \theta I_2(s), s, \gamma_{3-i}^t(s)) \, d\theta$$

$$\geq \lambda := - \sup_{0 < \theta < 1} \partial_I R(\lim_{t \to 0+} [(1-\theta)I_1(t) + \theta I_2(t)], 0, \lim_{t \to 0+} x_{3-i}^t) > 0, \tag{3.16}$$

where we used (3.14) to obtain the first equality, then used (3.15) to obtain the first inequality, and then used (U3)(ii) *for the final inequality. By invoking (3.16) instead of (3.6), one can repeat the exact same arguments in the proof of Lemma 11[1] and Theorem 1 to derive a contradiction.*

4   Proof of the Pessimization Principle

We prove Theorem 5 in this section.

Proof of Theorem 5. We start by choosing the right-continuous representative of I without loss of generality. For each $t > 0$, let x^t be a minimum point of $x \mapsto u(t, x)$ as before, and let γ^t be an associated minimizing curve ending up at x^t.

Step #1: $L(I(t-), x^t, 0) \leq 0$ **for all** t**.** It follows from the non-negativity constraint and the variational formulation (1.4) that

$$0 \leq u(s, \gamma^t(s)) = \int_0^s L(I(s'), \gamma^t(s'), \dot{\gamma}^t(s')) \, ds' + g(\gamma^t(0)) \quad \text{for } 0 < s < t,$$

and the equality holds when $s = t$. Hence, we deduce that $L(I(t-), \gamma^t(t-), \dot{\gamma}^t(t-)) \leq 0$. Since $\gamma^t(t-) = \gamma^t(t) = x^t$, we may use (L5) to deduce that

$$L(I(t-), x^t, 0) \leq L(I(t-), \gamma^t(t-), \dot{\gamma}^t(t-)) \leq 0.$$

Step #2: $I(t-) \leq I(t+)$ **for all** t**.** Fix $t > 0$, let x^t and γ^t be as above. We define $\gamma_1 : [0, t+1] \to \mathbb{R}^d$ by

$$\gamma_1(s) = \begin{cases} \gamma^t(s) & \text{for } 0 \leq s \leq t, \\ x^t & \text{for } s > t. \end{cases}$$

---

[1] Under assumption (U1), the term $|\ddot{\gamma}_1^t(s)|$ does not appear on the right hand side of (3.10), as $\partial_I L$ is independent of $\dot{\gamma}_1^t$. For this reason, the estimate $[\gamma_i^{t,x}]_{BV(0,t)} \to 0$ is not included in (U3)(iii).

Then $\gamma_1 \in AC[0, t+1]$ and for all $0 < h < 1$,

$$0 \le u(t+h, \gamma_1(t+h)) \le \int_0^{t+h} L(I(s'), \gamma_1(s'), \dot{\gamma}_1(s'))\, ds' + g(\gamma_1(0)).$$

Since $0 = u(t, \gamma_1(t)) = \int_0^t L(I(s'), \gamma_1(s'), \dot{\gamma}_1(s'))\, ds' + g(\gamma_1(0))$, we have

$$0 \le \int_t^{t+h} L(I(s'), \gamma_1(s'), \dot{\gamma}_1(s'))\, ds' \le \int_t^{t+h} L(I(s'), x^t, 0)\, ds'. \tag{4.1}$$

Dividing by h, and letting $h \to 0+$, we obtain $L(I(t+), x^t, 0) \ge 0$. Comparing with $L(I(t-), x^t, 0) \le 0$ (by Step #1), we deduce from the monotonicity of L in I (L2) that $I(t-) \le I(t+)$ for all $t > 0$.

Step #3: Conclusion. Suppose to the contrary that $I(t_2) < I(t_1)$ for some $t_1 < t_2$. Since I is right-continuous, there exists $t_3 > t_2$ such that $I(t) < I(t_1)$ for all $t \in [t_2, t_3)$. Let $t_0 = \sup\{t \in [t_1, t_3) : I(t) \ge I(t_1)\}$. Then $t_0 \le t_2 < t_3$, and

$$I(t) < I(t_1) \le I(t_0-) \quad \text{for } t \in (t_0, t_3). \tag{4.2}$$

Now, using Step #1 and (4.1) from Step #2, we have

$$0 \le \int_{t_0}^{t_3} \left[L(I(s'), x^{t_0}, 0) - L(I(t_0-), x^{t_0}, 0) \right]\, ds'.$$

However, this is in contradiction with (4.2), in view of the fact that L is strictly increasing in I, due to (L2). $\qquad\square$

A Variational and viscosity solutions coincide (proof of Theorem 2)

Given $I \in BV(0, T)$, let $V(t, x)$ denote the corresponding variational solution of (1.4). The purpose of this section is to show that $V(t, x)$ is the unique locally Lipschitz viscosity solution of (1.1). This can be achieved by establishing comparison theorem, i.e. $u \le V$ (resp. $u \ge V$) for all locally Lipschitz viscosity sub-solution (resp. super-solution) of (1.1). While there are PDE proofs for such comparison results among continuous, but not necessarily Lipschitz, super and sub-solutions of (1.1), they are usually proved under slightly different conditions than (L1) - (L4). For instance, in [13] (see also [16, Appendix A]), it is assumed that the Hamiltonian H is uniformly Lipschitz in $x \in \mathbb{R}^d$. Henceforth, we will adopt techniques in convex analysis to prove the comparison between the variational solution with Lipschitz continuous super and sub-solutions of (1.1), under exactly the assumptions (L1) - (L4).

As the Hamiltonian is convex with respect to p, sub-solutions in the almost everywhere sense, and viscosity sub-solutions in particular, lie automatically below the variational solution [4, 11]. We include a proof here for the sake of completeness.

Proposition 12. *Assume that u is locally Lipschitz, that $u(0, x) \le g(x)$ for all x, and that the following inequality holds for almost every $(t, x) \in (0, T) \times \mathbb{R}^d$,*

$$\partial_t u(t, x) + H(I(t), x, d_x u(t, x)) \le 0 \quad a.e. \tag{A.1}$$

Then, $u \le V$.

Proof. The proof is adapted from [11, Section 4.2]. A more direct proof can be found in [4, Section 9] but the latter assumes time continuity for H, which does not hold in the present case. A first observation is that (A.1) makes perfect sense as u is differentiable almost everywhere by Rademacher's theorem. We shall establish that

$$u(t_2, \gamma(t_2)) - u(t_1, \gamma(t_1)) \leq \int_{t_1}^{t_2} L(I(s), \gamma(s), \dot{\gamma}(s))\, ds\,, \tag{A.2}$$

for all curves $\gamma \in W^{1,\infty}$. Thus, the result will follow immediately by taking the infimum with respect to γ, since (2.1) of Lemma 9 says that any minimizer is indeed $W^{1,\infty}$.

To prove (A.2), we proceed by a density argument. The case of a linear curve $\gamma = x + (s - t_1)v$ is handled as follows: firstly, we deduce from (A.1) that

$$\partial_t u(t, x) + d_x u(t, x) \cdot v \leq L(I(t), x, v) \quad a.e. \tag{A.3}$$

Secondly, by Fubini's theorem one can find a sequence $x_n \to x$ such that (A.3) holds almost everywhere in the line $\{(s, x_n + (s - t_1)v)\}$ for each n. Therefore, we can apply the chain rule to $u(s, x_n + (s - t_1)v)$, so as to obtain:

$$\frac{d}{ds}\left(u(s, x_n + (s - t_1)v)\right) \leq L(I(s), x_n + (s - t_1)v, v) \quad a.e. \tag{A.4}$$

We deduce that (A.2) holds true for all linear curves by integrating (A.4) from t_1 to t_2 and taking the limit $n \to +\infty$.

Consequently, (A.2) holds true for any piecewise linear curve. The conclusion follows by a density argument of piecewise linear curves in the set of curves having bounded measurable derivatives. □

It remains to show that viscosity super-solutions lie above the variational solution. For completeness' sake, we give a definition of super-solution for time-measurable Hamiltonians. (See [12, 15] for various other equivalent definitions.)

Definition 13 (Viscosity super-solution). *Let $\phi \in \mathcal{C}^1(\mathbb{R}^d)$ be such that the minima of $u(t, \cdot) - \phi$ are reached in a ball of radius R for all $t \in [0, T]$. Let $\mathfrak{M}(t)$ be the set of minimum points of $u(t, \cdot) - \phi$, and $m(t) = \min u(t, \cdot) - \phi$. Then, it is required that the following inequality holds true in the distributional sense:*

$$m'(t) + \sup_{y \in \mathfrak{M}(t)} H(I(t), y, d_x \phi(y)) \geq 0 \quad in\ \mathcal{D}'(0, T)\,. \tag{A.5}$$

Rather than directly invoking Definition 13, we will only use the following two consequences of it in our proofs.

Remark 3. *In case $I(t)$ is continuous, then $\hat{H}(t, x, p) := H(I(t), x, p)$ defines a continuous Hamiltonian. In that case, the above definition is consistent with the usual one for viscosity super-solution based on the notion of sub-differential [4, Definition 3.2]. Namely, $u \in C((0, T) \times \mathbb{R}^d)$ is a super-solution of (1.1) if, for each (t, x),*

$$q + \hat{H}(t, x, p) \geq 0 \quad for\ all\ (q, p) \in D^{2,-}u(t, x),$$

where the set of sub-differential, $D^{2,-}u(t,x)$, is the subset of $\mathbb{R} \times \mathbb{R}^d$ given by

$$D^{2,-}u(t,x) = \{(q,p) : (\forall \mu, v) \quad u(t,x) - u(t-\mu, x-sv) \leq qs\mu + \langle p, sv \rangle + o(s)\}.$$

Remark 4. *Suppose that u is a viscosity super-solution of (1.1) for some $I(t)$ in the sense of Definition 13, and that $I(t) \geq \hat{I}(t)$ for some continuous $\hat{I}(t)$, then by monotonicity of H in I, it can be verified that u is a super-solution of (1.1), with $I(t)$ replaced by $\hat{I}(t)$, in the usual sense as in Remark 3. See [15] for details.*

Proposition 14. *Assume that u is a locally Lipschitz viscosity super-solution, in the sense of Definition 13, and that $u(0,x) \geq g(x)$ for all x. Then, $u \geq V$.*

Proof. We follow the lines of [8] which is essentially based on convex analysis. We adapt their proof in our context for the sake of completeness. We will first prove the proposition in the special case of $I \in W^{1,\infty}(0,T)$. This assumption will be relaxed to $I \in BV(0,T)$ at the end of the proof.

Step #1: Finding the backward velocity: setting of the problem. The key is to find, for each (t,x), a particular direction $\mathbf{v}(t,x)$, such that the following inequality holds true:

$$d_+u(t,x)(1,\mathbf{v}(t,x)) \geq L(I(t), x, \mathbf{v}(t,x)), \tag{A.6}$$

where $d_+u(t,x)(\mu, v)$ is the one-sided directional differentiation in the direction (μ, v):

$$d_+u(t,x)(\mu, v) = \limsup_{s \to 0+} \frac{u(t,x) - u(t-s\mu, x-sv)}{s}.$$

We can interpret (A.6) as follows: there exists an element which is common to the partial epigraph of $v \mapsto L(I(t), x, v)$:

$$\mathcal{E}_{t,x} = \mathbf{Epi}_v(L(I(t), x, v)) = \left\{ (v, \ell) \in \mathbb{R}^d \times \mathbb{R} : \ell \geq L(I(t), x, v) \right\},$$

and to the hypograph of $v \mapsto d_+u(t,x)(1,v)$:

$$\mathbf{Hypo}_v(d_+u(t,x)(1,v)) = \left\{ (v, \ell) \in \mathbb{R}^d \times \mathbb{R} : \ell \leq d_+u(t,x)(1,v) \right\}.$$

For technical reason, we consider the full hypograph $\mathcal{H}_{t,x} = \mathbf{Hypo}_{(\mu,v)}(d_+u(t,x)(\mu,v))$, taken with respect to variables $(\mu, v) \in \mathbb{R} \times \mathbb{R}^d$. Precisely,

$$\mathcal{H}_{t,x} = \left\{ (\mu, v, \ell) \in \mathbb{R} \times \mathbb{R}^d \times \mathbb{R} : \ell \leq \limsup_{s \to 0+} \frac{u(t,x) - u(t-s\mu, x-sv)}{s} \right\}. \tag{A.7}$$

In contrast with $\mathbf{Hypo}_v(d_+u(t,x)(1,v))$, $\mathcal{H}_{t,x}$ is a cone because the quantity in (A.7) is positively homogeneous with respect to (μ, v). In fact, it coincides with the definition of a contingent cone, up to a change of sign. If $S \subset \mathbb{R}^N$ is a non-empty subset, and $z \in \mathbb{R}^N$, recall that the contingent cone of S at z, denoted by $\mathcal{T}_S(z)$, is defined as follows [3, Definition 3.2.1]:

$$w \in \mathcal{T}_S(z) \quad \Longleftrightarrow \quad \liminf_{s \to 0^+} \frac{\text{dist}(z + sw, S)}{s} = 0.$$
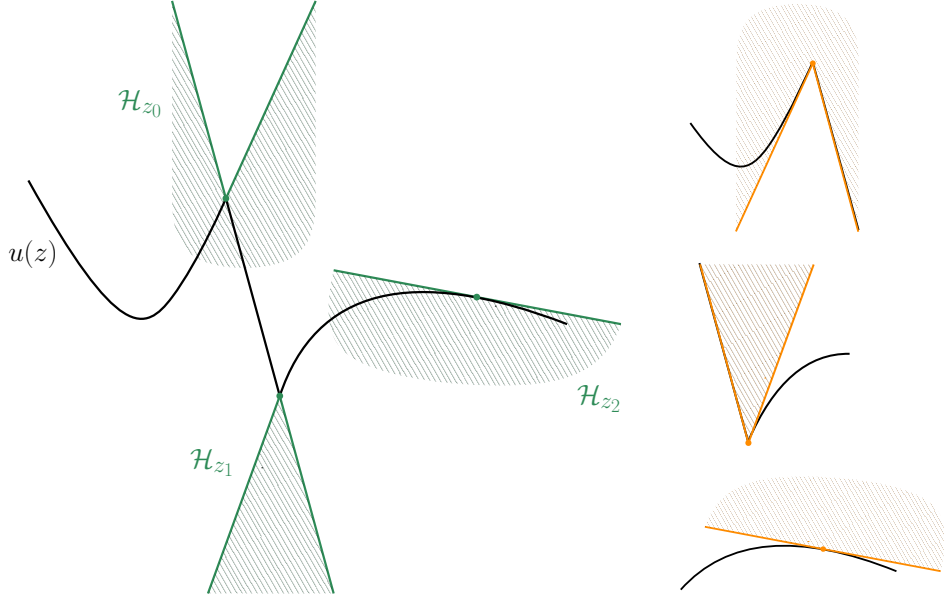
18

Figure 1: Illustration of various shapes of cones as they may appear for a scalar function u (in opposition to the text where the domain of u is genuinely multi-dimensional). The regions shaded green correspond to (translated) hypograph $(z_i, u(z_i)) + \mathcal{H}_{z_i}$, whereas the region shaded orange correspond to the (translated) contingent cones $(z_i, u(z_i)) + \mathcal{T}_{\mathbf{Epi}\, u}(z_i, u(z_i)) = (z_i, u(z_i)) - \mathcal{H}_{z_i}$. Note that we have translated the vertices of the cones to the respective points $(z_i, u(z_i))$ for illustrative purposes.

Then, we claim the following equivalence:

$$\mathcal{H}_{t,x} = -\mathcal{T}_{\mathbf{Epi}\, u}(t, x, u(t, x)), \tag{A.8}$$

where $\mathbf{Epi}\, u = \{(t, x, \ell) : \ell \geq u(t, x)\}$. For the convenience of readers, the equivalence (A.8) is illustrated in Figure 1 for a scalar function u. Now we show (A.8). Indeed, (μ, v, ℓ) belongs to $-\mathcal{T}_{\mathbf{Epi}\, u}(t, x, u(t, x))$ if and only if there exist subsequences $s_n \to 0+$ and (t_n, x_n, u_n) such that:

$$\begin{cases} t - s_n\mu = t_n + o(s_n) \\ x - s_nv = x_n + o(s_n) \\ u(t, x) - s_n\ell = u_n + o(s_n) \end{cases} \quad \text{and} \quad u_n \geq u(t_n, x_n).$$

The latter inequality is inherited from the choice $S = \mathbf{Epi}\, u$. Reorganizing the terms, and using the Lipschitz continuity of u, we obtain:

$$u(t, x) - s_n\ell \geq u(t - s_n\mu, x - s_nv) + o(s_n),$$

$$\frac{u(t, x) - u(t - s_n\mu, x - s_nv)}{s_n} \geq \ell + o(1).$$

The latter is precisely (A.7). i.e. $(\mu, v, \ell) \in \mathcal{H}_{t,x}$ and this proves (A.8).

Summarizing, we are seeking a vector $v \in \mathbb{R}^d$ so that the element $(1, v, L(I(t), x, v))$ is common to $\mathcal{H}_{t,x}$ and to $\{1\} \times \mathcal{E}_{t,x}$. The latter is a convex set, but the former is not necessarily convex. Therefore, we are led to consider its convex closure $\overline{\mathrm{co}}(\mathcal{H}_{t,x})$

in order to use the separation theorem. Next, we shall use the viability theory to remove the convex closure, exactly as in [8].

Step #2: Finding the backward velocity: the separation theorem. We wish to avoid separation of the two convex sets $\overline{\text{co}}(\mathcal{H}_{t,x})$ and $\{1\} \times \mathcal{E}_{t,x}$. We argue by contradiction. If the two sets are separated, then there exists a linear form $q \cdot + \langle p, \cdot \rangle$ such that (i) $\overline{\text{co}}(\mathcal{H}_{t,x})$ lies below the hyper-plane $\{(\mu, v, \ell) : \ell = q\mu + \langle p, v \rangle\}$, and (ii) $\{1\} \times \mathcal{E}_{t,x}$ lies strictly above it [21]. We deduce from the latter condition (ii) that $q + \langle p, v \rangle \leq L(I(t), x, v) - \delta$ for all $v \in \mathbb{R}^d$ and some $\delta > 0$. This can be recast as $q + H(I(t), x, p) \leq -\delta$ from the definition of the Legendre transform. On the other hand, we deduce from condition (i) that

$$\limsup_{s \to 0^+} \frac{u(t,x) - u(t - s\mu, x - sv)}{s} = d_+ u(t,x)(\mu, v) \leq q\mu + \langle p, v \rangle ,$$

for all $(\mu, v) \in \mathbb{R}^{d+1}$. Consequently, (q, p) belongs to the subdifferential of u at (t, x). By applying the usual criterion of viscosity super-solutions for continuous Hamiltonian functions (see Remark 3), we find that $q + H(I(t), x, p) \geq 0$. This is a contradiction. Thus, the two convex sets are not separated, i.e.

$$(\forall t, x) \quad \overline{\text{co}}(\mathcal{H}_{t,x}) \cap (\{1\} \times \mathcal{E}_{t,x}) \neq \emptyset . \tag{A.9}$$

Step #3: Finding the backward velocity: the viability theorem. Note that (A.9) is equivalent to

$$(\forall t, x) \quad \overline{\text{co}}(-\mathcal{T}_{\mathbf{Epi}\,u}(t, x, u(t,x))) \cap (\{1\} \times \mathcal{E}_{t,x}) \neq \emptyset . \tag{A.10}$$

We wish to use the viability theorem [3, p. 85] (see also [8, Theorem 2.3]):

Theorem 15 (Viability). *Suppose that $G : \mathbb{R}^N \rightsquigarrow \mathbb{R}^N$ is an upper semi-continuous set-valued map with compact convex values. Then for each closed set $S \subset \mathbb{R}^N$, the following statements are equivalent:*

(a) $(\forall z \in S) \quad \mathcal{T}_S(z) \cap G(z) \neq \emptyset$;

(b) $(\forall z \in S) \quad (\overline{\text{co}}\, \mathcal{T}_S(z)) \cap G(z) \neq \emptyset$.

Further compactness estimate is required in order to apply Theorem 15. We claim that we can restrict (A.9) to a compact set:

$$\overline{\text{co}}(\mathcal{H}_{t,x}) \cap (\{1\} \times \mathcal{E}_{t,x}) \cap \left(\{1\} \times B(0, R_{|x|}) \times [m, M]\right) \neq \emptyset ,$$

where for each $K > 0$, $R_K = \max\{1, r_K\}$, with r_K increasing in K such that

$$\Theta(r) > [u]_{\text{Lip}(\overline{[0,T] \times B(0,K)})}(1 + r) + C_\Theta \quad \text{for all } r \geq r_K , \tag{A.11}$$

(the choice of r_K is possible due to the superlinear growth of Θ), and m, M are respectively $m = \min L$, $M = \max L$ where both minimum and maximum are taken over the set $J \times \{x\} \times B(0, R_{|x|})$, where J is a compact set containing the values $\{I(t)\}_{t \in (0,T)}$.

To this end, consider the following two options: either the dual cone $(\mathcal{H}_{t,x})^-$ is empty or non-empty. In the first case, it implies $\overline{\text{co}}(\mathcal{H}_{t,x}) = \mathbb{R}^d$, so that any element of $\mathcal{E}_{t,x}$ is appropriate. In particular,

$$(1, 0, L(I(t), x, 0)) \in \overline{\text{co}}(\mathcal{H}_{t,x}) \cap (\{1\} \times \mathcal{E}_{t,x}) \cap (\{1\} \times B(0, 1) \times [m, M]) . \tag{A.12}$$

In the second case, $(\mathcal{H}_{t,x})^-$ is non-empty. Hence, there exists a linear form $q \cdot + \langle p, \cdot \rangle$ such that $\overline{\mathrm{co}}(\mathcal{H}_{t,x})$ lies below the linear set $\{q\mu + \langle p, v \rangle\}$ as in Step #2. Therefore, every common point $(1, v, \ell) \in \overline{\mathrm{co}}(\mathcal{H}_{t,x}) \cap (\{1\} \times \mathcal{E}_{t,x})$ (and there is at least one such point) must satisfy

$$L(I(t), x, v) \leq \ell \leq q + \langle p, v \rangle.$$

By the facts that (i) L grows uniformly super-linearly (by (L3)), and (ii) (q, p) is bounded as it belongs to the subdifferential of the locally Lipschitz function u, i.e. $\max\{|q|, |p|\} \leq [u]_{\mathrm{Lip}(\overline{(0,T) \times B(0,|x|)})}$, we deduce

$$\Theta(|v|) - C_\Theta \leq L(I(t), x, v) \leq \ell \leq [u]_{\mathrm{Lip}(\overline{(0,T) \times B(0,|x|)})}(1 + |v|).$$

By the choice of r_K in (A.11), we must have $|v| < r_K$ with $K = |x|$, that is

$$\overline{\mathrm{co}}(\mathcal{H}_{t,x}) \cap (\{1\} \times \mathcal{E}_{t,x}) \cap (\{1\} \times B(0, r_{|x|}) \times [m, M]) \neq \emptyset. \qquad (A.13)$$

By (A.12) and (A.13), and our choice of $R_{|x|} := \max\{1, r_{|x|}\}$, we find that

$$\overline{\mathrm{co}}\left(-\mathcal{T}_{\{\mathbf{Epi}\,u\}}(t, x, u(t, x))\right) \cap (-G(t, x)) = \overline{\mathrm{co}}(\mathcal{H}_{t,x}) \cap (-G(t, x)) \neq \emptyset,$$

where $G(t, x) = -(\{1\} \times \mathcal{E}_{t,x}) \cap (\{1\} \times B(0, R_{|x|}) \times [m, M])$ is a continuous set-valued map with compact convex values. In order to apply the viability theorem to the closed subset $S = \mathbf{Epi}\,u$, it remains to check that the statement (b) of Theorem 15, i.e.

$$\overline{\mathrm{co}}(\mathcal{T}_{\mathbf{Epi}\,u}(t, x, U)) \cap G(t, x) \neq \emptyset$$

holds for all $(t, x, U) \in \mathbf{Epi}\,u$, and not only for points $(t, x, u(t, x))$ on the graph of u. This is immediate, as $\mathcal{T}_{\{\mathbf{Epi}\,u\}}(t, x, U) = \mathbb{R}^{d+2}$ for $U > u(t, x)$.

Finally, all the assumptions of the viability theorem are met. As a consequence, we can remove the convex closure in (A.10), and thus in (A.9), so as to obtain:

$$(\forall t, x) \quad \mathcal{H}_{t,x} \cap (\{1\} \times \mathcal{E}_{t,x}) \neq \emptyset.$$

In particular, for each (t, x) there exists a vector $\mathbf{v}(t, x)$ such that (A.6) holds true.
Step #4: Building the backward trajectory up to the initial time. Now that we are able to make a small step backward at each (t, x), let $\epsilon > 0$ be given, and start from (t_0, x_0). There exists (s_0, v_0) such that

$$u(t_0, x_0) \geq s_0 L(t, x, v_0) + u(t_0 - s_0, x_0 - s_0 v_0) - \epsilon s_0.$$

By choosing s_0 small enough, we can even replace the right-hand-side by:

$$u(t_0, x_0) \geq \int_0^{s_0} L(t - s, x_0 - s v_0, v_0)\, ds + u(t_0 - s_0, x_0 - s_0 v_0) - 2\epsilon s_0. \qquad (A.14)$$

In particular, we have

$$u(t_0, x_0) \geq \inf_\gamma \int_{t_0 - s_0}^{t_0} L(s', \gamma(s'), \dot{\gamma}(s'))\, ds' + u(t_0 - s_0, \gamma(t_0 - s_0)) - 2\epsilon s_0,$$

where the infimum is taken over all $\gamma \in AC[0, t]$ such that $\gamma(t_0) = x_0$. As a result, the set

$$\Sigma = \left\{ \tau \in (0, t_0) : \ u(t_0, x_0) \geq \inf_\gamma \int_\tau^{t_0} L(s', \gamma(s'), \dot{\gamma}(s'))\, ds' + u(\tau, \gamma(\tau)) - 2\epsilon(t_0 - \tau) \right\}.$$

is non-empty, and $\tau_* := \inf \Sigma \in [0, t_0 - s_0]$ is well-defined. We wish to prove that $\tau_* = 0$. Suppose, for contradiction, that $\tau_* > 0$. Then, the estimates obtained in Lemma 9 allows to extract a converging sequence $\{\gamma_n\}$ such that γ_n is defined on the time span (τ_n, t), with $\tau_n \searrow \tau_*$, and $\{\dot\gamma_n\}$ is uniformly BV [2]. Hence, we can pass to the limit $\dot\gamma_n \to \dot\gamma$ a.e. by Helly's Selection Theorem, and then use Bounded Convergence Theorem to prove that

$$u(t_0, x_0) \geq \inf \int_{\tau_*}^{t_0} L(s', \gamma(s'), \dot\gamma(s')) \, ds' + u(\tau_*, \gamma(\tau_*)) - 2\epsilon(t_0 - \tau_*) \, .$$

By applying again the single step backward as in (A.14) at $(\tau_*, \gamma(\tau_*))$, we can push our lower estimate to an earlier time $\tau_{**} \in (0, \tau_*)$, and obtain thoroughly a contradiction. Thus, $\tau_* = 0$ and

$$u(t_0, x_0) \geq \inf_{\gamma} \int_0^{t_0} L(s', \gamma(s'), \dot\gamma(s')) \, ds' + g(\gamma(0)) - 2\epsilon t_0 = V(t_0, x_0) - 2\epsilon t_0.$$

By letting $\epsilon \to 0$, we have established that $u \geq V$ in the case when $t \mapsto I(t)$ is Lipschitz continuous.

To conclude, it remains to remove the additional continuity assumption on $I(t)$. Let $I \in BV(0, T)$. First of all, we approximate $I(t)$ from below by a sequence of Lipschitz functions $I_k(t) \nearrow I(t)$ converging pointwise [3]:

$$I_k(t) = \inf_{s > 0} \left(I(s) + k|t - s| \right) \leq I(t) \, .$$

It follows from (H2) and Remark 4 that u is also a super-solution associated with $I_k(t)$. Hence we have
$$u \geq V_k \quad \text{in } (0, T) \times \mathbb{R}^d, \tag{A.15}$$

where V_k is the variational solution associated with I_k.

On the other hand, the compactness estimates on minimizing curves obtained in Lemma 9 combined with Lebesgue's dominated convergence theorem guarantees that $V_k \nearrow V$. Thus, we may let $k \to \infty$ in (A.15) to deduce $u \geq V$. This completes the proof. $\qquad\square$

References

[1] L. Ambrosio, G. Dal Maso, A general chain rule for distributional derivatives, Proc. Amer. Math. Soc. 108, 691–792, 1990.

[2] L. Ambrosio, N. Fusco, D. Pallara, Functions of bounded variations and free discontinuity problems, Oxford University Press, 2000.

[3] J.-P. Aubin, Viability theory, Birkhäuser, 1991.

---

[2] The family $\{\dot\gamma_n\}$ is even uniformly Lipschitz as I is assumed to be Lipschitz here, see the proof of Lemma 9.

[3] Here, we choose the lower semi-continuous representative of I without loss of generality. Note that the criterion (A.5) is insensitive to the choice of the representative.

[4] G. Barles, An introduction to the theory of viscosity solutions for first-order Hamilton-Jacobi equations and applications, in Hamilton-Jacobi equations: approximations, numerical analysis and applications, Lecture Notes in Mathematics 2074, 2013.

[5] G. Barles, B. Perthame, Concentrations and constrained Hamilton-Jacobi equations arising in adaptive dynamics, Contemporary Mathematics 439, 57–68, 2007.

[6] G. Barles, S. Mirrahimi, B. Perthame, Concentration in Lotka-Volterra parabolic or integral equations: a general convergence result, Methods and Applications of Analysis 16, 321–340, 2009.

[7] M.G. Crandall, P.-L. Lions, Remarks on the existence and uniqueness of unbounded viscosity solutions of Hamilton-Jacobi equations, Illinois J. Math. 31, 665–688, 1987.

[8] G. Dal Maso, H. Frankowska, Value functions for Bolza problems with discontinuous Lagrangians and Hamilton-Jacobi inequalities, ESAIM: Control, Optimisation and Calculus of Variations 5, 369–393, 2000.

[9] O. Diekmann, A beginners guide to adaptive dynamics, Banach Center Publications 63, 47–86, 2003.

[10] O. Diekmann, P.-E. Jabin, S. Mischler, B. Perthame, The dynamics of adaptation: an illuminating example and a Hamilton-Jacobi approach, Theor. Pop. Biol. 67, 257–271, 2005.

[11] A. Fathi, Weak KAM theorem in Lagrangian dynamics, preliminary version number 10, 2008.

[12] H. Ishii, Hamilton-Jacobi equations with discontinuous Hamiltonians on arbitrary open sets, Bull. Fac. Sci. Engnrg Chuo Univ. 28, 33–77, 1985.

[13] H. Ishii, Comparison results for Hamilton-Jacobi equations without growth condition on solutions from above, Appl. Anal. 67 (1997) 357-372. `https://doi.org/10.1080/00036819708840617`

[14] Y. Kim, On the uniqueness for one-dimensional constrained Hamilton-Jacobi equations, preprint arXiv:1807.03432, 2018.

[15] P.-L. Lions, B. Perthame, Remarks on Hamilton-Jacobi equations with measurable time-dependent Hamiltonians, Non-linear Analysis TMA 11, 613–621, 1987.

[16] Q. Liu, S. Liu and K.-Y. Lam, Asymptotic spreading of interacting species with multiple fronts II: Exponentially decaying initial data, 50pp. arXiv:1908.05026 [math.AP]

[17] A. Lorz, S. Mirrahimi, B Perthame, Dirac mass dynamics in a multidimensional nonlocal parabolic equation, Communications in Partial Differential Equations 36, 1071–1098, 2011.

[18] S. Mirrahimi, J.-M. Roquejoffre, A class of Hamilton-Jacobi equations with constraint: uniqueness and constructive approach, J. Differential Equations, 260, 4717–4738, 2016.

[19] B. Perthame, Transport equations in biology, Birkhäuser, 2007.

[20] B. Perthame, G. Barles, Dirac concentrations in Lotka-Volterra parabolic PDEs, Indiana Univ. Math. J. 57, 3275–3301, 2008.

[21] R.T. Rockafellar, Convex analysis, Princeton University Press, 1970.

[22] A. Tourin, A comparison theorem for a piecewise Lipschitz continuous Hamiltonian and application to Shape-from-Shading problems, Numer. Math. 62 (1992) 75-85. `https://doi.org/10.1007/BF01396221`