Weak KAM Theory for a Weakly Coupled System of Hamilton-Jacobi Equations

A. Figalli, D. Gomes, and D. Marcon

Abstract

Here, we extend the weak KAM and Aubry-Mather theories to optimal switching problems. We consider three issues: the analysis of the calculus of variations problem, the study of a generalized weak KAM theorem for solutions of weakly coupled systems of Hamilton-Jacobi equations, and the long-time behavior of time-dependent systems. We prove the existence and regularity of action minimizers, obtain necessary conditions for minimality, extend Fathi's weak KAM theorem, and describe the asymptotic limit of the generalized Lax-Oleinik semigroup.

1 Introduction

Overview. Dynamical systems given by Tonelli Lagrangians have been extensively studied in recent years, and the deep connections between the calculus of variations, the weak KAM theory, and the Aubry-Mather theory have been a source of inspiration for researchers. Here, we extend these methods to optimal switching problems. We prove the existence and regularity of minimizers to a generalized Lagrangian action, extend the weak KAM theorem, and describe the long-time limit of weakly coupled Hamilton-Jacobi systems.

Optimal switching is a class of control problems where the running cost or the dynamics of a system can be modified by switching between different modes. The relation between viscosity solutions of quasivariational inequalities and optimal switching problems is well known. In [6], motivated by earlier publications [11, 7, 2], the authors extended the notion of viscosity solution to these problems and proved that their value functions are viscosity solutions of a weakly coupled system of Hamilton-Jacobi equations. Recently, several authors have investigated random switching problems, their weakly coupled Hamilton-Jacobi equations [19], the corresponding extensions of the weak KAM and Aubry-Mather theories [10, 17], the long-time behavior of solutions [3, 4, 18, 20, 22], and homogenization questions [21]. In these references, as in the present paper, the state of the systems has different modes. However, in those problems, the switching between modes is driven by a random process. In contrast, here, the switches occur at deterministic times as considered in [15]. Our results complement the earlier research in [15] and provide the counterpart of the aforementioned results for the optimal switching problem.

Setting of the problem. Let M be a compact, connected Riemannian manifold, TM its tangent bundle, and $\mathcal{I} = \{1, \dots, m\}$ a finite set of modes. The (multimodal) Lagrangian, $L : TM \times \mathcal{I} \to \mathbb{R}$, is, for each $i \in \mathcal{I}$, a Tonelli Lagrangian, $L(\cdot, \cdot, i) : TM \to \mathbb{R}$, prescribing the running cost at the mode "i". The switching cost is given by the function $\psi : M \times \mathcal{I} \times \mathcal{I} \to \mathbb{R}$. A trajectory, $\gamma = (\gamma_M, \gamma_{\mathcal{I}}) : [0, t] \to M \times \mathcal{I}$, determines both state and mode at each time. We denote by $AC([0, t]; M)$ the set of all absolutely continuous curves from $[0, t]$ into M, and by $\mathcal{P}([0, t]; \mathcal{I})$ the set of all piecewise constant functions on $[0, t]$ taking values on $\mathcal{I}$. More precisely, $\kappa \in \mathcal{P}([0, t]; \mathcal{I})$ if there exists a partition $P = \big\{0 < t_1 < \dots < t_N < t\big\}$ of $[0, t]$ such that

$$\kappa \text{ is constant on } (0, t_1) \cup \dots \cup (t_\ell, t_{\ell+1}) \cup \dots \cup (t_N, t). \tag{1}$$

In other words, κ can jump in the interval $[0, t]$ only at the times $\{0, t_1, \dots, t_n, t\}$ (note that κ can jump also at the initial and final times, 0 and t). The partition associated with κ is the smallest partition for which (1) holds. For simplicity, these spaces are denoted by AC and $\mathcal{P}$, respectively.

Let $t \in [0,+\infty)$ and $c \in \mathbb{R}$. The generalized Lagrangian action functional, $\mathcal{J}_t : AC \times \mathcal{P} \to \mathbb{R}$, is

$$\mathcal{J}_t[\gamma] := \int_0^t L\big(\gamma_M(s), \dot\gamma_M(s), \gamma_{\mathcal{I}}(s)\big)\ ds + \psi\big(\gamma_M(0),\ \gamma_{\mathcal{I}}(0),\ \gamma_{\mathcal{I}}(0^+)\big) \\ + \psi\big(\gamma_M(t),\ \gamma_{\mathcal{I}}(t^-),\ \gamma_{\mathcal{I}}(t)\big) + \sum_{\ell=1}^{N} \psi\big(\gamma_M(t_\ell),\ \gamma_{\mathcal{I}}(t_\ell^-),\ \gamma_{\mathcal{I}}(t_\ell^+)\big). \quad (2)$$

In other words, the action of the curve, $(\gamma_M, \gamma_{\mathcal{I}})$, is the sum of the action of the curves $\gamma_M|_{[t_\ell, t_{\ell+1}]}$ for the Lagrangian, $L(\cdot,\cdot,\kappa_k)$, plus the sum of the switching costs corresponding to changes of state.

Here, we investigate the following topics:

1. Properties of the calculus of variations problem associated with the action functional $\mathcal{J}_t$ - existence of minimizers, necessary conditions for minimality, and energy conservation.
2. A weak KAM-type theorem.
3. Extension of the Aubry-Mather Theory to system of Hamilton-Jacobi equations.

The existence of minimizers for $\mathcal{J}_t$ is a natural question that, to the best of our knowledge, was not addressed previously in the literature. Here, we prove the existence and the regularity of minimizers. Our assumptions are minimal in the sense that they are required in the "single" Lagrangian case.

For $t > 0$, we define $h_t : M \times \mathcal{I} \times M \times \mathcal{I} \to \mathbb{R}$ as

$$h_t(A,B) := \inf\Big\{\mathcal{J}_t[\gamma];\ \gamma \in AC \times \mathcal{P},\ \gamma(0) = A,\ \gamma(t) = B\Big\}, \qquad \forall\, A, B \in M \times \mathcal{I}. \quad (3)$$

Our first main result is:

Theorem 1.1. *Suppose that A1-A3 hold (see Section 3.1). Then, for every $A = (x,i)$ and $B = (y,j) \in M \times \mathcal{I}$, there exists $\gamma = (\gamma_M, \gamma_{\mathcal{I}}) \in AC \times \mathcal{P}$ with $\gamma(0) = A$ and $\gamma(t) = B$ for which*

$$\mathcal{J}_t[\gamma] = h_t(A,B).$$

Let $P = \big\{0 < t_1 < t_2 < \cdots < t_N < t\big\}$ be the partition associated with $\gamma_{\mathcal{I}}$. Then,

1. *γ_M is C^2 in $(0,t)\backslash\{t_1, t_2, \ldots, t_N\}$, and, in this set, it solves the Euler-Lagrange equation*

$$\frac{d}{ds}\left[\frac{\partial L}{\partial v}\big(\gamma_M(s), \dot\gamma_M(s), \gamma_{\mathcal{I}}(s)\big)\right] = \frac{\partial L}{\partial x}\big(\gamma_M(s), \dot\gamma_M(s), \gamma_{\mathcal{I}}(s)\big); \quad (4)$$

2. *Let $\dot\gamma_M(t_\ell^-)$ and $\dot\gamma_M(t_\ell^+)$ be the left and right derivatives of γ_M at t_ℓ, respectively. Then, for $\ell = 1, \ldots, N$,*

$$\frac{\partial L}{\partial v}\big(\gamma_M(t_\ell), \dot\gamma_M(t_\ell^-), \gamma_{\mathcal{I}}(t_\ell^-)\big) = \frac{\partial L}{\partial v}\big(\gamma_M(t_\ell), \dot\gamma_M(t_\ell^+), \gamma_{\mathcal{I}}(t_\ell^+)\big) + \partial_x \psi\big(\gamma_M(t_\ell), \gamma_{\mathcal{I}}(t_\ell^-), \gamma_{\mathcal{I}}(t_\ell^+)\big); \quad (5)$$

3. *The (generalized) energy functional*

$$E\big(\gamma_M(s), \dot\gamma_M(s), \gamma_{\mathcal{I}}(s)\big) := \frac{\partial L}{\partial v}\big(\gamma_M(s),\ \dot\gamma_M(s), \gamma_{\mathcal{I}}(s)\big) \cdot \dot\gamma_M(s) - L\big(\gamma_M(s), \dot\gamma_M(s), \gamma_{\mathcal{I}}(s)\big) \quad (6)$$

is constant on $(0,t)$.

Now, for $u : M \times \mathcal{I} \to \mathbb{R}$, we define $\Psi u : M \times \mathcal{I} \to \mathbb{R}$ as

$$\Psi u(x,i) = \min_{j \neq i} \big\{u(x,j) + \psi(x,i,j)\big\}. \quad (7)$$

The Hamiltonian, $H : T^*M \times \mathcal{I} \to \mathbb{R}$, is defined as

$$H(x,p,i) := \sup_{v \in T_x M} \big\{p(v) - L(x,v,i)\big\}. \quad (8)$$

Given a constant $c \in \mathbb{R}$, we consider the weakly coupled system of Hamilton-Jacobi equations

$$\max\Big\{H\big(x, du(x,i), i\big) - c,\ u(x,i) - \Psi u(x,i)\Big\} = 0 \quad \text{for every } i \in \mathcal{I}. \quad (9)$$

Definition 1.2. Let $u : M \times \mathcal{I} \to \mathbb{R}$ be a continuous function.

1. We say that u is a *viscosity subsolution* of (9) if for any test function $\phi \in C^1(M)$, and any $(x,i) \in M \times I$ such that $x \in \operatorname{argmax}_x u(\cdot, i) - \phi$, we have

$$\max\Big\{H\big(x, d\phi(x), i\big) - c, u(x,i) - \Psi u(x,i)\Big\} \leq 0. \tag{10}$$

2. We say that u is a *viscosity supersolution* of (9) if for any test function $\phi \in C^1(M)$, and any $(x,i) \in M \times I$ such that $x \in \operatorname{argmin}_x u(\cdot, i) - \phi$, we have

$$\max\Big\{H\big(x, d\phi(x), i\big) - c, u(x,i) - \Psi u(x,i)\Big\} \geq 0. \tag{11}$$

3. We say that u is a *viscosity solution* if it is both a viscosity subsolution and a viscosity supersolution.

Remark 1.3. A function $u : M \times \mathcal{I} \to \mathbb{R}$ is a viscosity subsolution of (9) if and only if the following two properties hold:

(i) $u(\cdot, i)$ satisfies $H\big(x, du(x,i), i\big) \leq c$ in the viscosity sense;

(ii) $u \leq \Psi u$.

In addition, $u : M \times \mathcal{I} \to \mathbb{R}$ is a viscosity supersolution of (9) if and only if *at least one* of the next two conditions holds:

(i) $u(\cdot, i)$ satisfies $H\big(x, du(x,i), i\big) \geq c$ in the viscosity sense;

(ii) $u \geq \Psi u$.

Our second main result is a weak KAM-type theorem, analog to the one of Fathi [12]:

Theorem 1.4 (Weak KAM). *Suppose that A1-A3 hold. Then, there exists a unique constant $c_0 \in \mathbb{R}$ for which there exists a viscosity solution $u : M \times \mathcal{I} \to \mathbb{R}$ of*

$$\max\Big\{H\big(x, du(x,i), i\big) - c_0, u(x,i) - \Psi u(x,i)\Big\} = 0 \qquad \forall\, i \in \mathcal{I}. \tag{12}$$

In addition, c_0 can be characterized as the unique constant such that the following holds: if (9) has a viscosity subsolution, then $c \geq c_0$.

Consider a Tonelli Lagrangian, L. For a continuous function $w : M \times \mathcal{I} \to \mathbb{R}$ and $t \geq 0$, the *Lax-Oleinik semigroup*, $T_t w : M \times \mathcal{I} \to \mathbb{R}$ is defined as

$$T_t w(A) := \inf\Big\{w(B) + h_t(B, A);\ B \in M \times \mathcal{I}\Big\}. \tag{13}$$

Inspired by [12], we prove Theorem 1.4 by looking at the long-time behavior of the Lax-Oleinik semigroup acting on a subsolution of (12). More precisely, we first show that u solves (12) if and only if it is a fixed point of $T_t u + c_0 t$ for all $t \geq 0$, see Proposition 4.7. Next, we analyze the long-time behavior of this Lax-Oleinik semigroup for any continuous initial condition $u_0 : M \times \mathcal{I} \to \mathbb{R}$, and show that long-time limit solves (12).

We define the *generalized Peierls barrier* $h : M \times \mathcal{I} \times M \times \mathcal{I} \to \mathbb{R}$ as

$$h(A, B) := \liminf_{t \to +\infty} h_t(A, B) + c_0 t, \tag{14}$$

and introduce the projected projected Aubry set as follows.

Definition 1.5. Let h be the Peierls barrier defined in (14). The *projected Aubry set* $\mathcal{A} \subset M \times \mathcal{I}$ is the set of points $B \in M \times \mathcal{I}$ for which $h(B, B) = 0$.

Then, we have the following convergence theorem:

Theorem 1.6. *Assume A1–A3 hold, let h be the Peierls barrier (14), and $u_0 : M \times \mathcal{I} \to \mathbb{R}$ be a continuous function. Then, as $t \to +\infty$, the Lax-Oleinik semigroup $T_t u_0$ converges to a viscosity solution v of (12). Moreover, v satisfies*

$$v(A) := \inf_{B \in \mathcal{A}}\Big\{h(B, A) + \inf_{C \in M \times \mathcal{I}}\Big\{u_0(C) + h(C, B)\Big\}\Big\}. \tag{15}$$

Structure of the paper. In Section 2, we discuss classical notions from the calculus of variations, including the lower semicontinuity of the action functional of Tonelli Lagrangians and the compactness for minimizers of this functional. These properties are used in the proof of Theorem 1.1. Next, in Section 3, we examine the optimal switching problem. We present the proof of Theorem 1.1 in Subsection 3.5, and we prove the semiconcavity of the cost function in Subsection 3.6. Finally, in Section 4, we present the proofs of Theorems 1.4 and 1.6: we prove the Weak KAM Theorem (Theorem 1.4) in Subsection 4.2, and we characterize the asymptotic limits of the Lax-Oleinik semigroup (Theorem 1.6) in Subsection 4.4.

2 Preliminaries

In this section, we recall some preliminary concepts. We define Tonelli Lagrangians and their action functional. Then, we state a compactness lemma and the lower semicontinuity of the action, whose proofs can be found in [13] and in [14, Appendix B]. Additional references include [1, 5, 8].

2.1 Definitions

Throughout this paper, (M, g) is a compact connected Riemannian manifold and TM its tangent bundle.

Definition 2.1 (Tonelli Lagrangian). $L : TM \to \mathbb{R}$ is a *Tonelli Lagrangian* if it satisfies the following conditions:

- *Regularity:* L belongs to $C^2(TM)$;
- *Convexity:* The second-order derivative in the fiber, $\dfrac{\partial^2 L}{\partial v^2}(x, v)$, is positive definite for every $(x, v) \in TM$;
- *Superlinearity:* For every $k \geq 0$ there exists a constant $C = C(k) \in \mathbb{R}$, such that

$$L(x, v) \geq k \|v\|_x - C \qquad \forall\, (x, v) \in TM,$$

  where $\|v\|_x := \sqrt{g_x(v, v)}$, g_x being the Riemannian metric on T_xM.

Consider a Tonelli Lagrangian L. The *action* of an absolutely continuous curve $\gamma : [a, b] \to M$ is

$$\mathcal{J}[\gamma] = \mathcal{J}_{a,b}[\gamma] := \int_a^b L\big(\gamma(s), \dot{\gamma}(s)\big)\, ds.$$

The action is well defined with values in $\mathbb{R} \cup \{+\infty\}$. An *action minimizer* is a curve $\gamma : [a, b] \to M$ satisfying $\mathcal{J}_{a,b}[\gamma] \leq \mathcal{J}_{a,b}[\alpha]$ for all absolutely continuous curves $\alpha : [a, b] \to M$ with the same endpoints. The next two propositions, stated here for convenience, are classical (see [13]).

Proposition 2.2 (Compactness). *Let $L : TM \to \mathbb{R}$ be a Lagrangian of class C^0 satisfying the Superlinearity condition. Suppose that $\gamma^k \in AC([0, t]; \mathbb{R}^d)$ is a sequence of curves such that $\sup_k \mathcal{J}[\gamma^k] < \infty$. Also, assume that $\sup_k \|\gamma_k(t_0)\| < \infty$ for some $t_0 \in [0, t]$. Then, up to a subsequence,*

1. *$\gamma^k \to \gamma$ uniformly for some $\gamma \in AC([0, t]; \mathbb{R}^d)$;*
2. *$\dot{\gamma}^k \rightharpoonup \dot{\gamma}$ weakly in L^1.*

Proposition 2.3 (Lower semicontinuity). *Suppose $L : TM \to \mathbb{R}$ is a Lagrangian of class C^0, bounded from below, with $L(x, \cdot) : T_xM \to \mathbb{R}$ convex for any $x \in M$. Assume $\gamma^k, \gamma \in AC([0, t]; M)$ are such that $\gamma^k \to \gamma$ uniformly in $[0, t]$ and $\dot{\gamma}^k \rightharpoonup \dot{\gamma}$ weakly in L^1. Then,*

$$\int_0^t L\big(\gamma(s), \dot{\gamma}(s)\big)\, ds \leq \liminf_k \int_0^t L\big(\gamma^k(s), \dot{\gamma}^k(s)\big)\, ds.$$

Tonelli's Theorem on the existence of minimizers is a corollary of the last two results.

Corollary 2.4 (Tonelli's Theorem). *Let $L : TM \to \mathbb{R}$ be a Tonelli Lagrangian. Then, for $x, y \in M$, there exists $\gamma \in AC([a, b]; M)$ such that $\gamma(a) = x$, $\gamma(b) = y$, and*

$$\mathcal{J}[\gamma] = \inf\big\{\mathcal{J}[\alpha];\ \alpha \in AC([a, b]; M) \text{ with } \alpha(a) = x,\ \alpha(b) = y\big\}.$$

Remark 2.5. If a Tonelli Lagrangian L is of class C^r, so is the minimizer (see [13]).

3 The optimal switching problem

Here, we study the optimal switching problem. In Subsection 3.1, we discuss the main assumptions. Next, in Subsection 3.2, we prove the existence of minimizers of $\mathcal{J}_t$. In Subsection 3.3, we obtain necessary optimality conditions for $(\gamma_M, \gamma_{\mathcal{I}})$ and show the regularity of γ_M. Then, in Subsection 3.4, we establish an energy conservation principle. In Subsection 3.5, we prove Theorem 1.1. Finally, in Subsection 3.6, we show that h_t is semiconcave.

3.1 Main assumptions

Let $L : TM \times \mathcal{I} \to \mathbb{R}$ be the Lagrangian and $\psi : M \times \mathcal{I} \times \mathcal{I} \to \mathbb{R}$ the switching cost. We suppose the following assumptions on L and ψ hold.

A1. For every $i \in \mathcal{I}$, $L(\cdot, \cdot, i)$ is a Tonelli Lagrangian, as in Definition 2.1;

A2. $\psi(\cdot, i, j)$ is continuous for all $i, j \in \mathcal{I}$, and satisfies the following inequality: for all $i, j, k \in \mathcal{I}$ with $i \neq k$ and $j \neq k$,

$$\psi(x, i, j) < \psi(x, i, k) + \psi(x, k, j) \qquad \forall\, x \in M;$$

A3. For any $i \in \mathcal{I}$, $\psi(\cdot, i, i) \equiv 0$;

A4. For all $i, j \in \mathcal{I}$, $\psi(\cdot, i, j) \in C^2(M)$.

Notice that we do not assume that the switching cost is nonnegative. Condition A2 is natural as it avoids that, to lower the action, a curve switches from i to j by using a double switch, over an infinitesimal amount of time, through an intermediate mode k. Condition A3 is also natural, as it states that the cost of not switching is zero.

Since M is compact and $\mathcal{I}$ is a finite set, it follows by A2 that there exists a constant $\delta > 0$ such that

$$\psi(x, i, k) + \psi(x, k, j) \geq \psi(x, i, j) + 2\delta \qquad \forall\, x \in M,\ i, j, k \in \mathcal{I} \text{ with } i \neq k \text{ and } j \neq k.$$

Hence, again by the compactness of M, since $\psi(\cdot, i, j)$ is continuous there exists $\eta > 0$ such that, for any $x, y \in M$, $i, j, k \in \mathcal{I}$ with $i \neq k$ and $j \neq k$,

$$\psi(x, i, k) + \psi(y, k, j) \geq \psi(x, i, j) + \delta, \qquad \text{provided } d_g(x, y) \leq \eta, \tag{16}$$

where $d_g : M \times M \to \mathbb{R}^+$ denotes the Riemannian distance on M.

Remark 3.1. Let $i_0, \dots, i_N \in \mathcal{I}$ be a sequence of states with $N \geq 2$, $i_0 = i_N$, and such that $i_j \neq i_{j+1}$. Let $x_0, \dots, x_N$ be a sequence of points such that $d_g(x_\ell, x_m) \leq \eta$ for all $\ell, m = 0, \dots, N-1$. Then, by applying (16) iteratively, we get

$$\begin{aligned} \psi(x_0, i_0, i_1) + \psi(x_1, i_1, i_2) + \cdots + \psi(x_{N-1}, i_{N-1}, i_N) &\geq \psi(x_0, i_0, i_2) + \psi(x_2, i_2, i_3) \\ + \cdots + \psi(x_{N-1}, i_{N-1}, i_N) + \delta &\geq \psi(x_0, i_0, i_{N-1}) + \psi(x_{N-1}, i_{N-1}, i_N) + (N-2)\delta \\ &\geq \psi(x_0, i_0, i_N) + (N-1)\delta \geq \delta, \end{aligned}$$

where the last inequality follows from A3 (since, by $i_0 = i_N$) and the fact that $N \geq 2$.

Hence, we proved that a loop in the state space always has a nontrivial cost greater than δ, provided all the switches happen at sufficiently nearby points.

Remark 3.2. Since $\mathcal{I}$ is a finite set, $L(\cdot, \cdot, i)$ is superlinear, uniformly with respect to $i \in \mathcal{I}$. In particular, we can find a superlinear, increasing, convex function $\theta : \mathbb{R}^+ \to \mathbb{R}^+$ with $\theta(0) = 0$, and a constant $\bar{C}_0 \in \mathbb{R}$, such that

$$L(x, v, i) \geq \theta(\|v\|_x) - \bar{C}_0 \qquad \forall\, (x, v, i) \in TM \times \mathcal{I}.$$

3.2 Existence of action minimizers

For fixed $t > 0$, define $\mathcal{J}_t$ and h_t as in (2) and (3), respectively. Observe that h_t is finite: given $A = (x, i)$ and $B = (y, j)$, consider a geodesic from x to y in M and one switch from i to j at an intermediate point. Because this curve has finite action, h_t is finite.

Using the direct method in the calculus of variations, we prove the existence of minimizers for $\mathcal{J}_t$ under fixed boundary conditions. We notice that the proof never requires $L(\cdot, \cdot, i)$ to be Tonelli, but only superlinearity and convexity in v, as in Propositions 2.2 and 2.3.

Theorem 3.3. *Assume A1–A3 hold, and let h_t be the cost function defined by (3). Then, for every $A, B \in M \times \mathcal{I}$, there exists $\gamma \in AC \times \mathcal{P}$ with $\gamma(0) = A$ and $\gamma(t) = B$ such that*

$$h_t(A,B) = \mathcal{J}_t[\gamma].$$

Proof. Let $\gamma^k = (\gamma_M^k,\ \gamma_{\mathcal{I}}^k)$ be a minimizing sequence for $\mathcal{J}_t$ with fixed endpoints: $\gamma^k(0) = A = (x,i)$, $\gamma^k(t) = B = (y,j)$, and

$$\mathcal{J}_t[\gamma^k] \to h_t(A,B). \tag{17}$$

Consider the partitions $P^k = \{0 < t_1^k < \dots < t_{N_k}^k < t\}$ associated with $\gamma_{\mathcal{I}}^k$.

Step 1. *We show that the total number of switches, N_k, is uniformly bounded with respect to k.*

Using Remark 3.2 and (17), we see that

$$\int_0^t \theta(\|\dot\gamma_M^k(s)\|_{\gamma_M^k(s)})\,ds + \psi\big(\gamma_M^k(0),\ \gamma_{\mathcal{I}}^k(0),\ \gamma_{\mathcal{I}}^k(0^+)\big) + \sum_{\ell=1}^{N_k} \psi\big(\gamma_M^k(t_\ell^k),\ \gamma_{\mathcal{I}}^k((t_\ell^k)^-),\ \gamma_{\mathcal{I}}^k((t_\ell^k)^+)\big)$$
$$+\, \psi\big(\gamma_M^k(t),\ \gamma_{\mathcal{I}}^k(t^-),\ \gamma_{\mathcal{I}}^k(t)\big) \le \mathcal{J}_t[\gamma^k] + \bar C_0 t \le \bar C_1 \tag{18}$$

for some constant $\bar C_1$ independent of k. We take η as in (16) and, for any k, we define a sequence of times $\{\tau_m^k\}_{m=0}^{\hat N_k}$ as follows:

$$\tau_0^k = 0, \qquad \tau_{\hat N_k} = t, \qquad \tau_{m+1}^k = \max\{s \in (\tau_m^k, t) \,:\, d_g\big(\gamma_M^k(s), \gamma_M^k(\tau_m^k)\big) \le \eta/2\} \quad \forall\, m = 0, \dots, \hat N_k - 2.$$

In other words, τ_{m+1}^k tells us when the curve $\gamma_M^k(s)$ exits the ball $B_{\eta/2}(\gamma_M^k(\tau_m^k))$. In particular, for any time interval $I_m^k := [\tau_m^k, \tau_{m+1}^k)$, it follows by the definition of I_m^k that

$$d_g\big(\gamma_M^k(s), \gamma_M^k(s')\big) \le \eta \qquad \forall\, s, s' \in I_m^k.$$

Thus, Remark 3.1 implies that every loop inside $\{\gamma_I(t_\ell^k)\}_{t_\ell^k \in I_m^k}$ has a total switching cost that is greater than δ. Consequently, consider the switching cost

$$\sum_{t_\ell^k \in I_m^k} \psi\big(\gamma_M^k(t_\ell^k),\ \gamma_{\mathcal{I}}^k((t_\ell^k)^-),\ \gamma_{\mathcal{I}}^k((t_\ell^k)^+)\big).$$

Every time we find a loop inside $\{\gamma_I(t_\ell^k)\}_{t_\ell^k \in I_m^k}$, we can remove it from the above sum, and the total switching cost will decrease by δ. Thus, we argue as follows. First, to simplify the notation, we set $I_0 := \#\mathcal{I}$ and $N_{k,m} := \#\{t_\ell^k \in I_m^k\}$. If $N_{k,m} \le I_0$, we do nothing and bound the switching cost by

$$\sum_{t_\ell^k \in I_m^k} \psi\big(\gamma_M^k(t_\ell^k),\ \gamma_{\mathcal{I}}^k((t_\ell^k)^-),\ \gamma_{\mathcal{I}}^k((t_\ell^k)^+)\big) \ge -I_0\,\|\psi\|_\infty.$$

Instead, if $N_{k,m} > I_0$, we consider the first $I_0 + 1$ indices inside the set $\{\gamma_{\mathcal{I}}(t_\ell^k) \,:\, t_\ell^k \in I_m^k\}$; that is, the indices corresponding to the first $I_0 + 1$ times. Then, we note that this set has to contain a loop because inside a set containing $I_0 + 1$ indices, at least two of them coincide. Observe that, by construction, this loop will have length at most $I_0 + 1$. We remove this loop; thus, the switching cost decreases by at least δ. Then, we repeat this procedure: we consider again the first $I_0 + 1$ times among the ones that remain, we find a loop there, and we remove it. We iterate this procedure until no loop remains. When this is achieved, the number of times remaining must be bounded by I_0 (otherwise, we could find another loop inside). We recall that the loops have length bounded by $I_0 + 1$. Because each time we remove a loop the set of times decreases by at most I_0 and because at the end we are left with at most I_0 times, the number of loops we remove from the set $\{\gamma_{\mathcal{I}}(t_\ell^k) \,:\, t_\ell^k \in I_m^k\}$ is at least $\left(\frac{N_{k,m}}{I_0} - 1\right)_+$. Hence, bounding the switching cost of the remaining times as in the case $N_{k,m} \le I_0$, we obtain

$$\sum_{t_\ell^k \in I_m^k} \psi\big(\gamma_M^k(t_\ell^k),\ \gamma_{\mathcal{I}}^k((t_\ell^k)^-),\ \gamma_{\mathcal{I}}^k((t_\ell^k)^+)\big) \ge \left(\frac{N_{k,m}}{I_0} - 1\right)_+ \delta - I_0\,\|\psi\|_\infty.$$

Notice that the bound above is valid both in the case $N_{k,m} > I_0$ and in the case $N_{k,m} \le I_0$. Therefore, summing this bound over $m \in \{0, \dots, \hat N_k - 1\}$, we conclude that

$$\sum_{\ell=1}^{N} \big(\gamma_M^k(t_\ell^k),\ \gamma_{\mathcal{I}}^k((t_\ell^k)^-),\ \gamma_{\mathcal{I}}^k((t_\ell^k)^+)\big) \ge \sum_{m=0}^{\hat N_k - 1} \left(\frac{N_{k,m}}{I_0} - 1\right)_+ \delta - \hat N_k\, I_0\,\|\psi\|_\infty.$$

The preceding estimate combined with (18) yields

$$\int_0^t \theta(\|\dot\gamma_M^k(s)\|_{\gamma_M^k(s)})\,ds - 2\,\|\psi\|_\infty + \sum_{m=0}^{\hat N_k-1}\left(\frac{N_{k,m}}{I_0}-1\right)_+\delta - \hat N_k\, I_0\,\|\psi\|_\infty \le \bar C_1. \tag{19}$$

To deal with the first term above, we apply Jensen's inequality to the increasing, convex function θ on each time interval I_m^k. Set $\Delta_m^k := |I_m^k| = \tau_{m+1}^k - \tau_m^k$. Then,

$$\begin{aligned}\int_0^t \theta(\|\dot\gamma_M^k(s)\|_{\gamma_M^k(s)})\,ds &= \sum_{m=0}^{\hat N_k-1}\int_{\tau_m^k}^{\tau_{m+1}^k}\theta(\|\dot\gamma_M^k(s)\|_{\gamma_M^k(s)})\,ds \ge \sum_{m=0}^{\hat N_k-1}\Delta_m^k\,\theta\left(\frac{1}{\Delta_m^k}\int_{\tau_m^k}^{\tau_{m+1}^k}\|\dot\gamma_M^k(s)\|_{\gamma_M^k(s)}\,ds\right)\\ &\ge \sum_{m=0}^{\hat N_k-1}\Delta_m^k\,\theta\left(\frac{d_g\big(\gamma_M(\tau_m^k),\gamma_M(\tau_{m+1}^k)\big)}{\Delta_m^k}\right) \ge \sum_{m=0}^{\hat N_k-2}\Delta_m^k\,\theta\left(\frac{\eta}{2\Delta_m^k}\right),\end{aligned}$$

where, for the last inequality, we used that $\theta \ge 0$ and that, from the definition of τ_j^k,

$$d_g\big(\gamma_M(\tau_m^k),\gamma_M(\tau_{m+1}^k)\big) = \eta/2 \qquad \forall\, j = 0,\dots,\hat N_k - 2.$$

We now set $\Delta^k := \sum_{m=0}^{\hat N_k-2}\Delta_m^k$. Notice that $\Delta^k \le t$. Then, by the superlinearity of θ, for any $\kappa > 0$ there exists a constant C_κ such that

$$\sum_{m=0}^{\hat N_k-2}\Delta_m^k\,\theta\left(\frac{\eta}{2\Delta_m^k}\right) \ge \sum_{m=0}^{\hat N_k-2}\Delta_m^k\left(\kappa\frac{\eta}{2\Delta_m^k} - C_\kappa\right) = \frac{\kappa\,\eta}{2}\big(\hat N_k - 1\big) - C_\kappa\,\Delta \ge \frac{\kappa\,\eta}{2}\big(\hat N_k - 1\big) - C_\kappa\, t.$$

Inserting this estimate in (19), we finally get

$$\sum_{m=0}^{\hat N_k-1}\left(\frac{N_{k,m}}{I_0}-1\right)_+\delta + \frac{\kappa\,\eta}{2}\big(\hat N_k - 1\big) - \hat N_k\, I_0\,\|\psi\|_\infty \le \bar C_1 + C_\kappa\, t + 2\,\|\psi\|_\infty.$$

Choosing κ large enough, we can reabsorb the negative term in the left-hand side into the second term, and we get

$$\sum_{m=0}^{\hat N_k-1}\left(\frac{N_{k,m}}{I_0}-1\right)_+\delta + \frac{\kappa\,\eta}{4}\hat N_k \le \bar C_2 \tag{20}$$

for some constant, $\bar C_2$, independent of k. Thanks to this bound, we deduce that $\hat N_k$, and, therefore, the total number of switches, $N_k = \sum_{m=0}^{\hat N_k-1} N_{k,m}$, is bounded independently of k.

Step 2. *We find a candidate minimizer.*

By the previous step and extracting a subsequence, if necessary, we can assume $N_k \equiv N$ for all k. Then,

$$P^k := \{0 < t_1^k < \dots < t_N^k < t\}. \tag{21}$$

For every $\ell = 1,\dots,N$, $(t_\ell^k)_k$ is a bounded sequence; so, up to a subsequence, it converges to some $t_\ell \in [0,t]$. We have $0 \le t_1 \le \dots \le t_N \le t$. Removing repeated points, if necessary, we obtain a limiting partition $\{0 = t_0^* < t_1^* < t_2^* < \ldots < t_m^* < t_{m+1}^* = t\}$. Observe that instantaneous jumps at the endpoints can occur.

We then define $\gamma_{\mathcal I}$ as follows: for $\varepsilon > 0$ small and k sufficiently large, we have that $\gamma_{\mathcal I}^k$ is constant in each subinterval $(0, t_1^* - \varepsilon), (t_1^* + \varepsilon, t_2^* - \varepsilon), \ \dots, (t_{m-1}^* + \varepsilon, t_m^* - \varepsilon), (t_m^* + \varepsilon, t)$. Hence, up to a subsequence, the values of $\gamma_{\mathcal I}^k$ will converge to some value $i_\ell \in \mathcal I$ inside the interval $(t_{\ell-1}^* + \varepsilon, t_\ell^* - \varepsilon)$, and we get

$$\gamma_{\mathcal I}^k \to \gamma_{\mathcal I} = \sum_{\ell=0}^{m} i_\ell\,\chi_{(t_\ell^*, t_{\ell+1}^*)} \tag{22}$$

pointwise inside $\cup_{\ell=0}^m (t_\ell^*, t_{\ell+1}^*)$. Also, recalling that by assumption $\gamma_{\mathcal I}^k(0) = i$ and $\gamma_{\mathcal I}^k(t) = j$, we set $\gamma_{\mathcal I}(0) = i$ and $\gamma_{\mathcal I}(t) = j$.

Concerning the curves γ_M^k, we note that, since the number of switching times is bounded, it follows from (18) that

$$\int_0^t \theta(\|\dot\gamma_M^k(s)\|_{\gamma_M^k(s)})\,ds \le \hat C_1$$

for some constant $\hat{C}_1$ independent of k. Since θ is superlinear, it follows from Proposition 2.2 applied with $L(x,v) = \theta(\|v\|_x)$ that, up to a subsequence, there exists an absolutely continuous curve $\gamma_M : [0,t] \to M$ such that

$$\gamma_M^k \to \gamma_M \text{ uniformly in } [0,t], \quad \text{and} \quad \dot{\gamma}_M^k \rightharpoonup \dot{\gamma}_M \text{ weakly in } L^1. \tag{23}$$

Step 3. *Lower semicontinuity and conclusion.*

Since L is bounded from below, we have

$$\begin{aligned}\int_0^t L(\gamma_M^k(s), \dot{\gamma}_M^k(s), \gamma_{\mathcal{I}}^k(s))\, ds &\geq \int_0^{t_1^*-\varepsilon} L(\gamma_M^k(s), \dot{\gamma}_M^k(s), \gamma_{\mathcal{I}}(0))\, ds \\ &+ \int_{t_1^*+\varepsilon}^{t_2^*-\varepsilon} L(\gamma_M^k(s), \dot{\gamma}_M^k(s), \gamma_{\mathcal{I}}(t_1^*))\, ds + \ldots + \int_{t_m^*+\varepsilon}^{t} L(\gamma_M^k(s), \dot{\gamma}_M^k(s), \gamma_{\mathcal{I}}(t_m^*))\, ds - C_0\,\varepsilon\end{aligned} \tag{24}$$

Note that some intervals $[t_\ell^k, t_{\ell+1}^k]$ can collapse when $k \to \infty$, and, in that case, we are removing the corresponding switches. Then, the triangle inequality $A2$ and the continuity of ψ give

$$\begin{aligned}&\psi\big(x, i, \gamma_{\mathcal{I}}^k(0^+)\big) + \sum_{\ell=1}^{N} \psi\big(\gamma_M^k(t_\ell^k), \gamma_{\mathcal{I}}^k((t_\ell^k)^-), \gamma_{\mathcal{I}}^k((t_\ell^k)^+)\big) + \psi\big(y, \gamma_{\mathcal{I}}^k(t^-), j\big) + \varepsilon \\ &\qquad \geq \psi\big(x, i, \gamma_{\mathcal{I}}(0^+)\big) + \sum_{\ell=1}^{m} \psi\big(\gamma_M(t_\ell^*), \gamma_{\mathcal{I}}((t_\ell^*)^-), \gamma_{\mathcal{I}}((t_\ell^*)^+)\big) + \psi\big(y, \gamma_{\mathcal{I}}(t^-), j\big)\end{aligned} \tag{25}$$

for all k sufficiently large. Next, we add (24) and (25), and, by taking the lim inf and using the lower semicontinuity (Proposition 2.3) with respect to the convergence in (23), we obtain

$$\begin{aligned}h_t(A,B) &= \liminf_k \left(\int_0^t L(\gamma_M^k(s), \dot{\gamma}_M^k(s), \gamma_{\mathcal{I}}^k(s))\, ds + \sum_{\ell=0}^{N-1} \psi\big(\gamma_M(t_{\ell+1}^k), \gamma_{\mathcal{I}}(t_\ell^k), \gamma_{\mathcal{I}}(t_{\ell+1}^k)\big) \right) \\ &\geq \int_0^{t_1^*-\varepsilon} L(\gamma_M(s), \dot{\gamma}_M(s), \gamma_{\mathcal{I}}(0))\, ds + \int_{t_1^*+\varepsilon}^{t_2^*-\varepsilon} L(\gamma_M(s), \dot{\gamma}_M(s), \gamma_{\mathcal{I}}(t_1^*))\, ds + \ldots + \\ &\quad + \int_{t_m^*+\varepsilon}^{t} L\big(\gamma_M(s), \dot{\gamma}_M(s), \gamma_{\mathcal{I}}(t_m^*)\big)\, ds + \psi\big(x, i, \gamma_{\mathcal{I}}(0^+)\big) + \sum_{\ell=1}^{m} \psi\big(\gamma_M(t_\ell^*), \gamma_{\mathcal{I}}((t_\ell^*)^-), \gamma_{\mathcal{I}}((t_\ell^*)^+)\big) \\ &\quad + \psi\big(y, \gamma_{\mathcal{I}}(t^-), j\big) - C_1\,\varepsilon\end{aligned} \tag{26}$$

for all $\varepsilon > 0$. Hence, the limit $\varepsilon \to 0$ yields

$$\begin{aligned}h_t(A,B) &\geq \int_0^t L(\gamma_M(s), \dot{\gamma}_M(s), \gamma_{\mathcal{I}}(s))\, ds + \psi\big(x, i, \gamma_{\mathcal{I}}(0^+)\big) \\ &\quad + \sum_{\ell=0}^{m-1} \psi\big(\gamma_M(t_\ell^*), \gamma_{\mathcal{I}}((t_\ell^*)^-), \gamma_{\mathcal{I}}((t_\ell^*)^+)\big) + \psi\big(y, \gamma_{\mathcal{I}}(t^-), j\big) = \mathcal{J}_t[\gamma].\end{aligned} \tag{27}$$

Therefore, $\gamma = (\gamma_M, \gamma_{\mathcal{I}})$ is a minimizer of $\mathcal{J}_t$. □

3.3 Euler-Lagrange equations

In this subsection, we obtain necessary conditions for minimality. First, we show that the Euler-Lagrange equation is satisfied, except possibly where $\gamma_{\mathcal{I}}$ has a jump. Furthermore, if $\gamma_{\mathcal{I}}$ has a jump, then (5) holds. Next, we obtain regularity for the minimizers of the optimal switching problem.

Proposition 3.4 (Euler-Lagrange equations). *Assume that A1–A3 hold. Let $\gamma = (\gamma_M, \gamma_{\mathcal{I}})$ be a minimizer of $\mathcal{J}_t$ and let $P = \{0 < t_1 < \cdots < t_N < t\}$ be the partition associated with $\gamma_{\mathcal{I}}$. Then, the Euler-Lagrange equations (4) hold in $[0,t] \backslash \{t_1, t_2, \ldots, t_N\}$.*

Proof. Observe that $\gamma_M|_{[t_\ell, t_{\ell+1}]}$ is a minimizer of

$$\int_{t_\ell}^{t_{\ell+1}} L\big(\alpha(s), \dot{\alpha}(s), \gamma_{\mathcal{I}}(t_\ell^+)\big)\, ds,$$

among all curves $\alpha \in AC([t_\ell, t_{\ell+1}], M)$ with $\alpha(t_\ell) = \gamma_M(t_\ell)$, $\alpha(t_{\ell+1}) = \gamma_M(t_{\ell+1})$. Then, (4) follows from the classical Euler-Lagrange equations for the Lagrangian $L(\cdot, \cdot, \gamma_{\mathcal{I}}(t_\ell))$. □

Remark 2.5 implies the following corollary.

Corollary 3.5. *Assume that A1–A3 hold. Let $\gamma = (\gamma_M, \gamma_{\mathcal{I}})$ be a minimizer of $\mathcal{J}_t$. Then γ_M is of class C^2 in $U = (0,t)\backslash\{t_1, \dots, t_N\}$, and the left and right derivatives $\dot\gamma_M(t_\ell^-)$ and $\dot\gamma_M(t_\ell^+)$ exist for every $0 \le \ell \le N+1$. Moreover, if all Lagrangians $L(\cdot,\cdot,i)$ are of class C^r, then so is γ_M in U.*

Now, by the Euler-Lagrange equations, we obtain further necessary minimality conditions.

Proposition 3.6 (Necessary conditions for minimality II). *Assume that A1–A4 hold, and let $\gamma = (\gamma_M, \gamma_{\mathcal{I}}) \in AC \times \mathcal{P}$ be a minimizer for the action $\mathcal{J}_t$ given by (2). Then (5) holds for $\ell = 1, \dots, N$.*

Proof. Let $\bar t = \min\{t_\ell - t_{\ell-1}, t_{\ell+1} - t_\ell\}$ and fix $\omega \in C^\infty([0,t]; M)$ such that $\omega \equiv 1$ in $[t_\ell - \bar t/2, t_\ell + \bar t/2]$ and $\omega \equiv 0$ outside $[t_\ell - \bar t, t_\ell + \bar t]$. Then,

$$\begin{aligned}
0 &= \frac{d}{d\sigma}|_{\sigma=0}\, \mathcal{J}_t[\gamma_M + \sigma\omega, \gamma_{\mathcal{I}}] = \int_{t_{\ell-1}}^{t_\ell} \left(\frac{\partial L}{\partial x}(\gamma_M, \dot\gamma_M, \gamma_{\mathcal{I}})\omega + \frac{\partial L}{\partial v}(\gamma_M, \dot\gamma_M, \gamma_{\mathcal{I}})\dot\omega \right) ds \\
&\quad + \int_{t_\ell}^{t_{\ell+1}} \left(\frac{\partial L}{\partial x}(\gamma_M, \dot\gamma_M, \gamma_{\mathcal{I}})\omega + \frac{\partial L}{\partial v}(\gamma_M, \dot\gamma_M, \gamma_{\mathcal{I}})\dot\omega \right) ds + \partial_x\psi\big(\gamma_M(t_\ell), \gamma_{\mathcal{I}}(t_\ell^-), \gamma_{\mathcal{I}}(t_\ell^+)\big)\omega(t_\ell) \\
&= \int_{t_{\ell-1}}^{t_{\ell+1}} \left(\frac{\partial L}{\partial x}(\gamma_M, \dot\gamma_M, \gamma_{\mathcal{I}}) - \frac{d}{ds}\frac{\partial L}{\partial v}(\gamma_M, \dot\gamma_M, \gamma_{\mathcal{I}}) \right)\omega \; ds + \frac{\partial L}{\partial v}\big(\gamma_M(t_\ell), \dot\gamma_M(t_\ell^-), \gamma_{\mathcal{I}}(t_\ell^-)\big) \\
&\quad - \frac{\partial L}{\partial v}\big(\gamma_M(t_\ell), \dot\gamma_M(t_\ell^+), \gamma_{\mathcal{I}}(t_\ell^+)\big) + \partial_x\psi\big(\gamma_M(t_\ell), \gamma_{\mathcal{I}}(t_\ell^-), \gamma_{\mathcal{I}}(t_\ell^+)\big) \\
&= \frac{\partial L}{\partial v}\big(\gamma_M(t_\ell), \dot\gamma_M(t_\ell^-), \gamma_{\mathcal{I}}(t_\ell^-)\big) - \frac{\partial L}{\partial v}\big(\gamma_M(t_\ell), \dot\gamma_M(t_\ell^+), \gamma_{\mathcal{I}}(t_\ell^+)\big) + \partial_x\psi\big(\gamma_M(t_\ell), \gamma_{\mathcal{I}}(t_\ell^-), \gamma_{\mathcal{I}}(t_\ell^+)\big),
\end{aligned} \tag{28}$$

where we used that, by construction, $\omega(t_\ell) = 1$ and $\omega(t_{\ell-1}) = \omega(t_{\ell+1}) = 0$, while the last identity follows from the Euler-Lagrange equations (4). □

Remark 3.7. Let ψ be a switching cost independent of the state variable; that is, $\psi(x,i,j) \equiv \psi(i,j)$ for any $x \in M$ and $i,j \in \mathcal{I}$. By Proposition 3.6, if γ is an action minimizer,

$$s \mapsto \frac{\partial L}{\partial v}\big(\gamma_M(s), \dot\gamma_M(s), \gamma_{\mathcal{I}}(s)\big) \text{ is continuous in } [0,t].$$

3.4 Energy Conservation

Here, we prove an energy conservation principle similar to the one for single Lagrangians.

Proposition 3.8 (Conservation of Energy). *Suppose that A1–A3 hold and let $\gamma = (\gamma_M, \gamma_{\mathcal{I}}) \in AC \times \mathcal{I}$ be an action minimizer for $\mathcal{J}_t$. Set $\gamma_{\mathcal{I}}(t_\ell) := \gamma_{\mathcal{I}}(t_\ell^-)$ for any switching time t_ℓ. Then the energy is conserved; that is,*

$$E(\gamma_M(s), \dot\gamma_M(s), \gamma_{\mathcal{I}}(s)) := \frac{\partial L}{\partial v}\big(\gamma_M(s), \dot\gamma_M(s), \gamma_{\mathcal{I}}(s)\big) \cdot \dot\gamma_M(s) - L\big(\gamma_M(s), \dot\gamma_M(s), \gamma_{\mathcal{I}}(s)\big) \tag{29}$$

is constant on $(0,t)$.

Proof. The energy is conserved on each subinterval $(t_\ell, t_{\ell+1})$ by (4). So, it is enough to check the energy conservation at the switching times. Let $\omega : [0,t] \to [0,t]$ be smooth and compactly supported in $[t_{\ell-1}, t_{\ell+1}]$ with $\omega(t_{\ell-1}) = \omega(t_{\ell+1}) = 0$ and $\omega \equiv 1$ in $[t_\ell - \delta, t_\ell + \delta]$ for some small $\delta > 0$. Define $\gamma^\varepsilon : [0,t] \to M \times \mathcal{I}$ as $\gamma^\varepsilon(s) = \gamma\big(s - \varepsilon\omega(s)\big)$, so that the function $f(\varepsilon) := \mathcal{J}_t[\gamma_\varepsilon] - \mathcal{J}_t[\gamma]$ has a minimum at $\varepsilon = 0$. Because γ_ε switches from one state to another at the same position as γ (just at a different time), this does not affect the switching cost. In other words, the part in the energy of γ_ε coming from the switching cost is the same as the one of γ. Therefore,

$$\begin{aligned}
f(\varepsilon) = \int_{t_{\ell-1}}^{t_\ell+\varepsilon} L\big(\gamma_M^\varepsilon(s), \dot\gamma_M^\varepsilon(s), \gamma_{\mathcal{I}}(t_\ell^-)\big)\, ds + \int_{t_\ell+\varepsilon}^{t_{\ell+1}} L\big(\gamma_M^\varepsilon(s), \dot\gamma_M^\varepsilon(s), \gamma_{\mathcal{I}}(t_\ell^+)\big)\, ds \\
- \int_{t_{\ell-1}}^{t_\ell} L\big(\gamma_M(s), \dot\gamma_M(s), \gamma_{\mathcal{I}}(t_\ell^-)\big)\, ds - \int_{t_\ell}^{t_{\ell+1}} L\big(\gamma_M(s), \dot\gamma_M(s), \gamma_{\mathcal{I}}(t_\ell^+)\big)\, ds,
\end{aligned}$$

and, by differentiating with respect to ε and setting $\varepsilon = 0$, we get

$$\begin{aligned}
0 = L\big(\gamma_M(t_\ell), \dot\gamma_M(t_\ell^-), \gamma_{\mathcal{I}}(t_\ell^-)\big) + \int_{t_{\ell-1}}^{t_\ell} \bigg(\frac{\partial L}{\partial x}\big(\gamma_M(s), \dot\gamma_M(s), \gamma_{\mathcal{I}}(t_\ell^-)\big) \cdot \dot\gamma_M \omega \\
+ \frac{\partial L}{\partial v}\big(\gamma_M(s), \dot\gamma_M(s), \gamma_{\mathcal{I}}(t_\ell^-)\big) \cdot (\ddot\gamma_M \omega + \dot\gamma_M \dot\omega) \bigg)\, ds - L\big(\gamma_M(t_\ell), \dot\gamma_M(t_\ell^+), \gamma_{\mathcal{I}}(t_\ell^+)\big) \\
- \int_{t_{\ell-1}}^{t_\ell} \bigg(\frac{\partial L}{\partial x}\big(\gamma_M(s), \dot\gamma_M(s), \gamma_{\mathcal{I}}(t_\ell^+)\big) \cdot \dot\gamma_M \omega + \frac{\partial L}{\partial v}\big(\gamma_M(s), \dot\gamma_M(s), \gamma_{\mathcal{I}}(t_\ell^+)\big) \cdot (\ddot\gamma_M \omega + \dot\gamma_M \dot\omega) \bigg)\, ds.
\end{aligned} \tag{30}$$

Integrating by parts and using the Euler-Lagrange equations (4), we get

$$\begin{aligned}
L\big(\gamma_M(t_\ell), \dot\gamma_M(t_\ell^-), \gamma_{\mathcal{I}}(t_\ell^-)\big) - \frac{\partial L}{\partial v}\big(\gamma_M(t_\ell), \dot\gamma_M(t_\ell^-), \gamma_{\mathcal{I}}(t_\ell^-)\big) \cdot \dot\gamma(t_\ell^-) \\
= L\big(\gamma_M(t_\ell), \dot\gamma_M(t_\ell^+), \gamma_{\mathcal{I}}(t_\ell^+)\big) - \frac{\partial L}{\partial v}\big(\gamma_M(t_\ell), \dot\gamma_M(t_\ell^+), \gamma_{\mathcal{I}}(t_\ell^+)\big) \cdot \dot\gamma(t_\ell^+),
\end{aligned}$$

as claimed. □

3.5 Proof of Theorem 1.1

Here, we gather the results from previous subsections to conclude the proof of Theorem 1.1.

Proof of Theorem 1.1. The existence of minimizers under fixed endpoint constraint is given in Theorem 3.3. Next, the Euler-Lagrange equations are proved in Proposition 3.4. Then, (5) corresponds to Proposition 3.6. Finally, the energy conservation follows from Proposition 3.8. □

3.6 Semiconcavity of the cost

In this subsection, we study the regularity of the value function. For convenience, we recall the definition of a semiconcave function (for example, see [5]).

Definition 3.9. Let $A \subseteq \mathbb{R}^d$. A function $f : A \to \mathbb{R}$ is *semiconcave* if there exists a modulus of continuity $\omega : \mathbb{R}^+ \to \mathbb{R}$ such that, for each $x \in A$, there exists $\ell_x \in \mathbb{R}^d$ for which

$$f(y) - f(x) \le \ell_x \cdot (y - x) + |y - x|\, \omega\big(|y - x|\big) \qquad \text{for all } y \in A.$$

If M is Riemannian manifold, we say that a function $f : M \to \mathbb{R}$ is semiconcave if, in any chart $\varphi : A \to \varphi(A) \subseteq M$ with $A \subseteq \mathbb{R}^d$, the function $f \circ \varphi : A \to \mathbb{R}$ is semiconcave.

Semiconcavity is well defined in the Riemannian setting because the composition with smooth functions preserves semiconcavity, see [5, 14].

Remark 3.10. We say that $f : A \to \mathbb{R}$ is semiconcave with a linear modulus of continuity if there exists $C > 0$ such that

$$f(y) - f(x) \le \ell_x \cdot (y - x) + C|y - x|^2 \qquad \text{for all } y \in A.$$

To study the semiconcavity of h, we use the conservation of energy to bound the norms of speed curves of minimizers. We consider this matter in the next lemma:

Lemma 3.11. *Assume that A1–A3 hold, and let $\gamma = (\gamma_M, \gamma_{\mathcal{I}})$ be a minimizer for $\mathcal{J}_t$ with $\gamma_M(0) = x$, $\gamma_M(t) = y$. Then, there exists a constant $C > 0$ such that $\|\dot\gamma_M(s)\|_{\gamma_M(s)} \le C$ for every $s \in [0, t]$.*

Proof. Fix a point $x_0 \in M$. Since M is compact, $d(x, y) \le R$ for all $x, y \in M$ for some $R > 0$. Thus, the minimizing constant speed geodesic $\alpha : [0, t] \to M$ connecting x to y satisfies

$$\|\dot\alpha(s)\|_{\alpha(s)} = \frac{d(x, y)}{t} \le \frac{R}{t}.$$

Note that the Lagrangian $L(\cdot, \cdot, i)$ is bounded on the compact set

$$\mathcal{L} := \left\{ (z, v) \ : \|v\|_z \le \frac{R}{t} \right\} \subset TM$$

for all $i \in \mathcal{I}$. Thus, there exists $B \geq 0$ such that $L(z, v, i) \leq B$ for every $(z, v, i) \in \mathcal{L} \times \mathcal{I}$. Then,

$$\int_0^t L\big(\gamma_M(s), \dot\gamma_M(s), \gamma_{\mathcal{I}}(s)\big)\ ds \leq \int_0^t L\big(\alpha(s), \dot\alpha(s), \gamma_{\mathcal{I}}(s)\big)\ ds \leq Bt. \tag{31}$$

According to the third condition in Definition 2.1, there exists $C \in \mathbb{R}$ such that

$$\int_0^t \|\dot\gamma_M(s)\|_{\gamma_M(s)}\ ds - Ct \leq \int_0^t L\big(\gamma_M(s), \dot\gamma_M(s), \gamma_{\mathcal{I}}(s)\big)\ ds. \tag{32}$$

Combining (31) and (32), we obtain

$$\frac{1}{t}\int_0^t \|\dot\gamma_M(s)\|_{\gamma_M(s)}\ ds \leq B + C =: C_1, \tag{33}$$

which implies the existence of $\bar s \in (0, t)$ such that $\|\dot\gamma_M(\bar s)\|_{\gamma_M(\bar s)} \leq C_1$.

Set

$$\theta := \max\big\{E(x, v, i)\ :\ (x, v, i) \in TM \times \mathcal{I},\ x \in M\ ,\ \|v\|_x \leq C_1\big\}.$$

Then $E\big(\gamma_M(\bar s), \dot\gamma_M(\bar s), \gamma_{\mathcal{I}}(\bar s)\big) \leq \theta$ and, by the conservation of energy (Proposition 3.8), we obtain $E\big(\gamma_M(s), \dot\gamma_M(s), \gamma_{\mathcal{I}}(s)\big) \leq \theta$ for every $s \in [0, t]$. Therefore $(\gamma_M, \dot\gamma_M)$ is contained in the compact set $\big\{(x, v) \in TM\ ;\ x \in M,\ E(x, v, i) \leq \theta\ \forall i\big\}$. Thus, there exists a constant $C' > 0$ such that $\|\dot\gamma_M(s)\|_{\gamma_M(s)} \leq C'$ for all $s \in [0, t]$. □

Our regularity result is the following.

Proposition 3.12 (Semiconcavity). *Assume that A1–A4 hold. Then,*

- *(i) For any $i, j \in \mathcal{I}$, the restricted cost function $\bar h_t(\cdot, \cdot) := h_t\big((\cdot, i), (\cdot, j)\big) : M \times M \to \mathbb{R}$ given by (3) is semiconcave on $M \times M$, with a linear modulus of semiconcavity.*
- *(ii) For any $x, y \in M$ and any $i, j, k \in \mathcal{I}$,*

$$h_t\big((x, i), (y, j)\big) - h_t\big((x, i), (y, k)\big) \leq \psi(x, k, j).$$

Remark 3.13. Proposition 3.12 implies that $\bar h_t$ is differentiable almost everywhere.

Proof of Proposition 3.12. First, we prove (i). We fix a local chart and identify M with an open subset of $\mathbb{R}^d$, for definiteness, $B_R(0)$, with $R > 0$ large enough. For $x_1, x_2 \in B_1(0)$, we consider $\gamma \in AC \times \mathcal{P}$ an action minimizer from (x_1, i) to (x_2, j). Hence,

$$\begin{aligned} \bar h_t(x_1, x_2) = &\int_0^t L\big(\gamma_M(s), \dot\gamma_M(s), \gamma_{\mathcal{I}}(s)\big)\ ds + \psi\big(x_1, i, \gamma_{\mathcal{I}}(0^+)\big) \\ &+ \sum_{\ell=0}^{N} \psi\big(\gamma_M(t_\ell), \gamma_{\mathcal{I}}(t_\ell^-), \gamma_{\mathcal{I}}(t_\ell^+)\big) + \psi\big(x_2, \gamma_{\mathcal{I}}(t^-), j\big). \end{aligned} \tag{34}$$

We claim that $\bar h_t$ is semiconcave in $B_1(x_1) \times B_1(x_2)$. By Lemma 3.11, there exists a constant C such that

$$\|\dot\gamma_M(s)\| \leq C \text{ for all } s \in [0, t].$$

Moreover, we know $\gamma_M\big([0, t]\big) \subset B_R(0)$. Next, we choose $\varepsilon > 0$ such that $C\varepsilon < 1$. This implies

$$\gamma_M([0, \varepsilon]) \subset B_2(x_1) \text{ and } \gamma_M([t - \varepsilon, t]) \subset B_2(x_2).$$

For $y_1, y_2 \in B_2$, we define

$$\tilde\gamma_M(s) = \begin{cases} \dfrac{\varepsilon - s}{\varepsilon}\ y_1 + \gamma_M(s) & \text{for } s \in [0, \varepsilon], \\ \gamma(s) & \text{for } s \in [\varepsilon, t - \varepsilon], \\ \dfrac{s - t + \varepsilon}{\varepsilon}\ y_2 + \gamma_M(s)\ , & \text{for } s \in [t - \varepsilon, t]. \end{cases} \tag{35}$$

Observe that $\tilde{\gamma}_M$ and $\dot{\tilde{\gamma}}_M$ are bounded as well. Because $\tilde{\gamma}_M(0) = x_1 + y_1$, $\tilde{\gamma}_M(t) = x_2 + y_2$, and we can take $\varepsilon > 0$ sufficiently small such that $\varepsilon < \min\{t_1, t - t_N\}$, we have

$$\begin{aligned}
&\bar{h}_t(x_1 + y_1, x_2 + y_2) - \bar{h}_t(x_1, x_2) \leq \mathcal{J}_t[\tilde{\gamma}_M] - \mathcal{J}_t[\gamma_M] \\
&\quad = \int_0^\varepsilon \Big[L\Big(\frac{\varepsilon - s}{\varepsilon} y_1 + \gamma_M(s), -\frac{1}{\varepsilon} y_1 + \dot{\gamma}_M(s), \gamma_{\mathcal{I}}(s) \Big) - L(\gamma_M, \dot{\gamma}_M, \gamma_{\mathcal{I}}) \Big] \, ds \\
&\quad + \int_{t-\varepsilon}^t \Big[L\Big(\frac{s - 1 + \varepsilon}{\varepsilon} y_2 + \gamma_M(s), \frac{1}{\varepsilon} y_2 + \dot{\gamma}_M(s), \gamma_{\mathcal{I}}(s) \Big) - L(\gamma_M, \dot{\gamma}_M, \gamma_{\mathcal{I}}) \Big] \, ds \\
&\quad + \psi\big(x_1 + y_1, i, \gamma_{\mathcal{I}}(0^+)\big) - \psi\big(x_1, i, \gamma_{\mathcal{I}}(0^+)\big) + \psi\big(x_2 + y_2, \gamma_{\mathcal{I}}(t^-), j\big) - \psi\big(x_2, \gamma_{\mathcal{I}}(t^-), j\big).
\end{aligned} \tag{36}$$

Since L and ψ are both of class C^2, we obtain

$$\bar{h}_t(x_1 + y_1, x_2 + y_2) - \bar{h}_t(x_1, x_2) \leq F_t(y_1, y_2) + C\Big(\|v\|^2 + \|z\|^2 \Big),$$

where

$$\begin{aligned}
F_t(y_1, y_2) := &\int_0^\varepsilon \Big(\frac{\varepsilon - s}{\varepsilon} \frac{\partial L}{\partial x}(\gamma_M, \dot{\gamma}_M, \gamma_{\mathcal{I}}) \cdot y_1 - \frac{1}{\varepsilon} \frac{\partial L}{\partial v}(\gamma_M, \dot{\gamma}_M, \gamma_{\mathcal{I}}) \cdot y_1 \Big) \, ds \\
+ &\int_{t-\varepsilon}^t \Big(\frac{s - 1 + \varepsilon}{\varepsilon} \frac{\partial L}{\partial x}(\gamma_M, \dot{\gamma}_M, \gamma_{\mathcal{I}}) \cdot y_2 + \frac{1}{\varepsilon} \frac{\partial L}{\partial v}(\gamma_M, \dot{\gamma}_M, \gamma_{\mathcal{I}}) \cdot y_2 \Big) \, ds \\
+ &\, \partial_x \psi\big(x_1, i, \gamma_{\mathcal{I}}(0^+)\big) \cdot y_1 + \partial_x \psi\big(x_2, \gamma_{\mathcal{I}}(t^-), j\big) \cdot y_2.
\end{aligned} \tag{37}$$

is a linear function. This proves that h_t is locally semiconcave. By the compactness of $M \times M$, h_t is semiconcave and (i) holds.

Next, we prove (ii). Let $\gamma \in AC([0, t]; M) \times \mathcal{P}([0, t]; \mathcal{I})$ be such that $\gamma(0) = (x, i)$, $\gamma(t) = (y, k)$, and

$$h_t\big((x, i), (y, k)\big) = \mathcal{J}_t[\gamma].$$

Define $\tilde{\gamma} : [0, t + \delta] \to M \times \mathcal{I}$ by

$$\tilde{\gamma}(s) := \begin{cases} \gamma(s) & \text{in } [0, t), \\ (y, j) & \text{in } [t, t + \delta]. \end{cases} \tag{38}$$

Then

$$h_{t+\delta}\big((x, i), (y, j)\big) \leq \mathcal{J}_{t+\delta}[\tilde{\gamma}] = h_t\big((x, i), (y, k)\big) + \delta \, L(x, 0, j) + \psi(x, k, j).$$

Letting $\delta \to 0$ ends the proof (note that h_t is continuous with respect to t, as it is easy to prove by a simple reparametrization argument). $\square$

Proposition 3.12 characterizes the superdifferential of $\bar{h}_t$, as we state next:

Corollary 3.14. *Assume that A1–A4 hold. For any action minimizer $\gamma = (\gamma_M, \gamma_{\mathcal{I}})$ satisfying $\gamma_M(0) = x$ and $\gamma_M(t) = y$, the linear functional F_t given by (37) is a superdifferential of $\bar{h}_t$ at (x, y), and it can be written as*

$$\begin{aligned}
F_t(y_1, y_2) = &\frac{\partial L}{\partial v}\big(\gamma_M(t), \dot{\gamma}_M(t), \gamma_{\mathcal{I}}(t^-)\big) \cdot y_2 - \frac{\partial L}{\partial v}\big(\gamma_M(0), \dot{\gamma}_M(0), \gamma_{\mathcal{I}}(0^+)\big) \cdot y_1 \\
&+ \partial_x \psi\big(\gamma_M(0), i, \gamma_{\mathcal{I}}(0^+)\big) \cdot y_1 + \partial_x \psi\big(\gamma_M(t), \gamma_{\mathcal{I}}(t^-), j\big) \cdot y_2.
\end{aligned} \tag{39}$$

In particular, if $\bar{h}_t$ is differentiable at (x, y), then

$$\begin{aligned}
d_{(x,y)} \bar{h}_t \cdot (y_1, y_2) = &\frac{\partial L}{\partial v}\big(\gamma_M(t), \dot{\gamma}_M(t), \gamma_{\mathcal{I}}(t^-)\big) \cdot y_2 - \frac{\partial L}{\partial v}\big(\gamma_M(0), \dot{\gamma}_M(0), \gamma_{\mathcal{I}}(0^+)\big) \cdot y_1 \\
&+ \partial_x \psi\big(\gamma_M(0), i, \gamma_{\mathcal{I}}(0^+)\big) \cdot y_1 + \partial_x \psi\big(\gamma_M(t), \gamma_{\mathcal{I}}(t^-), j\big) \cdot y_2.
\end{aligned} \tag{40}$$

Proof. Clearly, F_t given by (37) is in the superdifferential of $\bar{h}_t$. To prove the representation formula (39), it is enough to use the Euler-Lagrange equations (4) and integrate by parts in (37). $\square$

Remark 3.15. The terms involving $\partial_x \psi$ disappear if there are no switches at the endpoints.

4 Weakly coupled systems of Hamilton-Jacobi equations

Here, we prove Theorem 1.4 and Theorem 1.6. First, in Subsection 4.1, we introduce time-dependent systems and begin our study of the Lax-Oleinik semigroup (13). We define viscosity solutions of (9) and explain their relation to the Lax-Oleinik semigroup. Next, in Subsection 4.2, we prove Theorem 1.4. Then, in Subsection 4.3, we define the Aubry set and study its properties. Finally, in Subsection 4.4, we examine the long-time behavior of the Lax-Oleinik semigroup.

4.1 Lax-Oleinik semigroup and viscosity solutions

We investigate the connection between the optimal switching problem and weakly coupled systems of Hamilton-Jacobi equations (see Propositions 4.1, 4.6, and 4.7).

Let $L : TM \times \mathcal{I} \to \mathbb{R}$ be a Lagrangian and $\psi : M \times \mathcal{I} \times \mathcal{I} \to \mathbb{R}$ be a switching cost. Let $H : T^*M \times \mathcal{I} \to \mathbb{R}$ be the Hamiltonian (8). Consider the time-dependent system of weakly coupled Hamilton-Jacobi equations:

$$\max\left\{\partial_t u(t,x,i) + H\big(x, \partial_x u(t,x,i), i\big),\ u(t,x,i) - \min_{j\neq i}\{u(t,x,j) + \psi(x,i,j)\}\right\} = 0, \tag{41}$$

on $[0,+\infty) \times M \times \mathcal{I}$. Here, we examine viscosity solutions of (41) according to Definition 1.2, with obvious modifications.

As in the unimodal case, the Lax-Oleinik semigroup $T_t u$ defined by (13) gives a viscosity solution of the time-dependent system (41).

Proposition 4.1. *Assume that A1–A3 hold. Let $u_0 \in C(M \times \mathcal{I})$ and $T_t u_0$ be as in (13). Then the function $u : [0,+\infty) \times M \times \mathcal{I} \to \mathbb{R}$ defined by*

$$u(t,A) := T_t u_0(A), \qquad t \geq 0,\ A \in M \times \mathcal{I},$$

is a viscosity solution of (41) with $u(0,A) = u_0(A)$.

Proof. The proof is standard and can be found in [16] (see also [1, 15]). □

Proposition 4.2. *Assume that A1–A4 hold. For $u \in C(M\times\mathcal{I})$, $t > 0$, and $i \in \mathcal{I}$, the map $x \mapsto T_t u(x,i)$ is semiconcave. More precisely, suppose that $\gamma = (\gamma_M, \gamma_{\mathcal{I}})$ is an action minimizer from (x,i) to $(x+h,i)$. Then, choosing a local chart for M, we have*

$$T_t u(x+h,i) - T_t u(x,i) \leq \frac{\partial L}{\partial v}\big(\gamma_M(t), \dot\gamma_M(t), \gamma_{\mathcal{I}}(t^-)\big)\cdot h + C\,\|h\|^2. \tag{42}$$

Consequently,

$$\frac{\partial L}{\partial v}\big(\gamma_M(t), \dot\gamma_M(t), \gamma_{\mathcal{I}}(t^-)\big) \in \partial^+\big(T_t u\big)(x,i).$$

Proof. By compactness, there exists $B \in M \times \mathcal{I}$ such that

$$T_t u(A) = u(B) + h_t(B,A).$$

Thus,

$$T_t u(x+h,i) - T_t u(x,i) \leq h_t\big(B,(x+h,i)\big) - h_t(B,A).$$

Reasoning as in the proof of Proposition 3.12, we have

$$T_t u(x+h,i) - T_t u(x,i) \leq \frac{\partial L}{\partial v}\big(\gamma_M(t), \dot\gamma_M(t), \gamma_{\mathcal{I}}(t^-)\big)\cdot h + K\,\|h\|^2, \tag{43}$$

as desired. □

Before discussing the Lax-Oleinik semigroup and the solutions of (9), we prove some auxiliary results:

Lemma 4.3. *Assume that A1–A4 hold. Let $v \in C(M\times\mathcal{I})$ and $A = (x,i) \in M \times \mathcal{I}$. Then there exists a minimizer $\gamma = (\gamma_M, \gamma_{\mathcal{I}}) \in AC \times \mathcal{P}$ for the Lax-Oleinik operator (13). In other words, there exists γ such that*

$$T_t v(A) = v\big(\gamma(0)\big) + \mathcal{J}_t[\gamma].$$

Moreover, if v is a subsolution of (9) for some constant $c \in \mathbb{R}$, we can choose γ to have no switch at $t = 0$.

Proof. Let $A = (x, i) \in M \times \mathcal{I}$. By compactness, there exists $B = (y, j) \in M \times \mathcal{I}$ such that $T_t v(A) = v(B) + h_t(B, A)$. Furthermore, Theorem 1.1 implies the existence of a curve, $\gamma : [0, t] \to M \times \mathcal{I}$, with $\gamma(0) = B$ and $\gamma(t) = A$, for which

$$
\begin{aligned}
T_t v(A) = v(\gamma(0)) + \int_0^t L(\gamma_M, \dot{\gamma}_M, \gamma_{\mathcal{I}}) + \psi\big(\gamma_M(0), \gamma_{\mathcal{I}}(0), \gamma_{\mathcal{I}}(0^+)\big) \\
+ \sum_{\ell=1}^{N} \psi\big(\gamma_M(t_\ell), \gamma_{\mathcal{I}}(t_\ell^-), \gamma_{\mathcal{I}}(t_\ell^+)\big) + \psi(x, \gamma_{\mathcal{I}}(t^-), \gamma_{\mathcal{I}}(t)).
\end{aligned} \tag{44}
$$

This proves the existence of a minimizing curve.

Now we prove that, if v is a subsolution, γ can be chosen such that $\gamma_{\mathcal{I}}(0) = \gamma_{\mathcal{I}}(0^+)$. Let $\tilde{B} = \big(\gamma_M(0), \gamma_{\mathcal{I}}(0^+)\big)$. Thus,

$$
\begin{aligned}
T_t v(A) \le v(\tilde{B}) + h_t(\tilde{B}, A) = v\big(\gamma_M(0), \gamma_{\mathcal{I}}(0^+)\big) + \int_0^t L(\gamma_M, \dot{\gamma}_M, \gamma_{\mathcal{I}}) \\
+ \sum_{\ell=1}^{N} \psi\big(\gamma_M(t_\ell), \gamma_{\mathcal{I}}(t_\ell^-), \gamma_{\mathcal{I}}(t_\ell^+)\big) + \psi(x, \gamma_{\mathcal{I}}(t^-), \gamma_{\mathcal{I}}(t)).
\end{aligned} \tag{45}
$$

According to (44) and (45), we have

$$
v\big(\gamma_M(0), \gamma_{\mathcal{I}}(0)\big) + \psi\big(\gamma_M(0), \gamma_{\mathcal{I}}(0), \gamma_{\mathcal{I}}(0^+)\big) \le v\big(\gamma_M(0), \gamma_{\mathcal{I}}(0^+)\big).
$$

Because v is a subsolution, the opposite inequality holds (recall Remark 1.3(*ii*)). Hence,

$$
v\big(\gamma_M(0), \gamma_{\mathcal{I}}(0)\big) + \psi\big(\gamma_M(0), \gamma_{\mathcal{I}}(0), \gamma_{\mathcal{I}}(0^+)\big) = v\big(\gamma_M(0), \gamma_{\mathcal{I}}(0^+)\big).
$$

Consequently, the switch at $t = 0$ is unnecessary, and we can set $\gamma_{\mathcal{I}}(0) := \gamma_{\mathcal{I}}(0^+)$. ☐

Lemma 4.4. *Assume that A1–A4 hold. Let v is a subsolution of (9) for some constant $c \in \mathbb{R}$. Fix $t > 0$, and let $\gamma = (\gamma_M, \gamma_{\mathcal{I}}) \in AC \times \mathcal{P}$ be a minimizer for the Lax-Oleinik operator as in Lemma 4.3. Then, if $t > 0$ sufficiently small, γ has at most one switch.*

Proof. First, by the same argument as in the unimodal case, viscosity subsolutions are Lipschitz. Hence, there exists $K_1 > 0$ such that

$$
|v(z, i) - v(y, i)| \le K_1 d(z, y) \quad \text{for any } z, y \in M \text{ and for all } i \in \mathcal{I}.
$$

Fix $K_2 > 0$ to be selected later. By the superlinearity of L, there exists a constant $C \ge 0$ such that

$$
(K_1 + K_2) \|v\|_x - C \le L(x, v, i).
$$

Let $K_3 > 0$ be the Lipschitz constant of ψ, that is,

$$
|\psi(z, i, j) - \psi(y, i, j)| \le K_3 d(z, y) \quad \text{for any } y, z \in M \text{ and } i, j \in \mathcal{I}.
$$

Let γ be a minimizer. Because v is a subsolution we have $v\big(x, \gamma_{\mathcal{I}}(0)\big) \ge v(x, i) - \psi(x, \gamma_{\mathcal{I}}(0), i)$ (see

Remark 1.3(*ii*)). In addition, by Lemma 4.3, we can assume that no switch occurs at $t=0$. Therefore,

$$\begin{aligned}
T_t v(A) &\geq v\big(x, \gamma_{\mathcal{I}}(0)\big) + (K_1+K_2)\int_0^t \|\dot\gamma_M(\tau)\|_{\gamma_M(\tau)}\, d\tau - Ct + v\big(\gamma_M(0), \gamma_{\mathcal{I}}(0)\big) - v\big(x, \gamma_{\mathcal{I}}(0)\big)\\
&\quad + \sum_{\ell=1}^N \psi\big(\gamma_M(t_\ell), \gamma_{\mathcal{I}}(t_\ell^-), \gamma_{\mathcal{I}}(t_\ell^+)\big) + \psi(x, \gamma_{\mathcal{I}}(t^-), i)\\
&\geq v(x,i) - K_1 d\big(\gamma_M(0), x\big) + (K_1+K_2)\int_0^t \|\dot\gamma_M(\tau)\|_{\gamma_M(\tau)}\, d\tau - Ct\\
&\quad + \sum_{\ell=1}^N \psi\big(\gamma_M(t_\ell), \gamma_{\mathcal{I}}(t_\ell^-), \gamma_{\mathcal{I}}(t_\ell^+)\big) + \psi(x, \gamma_{\mathcal{I}}(t^-), i) - \psi(x, \gamma_{\mathcal{I}}(0), i)\\
&\geq v(x,i) + K_2\int_0^t \|\dot\gamma_M(\tau)\|_{\gamma_M(\tau)}\, d\tau - Ct + \sum_{\ell=1}^N \psi\big(x, \gamma_{\mathcal{I}}(t_\ell^-), \gamma_{\mathcal{I}}(t_\ell^+)\big) + \psi(x, \gamma_{\mathcal{I}}(t^-), i)\\
&\quad - \psi(x, \gamma_{\mathcal{I}}(0), i) + \sum_{\ell=1}^N \Big[\psi\big(\gamma_M(t_\ell), \gamma_{\mathcal{I}}(t_\ell^-), \gamma_{\mathcal{I}}(t_\ell^+)\big) - \psi\big(x, \gamma_{\mathcal{I}}(t_\ell), \gamma_{\mathcal{I}}(t_{\ell+1})\big)\Big]\\
&\geq v(x,i) + K_2\int_0^t \|\dot\gamma_M(\tau)\|_{\gamma_M(\tau)}\, d\tau - Ct + \sum_{\ell=1}^N \psi\big(x, \gamma_{\mathcal{I}}(t_\ell^-), \gamma_{\mathcal{I}}(t_\ell^+)\big)\\
&\quad + \psi(x, \gamma_{\mathcal{I}}(t^-), i) - \psi(x, \gamma_{\mathcal{I}}(0), i) - K_3\sum_{\ell=1}^N d\big(\gamma_M(t_\ell), x\big).
\end{aligned} \tag{46}$$

Select $K_2 \geq N K_3$. Because the number of switches of any minimizer is universally bounded (see the first step in the proof of Theorem 3.3), K_2 is bounded by a universal constant. Consequently, we have

$$T_t v(A) \geq v(A) - Ct + \sum_{\ell=1}^N \psi\big(x, \gamma_{\mathcal{I}}(t_\ell^-), \gamma_{\mathcal{I}}(t_\ell^+)\big) + \psi(x, \gamma_{\mathcal{I}}(t^-), i) - \psi(x, \gamma_{\mathcal{I}}(0), i). \tag{47}$$

Because $\gamma_{\mathcal{I}}(t_1^-) = \gamma_{\mathcal{I}}(0)$ (recall that no switch occurs at $t=0$), if γ has two or more switches, it follows from A2 that

$$\sum_{\ell=1}^N \psi\big(x, \gamma_{\mathcal{I}}(t_\ell^-), \gamma_{\mathcal{I}}(t_\ell^+)\big) + \psi(x, \gamma_{\mathcal{I}}(t^-), i) - \psi(x, \gamma_{\mathcal{I}}(0), i) \geq \delta > 0. \tag{48}$$

Therefore,

$$T_t v(A) \geq v(A) + \delta - Ct. \tag{49}$$

On the other hand, by choosing $\bar\gamma(s) := A$ for $s \in [0,t]$, we get

$$T_t v(A) \leq v(A) + \mathcal{J}_t[\bar\gamma] \leq v(A) + Ct.$$

The previous inequality combined with (49) gives a contradiction if $t>0$ is sufficiently small. □

Lemma 4.5. *Assume that A1–A4 hold. Let v be a subsolution of (9) and fix $(x,i) \in M \times \mathcal{I}$. Assume further that*

$$v(x,i) < \min_{j\neq i}\big\{v(x,j) + \psi(x,i,j)\big\}. \tag{50}$$

Then, there exists $s_0 = s_0(x)$ such that

$$T_s v(x,i) = \inf_{\gamma_M(s)=x}\left\{v(\gamma_M(0), i) + \int_0^s L(\gamma_M, \dot\gamma_M, i)\right\} \qquad \forall\, s \in [0, s_0]. \tag{51}$$

Proof. Fix a neighborhood N of x in M such that (50) holds in N. We claim that we can select $s>0$ so that any minimizing curve γ for $T_s v(x,i)$ is contained in N. To see this, we recall that $T_s v(y,i) \leq C$ for $(s,y) \in [0,s_0] \times M$. Accordingly, by the superlinearity of $L(\cdot,\cdot,i)$, for any $\varepsilon>0$ there exists $C_\varepsilon \geq 0$ such that

$$-C + \frac{1}{\varepsilon}\int_0^s \|\dot\gamma_M\|_{\gamma_M} - C_\varepsilon s \leq v\big(\gamma(0)\big) + \int_0^s L\big(\gamma_M, \dot\gamma_M, \gamma_{\mathcal{I}}\big) \leq C. \tag{52}$$

The prior inequality implies that

$$d\big(\gamma_M(s'),x\big) \leq \int_0^s \|\dot\gamma_M\|_{\gamma_M} \leq C\varepsilon + C_\varepsilon \varepsilon s \qquad \forall\, s' \in [0,s].$$

Given $\rho > 0$, we choose $\varepsilon < \delta/2C$; then, we fix $s > 0$ sufficiently small such that $C_\varepsilon \varepsilon s \leq \delta/2$. Therefore we have

$$d\big(\gamma_M(s'),x\big) \leq \rho \qquad \forall\, s' \in [0,s]. \tag{53}$$

Now, suppose $\gamma_{\mathcal{I}}(0) = j \neq i$. By Lemma 4.4, we have only one switch if $s > 0$ is small enough. Thanks to (50) and (53), by choosing ρ sufficiently small, we gather that

$$v\big(\gamma_M(0),j\big) + \psi\big(\gamma_M(0),j,i\big) \geq v\big(\gamma_M(0),i\big) + \delta_0 \tag{54}$$

for some $\delta_0 > 0$. Then, arguing as in the proof of Lemma 4.4 and using (54) instead of (48), we get that

$$T_s v(A) \geq v(A) + \delta_0 - Cs,$$

which is a contradiction if s is sufficiently small. Consequently, we cannot have switches for $0 < s < s_0(x)$. □

Next, we investigate the connection between the Lax-Oleinik semigroup and viscosity solutions of (9).

Proposition 4.6. *Assume that A1–A4 hold. A Lipschitz function $u : M \times \mathcal{I} \to \mathbb{R}$ is a viscosity subsolution of (9) if and only if $u \leq T_t u + ct$ for all $t \geq 0$.*

Proof. To simplify, we assume that $c = 0$. This can be achieved, without loss of generality, by subtracting c from L if necessary. If u is a subsolution then, as in the unimodal case (see for instance [13]),

$$u(\gamma_M(\tau),i) - u(\gamma_M(\tau'),i) \leq \int_{\tau'}^{\tau} L\big(\gamma_M(s),\dot\gamma_M(s),i\big)\ ds$$

for any $0 \leq \tau' \leq \tau$ and any $\gamma_M : [\tau',\tau] \to M$.

We fix a curve $\gamma : [0,t] \to M \times \mathcal{I}$ with $\gamma(t) = A$, and we recall that $\gamma_{\mathcal{I}}(t_\ell^+) = \gamma_{\mathcal{I}}(t_{\ell+1}^-)$. Using the convention $\gamma_{\mathcal{I}}(t_{N+1}^+) = \gamma_{\mathcal{I}}(t)$, we get

$$\begin{aligned} u(A) - u\big(\gamma(0)\big) &= \sum_{\ell=0}^{N}\Big[u\big(\gamma(t_{\ell+1}^+)\big) - u\big(\gamma(t_\ell^+)\big)\Big] = \sum_{\ell=0}^{N}\Big[u\big(\gamma_M(t_{\ell+1}),\gamma_{\mathcal{I}}(t_\ell^+)\big) - u\big(\gamma_M(t_\ell),\gamma_{\mathcal{I}}(t_\ell^+)\big)\Big] \\ &\quad + \sum_{\ell=0}^{N}\Big[u\big(\gamma_M(t_{\ell+1}),\gamma_{\mathcal{I}}(t_{\ell+1}^+)\big) - u\big(\gamma_M(t_{\ell+1}),\gamma_{\mathcal{I}}(t_\ell^+)\big)\Big] \\ &\leq \sum_{\ell=0}^{N}\int_{t_\ell}^{t_{\ell+1}} L\big(\gamma_M(s),\dot\gamma_M(s),\gamma_{\mathcal{I}}(t_\ell^+)\big)\ ds + \sum_{\ell=0}^{N}\psi\big(\gamma_M(t_{\ell+1}),\gamma_{\mathcal{I}}(t_{\ell+1}^-),\gamma_{\mathcal{I}}(t_{\ell+1}^+)\big). \end{aligned} \tag{55}$$

Therefore, $u \leq T_t u$.

To prove the converse implication, we fix (x_0,i_0) and take γ satisfying $\gamma(t_0) = (x_0,i_0)$, $\gamma_M(t) = x$, $\gamma_{\mathcal{I}} \equiv i_0$, and $\dot\gamma_M(t_0) = v$. Then,

$$u(x_0,i_0) \leq u\big(\gamma_M(t),i_0\big) + \int_t^{t_0} L\big(\gamma_M,\dot\gamma_M,i_0\big)\ ds.$$

Fix a C^1 function ϕ such that $u(\cdot,i_0) - \phi$ has a maximum at x_0. Then,

$$\frac{\phi(x_0) - \phi(\gamma_M(t))}{t_0 - t} \leq \frac{1}{t_0 - t}\int_t^{t_0} L\big(\gamma_M,\dot\gamma_M,\gamma_{\mathcal{I}}\big)\ ds.$$

Hence,

$$d\phi(x_0)\cdot v \leq L(x_0,v,i_0).$$

Because v is arbitrary, we get

$$H\big(x_0,d\phi(x_0),i_0\big) \leq 0.$$

To complete the proof, we observe that the Lax-Oleinik semigroup solves (41) (see Proposition 4.1). Consequently,

$$T_t u(x,i) \le T_t u(x,j) + \psi(x,i,j).$$

We let $t \to 0$ to obtain

$$u(x,i) \le u(x,j) + \psi(x,i,j).$$

□

Proposition 4.7. *Suppose A1–A4 hold. A semiconcave function $u : M \times \mathcal{I} \to \mathbb{R}$ is a viscosity solution of (9) if and only if $u = T_t u + ct$ for all $t \ge 0$.*

Proof. Again, we assume for simplicity that $c = 0$. The 'if' part is the same as in the classical case (see Proposition 4.1).

To prove the converse statement, assume that $u : M \times \mathcal{I} \to \mathbb{R}$ solves (9). Then, for any $i \in \mathcal{I}$,

$$\max \Big\{ H\big(x, du(x), i\big) - c_0, u(x,i) - \Psi u(x,i) \Big\} = 0 \tag{56}$$

in the viscosity sense (see Definition 1.2). We claim that, for each $(x,i) \in M \times \mathcal{I}$, there exists $s_0 = s_0(x,i) > 0$ such that $T_s u = u$ for all $0 \le s < s_0(x,i)$. This claim and the semigroup property give that the set of times s such that $T_s u = u$ is open. By continuity, this set is also closed; this proves the result.

To establish the claim, we consider two cases.

Case 1: $u(x,i) < \Psi u(x,i)$. Select a neighborhood $\tilde{U}$ of x such that $u < \Psi u$ in $\tilde{U}$. Next, we apply Lemma 4.5 as follows. If $s > 0$ is sufficiently small, there exists $U \subseteq \tilde{U}$ such that $T_s u(y,i)$ cannot have switches for any $y \in U \subset \tilde{U}$. Hence, because u is semiconcave and $H\big(y, du(y), i\big) = c_0$ on U, the classical theory implies $u(y,i) = T_s u(y,i)$ for all $y \in U$; this proves the claim in this case.

Case 2: $u(x,i) = \Psi u(x,i)$. This means that there exists $j \in \mathcal{I} \backslash \{i\}$ such that

$$u(x,i) = u(x,j) + \psi(x,j,i). \tag{57}$$

We claim that

$$u(x,j) < \Psi u(x,j). \tag{58}$$

To verify the claim, we argue by contradiction. If the preceding inequality fails, we have $u(x,j) = \Psi u(x,j) = u(x,k) + \psi(x,k,j)$ for some $k \ne i,j$. Hence, (57) and Assumption A2 imply that

$$\Psi u(x,i) = u(x,i) = u(x,k) + \psi(x,k,j) + \psi(x,j,i) > u(x,k) + \psi(x,k,i),$$

which is a contradiction. Therefore, (58) holds.

By the argument in Case 1, we have $u(x,j) = T_s u(x,j)$ for $s > 0$ small. This, together with (57) and Proposition 4.1, implies

$$u(x,i) = u(x,j) + \psi(x,j,i) = T_s u(x,j) + \psi(x,j,i) \ge T_s u(x,i).$$

Because u is a subsolution, the other inequality follows from the previous proposition. □

To end this section, we prove two more properties of the Lax-Oleinik semigroup.

Proposition 4.8. *Assume that A1–A4 hold. For every $i \in \mathcal{I}$, the map $u \mapsto T_t u(\cdot, i)$ is a weak contraction in the $L^\infty(M)$-norm; that is, for any two continuous functions $g, h : M \times \mathcal{I} \to \mathbb{R}$, we have*

$$\|T_t g(\cdot,i) - T_t h(\cdot,i)\|_{L^\infty(M)} \le \|g(\cdot,i) - h(\cdot,i)\|_{L^\infty(M)}.$$

Proof. From (13), we have $T_t(g+c)(\cdot,i) = T_t g(\cdot,i) + c$ for $c \in \mathbb{R}$. Moreover, if $g(\cdot,i) \le h(\cdot,i)$, for every $i \in \mathcal{I}$, then $T_t g(\cdot,i) \le T_t h(\cdot,i)$. Thus, from

$$h(\cdot,i) - \|g(\cdot,i) - h(\cdot,i)\|_{L^\infty(M)} \le g(\cdot,i) \le h(\cdot,i) + \|g(\cdot,i) - h(\cdot,i)\|_{L^\infty(M)},$$

we conclude that

$$T_t h(\cdot,i) - \|g(\cdot,i) - h(\cdot,i)\|_{L^\infty(M)} \le T_t g(\cdot,i) \le T_t h(\cdot,i) + \|g(\cdot,i) - h(\cdot,i)\|_{L^\infty(M)},$$

as desired. □

Proposition 4.9. *Assume that A1–A4 hold. The Lax-Oleinik semigroup* (13) *is a semigroup; that is,*

$$T_t(T_s u) = T_{t+s}u \qquad \forall\ t, s \geq 0. \tag{59}$$

Proof. First, the same argument as in the classical unimodal case (see, for instance, [13]) shows that the function h_t given by (3) satisfies

$$h_{t+s}(A, B) = \inf_{C \in M \times \mathcal{I}} \big\{ h_t(A, C) + h_s(C, B) \big\}. \tag{60}$$

Hence, using (60), for $A \in M \times \mathcal{I}$ we have

$$\begin{aligned} T_{t+s}u(A) &= \inf_{B \in M \times \mathcal{I}} \big\{ u(B) + h_{t+s}(B, A) \big\} = \inf_{B \in M \times \mathcal{I}} \big\{ u(B) + \inf_{C \in M \times \mathcal{I}} \big[h_t(B, C) + h_s(C, A) \big] \big\} \\ &= \inf_{C \in M \times \mathcal{I}} \big\{ h_s(C, A) + \inf_{B \in M \times \mathcal{I}} \big[u(B) + h_t(B, C) \big] \big\} = \inf_{C \in M \times \mathcal{I}} \big\{ T_t(C) + h_s(C, A) \big\} = T_s(T_t u)(A), \end{aligned} \tag{61}$$

which proves (59). □

4.2 The Weak KAM theorem for optimal switching problems

Here we prove Theorem 1.4. As in the original Weak KAM Theorem, see [12], we show that the Lax-Oleinik semigroup converges to a viscosity solution of (12).

We begin by proving that the critical value c_0 can be defined as the infimum among all constants $c \in \mathbb{R}$ for which (9) admits a viscosity subsolution.

Proposition 4.10. *Assume that A1–A4 hold. Then, there exists a unique constant $c_0 \in \mathbb{R}$ for which:*

(*i*) *the equation* (12) *has a subsolution;*

(*ii*) *for any $c \in \mathbb{R}$ such that* (9) *has a subsolution, $c \geq c_0$.*

Proof. Define $c_0 \in \mathbb{R}$ by

$$c_0 := \inf \big\{ c \in \mathbb{R} : (9) \text{ has a subsolution} \big\}.$$

For c large enough, $u \equiv 0$ is a subsolution. Therefore, the previous infimum is taken over a nonempty set. Moreover $c_0 \geq \min_{(x,p,i) \in T^*M \times \mathcal{I}} H(x, p, i) > -\infty$, hence c_0 cannot be $-\infty$.

To show that this infimum is a minimum, let $c^j \in \mathbb{R}$ be a minimizing sequence for c_0, that is, $c^j \to c_0$. Let u^j be a subsolution of (9) for $c = c^j$. Because $u^j(\cdot, i)$ is uniformly Lipschitz, up to adding a constant to u^j and extracting a subsequence, there exists u such that $u^j(\cdot, i) \to u(\cdot, i)$. The stability of viscosity subsolutions implies that u is a subsolution of (12). □

Lemma 4.11. *Assume that A1–A4 hold and that $u : M \times \mathcal{I} \to \mathbb{R}$ is a subsolution of* (12)*. Then, for any $t \in [0, \infty)$, there exists $A_t = (x_t, i_t) \in M \times \mathcal{I}$ such that $u(A_t) = T_t u(A_t) + c_0 t$.*

Proof. By contradiction, assume that the statement is false. Thus, there exists $t_0 > 0$ such that, for every $A \in M \times \mathcal{I}$,

$$u(A) < T_{t_0} u(A) + c_0 t_0.$$

Because M is compact, there exists $\varepsilon > 0$ such that, for all $A \in M \times \mathcal{I}$,

$$u(A) < T_{t_0} u(A) - \varepsilon + c_0 t_0 =: \tilde{T}_{t_0} u(A), \tag{62}$$

where

$$\tilde{T}_t u(A) := \inf \left\{ u(\gamma(0)) + \int_0^t \Big(L\big(\gamma_M(s), \dot{\gamma}_M(s), \gamma_{\mathcal{I}}(s)\big) + \tilde{c} \Big)\, ds + \sum_{\ell=1}^{N+1} \psi\big(\gamma_M(t_\ell),\ \gamma_{\mathcal{I}}(t_\ell^-),\ \gamma_{\mathcal{I}}(t_\ell^+)\big) \right\},$$

with $t_{N+1}^+ := t$ and $\tilde{c} := c_0 - \varepsilon / t_0$.

Let

$$\tilde{u} := \inf_{\tau \geq 0} \tilde{T}_\tau u.$$

Note that by (62) and the semigroup property, $\tilde{T}_{nt_0} u \geq u$ for all $n \geq 1$. Thanks to this fact, it follows easily that $\tilde{u}$ is finite. Next, for any $h > 0$, we have

$$\tilde{T}_h \tilde{u} = \inf_{\tau \geq 0} \tilde{T}_{\tau + h} u = \inf_{\tau \geq h} \tilde{T}_\tau u \geq \tilde{u}. \tag{63}$$

Consequently, Proposition 4.6 shows that $\tilde{u}$ is a subsolution of

$$\max\Big\{H\big(x, du(x,i), i\big) - \tilde{c},\ u(x,i) - \Psi u(x,i)\Big\} = 0 \qquad \forall\ i. \tag{64}$$

On the other hand, by the definition of c_0, we have

$$c_0 < \tilde{c} = c_0 - \varepsilon/t_0,$$

a contradiction. □

Proposition 4.12. *Assume that A1–A4 hold and that $u : M \times \mathcal{I} \to \mathbb{R}$ is a subsolution of (12). Then, for every $i \in \mathcal{I}$, $T_t u(\cdot, i) + c_0 t$ is uniformly bounded in t; that is,*

$$\sup_{t \geq 0} \|T_t u(\cdot, i) + c_0 t\|_{L^\infty(M)} < +\infty.$$

Proof. Without loss of generality, by subtracting c_0 from L if necessary, we assume that $c_0 = 0$.

Thanks to Lemma 4.11, for any $t \geq 0$ there exists $A_t = (x_t, i_t) \in M \times \mathcal{I}$ such that $(T_t u - u)(A_t) = 0$. Since $T_t u(\cdot, i_t) - u(\cdot, i_t)$ is uniformly Lipschitz and vanishes at one point, we have

$$\begin{aligned} |T_t u(x, i_t) - u(x, i_t)| &= |(T_t u - u)(x, i_t) - (T_t u - u)(A_t)| \\ &\leq 2C d(x, x_t^i) \leq 2C \operatorname{diam}(M) \equiv C_M. \end{aligned} \tag{65}$$

By Proposition 4.1, $T_t u$ solves (41) for any $i \in \mathcal{I}$. Consequently, we have that

$$T_t u(x, i_t) - \psi(x, i_t, i) \leq T_t u(x, i) \leq T_t u(x, i_t) + \psi(x, i, i_t).$$

This implies that

$$|T_t u(\cdot, i)| \leq \|u(\cdot, i_t)\|_{L^\infty(M)} + C_M + \sup_{x \in M, i \neq j} \psi(x, i, j),$$

proving the result. □

Proposition 4.13. *Assume that A1–A4 hold. If $u : M \times \mathcal{I} \to \mathbb{R}$ is a subsolution of (12), then $T_t u + c_0 t$ converges to a fixed point of $T_t + c_0 t$.*

Proof. Again, without loss of generality, we assume that $c_0 = 0$. Because u is a subsolution of (9), $u \leq T_t u$, for all $t \geq 0$. By the monotonicity of the Lax-Oleinik semigroup, we have

$$T_s u \leq T_{s+t} u \quad \text{for all } t \geq 0.$$

By Proposition 4.12, $T_t u$ is also uniformly bounded in t. Thus, the pointwise limit

$$u^\infty(x, i) := \lim_{t \to +\infty} T_t u(x, i)$$

exists everywhere in M. Since T_s is continuous, we conclude that, for any $s \geq 0$,

$$T_s u^\infty(x, i) = \lim_{t \to +\infty} T_{s+t} u(x, i) = u^\infty(x, i).$$

□

Proof of Theorem 1.4. By Proposition 4.7, the proof follows by combining Propositions 4.10 and 4.13. □

4.3 The Aubry set

In this section we examine the Aubry set $\mathcal{A}$ and investigate various properties used in the analysis of the Lax-Oleinik semigroup. The Aubry set for optimal switching problems was first defined in [15]. Our definition, Definition 1.5, is distinct from that one. There, $\mathcal{A}$ is the set of all points B in $M \times \mathcal{I}$ for which property (ii) of the next proposition is satisfied. However, these two definitions are equivalent as we prove next.

Proposition 4.14. *Assume that A1–A4 hold. The following are equivalent:*

(i) $B \in \mathcal{A}$;

(ii) there exists $\delta > 0$ such that $\inf\left\{\mathcal{J}_t[\gamma] + c_0 t \;\middle|\; t \geq \delta,\ \gamma(0) = \gamma(t) = B\right\} = 0$*;*

(iii) for every $\delta > 0$, $\inf\left\{\mathcal{J}_t[\gamma] + c_0 t \;\middle|\; t \geq \delta,\ \gamma(0) = \gamma(t) = B\right\} = 0$.

Proof. $(i) \implies (ii)$. Let $\varepsilon > 0$. If $B \in \mathcal{A}$, then there exists a sequence $t_k \to +\infty$ for which $h_{t_k}(B, B) + c_0 t_k \to 0$. Let $\gamma^k : [0, t_k] \to M \times \mathcal{I}$ with $\gamma^k(0) = \gamma^k(t_k) = B$ be such that

$$h_{t_k}(B, B) + c_0 t_k + \frac{\varepsilon}{2} \geq \mathcal{J}_{t_k}[\gamma^k] + c_0 t_k.$$

Without loss of generality, we can assume $t_k \geq 1$ for all k. Since $h_{t_k}(B, B) + c_0 t_k \to 0$, if k is large enough we get $\varepsilon \geq \mathcal{J}_{t_k}[\gamma^k] + c_0 t_k$. Thus, (ii) holds with $\delta = 1$.

Before proceeding further we recall that, by Proposition 4.6, any critical subsolution satisfies

$$u\big(\gamma(t)\big) - u\big(\gamma(0)\big) \leq \mathcal{J}_t[\gamma] + c_0 t \quad \forall\, \gamma \in AC([0, t]; M) \times \mathcal{P}([0, t]; \mathcal{I}).$$

Hence, the action on loops is always nonnegative.

$(ii) \implies (iii)$. Let $\delta_0 > 0$ be such that the infimum in (ii) is zero. To show (iii), fix any $\delta > 0$. If $\delta < \delta_0$,

$$0 \leq \inf\left\{J_t[\gamma] + c_0 t \;\middle|\; t \geq \delta,\ \gamma(0) = \gamma(t) = B\right\} \leq \inf\left\{J_t[\gamma] + c_0 t \;\middle|\; t \geq \delta_0,\ \gamma(0) = \gamma(t) = B\right\} = 0,$$

where the first inequality follows from the fact that the action on loops is nonnegative.

Otherwise, if $\delta > \delta_0$, we fix $m \in \mathbb{N}$ such that $\delta < m\delta_0$. For any $\gamma : [0, t] \to M \times \mathcal{I}$ with period $t \geq \delta_0$, we denote the concatenation of γ with itself m times by $\bar{\gamma} : [0, mt] \to M \times \mathcal{I}$. Naturally,

$$J_{mt}[\bar{\gamma}] + c_0 mt = m\big(J_t[\gamma] + c_0 t\big).$$

The concatenation does not create new switches since γ is a loop. Since $mt \geq m\delta_0 \geq \delta$, we know

$$0 \leq \inf\left\{J_t[\gamma] + c_0 t \;\middle|\; t \geq \delta, \gamma(0) = \gamma(t) = B\right\} \leq \inf\left\{J_t[\gamma] + c_0 t \;\middle|\; t \geq m\delta_0, \gamma(0) = \gamma(t) = B\right\} = 0,$$

so (iii) holds in the general case.

$(iii) \implies (i)$. Take $k \in \mathbb{N}$ and $\delta = k$. By (iii), there exist $t_k \geq k$ and γ_k with $\gamma_k(0) = \gamma_k(t_k) = B$ such that

$$J_{t_k}[\gamma_k] + c_0 t_k < \frac{1}{k}.$$

Hence, by (14), $B \in \mathcal{A}$. □

Definition 4.15 (Critical Curve). $\gamma : \mathbb{R} \to M \times \mathcal{I}$ is a critical curve if, for any subsolution $u : M \times \mathcal{I} \to \mathbb{R}$ of (12) and all $t_1 < t_2$,

$$u\big(\gamma(t_2)\big) - u\big(\gamma(t_1)\big) = \int_{t_1}^{t_2} \left[L\big(\gamma_M(s), \dot{\gamma}_M(s), \gamma_{\mathcal{I}}(s)\big) + c_0\right] ds + \sum_{\ell=1}^{N+1} \psi\big(\gamma_M(s_\ell),\ \gamma_{\mathcal{I}}(s_\ell^-),\ \gamma_{\mathcal{I}}(s_\ell^+)\big),$$

where the sum above is taken for all ℓ such that $t_1 < s_\ell < t_2$, with the convention $s_{N+1}^+ = t_2$.

Proposition 4.16 (Existence of critical curves). *Assume that A1–A4 hold. Given $B = (y, j) \in \mathcal{A}$, there exists a critical curve $\gamma : \mathbb{R} \to M \times \mathcal{I}$ with $\gamma(0) = B$.*

Proof. Without loss of generality, assume $c_0 = 0$. Let $\eta^k : [0, t_k] \to M \times \mathcal{I}$, with $\eta^k(0) = B = \eta^k(t_k)$ and $t_k \geq k$, be such that

$$\mathcal{J}_{t_k}[\eta^k] = \int_0^{t_k} L\big(\eta_M^k(s), \dot{\eta}_M^k(s), \eta_{\mathcal{I}}(s)\big)\, ds + \sum_{\ell=1}^{N_k} \psi\big(\eta_M^k(s_\ell), \eta_{\mathcal{I}}^k(s_\ell^-), \eta_{\mathcal{I}}^k(s_\ell^+)\big) \to 0.$$

We set $\gamma^k : [-t_k/2, t_k/2] \to M \times \mathcal{I}$ as $\gamma^k(s) := \eta^k(s + t_\ell/2)$. Then,

$$\int_{-t_k/2}^{t_k/2} L\big(\gamma_M^k(s), \dot{\gamma}_M^k(s), \gamma_{\mathcal{I}}^k(s)\big)\, ds + \sum_{\ell=1}^{N_k} \psi\big(\gamma_M^k(s_\ell), \gamma_{\mathcal{I}}^k(s_\ell^-), \gamma_{\mathcal{I}}^k(s_\ell^+)\big) \to 0.$$

It follows from the argument in Theorem 3.3 that the number of switches grows at most linearly, otherwise $\mathcal{J}_{t_k}[\eta^k]$ would diverge. Also, we can take γ^k to be an action minimizer. Thus,

$$0 \leq \frac{1}{t_k} \int_{-t_k/2}^{t_k/2} L\big(\gamma_M^k(s), \dot\gamma_M^k(s), \gamma_{\mathcal{I}}^k(s)\big)\, ds + \frac{1}{t_k} \sum_{\ell=1}^{N_k} \psi\big(\gamma_M^k(s_\ell), \gamma_{\mathcal{I}}^k(s_\ell)^-, \gamma_{\mathcal{I}}^k(s_\ell^+)\big) \to 0.$$

Furthermore, there exists $t_0^k \in [-t_k/2, t_k/2]$ for which $L\big(\gamma_M^k(t_0^k), \dot\gamma_M^k(t_0^k), \gamma_{\mathcal{I}}^k(t_0^k)\big) \leq C$. Consequently, by the superlinearity of L,

$$\|\dot\gamma_M^k(t_0^k)\|_{\gamma_M^k(t_0^k)} \leq L\big(\gamma_M^k(t_0^k), \dot\gamma_M^k(t_0^k), \gamma_{\mathcal{I}}^k(t_0^k)\big) + C \leq C.$$

By energy conservation (see Proposition 3.8), the curve $(\gamma_M^k, \dot\gamma_M^k)$ is contained in a compact subset of TM. Thus, by a diagonal argument, we can build a curve $\gamma_M : \mathbb{R} \to M$ such that, for every fixed interval $[-T, T]$,

$$\gamma_M^k \to \gamma_M, \text{ uniformly in } [-T,T], \quad \text{and} \quad \dot\gamma_M^k \rightharpoonup \dot\gamma_M, \text{ weakly in } L^1\big([-T,T]\big). \tag{66}$$

The argument in the proof of Theorem 3.3 shows that we can define a limit $\gamma_{\mathcal{I}}$ of $\gamma_{\mathcal{I}}^k$ (see (22)) for which

$$\begin{aligned}\int_a^b L(\gamma_M(s), \dot\gamma_M(s), \gamma_{\mathcal{I}}(s))\, ds &+ \sum_{\ell=0}^{N} \psi\big(\gamma_{\mathcal{I}}(t_\ell), \gamma_{\mathcal{I}}(t_\ell^-), \gamma_{\mathcal{I}}(t_\ell^+)\big)\\ &\leq \liminf_k \left(\int_a^b L(\gamma_M^k(s), \dot\gamma_M^k(s), \gamma_{\mathcal{I}}^k(s))\, ds + \sum_{\ell=0}^{N} \psi\big(\gamma_{\mathcal{I}}(t_\ell^k), \gamma_{\mathcal{I}}((t_\ell^k)^-), \gamma_{\mathcal{I}}((t_\ell^k)^+)\big)\right).\end{aligned}$$

Finally, we prove γ is a critical curve. Consider a subsolution u of (12), and $a \leq b$. We have

$$0 \leq u(\gamma^k(a)) - u(\gamma^k(b)) + \int_a^b L(\gamma_M^k, \dot\gamma_M^k, \gamma_{\mathcal{I}}^k)\, ds + \sum_{[a,b]} \psi\big(\gamma_M^k, \gamma_{\mathcal{I}}^{k-}, \gamma_{\mathcal{I}}^{k+}\big) =: I_1^k. \tag{67}$$

Since $\gamma^k : [-t_k/2, t_k/2] \to M$ is a loop, assuming that k is large enough so that $[a,b] \subset [-t_k/2, t_k/2]$, using again that u is a subsolution u of (12) we have

$$\begin{aligned}0 \leq u(\gamma^k(b)) - u(\gamma^k(a)) + \int_b^{t_k/2} L(\gamma_M^k, \dot\gamma_M^k, \gamma_{\mathcal{I}}^k) + \int_{-t_k/2}^{a} L(\gamma_M^k, \dot\gamma_M^k, \gamma_{\mathcal{I}}^k)\\ + \sum_{[-t_k/2, t_k/2]\setminus[a,b]} \psi(\gamma_M^k, \gamma_{\mathcal{I}}^{k-}, \gamma_{\mathcal{I}}^{k+}) =: I_2^k.\end{aligned} \tag{68}$$

By adding these inequalities, we obtain

$$0 \leq \int_{-t_k/2}^{t_k/2} L(\gamma_M^k, \dot\gamma_M^k, \gamma_{\mathcal{I}}^k) + \sum_{[-t_k/2, t_k/2]} \psi(\gamma_M^k, \gamma_{\mathcal{I}}^{k-}, \gamma_{\mathcal{I}}^{k+}) \to 0.$$

This proves that the numbers $I_1^k \geq 0$ and $I_2^k \geq 0$ satisfy $I_1^k + I_2^k \to 0$. In particular, each of them converges to zero. Therefore,

$$\lim_k \left\{ u(\gamma^k(a)) - u(\gamma^k(b)) + \int_a^b L(\gamma_M^k, \dot\gamma_M^k, \gamma_{\mathcal{I}}^k)\, ds + \sum_{[a,b]} \psi\big(\gamma_M^k, \gamma_{\mathcal{I}}^{k-}, \gamma_{\mathcal{I}}^{k+}\big)\right\} = 0.$$

To conclude, we note that the lower semicontinuity implies

$$\begin{aligned}u(\gamma(b)) - u(\gamma(a)) &= \liminf_k \left\{ \int_a^b L(\gamma_M^k, \dot\gamma_M^k, \gamma_{\mathcal{I}}^k)\, ds + \sum_{[a,b]} \psi\big(\gamma_M^k, \gamma_{\mathcal{I}}^{k-}, \gamma_{\mathcal{I}}^{k+}\big)\right\}\\ &\geq \int_a^b L(\gamma_M, \dot\gamma_M, \gamma_{\mathcal{I}}) + \sum_{[a,b]} \psi\big(\gamma_M, \gamma_{\mathcal{I}}^-, \gamma_{\mathcal{I}}^+\big)\end{aligned}$$

while the opposite inequality follows by Proposition 4.6. □

Proposition 4.17. *Let $U \subset M \times \mathcal{I}$ be an open set such that $\overline{U} \subseteq (M \times \mathcal{I})\backslash\mathcal{A}$. Then there exists a subsolution v of* (12) *that strict in U; that is,*

(i) for all $(y, j) \notin \mathcal{A}$ and $p \in \partial^+ v_j(y)$, we have $H(y, p, j) < 0$;

(ii) for all $i \neq j$, $v_j(y) - v_i(y) < \psi(i, j)$.

Furthermore, v is smooth in $(M \times \mathcal{I})\backslash\mathcal{A}$.

Proof. See [15, Lemma 4.2]. □

Proposition 4.18. *Assume A1–A4 hold. Then, every critical curve is contained in the projected Aubry set $\mathcal{A}$.*

Proof. From the remarks after [15, Theorem 3.6], we have that for every $(x, i) \in \mathcal{A}$ and every subsolution u of (12), $u(\cdot, i)$ cannot be a strict subsolution at x. Accordingly, at least one of the following holds:

(i) either $H\big(x, du(x, i), i\big) = 0$, in the viscosity sense;

(ii) or $v(x, i) - v(x, j) = \psi(x, i, j)$ for some $j \neq i$.

Then the result follows from the previous proposition. □

Consider a subsolution u of (12). The proof of Proposition 4.17 ensures that, given U with $\bar{U} \subset (M \times \mathcal{I})\backslash\mathcal{A}$, we can approximate u uniformly by subsolutions v that are strict and smooth in U. We apply this result to prove a comparison principle:

Corollary 4.19 (Comparison Principle). *Suppose that A1–A4 hold. Let u be a subsolution of* (12)*, w be a supersolution of* (12)*, and assume that $u \leq w$ on $\mathcal{A}$. Then, $u \leq w$ in $M \times \mathcal{I}$.*

Proof. By contradiction, suppose that

$$\min_{M \times \mathcal{I}} (w - u) < 0, \tag{69}$$

and consider a set U, with $\bar{U} \subset (M \times \mathcal{I})\backslash\mathcal{A}$, such that $\min_{M \times \mathcal{I}}(w - u) = \min_U(w - u)$. Then, consider a sequence v_n of subsolutions, strict and smooth in U, that converges uniformly to u and such that

$$w(x_n, i_n) - v_n(x_n, i_n) = \min_{M \times \mathcal{I}} (w - v_n), \tag{70}$$

with $(x_n, i_n) \in U$ for $n \gg 1$. By (70), it follows that $d_{x_n} v_n \in \partial^- w(x_n, i_n)$. Since w is a supersolution, we have:

- either $H(x_n, d_{x_n} v_n, i_n) \geq 0$;
- or $v_n(x_n, i_n) - v_n(x_n, j) = \psi(x_n, i_n, j)$ for some $j \neq i_n$.

Because both alternatives contradict Proposition 4.17, the result follows. □

Next, we examine properties of a particular solution of (12) that we use later to study the large time behavior of the Lax-Oleinik semigroup.

Given $u_0 \in C(M \times \mathcal{I})$, we consider

$$v(A) := \inf_{B \in \mathcal{A}} \left\{ h(B, A) + \inf_{C \in M \times \mathcal{I}} \Big\{ u_0(C) + h(C, B) \Big\} \right\}, \tag{71}$$

where h is the Peierls barrier defined in (14). Next, we prove that v is a critical solution in $M \times \mathcal{I}$. We have (compare to [9, Theorem 3.1]) the following result.

Proposition 4.20. *Suppose that A1–A4 hold and set*

$$v_0(B) := \inf_{C \in M \times \mathcal{I}} \big\{ u_0(C) + h(C, B) \big\} \qquad \forall\, B \in M \times \mathcal{I}.$$

Let v be as in (71)*. Then*

$$v(A) = \inf_{B \in \mathcal{A}} \big\{ v_0(B) + h(B, A) \big\}$$

and the following hold:

1. *v_0 is the maximal subsolution with $v_0 \le u_0$ on $M \times \mathcal{I}$;*

2. *v is a solution and $v = v_0$ on $\mathcal{A}$;*

3. *If $u_0(B) - u_0(A) \le h(B, A)$ for all $A, B \in M \times \mathcal{I}$, then*

$$v(A) = \inf_{B \in \mathcal{A}} \{u_0(B) + h(B, A)\} \qquad \text{in } M \times \mathcal{I}, \tag{72}$$

and $v_0 = u_0$ on $\mathcal{A}$.

Proof. 1. By setting $C = B$, we get that $v_0 \le u_0$ on $M \times \mathcal{I}$. Moreover, because the infimum of subsolutions is a subsolution, v_0 is a subsolution. Consequently, the first property holds.

2. Next, select $C \in M \times \mathcal{I}$ such that

$$v_0(B) = u_0(C) + h(C, B).$$

Then,

$$v_0(A) - v_0(B) \le h(C, A) - h(C, B) \le h(B, A).$$

Furthermore, because $h(C, B) + h(B, A) \ge h(C, A)$, we have $v(A) = v_0(A)$ for $A \in \mathcal{A}$. Thus, the second claim follows from [15, Proposition 4.3].

3. If $u_0(B) - u_0(A) \le h(B, A)$ for all $A, B \in M \times \mathcal{I}$, then u_0 is a subsolution. Therefore $u_0 = v_0$ and thus (72) holds. □

4.4 Large-time behavior of the generalized Lax-Oleinik semigroup

Let $u_0 : M \times \mathcal{I} \to \mathbb{R}$ be a subsolution of (12) and $T_t w$ be given by (13). Theorem 1.4 combined with Proposition 4.1 gives that the solution $u(t, x) = T_t u_0 + c_0 t$ of (41) converges, as $t \to +\infty$, to a solution of (12). Here, we investigate the convergence of the Lax-Oleinik semigroup for arbitrary initial conditions. The long-time behavior of different but related systems was studied in [3]. The proof of Theorem 1.6 is inspired by [9].

Without loss of generality, we assume that $c_0 = 0$ and we set

$$\omega(u_0) := \{\psi : M \times \mathcal{I} \to \mathbb{R} \text{ such that } \psi = T_{t_n} u_0 \text{ for some } t_n \to +\infty\}. \tag{73}$$

Next, we define the semilimits:

$$\underline{u}(A) = \sup \{\psi(A) \mid \psi \in \omega(u_0)\} \tag{74}$$

and

$$\overline{u}(A) = \inf \{\psi(A) \mid \psi \in \omega(u_0)\}. \tag{75}$$

Because the family $\{T_t u_0\}_{t>0}$ is uniformly bounded and uniformly Lipschitz, $\underline{u}$ and $\overline{u}$ are well defined.

Proposition 4.21. *Suppose that A1–A4 hold. Let $\underline{u}$ and $\overline{u}$ be given by (74) and (75), respectively. Then, $\underline{u}$ is a subsolution of (12) and $\overline{u}$ is a supersolution of (12).*

Proof. To prove that $\underline{u}$ is a subsolution of (12), we use Propositions 4.6 and 4.7. A similar argument shows that $\overline{u}$ is a supersolution. Fix $t > 0$. From (74), $\underline{u} \ge \varphi$ for any $\varphi \in \omega(u_0)$. Consequently,

$$T_t \underline{u} \ge T_t \varphi \qquad \forall\, \varphi \in \omega(u_0).$$

Thus,

$$T_t \underline{u} \ge \sup \{T_t \varphi \mid \varphi \in \omega(u_0)\}.$$

By Proposition 4.6, to prove the statement, it suffices to show that

$$\{T_t \varphi \mid \varphi \in \omega(u_0)\} = \{\varphi \mid \varphi \in \omega(u_0)\}.$$

If $\varphi \in \omega(u_0)$, we have

$$\varphi = \lim_n T_{t_n} u_0$$

for some $t_n \to +\infty$. Also, thanks to Proposition 4.12, up to a subsequence,

$$T_{t_n - t} u_0 \to \tilde{\varphi} \in \omega(u_0)$$

Thus, $T_t \tilde{\varphi} = \lim T_{t_n} u_0 = \varphi$, as desired. □

To prove Theorem 1.6, we establish the identity $\underline{u} = \overline{u} = v$ on $M \times \mathcal{I}$.

Proposition 4.22. *Suppose that A1–A4 hold. Let $\underline{u}$ and $\overline{u}$ be given by (74) and (75), respectively, and let v be the solution of (12) given by (71). We have*

$$v \leq \overline{u} \leq \underline{u} \text{ on } M \times \mathcal{I}. \tag{76}$$

Proof. Set

$$v_0(B) := \inf_{C \in M \times \mathcal{I}} \{u_0(C) + h(C, B)\}.$$

By Proposition 4.20, v_0 is the maximal subsolution below u_0; that is, $v_0 \leq u_0$ on $M \times \mathcal{I}$. Therefore, $T_t v_0 \leq T_t u_0$ on $M \times \mathcal{I}$.

We claim that, because v_0 is a subsolution, we have $T_t v_0 \to v$. First, note that $v_0 \leq T_t v_0$. Because v is the maximal subsolution with $v = v_0$ on $\mathcal{A}$, we have $v_0 \leq v$ on $M \times \mathcal{I}$. Hence, since $v = T_t v$ for any $t > 0$, we get

$$v_0 \leq T_t v_0 \leq v \text{ on } M \times \mathcal{I}.$$

Now, $v = v_0$ on $\mathcal{A}$ implies $T_t v_0 = v$ on $\mathcal{A}$ for all $t > 0$. In particular,

$$v = \lim_{t \to +\infty} T_t v_0 \text{ on } \mathcal{A}.$$

Next, the Comparison Principle (Corollary 4.19) implies the same equality on $M \times \mathcal{I}$, as claimed.

This claim, together with $T_t v_0 \leq T_t u_0$ on $M \times \mathcal{I}$, implies that $v \leq \overline{u}$ and $v \leq \underline{u}$. Since $\overline{u} \leq \underline{u}$ is clear, the result follows. □

The last step of the proof of Theorem 1.6 is to show that $v \geq \underline{u}$ on $\mathcal{A}$.

Proposition 4.23. *Suppose A1–A4 hold. Let $\underline{u}$ be given by (74) and let v be the solution of (9) given by (71). Then, $\underline{u} \leq v$ on $\mathcal{A}$.*

Proof. Let $\phi : M \times \mathcal{I} \to \mathbb{R}$ be in the ω-limit set of u_0; that is, $\phi = \lim_{n \to +\infty} T_{t_n} u_0$ for some sequence $t_n \to +\infty$. By extracting a subsequence, we can assume that $s_n := t_{n+1} - t_n \to \infty$.

We claim that

$$\phi = \lim_{n \to +\infty} T_{s_n} \phi.$$

Indeed, given $\epsilon > 0$ there exists an integer, n_ϵ, such that

$$\phi - \epsilon \leq T_{t_n} u_0 \leq \phi + \epsilon \qquad \forall n \geq n_\epsilon$$

Then, since $s_n = t_{n+1} - t_n$, we get

$$T_{s_n} \phi - \epsilon \leq T_{t_{n+1}} u_0 \leq T_{s_n} \phi + \epsilon,$$

and our claim follows.

Let $A \in \mathcal{A}$. By Proposition 4.16 there exists a critical curve γ such that $A = (x, i) \in \omega(\gamma)$; that is, $A = \lim_{\tilde{t}_n \to +\infty} \gamma(\tilde{t}_n)$ for some sequence $\tilde{t}_n \to +\infty$. By extracting a suitable subsequence of $\tilde{t}_n$, if necessary, we can assume $\tau_n := \tilde{t}_n - s_n \to +\infty$. Because γ is a critical curve and v is a subsolution, we have

$$\begin{aligned} T_{s_n} \phi(\gamma(\tilde{t}_n)) - \phi(\gamma(t + \tau_n)) &\leq \int_{t+\tau_n}^{\tilde{t}_n} L(\gamma_M, \dot{\gamma}_M, \gamma_{\mathcal{I}}) + \sum_{\ell=1}^{N+1} \psi\big(\gamma_M(t_\ell), \gamma_{\mathcal{I}}(t_\ell^-), \gamma_{\mathcal{I}}(t_\ell^+)\big) \\ &= v(\gamma(\tilde{t}_n)) - v(\gamma(t + \tau_n)). \end{aligned}$$

Set

$$\eta(t) := \lim_n \gamma(t + \tau_n),$$

through some suitable subsequence. Then, η is also a critical curve. Moreover, by letting $n \to +\infty$, we obtain

$$\phi(A) - \phi(\eta(t)) \leq v(A) - v(\eta(t)).$$

Hence, to conclude the proof, it remains to show that

$$\liminf_{t \to \infty} \{\phi(\eta(t)) - v(\eta(t))\} \leq 0.$$

For that, we observe

$$v(\eta(t)) - v(\eta(0)) = \int_0^t L(\gamma_M, \dot\gamma_M, \gamma_{\mathcal{I}}) + \sum_{\ell=1}^{N+1} \psi\big(\gamma_M(t_\ell), \gamma_{\mathcal{I}}(t_\ell^-), \gamma_{\mathcal{I}}(t_\ell^+)\big) \geq T_t u_0(\eta(t)) - u_0(\eta(0)).$$

Because $\eta(\mathbb{R}) \subset \mathcal{A}$ and $v = u_0$ on $\mathcal{A}$, we have

$$\phi(\eta(t)) - v(\eta(t)) \leq \phi(\eta(t)) - T_t u_0(\eta(t)) + u_0(\eta(0)) - v(\eta(0)) \leq \max_{B\in\mathcal{A}} \big|\phi(B) - T_t u_0(B)\big|. \tag{77}$$

Since $\phi = \lim_{n\to+\infty} T_{t_n} u_0$, (77) implies the claim. □

Proof of Theorem 1.6. Let $\underline{u}$ and $\overline{u}$ be given, respectively, by (74) and (75). Proposition 4.21 ensures these are, respectively, a subsolution and a supersolution of (12). By Propositions 4.19, 4.22, and 4.23, we have

$$v \leq \overline{u} \leq \underline{u} \leq v \text{ on } M \times \mathcal{I}.$$

Thus, these are all identities. Therefore, $\omega(u_0) = \{v\}$; that is, $\lim_{t\to+\infty} T_t u_0 = v$. □

Acknowledgments. A.Figalli is partially supported by the NSF Grants DMS-1262411 and DMS-1361122. D. Gomes was partially supported by KAUST baseline and start-up funds. D. Marcon was partially supported by the UT Austin-Portugal partnership through the FCT doctoral fellowship SFRH/BD/33919/2009.

References

[1] M. Bardi and I. Capuzzo-Dolcetta. *Optimal control and viscosity solutions of Hamilton-Jacobi-Bellman equations.* Systems & Control: Foundations & Applications. Birkhäuser Boston Inc., Boston, MA, 1997. With appendices by Maurizio Falcone and Pierpaolo Soravia.

[2] S.A. Belbas. Optimal switching control of diffusion processes: The associated implicit variational problems. In *Decision and Control including the Symposium on Adaptive Processes, 1981 20th IEEE Conference on*, volume 20, pages 1380–1383, 1981.

[3] Filippo Cagnetti, Diogo Gomes, Hiroyoshi Mitake, and Hung V. Tran. A new method for large time behavior of degenerate viscous Hamilton-Jacobi equations with convex Hamiltonians. *Annales de l'Institut Henri Poincaré. Analyse Non Linéaire*, 32(1):183–200, 2015.

[4] Fabio Camilli, Olivier Ley, Paola Loreti, and Vinh Duc Nguyen. Large time behavior of weakly coupled systems of first-order Hamilton-Jacobi equations. *NoDEA. Nonlinear Differential Equations and Applications*, 19(6):719–749, 2012.

[5] P. Cannarsa and C. Sinestrari. *Semiconcave Functions, Hamilton-Jacobi Equations, and Optimal Control.* Progress in Nonlinear Differential Equations and Their Applications. Birkhäuser Boston, 2004.

[6] I. Capuzzo Dolcetta and L. C. Evans. Optimal switching for ordinary differential equations. *SIAM Journal of Control and Optimization*, 22(1):143–161, 1984.

[7] I. Capuzzo Dolcetta, M. Matzeu, and J.-L. Menaldi. On a system of first order quasi-variational inequalities connected with optimal switching problem. *Systems and Control Letters*, 3:113–116, 1983.

[8] F. H. Clarke. *Methods of dynamic and nonsmooth optimization*, volume 57 of *CBMS-NSF Regional Conference Series in Applied Mathematics.* Society for Industrial and Applied Mathematics (SIAM), Philadelphia, PA, 1989.

[9] A. Davini and A. Siconolfi. A generalized dynamical approach to the large time behavior of solutions of Hamilton-Jacobi equations. *SIAM Journal of Mathematical Analysis*, 38(2):478–502, 2006.

[10] A. Davini and M. Zavidovique. Aubry sets for weakly coupled systems of Hamilton-Jacobi equations.

[11] L.C. Evans and A. Friedman. Optimal stochastic switching and the Dirichlet problem for the Bellman equation. *Transactions of American Mathematical Society*, 253:365–389, 1979.

[12] A. Fathi. Théorème KAM faible et théorie de Mather sur les systèmes legrangiens. *Comptes Rendus de l'Académie des Sciences*, 324:1043–1046, 1997.

[13] A. Fathi. *Weak KAM Theorem in Lagrangian Dynamics.* Preliminary Version Number 10, June 2008.

[14] A. Fathi and A. Figalli. Optimal transportation on non-compact manifolds. *Israel Journal of Mathematics*, 175:1–59, 2010.

[15] D. Gomes and A. Serra. Systems of weakly coupled Hamilton-Jacobi equations with implicit obstacles. *Canadian Journal of Mathematics*, 64:1289–1309, 2012.

[16] D. Marcon. *Weak KAM and Aubry–Mather theories in an optimal switching setting.* PhD thesis, Instituto Superior Técnico, 2013.

[17] H. Mitake, A. Siconolfi, H.V. Tran, and N. Yamada. A Lagrangian approach to weakly coupled Hamilton-Jacobi systems. *ArXiv preprint*, 2015. arxiv:1503.00521v1[math.AP].

[18] H. Mitake and H. V. Tran. A dynamical approach to the large-time behavior of solutions to weakly coupled systems of Hamilton-Jacobi equations. *Journal de Mathématiques Pures et Appliquées. Neuvième Série*, 101(1):76–93, 2014.

[19] H. Mitake and H. V. Tran. Weakly coupled systems of infinity laplace equations. *Preprint*, 2014.

[20] Hiroyoshi Mitake and Hung V. Tran. Remarks on the large time behavior of viscosity solutions of quasi-monotone weakly coupled systems of Hamilton-Jacobi equations. *Asymptotic Analysis*, 77(1-2):43–70, 2012.

[21] Hiroyoshi Mitake and Hung V. Tran. Homogenization of weakly coupled systems of Hamilton-Jacobi equations with fast switching rates. *Archive for Rational Mechanics and Analysis*, 211(3):733–769, 2014.

[22] Vinh Duc Nguyen. Some results on the large-time behavior of weakly coupled systems of first-order Hamilton-Jacobi equations. *Journal of Evolution Equations*, 14(2):299–331, 2014.

Alessio Figalli

Department of Mathematics
The University of Texas at Austin
1 University Station, C1200
Austin TX 78712, USA
email: `figalli@math.utexas.edu`

Diogo Gomes

King Abdullah University of Science
and Technology (KAUST)
CEMSE Division
Thuwal 23955-6900, Saudi Arabia.
email: `diogo.gomes@kaust.edu.sa`

Diego Marcon

Instituto de Matemática e Estatística
Universidade Federal do Rio Grande do Sul (UFRGS)
Porto Alegre 91509-900, Brazil
email: `diego.marcon@ufrgs.br`