RANDOM LAX–OLEINIK SEMIGROUPS FOR HAMILTON–JACOBI SYSTEMS

ANDREA DAVINI, ANTONIO SICONOLFI AND MAXIME ZAVIDOVIQUE

Abstract. Following the random approach of [27], we define a Lax–Oleinik formula adapted to evolutive weakly coupled systems of Hamilton–Jacobi equations. It is reminiscent of the corresponding scalar formula, with the relevant difference that it has a stochastic character since it involves, loosely speaking, random switchings between the various associated Lagrangians. We prove that the related value functions are viscosity solutions to the system, and establish existence of minimal random curves under fairly general hypotheses. Adding Tonelli like assumptions on the Hamiltonians, we show differentiability properties of such minimizers, and existence of adjoint random curves. Minimizers and adjoint curves are trajectories of a twisted generalized Hamiltonian dynamics.

Introduction

The aim of the paper is to define a Lax–Oleinik formula adapted to evolutive weakly coupled Hamilton–Jacobi systems and study its main properties. The system can be written as

$$\partial_t u_i + H_i(x, D_x u_i) + \big(B\mathbf{u}(t,x)\big)_i = 0 \qquad \text{in } (0,+\infty)\times\mathbb{R}^N \tag{1}$$

for $i \in \{1,\dots,m\}$, where $\mathbf{u} = (u_1,\cdots,u_m)$ is the unknown function, and the H_i are unrelated Hamiltonians satisfying rather standard conditions, see Section 2. The hypotheses taken on the coupling matrix $B = (b_{ij})$ correspond to suitable monotonicity properties of the equations with respect to the entries u_j, see Remark 2.1. They are complemented by a degeneracy condition requiring all the rows of B to sum to 0, yielding that $-B$ is generator of a semigroup of stochastic matrices.

It is worth pointing out the relevance of such a formula in the case of a single equation. Besides providing a variational way to represent viscosity solutions of related evolutionary or stationary equations, it enters crucially into play in a variety of theoretical constructions and problems. Just to give some examples of its range of application, we mention that the Weak KAM Theory, as developed by Fathi [18], is founded on the Lax–Oleinik formula. Bernard's construction of regular subsolutions relies on a perturbation of a suitable initial datum via alternate application of the two conjugate Lax–Oleinik semigroups [3]. The variational representation formulae for solutions of Hamilton–Jacobi equations was exploited as a key tool to establish several asymptotic results, such as homogenization in random media [33, 31], large–time behavior of solutions [17, 13, 22], selection principles in the ergodic approximation of the Hamilton–Jacobi equation [12].

A dynamical interpretation of the system setting is illuminating and provides some insight on our method. At least when the H_i satisfy Tonelli like regularity assumptions,

---

Date: August 5, 2016.

2010 *Mathematics Subject Classification.* 35F21, 49L25, 37J50.

Key words and phrases. weakly coupled systems of Hamilton–Jacobi equations, viscosity solutions, weak KAM Theory.

the m Hamiltonian dynamics related to the H_i can be viewed as possible evolutions of a system, with coupling term providing random switching criteria. Randomness being governed by the continuous–time Markov chain with $-B$ as transition matrix. In this context, the adapted Lax–Oleinik formula is devised to define the associated expectation semigroup.

The pattern can be thought as a nonlinear version of the so–called random evolution, a topic initiated by Reuben Hersh at the end of the sixties and pursued by several authors as Griego and Pinsky, see [19, 30]. The theory provides a mathematical frame to models where evolving systems modify the mode of motion because of random changes in the environment.

Mostly using a pure PDE approach, weakly coupled systems have been recently widely investigated, in the stationary as well as in the time–dependent version. The main purpose being to find parallels, under the aforementioned degeneracy assumption on the coupling matrix, with Weak KAM theory for scalar Eikonal equations.

This stream of research was initiated in [7], where the authors studied homogenization à la Lions–Papanicolaou–Varadhan [26], and pursued in [8] with the proof of time convergence results for solutions of evolutive problems, under hypotheses close to [28]. Other outputs in this vein can be found in a series of works including [29, 6]. The links with weak KAM theory were further made precise by two of the authors of the present paper (AD and MZ) in [14] where, among other things, an appropriate notion of Aubry set for systems was given and some relevant properties of it were generalized from the scalar case. This study partially relies on the properties of the semigroup associated to the evolutive system (1), but, due to the inability to provide a variational representation for it, such properties are established purely by means of PDE tools and viscosity solution techniques. Weak KAM Theory relies instead on the intertwining of PDE techniques, variational arguments and ideas borrowed from dynamical systems via Lax–Oleinik formula.

A dynamical and variational point of view of the matter, integrating the PDE methods, was brought in by the third author (AS) with collaborators in [27, 20]. This angle allowed detecting the stochastic character of the problem, displayed by the random switching nature of the dynamics related to systems. This approach led to the definition of an adapted random action functional, which constitutes a key tool of our analysis as well. Representation formulae for viscosity (sub)solutions to stationary systems and a cycle characterization for the Aubry set were derived.

A key output of the present paper is to provide another piece of the dictionary between Weak KAM Theory and weakly coupled systems of Hamilton–Jacobi equations. The study herein carried on is to be regarded as a step in the direction of a deeper understanding of the phenomena taking place at the critical level of the stationary version of (1). In this line of research, further issues to be addressed certainly include an investigation on the dynamical and geometric properties of the Aubry set and on the differentiability properties of critical subsolutions on this set, as well as an extension to systems of the theory of minimizing Mather measures. The variational formulae provided in this paper seem to be the right tools for this kind of analysis.

Presentation of the main results. All the results can be localized by classical arguments of finite speed of propagation. Therefore, we have preferred to state them in the Euclidean space $\mathbb{R}^N$, keeping in mind that they remain valid on any manifold.

The Lax–Oleinik formula for systems is obtained in the form of infima of expectations, where the infimum is over a suitable class of random curves. The admissible random curves for this procedure are defined in Section 4. The nonlinear character of the setting makes the structure of the formula more involved than in the original linear random evolution models. In this case, in fact, the expectation semigroup is simply obtained via concatenation on any sample path of the deterministic semigroups related to the switching operators plus averaging. The nonlinearity brings in a sense a commutation problem between infimum and expectation. In this framework, we perform a key step in the analysis, notably in view of studying viscosity solutions of the evolutive system, by establishing a differentiation formula for Lipschitz–continuous functions on admissible curves, see Theorem 4.7.

Due to its random character, the formula is painful to handle directly. It is not easy to show for instance that it defines a semigroup of operators on suitable functional spaces or that the associated value functions are continuous or even semicontinuous in (t, x). For this reason, we resort to a rather indirect approach putting it in relation to the system via a sub–optimality principle and showing first that the value function, for a suitable initial datum, is a viscosity subsolution to the system in the discontinuous sense, see Section 5. The procedure is not new, but the vectorial character of the problem and the random setting add a number of additional difficulties. The implementation therefore requires some new tools and ideas.

Under mild regularity conditions on the initial datum, we moreover prove in Section 6 the existence of minimizing random curves, namely curves realizing the infimum in the Lax–Oleinik formula. This is somehow surprising since in general the presence of expectation operators makes such an output quite difficult to obtain. Our strategy is composed of two steps. We first untangle the randomness and tackle the optimization problem on any sample path, obtaining in this way, in general, multiple deterministic minimizers, and then build the sought random minimizer by performing a measurable selection.

We get in Section 7 more information on the regularity of minimizers assuming Tonelli like conditions on the H_i. Given any such minimal random curve, we prove, for almost all fixed sample path, differentiability in any bounded interval up to a finite number of points. We derive differentiability of the solution of the system on such curves plus existence of an adjoint random curve. Minimizers and adjoint curves are governed by a twisted generalized Hamiltonian dynamics. As extra consequence, we recover from the scalar case regularizing properties of the action of the associated semigroup on bounded Lipschitz–continuous initial data.

To complete the outline of the paper, we further point out that Section 1 contains preliminary material plus notations and terminology, Section 2 collects some basic facts and definitions on systems, and Appendix A is devoted to the proofs of some needed results for both systems and time–dependent equations.

We would finally like to stress that in the random part, see Section 3, we avoid as much as possible technicalities and advanced probabilistic notions working on spaces of càdlàg and continuous paths. Hopefully, it makes the presentation palatable for PDE oriented readers.

1. Preliminaries

With the symbols $\mathbb{N}$ and $\mathbb{R}_+$ we will refer to the sets of positive integer numbers and nonnegative real numbers, respectively. Given $k \in \mathbb{N}$, we denote by $\mathcal{L}^k$ the Lebesgue measure in $\mathbb{R}^k$. Given $E \subset \mathbb{R}^k$, we say that a property holds *almost everywhere* (*a.e.* for short) in E if it holds up to a subset of E with vanishing $\mathcal{L}^k$ measure. We say that E has *full measure* if $\mathcal{L}^k(\mathbb{R}^k \setminus E) = 0$. We write $\langle \cdot, \cdot \rangle$ for the scalar product in $\mathbb{R}^k$. We will denote by $\overline{E}$ the closure of E. We will denote by $B_r(x)$ and B_r the open Euclidean ball of radius

$r > 0$ centered at $x \in \mathbb{R}^k$ and 0, respectively. By the term *curve*, we mean throughout the paper a locally absolutely continuous curve.

Let E be a Borel subset E of $\mathbb{R}^k$. Given a measurable function $g : E \to \mathbb{R}$, we will denote by $\|g\|_{L^\infty(E)}$ the usual L^∞–norm of g. When $\mathbf{g}$ is vector–valued, i.e. $\mathbf{g} : E \to \mathbb{R}^d$, we will write

$$\|\mathbf{g}\|_{L^\infty(E)} := \max_{1 \leqslant i \leqslant d} \|g_i\|_{L^\infty(E)}.$$

The above notation will be mostly used in the case when either $E = [0,T] \times \mathbb{R}^N$ or $E = \mathbb{R}^N$. In the latter case, we will often write $\|\mathbf{g}\|_\infty$ in place of $\|\mathbf{g}\|_{L^\infty(\mathbb{R}^N)}$.

We will denote by $\big(\mathrm{BUC}(\mathbb{R}^N)\big)^m$ the space of bounded uniformly continuous functions $\mathbf{u} = (u_1, \dots, u_m)^T$ from $\mathbb{R}^N$ to $\mathbb{R}^m$ (where the upper–script symbol T stands for the transpose). A function $\mathbf{u} : \mathbb{R}^N \to \mathbb{R}^m$ will be termed Lipschitz continuous if each of its components is κ–Lipschitz continuous, for some $\kappa > 0$. Such a constant κ will be called a *Lipschitz constant* for $\mathbf{u}$. The space of all such functions will be denoted by $\big(\mathrm{Lip}(\mathbb{R}^N)\big)^m$. Analogously, for every fixed $T > 0$, we will denote by $\big(\mathrm{BUC}([0,T] \times \mathbb{R}^N)\big)^m$ and $\big(\mathrm{Lip}([0,T] \times \mathbb{R}^N)\big)^m$ the space of bounded uniformly continuous functions and Lipschitz continuous functions from $(0,T) \times \mathbb{R}^N$ to $\mathbb{R}^m$, respectively.

We will denote by $\mathbb{1} = (1, \cdots, 1)^T$ the vector of $\mathbb{R}^m$ with all components equal to 1. We consider the following partial relations between elements $\mathbf{a}, \mathbf{b} \in \mathbb{R}^m$: $\mathbf{a} \leqslant \mathbf{b}$ (respectively, $\mathbf{a} < \mathbf{b}$) if $a_i \leqslant b_i$ (resp., $<$) for every $i \in \{1, \dots, m\}$. Given two functions $\mathbf{u}, \mathbf{v} : \mathbb{R}^N \to \mathbb{R}^m$, we will write $\mathbf{u} \leqslant \mathbf{v}$ in $\mathbb{R}^N$ (respectively, $<$) to mean that $\mathbf{u}(x) \leqslant \mathbf{v}(x)$ $\big($resp., $\mathbf{u}(x) < \mathbf{v}(x)\big)$ for every $x \in \mathbb{R}^N$.

Given n subsets A_i of $\mathbb{R}^k$ and n scalars λ_i, $i = 1, \cdots, n$, we define

$$\sum_{i=1}^n \lambda_i \, A_i = \left\{ x = \sum_i \lambda_i, y_i \mid y_i \in A_i \text{ for any } i \right\}. \tag{1.2}$$

We give some definitions and results of set–valued analysis we will need in what follows, the material is taken from [10]. Let X, Y be Polish spaces, namely complete, separable metric spaces, endowed with the Borel σ–algebras $\mathcal{F}_X$, $\mathcal{F}_Y$. We denote by Z a map from X to the compact (nonempty) subsets of Y. Given $E \subset Y$, we set

$$Z^{-1}(E) = \{x \in X \mid Z(x) \cap E \neq \emptyset\}.$$

Definition 1.1. *The set–valued map Z is said* upper semicontinuous *if $Z^{-1}(E)$ is closed for any closed subset E of Y; Z is said* measurable *if $Z^{-1}(E) \in \mathcal{F}_X$ for any closed (or alternatively open) subset E of Y.*

The next selection result is a simplified version, adapted to our needs, of Theorem III.8 in [10].

Theorem 1.2. *If the compact–valued map Z is measurable then it admits a measurable selection, namely there exists a measurable function $f : X \to Y$ with $f(x) \in Z(x)$ for any x and $f^{-1}(\mathcal{F}_Y) \subset \mathcal{F}_X$.*

Given a locally Lipschitz continuous function $u : \mathbb{R}^k \to \mathbb{R}$ and $x \in \mathbb{R}^k$ we define the *Clarke generalized gradient* at x as

$$\partial^C u(x) = \mathrm{co}\{p \in \mathbb{R}^k \mid p = \lim_n Du(x_n), \, x_n \to x\}$$

where co stands for the convex envelope and the approximating sequences x_n are made up by differentiability points of u. Recall that the function u is differentiable in a set of full $\mathcal{L}^k$ measure thanks to Rademacher Theorem. We record for later use

Proposition 1.3. *Given a locally Lipschitz continuous function $u : \mathbb{R}^k \to \mathbb{R}$, the map $x \mapsto \partial^C u(x)$ is convex compact valued and upper semicontinuous.*

Even if the following statement is well known, we provide a proof for reader's convenience.

Lemma 1.4. *Let $u : \mathbb{R}^k \to \mathbb{R}$, $\eta : \mathbb{R}_+ \to \mathbb{R}^k$ be a locally Lipschitz continuous function and a locally absolutely continuous curve, respectively. Let $s \geqslant 0$ be such that $t \mapsto u\big(\eta(t)\big)$ and $t \mapsto \eta(t)$ are both differentiable at s. Then*

$$\frac{\mathrm{d}}{\mathrm{d}t} u\big(\eta(t)\big)\Big|_{t=s} = \langle p, \dot\eta(s)\rangle \quad \textit{for some } p \in \partial^C u\big(\eta(s)\big).$$

Proof. The function $u \circ \eta$ is clearly locally absolutely continuous. We start from the relation

$$\limsup_{\substack{y \to x \\ h \to 0^+}} \frac{u(y + h\, q) - u(y)}{h} = \max\{\langle p, q\rangle \mid p \in \partial^C u(x)\} \tag{1.3}$$

which holds true for any x, q in $\mathbb{R}^k$, see [11, pp. 195–196 and 208]. If s satisfies the assumptions then

$$\frac{\mathrm{d}}{\mathrm{d}t} u\big(\eta(t)\big)\Big|_{t=s} = \lim_{h \to 0} \frac{u\big(\eta(s+h)\big) - u\big(\eta(s)\big)}{h}$$

and

$$\lim_{h \to 0} \frac{u\big(\eta(s+h)\big) - u\big(\eta(s) + h\,\dot\eta(s)\big)}{h} = 0$$

which implies

$$\frac{\mathrm{d}}{\mathrm{d}t} u\big(\eta(t)\big)\Big|_{t=s} = \lim_{h \to 0^+} \frac{u\big(\eta(s) + h\,\dot\eta(s)\big) - u\big(\eta(s)\big)}{h}$$

and taking into account (1.3) we get

$$\frac{\mathrm{d}}{\mathrm{d}t} u\big(\eta(t)\big)\Big|_{t=s} \leqslant \max\{\langle p, \dot\eta(s)\rangle \mid p \in \partial^C u(x)\}. \tag{1.4}$$

We further have

$$-\frac{\mathrm{d}}{\mathrm{d}t} u\big(\eta(t)\big)\Big|_{t=s} = \lim_{h \to 0^+} \frac{u\big(\eta(s) + h\,(-\dot\eta(s))\big) - u\big(\eta(s)\big)}{h}$$

and consequently

$$\begin{aligned} \frac{\mathrm{d}}{\mathrm{d}t} u\big(\eta(t)\big)\Big|_{t=s} &\geqslant -\max\big\{\big\langle p, \big(-\dot\eta(s)\big)\big\rangle \mid p \in \partial^C u(x)\big\} \\ &= \min\big\{\big\langle p, \dot\eta(s)\big\rangle \mid p \in \partial^C u(x)\big\} \end{aligned} \tag{1.5}$$

Bearing in mind that $\partial^C u\big(\eta(s)\big)$ is convex, we directly deduce the assertion from (1.4) and (1.5). □

We write down, in view of future use, a version of Denjoy–Young–Saks Theorem, see [32, pp. 17–19]. A definition is preliminarily needed: for a real valued function f, the upper right and lower right Dini derivative at a point s are given, respectively, by

$$\limsup_{h\to 0^+} \frac{f(s+h)-f(s)}{h}, \qquad \liminf_{h\to 0^+} \frac{f(s+h)-f(s)}{h}.$$

Theorem 1.5. *Let f be a real valued function defined on an interval. Then outside a set of vanishing $\mathcal{L}^1$ measure the following condition holds true: if f is not differentiable at s then one of the two right Dini derivatives must be infinite.*

As at a point s where f admits a right derivative, both right Dini derivatives are finite (and equal), an immediate corollary is:

Corollary 1.6. *Let f be a real valued function defined on an interval. If f is right differentiable a.e. then it is differentiable a.e.*

2. Weakly coupled systems

We consider the evolutionary weakly coupled system

$$\partial_t u_i + H_i(x, D_x u_i) + \big(B\mathbf{u}(t,x)\big)_i = 0 \quad \text{in } (0,+\infty)\times\mathbb{R}^N \quad i\in\{1,\dots,m\}, \tag{HJS}$$

where we have denoted by $\mathbf{u}(t,x) = \big(u_1(t,x),\dots,u_m(t,x)\big)^T$ the vector–valued unknown function. We assume the Hamiltonians H_i to satisfy, for $i\in\{1,\dots,m\}$

(H1) $\quad H_i \in \mathrm{UC}(\mathbb{R}^N\times B_R)\quad$ for every $R>0$;

(H2) $\quad p\mapsto H_i(x,p)\quad$ is convex on $\mathbb{R}^N$ for any $x\in\mathbb{R}^N$;

(H3) $\quad$ there exist two superlinear functions $\alpha,\beta:\mathbb{R}_+\to\mathbb{R}$ such that

$$\alpha(|p|)\leqslant H_i(x,p)\leqslant \beta(|p|) \quad \text{for every } (x,p)\in\mathbb{R}^N\times\mathbb{R}^N.$$

By *superlinear* we mean that

$$\lim_{h\to+\infty}\frac{\alpha(h)}{h} = \lim_{h\to+\infty}\frac{\beta(h)}{h} = +\infty.$$

It is easily seen that the continuity modulus of H_i in $\mathbb{R}^N\times B_R$ and the functions α, β can be chosen independently of $i\in\{1,\dots,m\}$.

We define the *Fenchel transform* $L_i:\mathbb{R}^N\times\mathbb{R}^N\to\mathbb{R}$ of H_i via

$$L_i(x,q) := \sup_{p\in\mathbb{R}^N}\{\langle p,q\rangle - H_i(x,p)\}. \tag{2.1}$$

The function L_i is called the Lagrangian associated with the Hamiltonian H_i and satisfies properties analogous to (H1)–(H3).

The *coupling matrix* $B=(b_{ij})$ has dimensions $m\times m$ and satisfies

(B) $\quad b_{ij}\leqslant 0 \quad \text{for } j\neq i, \qquad \sum_{j=1}^m b_{ij} = 0 \qquad$ for every $i\in\{1,\dots,m\}$.

We will denote by $\big(B\mathbf{u}(t,x)\big)_i$ the i–th component of the vector $B\mathbf{u}(t,x)$, i.e.

$$\big(B\mathbf{u}(t,x)\big)_i = \sum_{j=1}^{m} b_{ij} u_j(t,x).$$

Remark 2.1. The weakly coupled system (HJS) is a particular type of *monotone system*, i.e. a system of the form $G_i\big(t,x,u_1(x),\dots,u_m(x),D_x u_i\big) = 0$ in $\mathbb{R}^N$ for every $i \in \{1,\dots,m\}$, where suitable monotonicity conditions with respect to the u_j–variables are assumed on the functions G_i, see [7, 15, 21, 23, 25]. In the case under investigation, the conditions assumed on the coupling matrix imply, in particular, that each function G_i is non–decreasing in u_i, and non–increasing in u_j for $j \neq i$, for every $i \in \{1,\dots,m\}$.

Given a function u on $(0,+\infty)\times\mathbb{R}^N$, we will call *subtangent* (respectively, *supertangent*) of u at $(t_0,x_0) \in (0,+\infty)\times\mathbb{R}^N$ a function ϕ of class C^1 in a neighborhood of (t_0,x_0) such that $u-\phi$ has a local minimum (resp., maximum) at x_0. The differentials of subtangents (resp. supertangents) $\big(\partial_t\phi(t_0,x_0), D_x\phi(t_0,x_0)\big)$ make up the *subdifferential* (resp. *superdifferential*) of u at (t_0,x_0), denoted $D^-u(t_0,x_0)$ $\big($resp. $D^+u(t_0,x_0)\big)$. The function ϕ will be furthermore termed *strict subtangent* (resp., *strict supertangent*) if $u-\phi$ has a *strict* local minimum (resp., maximum) at (t_0,x_0).

Given a function $\mathbf{u} : \mathbb{R}_+ \times \mathbb{R}^N \to \mathbb{R}^m$ locally bounded from above (resp. from below), we define its *upper semicontinuous envelope* $\mathbf{u}^*$ (resp. *lower semicontinuous envelope* $\mathbf{u}_*$) as follows:

$$u_i^*(t,x) := \limsup_{(s,y)\to(t,x)} u_i(s,y) \qquad \text{for every } (t,x)\in\mathbb{R}_+\times\mathbb{R}^N \text{ and } i\in\{1,\dots,m\},$$

(resp. $u_{i*}(t,x) := \liminf_{(s,y)\to(t,x)} u_i(s,y)$ for every $(t,x)\in\mathbb{R}_+\times\mathbb{R}^N$ and $i\in\{1,\dots,m\}$).

Definition 2.2. *We will say that* $\mathbf{u} : (0,+\infty)\times\mathbb{R}^N \to \mathbb{R}^m$ *locally bounded from above is a* viscosity subsolution *of* (HJS) *if*

$$p_t + H_i(x,p_x) + \big(B\mathbf{u}^*(t,x)\big)_i \leqslant 0$$

for every $(t,x,i) \in (0,+\infty)\times\mathbb{R}^N\times\{1,\dots,m\}$, $(p_t,p_x)\in D^+u_i^*(t,x)$.

We will say that $\mathbf{u} : (0,+\infty)\times\mathbb{R}^N \to \mathbb{R}^m$ *locally bounded from below is a* viscosity supersolution *of* (HJS) *if*

$$p_t + H_i(x,p_x) + \big(B\mathbf{u}_*(t,x)\big)_i \geqslant 0$$

for every $(t,x,i) \in (0,+\infty)\times\mathbb{R}^N\times\{1,\dots,m\}$, $(p_t,p_x)\in D^-u_{i*}(t,x)$.

We will say that a locally bounded function $\mathbf{u}$ *is a* viscosity solution *if it is both a sub and a supersolution.*

In the sequel, solutions, subsolutions and supersolutions will be always meant in the viscosity sense, hence the adjective *viscosity* will be omitted.

Due to the continuity and convexity properties of the Hamiltonians H_i, we have:

Proposition 2.3. *Let* $\mathbf{u}$ *be locally Lipschitz in* $(0,+\infty)\times\mathbb{R}^N$. *The following properties are equivalent*

(i) $\mathbf{u}$ *is a (viscosity) subsolution of* (HJS)*;*

(ii) $\mathbf{u}$ *is an almost everywhere subsolution, i.e. for any* $i\in\{1,\dots,m\}$

$$\partial_t u_i(t,x) + H_i\big(x, D_x u_i(t,x)\big) + \big(B\mathbf{u}(t,x)\big)_i \leqslant 0 \quad \textit{for a.e. } (t,x)\in(0,+\infty)\times\mathbb{R}^N;$$

(iii) $\mathbf{u}$ *is a Clarke subsolution, i.e.*

$$r + H_i\big(x,p\big) + \big(B\mathbf{u}(t,x)\big)_i \leqslant 0 \qquad \text{for every } (r,p) \in \partial^C u_i(t,x),$$

for every $(t,x,i) \in (0,+\infty) \times \mathbb{R}^N \times \{1,\ldots,m\}$.

The following holds:

Theorem 2.4. *Let* $\mathbf{u}^0 \in \big(\mathrm{BUC}(\mathbb{R}^N)\big)^m$. *There exists a unique solution* $\mathbf{u}(t,x)$ *of* (HJS) *in* $(0,+\infty)\times\mathbb{R}^N$ *agreeing with* $\mathbf{u}^0$ *at* $t=0$, *which belongs to* $\big(\mathrm{BUC}([0,T]\times\mathbb{R}^N)\big)^m$ *for any* $T>0$. *If* $\mathbf{u}^0$ *is furthermore assumed Lipschitz continuous, then, for every* $T>0$, $\mathbf{u} \in \big(\mathrm{Lip}([0,T]\times\mathbb{R}^N)\big)^m$, *with Lipschitz constant solely depending on* $H_1,\ldots,H_m$ *and* B, *on* $\|\mathbf{u}^0\|_{L^\infty(\mathbb{R}^N)}$, $\|D\mathbf{u}^0\|_{L^\infty(\mathbb{R}^N)}$, *and on* T.

The uniqueness of the solutions provided by the previous theorem is in fact a consequence of the following comparison principle:

Proposition 2.5. *Let* $\mathbf{v}$, $\mathbf{w} : [0,+\infty)\times\mathbb{R}^N \to \mathbb{R}^m$ *be an upper semicontinuous subsolution and a lower semicontinuous supersolution of* (HJS), *respectively. Let us assume that* $\mathbf{v}$ *and* $\mathbf{w}$ *are bounded in* $[0,T]\times\mathbb{R}^N$, *for every* $T>0$, *and that either* $\mathbf{v}(0,\cdot)$ *or* $\mathbf{w}(0,\cdot)$ *are in* $\mathrm{BUC}(\mathbb{R}^N)$. *Then*

$$v_i(t,x) - w_i(t,x) \leqslant \max_{1\leqslant i\leqslant m} \sup_{\mathbb{R}^N} \big(v_i(0,\cdot) - w_i(0,\cdot)\big)$$

for all $(t,x) \in [0,+\infty)\times\mathbb{R}^N$ *and* $i \in \{1,\ldots,m\}$.

The above stated results are, essentially, a consequence of what is proved in [7], see Appendix A for more details.

3. Random frame

3.1. **Definitions and terminology.** In this subsection we make precise the random frame in which our analysis takes place. This is basically a way of introducing the Markov chain generated by $-B$.

Following the constructive approach of [27], we take as sample space the space of paths

$$\omega : \mathbb{R}_+ \to \{1,\ldots,m\}$$

that are right–continuous and possess left–hand limits, denoted by Ω. These are known in literature as *càdlàg paths*, a French acronym for *continu à droite, limite à gauche*.

We refer the reader to the magnificent book of Billingsley [4] for a detailed treatment of the topics. By càdlàg property and the fact that the range of $\omega \in \Omega$ is finite, the points of discontinuity of any such path are isolated and consequently finite in compact intervals of $\mathbb{R}_+$ and countable (possibly finite) in the whole of $\mathbb{R}_+$. We call them *jump times* of ω.

The space Ω is endowed with a distance, named after *Skorohod*, see [4], which turns it into a Polish space. We denote by $\mathcal{F}$ the corresponding Borel σ–algebra and, for every $t \geqslant 0$, by $\pi_t : \Omega \to \{1,\ldots,m\}$ the map that evaluates each ω at t, i.e. $\pi_t(\omega) = \omega(t)$ for every $\omega \in \Omega$. It is known that $\mathcal{F}$ is the minimal σ–algebra that makes all the functions π_t measurable, i.e. $\pi_t^{-1}(i) \in \mathcal{F}$ for every $i \in \{1,\ldots,m\}$ and $t \geqslant 0$. In other terms, the family $\mathcal{C}$ of *cylinders*

$$\mathcal{C}(t_1,\cdots,t_k; i_1,\cdots,i_k) = \{\omega \in \Omega \ \mid \omega(t_1) = i_1,\ldots,\omega(t_k) = i_k\},$$

with $0 \leqslant t_1 < t_2 < \cdots < t_k$, $i_1,\ldots,i_k \in \{1,\ldots,m\}$ and $k \in \mathbb{N}$, generates $\mathcal{F}$. Given $t \geqslant 0$, the σ–algebra generated by the cylinders of $\mathcal{C}$ enjoying the additional property that $t_k \leqslant t$, is denoted by $\mathcal{F}_t$. Then $\{\mathcal{F}_t\}_{t\geqslant 0}$ is a *filtration of* $\mathcal{F}$, i.e. $\mathcal{F}_s \subseteq \mathcal{F}_t$ for every $0 \leqslant s < t$ and $\cup_{t\geqslant 0}\mathcal{F}_t = \mathcal{F}$. It is in addition *right–continuous* in the sense that $\mathcal{F}_t = \cap_{s>t}\mathcal{F}_s$ for any t. Note that $\mathcal{F}_0$ comprises a finite number of sets, namely Ω, $\varnothing$, $\Omega_i := \{\omega \in \Omega \,:\, \omega(0) = i\,\}$ for every $i \in \{1, \dots, m\}$, and unions of such sets. The cylinders constitute a *separating class*, in the sense that any probability measure on $\mathcal{F}$ is identified by the values taken on $\mathcal{C}$, see Theorem 16.6 in [4].

Let μ be a probability measure on $(\Omega, \mathcal{F})$. Given $E \in \mathcal{F}$, we define the *restriction* of μ to E as $\mu \llcorner E(F) = \mu(E \cap F)$ for any $F \in \mathcal{F}$. The probability μ conditioned to the event $E \in \mathcal{F}$ is defined as

$$\mu(F \,|\, E) := \frac{\mu(F \cap E)}{\mu(E)} \qquad \text{for every } F \in \mathcal{F},$$

where we agree that $\mu(F \,|\, E) = 0$ whenever $\mu(E) = 0$.

Let us now fix an $m \times m$ matrix B satisfying assumption (B). We record that e^{-tB} is a stochastic matrix for every $t \geqslant 0$, namely a matrix with nonnegative entries and with each row summing to 1, see for instance Appendix A in [27]. We endow Ω of a probability measure $\mathbb{P}$ defined on the σ–algebra $\mathcal{F}$ in such a way that the right–continuous process $(\pi_t)_{t\geqslant 0}$ is a *Markov chain with generator matrix* $-B$, i.e. it satisfies the Markov property

$$\mathbb{P}\big(\omega(t_k) = i_k \,|\, \omega(t_1) = i_1, \dots, \omega(t_{k-1}) = i_{k-1}\,\big) = \Big(\mathrm{e}^{-B(t_k - t_{k-1})}\Big)_{i_{k-1} i_k} \tag{3.1}$$

for all times $0 \leqslant t_1 < t_2 < \cdots < t_k$, states $i_1, \dots, i_k \in \{1, \dots, m\}$ and $k \in \mathbb{N}$. For the existence and an explicit construction of such a probability measure, we refer the reader to [27]. We will denote by $\mathbb{P}_i$ the probability measure $\mathbb{P}$ conditioned to the event Ω_i and write $\mathbb{E}_i$ for the corresponding expectation operators. These entities will constitute the basic building blocks of our analysis. It is easily seen that the Markov property (3.1) holds with $\mathbb{P}_i$ in place of $\mathbb{P}$, for every $i \in \{1, \dots, m\}$.

We proceed by introducing some more notations and terminology. We call *random variable* a map $X : (\Omega, \mathcal{F}) \to \big(\mathbb{F}, \mathscr{B}(\mathbb{F})\big)$, where $\mathbb{F}$ is a Polish space and $\mathscr{B}(\mathbb{F})$ its Borel σ–algebra, satisfying $X^{-1}(A) \in \mathcal{F}$ for every $A \in \mathscr{B}(\mathbb{F})$.

Given a probability measure μ on $(\Omega, \mathcal{F})$, we denote by $X_\#\mu$ the *push–forward* of μ through the map X, i.e. the probability measure on $\mathscr{B}(\mathbb{F})$ defined as

$$(X_\#\mu)\,(A) := \mu\,(\{\omega \in \Omega \,:\, X(\omega) \in A\,\}) \quad \text{for every } A \in \mathscr{B}(\mathbb{F}).$$

A probability measure ν on $\mathcal{I} := \{1, \dots, m\}$ will be identified with a *probability vector* $\mathbf{a} \in \mathbb{R}^m$, i.e. a vector with nonnegative components summing to 1, via the formula $\mathbf{a} \cdot \mathbf{b} := \int_{\mathcal{I}} \mathbf{b} \, \mathrm{d}\nu(i)$ for every $\mathbf{b} \in \mathbb{R}^m$.

3.2. **Basic facts on $\mathbb{P}_i$ and $\mathbb{E}_i$.** In this subsection we gather for later use some properties of probability measures $\mathbb{P}_i$ and the corresponding expectation operators.

Lemma 3.1. *There exists a constant C such that*

$$\mathbb{P}_i\big(\{\omega \mid \omega(t_1) \neq \omega(t_2)\}\big) \leqslant C\,|t_1 - t_2| \qquad \textit{for any } i \in \{1, \dots, m\},\ t_1,\ t_2 \textit{ in } \mathbb{R}_+.$$

Proof. We fix an index i and assume $t_2 > t_1$. We denote by F the set in the statement. Then $F = \bigcup_{j=1}^m \bigcup_{k\neq j} \mathcal{C}(t_1, t_2; k, j)$ and

$$\mathbb{P}_i\big(\mathcal{C}(t_1, t_2; k, j)\big) = \mathbb{P}_i\big(\omega(t_2) = j \mid \omega(t_1) = k\big)\,\mathbb{P}_i\big(\omega(t_1) = k\big).$$

By making use of the Markov property (3.1) for $\mathbb{P}_i$ and of the fact that e^{-tB} is a stochastic matrix we infer

$$\mathbb{P}_i(F) \leqslant \sum_{\substack{j,k=1\\ j\neq k}}^{m} (\mathrm{e}^{-t_1 B})_{ik}\, (\mathrm{e}^{-(t_2-t_1)B})_{kj} \leqslant \sum_{\substack{j,k=1\\ j\neq k}}^{m} (\mathrm{e}^{-(t_2-t_1)B})_{kj}.$$

The assertion is obtained by taking into account that the rightmost term in the above inequality is 0 when $t_2 = t_1$ and all the entries of the matrix e^{-tB} are Lipschitz continuous functions in $\mathbb{R}_+$. $\square$

Given $t \geqslant 0$ and an index i, the components of $\pi_{t\#}\mathbb{P}_i$ are equal to $\mathbb{P}_i\big(\mathcal{C}(t;j)\big)$, for any i, and we deduce from the definition of $\mathbb{P}_i$

$$\pi_{t\#}\mathbb{P}_i = \mathbf{e}_i\, \mathrm{e}^{-Bt}.$$

This implies for $0 < s < t$

$$\mathbb{E}_i v_{\omega(t)} = \sum_j v_j\, \pi_{t\#}\mathbb{P}_i(j) = \big(\mathrm{e}^{-Bt}\mathbf{v}\big)_i = \sum_j (\mathrm{e}^{-B(t-s)}\mathbf{v})_j\, \pi_{s\#}\mathbb{P}_i(j) \tag{3.2}$$

for any $\mathbf{v} \in \mathbb{R}^m$. We aim to extend the above formula with a random variable taking values in $\mathbb{R}^m$ in place of a constant vector. The task will be performed via approximation by simple random variables, a difficulty is that while a constant vector is trivially $\mathcal{F}_0$–measurable, a general $\mathbb{R}^m$–valued random variable is related in a more involved way to the filtration $\mathcal{F}_t$. As a preliminary step, we recall from [27, Lemma 3.4].

Lemma 3.2. *Let $s \geqslant 0$ and $E \in \mathcal{F}_s$. Then for every $i \in \{1, \dots, m\}$ and $t \geqslant s$*

$$\pi_{t\#}\,(\mathbb{P}_i \llcorner E) = \big(\pi_{s\#}\,(\mathbb{P}_i \llcorner E)\big)\, \mathrm{e}^{-B(t-s)}.$$

Proposition 3.3. *Let $s \geqslant 0$ and $i \in \{1, \dots, m\}$. Let $\mathbf{g} = (g_1, \cdots, g_m)$ be an $\mathcal{F}_s$–measurable random variable taking values in $\mathbb{R}^m$, which is, in addition, bounded in Ω_i. Then*

$$\mathbb{E}_i\left[g_{\omega(t)}(\omega)\right] = \mathbb{E}_i\Big[\big(\mathrm{e}^{-B(t-s)}\mathbf{g}(\omega)\big)_{\omega(s)}\Big] \qquad \textit{for every } t \geqslant s. \tag{3.3}$$

Proof. We first assume $\mathbf{g}$ to be simple, namely $\mathbf{g} = \sum_{k=1}^{l} \xi^k\, \chi_{E_k}$ for some $l \in \mathbb{N}$, vectors $\xi^k \in \mathbb{R}^m$ and $\mathcal{F}_s$–measurable sets $E_k \subset \Omega$. By exploiting Lemma 3.2, we get

$$\begin{aligned}
\mathbb{E}_i\left[g_{\omega(t)}(\omega)\right] &= \sum_k \mathbb{E}_i[\xi^k_{\omega(t)}\, \chi_{E_k}] = \sum_k \sum_j \xi^k_j\, (\mathbb{P}_i \llcorner E_k)\big(\mathcal{C}(t;j)\big) \\
&= \sum_k \big(\pi_{t\#}\,(\mathbb{P}_i \llcorner E_k)\big) \cdot \xi^k = \sum_k \Big(\pi_{s\#}\,(\mathbb{P}_i \llcorner E_k)\ \mathrm{e}^{-B(t-s)}\Big) \cdot \xi^k \\
&= \sum_k \sum_j \Big(\mathrm{e}^{-B(t-s)}\, \xi^k\Big)_j\ (\mathbb{P}_i \llcorner E_k)\big(\mathcal{C}(s;j)\big) = \sum_k \mathbb{E}_i\left[\Big(\mathrm{e}^{-B(t-s)}\, \xi^k\Big)_{\omega(s)}\, \chi_{E_k}\right] \\
&= \mathbb{E}_i\Big(\mathrm{e}^{-B(t-s)}\mathbf{g}(\omega)\Big)_{\omega(s)}.
\end{aligned}$$

This shows the assertion for simple random variables. For a general $\mathbf{g}$, there exists, see [24, Theorem 1.4.4, Chapter 1], a sequence of $\mathbb{R}^m$–valued $\mathcal{F}_s$–measurable random variables $\mathbf{g}_n$ with $\mathbf{g}_n(\omega) \to \mathbf{g}(\omega)$ for all $\omega \in \Omega$ and $\mathbf{g}_n$ bounded in Ω_i. Since (3.3) holds true for $\mathbf{g}_n$ thanks to the first part of the proof, we pass to the limit on both side of the formula (3.3) exploiting the boundedness of $\mathbf{g}$ on Ω_i and using the Dominated Convergence Theorem. This ends the proof. $\square$

Differentiating under the integral sign, we derive from (3.3):

Proposition 3.4. *Let $s \geqslant 0$ and $i \in \{1, \dots, m\}$. Let $\mathbf{g}$ be an $\mathcal{F}_s$–measurable random variable taking values in $\mathbb{R}^m$, bounded in Ω_i. Then the function $t \mapsto \mathbb{E}_i\left[g_{\omega(t)}(\omega)\right]$ is differentiable in $(s, +\infty)$ and right–differentiable at s. Moreover*

$$\frac{\mathrm{d}}{\mathrm{d}t}\mathbb{E}_i\left[g_{\omega(t)}(\omega)\right] = -\mathbb{E}_i\left[\left(\mathrm{e}^{-B(t-s)}\, B\mathbf{g}(\omega)\right)_{\omega(s)}\right] \quad \textit{for every } t \geqslant s,$$

where the above formula must be understood in the sense of right differentiability at $t = s$.

It is worth pointing out that what matters most in the later application of the above result is actually the right differentiability of the expectations at the initial time s.

4. Admissible curves

4.1. **Definition and basic properties.** A major role in our construction will be played by the notion of *admissible curve.* In the sequel and throughout the paper, we will denote by $\mathrm{C}\left(\mathbb{R}_+; \mathbb{R}^N\right)$ the Polish space of continuous paths taking values in $\mathbb{R}^N$, endowed with a metric that induces the topology of local uniform convergence in $\mathbb{R}_+$.

Definition 4.1. *We call admissible curve a random variable $\gamma : \Omega \to \mathrm{C}\left(\mathbb{R}_+; \mathbb{R}^N\right)$ such that*

(i) it is uniformly (in ω) locally (in t) absolutely continuous, i.e. given any bounded interval I and $\varepsilon > 0$, there is $\delta_\varepsilon > 0$ such that

$$\sum_j (b_j - a_j) < \delta_\varepsilon \;\Rightarrow\; \sum_j |\gamma(b_j, \omega) - \gamma(a_j, \omega)| < \varepsilon \tag{4.1}$$

for any finite family $\{(a_j, b_j)\}$ of pairwise disjoint intervals contained in I and for any $\omega \in \Omega$;

(ii) it is nonanticipating, *i.e. for any $t \geqslant 0$*

$$\omega_1 \equiv \omega_2 \text{ in } [0, t] \quad \Rightarrow \gamma(\cdot, \omega_1) \equiv \gamma(\cdot, \omega_2) \text{ in } [0, t]. \tag{4.2}$$

The latter condition, with $t = 0$, implies that for any admissible curve γ, $\gamma(0, \omega)$ is constant on Ω_i, $i \in \{1, \dots, m\}$. We refer to this value as the starting point *of γ on Ω_i.*

We will say that γ is an admissible curve starting at $x \in \mathbb{R}^N$ when $\gamma(0, \omega) = x$ for every $\omega \in \Omega$.

Remark 4.2. Item (i) in Definition 4.1 is equivalent to one of the following statements, see [5, Theorem 2.12]:

(a) the derivatives $\dot\gamma(t, \omega)$ have locally equi–absolutely continuous integrals, i.e. for any bounded interval I in $\mathbb{R}_+$ and $\varepsilon > 0$ there is $\delta > 0$ such that

$$\sup_{\omega \in \Omega} \int_J |\dot\gamma(t, \omega)|\, \mathrm{d}t < \varepsilon \qquad \text{for any } J \subset I \text{ with } |J| < \delta;$$

(b) there exists a superlinear function $\Theta : \mathbb{R}_+ \to \mathbb{R}$ (that can be taken convex and increasing as well) such that, for every bounded interval I in $\mathbb{R}_+$,

$$\sup_{\omega \in \Omega} \int_I \Theta\big(|\dot\gamma(t, \omega)|\big)\, \mathrm{d}t < +\infty.$$

This in particular implies that lengths of the curves $t \mapsto \gamma(t,\omega)$ in I are equi–bounded with respect to ω. Item (ii) will be crucial in the subsequent analysis and can be equivalently rephrased by requiring that $\gamma(t,\cdot)$ is *adapted*, for any t, to the filtration $\mathcal{F}_t$, meaning that

$$\gamma(t,\cdot) : \Omega \to \mathbb{R}^N \quad \text{is } \mathcal{F}_t\text{–measurable for any } t.$$

Being the paths $s \mapsto \gamma(s,\omega)$ continuous, this is in turn equivalent to a joint–measurability condition that will be essentially exploited in what follows. More precisely, γ is *progressively measurable*, in the sense that for any $t \geqslant 0$ the map

$$\gamma : [0,t] \times \Omega \to \mathbb{R}^N \quad \text{is } \mathscr{B}([0,t]) \otimes \mathcal{F}_t\text{–measurable.} \tag{4.3}$$

It is understood that in all the previous measurability conditions $\mathbb{R}^N$ is equipped with the Borel σ–algebra corresponding to the natural topology.

It is clear that the admissible curves make up a vector space with the natural sum and product by a scalar. We proceed by establishing some differentiability properties for this kind of curves.

Lemma 4.3. *For any admissible curve γ the set*

$$\{(t,\omega) \in \mathbb{R}_+ \times \Omega \,:\, \gamma(\cdot,\omega) \text{ is not differentiable at } t\}$$

belongs to the product σ–algebra $\mathscr{B}(\mathbb{R}_+) \otimes \mathcal{F}$ and has vanishing $\mathcal{L}^1 \times \mathbb{P}$ measure.

Proof. Since measurability properties of a vector valued map and those of its components are equivalent, we can assume without loosing generality that $m = 1$. We set for any (t,ω)

$$\begin{aligned}
D^+\gamma(t,\omega) &= \limsup_{s\to 0} \frac{\gamma(t+s,\omega)-\gamma(t,\omega)}{s} = \lim_{k\to+\infty} \sup_{\substack{|s|<1/k\\ s\in\mathbb{Q}\setminus\{0\}}} \frac{\gamma(t+s,\omega)-\gamma(t,\omega)}{s}\\
D^-\gamma(t,\omega) &= \liminf_{s\to 0} \frac{\gamma(t+s,\omega)-\gamma(t,\omega)}{s} = \lim_{k\to+\infty} \inf_{\substack{|s|<1/k\\ s\in\mathbb{Q}\setminus\{0\}}} \frac{\gamma(t+s,\omega)-\gamma(t,\omega)}{s}.
\end{aligned}$$

From the above formulae and the fact that the admissible curves make up a vector space, we derive that both $D^+\gamma(t,\omega)$, $D^-\gamma(t,\omega)$ are $\mathscr{B}(\mathbb{R}_+) \otimes \mathcal{F}$ measurable, then the set in the statement belongs to $\mathscr{B}(\mathbb{R}_+) \otimes \mathcal{F}$ as well, since it can be expressed as

$$\{(t,\omega) \mid D^+\gamma(t,\omega) > D^-\gamma(t,\omega)\}.$$

Moreover its ω–sections have zero Lebesgue measure, being $\gamma(\cdot,\omega)$ an absolutely continuous curve for a.e. $\omega \in \Omega$. This implies that it has vanishing $\mathcal{L}^1 \times \mathbb{P}$ measure, as it was claimed. □

Thanks to the previous result, the map

$$\dot\gamma : [0,+\infty) \times \Omega \to \mathbb{R}^N$$

associating to any (t,ω) the derivative of $\gamma(\cdot,\omega)$ at t is well defined, up to giving an arbitrary value on the $\mathcal{L}^1 \times \mathbb{P}$–null set where the derivative does not exist. Such a map is progressively measurable, as it is clarified by the next

Proposition 4.4. *The map $\dot\gamma$ is $\mathscr{B}([0,t]) \otimes \mathcal{F}_t$–measurable.*

Proof. Looking at the definition of $D^+\gamma$, $D^-\gamma$ provided in the proof of Lemma 4.3, and taking into account (4.3) and that the filtration $\{\mathcal{F}_t\}$ is right–continuous, we see that both $D^+\gamma$, $D^-\gamma$ are $\mathscr{B}([0,t]) \otimes \mathcal{F}_t$ progressively measurable and that for any fixed t

$$\dot\gamma = D^+\gamma = D^-\gamma \qquad \text{in } [0,t] \times \Omega$$

up to a set of vanishing $\mathcal{L}^1 \times \mathbb{P}$ measure belonging to $\mathscr{B}([0,t]) \otimes \mathcal{F}_t$. This gives the assertion. $\square$

Corollary 4.5. *For any $i \in \{1, \dots, m\}$, a.e. $s \in \mathbb{R}_+$ we have*

$$\lim_{h \to 0^+} \mathbb{E}_i \left[\frac{1}{h} \int_s^{s+h} |\dot{\gamma}(t,\omega)| \, \mathrm{d}t \right] = \mathbb{E}_i \Big[|\dot{\gamma}(s,\omega)| \Big]. \tag{4.4}$$

Proof. Due to the joint measurability property proved in the previous proposition, we get for any $T > 0$ via Fubini's Theorem

$$\int_0^T \mathbb{E}_i |\dot{\gamma}(t,\omega)| \, \mathrm{d}t = \mathbb{E}_i \left[\int_0^T |\dot{\gamma}(t,\omega)| \, \mathrm{d}t \right]$$

and the integral in the right–hand side is finite because $\int_0^T |\dot{\gamma}(t,\omega)| \, \mathrm{d}t$ is bounded uniformly in ω, see Remark 4.2. This implies that the function $t \mapsto \mathbb{E}_i \Big[|\dot{\gamma}(s,\omega)| \Big]$ is locally summable in $\mathbb{R}_+$, so that by Lebesgue Differentiation Theorem adapted to the Lebesgue measure, namely not requiring the shrinking neighborhoods of a given time s to be centered at s, we get

$$\lim_{h \to 0^+} \frac{1}{h} \int_s^{s+h} \mathbb{E}_i \big[|\dot{\gamma}(t,\omega)| \big] \, \mathrm{d}t = \mathbb{E}_i \big[|\dot{\gamma}(s,\omega)| \big] \quad \text{for a.e. } s \in \mathbb{R}_+, \tag{4.5}$$

again applying Fubini Theorem we have

$$\mathbb{E}_i \left[\frac{1}{h} \int_s^{s+h} |\dot{\gamma}(t,\omega)| \, \mathrm{d}t \right] = \frac{1}{h} \int_s^{s+h} \mathbb{E}_i |\dot{\gamma}(t,\omega)| \, \mathrm{d}t. \tag{4.6}$$

The assertion is a direct consequence of (4.5), (4.6). $\square$

4.2. **Lipschitz continuous functions and admissible curves.** We proceed by studying the behavior of a Lipschitz continuous function on an admissible curve. The first result is

Proposition 4.6. *Let $\mathbf{u} : \mathbb{R}_+ \times \mathbb{R}^N \to \mathbb{R}^m$ be a locally Lipschitz function and γ an admissible curve. For every index $i \in \{1, \dots, m\}$, the function*

$$t \mapsto \mathbb{E}_i \big[u_{\omega(t)} \big(t, \gamma(t,\omega) \big) \big]$$

is locally absolutely continuous in $\mathbb{R}_+$.

Proof. We denote by f the function in object. We fix an index i and $\varepsilon > 0$. We consider a bounded interval I and a finite family of pairwise disjoint intervals $\{(a_j, b_j)\}$ contained in I.

Taking into account that $\gamma(0,\omega)$ is constant in Ω_i and item (i) in Definition 4.1, we see that the curve γ lies in a given bounded set B for $t \in I$ and $\omega \in \Omega_i$. We denote by R an upper bound of $\mathbf{u}$ in $I \times B$. Owing to item (i) in Definition 4.1 and the fact that $\mathbf{u}$ is locally Lipschitz continuous, the functions $t \mapsto u_k\big(t, \gamma(t,\omega)\big)$ are equi–absolutely continuous in I, for $k \in \{1, \dots, m\}$, $\omega \in \Omega$. We can therefore determine a positive constant δ with

$$\sum_j (b_j - a_j) < \delta \quad \Rightarrow \quad \sum_j \big| u_k\big(b_j, \gamma(b_j,\omega)\big) - u_k\big(a_j, \gamma(a_j,\omega)\big) \big| < \frac{\varepsilon}{2} \tag{4.7}$$

$$\delta \quad < \quad \frac{1}{2\,R\,C}, \tag{4.8}$$

where C is the constant appearing in the statement of Lemma 3.1. We claim that

$$\sum_j (b_j - a_j) < \delta \ \Rightarrow \ \sum_j |f(b_j) - f(a_j)| < \varepsilon. \tag{4.9}$$

We set

$$F_j = \{\omega \mid \omega(a_j) \neq \omega(b_j)\}.$$

We know from Lemma 3.1 that

$$\mathbb{P}_i(F_j) \leqslant C\,(b_j - a_j) \qquad \text{for any } j. \tag{4.10}$$

We have

$$\begin{aligned}
\sum_j |f(b_j) - f(a_j)| \ &\leqslant \ \sum_j \mathbb{E}_i \left[\left| u_{\omega(b_j)}\big(b_j, \gamma(b_j,\omega)\big) - u_{\omega(a_j)}\big(a_j, \gamma(a_j,\omega)\big) \right| \right] \\
&\leqslant \ \sum_j \int_{\Omega \setminus F_j} \left[\left| u_{\omega(b_j)}\big(b_j, \gamma(b_j,\omega)\big) - u_{\omega(b_j)}\big(a_j, \gamma(a_j,\omega)\big) \right| \right] \mathrm{d}\mathbb{P}_i \\
&+ \ \int_{F_j} \left[\left| u_{\omega(b_j)}\big(b_j, \gamma(b_j,\omega)\big) - u_{\omega(a_j)}\big(a_j, \gamma(a_j,\omega)\big) \right| \right] \mathrm{d}\mathbb{P}_i
\end{aligned}$$

and we conclude, recalling the role of R and (4.7), (4.8), (4.10)

$$\sum_j |f(b_j) - f(a_j)| < \frac{\varepsilon}{2} + 2\,R\,C \sum_j (b_j - a_j) < \frac{\varepsilon}{2} + \frac{\varepsilon}{2} = \varepsilon.$$

This proves (4.9) and concludes the proof. □

We go on proving that time derivative of a locally Lipschitz continuous function on an admissible curve and expectations $\mathbb{E}_i$ commute, up to a term which, roughly speaking, records the indices jumps on the underlying paths and contains the coupling matrix.

To comment on it, let us take for simplicity γ deterministic and $\mathbf{u}$, γ both of class C^1. By linearity the difference quotient of $t \mapsto \mathbb{E}_i\big[u_{\omega(t)}\big(t, \gamma(t)\big)\big]$ is given by

$$\mathbb{E}_i\Big[\frac{u_{\omega(t+h)}\big(t+h, \gamma(t+h)\big) - u_{\omega(t)}\big(t, \gamma(t)\big)}{h}\Big].$$

Owing to right continuity of ω, the integrand ω–pointwise converges to the time derivative of $u_{\omega(t)}$ on γ at t but, due to indices jumps, it is not bounded in Ω so that the Dominated Convergence Theorem cannot be applied to get the corresponding convergence of expectations. In this framework the extra term with the coupling matrix pops up.

This is the main output of the section and will be exploited to prove some properties of the Lax–Oleinik formula in Section 5.

Theorem 4.7. *Let $\mathbf{u} : \mathbb{R}_+ \times \mathbb{R}^N \to \mathbb{R}^m$ be a locally Lipschitz function and γ an admissible curve. Then, for every index $i \in \{1, \dots, m\}$, we have*

$$\frac{\mathrm{d}}{\mathrm{d}t}\mathbb{E}_i\big[u_{\omega(t)}\big(t, \gamma(t,\omega)\big)\big]\Big|_{t=s} = \mathbb{E}_i\Big[-\big(B\mathbf{u}\big)_{\omega(s)}\big(s, \gamma(s,\omega)\big) + \frac{\mathrm{d}}{\mathrm{d}t} u_{\omega(s)}\big(t, \gamma(t,\omega)\big)\Big|_{t=s}\Big] \tag{4.11}$$

for a.e. $s \in \mathbb{R}_+$.

For the proof we need some preliminary material. We consider the map

$$(t,\omega) \mapsto \mathbf{u}\big(t, \gamma(t,\omega)\big),$$

for some admissible curve γ. Thanks to the fact that $\mathbf{u}$ is (Lipschitz) continuous and γ jointly measurable, we derive that such map is also measurable from $\mathscr{B}(\mathbb{R}_+) \otimes \mathcal{F}$ to $\mathscr{B}(\mathbb{R}^m)$. We can therefore argue as in Lemma 4.3 to get:

Lemma 4.8. *For any locally Lipschitz continuous function $\mathbf{u} : \mathbb{R}_+ \times \mathbb{R}^N \to \mathbb{R}^m$ and any admissible curve γ the set*

$$\big\{(t,\omega) \mid t \mapsto \mathbf{u}\big(t, \gamma(t,\omega)\big) \text{ is not differentiable at } t\,\big\}$$

belongs to the product σ–algebra $\mathscr{B}(\mathbb{R}_+) \otimes \mathcal{F}$ and has vanishing $\mathcal{L}^1 \times \mathbb{P}$ measure.

Proof of Theorem 4.7. The difference quotient of $t \mapsto \mathbb{E}_i\big[u_{\omega(t)}\big(t,\gamma(t)\big)\big]$ at s is equal to $\mathbb{E}_i\left[\psi_h(\omega) + \varphi_h(\omega)\right]$, with

$$\begin{aligned}
\psi_h(\omega) \quad &:= \quad \frac{u_{\omega(s+h)}\big(s,\gamma(s,\omega)\big) - u_{\omega(s)}\big(s,\gamma(s,\omega)\big)}{h} \\
\varphi_h(\omega) \quad &:= \quad \frac{u_{\omega(s+h)}\big(s+h,\gamma(s+h,\omega)\big) - u_{\omega(s+h)}\big(s,\gamma(s,\omega)\big)}{h}.
\end{aligned}$$

Due to right continuity of ω, we further have

$$\varphi_h(\omega) = \frac{u_{\omega(s)}\big(s+h,\gamma(s+h,\omega)\big) - u_{\omega(s)}\big(s,\gamma(s,\omega)\big)}{h}, \tag{4.12}$$

for $h > 0$ small enough, with smallness depending on ω.

Keeping s frozen, we apply Proposition 3.4 to $\mathbf{g}(\omega) = \mathbf{u}\big(s,\gamma(s,\omega)\big)$, to get

$$\lim_{h\to 0^+} \mathbb{E}_i\big[\psi_h(\omega)\big] = \frac{\mathrm{d}^+}{\mathrm{d}t}\left(\mathbb{E}_i\big[u_{\omega(t)}\big(s,\gamma(s,\omega)\big)\big]\right)\Big|_{t=s} = -\mathbb{E}_i\big[(B\mathbf{u})_{\omega(s)}\big(s,\gamma(s,\omega)\big)\big], \tag{4.13}$$

where the symbol $\dfrac{\mathrm{d}^+}{\mathrm{d}t}$ stands for the right derivative. The assumptions in Proposition 3.4 are actually satisfied: in fact $\omega \mapsto \mathbf{u}\big(s,\gamma(s,\omega)\big)$ is bounded in Ω_i because of item (i) in Definition 4.1 and the fact that γ has constant value on Ω_i at $t = 0$. It is in addition $\mathcal{F}_s$–measurable because $\mathbf{u}$ is continuous and γ adapted to the filtration $\{\mathcal{F}_t\}$.

To handle the term φ_h, we restrict the choice of s. By Lemma 4.8, we know that the set

$$N := \big\{(t,\omega) \mid t \mapsto \mathbf{u}\big(t,\gamma(t,\omega)\big) \text{ is not differentiable at } t\,\big\}$$

has vanishing $\mathcal{L}^1 \times \mathbb{P}$ measure, and consequently its s–sections N_s have vanishing probability for s varying in a set J of full measure in $\mathbb{R}_+$. We therefore deduce from (4.12) that

$$\varphi_h(\omega) \underset{h\to 0^+}{\longrightarrow} \frac{\mathrm{d}}{\mathrm{d}t} u_{\omega(s)}\big(t,\gamma(t,\omega)\big)\Big|_{t=s} \qquad \text{for } s \in J,\ \omega \in \Omega \setminus N_s. \tag{4.14}$$

Due to Corollary 4.5, we can assume, without any loss to generality, that the $s \in J$ also satisfy the limit relation (4.4). We have for $h \leqslant 1$

$$\begin{aligned}
|\varphi_h(\omega)| \quad &= \quad \left|\frac{u_{\omega(s+h)}\big(s+h,\gamma(s+h,\omega)\big) - u_{\omega(s+h)}\big(s,\gamma(s,\omega)\big)}{h}\right| \\
&\leqslant \quad 1 + \frac{1}{h}\,\kappa\left(\big|\gamma(s+h,\omega) - \gamma(s,\omega)\big|\right) \leqslant 1 + \frac{1}{h}\,\kappa\left(\int_s^{s+h} |\dot{\gamma}(t,\omega)|\,\mathrm{d}t\right)
\end{aligned} \tag{4.15}$$

where κ a Lipschitz constant for $\mathbf{u}$ in $[s,s+1]\times B$, and B stands for the bounded set containing the curves $\gamma(\cdot,\omega)$, for $t\in[s,s+1]$, as ω varies in Ω, see item (i) in Definition 4.1. In addition, by (4.4)

$$\mathbb{E}_i\left[\frac{1}{h}\int_s^{s+h}|\dot\gamma(t,\omega)|\,\mathrm{d}t\right] \qquad \text{is convergent as } h\to 0^+. \tag{4.16}$$

Therefore, when $s\in J$, the sequence φ_h is a.e. pointwise convergent thanks to (4.14), and dominated by another sequence with convergent $\mathbb{E}_i$ expectation in force of (4.15), (4.16). This allows using the variant of Dominated Convergence Theorem (see for instance [16, Theorem 4, Chapter 1.3]) to get

$$\lim_{h\to 0^+}\mathbb{E}_i\left[\varphi_h(\omega)\right]=\mathbb{E}_i\left[\frac{\mathrm{d}}{\mathrm{d}t}u_{\omega(s)}\big(t,\gamma(t,\omega)\big)_{\big|t=s}\right] \quad \text{for } s\in J. \tag{4.17}$$

Owing to (4.13), (4.17), the function

$$s\mapsto \mathbb{E}_i\big[u_{\omega(s)}\big(s,\gamma(s,\omega)\big)\big]=\mathbb{E}_i[\psi_h(\omega)+\varphi_h(\omega)]$$

is a.e. right–differentiable in $\mathbb{R}_+$, and so a.e. differentiable in view of Denjoy–Young–Saks Theorem, see Corollary 1.6. Formula (4.11) directly comes from (4.13), (4.17). □

Given the set

$$\big\{(t,\omega)\in(0,+\infty)\times\Omega \mid t\mapsto \mathbf{u}\big(t,\gamma\big(t,\omega\big)\big) \text{ and } t\mapsto\gamma(t,\omega) \text{ are not differentiable at } t\big\} \tag{4.18}$$

we denote by J the set of points $s>0$ such that the s–section of the set in (4.18) has probability 0 and (4.4) holds at s. Note that J has full measure in $\mathbb{R}_+$ because of Lemma 4.3, Lemma 4.8 and Corollary 4.5.

Lemma 4.9. *Let* $\mathbf{u}$, γ *be as in Theorem 4.7 and let* $s\in J$. *The compact–valued map*

$$Z(\omega)=\left\{(r,p)\in\partial^C u_{\omega(s)}\big(s,\gamma(s,\omega)\big)\mid r+\langle p,\dot\gamma(s,\omega)\rangle=\frac{\mathrm{d}}{\mathrm{d}t}u_\omega(s)\big(s,\gamma(s,\omega)\big)\right\}$$

is measurable.

Proof. Since the lengths of the curves $t\mapsto\gamma(t,\omega)$ in $[0,s]$ are equibounded with respect to ω, see Remark 4.2, and the elements $\gamma(0,\omega)$ are finite as ω varies in Ω, we deduce that the set $\big\{\big(s,\gamma(s,\omega)\big)\mid\omega\in\Omega\big\}$ is bounded. We denote by R a Lipschitz constant for all the u_i in such a set.

We claim that the function

$$\omega\mapsto\frac{\mathrm{d}}{\mathrm{d}t}u_{\omega(s)}\big(s,\gamma(s,\omega)\big) \text{ is measurable.} \tag{4.19}$$

Indeed, it is obtained as the composition of $\omega\mapsto(\omega,\omega(s))$, which is a measurable map from $(\Omega,\mathcal{F})$ to $\big(\Omega\times\{1,\dots,m\},\mathcal{F}\otimes\mathscr{P}(\{1,\dots,m\})\big)$ by definition of the σ–algebra $\mathcal{F}$, with $(\omega,i)\mapsto\frac{\mathrm{d}}{\mathrm{d}t}u_i\big(s,\gamma(s,\omega)\big)$, which is a $\mathscr{B}([0,s])\otimes\mathscr{P}\{\{1,\dots,m\}\}$–measurable real function, as it can be easily checked by arguing as in Lemma 4.3 and Proposition 4.4. Furthermore, since the map $\omega\mapsto\dot\gamma(s,\omega)$ is measurable in force of Proposition 4.4, we deduce that

$$\omega\mapsto\langle p,\dot\gamma(s,\omega)\rangle \text{ is measurable for any } p\in\mathbb{R}^N. \tag{4.20}$$

We proceed by showing that the compact–valued map

$$\widetilde{Z}(\omega)=\left\{(r,p)\in\mathbb{R}\times\mathbb{R}^N \mid |r|+|p|\leqslant R,\ r+\langle p,\dot\gamma(s,\omega)\rangle-\frac{\mathrm{d}}{\mathrm{d}t}u_{\omega(s)}\big(s,\gamma(s,\omega)\big)=0\right\}$$

is measurable. Taking into account Definition 1.1, it is enough to prove that $\widetilde{Z}^{-1}(K) \in \mathcal{F}$ for any compact subset K of $\mathbb{R}^{N+1}$. Let $(r_n, p_n)_n$ be a dense sequence in K, then

$$\widetilde{Z}^{-1}(K) = \bigcap_{h=1}^{\infty} \bigcup_{n=1}^{\infty} \left\{ \omega \mid \ \left| r_n + \langle p_n, \dot\gamma(s,\omega)\rangle - \frac{\mathrm{d}}{\mathrm{d}t} u_{\omega(s)}\big(s,\gamma(s,\omega)\big) \right| < \frac{1}{h} \right\}$$

and the sets appearing in the above formula belong to $\mathcal{F}$ thanks to (4.19), (4.20). This concludes the proof of the claim.

Now notice that the set–valued map $\omega \mapsto \partial^C u_{\omega(s)}\big(s,\gamma(s,\omega)\big)$ is $\mathcal{F}$–measurable. Indeed, it is obtained as the composition of $\omega \mapsto (\gamma(s,\omega), \omega(s))$, which is a measurable map from $(\Omega, \mathcal{F})$ to $\big(\mathbb{R}^N \times \{1,\dots,m\}, \mathscr{B}(\mathbb{R}^N) \otimes \mathscr{P}(\{1,\dots,m\})\big)$, with $(x,i) \mapsto \partial^C u_i\big(s,x\big)$, which is an uppersemicontinuous set–valued map defined on $\mathbb{R}^N \times \{1,\dots,m\}$.

Bearing in mind the definition of R, we find that

$$Z(\omega) = \widetilde{Z}(\omega) \cap \partial^C u_{\omega(s)}\big(s,\gamma(s,\omega)\big).$$

The map Z has nonempty compact images thanks to Propositions 1.3 and 1.4, and it is measurable as an intersection of measurable set–valued maps, see [10]. $\square$

We invoke Theorem 1.2 to derive

Corollary 4.10. *Let $\mathbf{u}$, γ, s, $Z(\omega)$ be as in Lemma 4.9, then there is a measurable selection $\omega \mapsto \big(r(s,\omega), p(s,\omega)\big)$ of $Z(\omega)$.*

Taking into account Theorem 4.7, Lemma 1.4, and the above Corollary, we get

Corollary 4.11. *Let $\mathbf{u}$, γ, i be as in Theorem 4.7, and let $s > 0$ be in J. Then the map $t \mapsto \mathbb{E}_i\big[u_{\omega(t)}\big(t,\gamma(t,\omega)\big)\big]$ is differentiable at s, and, for every $\omega \in \Omega$, one can find a measurable selection $\big(r(s,\omega), p(s,\omega)\big) \in \partial^C u_{\omega(s)}\big(s,\gamma(s,\omega)\big)$ satisfying*

$$\frac{\mathrm{d}}{\mathrm{d}t}\mathbb{E}_i\big[u_{\omega(t)}\big(t,\gamma(t,\omega)\big)\big]\Big|_{t=s} = \mathbb{E}_i\Big[- \big[B\mathbf{u}\big(s,\gamma(s,\omega)\big)\big]_{\omega(s)} + r(s,\omega) + \langle p(s,\omega), \dot\gamma(s,\omega)\rangle \Big].$$

We finally state, for the reader's convenience, a differentiability property of $t \mapsto \mathbb{E}_i\big[u_{\omega(t)}\big(t,\gamma(t)\big)\big]$, easily descending from Theorem 4.7, in the way we are going to use it in Theorem 5.4.

Corollary 4.12. *Let $\mathbf{u}$ be a C^1 function and γ a deterministic curve of class C^1. For every index $i \in \{1,\dots,m\}$, the map $t \mapsto \mathbb{E}_i\big[u_{\omega(t)}\big(t,\gamma(t)\big)\big]$ is right differentiable at $t = 0$ and*

$$\frac{\mathrm{d}^+}{\mathrm{d}t}\mathbb{E}_i\big[u_{\omega(t)}\big(t,\gamma(t)\big)\big]\Big|_{t=0} = -\big[B\mathbf{u}\big(0,\gamma(0)\big)\big]_i + \partial_t u_i\big(0,\gamma(0)\big) + \big\langle Du_i\big(0,\gamma(0)\big), \dot\gamma(0)\big\rangle.$$

5. The random Lax–Oleinik formula and its PDE counterpart

The random Lax–Oleinik formula is given by

$$\big(\mathcal{S}(t)\mathbf{u}^0\big)_i(x) = \inf_{\gamma(0,\omega)=x} \mathbb{E}_i\left[u^0_{\omega(t)}\big(\gamma(t,\omega)\big) + \int_0^t L_{\omega(s)}\big(\gamma(s,\omega), -\dot\gamma(s,\omega)\big)\,\mathrm{d}s \right] \tag{LO}$$

for every $(t,x) \in (0,+\infty) \times \mathbb{R}^N$ and $i \in \{1,\dots,m\}$, and for any bounded initial datum $\mathbf{u}^0 : \mathbb{R}^N \to \mathbb{R}^m$. Some few properties can be recovered via direct inspection of the formula. For every $(t,x) \in (0,+\infty) \times \mathbb{R}^N$, we have

$$\big(- \|\mathbf{u}^0\|_\infty + t\mu\big)\,\mathbb{1} \leqslant \mathcal{S}(t)\mathbf{u}^0(x) \leqslant \mathrm{e}^{-Bt}\mathbf{u}^0(x) + tM\,\mathbb{1}, \tag{5.1}$$

where $\mu := \inf_i \inf_{\mathbb{R}^N \times \mathbb{R}^N} L_i$ and $M := \sup_i \sup_{x \in \mathbb{R}^N} L_i(x,0)$. The leftmost inequality in (5.1) is immediate, while the second follows by taking a constant curve and by applying (3.2).

We furthermore derive from the definition

$$\|\mathcal{S}(t)\mathbf{u}^0 - \mathcal{S}(t)\mathbf{v}^0\|_\infty \leqslant \|\mathbf{u}^0 - \mathbf{v}^0\|_\infty \qquad \text{for } t \geqslant 0, \tag{5.2}$$

for any given pair of bounded functions $\mathbf{u}^0$, $\mathbf{v}^0$.

We proceed by introducing a sub–optimality principle that will allow us to link (5) to systems, and to show in this way the semigroup and continuity properties of the related value function.

Definition 5.1. *We say that a function $\mathbf{u} : \mathbb{R}_+ \times \mathbb{R}^N \to \mathbb{R}^m$ satisfies the* sub–optimality principle *if*

$$u_i\big(t_0 + h, \gamma(0)\big) - \mathbb{E}_i\big[u_{\omega(h)}\big(t_0, \gamma(h)\big)\big] \leqslant \mathbb{E}_i\left[\int_0^h L_{\omega(s)}\big(\gamma(s), -\dot\gamma(s)\big)\,\mathrm{d}s\right]$$

for any t_0, $h \geqslant 0$, $i \in \{1,\dots,m\}$ and any deterministic curve γ.

The link with the Lax–Oleinik semigroup is given by

Proposition 5.2. *Let $\mathbf{u}^0 : \mathbb{R}^N \to \mathbb{R}^m$ bounded and continuous. The function $(t,x) \mapsto \mathcal{S}(t)\mathbf{u}^0(x)$ satisfies the sub–optimality principle.*

To prove the proposition, we need some preliminary material. We denote, for any $h > 0$, by Φ_h the shift operator defined via

$$\Phi_h(\omega) = \omega(\cdot + h) \qquad \text{for any } \omega \in \Omega.$$

We recall that it is a measurable map from Ω to Ω, see [27].

Lemma 5.3. *For any index i, any $h > 0$ we have*

$$\Phi_h \# \mathbb{P}_i = \sum_{k=1}^m \left(\mathrm{e}^{-Bh}\right)_{ik} \mathbb{P}_k.$$

Proof. If $\mathcal{C}(t_1, \cdots, t_k; i_1, \cdots i_k)$ is a cylinder, then

$$\begin{aligned}
\Phi_h \# \mathbb{P}_i\big(\mathcal{C}(t_1, \cdots, t_k; i_1, \cdots i_k)\big) &= \mathbb{P}_i \Phi_h^{-1}\big(\mathcal{C}(t_1, \cdots, t_k; i_1, \cdots i_k)\big) \\
&= \mathbb{P}_i\big(\mathcal{C}(t_1 + h, \cdots, t_k + h; i_1, \cdots i_k)\big) \\
&= \sum_{k=1}^m \mathbb{P}_i\big(\mathcal{C}(h, t_1 + h, \cdots, t_k + h; k, i_1, \cdots i_k)\big) \\
&= \sum_{k=1}^m \left(\mathrm{e}^{-Bh}\right)_{ik} \mathbb{P}_k\big(\mathcal{C}(t_1, \cdots, t_k; i_1, \cdots i_k)\big).
\end{aligned}$$

The above computation proves the assertion, because the family of cylinders is a separating class, as it was pointed out in Section 3.1. □

Proof of Proposition 5.2. We select positive times h, t_0, a deterministic curve γ and $i \in \{1,\dots,m\}$, we set $x = \gamma(0)$, $y = \gamma(h)$. We fix an $\varepsilon > 0$ devoted to become infinitesimal, and pick for any $j \in \{1,\dots,m\}$ an admissible random curve ξ_j, with initial point y, such that

$$(\mathcal{S}(t_0)\mathbf{u}_0)_j(y) \geqslant \mathbb{E}_j\left[u^0_{\omega(t_0)}\big(\xi_j(t_0,\theta)\big) + \int_0^{t_0} L_{\omega(s)}(\xi_j, -\dot\xi_j)\,\mathrm{d}s\right] - \varepsilon. \tag{5.3}$$

We proceed by defining for any $(t,\omega) \in [0,+\infty) \times \Omega$

$$\eta(t,\omega) = \begin{cases} \gamma(t) & t \in [0,h) \\ \xi_{\omega(h)}\big(t-h, \Phi_h(\omega)\big) & t \in [h, t_0+h) \\ \xi_{\omega(h)}\big(t_0, \Phi_h(\omega)\big) & t \in [t_0+h, +\infty) \end{cases}$$

We claim that η is an admissible curve. We first show that for any $t \geqslant 0$, any Borel set E in $\mathbb{R}^N$

$$\Omega_E := \{\omega \mid \eta(t,\omega) \in E\} \in \mathcal{F}_t. \tag{5.4}$$

Clearly Ω_E is equal either to the whole Ω or to the empty set when $t \in [0,h]$. We focus on the case where $t \in (h, t_0+h]$, the same argument will give the property when $t > t_0 + h$. We have

$$\Omega_E = \bigcup_j \big[\{\omega \mid \xi_j\big(t-h, \Phi_h(\omega)\big) \in E\} \cap \mathcal{C}(h,j)\big]. \tag{5.5}$$

Owing to the fact that ξ_j is $\mathcal{F}_t$ adapted, for any j, and to the relation $\Phi_h^{-1}(\mathcal{F}_{t-h}) \subset \mathcal{F}_t$, see for instance [27, Proposition B.5], we further get

$$\{\omega \mid \xi_j\big(t-h, \Phi_h(\omega)\big) \in E\} = \Phi_h^{-1}\big(\{\theta \mid \xi_j(t-h,\theta) \in E\}\big) \in \Phi_h^{-1}(\mathcal{F}_{t-h}) \subset \mathcal{F}_t.$$

This gives (5.4) taking into account (5.5) and that $\mathcal{C}(h,j) \in \mathcal{F}_h \subset \mathcal{F}_t$, for any j. Being the continuous concatenation of a deterministic curve and m random admissible curves, we also see that η satisfies item (i) in Definition 4.1. We have therefore proved that it is an admissible curve, as it was claimed.

We have by Lax–Oleinik formula

$$\begin{aligned} &(\mathcal{S}(t_0+h)\mathbf{u}^0)_i(x) \leqslant \mathbb{E}_i \left[u^0_{\omega(t_0+h)}\big(\eta(t_0+h,\omega)\big) + \int_0^{t_0+h} L_{\omega(s)}\big(\eta(s,\omega), -\dot\eta(s,\omega)\big)\,\mathrm{d}s \right] \\ &= \mathbb{E}_i \left[u^0_{\Phi_h(\omega)(t_0)}\big(\xi_{\Phi_h(\omega)(0)}\big(t_0, \Phi_h(\omega)\big)\big) + \int_0^h L_{\omega(s)}\big(\gamma(s), -\dot\gamma(s)\big)\,\mathrm{d}s + \right. \\ &\left. \int_0^{t_0} L_{\Phi_h(\omega)(s)}\Big(\xi_{\Phi_h(\omega)(0)}\big(s,\Phi_h(\omega)\big), -\dot\xi_{\Phi_h(\omega)(0)}\big(s,\Phi_h(\omega)\big)\Big)\,\mathrm{d}s \right]. \end{aligned} \tag{5.6}$$

We apply the change of variable formula with $\theta = \Phi_h(\omega)$ and Lemma 5.3 to get for any $j \in \{1,\dots,m\}$

$$\begin{aligned} &\int_\Omega \Big[u^0_{\Phi_h(\omega)(t_0)}\big(\xi_{\Phi_h(\omega)(0)}\big(t_0, \Phi_h(\omega)\big)\big) \\ &+ \int_0^{t_0} L_{\Phi_h(\omega)(s)}\Big(\xi_{\Phi_h(\omega)(0)}\big(s,\Phi_h(\omega)\big), -\dot\xi_{\Phi_h(\omega)(0)}\big(s,\Phi_h(\omega)\big)\Big)\,\mathrm{d}s \Big]\mathrm{d}\mathbb{P}_i(\omega) \\ &= \int_\Omega \left[u^0_{\theta(t_0)}\big(\xi_{\theta(0)}(t_0,\theta)\big) + \int_0^{t_0} L_{\theta(s)}\big(\xi_{\theta(0)}(s,\theta), -\dot\xi_{\theta(0)}(s,\theta)\big)\,\mathrm{d}s \right] \mathrm{d}\Phi_h \# \mathbb{P}_i(\theta) \\ &= \sum_{j=1}^m \left(\mathrm{e}^{-Bh}\right)_{ij} \mathbb{E}_j \left[u^0_{\theta(t_0)}\big(\xi_j(t_0,\theta)\big) + \int_0^{t_0} L_{\theta(s)}\big(\xi_j(s,\theta), -\dot\xi_j(s,\theta)\big)\,\mathrm{d}s \right]. \end{aligned}$$

We plug the previous relation in (5.6) and use (5.3), (3.2) to get

$$
\begin{aligned}
&(\mathcal{S}(t_0+h)\mathbf{u}^0)_i(x) \leqslant \mathbb{E}_i\Big[\int_0^h L_{\omega(s)}\big(\gamma(s,\omega),-\dot\gamma(s,\omega)\big)\,\mathrm{d}s\Big]+\\
&\sum_{j=1}^m \Big(\mathrm{e}^{-Bh}\Big)_{ij}\ \mathbb{E}_j\Big[u^0_{\theta(t_0)}\big(\xi_j(t_0,\theta)\big)+\int_0^{t_0} L_{\theta(s)}\big(\xi_j(s,\theta),-\dot\xi_j(s,\theta)\big)\,\mathrm{d}s\Big] \leqslant\\
&\mathbb{E}_i\Big[\int_0^h L_{\omega(s)}\big(\gamma(s,\omega),-\dot\gamma(s,\omega)\big)\,\mathrm{d}s\Big]+\sum_{j=1}^m \Big(\mathrm{e}^{-Bh}\Big)_{ij}\ \Big((\mathcal{S}(t_0)\mathbf{u}_0)_j(y)+\varepsilon\Big)=\\
&\mathbb{E}_i\left[\int_0^h L_{\omega(s)}\big(\gamma(s,\omega),-\dot\gamma(s,\omega)\big)\,\mathrm{d}s+\big(\mathcal{S}(t_0)\mathbf{u}_0\big)_{\omega(h)}(y)+\varepsilon\right],
\end{aligned}
$$

and the assertion follows because ε is arbitrary. □

In the next result we link the sub–optimality principle with the property of being subsolution to the system (HJS). It is worth pointing out that we are not assuming any continuity or semicontinuity condition on the function appearing in the statement.

Theorem 5.4. *Let $\mathbf{u}:\mathbb{R}_+\times\mathbb{R}^N\to\mathbb{R}^m$ be a function locally bounded from above satisfying the sub–optimality principle. Then it is a subsolution to* (HJS) *in* $(0,+\infty)\times\mathbb{R}^N$.

Proof. Recall that we denote by $\mathbf{u}^*$ the upper semicontinuous envelope of $\mathbf{u}$.

Let us fix an index $i\in\{1,\dots,m\}$, $q\in\mathbb{R}^N$, $\delta>0$. Let ψ be a C^1 supertangent to u_i^* at a point $(s_0,x_0)\in(0,+\infty)\times\mathbb{R}^N$. We define a C^1, $\mathbb{R}^m$–valued function ϕ setting $\phi_i=\psi$, and choosing the other components ϕ_j of class C^1 such that

$$\phi_j(s_0,x_0)=u_j^*(s_0,x_0)+\delta \qquad \text{for } j\neq i. \tag{5.7}$$

The definition implies

$$u_i^*(s_0,x_0)=\psi(s_0,x_0)=\phi_i(s_0,x_0) \tag{5.8}$$

and

$$\phi_k\geqslant u_k^*\geqslant u_k \qquad \text{for any } k\in\{1,\dots,m\},\text{ in some neighborhood of } (s_0,x_0). \tag{5.9}$$

The argument being local, we can assume ϕ uniformly continuous with corresponding modulus denoted by ν. We consider (s_n,x_n) converging to (s_0,x_0) with $\lim_n u_i(s_n,x_n)=u_i^*(s_0,x_0)$, and set

$$\varepsilon_n=\nu(|x_n-x_0|+|s_n-s_0|)+|u_i(s_n,x_n)-u_i^*(s_0,x_0)|.$$

We proceed by selecting an infinitesimal positive sequence h_n with

$$\lim_n\frac{\varepsilon_n}{h_n}=0 \tag{5.10}$$

and define $\gamma_n(s)=x_n+s\,q$, $\gamma(s)=x_0+s\,q$. From the supertangency condition, the relation $|\gamma_n(s)-\gamma(s)|=|x_n-x_0|$ for any s, and the definition of ε_n, we derive for n large enough

$$
\begin{aligned}
\psi(s_0,x_0)-\mathbb{E}_i\big[\phi_{\omega(h_n)}\big(-h_n+s_0,\gamma(h_n)\big)\big] &\leqslant\\
&\leqslant u_i(s_n,x_n)-\mathbb{E}_i\big[\phi_{\omega(h_n)}\big(-h_n+s_n,\gamma_n(h_n)\big)\big]+\varepsilon_n\\
&\leqslant u_i(s_n,x_n)-\mathbb{E}_i\big[u_{\omega(h_n)}\big(-h_n+s_n,\gamma_n(h_n)\big)\big]+\varepsilon_n.
\end{aligned}
\tag{5.11}
$$

Further, we have by the sub–optimality principle

$$u_i(s_n, x_n) - \mathbb{E}_i\big[u_{\omega(h)}\big(-h_n + s_n, \gamma_n(h_n)\big)\big] \leqslant \mathbb{E}_i\left[\int_0^{h_n} L_{\omega(s)}(\gamma_n(s), -q)\,\mathrm{d}s\right]. \quad (5.12)$$

We apply Corollary 4.12 to the function $\widetilde{\phi}(s,x) := \phi(s_0 - s, x)$ and the curve γ. We derive from (5.7), (5.8), (5.9), (5.10), (5.11), (5.12), and assumption (B) on the coupling matrix,

$$\begin{aligned}
&\liminf_{n\to+\infty} \frac{1}{h_n}\mathbb{E}_i\left[\int_0^{h_n} L_{\omega(s)}(\gamma_n(s), -q)\,\mathrm{d}s + \varepsilon_n\right] \\
&\geqslant \lim_{n\to+\infty} \frac{1}{h_n}\left(\psi(s_0,x_0) - \mathbb{E}_i\big[\phi_{\omega(h_n)}\big(-h_n + s_0, \gamma(h_n)\big)\big]\right) \\
&= -\frac{\mathrm{d}^+}{\mathrm{d}t}\mathbb{E}_i\big[\widetilde{\phi}_{\omega(t)}\big(t,\gamma(t)\big)\big]\Big|_{t=0} \qquad (5.13) \\
&= \big[B\widetilde{\phi}\big(0,\gamma(0)\big)\big]_i - \partial_t\widetilde{\phi}_i\big(0,\gamma(0)\big) - \big\langle D_x\widetilde{\phi}_i\big(0,\gamma(0)\big), \dot{\gamma}(0)\big\rangle. \\
&= \big[B\phi(s_0,x_0)\big]_i + \partial_t\psi(s_0,x_0) - \langle D_x\psi(s_0,x_0), q\rangle \\
&\geqslant \big[B\mathbf{u}^*(s_0,x_0)\big]_i + \partial_t\psi(s_0,x_0) + \langle D_x\psi(s_0,x_0), -q\rangle - \delta\, C,
\end{aligned}$$

where $C = -\sum_{j\neq i} b_{ij}$. We know that $\mathbb{P}_i$–a.e. path ω takes the value i in a suitable right neighborhood of 0, depending on ω. From this and the continuity of L_i, we deduce for $\mathbb{P}_i$–a.e. ω

$$\begin{aligned}
&\lim_n \frac{1}{h_n}\int_0^{h_n} L_{\omega(s)}(\gamma_n(s), -q)\,\mathrm{d}s = \\
&\lim_n \frac{1}{h_n}\int_0^{h_n} L_i(\gamma_n(s), -q)\,\mathrm{d}s \\
&\lim_n \frac{1}{h_n}\int_0^{h_n} L_i(\gamma(s), -q)\,\mathrm{d}s + \frac{1}{h_n}\int_0^{h_n} \big(L_i(\gamma_n(s), -q) - L_i(\gamma(s), -q)\big)\,\mathrm{d}s \\
&= L_i(x_0, -q).
\end{aligned}$$

By the Dominated Convergence Theorem and (5.10), we thus infer

$$\lim_{n\to+\infty} \frac{1}{h_n}\,\mathbb{E}_i\left[\frac{1}{h_n}\int_0^{h_n} L_{\omega(s)}(\gamma_n(s), -q)\,\mathrm{d}s + \varepsilon_n\right] = L_i(x_0, -q).$$

We further derive from (5.10)

$$\big(B\mathbf{u}^*(s_0,x_0)\big)_i + \partial_t\psi(s_0,x_0) + \langle D_x\psi(s_0,x_0), -q\rangle - L_i(x_0,-q) - \delta\, C \leqslant 0,$$

and, being q, δ arbitrary, we finally obtain

$$\big(B\mathbf{u}^*(s_0,x_0)\big)_i + \partial_t\psi(s_0,x_0) + H_i\big(x_0, D_x\psi(s_0,x_0)\big) \leqslant 0,$$

which proves the claimed subsolution property for $\mathbf{u}^*$. □

By combining Proposition 5.2 and Theorem 5.4 we obtain

Proposition 5.5. *Let $\mathbf{u}^0 : \mathbb{R}^N \to \mathbb{R}^m$ be bounded and continuous. The function $\mathbf{v}(t,x) := \mathcal{S}(t)\mathbf{u}^0(x)$ is a subsolution of* (HJS) *satisfying $\mathbf{v}^*(0,\cdot) \leqslant \mathbf{u}^0$ on $\mathbb{R}^N$.*

Proof. The asserted subsolution property comes from $\mathbf{v}$ being locally bounded in force of (5.1), and Proposition 5.2, Theorem 5.4. The relation at $t = 0$ is readily obtained passing to the limit in the rightmost inequality of (5.1) as t goes to 0. □

We proceed by showing a sort of maximality property of the function given by Lax–Oleinik formula.

Proposition 5.6. *Let* $\mathbf{u}$ *be a locally Lipschitz continuous subsolution of* (HJS) *with* $\mathbf{u}^0 := \mathbf{u}(0,\cdot)$ *bounded, then*

$$\mathbf{u}(t,x) \leqslant \mathcal{S}(t)\mathbf{u}^0(x) \qquad \text{for any } (t,x) \in \mathbb{R}_+ \times \mathbb{R}^N.$$

Proof. We fix (t_0,x_0) and pick an admissible curve γ with initial point x. By applying Corollary 4.11, we get

$$\begin{aligned}&\frac{\mathrm{d}}{\mathrm{d}t}\mathbb{E}_i\left[u_{\omega(t)}\big(t_0-t,\gamma(t,\omega)\big)\right]\Big|_{t=s}\\ &\qquad\qquad = \mathbb{E}_i\Big[-\big[B\mathbf{u}\big(t_0-s,\gamma(s,\omega)\big)\big]_{\omega(s)} - r(s,\omega) + \langle p(s),\dot\gamma(s,\omega)\rangle\Big]\end{aligned}$$

for a.e. s and some $\big(r(s,\omega),p(s,\omega)\big) \in \partial^C u_{\omega(s)}\big(t_0+h-s,\gamma(s,\omega)\big)$. By exploiting the subsolution property of $\mathbf{u}$ and the Fenchel inequality, we further get

$$\begin{aligned}
&-\frac{\mathrm{d}}{\mathrm{d}t}\mathbb{E}_i\big[u_{\omega(t)}\big(t_0-t,\gamma(t,\omega)\big)\big]\Big|_{t=s}\\
=\;& \mathbb{E}_i\Big[\big[B\mathbf{u}\big(t_0-s,\gamma(s,\omega)\big)\big]_{\omega(s)} + r(s,\omega) + \langle p(s,\omega),-\dot\gamma(s,\omega)\rangle\Big]\\
\leqslant\;& \mathbb{E}_i\Big[\big[B\mathbf{u}(t_0-s,\gamma(s,\omega))\big]_{\omega(s)} + r(s,\omega) + H_{\omega(s)}\big(\gamma(s,\omega),p(s,\omega)\big)\\
&\qquad + L_{\omega(s)}\big(\gamma(s,\omega),-\dot\gamma(s,\omega)\big)\Big]\\
\leqslant\;& \mathbb{E}_i\left[L_{\omega(s)}\big(\gamma(s,\omega),-\dot\gamma(s,\omega)\big)\right]. \qquad (5.14)
\end{aligned}$$

We finally obtain by integrating between 0 and t_0 and by commuting integrals, which can be done by joint measurability properties of $(\gamma,\dot\gamma)$

$$u_i\big(t_0,x_0\big) - \mathbb{E}_i\big[u^0_{\omega(t_0)}\big(\gamma(t_0,\omega)\big)\big] \leqslant \mathbb{E}_i\left[\int_0^{t_0} L_{\omega(s)}\big(\gamma(s,\omega),-\dot\gamma(s,\omega)\big)\,\mathrm{d}s\right].$$

This gives the assertion for γ is an arbitrary admissible curve starting at x_0. □

We finally provide the announced PDE characterization of the random Lax–Oleinik formula.

Theorem 5.7. *Let* $\mathbf{u}^0 \in \big(\mathrm{BUC}(\mathbb{R}^N)\big)^m$. *Then* $(x,t) \mapsto \big(\mathcal{S}(t)\mathbf{u}^0\big)(x)$ *is the unique solution of* (HJS) *in* $(0.+\infty)\times\mathbb{R}^N$ *agreeing with* $\mathbf{u}^0$ *at* $t=0$ *and belonging to* $\big(\mathrm{BUC}([0,T]\times\mathbb{R}^N)\big)^m$ *for every* $T>0$.

Proof. We denote by $\mathbf{u}$ the unique solution of the system taking the value $\mathbf{u}^0$ at $t=0$ and belonging to $\big(\mathrm{BUC}([0,T]\times\mathbb{R}^N)\big)^m$ for every $T>0$, see Theorem 2.4.

Let us first assume $\mathbf{u}^0$ Lipschitz continuous on $\mathbb{R}^N$. Then $\mathbf{u}$ is Lipschitz continuous in $[0,T]\times\mathbb{R}^N$, for every $T>0$, according to Theorem 2.4. In view of Proposition 5.5 and the comparison principle for (HJS) stated in Proposition 2.5, we infer

$$\big(\mathcal{S}(t)\mathbf{u}^0\big)(x) \leqslant \mathbf{u}(t,x) \qquad \text{for every } (t,x) \in [0,T]\times\mathbb{R}^N.$$

The opposite inequality holds as well by Proposition 5.6. The assertion is then proved when the initial datum is additionally assumed Lipschitz continuous.

Let us now consider the general case $\mathbf{u}^0 \in \big(\mathrm{BUC}(\mathbb{R}^N)\big)^m$. Let $\mathbf{v}^0$ be a Lipschitz function in $\big(\mathrm{BUC}(\mathbb{R}^N)\big)^m$. By what was just proved, we know that the map $(t,x) \mapsto \big(\mathcal{S}(t)\mathbf{v}^0\big)(x)$

is a Lipschitz solution of (HJS) in $(0,T)\times\mathbb{R}^N$ taking the initial value $\mathbf{v}^0$ at $t=0$. From the comparison principle stated in Proposition 2.5 and (5.2), we infer

$$\|\mathbf{u}(t,\cdot)-\mathcal{S}(t)\mathbf{u}^0\|_\infty \leqslant \|\mathbf{u}(t,\cdot)-\mathcal{S}(t)\mathbf{v}^0\|_\infty + \|\mathcal{S}(t)\mathbf{u}^0-\mathcal{S}(t)\mathbf{v}^0\|_\infty \leqslant 2\,\|\mathbf{u}^0-\mathbf{v}^0\|_\infty$$

for every $t>0$. Using the fact that Lipschitz initial data are dense in $\big(\mathrm{BUC}(\mathbb{R}^N)\big)^m$, we eventually get the asserted identity $\mathbf{u}(t,x)=\mathcal{S}(t)\mathbf{u}^0(x)$ for every $(t,x)\in[0,+\infty)\times\mathbb{R}^N$. □

We directly derive from the previous results of the section

Corollary 5.8. *The Lax–Oleinik formula defines a semigroup of operators on both* $\big(\mathrm{BUC}(\mathbb{R}^N)\big)^m$ *and* $\big(\mathrm{Lip}(\mathbb{R}^N)\big)^m\bigcap\big(\mathrm{BUC}(\mathbb{R}^N)\big)^m$.

6. Minimal admissible curves

In this section we aim to prove the following result:

Theorem 6.1. *Let* $\mathbf{u}^0\in\big(\mathrm{BUC}(\mathbb{R}^N)\big)^m$ *and* $T>0$. *Assume that the function* $(t,x)\mapsto\mathcal{S}(t)\mathbf{u}^0(x)$ *is locally Lipschitz in* $(0,T]\times\mathbb{R}^N$. *Then, for every* $x\in\mathbb{R}^N$, $i\in\{1,\dots,m\}$, *there exists an admissible curve* $\eta:\Omega\to\mathrm{C}\big([0,T];\mathbb{R}^N\big)$, *starting at* x, *realizing the minimum for* $(\mathcal{S}(T)\mathbf{u}^0)_i(x)$ *in the Lax–Oleinik formula* (LO).

Remark 6.2. The assumption of the above theorem is always satisfied whenever $\mathbf{u}^0$ is Lipschitz continuous on $\mathbb{R}^N$, in view of Theorem 2.4.

6.1. Deterministic minimization.

Let $\mathbf{u}^0\in\big(\mathrm{BUC}(\mathbb{R}^N)\big)^m$ and $T>0$ be fixed, and denote by $\mathbf{u}(t,x)$ the unique function in $\big(\mathrm{BUC}([0,T]\times\mathbb{R}^N)\big)^m$ that solves the system (HJS) in $(0,T)\times\mathbb{R}^N$ subject to the initial condition $\mathbf{u}(0,\cdot)=\mathbf{u}^0$ in $\mathbb{R}^N$. We know that, for every $j\in\{1,\dots,m\}$, the j–th component u_j of $\mathbf{u}$ is a solution to

$$\frac{\partial u}{\partial t}+G_j(t,x,D_xu)=0 \qquad \text{in } (0,+\infty)\times\mathbb{R}^N. \tag{6.1}$$

with initial datum u_j^0, where

$$G_j(t,x,p)=H_j(x,p)+(B\mathbf{u})_j(t,x).$$

Let us denote by $L_{G_j}=L_j-(B\mathbf{u})_j$ the Lagrangian associated with G_j via the Fenchel transform. The following result holds:

Proposition 6.3. *Let* u_j *and* G_j *be as above. Then, for every* $0\leqslant a\leqslant T$ *and* $y\in\mathbb{R}^N$, *the following identity holds:*

$$u_j(T-a,y)=\inf_{\xi(a)=y}\Big(u_j^0\big(\xi(T)\big)+\int_a^T L_{G_j}\big(T-t,\xi(t),-\dot\xi(t)\big)\mathrm{d}t\Big), \tag{6.2}$$

where the infimum is taken by letting ξ *vary in the family of absolutely continuous curves from* $[a,T]$ *to* $\mathbb{R}^N$. *Moreover, such an infimum is a minimum.*

Remark 6.4. Note that there is a slight difference between (6.2) and the other deterministic formula given in Proposition A.2 of the appendix. However, both formulas are equivalent up to the change of variables $s=T-t$.

Proof. Let us first assume $\mathbf{u}^0$ Lipschitz continuous, so that u_j is Lipschitz in $[0,T]\times\mathbb{R}^N$. Then the result is a direct application of Proposition A.2.

In the general case, let $\mathbf{v}^n : [0,T]\times\mathbb{R}^N \to \mathbb{R}^m$ be a sequence of solutions of (HJS) with Lipschitz initial data $\mathbf{v}^n(0,\cdot)$, uniformly converging to $\mathbf{u}^0$. By Proposition 2.5, $\mathbf{v}^n$ uniformly converges to $\mathbf{u}$ on $[0,T]\times\mathbb{R}^N$. Denote by $U_j(T-a,y)$ the right hand side of (6.2) and denote by L_j^n the Lagrangian $L_j-(B\mathbf{v}^n)_j$. It is readily verified from the formula and the first part of the proof that, for $0\leqslant a\leqslant T$,

$$\begin{aligned}|U_j(T-a,y)-v_j^n(T-a,y)| &\leqslant \|\mathbf{u}_0-\mathbf{v}^n(0,\cdot)\|_\infty+(T-a)\|L_{G_j}-L_j^n\|_{L^\infty([0,T]\times\mathbb{R}^N)}\\ &= \|\mathbf{u}_0-\mathbf{v}^n(0,\cdot)\|_\infty+(T-a)\|(B\mathbf{u}-B\mathbf{v}^n)_j\|_{L^\infty([0,T]\times\mathbb{R}^N)}.\end{aligned}$$

It follows that $\mathbf{v}^n$ uniformly converges to $\mathbb{U}:=(U_j)_{1\leqslant j\leqslant m}$, hence $\mathbb{U}=\mathbf{u}$, as it was to be shown. $\square$

Given $0\leqslant a\leqslant T$, $y\in\mathbb{R}^N$ and $j\in\{1,\dots,m\}$, let us denote by $\Gamma_j(a,y)$ the family of absolutely continuous curves $\xi:[0,T]\to\mathbb{R}^N$ such that $\xi(s)=y$ for every $s\in[0,a]$ and $\xi_{|_{[a,T]}}$ realizes the infimum in (6.2). In what follows, the space $\mathrm{C}\big([0,T];\mathbb{R}^N\big)$ of continuous curves from the interval $[0,T]$ to $\mathbb{R}^N$ is endowed with the uniform norm, which makes it a Polish space, and the corresponding Borel σ–algebra. The following holds:

Proposition 6.5. *Let $j\in\{1,\dots,m\}$. Then*

(i) *the set $X_j^T:=\{\xi\,:\,\xi\in\Gamma_j(a,y)$ for some $(a,y)\in[0,T]\times\mathbb{R}^N\}$ is a family of equi-absolutely continuous curves in $\mathrm{C}\big([0,T];\mathbb{R}^N\big)$;*

(ii) *$\Gamma_j(a,y)$ is a compact subset of $\mathrm{C}\big([0,T];\mathbb{R}^N\big)$, for every $(a,y)\in(0,T)\times\mathbb{R}^N$;*

(iii) *the set–valued map $(a,y)\mapsto\Gamma_j(a,y)$ from $[0,T)\times\mathbb{R}^N$ to $\mathrm{C}\big([0,T];\mathbb{R}^N\big)$ is upper semicontinuous in the sense of Definition 1.1;*

(iv) *for every $y\in\mathbb{R}^N$, $0\leqslant a\leqslant T$ and $\xi\in\Gamma_j(a,y)$ we have*

$$u_j\big(T-t,\xi(t)\big)=u_j^0\big(\xi(T)\big)+\int_t^T L_{G_j}\big(T-s,\xi(s),-\dot\xi(s)\big)\mathrm{d}s$$

for every $t\in[a,T]$.

Proof. The first point is standard in Calculus of Variations and is the first step in establishing Tonelli's existence Theorem. Let us denote by $\Theta:\mathbb{R}_+\to\mathbb{R}$ a superlinear function such that

$$L_i(x,q)\geqslant\Theta(|q|)\qquad\text{for every } (x,q)\in\mathbb{R}^N\times\mathbb{R}^N \text{ and } i\in\{1,\dots,m\}.$$

Such a function Θ does exist for the Hamiltonians H_i satisfy condition (H3). Since any $\xi\in X_j^T$ is a minimizer of (6.2) for some $(a,y)\in[0,T]\times\mathbb{R}^N$, we infer

$$\int_a^T\Theta\big(|\dot\xi(t)|\big)\,\mathrm{d}t\leqslant\int_a^T L_{G_j}\big(T-t,\xi(t),-\dot\xi(t)\big)\mathrm{d}t\leqslant 2\|u_j\|_\infty,$$

yielding

$$\int_0^T\Theta\big(|\dot\xi(t)|\big)\,\mathrm{d}t=\int_0^a\Theta\big(|\dot\xi(t)|\big)\,\mathrm{d}t+\int_a^T\Theta\big(|\dot\xi(t)|\big)\,\mathrm{d}t\leqslant|\Theta(0)|T+2\|u_j\|_\infty.$$

This readily implies (i) in view of [5, Theorem 2.12].

We will prove items (ii) and (iii) by using Arzelà-Ascoli, Dunford–Pettis theorems (notice that by (i) the elements of X_j^T are equicontinuous) and the lower semicontinuity of $\xi\mapsto\int_a^T L_{G_j}\big(T-t,\xi(t),-\dot\xi(t)\big)\mathrm{d}t$, see [5, Theorem 3.6].

Let us prove (iii) first. We have to check that Γ_j satisfies Definition 1.1. Let C be a closed subset of $\mathrm{C}([0,T];\mathbb{R}^N)$, (a_n, y_n) a sequence of $\Gamma_j^{-1}(C)$ converging to some (a,y). We consider a sequence $\xi_n \in \Gamma_j(a_n,y_n)\cap C$. By the first part of the proof, we can assume that ξ_n converges, up to extracting a subsequence, to some ξ uniformly in $[0,T]$ and that $\dot\xi_n$ converges to $\dot\xi$ weakly in $\left(L^1([0,T])\right)^N$, see [5, Theorem 2.13]. Hence, passing to the limit in the equalities

$$u_j(T-a_n,y_n) - u_j^0\big(\xi_n(T)\big) = \int_{a_n}^T L_{G_j}\big(T-t,\xi_n(t),-\dot\xi_n(t)\big)\mathrm{d}t,$$

and using the lower semicontinuity of the integral functional, it follows that

$$u_j(T-a,y) - u_j^0\big(\xi(T)\big) \geqslant \int_a^T L_{G_j}\big(T-t,\xi(t),-\dot\xi(t)\big)\mathrm{d}t.$$

Since (6.2) gives the reverse inequality, we deduce that ξ belongs to $\Gamma_j(a,y)$ and clearly also to C. Therefore $(a,y)\in\Gamma_j^{-1}(C)$, so that $\Gamma_j^{-1}(C)$ is closed. This shows assertion (iii). Item (ii) follows arguing as above and taking $a_n = a$ and $y_n = y$ for every $n\in\mathbb{N}$.

Item (iv) is a consequence of the sub–optimality principle, see Proposition A.2. Indeed, we have

$$\begin{aligned}
&u_j\big(T-a,\xi(a)\big) - u_j^0\big(\xi(T)\big)\\
&= \Big(u_j\big(T-a,\xi(a)\big) - u_j\big(T-t,\xi(t)\big)\Big) + \Big(u_j\big(T-t,\xi(t)\big) - u_j^0\big(\xi(T)\big)\Big)\\
&\leqslant \int_a^t L_{G_j}\big(T-s,\xi(s),-\dot\xi(s)\big)\,\mathrm{d}s + \int_t^T L_{G_j}\big(T-s,\xi(s),-\dot\xi(s)\big)\,\mathrm{d}s.
\end{aligned}$$

The previous inequality, which is a sum of two inequalities, is actually an equality in view of (6.2). Hence both inequalities were equalities to start with, as it was to be proved. $\square$

We define a set–valued map $\Gamma : [0,T)\times\mathbb{R}^N\times\{1,\dots,m\}\to \mathrm{C}([0,T];\mathbb{R}^N)$ by setting $\Gamma(a,y,j) := \Gamma_j(a,y)$. It is compact–valued and upper semicontinuous, hence measurable. We are thus in the hypotheses of Theorem 1.2, so that there exists a measurable selection Ξ for Γ, i.e. a measurable function

$$\Xi : [0,T]\times\mathbb{R}^N\times\{1,\dots,m\}\to \mathrm{C}([0,T];\mathbb{R}^N) \tag{6.3}$$

such that

$$\Xi(a,y,j)\in\Gamma(a,y,j) \qquad \text{for every } (a,y,j)\in[0,T]\times\mathbb{R}^N\times\{1,\dots,m\}.$$

Notice that the measurability condition can be equivalently rephrased requiring

$$(t,a,y,j)\mapsto \Xi(a,y,j)(t)$$

to be measurable from $[0,T]\times[0,T]\times\mathbb{R}^N\times\{1,\dots,m\}$ to $\mathbb{R}^N$ with the natural Borel σ–algebras.

6.2. **Random minimization.** Given given $T>0$, $x\in\mathbb{R}^N$, $\mathbf{u}^0\in\big(\mathrm{BUC}(\mathbb{R}^N)\big)^m$, $i\in\{1,\dots,m\}$, we aim to show the existence of a minimizing admissible curve η for $(\mathcal{S}(T)\mathbf{u}^0)_i(x)$. We will provide a rather explicit construction via concatenation of minimizers of (6.2).

We first give a rough picture of it, just to contribute some insight. We fix ω and minimize at the initial step the deterministic functional in (6.2) in the interval $[a,T]$ with $a=0$, $j=\omega(0)$ and $y=x$. Clearly we get multiple minimizers, but, according to the results of the previous subsection, we can select one in a measurable way with respect (a,y,j). This will be crucial to show the random character of the curve obtained as output. At the first jump time of ω, say $\tau_1(\omega)$, we switch to index j accordingly and restart the procedure minimizing on $[\tau_1(\omega), T]$, with y equal to the position reached by the selected minimizing curve in the previous step at τ_1. Notice that the final time T stays untouched, which is needed to get in the end a nonanticipating random curve. We go on iterating the above procedure at any jump time of ω belonging to $[0, T]$.

We proceed by presenting a full description of the construction. We point out that this part is independent of the local Lipschitz continuity assumption on $(x,t) \mapsto \mathcal{S}(t)\mathbf{u}^0(x)$. This condition will come into play only in the proof of Theorem 6.1, to apply the derivation formula given in Theorem 4.7.

For any fixed $\omega \in \Omega$, we set $\tau_0(\omega) = 0$ and we define inductively a sequence $\big(\tau_k(\omega)\big)_k$ by setting

$$\tau_k(\omega) = \begin{cases} k\text{–th jump time} & \text{if } \omega \text{ has at least } k \text{ jump times in } [0,T] \\ T & \text{otherwise} \end{cases}$$

Let us denote by $\Xi : [0,T] \times \mathbb{R}^N \times \{1, \dots, m\} \to \mathrm{C}\big([0,T]; \mathbb{R}^N\big)$ the measurable selection introduced at the end of Section 6.1, and define inductively a sequence $\big(x_k(\omega)\big)_{k \geqslant 0}$ of points in $\mathbb{R}^N$ by setting $x_0 = x$ and

$$x_k(\omega) = \Xi\big(\tau_{k-1}(\omega), x_{k-1}(\omega), \omega(\tau_{k-1}(\omega))\big)\big(\tau_k(\omega)\big) \quad \text{for every } k \geqslant 1.$$

The following holds:

Lemma 6.6. *For every $k \in \mathbb{N}$, the maps $\omega \mapsto \tau_k(\omega)$, $\omega \mapsto \omega\big(\tau_k(\omega)\big)$ and $\omega \mapsto x_k(\omega)$ are random variables.*

Proof. We start by proving the assertion for the maps $\omega \mapsto \tau_k(\omega)$ and $\omega \mapsto \omega(\tau_{k-1}\big(\omega)\big)$. The argument is by induction on $k \geqslant 1$. Let us denote by $(t_n)_{n \in \mathbb{N}}$ a dense sequence in $(0,T)$. For every $t \in (0,T]$, we have

$$\{\omega \,:\, \tau_1(\omega) < t\} = \bigcup_{t_n < t} \{\omega(t_n) \neq \omega(0)\} \in \mathcal{F},$$

which clearly gives the asserted measurability of $\omega \mapsto \tau_1(\omega)$ (note that $\tau_k(\omega) \leqslant T$ for every $k \geqslant 1$). The fact that $\omega \mapsto \omega\big(\tau_0(\omega)\big) = \omega(0)$ is a random variable is trivial by the definition of $\mathcal{F}$. Assume now that $\omega \mapsto \tau_j(\omega)$ and $\omega \mapsto \omega\big(\tau_{j-1}(\omega)\big)$ are random variables for every $j \leqslant k$. By the inductive step we have that, for every $t \in (0,T)$,

$$\{\omega \,:\, \tau_{k+1}(\omega) < t\} = \bigcup_{t_n < t} \{\omega \,:\, \tau_k(\omega) < t_n\} \cap \big\{\omega \,:\, \omega(t_n) \neq \omega\big(\tau_k(\omega)\big)\big\} \in \mathcal{F},$$

thus showing the asserted measurability of $\omega \mapsto \tau_{k+1}(\omega)$. The fact that $\omega \mapsto \omega\big(\tau_k(\omega)\big)$ is a random variable follows from the fact that, for every $i \in \{1, \dots, m\}$,

$$\big\{\omega \,:\, \omega\big(\tau_k(\omega)\big) = i\,\big\} = \bigcup_n \{\omega \,:\, \omega(t_n) = i\,\} \cap \{\omega \,:\, \tau_k(\omega) \leqslant t_n < \tau_{k+1}(\omega)\,\}.$$

Last, the fact that the map $\omega \mapsto x_k(\omega)$ is a random variable for every $k \geqslant 0$ is again by induction on k. The measurability for $k = 0$ is trivial. Let us assume that $\omega \mapsto x_j(\omega)$ is measurable for every $j \leqslant k$. Then the map $\omega \mapsto x_{k+1}(\omega)$ is a random variable since it is the composition of the $\mathcal{F}$–measurable function $\omega \mapsto \Big(\tau_{k+1}(\omega), \tau_k(\omega), x_k(\omega), \omega\big(\tau_k(\omega)\big)\Big)$ with $(t,a,y,j) \mapsto \Xi(a,y,j)(t)$, which is measurable from $[0,T] \times [0,T] \times \mathbb{R}^N \times \{1, \dots, m\}$ to $\mathbb{R}^N$ with the natural Borel σ–algebras. $\square$

The sought curve is defined by setting $\eta(0,\omega) = x$ and, for every $k \geqslant 0$,

$$\eta(t,\omega) := \Xi\Big(\tau_k(\omega), x_k(\omega), \omega\big(\tau_k(\omega)\big)\Big)(t \wedge T) \qquad \text{if} \quad t \wedge T \in (\tau_k(\omega), \tau_{k+1}(\omega)],$$

for every $\omega \in \Omega$, where $t \wedge T := \min\{t, T\}$. Note that the curve $\eta(\cdot,\omega)$ is constant in $[T, +\infty)$, for any fixed $\omega \in \Omega$. The following holds:

Proposition 6.7. *The curve $\eta : \Omega \to \mathrm{C}\big(\mathbb{R}_+; \mathbb{R}^N\big)$ is admissible.*

Proof. For every fixed ω, the map $t \mapsto \eta(t,\omega)$ is constructed as a concatenation of equi–absolutely continuous curves, see Proposition 6.5, so it is clear that η satisfies item (i) of Definition 4.1. Its non–anticipating character is also clear by definition. It is left to show that $(t,\omega) \mapsto \eta(t,\omega)$ is jointly measurable from $\mathbb{R}_+ \times \Omega$ to $\mathbb{R}^N$ with respect to the product σ–algebra $\mathscr{B}(\mathbb{R}_+) \otimes \mathcal{F}$. To this aim, we remark that

$$\eta(t,\omega) = x\chi_{\{0\}}(t) + \sum_{k=0}^{+\infty} \Xi\Big(\tau_k(\omega), x_k(\omega), \omega\big(\tau_k(\omega)\big)\Big)(t \wedge T)\, \chi_{(\tau_k(\omega), \tau_{k+1}(\omega)]}(t \wedge T),$$

where we agree that the characteristic function $\chi_\varnothing(\cdot)$ of the empty set is identically 0. For each $k \geqslant 0$, the map $(t,\omega) \mapsto \Xi\Big(\tau_k(\omega), x_k(\omega), \omega\big(\tau_k(\omega)\big)\Big)(t \wedge T)$ is measurable as a composition of the $\mathscr{B}(\mathbb{R}_+) \otimes \mathcal{F}$–measurable function $(t,\omega) \mapsto \Big(t \wedge T, \tau_k(\omega), x_k(\omega), \omega\big(\tau_k(\omega)\big)\Big)$ with the Borel map $(t, a, y, j) \mapsto \Xi(a, y, j)(t)$ from $[0,T] \times [0,T] \times \mathbb{R}^N \times \{1, \dots, m\}$ to $\mathbb{R}^N$. The joint measurability of $(t,\omega) \mapsto \chi_{(\tau_k(\omega), \tau_{k+1}(\omega)]}(t \wedge T)$ follows from the fact that $\chi_{(\tau_k(\omega), \tau_{k+1}(\omega)]}(t \wedge T) = \chi_{F_k}(t,\omega)$ and

$$F_k := \{(t,\omega) \in \mathbb{R}_+ \times \Omega \,:\, \tau_k(\omega) < t \wedge T \leqslant \tau_{k+1}(\omega)\} \in \mathcal{F}.$$

As a countable sum of $\mathscr{B}(\mathbb{R}_+) \otimes \mathcal{F}$–measurable functions, we conclude that η has the required measurability property. $\square$

We proceed by showing a further property enjoyed by the curve η defined above.

Lemma 6.8. *Let $\omega \in \Omega$. For $\mathcal{L}^1$–a.e. $s \in (0,T)$ the following holds:*

(i) *$t \mapsto \eta(t,\omega)$ is differentiable at s;*

(ii) $$\lim_{h \to 0^+} \frac{1}{h} \int_s^{s+h} L_{G_{\omega(t)}}\big(T - t, \eta(t,\omega), -\dot\eta(t,\omega)\big)\,\mathrm{d}t = L_{G_{\omega(s)}}\big(T - s, \eta(s,\omega), -\dot\eta(s,\omega)\big);$$

(iii) *$t \mapsto u_{\omega(s)}\big(T - t, \eta(t,\omega)\big)$ is differentiable at s and*

$$-\frac{\mathrm{d}}{\mathrm{d}t} u_{\omega(s)}\big(T - t, \eta(t,\omega)\big)\Big|_{t=s} = L_{G_{\omega(s)}}\big(T - s, \eta(s,\omega), -\dot\eta(s,\omega)\big). \tag{6.4}$$

Proof. Let us fix $\omega \in \Omega$. The càdlàg path $t \mapsto \omega(t)$ has a finite number of jump times in $(0,T)$, let us say $0 < s_1 < \dots < s_n < T$. Let us set $s_0 = 0$, $s_{n+1} = T$ and pick $k \in \{0, \dots, n\}$. By definition of η, we have $\eta(\cdot,\omega) = \xi(\cdot)$ in $[s_k, s_{k+1}]$, where

$$\xi(t) = \Xi\big(s_k, \eta(s_k,\omega), \omega(s_k)\big)(t) \qquad \text{for every } t \in [0,T].$$

In view of Proposition 6.5, we have for every $t \in [s_k, T]$

$$u_{\omega(s_k)}\big(T - t, \xi(t)\big) - u^0_{\omega(s_k)}\big(\xi(T)\big) = \int_t^T L_{G_{\omega(s_k)}}\big(T - r, \xi(r), -\dot\xi(r)\big)\,\mathrm{d}r.$$

Choose $s_k \leqslant t_0 < t_1 \leqslant s_{k+1}$. By plugging $t = t_0$ and $t = t_1$ in the above equality and by subtracting the corresponding relations, we end up with

$$u_{\omega(s_k)}\big(T - t_0, \xi(t_0)\big) - u_{\omega(s_k)}\big(T - t_1, \xi(t_1)\big) = \int_{t_0}^{t_1} L_{G_{\omega(s_k)}}\big(T - r, \xi(r), -\dot{\xi}(r)\big)\,\mathrm{d}r. \tag{6.5}$$

By summing the equalities (6.5) with $t_0 = s_k$, $t_1 = s_{k+1}$ for $k = 0, \ldots, n+1$, we get

$$\int_0^T L_{G_{\omega(t)}}\big(T - t, \eta(t,\omega), -\dot{\eta}(t,\omega)\big)\,\mathrm{d}t \leqslant (n+1)\|\mathbf{u}\|_{L^\infty([0,T]\times\mathbb{R}^N)}. \tag{6.6}$$

Since the functions L_{G_i} are bounded from below, this tells us that the map

$$t \mapsto L_{G_{\omega(t)}}\big(T - t, \eta(t,\omega), -\dot{\eta}(t,\omega)\big)$$

is integrable in $[0,T]$. Therefore assertions (i) and (ii) hold whenever s is a differentiability point of the curve $t \mapsto \eta(t,\omega)$ and a Lebesgue point for $t \mapsto L_{G_{\omega(t)}}\big(T - t, \eta(t,\omega), -\dot{\eta}(t,\omega)\big)$, namely for $\mathcal{L}^1$–a.e. $s \in (0,T)$. Plug $t_0 = s$ and $t_1 = s + h$ in (6.5) for any such point $s \in (0,T)$ and for $h > 0$ small enough. By dividing the corresponding equality by h and by passing to the limit, we finally obtain (iii). □

Proof of Theorem 6.1. We introduce $h \in (0,T)$, devoted to become infinitesimal. Since $\mathbf{u}(t,x) := \mathcal{S}(t)\mathbf{u}^0(x)$ is assumed locally Lipschitz continuous in $(0,T] \times \mathbb{R}^N$, $\mathbf{u}$ is locally Lipschitz in $[h,T] \times \mathbb{R}^N$.

We can apply Theorem 4.7 to the absolutely continuous function $t \mapsto \mathbb{E}_i\big[u_{\omega(t)}\big(T - t, \eta(t,\omega)\big)\big]$ for $t \in [h,T]$. By taking into account Lemma 6.8 and the definition of the functions L_{G_j} we get

$$\begin{aligned}
&\mathbb{E}_i\big[u_{\omega(h)}\big(T - h, \eta(h,\omega)\big)\big] - \mathbb{E}_i\big[u^0_{\omega(T)}\big(\eta(T,\omega)\big)\big] \qquad (6.7)\\
&= -\int_h^T \mathbb{E}_i\left[-(B\mathbf{u})_{\omega(s)}\big(T - s, \eta(s,\omega)\big) + \frac{\mathrm{d}}{\mathrm{d}t} u_{\omega(s)}\big(T - t, \eta(t,\omega)\big)\Big|_{t=s}\right]\mathrm{d}s\\
&= \int_h^T \mathbb{E}_i\big[L_{\omega(s)}\big(\eta(s,\omega), -\dot{\eta}(s,\omega)\big)\big]\,\mathrm{d}s = \mathbb{E}_i\left[\int_h^T L_{\omega(s)}\big(\eta(s,\omega), -\dot{\eta}(s,\omega)\big)\,\mathrm{d}s\right].
\end{aligned}$$

By sending $h \to 0^+$, we get

$$\lim_{h\to 0^+} u_{\omega(h)}\big(T - h, \eta(h,\omega)\big) = u_{\omega(0)}(T,x) \quad \text{for every } \omega \in \Omega. \tag{6.8}$$

Moreover, since $\mathbf{u}$ is bounded in $[0,T] \times \mathbb{R}^N$, we obtain via Dominated Convergence Theorem

$$\lim_{h\to 0^+} \mathbb{E}_i\big[u_{\omega(h)}\big(T - h, \eta(h,\omega)\big)\big] = \mathbb{E}_i\big[u_{\omega(0)}\big(T, x\big)\big] = u_i(T,x). \tag{6.9}$$

Further, being the Lagrangians L_j bounded from below, we get via a standard application of the Monotone Convergence Theorem

$$\begin{aligned}
&\lim_{h\to 0^+} \mathbb{E}_i\left[\int_h^T L_{\omega(s)}\big(\eta(s,\omega), -\dot{\eta}(s,\omega)\big)\,\mathrm{d}s\right]\\
&\qquad = \lim_{h\to 0^+} \int_0^T \mathbb{E}_i\left[L_{\omega(s)}\big(\eta(s,\omega), -\dot{\eta}(s,\omega)\big)\right]\,\chi_{[h,T]}(s)\,\mathrm{d}s\\
&= \int_0^T \mathbb{E}_i\left[L_{\omega(s)}\big(\eta(s,\omega), -\dot{\eta}(s,\omega)\big)\right]\,\mathrm{d}s = \mathbb{E}_i\left[\int_0^T L_{\omega(s)}\big(\eta(s,\omega), -\dot{\eta}(s,\omega)\big)\,\mathrm{d}s\right].
\end{aligned}$$

By putting together the above relation plus (6.7), (6.8), (6.9), we have

$$u_i(x,T) - \mathbb{E}_i\big[u^0_{\omega(T)}\big(\eta(T,\omega)\big)\big] = \mathbb{E}_i\left[\int_0^T L_{\omega(s)}\big(\eta(s,\omega), -\dot\eta(s,\omega)\big)\,\mathrm{d}s\right],$$

which shows the claimed minimality property of $\eta(t,\omega)$. $\square$

7. Properties of minimizing random curves

In the final section, we want to establish further properties of arbitrary minimizing curves and of solutions of the evolution equation. We start showing that any minimizing curve has a similar structure as the one constructed in the previous section, up to a set of negligible probability.

We consider a solution $\mathbf{u}$ of (HJS) in $(0,+\infty)\times\mathbb{R}^N$ taking an initial value $\mathbf{u}^0$ bounded and Lipschitz continuous in $\mathbb{R}^N$. The function $\mathbf{u}$ is consequently Lipschitz continuous in $[0,T]\times\mathbb{R}^N$, for any $T>0$, by Theorem 2.4. We fix $T>0$, $x\in\mathbb{R}^N$, $i\in\{1,\dots,m\}$, and denote by $\eta:\Omega\to \mathrm{C}\big([0,T];\mathbb{R}^N\big)$ an admissible curve realizing the minimum for $u_i(x,T) = \big(\mathcal{S}(T)\mathbf{u}^0\big)_i(x)$. These notations will stay in place throughout the section.

Lemma 7.1. *There is a full measure set $\Omega_i'\subset\Omega_i$ such that for all $\omega\in\Omega_i'$, if $a<b\in[0,T]$ such that ω is constantly equal to some $j\in\{1,\dots,m\}$ in $[a,b)$, then*

$$u_j\big(b,\eta(T-b,\omega)\big) - u_j\big(a,\eta(T-a,\omega)\big) = \int_a^b L_{G_j}\big(s,\eta(T-s,\omega),-\dot\eta(T-s,\omega)\big)\mathrm{d}s.$$

Proof. Using that we have equality in the proof of Theorem 5.6, one concludes from (4.11) and (5.14) that for a.e. ω and s,

$$\big(B\mathbf{u}\big)_{\omega(s)}\big(s,\eta(T-s,\omega)\big) + \frac{\mathrm{d}}{\mathrm{d}t}u_{\omega(s)}\big(t,\eta(T-t,\omega)\big)\Big|_{t=s} = L_{\omega(s)}\big(\eta(T-s,\omega),-\dot\eta(T-s,\omega)\big).$$

It follows by Fubini's theorem that there exists a set $\Omega_i'\subset\Omega_i$ such that for all $\omega\in\Omega_i'$, the above relation holds for almost every $s\in[0,T]$. By integrating, for $\omega\in\Omega_i'$, if $0<a<b<T$ are such that ω is constantly equal to j on $[a,b)$, then

$$\begin{aligned} u_j\big(b,\eta(T-b,\omega)\big) - u_j\big(a,\eta(T-a,\omega)\big) &= \int_a^b \Big[-(B\mathbf{u})_j\big(s,\eta(T-s,\omega)\big) \\ &\qquad + L_j\big(\eta(T-s,\omega),-\dot\eta(T-s,\omega)\big)\Big]\mathrm{d}s \\ &= \int_a^b L_{G_j}\big(s,\eta(T-s,\omega),-\dot\eta(T-s,\omega)\big)\mathrm{d}s. \end{aligned}$$

$\square$

Remark 7.2. Notice that for the particular minimizing curve constructed in the previous section the exceptional negligible set is empty.

When the Hamiltonians enjoy stronger regularity properties we will accordingly get further regularity information on the minimizing curves as well as on the solutions on such curves.

We assume in the remainder of the section $H_1,\dots,H_m$ to satisfy, besides (H1)–(H3), the following further assumptions:

(H4) $\quad p\mapsto H(x,p)\quad$ is strictly convex on $\mathbb{R}^N$ for any $x\in\mathbb{R}^N$;

(H5) $\quad H \in \mathrm{C}^1(\mathbb{R}^N \times \mathbb{R}^N)$.

In this case, the associated Lagrangian L is of class C^1 in $\mathbb{R}^{\mathbb{N}} \times \mathbb{R}^N$ and is strictly convex in q. Moreover, the map $(x,q) \mapsto \big(x, \partial_q L(x,q)\big)$ is a homeomorphism of $\mathbb{R}^N \times \mathbb{R}^N$ onto itself, with continuous inverse given by $(x,p) \mapsto \big(x, \partial_p H(x,p)\big)$, see for instance [9, Appendix A.2].

Theorem 7.3. *For any fixed $\omega \in \Omega_i'$ we have*

(i) *the curve $\eta(\cdot,\omega)$ is continuously differentiable in $(0,T)$ outside the jump times of ω;*

(ii) *if t is a jump time of ω then $\eta(\cdot,\omega)$ is right and left–differentiable at t with*

$$\lim_{s\to t^+} \dot\eta(s,\omega) = \frac{\mathrm{d}^+}{\mathrm{dt}}\eta(t,\omega) \quad \textit{and} \quad \lim_{s\to t^-} \dot\eta(s,\omega) = \frac{\mathrm{d}^-}{\mathrm{dt}}\eta(t),$$

where $\frac{\mathrm{d}^+}{\mathrm{dt}}$ and $\frac{\mathrm{d}^-}{\mathrm{dt}}$ indicates right and left derivatives, respectively;

(iii) *$\eta(\cdot,\omega)$ is right differentiable at 0 and left differentiable at T with*

$$\lim_{s\to 0^+} \dot\eta(s,\omega) = \frac{\mathrm{d}^+}{\mathrm{dt}}\eta(0,\omega) \quad \textit{and} \quad \lim_{s\to T^-} \dot\eta(s,\omega) = \frac{\mathrm{d}^-}{\mathrm{dt}}\eta(T,\omega).$$

Proof. Let $\omega \in \Omega_i'$, and $0 \leqslant \hat a < \hat b \leqslant T$ such that ω is constantly equal to $j \in \{1,\dots,m\}$ on $[\hat a,\hat b]$. By recalling that u_j is a solution of (6.1) in $(0,+\infty)\times\mathbb{R}^N$ and exploiting (6.2), we get that the curve $s \mapsto \eta(s,\omega)$ is a minimizer of

$$\xi \mapsto u_j\big(T-\hat b, \xi(\hat b)\big) - u_j\big(T-\hat a, \xi(\hat a)\big) + \int_{\hat a}^{\hat b} L_{G_j}\big(T-s, \xi(s), -\dot\xi(s)\big)\mathrm{d}s$$

on the space of absolutely continuous curves $\xi : [\hat a,\hat b] \to \mathbb{R}^N$ taking the value $\eta(\hat a,\omega)$ at $t = \hat a$. Theorem 18.1 in [11] thus establishes, among other things, that the map

$$t \mapsto \partial_q L_j\big(\eta(t,\omega), -\dot\eta(t,\omega)\big) \tag{7.1}$$

which in principle is defined for a.e. $t \in [\hat a,\hat b]$, can be extended to an absolutely continuous curve on $[\hat a,\hat b]$, denoted by $p(\cdot)$. Next, we use the regularity assumptions on the Hamiltonians to invert the relation in (7.1) and get

$$-\dot\eta(t,\omega) = \partial_p H_j\big(\eta(t,\omega), p(t)\big) \qquad \text{for a.e. } t \in [\hat a,\hat b]. \tag{7.2}$$

From the continuous character of $\eta(\cdot,\omega)$ and $p(\cdot)$, it follows that $\dot\eta(\cdot,\omega)$ can be continuously extended on $[\hat a,\hat b]$. We deduce that $\eta(\cdot,\omega)$ is Lipschitz continuous in $[\hat a,\hat b]$ and, by the above continuity properties of $\dot\eta$, is in addition continuously differentiable in $(\hat a,\hat b)$. This gives (i). We moreover have

$$\frac{\eta(t,\omega)-\eta(\hat a,\omega)}{t-\hat a} = \frac{1}{t-\hat a}\int_{\hat a}^{t} \dot\eta(s)\,\mathrm{d}s$$

for $t \in (\hat a,\hat b)$. Taking into account that $\dot\eta(\cdot,\omega)$ is continuous in $(\hat a,\hat b)$ and can be continuously extended up to the boundary, we deduce that

$$\frac{\mathrm{d}^+}{\mathrm{d}t}\eta(\hat a,\omega) = \lim_{t\to\hat a^+} \frac{\eta(t,\omega)-\eta(\hat a,\omega)}{t-\hat a} = \lim_{t\to\hat a^+} \dot\eta(t).$$

The above argument, with obvious adaptations, gives items (ii) and (iii), and concludes the proof. $\square$

Corollary 7.4. *For any fixed $\omega \in \Omega_i'$, the function $u_{\omega(t)}$ is differentiable at $\big(T-t, \eta(t,\omega)\big)$, whenever t is not a jump time of ω in $(0,T)$ and*

$$\partial_x u_{\omega(t)}(T-t, \eta(t,\omega)\big) = \partial_q L_{\omega(t)}\big(\eta(t), -\dot\eta(t,\omega)\big) \tag{7.3}$$

$$\partial_t u_{\omega(t)}(T-s, \eta(t,\omega)\big) = -H_{\omega(t)}\big(\eta(t,\omega), D_x u_{\omega(t)}\big(T-t, \eta\big(t,\omega)\big)\big) - (B\,\mathbf{u})_{\omega(t)}\big(T-t, \eta(t,\omega)\big). \tag{7.4}$$

Moreover if t is a jump time, the same holds by replacing $\omega(t)$ with $\omega(t^-)$ and $\dot\eta$ with $\frac{\mathrm{d}^-}{\mathrm{dt}}\eta$.

Proof. The assertion directly comes from the previous result and Corollary A.4 with $t = T - \hat a$, $a = T - \hat b$, $\gamma(s) = \eta(T-s,\omega)$ for $s \in [a,t]$, and by making the change of variables from s to $\tau = T - s$ in the integral appearing in the representation formula of $u_{\omega(t)}$, see (6.2). □

In the sequel, we will denote by $\mathrm{D}\left(0,T;\mathbb{R}^N\right)$ the Polish space of càdlàg paths taking values in $\mathbb{R}^N$, endowed with the Prohorov metric, see [4].

Keeping in mind Theorem 7.3, we extend $\dot\eta(\cdot,\omega)$ on the whole $[0,T]$ setting

$$\dot\eta(t,\omega) = \begin{cases} \frac{\mathrm{d}^+}{\mathrm{d}t}\eta(t,\omega) & \text{if } t \text{ is a jump time of } \omega \text{ or } t = 0 \\ \frac{\mathrm{d}^-}{\mathrm{d}t}\eta(t,\omega) & \text{if } t = T. \end{cases}$$

We further define for $t \in [0,T]$ the adjoint curve

$$P(t,\omega) = \partial_q L_{\omega(t)}\big(\eta(t,\omega), -\dot\eta(t,\omega)\big).$$

Note that thanks to (7.3), $P(t,\omega) \in \partial_x^C u_{\omega(t)}\big(T-t, \eta(\omega,t)\big)$ for all t and $\omega \in \Omega_i'$, where ∂^C stands for the Clarke generalized grandient.

We deduce from the proof of Theorem 7.3

Corollary 7.5. *For any fixed ω, the curve $P(\cdot,\omega)$ is absolutely continuous on intervals of $[0,T]$ where ω is constant.*

Proposition 7.6. *The maps $\omega \to \dot\eta(\cdot,\omega)$, $\omega \to P(\cdot,\omega)$ are nonanticipating random variables from Ω to $\mathrm{D}\left(0,T;\mathbb{R}^N\right)$. In addition, the jump times of $\dot\eta(\cdot,\omega)$, $P(\cdot,\omega)$ and ω coincide, for any $\omega \in \Omega_i'$, with the possible exception of T.*

Proof. For any ω, the curves $\dot\eta(\cdot,\omega)$, $P(\cdot,\omega)$ are càdlàg by construction, with discontinuity points corresponding to the jump times of ω, with the possible exception of T where $\dot\eta(\cdot,\omega)$ and $P(\cdot,\omega)$ are continuous.

Thanks to Proposition 4.4, $\dot\eta$ is in addition $\mathscr{B}([0,t]) \otimes \mathcal{F}_t$–progressively measurable, for $t \in [0,T]$, which is, due to its càdlàg character, is equivalent of being nonanticipating. The measurability properties of η, $\dot\eta$ and the fact that $\partial_q L$ is continuous in both arguments implies that P is a random variable. It also inherits the nonanticipating character of η, $\dot\eta$. □

We know, thanks to Corollary 7.5, that $P(\cdot,\omega)$ is a.e. differentiable in $[0,T]$, for any fixed ω. We derive from [11, Theorem 18.1] that it satisfies a suitable differential inclusion in $[0,T]$. Combining this information with (7.1), (7.2) and the very definition of P, we moreover get that the pair $\big(\eta(\cdot,\omega), P(\cdot,\omega)\big)$ is a trajectory of a twisted generalized Hamiltonian dynamics, where the equation related to P is multivalued and contains a coupling term.

Corollary 7.7. *Given $\omega \in \Omega$, we have*

$$-\dot\eta(t,\omega) = \partial_p H_{\omega(t)}\big(\eta(t,\omega), P(t,\omega)\big) \quad \text{for any } t \in (0,T), \text{ not jump time of } \omega$$

and

$$\begin{aligned} \dot P(t,\omega) &\in -\partial_x H_{\omega(t)}\big(\eta(\omega,t), P(\omega,t)\big) - \sum_{j=1}^m b_{\omega(t)j}\, \partial_x^C u_j\big(T-t, \eta(\omega,t)\big) \\ &= \partial_x L_{\omega(t)}\big(\eta(\omega,t), -\dot\eta(\omega,t)\big) - \sum_{j=1}^m b_{\omega(t)j}\, \partial_x^C u_j\big(T-t, \eta(\omega,t)\big) \end{aligned}$$

for a.e. $t \in [0,T]$, where $\partial_x^C u_j$ indicates the Clarke generalized gradient of u_j with respect to the state variable. The multivalued linear combination in the formula must be understood in the sense of (1.2).

By combining Corollaries 7.5, 7.7 and the continuity properties of η, $\dot\eta$, we further get

Corollary 7.8. *Given ω, the curve $P(\cdot,\omega)$ is Lipschitz continuous on intervals of $[0,T]$ where ω is constant.*

We conclude the section by showing that when the Hamiltonians are of Tonelli type, the Lax–Oleinik semigroup has a regularizing effect, similar to the one well known for scalar Hamilton–Jacobi equations.

Given an open convex set $U \subset \mathbb{R}^N$ and $C > 0$, we recall that a function $f : U \to \mathbb{R}$ is said semiconcave with semiconcavity constant C if

$$f(\lambda\, x + (1-\lambda)\, y \geq \lambda\, f(x) + (1-\lambda)\, f(y) - C\, |x-y|^2$$

Theorem 7.9. *Assume, in addition to conditions (H1)–(H5), that the Hamiltonians $H_1, \ldots, H_m$ are of class C^2 in $\mathbb{R}^N \times \mathbb{R}^N$ with positive definite Hessian. Then, for any fixed $t > 0$, the function $\mathbf{u}(t,\cdot)$ is locally semiconcave with linear modulus on $\mathbb{R}^N$.*

Proof. We first remark that, under the conditions assumed on the Hamiltonians, the associated Lagrangians are locally semiconcave in (x,q) with a linear modulus, see [18]. Let $t > 0$ and $i \in \{1,\ldots,m\}$ be fixed. We want to prove that the function $u_i(t,\cdot)$ is locally semiconcave with linear modulus. Indeed, consider η realizing the minimum in the Lax–Oleinik formula (LO) for $u_i(t,x)$, for some $x \in \mathbb{R}^N$.

Given $z \in \mathbb{R}^N$, we define

$$\eta_\pm(s,\omega) = \eta(s,\omega) \pm \frac{t-s}{t} z.$$

Those two curves are admissible and start at $x \pm z$. We can estimate

$$
\begin{aligned}
&u_i(t, x+z) + u_i(t, x-z) - 2u_i(t,x) \\
&\quad \leqslant \mathbb{E}_i\Big[u^0_{\omega(t)}\big(\eta_+(t,\omega)\big) + \int_0^t L_{\omega(s)}\big(\eta_+(s,\omega), -\dot\eta_+(s,\omega)\big)\,\mathrm{d}s \\
&\qquad\quad + u^0_{\omega(t)}\big(\eta_-(t,\omega)\big) + \int_0^t L_{\omega(s)}\big(\eta_-(s,\omega), -\dot\eta_-(s,\omega)\big)\,\mathrm{d}s \\
&\qquad\quad - 2u^0_{\omega(t)}\big(\eta(t,\omega)\big) - 2\int_0^t L_{\omega(s)}\big(\eta(s,\omega), -\dot\eta(s,\omega)\big)\,\mathrm{d}s\Big] \\
&\quad = \mathbb{E}_i\Big[\int_0^t \Big[L_{\omega(s)}\big(\eta(s,\omega) - \frac{t-s}{t}z, -\dot\eta(s,\omega) - \frac{1}{t}z\big) \\
&\qquad\qquad + L_{\omega(s)}\big(\eta(s,\omega) + \frac{t-s}{t}z, -\dot\eta(s,\omega) + \frac{1}{t}z\big) \\
&\qquad\qquad - 2L_{\omega(s)}\big(\eta(s,\omega), -\dot\eta(s,\omega)\big)\Big]\,\mathrm{d}s\Big] \\
&\qquad \leqslant \mathbb{E}_i\Big[\int_0^t 2\,C\frac{(t-s)^2+1}{t^2}|z|^2\,\mathrm{d}s\Big] = 2\,C\Big(\frac{t}{3} + \frac{1}{t}\Big)|z|^2,
\end{aligned}
$$

where C is a constant of semiconcavity of the L_j, $j \in \{1,\dots,m\}$ restricted to a neighborhood of the $(\eta,\dot\eta)$ and $(\eta_\pm,\dot\eta_\pm)$ that are relatively compact by (7.3) (uniformly with respect to ω).

□

Appendix A. PDE material

A.1. **For systems.** This section is devoted to the proofs of Theorem 2.4 and Proposition 2.5. We prove a preliminary comparison result first:

Proposition A.1. *Let $T > 0$ and $\mathbf{v}$, $\mathbf{w} : [0,T]\times\mathbb{R}^N \to \mathbb{R}^m$ be a bounded upper semicontinuous subsolution and a bounded lower semicontinuous supersolution of* (HJS)*, respectively. Let us furthermore assume that either $\mathbf{v}$ or $\mathbf{w}$ are in $\big(\mathrm{Lip}([0,T]\times\mathbb{R}^N)\big)^m$. Then*

$$
v_i(t,x) - w_i(t,x) \leqslant \max_{1\leqslant i\leqslant m}\,\sup_{\mathbb{R}^N}\big(v_i(0,\cdot) - w_i(0,\cdot)\big)
$$

for all $(t,x) \in [0,+\infty)\times\mathbb{R}^N$ and $i \in \{1,\dots,m\}$.

Proof. The result essentially follows from Proposition 3.1 in [7], which covers our case under the additional assumption that there exists a continuity modulus $\widetilde\nu$ such that

$$
\max_{i\in\{1,\dots,m\}} |H_i(x,p) - H_i(y,p)| \leqslant \widetilde\nu\big((1+|p|)|x-y|\big) \quad \text{for all } (x,p) \in \mathbb{R}^N\times\mathbb{R}^N. \tag{A.1}
$$

When either $\mathbf{v}$ or $\mathbf{w}$ is Lipschitz continuous in $[0,T]\times\mathbb{R}^N$, such hypothesis can be safely removed. Indeed, with the notation used in [7], we see that either $\overline{p} + 2\beta\overline{x}$ or $\overline{p} - 2\beta\overline{y}$ is bounded, uniformly with respect to the parameters α,β,η,μ, since it belongs to the super differential of $v_{\overline{j}}(\overline{t},\cdot)$ at $\overline{x}$ or to the subdifferential of $w_{\overline{j}}(\overline{s},\cdot)$ at $\overline{y}$. Using the estimates (3.3) and (3.4) in [7], we conclude that both $\overline{p} + 2\beta\overline{x}$ and $\overline{p} - 2\beta\overline{y}$ are bounded, uniformly with respect to the parameters, and the result follows without invoking condition (A.1). □

We now proceed by proving the existence part in the statement of Theorem 2.4.

Proof of Theorem 2.4 (Existence of solutions). Let us first assume $\mathbf{u}^0$ bounded and Lipschitz on $\mathbb{R}^N$ and fix $T>0$. Let us denote by $b_\infty := \max_i \sum_{j=1}^m |b_{ij}|$ and pick a constant C such that

$$C > \max_{i\in\{1,\dots,m\}} \|H_i(x, Du_i^0(x))\|_\infty + b_\infty \|\mathbf{u}^0\|_\infty. \tag{A.2}$$

Set $M := C + b_\infty(\|\mathbf{u}^0\|_\infty + CT)$ and choose $n\in\mathbb{N}$ large enough so that the Hamiltonians

$$\widetilde{H}_i(x,p) := \min\{H_i(x,p), |p|+n\} \qquad (x,p)\in\mathbb{R}^N\times\mathbb{R}^N \text{ and } i\in\{1,\dots,m\}$$

satisfy

$$\widetilde{H}_i = H_i \quad \text{on } \big\{(x,p) : \max_i \widetilde{H}_i(x,p) < M+1\big\} \qquad \text{for all } i\in\{1,\dots,m\}. \tag{A.3}$$

The modified Hamiltonians $\widetilde{H}_i$ satisfy the additional condition (A.1), thus we can apply Proposition 3.1 in [7] and infer the existence of a function $\mathbf{u}\in\big(\mathrm{BUC}([0,T]\times\mathbb{R}^N)\big)^m$ which solves (HJS) in $(0,T)\times\mathbb{R}^N$ with $\widetilde{H}_i$ in place of H_i and satisfying the initial condition $\mathbf{u}(0,\cdot)=\mathbf{u}^0$ on $\mathbb{R}^N$.

We shall now prove that $\mathbf{u}$ is Lipschitz in $[0,T]\times\mathbb{R}^N$. To this aim, first notice that the functions $\mathbf{u}^+(t,x) := \mathbf{u}^0(x) + tC\mathbb{1}$ and $\mathbf{u}^-(t,x) := \mathbf{u}^0(x) - tC\mathbb{1}$ are a Lipschitz super and subsolution to (HJS) in $(0,T)\times\mathbb{R}^N$ with $\widetilde{H}_i$ in place of H_i. Therefore, by Proposition 3.1 in [7], we get in particular

$$|u_i(t,x) - u_i^0(x)| \leqslant Ct \qquad \text{for every } (t,x,i)\in[0,T]\times\mathbb{R}^N\times\{1,\dots,m\}. \tag{A.4}$$

By applying Proposition 3.1 in [7] again to $\mathbf{u}(h+\cdot,\cdot)$ and to the solutions $\mathbf{w}^\pm := \mathbf{u}\pm\|\mathbf{u}(h,\cdot)-\mathbf{u}^0\|_\infty$, we get

$$\|\mathbf{u}(t+h,\cdot)-\mathbf{u}(t,\cdot)\|_\infty \leqslant \|\mathbf{u}(h,\cdot)-\mathbf{u}^0\|_\infty \leqslant C\,h \qquad \text{for every } h>0.$$

This shows that the function $\mathbf{u}$ is C–Lipschitz in t. By making use of the fact that $\mathbf{u}$ is a solution to (HJS) in $(0,T)\times\mathbb{R}^N$ with $\widetilde{H}_i$ in place of H_i and of the estimate (A.4), we get

$$\widetilde{H}_i(x, D_x u_i(t,x)) \leqslant -\partial_t u_i(t,x) - \big(B\mathbf{u}(t,x)\big)_i \leqslant C + b_\infty(\|\mathbf{u}^0\|_\infty + CT) = M$$

in the viscosity sense in $(0,T)\times\mathbb{R}^N$. By coercivity of $\widetilde{H}_i$ in p, we infer that $u_i(t,\cdot)$ is Lipschitz for every $t\in(0,T)$ with $(x, D_x u_i(t,x))\in\{\widetilde{H}_i(x,p)\leqslant M\}$ for a.e. $x\in\mathbb{R}^N$, see for instance Lemma 2.5 in [2]. In view of (A.3), this finally implies that $\mathbf{u}$ is a solution of (HJS) in $(0,T)\times\mathbb{R}^N$ as well.

Let now assume that $\mathbf{u}^0\in\big(\mathrm{BUC}(\mathbb{R}^N)\big)^m$. Let $\mathbf{g}^n$ be a sequence of Lipschitz functions in $\big(\mathrm{BUC}(\mathbb{R}^N)\big)^m$ uniformly converging to $\mathbf{u}^0$ on $\mathbb{R}^N$, and denote by $\mathbf{u}^n$ the corresponding bounded and Lipschitz solution to (HJS) in $(0,T)\times\mathbb{R}^N$ with initial datum $\mathbf{g}^n$. By Proposition A.1 we have $\|\mathbf{u}^m-\mathbf{u}^n\|_{L^\infty([0,T]\times\mathbb{R}^N)} \leqslant \|\mathbf{g}^m-\mathbf{g}^n\|_{L^\infty(\mathbb{R}^N)}$, that is, $(\mathbf{u}^n)_n$ is a Cauchy sequence in $[0,T]\times\mathbb{R}^N$ with respect to the sup–norm. Hence the Lipschitz continuous functions $\mathbf{u}^n$ uniformly converge to a function $\mathbf{u}$ on $[0,T]\times\mathbb{R}^N$, which is therefore bounded and uniformly continuous on $[0,T]\times\mathbb{R}^N$. By the stability of the notion of viscosity solution, we conclude that $\mathbf{u}$ is a solution of (HJS) with initial datum $\mathbf{u}^0$. This completes the proof since $T>0$ was arbitrarily chosen. □

The uniqueness part in Theorem 2.4 is guaranteed by the comparison principle stated in Proposition 2.5, that we prove now. The proof makes use of the existence result just established and of Proposition A.1.

Proof of Proposition 2.5. Up to trivial cases and up to adding a constant vector of the form $C\mathbb{1}$ to $\mathbf{w}$, we reduce the assertion to proving that $\mathbf{v} \leqslant \mathbf{w}$ in $[0,T]\times\mathbb{R}^N$ for every fixed $T>0$ when $\mathbf{v}(0,\cdot) \leqslant \mathbf{w}(0,\cdot)$ in $\mathbb{R}^N$. Let us fix $T>0$ and, for every $\varepsilon>0$, set $\mathbf{w}^\varepsilon := \mathbf{w}+\varepsilon\mathbb{1}$. Since either $\mathbf{v}(0,\cdot)$ or $\mathbf{w}^\varepsilon(0,\cdot)$ are in $\big(\mathrm{BUC}(\mathbb{R}^N)\big)^m$ and $\mathbf{w}^\varepsilon(0,\cdot)-\mathbf{v}(0,\cdot) \geqslant \varepsilon\mathbb{1}$, we can find $\mathbf{u}^0 \in \big(\mathrm{BUC}(\mathbb{R}^N)\big)^m \cap \big(\mathrm{Lip}(\mathbb{R}^N)\big)^m$ such that $\mathbf{v}(0,\cdot) \leqslant \mathbf{u}^0 \leqslant \mathbf{w}^\varepsilon(0,\cdot)$ in $\mathbb{R}^N$. According to the existence part of Theorem 2.4 proved above, we know that there exists a Lipschitz function $\mathbf{u} \in \big(\mathrm{BUC}([0,T]\times\mathbb{R}^N)\big)^m$ which solves (HJS) in $(0,T)\times\mathbb{R}^N$ with initial datum $\mathbf{u}_0$. By applying Proposition A.1 to the pair $\mathbf{v}$, $\mathbf{u}$ and $\mathbf{u}$, $\mathbf{w}^\varepsilon$, respectively, we get $\mathbf{v} \leqslant \mathbf{u} \leqslant \mathbf{w}^\varepsilon$ in $[0,T]\times\mathbb{R}^N$. The assertion follows by sending $\varepsilon \to 0^+$. □

A.2. **For a single equation.** We now turn back to results concerning a single equation. Let $G : [0,T]\times\mathbb{R}^N\times\mathbb{R}^N \to \mathbb{R}$ be a continuous Hamiltonian such that, for every fixed $(t,x) \in [0,T]\times\mathbb{R}^N$, $G(t,x,\cdot)$ is convex in $\mathbb{R}^N$, and there exist two superlinear functions $\alpha,\beta : \mathbb{R}_+ \to \mathbb{R}$ such that

$$\alpha(|p|) \leqslant G(t,x,p) \leqslant \beta(|p|) \qquad \text{for all } (t,x,p) \in [0,T]\times\mathbb{R}^N\times\mathbb{R}^N.$$

We will denote by $L_G : [0,T]\times\mathbb{R}^N\times\mathbb{R}^N \to \mathbb{R}$ the Lagrangian associated with G through the Fenchel transform. The following holds:

Proposition A.2. *Let $u : [0,T]\times\mathbb{R}^N \to \mathbb{R}$ be a Lipschitz solution of*

$$\frac{\partial u}{\partial t} + G(t,x,D_x u) = 0 \qquad \textit{in } (0,T)\times\mathbb{R}^N. \tag{A.5}$$

Then for every $0 \leqslant a < t \leqslant T$, the following identity holds:

$$u(t,x) = \inf_{\xi(t)=x} \Big(u\big(a,\xi(a)\big) + \int_a^t L_G\big(s,\xi(s),\dot\xi(s)\big)\mathrm{d}s \Big), \qquad x \in \mathbb{R}^N, \tag{A.6}$$

where the infimum is taken by letting ξ vary in the family of absolutely continuous curves from $[a,t]$ to $\mathbb{R}^N$. Moreover, such an infimum is a minimum.

Proof. Let us first prove the assertion for $a=0$. It is always true that

$$u(t,x) \leqslant V(t,x) := \inf_{\xi(t)=x} \Big(u\big(0,\xi(0)\big) + \int_0^t L_G\big(s,\xi(s),\dot\xi(s)\big)\mathrm{d}s \Big), \tag{A.7}$$

for every $x \in \mathbb{R}^N$. Indeed, let $\xi : [0,t] \to \mathbb{R}^N$ be any absolutely continuous curve with $\xi(t)=x$. Then, for almost every $s \in [0,T]$, we have

$$\frac{\mathrm{d}}{\mathrm{d}\tau} u\big(\tau,\xi(\tau)\big)\Big|_{\tau=s} = p_s + \langle p_{\xi(s)}, \dot\xi(s)\rangle \leqslant p_s + G(s,\xi(s),p_{\xi(s)}) + L_G\big(s,\xi(s),\dot\xi(s)\big),$$

where $\big(p_s,p_{\xi(s)}\big)$ is a suitable element of the Clarke generalized gradient of u at $\big(s,\xi(s)\big)$, chosen according to Lemma 1.4. By integrating the above inequality between 0 and t and by taking into account that u is a subsolution of (6.1) and ξ was arbitrarily chosen, we readily get (A.7).

We therefore have to prove the converse inequality. Let us fix $R>0$ and define $L_R = L_G$ on $[0,T]\times\mathbb{R}^N\times B_R$ and $+\infty$ elsewhere. Let us set

$$V_R(t,x) := \inf_{\xi(t)=x} \Big(u\big(0,\xi(0)\big) + \int_0^t L_R\big(s,\xi(s),\dot\xi(s)\big)\mathrm{d}s \Big), \qquad (t,x) \in (0,T]\times\mathbb{R}^N,$$

where the infimum is taken by letting ξ vary in the family of absolutely continuous curves from $[0,t]$ to $\mathbb{R}^N$. Clearly, it is the same to take the infimum over R–Lipschitz curves. Moreover, $V_R \geqslant V$. As the curves take velocities in a compact set, by basic results of

Optimal Control Theory (see [1, Theorem 3.17 and Exercise 3.7]), the function V_R is a Lipschitz continuous solution of

$$\frac{\partial V_R}{\partial t} + G_R(t, x, D_x V_R) = 0 \qquad \text{in } (0,T) \times \mathbb{R}^N \tag{A.8}$$

where G_R is the convex dual of L_R. It is easily checked from its very definition that G_R is R–Lipschitz in p, and uniformly continuous in $[0,T] \times K \times \mathbb{R}^N$, for every compact set $K \subset \mathbb{R}^N$. Let us denote by κ a Lipschitz constant of u in $[0,T] \times \mathbb{R}^N$ and choose R big enough such that $G = G_R$ on $[0,T] \times \mathbb{R}^N \times B_{\kappa+1}$. From the definition of (viscosity) solutions and the fact that sub and supertangents to u have norms bounded by κ, it follows that u also solves (A.8) in $[0,T] \times \mathbb{R}^N$. Now u and V_R are both Lipschitz solutions to (A.8) with same initial data, hence, by applying [1, Theorem 3.12], we conclude that $u \equiv V_R$ in $[0,T] \times \mathbb{R}^N$. Since $V_R \geqslant V$, we finally get $u \equiv V$.

The fact that (A.6) holds for $0 \leqslant a < t \leqslant T$ is due to the fact that the function V defined by (A.7) satisfies the Dynamic Programming Principle. The fact that the infimum in (A.6) is attained follows from classical results of the Calculus of Variations. $\square$

We proceed by proving some differentiability properties of the solution u at points belonging to the support of a minimizing curve for (A.6).

Proposition A.3. *Let G be as above, and assume moreover that $G(t,x,p)$ is strictly convex in p and Lipschitz in x, locally with respect to (t,x,p). Let u be a Lipschitz solution of the evolutive Hamilton–Jacobi equation* (6.1) *in $[0,T] \times \mathbb{R}^N$. Let $0 < a < t < T$ and $\gamma : [a,t] \to \mathbb{R}^N$ a curve that realizes the infimum in* (A.6). *Assume a is a differentiability point of γ and a Lebesgue point of $s \mapsto L_G\big(s, \gamma(s), \dot\gamma(s)\big)$. Then u is differentiable in $\big(a, \gamma(a)\big)$ and*

$$\partial_x u\big(a, \gamma(a)\big) = \partial_q L_G\big(a, \gamma(a), \dot\gamma(a)\big), \qquad \partial_t u\big(a, \gamma(a)\big) = -G\Big(a, \gamma(a), D_x u\big(a, \gamma(a)\big)\Big).$$

Proof. Under the above hypotheses, the function u is locally semiconcave on $(0, +\infty) \times \mathbb{R}^N$ (see [9, Theorem 5.3.8]) and γ is Lipschitz. The proof is borrowed from [9, Theorem 6.4.7]. Let (p_t, p_x) be a superdifferential to u at $\big(a, \gamma(a)\big)$. By definition of (viscosity) solutions, $p_t + G(a, \gamma(a), p_x) \leqslant 0$. We will prove that equality holds. As for $t - a > h > 0$, we have $u\big(a+h, \gamma(a+h)\big) - u\big(a, \gamma(a)\big) = \int_a^{a+h} L_G\big(s, \gamma(s), \dot\gamma(s)\big)\mathrm{d}s$, since a is a Lebesgue point of $s \mapsto L_G\big(s, \gamma(s), \dot\gamma(s)\big)$, it follows that

$$\lim_{h \to 0^+} \frac{u\big(a+h, \gamma(a+h)\big) - u\big(a, \gamma(a)\big)}{h} = L_G\big(a, \gamma(a), \dot\gamma(a)\big).$$

On the other hand, by properties of superdifferentials,

$$\lim_{h \to 0^+} \frac{u\big(a+h, \gamma(a+h)\big) - u\big(a, \gamma(a)\big)}{h} \leqslant p_t + \langle p_x, \dot\gamma(a) \rangle.$$

It follows from the Fenchel inequality that $p_t + G(a, \gamma(a), p_x) \geqslant 0$ hence the claimed equality holds.

Finally, as the superdifferential is convex and G is strictly convex in the last argument, $D^+ u\big(a, \gamma(a)\big)$ cannot contain more than one element. It is moreover not empty by properties of semiconcave functions, hence $D^+ u\big(a, \gamma(a)\big)$ is reduced to a singleton and u is differentiable at $\big(a, \gamma(a)\big)$. Moreover, as p_x realize the equality in the Fenchel inequality it follows that $p_x = \partial_q L_G\big(a, \gamma(a), \dot\gamma(a)\big)$ and $p_t = -G\Big(a, \gamma(a), D_x u\big(a, \gamma(a)\big)\Big)$.

$\square$

We finally state a consequence of the previous results and of Theorem 18.1 in [11]:

Corollary A.4. *Let us assume that the hypotheses of Proposition A.3 are in force, and furthermore that $G(t,x,\cdot)$ is of class C^1, for every fixed $(t,x) \in [0,T] \times \mathbb{R}^N$. Then the curve γ is C^1 and u is differentiable at $\big(s,\gamma(s)\big)$ for every $s \in [a,t)$, with*

$$\partial_x u\big(s,\gamma(s)\big) = \partial_q L_G\big(s,\gamma(s),\dot\gamma(s)\big), \qquad \partial_t u\big(s,\gamma(s)\big) = -G\big(s,\gamma(s),D_x u\big(s,\gamma(s)\big)\big).$$

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] G. Barles, *Solutions de viscosité des équations de Hamilton-Jacobi*, vol. 17 of Mathématiques & Applications (Berlin) [Mathematics & Applications], Springer-Verlag, Paris, 1994.

[3] P. Bernard, *Existence of $C^{1,1}$ critical sub-solutions of the Hamilton-Jacobi equation on compact manifolds*, Ann. Sci. École Norm. Sup. (4), 40 (2007), pp. 445–452.

[4] P. Billingsley, *Convergence of probability measures*, Wiley Series in Probability and Statistics: Probability and Statistics, John Wiley & Sons, Inc., New York, second ed., 1999. A Wiley-Interscience Publication.

[5] G. Buttazzo, M. Giaquinta, and S. Hildebrandt, *One-dimensional variational problems*, vol. 15 of Oxford Lecture Series in Mathematics and its Applications, The Clarendon Press, Oxford University Press, New York, 1998. An introduction.

[6] F. Cagnetti, D. Gomes, H. Mitake, and H. V. Tran, *A new method for large time behavior of degenerate viscous Hamilton-Jacobi equations with convex Hamiltonians*, Ann. Inst. H. Poincaré Anal. Non Linéaire, 32 (2015), pp. 183–200.

[7] F. Camilli, O. Ley, and P. Loreti, *Homogenization of monotone systems of Hamilton-Jacobi equations*, ESAIM Control Optim. Calc. Var., 16 (2010), pp. 58–76.

[8] F. Camilli, O. Ley, P. Loreti, and V. D. Nguyen, *Large time behavior of weakly coupled systems of first-order Hamilton–Jacobi equations*, NoDEA Nonlinear Differential Equations Appl., 19 (2012), pp. 719–749.

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

[10] C. Castaing and M. Valadier, *Convex analysis and measurable multifunctions*, Lecture Notes in Mathematics, Vol. 580, Springer-Verlag, Berlin-New York, 1977.

[11] F. Clarke, *Functional analysis, calculus of variations and optimal control*, vol. 264 of Graduate Texts in Mathematics, Springer, London, 2013.

[12] A. Davini, A. Fathi, R. Iturriaga, and M. Zavidovique, *Convergence of the solutions of the discounted Hamilton–Jacobi equation*, Inventiones mathematicae, (2016), pp. 1–27.

[13] A. Davini and A. Siconolfi, *A generalized dynamical approach to the large time behavior of solutions of Hamilton-Jacobi equations*, SIAM J. Math. Anal., 38 (2006), pp. 478–502 (electronic).

[14] A. Davini and M. Zavidovique, *Aubry sets for weakly coupled systems of Hamilton-Jacobi equations*, SIAM J. Math. Anal., 46 (2014), pp. 3361–3389.

[15] H. Engler and S. M. Lenhart, *Viscosity solutions for weakly coupled systems of Hamilton-Jacobi equations*, Proc. London Math. Soc. (3), 63 (1991), pp. 212–240.

[16] L. C. Evans and R. F. Gariepy, *Measure theory and fine properties of functions*, Studies in Advanced Mathematics, CRC Press, Boca Raton, FL, 1992.

[17] A. Fathi, *Sur la convergence du semi-groupe de Lax-Oleinik*, C. R. Acad. Sci. Paris Sér. I Math., 327 (1998), pp. 267–270.

[18] ———, *Weak KAM Theorem in Lagrangian Dynamics, preliminary version 10, Lyon.* unpublished, June 15 2008.

[19] R. Griego and R. Hersh, *Theory of random evolutions with applications to partial differential equations*, Trans. Amer. Math. Soc., 156 (1971), pp. 405–418.

[20] H. Ibrahim, A. Siconolfi, and S. Zabad, *Cycle Chacterization of the Aubry Set for Weakly Coupled Hamilton-Jacobi Systems*, ArXiv e-prints, (2016).

[21] H. Ishii, *Perron's method for monotone systems of second-order elliptic partial differential equations*, Differential Integral Equations, 5 (1992), pp. 1–24.

[22] ———, *Asymptotic solutions for large time of Hamilton-Jacobi equations in Euclidean n space*, Ann. Inst. H. Poincaré Anal. Non Linéaire, 25 (2008), pp. 231–266.
[23] H. Ishii and S. Koike, *Viscosity solutions for monotone systems of second-order elliptic PDEs*, Comm. Partial Differential Equations, 16 (1991), pp. 1095–1128.
[24] K. Kuttler, *Notes for Seminar*, 2013.
[25] S. M. Lenhart, *Viscosity solutions for weakly coupled systems of first-order partial differential equations*, J. Math. Anal. Appl., 131 (1988), pp. 180–193.
[26] P.-L. Lions, G. Papanicolaou, and S. Varadhan, *Homogenization of Hamilton-Jacobi equation.* unpublished preprint, 1987.
[27] H. Mitake, A. Siconolfi, H. V. Tran, and N. Yamada, *A Lagrangian approach to weakly coupled Hamilton–Jacobi systems*, SIAM Journal on Mathematical Analysis, 48 (2016), pp. 821–846.
[28] G. Namah and J.-M. Roquejoffre, *Convergence to periodic fronts in a class of semilinear parabolic equations*, NoDEA Nonlinear Differential Equations Appl., 4 (1997), pp. 521–536.
[29] V. Nguyen, *Some results on the large–time behavior of weakly coupled systems of first-order Hamilton–Jacobi equations*, Journal of Evolution Equations, (2013), pp. 1–33.
[30] M. A. Pinsky, *Lectures on random evolution*, World Scientific Publishing Co., Inc., River Edge, NJ, 1991.
[31] F. Rezakhanlou and J. E. Tarver, *Homogenization for stochastic Hamilton-Jacobi equations*, Arch. Ration. Mech. Anal., 151 (2000), pp. 277–309.
[32] F. Riesz and B. Sz.-Nagy, *Functional analysis*, Dover Books on Advanced Mathematics, Dover Publications, Inc., New York, 1990. Translated from the second French edition by Leo F. Boron, Reprint of the 1955 original.
[33] P. E. Souganidis, *Stochastic homogenization of Hamilton-Jacobi equations and some applications*, Asymptot. Anal., 20 (1999), pp. 1–11.

Dip. di Matematica, Sapienza Università di Roma, P.le Aldo Moro 2, 00185 Roma, Italy
E-mail address: davini@mat.uniroma1.it

Dip. di Matematica, Sapienza Università di Roma, P.le Aldo Moro 2, 00185 Roma, Italy
E-mail address: siconolf@mat.uniroma1.it

IMJ (projet Analyse Algébrique), UPMC, 4, place Jussieu, Case 247, 75252 Paris Cédex 5, France
E-mail address: zavidovique@math.jussieu.fr