SIAM J. Numer. Anal.
Vol. 58, No. 6, pp. 3165–3196

A SEMI-LAGRANGIAN SCHEME FOR HAMILTON–JACOBI–BELLMAN EQUATIONS ON NETWORKS*

E. CARLINI†, A. FESTA‡, and N. FORCADEL§

Abstract. We present a semi-Lagrangian scheme for the approximation of a class of Hamilton–Jacobi–Bellman (HJB) equations on networks. The scheme is explicit, consistent, and stable for large time steps. We prove a convergence result and two error estimates. For an HJB equation with space-independent Hamiltonian, we obtain a first order error estimate. In the general case, we provide, under a hyperbolic CFL condition, a convergence estimate of order one half. The theoretical results are discussed and validated in a numerical tests section.

Key words. Hamilton–Jacobi–Bellman equations, semi-Lagrangian scheme, networks

AMS subject classifications. 65M15, 65M25, 49L25, 90B20

DOI. 10.1137/19M1260931

1. Introduction.

Interest in the study of linear and nonlinear partial differential equations on networks has risen steadily in recent decades, motivated by the modeling of various networked systems like roads, pipelines, and electronic and information networks. In particular, an extensive literature has been developed for vehicular traffic systems modeled through conservation laws. Existence results can be found in [19], and some partial uniqueness results (for a limited number of intersecting roads) in [18, 2]. In many cases, the lack of uniqueness on the junction points obliges one to add some special conditions, which may be ambiguous or difficult to derive. More recently, models based on Hamilton–Jacobi (HJ) equations have been proposed. In these models, the density of the cars is obtained as the derivative of the solution of the HJ equation (see [25]). The main advantage of this framework is the ability to include an optimality principle in the model, solving some of the ambiguities in the junction points without the introduction of additional conditions. However, the relationship between the two approaches is still under investigation.

The theory of HJ equations on networks is very recent. In general, these equations do not have regular solutions, and the notion of weak solution (*viscosity solution*) needs to be extended on the junction points. So far several proposals have been made. The early attempts are contained in the works [1, 7, 21, 22, 27], where the authors introduce new definitions of weak solutions and prove the well-posedness of the problem. We highlight the paper [10], where the authors discuss the differences among the models. We also refer the reader to the most recent works [5, 24] for simplified proofs of uniqueness.

*Received by the editors May 10, 2019; accepted for publication (in revised form) July 23, 2020; published electronically November 9, 2020.
https://doi.org/10.1137/19M1260931

Funding: The work of the first author was supported by INdAM GNCS project "Approssimazione numerica di problemi di natura iperbolica ed applicazioni." The work of the second author was partially supported by MIUR grant "Dipartimenti Eccellenza 2018- 2022" CUP: E11G18000350001, by DISMA, and by Politecnico di Torino. The work of the third author was partially supported by the European Union with the European Regional Development Fund (ERDF) through grants HN0002137, 18P03390, 18E01750, and 18P02733 and by the Normandie Regional Council via projects M2NUM, M2SINUM.

†Department of Mathematics, Sapienza University of Rome, 00185 Roma, Italy (carlini@mat.uniroma1.it).
‡Politecnico di Torino, 10129 Turin, Italy (adriano.festa@polito.it).
§INSA Rouen, 76800 Saint-Étienne-du-Rouvray, France (nicolas.forcadel@insa-rouen.fr).

Regarding the numerical approximation, there are very few schemes, and only some of them are supported by theoretical results. Let us mention the finite differences scheme proposed in [7, 11] and the paper [20] in which some error estimates are proved.

In this paper, we adopt the framework and the notion of weak solution as introduced in [21]. This framework has the advantage of including very general models. Here, the Hamiltonian is convex with respect to the gradient variable. At the junction, it can be discontinuous with respect to the space variable and may depend on a *flux limiter*.

We propose a semi-Lagrangian scheme for this kind of equation by discretizing the *dynamic programming principle* presented in [21]. The scheme generalizes that introduced in [8] and enables discrete characteristics to cross the junctions. This property makes the scheme unconditionally stable, allowing for large time steps. This is the main advantage compared to finite differences and finite element schemes. With almost standard techniques, it is possible to prove consistency and monotonicity, which imply the convergence of the scheme.

We prove two convergence error estimates: for state-independent Hamiltonians, where optimal controls are constant in time, and for more general Hamiltonians. In the first case, we obtain a first order convergence estimate depending only on the space step. In the second case, we prove a general convergence result and, in the case of Courant number less than one, an error estimate which, for constant Courant number, gives order of convergence 1/2. The proof is obtained applying some techniques derived from papers on regional optimal control problems [4, 5], and it improves the results presented in [20] in the case of finite differences schemes.

For the sake of clarity, we consider a simplified network (a *junction*), but the result can be extended to more general networks, with more than one junction, as we show in the last numerical test.

Structure of the paper. In section 2 we recall some basic notions for junctions and build the optimal control problem on these domains. In section 3, we derive the scheme and prove its basic properties: consistency, monotonicity, and regularity. In section 4, we present the main results concerning convergence and error estimates. Finally, in section 5, we show some numerical simulations.

2. An optimal control problem on networks.

A network is a domain composed of a finite number of nodes connected by a finite number of edges. To simplify the description of such a system we focus on the case of a *junction*, which is a network composed of one node and a finite number of edges. We follow [21] and the notation therein to describe the problem.

Given a positive number N, a junction $J \in \mathbb{R}^2$ is a network of N half-lines $J_i := \{k\, e_i, k \in \mathbb{R}^+\}$ (where each line is isometric to $[0, +\infty)$ and e_i is a unitary vector centered at the origin) connected at a *junction point* that we conventionally place at the origin (see Figure 1). We then have

$$J := \bigcup_{i=1,\dots,N} J_i, \quad J_i \cap J_j = \{0\} \quad \forall i \neq j, \quad i, j \in \{1, \dots, N\}.$$

We consider the geodesic distance function on J given by

$$d(x, y) = \begin{cases} |x - y| & \text{if } x, y \in J_i \text{ for one } i \in \{1, \dots, N\}, \\ |x| + |y| & \text{otherwise.} \end{cases}$$

For a real-valued function u defined on J, $\partial_i u(x)$ denotes the (spatial) derivative of u

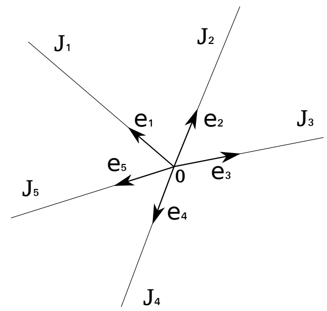

FIG. 1. *Junction with $N = 5$ edges.*

at $x \in J_i$ and the gradient of u is defined as

$$u_x := \begin{cases} \partial_i u(x) & \text{if } x \in J_i^* := J_i \setminus \{0\}, \\ (\partial_1 u(x), \partial_2 u(x), \ldots, \partial_N u(x)) & \text{if } x = 0. \end{cases} \tag{1}$$

We describe a finite-horizon optimal control problem on the network J. For a more extensive description of the problem, see [21].

Let us define the set of admissible dynamics on the network J connecting point (s, y) to point (t, x) as

$$\Gamma_{s,y}^{t,x} := \begin{cases} (X(\cdot), \alpha(\cdot)) \in \mathrm{Lip}([s,t]; J) \times L^\infty([s,t]; \mathbb{R}^{N+1}), \\ \dot{X}(\tau) = U(X(\tau), \alpha(\tau)), \quad \tau \in [s,t], \\ X(s) = y, \quad X(t) = x, \end{cases} \tag{2}$$

where for any $(t, x) \in [0, T] \times J$ and $\alpha = (\alpha_0, \alpha_1, \ldots, \alpha_N) \in \mathbb{R}^{N+1}$,

$$U(x, \alpha) = \begin{cases} \alpha_i & \text{if } x \in J_i^*, \\ \alpha_0 & \text{if } x = 0. \end{cases}$$

We define the *cost function*,

$$L(x, \alpha) := \begin{cases} L_i(x, \alpha_i) & \text{if } x \in J_i, \\ L_0(\alpha_0) & \text{if } x = 0, \end{cases}$$

where the functions $L_i : \mathbb{R}^+ \times \mathbb{R} \to \mathbb{R}$ for $i = 1, \ldots, N$ satisfy the following:

(A1) L_i are *strictly convex* (with respect to the second argument) and uniformly Lipschitz continuous.

(A2) L_i are *strongly coercive* with respect to the second argument uniformly in x ($L_i(x, \alpha_i)/|\alpha_i| \to +\infty$ for $|\alpha_i| \to +\infty$ uniformly in $x \in \mathbb{R}^+$).

(A3) For all $\mu > 0$ there exists $C_\mu > 0$ such that

$$\sup_{x \in J_i} \left| \inf_{\alpha_i \in [-\mu, \mu]} L_i(x, \alpha_i) \right| \le C_\mu.$$

In addition, $L_0 : \mathbb{R} \to \mathbb{R}$ is defined as

$$L_0(\alpha_0) := \begin{cases} \bar{L}_0 & \text{if } \alpha_0 = 0, \\ +\infty & \text{otherwise} \end{cases}$$

for a given $\bar{L}_0 \in \mathbb{R}$. The *value function* of the optimal control problem is

$$u(t, x) = \inf_{y \in J} \inf_{(X(\cdot), \alpha(\cdot)) \in \Gamma_{0,y}^{t,x}} \left\{ u_0(y) + \int_0^t L(X(\tau), \alpha(\tau)) d\tau \right\}. \tag{3}$$

Remark. We stress the generality of the model in treating the junction: an optimal trajectory chosen in (2) is evaluated by the functional (3) where the cost does not have any regularity passing through the junction. The trajectories have also the possibility of waiting in the junction, paying a specific constant cost L_0 per time unity.

It has been proved in [21] that the following *dynamic programming principle* holds.

PROPOSITION 2.1 (dynamic programming principle). *For all $x \in J$, $t \in (0,T]$, $s \in [0,t)$, the value function u defined in* (3) *satisfies*

$$u(t,x) = \inf_{y\in J} \inf_{(X(\cdot),\alpha(\cdot))\in\Gamma^{t,x}_{s,y}} \left\{ u(s,X(s)) + \int_s^t L(X(\tau),\alpha(\tau))d\tau \right\}. \tag{4}$$

A direct approximation of the dynamic programming principle (4) is the basis for the scheme which we describe in the next section.

The following theorem characterizes the value function (3) as the solution of a Hamilton–Jacobi–Bellman (HJB) equation (for the definition of viscosity solution and the proof, see Appendix A and [21]).

THEOREM 2.2 (HJB equation satisfied by the value function u). *Given a function u_0, globally Lipschitz continuous on J, the value function u defined in* (3) *is the unique viscosity solution of*

$$\begin{cases} \partial_t u(t,x) + H_i(x,u_x(t,x)) = 0 & \text{in } (0,T)\times J_i^*, \\ \partial_t u(t,x) + F_A(u_x(t,x)) = 0 & \text{in } (0,T)\times\{0\}, \end{cases} \tag{5}$$

with initial condition $u(0,x) = u_0(x)$ for $x \in J$, where

$$H_i(x,p) := \sup_{\alpha_i\in\mathbb{R}} \{\alpha_i\, p - L_i(x,\alpha_i)\},$$

with $\hat{p}_i(x)$ chosen such that H_i is nonincreasing in $(-\infty,\hat{p}_i(x)]$ and nondecreasing in $[\hat{p}_i(x),\infty)$. The operator $F_A : \mathbb{R}^N \to \mathbb{R}$ on the junction point is

$$F_A(p) := \max\left(A, \max_{i=1,\dots,N} H_i^-(0,p_i)\right), \ \text{with } A = -\bar{L}_0, \tag{6}$$

where

$$H_i^-(x,p) := \begin{cases} H_i(x,p) & \text{for } p \le \hat{p}_i(x), \\ H_i(x,\hat{p}_i) & \text{for } p > \hat{p}_i(x). \end{cases}$$

Remark. In the vehicular traffic flow models, the function H^- can be related, with suitable transformations, to the demand and supply functions, introduced in [23]. This relation was observed in [22]. In this setting, the constant A is known as the flux limiter at the junction. In fact, the lower the cost at the junction $\bar{L}_0$, the longer the vehicles will stay at the junction, and so the bigger the flux limiter.

We now gives some useful results on the Hamiltonian H.

PROPOSITION 2.3 (properties on H). *Under assumptions* (A1)–(A3)*, the following assertions hold true:*

(i) *For every $x \in \mathbb{R}^+ \cup \{+\infty\}$ and bounded p, $\alpha_i \in \arg\sup_{\alpha_i\in\mathbb{R}}\{\alpha_i p - L_i(x,\alpha_i)\}$ is bounded.*

(ii) *The nonincreasing part of $H_i(x,p)$ with respect to p_i is given by*

$$H_i^-(x,p_i) = \sup_{\alpha_i \le 0}\{\alpha_i p_i - L_i(x,\alpha_i)\}.$$

(iii) *(Regularity.) For all $M>0$ there exists a modulus of continuity ω_M such that for all $|p|,|q|\le M$ and $x\in J_i$*

$$|H_i(x,p) - H_i(x,q)| \le \omega_M(|p-q|);$$

in addition, $H_i(\cdot,p)$ is Lipschitz continuous with respect to the space variable.

(iv) *(Uniform coercivity.) $H_i(x,p)\to+\infty$ for $|p|\to+\infty$ uniformly for every $x\in J_i\cup\{+\infty\}$, $i=1,\dots,N$.*

(v) *(Convexity.) $p\mapsto H_i(x,p)$ is convex for every $x\in J$.*

(vi) *(Uniform bound of the Hamiltonian for bounded gradient.) For all $M>0$, there exists $C_M>0$ such that*

$$\sup_{p\in[-M,M],\ x\in J^*} |H(x,p)| \le C_M,$$

where $J^* := J\setminus\{0\}$.

Remark. The previous proposition implies in particular the well-posedness of (5) (see [21]).

Proof. From assumptions (A1)–(A2), $\alpha_i p - L_i(x,\alpha_i)$ is a continuous function (negatively) coercive; therefore there exists a compact interval $[-\mu,\mu]$, $\mu\in\mathbb{R}$, such that

$$\sup_{\alpha_i\in\mathbb{R}}\{\alpha_i\, p - L_i(x,\alpha_i)\} = \sup_{\alpha_i\in[-\mu,\mu]}\{\alpha_i\, p - L_i(x,\alpha_i)\}.$$

Since p is bounded, (i) holds. Assertion (ii) follows from Lemma 6.2 in [21].
From (i), we have

$$H_i(x,p) - H_i(x,q) \le \bar\alpha|p-q|,$$

where $\bar\alpha$ is the minimizer in $H_i(x,q)$. Exchanging the role of p,q, we get (iii).

Taking $\alpha=1$ in the Hamiltonian, we have

$$H_i(x,p) \ge p - L_i(x,1).$$

The same argument for $\alpha=-1$ gives $H_i(x,p)\to+\infty$ for $|p|\to+\infty$; then (iv) holds. Finally (v) holds since H_i is the upper envelope of convex functions, and (vi) follows directly from assumption (A3). ∎

We now give regularity results for the value functions.

Proposition 2.4 (regularity of the value function). *Under assumptions* (A1)–(A2)*, the value function u defined in* (3) *is Lipschitz continuous in space and time.*

Proof. First, we remark that for $C\ge C_0$, with C_0 defined as in (vi) of Proposition 2.3 with $L=0$, $u_0(x)\pm Ct$ are, respectively, the subsolution and supersolution of (5). Using the comparison principle, we deduce that

$$u_0(x) - Ct \le u(t,x) \le u_0(x) + Ct.$$

Let $h\ge 0$ and define $u^h(t,x) = u(t+h,x) - Ch$. The previous inequalities imply that

$$u^h(0,x) = u(h,x) - Ch \le u(0,x).$$

The equation is invariant by translation in time and by the addition of a constant (which implies that u^h is a subsolution of (5)); we then get by the comparison principle that

$$u^h(t,x) \le u(t,x).$$

This implies that

$$\frac{u(t+h,x)-u(t,x)}{h} \le C.$$

The reverse inequality can be proved in the same way (using that $u(t+h,x)+Ch$ is a supersolution), so we deduce that

$$|u_t| \le C,$$

since h can be chosen arbitrarily small. Hence u is Lipschitz continuous in time. We know that u is a viscosity solution of (5), and therefore it satisfies in particular (in the viscosity sense) for each $i \in \{1,\dots,N\}$

$$H_i(x,u_x) \le C \quad \text{on } (0,T)\times J_i^*.$$

Using the coercivity of H, this implies the existence of a constant $\tilde{C}$ such that (in the viscosity sense)

$$|u_x| \le \tilde{C} \quad \text{on } (0,T)\times J_i^*.$$

Therefore, u is Lipschitz continuous also with respect to the space variable. ∎

Remark. Note that in the classical setting (i.e., without junction), this type of result can be obtained more directly using the definition of the value function and the dynamic programming principle (obtaining first regularity in space and then in time). But the arguments used in this setting rely on the Lipschitz continuity of the coset L, which is no longer true at the junction. This is the reason why, in this proof, we have to use viscosity techniques.

3. A semi-Lagrangian scheme for the approximation of the solution. Let us introduce a uniform discretization of the network $(0,T)\times J$. The choice of a uniform discretization is not restrictive, and the scheme can be easily extended to nonuniform grids. Given Δt and Δx in $\mathbb{R}^+$, we define $\Delta = (\Delta x, \Delta t)$, $N_T = \lfloor T/\Delta t \rfloor$ ($\lfloor\cdot\rfloor$ is the integer part), and

$$\mathcal{G}^\Delta := \{t_n : n = 0,\dots,N_T\} \times J^{\Delta x},$$

where $J^{\Delta x} := \bigcup_{i=1,\dots,N} J_i^{\Delta x}$, $J_i^{\Delta x} = \{k\Delta x\, e_i : k \in \mathbb{N}\}$. We define $t_n = n\Delta t$ for $n = 0,\dots,N_T$ and derive a discrete version of the dynamic programming principle (5) defined on the grid $\mathcal{G}^\Delta$. To do so, as usual in first order semi-Lagrangian schemes, we discretize the trajectories in $\Gamma_{t_n,y}^{t_{n+1},x}$ by one step of the Euler scheme. For $i \in \{1,\dots,N\}$, let $x \in J_i$ and let $\alpha \in \mathbb{R}^{N+1}$ be such that $\alpha_i \Delta t \le |x|$; then the approximated trajectory gets

$$x \simeq y + \alpha_i \Delta t.$$

In this case, the discrete backward trajectory $x - \Delta t \alpha_i$ remains on J_i and, applying a quadrature formula, a discrete version of (4) at the point (t_{n+1},x) is

$$u(t_{n+1},x) \simeq u(t_n, x-\alpha_i \Delta t e_i) + \Delta t L_i(x,\alpha_i).$$

Instead, if $\alpha_i \Delta t > |x|$, the discrete trajectory reaches the junction at a time included in the interval $[0,\Delta t]$. Denoting by $s_0 \in [0, \Delta t - \frac{|x|}{\alpha_i}]$ the time spent by the trajectory at the junction point, J_j the arc from which the trajectory comes, and $\hat{t} := \left(\Delta t - s_0 - \frac{|x|}{\alpha_i}\right)$ the remaining time on a new arc J_j, the approximation of (4) at the point (t_{n+1}, x) becomes

$$u(t_{n+1}, x) \simeq u\left(t_n, -\alpha_j \hat{t} e_j\right) + \hat{t}\, L_j(0, \alpha_j) + s_0 L_0(\alpha_0) + \frac{|x|}{\alpha_i} L_i(x, \alpha_i).$$

We denote by $B(J^{\Delta x})$ and $B(\mathcal{G}^{\Delta})$ the spaces of bounded functions defined, respectively, on $J^{\Delta x}$ and on $\mathcal{G}^{\Delta}$. We approximate the value function on the feet of the discrete trajectories, which in general are not grid nodes, by a standard piecewise linear Lagrange interpolation $\mathbb{I}[\hat{u}](z)$, where $\hat{u} \in B(J^{\Delta x})$ and $z \in J_j$, i.e.,

$$\mathbb{I}[\hat{u}](z) := \hat{u}(i\Delta x e_j) + (z - i\Delta x e_j)\frac{\hat{u}((i+1)\Delta x e_j) - \hat{u}(i\Delta x e_j)}{\Delta x e_j}$$

for $z \in [i\Delta x e_j, (i+1)\Delta x e_j)$.

Finally, we define a fully discrete numerical operator $S : B(\mathcal{G}^{\Delta}) \times J^{\Delta x} \to \mathbb{R}$ as, if $x \in J_i$,

$$S[\hat{v}](x) := \min \begin{cases} \displaystyle\min_{\alpha_i < \frac{|x|}{\Delta t}} \mathbb{I}[\hat{v}](x - \alpha_i \Delta t e_i) + \Delta t L_i(x, \alpha_i), \\ \displaystyle\min_{\alpha_i \geq \frac{|x|}{\Delta t}} \min_{s_0 \in [0, \Delta t - \frac{|x|}{\alpha_i}]} \min_{j, \alpha_j \leq 0} \Big\{ \mathbb{I}[\hat{v}]\left(-\left(\Delta t - s_0 - \tfrac{|x|}{\alpha_i}\right)\alpha_j e_j\right) \\ \qquad + \left(\Delta t - s_0 - \tfrac{|x|}{\alpha_i}\right) L_j(0, \alpha_j) + s_0 L_0(\alpha_0) + \tfrac{|x|}{\alpha_i} L_i(x, \alpha_i) \Big\} \end{cases}$$

and, if $x = 0$,

$$S[\hat{v}](x) := \min_{j, \alpha_j \leq 0} \min_{s_0 \in [0, \Delta t]} \left\{ \mathbb{I}[\hat{v}]\left(-(\Delta t - s_0)\alpha_j e_j\right) + (\Delta t - s_0) L_j(0, \alpha_j) + s_0 L_0(\alpha_0) \right\}.$$

Then the *discrete solution* $w \in B(\mathcal{G}^{\Delta})$ solves

$$w(t_{n+1}, x) = S[\hat{w}^n](x), \quad n = 0, \ldots, N_T - 1,\ x \in J^{\Delta x}, \tag{7}$$

where $\hat{w}^n := \{w(t_n, x)\}_{x \in J^{\Delta x}}$ for $n = 0, \ldots, N_T - 1$ and $\hat{w}^0 = \{u_0(x)\}_{x \in J^{\Delta x}}$.

3.1. Basic properties of the scheme. We prove some basic properties of (7).

PROPOSITION 3.1 (monotonicity and stability of the scheme). *We assume that* (A1)–(A3) *hold. Then the numerical scheme* (7) *is*

(i) monotone, *i.e., given two discrete functions $v_1, v_2 \in B(J^{\Delta x})$ such that $v_1 \leq v_2$ we have*
$$S[\hat{v}_1](x) \leq S[\hat{v}_2](x) \quad \forall x \in J^{\Delta x};$$

(ii) invariant by addition of constants, *i.e., $S[\hat{\varphi} + C](z) = S[\hat{\varphi}](z) + C$ for any constant C;*

(iii) stable, *i.e., there exists a positive constant K such that for any $(t_n, x) \in \mathcal{G}^{\Delta}$*
$$|w(t_n, x) - u_0(x)| \leq K t_n.$$

Proof. To prove monotonicity, let us fix an $x \in J_i^{\Delta x}$. We focus on the difficulty due to the junction. More precisely, we assume that the trajectory related to v_1 passes through the junction and the one related to v_2 does not. The other cases are easier and can be treated in a similar way. Let us denote by $(\bar{\alpha}_i, \bar{s}_0, \bar{j}, \bar{\alpha}_{\bar{j}}, \bar{\alpha}_0)$ the optimal strategy contained in v_1, and let us denote by $\hat{\alpha}_i$ the optimal control of v_2. The optimal controls are bounded by Proposition 2.3. We have

$$S[\hat{v}_1](x) = \mathbb{I}[\hat{v}_1]\left(-\left(\Delta t - \bar{s}_0 - \frac{|x|}{\bar{\alpha}_i}\right)\bar{\alpha}_{\bar{j}} e_{\bar{j}}\right) + \left(\Delta t - \bar{s}_0 - \frac{|x|}{\bar{\alpha}_i}\right) L_j(0, \bar{\alpha}_{\bar{j}})$$
$$+ \bar{s}_0 L_0(\bar{\alpha}_0) + \frac{|x|}{\alpha_i} L_i(x, \hat{\alpha}_i) \leq \mathbb{I}[\hat{v}_1](x - \hat{\alpha}_i \Delta t e_i) + \Delta t L_i(x, \hat{\alpha}_i) = S[\hat{v}_2](x),$$

which proves the monotonicity.

Point (ii) is a straightforward verification. The stability property (iii) follows directly from (i) and (ii) with $K \geq \sup_{x\in J^{\Delta x}} \frac{|S[\hat{u}^0](x) - u_0(x)|}{\Delta t}$. For the proof, see [12]. $\square$

We now give a regularity result for the solution of the scheme. This is the discrete analogue of the Lipschitz estimate in the space of the value function and will be used in the proof of the error estimate.

Proposition 3.2 (almost Lipschitz Regularity in the space of w). *Let $w(t_n, x)$ be a solution of (7). If u_0 is uniformly Lipschitz continuous, then for $x, y \in J^{\Delta x}$ there exists a $C > 0$ such that*

$$|w(t_n, x) - w(t_n, y)| \leq C\,(\Delta t + d(x, y)), \quad n = 0, \ldots, N_t.$$

The proof is postponed to Appendix B.

Remark (bounded control). By Proposition 3.1(iii) the w solution of (7) is bounded and then the discrete problem (7) is well-posed. We observe also that the same argument of Proposition 2.3 (based on (A2)) can be used to prove that the control α in (7) is bounded. We define

$$\mu = \sup_{(x,t)\in J\times(0,T]} \max_{i=1,\ldots,N} |\alpha_i^*|, \tag{8}$$

the maximal absolute value of the optimal control.

3.2. Consistency of the scheme. We now focus on the study of the consistency properties of the scheme. First of all, we recall the definition of consistency (the class of test functions $C^2(J)$ is defined in Appendix A).

Definition 3.3 (consistency). *Let $x \in J$ and $(\Delta x_m, \Delta t_m) \to 0$ as $m \to \infty$. Let $y_m \in J^{\Delta x_m}$ be a sequence of grid points such that $y_m \to x$ as $m \to \infty$. The scheme S is said to be* consistent *with* (5) *if the following properties hold:*

(i) *If $x \in J_i$, for all test functions $\varphi \in C^2(J)$, we have*

$$\frac{\varphi(y_m) - S[\hat{\varphi}](y_m)}{\Delta t_m} \to H_i(x, \varphi_x(x)) \quad \text{as } m \to \infty. \tag{9}$$

(ii) *If $x = 0$, for all test functions $\varphi \in C^2(J)$ such that $\partial_i\varphi(0) = p_i^{L_0}$ for $i = 1, \ldots, N$, where $p_i^{L_0} \in \mathbb{R}$ are such that $H_i(0, p_i^{L_0}) = H_i^+(0, p_i^{L_0}) = -L_0$ and $H_i^+(x, p) := \sup_{\alpha_i \geq 0}(\alpha_i p - L_i(x, \alpha))$, we have*

$$\frac{\varphi(y_m) - S[\hat{\varphi}](y_m)}{\Delta t_m} \to F_{-L_0}(\varphi_x(x)) = -L_0 \quad \text{as } m \to \infty. \tag{10}$$

Definition 3.4 (consistency estimate). *Let $x \in J^{\Delta x}$ and $\Delta x, \Delta t > 0$. We say that the scheme S satisfies a* consistency estimate $\mathcal{E}(\Delta x, \Delta t) > 0$ *if for all test functions $\varphi \in C^2(J)$ with bounded second order derivatives, the following hold:*

(i) *If* $x \in J_i^{\Delta x} \setminus \{0\}$, *we have*

$$\left| \frac{\varphi(x) - S[\hat{\varphi}](x)}{\Delta t} - H_i(x, \varphi_x(x)) \right| \le \|\varphi_{xx}\|_\infty \, \mathcal{E}(\Delta x, \Delta t). \tag{11}$$

(ii) *If* $x = 0$, *we have*

$$\left| \frac{\varphi(x) - S[\hat{\varphi}](x)}{\Delta t} - F_{-L_0}(\varphi_x(x)) \right| \le \|\varphi_{xx}\|_\infty \, \mathcal{E}(\Delta x, \Delta t). \tag{12}$$

Remark. Let us remark that, due to the particular form of the test function in (10), if the scheme admits a consistency estimate $\mathcal{E}(\Delta x, \Delta t) \to 0$, then the scheme is consistent in the sense of Definition 3.3. Indeed, if $y_m \to 0$ as $m \to +\infty$, with $y_m \in J_i^*$ and $\varphi \in C^2(J)$ with $\partial_i \varphi(0) = p_i^{L_0}$, then the consistency estimate implies

$$\frac{\varphi(y_m) - S[\hat{\varphi}](y_m)}{\Delta t_m} \to H_i(0, \varphi_x(0)) = H_i(0, p_i^{L_0}) = -L_0.$$

We begin to prove some *consistency estimates* for the numerical operators.

PROPOSITION 3.5. *Given* $\Delta t > 0$ *and* $\Delta x > 0$, *let us assume the CFL condition*

$$\mu \frac{\Delta t}{\Delta x} \le 1 \quad (\textit{with } \mu \textit{ as in } (8)). \tag{13}$$

Then for any $\varphi \in C^2(J)$ *the following estimates hold for* (7):

(i) *If* $x \in J_i^{\Delta x} \setminus \{0\}$, *then*

$$\left| \frac{\varphi(x) - S[\hat{\varphi}](x)}{\Delta t} - H_i(x, \varphi_x(x)) \right| \le K \|\varphi_{xx}\|_\infty \, \left(\Delta t + \min\left(\Delta x, \frac{\Delta x^2}{\Delta t} \right) \right);$$

(ii) *if* $x = 0$,

$$\left| \frac{\varphi(x) - S[\hat{\varphi}](x)}{\Delta t} - F_{-L_0}(\varphi_x(x)) \right| \le K \|\varphi_{xx}\|_\infty \, \left(\Delta t + \min\left(\Delta x, \frac{\Delta x^2}{\Delta t} \right) \right),$$

where K *is a positive constant.*

Remark (small Courant number). In the case when very small Courant numbers are considered, $\mu \frac{\Delta t}{\Delta x} \le \Delta x$, the estimates in Proposition 3.5 ensure consistency error of order 1. These estimates improve the classical estimate $\frac{\Delta x^2}{\Delta t} + \Delta t$ for first order semi-Lagrangian scheme and were first proved in [14].

Proof. (i) Let $x \in J_i^{\Delta x} \setminus \{0\}$. We remark that condition (13) implies in particular that the scheme reads

$$S[\hat{\varphi}](x) = \min_{\alpha_i < \frac{|x|}{\Delta t}} \mathbb{I}[\hat{\varphi}](x - \Delta t \alpha_i e_i) + \Delta t L_i(x, \alpha_i) = \min_{\alpha_i \in \mathbb{R}} \mathbb{I}[\hat{\varphi}](x - \Delta t \alpha_i e_i) + \Delta t L_i(x, \alpha_i).$$

By using recent estimates proved in [14, 15], we have

$$\mathbb{I}[\hat{\varphi}](x - \Delta t \alpha_i e_i) = \varphi(x - \Delta t \alpha_i e_i) + K \|\varphi_{xx}\|_\infty \, \min(\Delta x^2, \Delta t \Delta x). \tag{14}$$

Then by standard Taylor expansion we get the result.

(ii) Let $x = 0$. In this case

$$S[\hat{\varphi}](0) = \min_{s_0 \in [0,\Delta t]} \min_{j,\alpha_j \le 0} \left\{ \mathbb{I}[\hat{\varphi}](-(\Delta t - s_0)\alpha_j e_j) + (\Delta t - s_0) L_j(0,\alpha_j) + s_0 L_0(\alpha_0) \right\}.$$

Let us define $K_{\Delta t} := \frac{s_0}{\Delta t}$; since $s_0 \in [0, \Delta t]$ we have $K_{\Delta t} \in [0,1]$. Again by Taylor expansion, by Proposition 2.3, and by the interpolation error (14), we have

$$\begin{aligned}
&\max \frac{\varphi(0) - S[\hat{\varphi}](0)}{\Delta t} + K\|\varphi_{xx}\|_\infty \left(\Delta t + \min\left(\Delta x, \frac{\Delta x^2}{\Delta t}\right)\right) \\
&= -\min_{K_{\Delta t}\in[0,1]} \min_{j,\alpha_j \le 0} \left(-(1-K_{\Delta t})\alpha_j \partial_j \varphi(0) + (1-K_{\Delta t}) L_j(0,\alpha_j) + K_{\Delta t} L_0(\alpha_0)\right) \\
&= -\min_{K_{\Delta t}\in[0,1]} \left[(1-K_{\Delta t}) \min_{j,\alpha_j\le 0} \left(-\alpha_j \partial_j\varphi(0) + L_j(0,\alpha_j)\right) + K_{\Delta t} \min_{\alpha_0} \left(L_0(\alpha_0)\right)\right] \\
&= \max_{K_{\Delta t}\in[0,1]} \left[(1-K_{\Delta t}) \max_{j,\alpha_j\le 0} \left(\alpha_j \partial_j\varphi(0) - L_j(0,\alpha_j)\right) + K_{\Delta t} \max_{\alpha_0} \left(-L_0(\alpha_0)\right)\right] \\
&= \max_{K_{\Delta t}\in[0,1]} \left\{ (1-K_{\Delta t}) \max_j H_j^-(0, \partial_j \varphi(0)) - K_{\Delta t} L_0 \right\} \\
&= \max\left(\max_j H_j^-(0,\partial_j\varphi(0)), -L_0 \right).
\end{aligned}$$

This ends the proof of the proposition. ∎

The case that we study behaves differently from classic semi-Lagrangian schemes, where the consistency error estimate is not limited by a CFL condition. This difference is due to the presence of discontinuities on the Hamiltonians at the junction point.

It is worthwhile to underline that consistency (in the sense of Definition 3.3) holds *even without* (13), and consequently the scheme is convergent without any CFL condition, as we show at the beginning of section 4.

PROPOSITION 3.6 (consistency of the scheme). *Assume* $\min(\frac{\Delta x^2}{\Delta t}, \Delta x) \to 0$. *Then the scheme* (7) *is consistent according to Definition* 3.3.

Proof. Let us consider a sequence y_m such that $y_m \to x$ as $\Delta_m = (\Delta x_m, \Delta t_m) \to (0,0)$. For notational convenience we drop the index m of the sequence of grid points. In case the limit point x is not on the junction since x is fixed for every sequence $(\Delta x, \Delta t) \to (0,0)$, y eventually verifies $|y| > \mu \Delta t$ *independently* from the rate $\Delta t/\Delta x$. Then the consistency follows as Case 1 in the proof of Proposition 3.5 (without the condition $\Delta t/\Delta x \le 1/\mu$).

The situation is more complex when the limit point x is 0. If $y \equiv 0$, this case is equivalent to Case 2 in the proof of Proposition 3.5. If y is such that $y \to 0$ and $y \neq 0$, up to a subsequence, we can assume that $y \in J_i$ for some i independent of m. In that case, the optimal trajectory can cross the junction in one time step. Let $\varphi \in C^2(J)$ such that $\partial_i \varphi(0) = p_i^A$ for $i = 1, \dots, N$, and let us define the two quantities:

$$\begin{aligned}
\mathcal{I}_1 &:= \min_{\alpha_i < \frac{|y|}{\Delta t}} \left(\mathbb{I}[\hat{\varphi}](y - \Delta t \alpha_i e_i) + \Delta t L_i(y,\alpha_i)\right), \\
\mathcal{I}_2 &:= \min_{\alpha,\alpha_i \ge \frac{|y|}{\Delta t}} \min_{s_0 \in [0, \Delta t - \frac{|y|}{\alpha_i}]} \min_{j,\alpha_j\le 0} \left\{ \mathbb{I}[\hat{\varphi}] \left(-\left(\Delta t - s_0 - \frac{|y|}{\alpha_i}\right)\alpha_j e_j \right) \right. \\
&\qquad \left. + \left(\Delta t - s_0 - \frac{|y|}{\alpha_i}\right) L_j(0,\alpha_j) + s_0 L_0(\alpha_0) + \frac{|y|}{\alpha_i} L_i(y,\alpha_i) \right\}.
\end{aligned}$$

We remark that $S[\varphi](y) = \min(\mathcal{I}_1, \mathcal{I}_2)$. We begin with the term $\mathcal{I}_1$. Evaluating the interpolation error and using a Taylor expansion, we get

(15)
$$\begin{aligned}\mathcal{I}_1 &= \min_{\alpha_i \le \frac{|y|}{\Delta t}} \{\varphi(y) - \alpha_i \Delta t \partial_i \varphi(y) + \Delta t L_i(y,\alpha_i)\} + K\|\varphi_{xx}\|_\infty \left(\min(\Delta x^2, \Delta t \Delta x) + \Delta t^2\right)\\ &= \varphi(y) - \Delta t \max_{\alpha_i \le \frac{|y|}{\Delta t}} \{\alpha_i \partial_i \varphi(y) - L_i(y,\alpha_i)\} + K\|\varphi_{xx}\|_\infty \left(\min(\Delta x^2, \Delta t \Delta x) + \Delta t^2\right).\end{aligned}$$

Using the inequality

$$\begin{aligned}\max_{\alpha_i \le \frac{|y|}{\Delta t}} \{\alpha_i \partial_i \varphi(y) - L_i(y,\alpha_i)\} &\le \max_{\alpha_i \in \mathbb{R}} \{\alpha_i \partial_i \varphi(y) - L_i(y,\alpha_i)\}\\ &= H_i(y, \partial_i \varphi(y)) = -L_0 + o(1),\end{aligned}$$

we deduce that

$$\mathcal{I}_1 \ge \varphi(y) - \Delta t A + \Delta t\, o(1) + K\|\varphi_{xx}\|_\infty \left(\min(\Delta x^2, \Delta t \Delta x) + \Delta t^2\right). \tag{16}$$

For the term $\mathcal{I}_2$, we add into the argument of φ the term $y - \frac{|y|}{\alpha_i}\alpha_i e_i = 0$. Using the Taylor expansion twice and the interpolation accuracy, we obtain

$$\begin{aligned}\mathbb{I}[\varphi]\left(-\left(\Delta t - s_0 - \frac{|y|}{\alpha_i}\right)\alpha_j e_j\right) = \varphi(y) - \frac{|y|}{\alpha_i}\alpha_i \partial_i \varphi(y) - \left(\Delta t - s_0 - \frac{|y|}{\alpha_i}\right)\alpha_j \partial_j \varphi(0)\\ + K\|\varphi_{xx}\|_\infty \left(\min(\Delta x^2, \Delta t \Delta x) + \Delta t^2\right).\end{aligned}$$

The equation above implies

$$\begin{aligned}&\mathcal{I}_2 + K\|\varphi_{xx}\|_\infty \left(\min(\Delta x^2, \Delta t \Delta x) + \Delta t^2\right)\\ &= \min_{\alpha_i \ge \frac{|y|}{\Delta t}} \min_{s_0 \in [0, \Delta t - \frac{|y|}{\alpha_i}]} \left\{ \min_j \min_{\alpha_j \le 0} \left\{ -\left(\Delta t - s_0 - \frac{|y|}{\alpha_i}\right)(\alpha_j \partial_j \varphi(0) - L_j(0,\alpha_j)) \right\}\right.\\ &\qquad\qquad \left. + \varphi(y) - \frac{|y|}{\alpha_i}(\alpha_i \partial_i \varphi(y) - L_i(y,\alpha_i)) + s_0 L_0(\alpha_0) \right\}\\ &= \varphi(y) + \min_{\alpha_i \ge \frac{|y|}{\Delta t}} \min_{s_0 \in [0, \Delta t - \frac{|y|}{\alpha_i}]} \left\{ -\frac{|y|}{\alpha_i}(\alpha_i \partial_i \varphi(y) - L_i(y,\alpha_i)) + s_0 L_0(\alpha_0)\right.\\ &\qquad\qquad \left. - \left(\Delta t - s_0 - \frac{|y|}{\alpha_i}\right) \max_j \max_{\alpha_j \le 0} \{(\alpha_j \partial_j \varphi(0) - L_j(0,\alpha_j))\} \right\}\\ &= \varphi(y) + \min_{\alpha_i \ge \frac{|y|}{\Delta t}} \min_{s_0 \in [0, \Delta t - \frac{|y|}{\alpha_i}]} \left\{ -\frac{|y|}{\alpha_i}(\alpha_i \partial_i \varphi(y) - L_i(y,\alpha_i)) + s_0 L_0(\alpha_0)\right.\\ &\qquad\qquad \left. - \left(\Delta t - s_0 - \frac{|y|}{\alpha_i}\right) \max_j H_j^-(0, \partial_j \varphi(0)) \right\}.\end{aligned}$$

Using $\max_j H_j^-(0, \partial_j \varphi(0)) = \max_j \min_p H_j(0,p) =: \bar{H}^0$ and $L_0(\alpha_0) = L_0$, we deduce that (we use $\bar{H}^0 \le -L_0$)

$$
\begin{aligned}
&\mathcal{I}_2 + K\|\varphi_{xx}\|_\infty \left(\min(\Delta x^2, \Delta t\Delta x) + \Delta t^2\right) \\
&= \varphi(y) + \min_{\alpha_i \ge \frac{|y|}{\Delta t}} \Bigg\{ -\frac{|y|}{\alpha_i}(\alpha_i \partial_i \varphi(y) - L_i(y,\alpha_i)) \\
&\qquad\qquad + \min_{s_0 \in [0, \Delta t - \frac{|y|}{\alpha_i}]} \left\{ s_0(\bar{H}^0 + L_0) - \left(\Delta t - \frac{|y|}{\alpha_i}\right)\bar{H}^0 \right\} \Bigg\} \\
&= \varphi(y) + \min_{\alpha_i \ge \frac{|y|}{\Delta t}} \left\{ -\frac{|y|}{\alpha_i}(\alpha_i \partial_i \varphi(y) - L_i(y,\alpha_i)) + \left(\Delta t - \frac{|y|}{\alpha_i}\right) L_0 \right\} \qquad (17)\\
&= \Delta t\, L_0 + \varphi(y) - \max_{\alpha_i \ge \frac{|y|}{\Delta t}} \left\{ \frac{|y|}{\alpha_i}(\alpha_i \partial_i \varphi(y) - L_i(y,\alpha_i) + L_0) \right\}.
\end{aligned}
$$

We use $\frac{|y|}{\alpha_i} \le \Delta t$ in the last sup, and we observe that $\alpha_i \partial_i \varphi(y) - L_i(y,\alpha_i) + L_0 \le o(1)$ getting

$$
\mathcal{I}_2 + K\|\varphi_{xx}\|_\infty \left(\min(\Delta x^2, \Delta t\Delta x) + \Delta t^2\right) \ge +\Delta t\, L_0 + \varphi(y) + \Delta t o(1). \qquad (18)
$$

Finally, via (16) and (18), we obtain

$$
\begin{aligned}
S[\varphi](y) &= \min(\mathcal{I}_1, \mathcal{I}_2) \qquad (19)\\
&\ge +\Delta t L_0 + \varphi(y) + \Delta t o(1) + K\|\varphi_{xx}\|_\infty \left(\min(\Delta x^2, \Delta t\Delta x) + \Delta t^2\right).
\end{aligned}
$$

Now, we need to show that this inequality is in fact an equality. We denote by $\bar{\alpha}_i$ the solution of

$$
\max_{\alpha_i \in \mathbb{R}} \{\alpha_i \partial_i \varphi(y) - L_i(y,\alpha_i)\}
$$

and distinguish two cases. First, we consider the case $\bar{\alpha}_i \le \frac{|y|}{\Delta t}$. This implies in particular that

$$
\begin{aligned}
\max_{\alpha_i \le \frac{|y|}{\Delta t}} \{\alpha_i \partial_i \varphi(y) - L_i(y,\alpha_i)\} &= \max_{\alpha_i \in \mathbb{R}} \{\alpha_i \partial_i \varphi(y) - L_i(y,\alpha_i)\} \\
&= H_i(y, \partial_i \varphi(y)) = -L_0 + o(1).
\end{aligned}
$$

Using (15), we deduce that

$$
\mathcal{I}_1 = \Delta t L_0 + \varphi(y) + \Delta t o(1) + K\|\varphi_{xx}\|_\infty \left(\min(\Delta x^2, \Delta t\Delta x) + \Delta t^2\right),
$$

and so (19) is an equality.

We now consider the case $\bar{\alpha}_i \ge \frac{|y|}{\Delta t}$. We define

$$
\bar{\mathcal{I}}_2 := \max_{\alpha_i \ge \frac{|y|}{\Delta t}} \left\{ \frac{|y|}{\alpha_i}(\alpha_i \partial_i \varphi(y) - L_i(y,\alpha_i) + L_0) \right\}.
$$

Clearly, $0 \le \frac{|y|}{\alpha_i} \le \Delta t$ and $\alpha_i \partial_i \varphi(y) - L_i(y,\alpha_i) + L_0 \le o(1)$; therefore we can say

$$
\begin{aligned}
\Delta t\, o(1) \ge \bar{\mathcal{I}}_2 &\ge \Delta t \left\{ \max_{\alpha_i \ge \frac{|y|}{\Delta t}} \left\{ \frac{|y|}{\alpha_i}(\alpha_i \partial_i \varphi(y) - L_i(y,\alpha_i) \right\} + L_0 \right\} \\
&= \Delta t(H_i(y, \partial_i \varphi(y)) + L_0) = \Delta t\, o(1).
\end{aligned}
$$

This implies again that (19) is an equality and completes the proof. $\square$

4. Convergence and convergence estimates. In this section, we introduce the main results of the paper. First of all, the convergence of the scheme can be proven with a standard argument based on the monotonicity.

THEOREM 4.1 (convergence). *Assume that* $\min(\Delta x^2/\Delta t, \Delta x) \to 0$*, and let* $T > 0$ *and let* u_0 *be a Lipschitz continuous function on* J*. Then the numerical solution* w *of* (7) *converges uniformly on any compact set* $\mathcal{K}$ *of* $(0,T) \times J$ *as* $\Delta \to (0,0)$ *to the unique viscosity solution* u *of* (5)*, i.e.,*

$$\limsup_{\Delta x, \Delta t \to 0} \sup_{(t,x) \in \mathcal{K} \cap \mathcal{G}^{\Delta}} |w(t,x) - u(t,x)| = 0.$$

Proof. Since the scheme is consistent (Proposition 3.6) for a subsequence verifying $\min(\Delta x^2/\Delta t, \Delta x) \to 0$ and is monotone and stable, we can follow [6, 11, 21] and obtain the result. Note that the choice of the test functions in the definition of the consistency at the junction uses Theorem A.2(ii). ∎

Once the convergence of the scheme is shown, we want to provide also some convergence estimates. This is a less easy task. We need to take into account two different scenarios: For the special case of space-independent Hamiltonians, i.e., assuming the additional property

(A4) the Lagrangians $L_i(x, \alpha_i) \equiv L_i(y, \alpha_i) =: L_i(\alpha_i)$ for every choice of $x, y \in J_i$,

it is possible to prove an error bound, independent of the time step. We observe that, as a consequence of the structure of the costs, the optimal control $\bar{\alpha}_i$ is constant in time and no restriction on the time step is required.

THEOREM 4.2 (rate of convergence in the case of space-independent Hamiltonians). *Let* (A1)*,* (A2)*,* (A4) *be verified. Consider* u *a viscosity solution of* (5)*, and let* w *be a solution of the scheme* (7)*. Then there exists a positive constant* C *depending only on the Lipschitz constant of* u *such that*

$$\sup_{(t,x) \in \mathcal{G}^{\Delta}} |u(t,x) - w(t,x)| \le CT\Delta x. \tag{20}$$

The proof is contained in Appendix C.

For more general Hamiltonians (i.e., without assuming (A4)), we prove an error bound that applies to any stable, monotone scheme for which a consistency estimate is valid.

THEOREM 4.3 (rate of convergence). *Assume* (A1)–(A3)*. Let* u *be the viscosity solution of* (5)*, let* w *be the solution of a scheme for which Proposition* 3.1 *holds, and assume that the scheme satisfies a consistency estimate* $\mathcal{E}(\Delta t, \Delta x)$ *as in Definition* 3.4*. Then there exists a positive constant* C *independent of* Δt *and* Δx *such that*

$$\sup_{(t,x) \in \mathcal{G}^{\Delta}} |u(t,x) - w(t,x)| \le CT\left(\frac{\mathcal{E}(\Delta t, \Delta x)}{\sqrt{\Delta t}} + \sqrt{\Delta t}\right) + \sup_{x \in J^{\Delta x}} |u_0(x) - w(0,x)|. \tag{21}$$

By applying the previous theorem, we get an error estimate for scheme (7) under a restriction on the time step, given by assumption (13).

COROLLARY 4.4 (rate of convergence for (7)). *In the specific case of the scheme* (7)*, assuming* (13)*, we have*

$$\sup_{(t,x) \in \mathcal{G}^{\Delta}} |u(t,x) - w(t,x)| \le CT\left(\sqrt{\Delta t} + \frac{1}{\sqrt{\Delta t}} \min\left(\frac{\Delta x^2}{\Delta t}, \Delta x\right)\right). \tag{22}$$

Remark (CFL condition and error estimates). The CFL condition (13) is needed to prove the error estimate (22), and it is not assumed to prove the convergence result in Theorem 4.1. If the assumption (A4) holds, no CFL condition is needed to prove the error estimate (20).

Proof. As is standard in this kind of proof, we only prove that

$$u(t,x) - w(t,x) \le C\left(\frac{\mathcal{E}(\Delta t, \Delta x)}{\sqrt{\Delta t}} + \sqrt{\Delta t}\right) + \sup_{x\in J^{\Delta x}} |u_0(x) - w(0,x)| \quad \text{in } \mathcal{G}^{\Delta}, \tag{23}$$

since the reverse inequality is obtained with small modifications. Assume that $T \le 1$ (the case $T \ge 1$ is obtained by induction).

For $i \in \{1, \dots, N\}$ and $j \in \mathbb{N}$, we set $x_j^i = j\Delta x e_i$, and we define the extension in the continuous space of w as

$$w_{\#}(t_n, x) = \mathbb{I}[\hat{w}(t_n, \cdot)](x).$$

First, we assume that

$$u_0(x_j^i) \ge w_{\#}(0, x_j^i) \quad \forall i \in \{0, \dots, N\} \text{ and } j \in \mathbb{N},$$

and we define

$$0 \le \mu_0 := \sup_{x\in J} \{|u_0(x) - w_{\#}(0,x)|\},$$

assuming without any restriction that $\mu_0 \le K$. For every $\beta, \eta \in (0,1)$ and $\sigma > 0$, we define an auxiliary function: for $(t,s,x) \in [0,T) \times \{t_n : n = 0, \dots, N_T\} \times J$,

$$\psi(t,s,x) := u(t,x) - w_{\#}(s,x) - \frac{(t-s)^2}{2\eta} - \beta|x|^2 - \sigma t.$$

Using Proposition 3.1(iii) and the inequality $|u(x,t) - u_0(x)| \le C_T$ (which holds for the continuous solution; see Theorem 2.14 in [21]), we deduce that $\psi(t,s,x) \to -\infty$ as $|x| \to +\infty$, and then the function ψ achieves its maximum at some point $(\bar{t}_\beta, \bar{s}_\beta, \bar{x}_\beta)$. In particular, we have

$$\psi(\bar{t}_\beta, \bar{s}_\beta, \bar{x}_\beta) \ge \psi(0,0,0) = u_0(0) - w_{\#}(0,0) \ge 0.$$

In the following, we denote by K several positive constants depending only on the Lipschitz constants of u.

Case 1: $\bar{x}_\beta \in J_i \setminus \{0\}$. In this case, we duplicate the space variable by considering, for $\varepsilon \in (0,1)$,

$$\begin{aligned}\psi_1(t,s,x,y) = u(t,x) - w_{\#}(s,y) - \frac{(t-s)^2}{2\eta} - \frac{d(x,y)^2}{2\varepsilon} - \frac{\beta}{2}(|x|^2 + |y|^2) - \sigma t \\ - \frac{\beta}{2}|x - \bar{x}_\beta|^2 - \frac{\beta}{2}|y - \bar{x}_\beta|^2 - \frac{\beta}{2}|t - \bar{t}_\beta|^2 - \frac{\beta}{2}|s - \bar{s}_\beta|^2 \\ \text{for } (t,s,x,y) \in [0,T) \times \{t_n : n = 0, \dots, N_T\} \times J \times J.\end{aligned}$$

Using Proposition 3.1(iii) again, the inequality $|u(x,t) - u_0(x)| \le C_T$, and the fact that u_0 is Lipschitz continuous, we deduce that $\psi_1(t,s,x,y) \to -\infty$ as $|x|, |y| \to +\infty$, and then the function ψ_1 achieves its maximum at some point $(\bar{t}, \bar{s}, \bar{x}, \bar{y})$, i.e.,

$$\psi_1(\bar{t}, \bar{s}, \bar{x}, \bar{y}) \ge \psi_1(t,s,x,y) \quad \forall (t,x), (s,y) \in [0,T) \times J.$$

It is also easy to show that $(\bar{t},\bar{s},\overline{x},\overline{y}) \to (\bar{t}_\beta,\bar{s}_\beta,\overline{x}_\beta,\overline{x}_\beta)$ as ε goes to zero and so $\overline{x},\overline{y} \in J_i \setminus \{0\}$ for ε small enough.

Step 1 *(basic estimates).* The maximum point of ψ_1 satisfies the following estimates:

$$d(\overline{x},\overline{y}) \leq K\varepsilon, \quad |\bar{t}-\bar{s}| \leq K\eta, \tag{24}$$

$$\beta\left(|\overline{x}|^2+|\overline{y}|^2\right) \leq K, \quad \beta\left(|\bar{x}-\bar{x}_\beta|^2+|\bar{y}-\bar{x}_\beta|^2+|\bar{t}-\bar{t}_\beta|^2+|\bar{s}-\bar{s}_\beta|^2\right) \leq K. \tag{25}$$

From

$$\psi_1(\bar{t},\bar{s},\overline{x},\overline{y}) \geq \psi_1(\bar{t}_\beta,\bar{s}_\beta,\overline{x}_\beta,\overline{x}_\beta) = \psi(\bar{t}_\beta,\bar{s}_\beta,\overline{x}_\beta) \geq 0,$$

we get (using $0 \geq -(\bar{t}-\bar{s})^2/2\eta - d(\bar{x},\bar{y})^2/2\varepsilon - \sigma\bar{t}$)

$$\begin{aligned} &\frac{\beta}{2}(|\bar{x}|^2+|\bar{y}|^2)+\frac{\beta}{2}\left(|\bar{x}-\bar{x}_\beta|^2+|\bar{y}-\bar{x}_\beta|^2+|\bar{t}-\bar{t}_\beta|^2+|\bar{s}-\bar{s}_\beta|^2\right)\\ &\leq u(\bar{t},\bar{x})-w_\#(\bar{s},\bar{y}) \leq u_0(\bar{x})-w_\#(0,\bar{y})+K\bar{t}+K\bar{s} \leq K(1+|\bar{x}|+|\bar{y}|), \end{aligned} \tag{26}$$

where we used Proposition 3.1(i) (extended to all the points of J thanks to the monotonicity of the interpolation operator), [21, Theorem 2.14] for the second inequality, and the fact that $T \leq 1$ for the last one. Using Young's inequality (i.e., the fact that $|\bar{x}| \leq 1/\beta + \beta/4|\bar{x}|^2$ since $(\beta/2|\bar{x}|-1)^2 \geq 0$) (26) implies in particular that

$$\frac{\beta}{2}(|\bar{x}|^2+|\bar{y}|^2) \leq K\left(1+\frac{2}{\beta}+\frac{\beta}{4}(|\bar{x}|^2+|\bar{y}|^2)\right).$$

Multiplying by β and using $\beta \leq 1$, we finally deduce that

$$\beta|\bar{x}|,\ \beta|\bar{y}| \leq K.$$

Then, using this in (26), we have

$$\beta\left(|\bar{x}-\bar{x}_\beta|^2+|\bar{y}-\bar{x}_\beta|^2+|\bar{t}-\bar{t}_\beta|^2+|\bar{s}-\bar{s}_\beta|^2\right) \leq K\left(1+\frac{1}{\beta}\right)$$

and, in particular,

$$\beta\left(|\bar{x}-\bar{x}_\beta|+|\bar{y}-\bar{x}_\beta|+|\bar{t}-\bar{t}_\beta|+|\bar{s}-\bar{s}_\beta|\right) \leq K.$$

From $\psi_1(\bar{t},\bar{s},\overline{x},\overline{y}) \geq \psi_1(\bar{t},\bar{s},\overline{y},\overline{y})$ we get

$$\begin{aligned} \frac{d(\bar{x},\bar{y})^2}{2\varepsilon} &\leq u(\bar{t},\bar{x})-u(\bar{t},\bar{y})+\frac{\beta}{2}(|\bar{y}|^2-|\bar{x}|^2)+\frac{\beta}{2}(|\bar{y}-\bar{x}_\beta|^2-|\bar{x}-\bar{x}_\beta|^2)\\ &\leq Kd(\bar{x},\bar{y})+\frac{\beta}{2}(|\bar{x}|+|\bar{y}|)d(\bar{x},\bar{y})+\frac{\beta}{2}(|\bar{x}-\bar{x}_\beta|+|\bar{y}-\bar{x}_\beta|)d(\bar{x},\bar{y}) \leq Kd(\bar{x},\bar{y}), \end{aligned} \tag{27}$$

which implies the first estimate of (24). The second bound in (24) is deduced from $\psi(\bar{t},\bar{s},\overline{x},\overline{y}) \geq \psi(\bar{s},\bar{s},\overline{x},\overline{y})$ in the same way.

If we include the estimate

$$u(\bar{t},\bar{x})-w_\#(\bar{s},\bar{y}) \leq u_0(\bar{x})+K\bar{t}-w_\#(0,\bar{y})+K\bar{s} \leq K(\mu_0+d(\bar{x},\bar{y})+1) \leq K$$

in the first part of (26), we finally deduce (25).

Step 2 *(viscosity inequalities).* We claim that for σ large enough, the supremum of ψ_1 is achieved for $\bar{t} = 0$ or $\bar{s} = 0$. We prove the assertion by contradiction. Suppose $\bar{t} > 0$ and $\bar{s} > 0$. Using the fact that $(t, x) \to \psi_1(t, \bar{s}, x, \bar{y})$ has a maximum in $(\bar{x}, \bar{t})$ and that u is a subsolution, we get

$$\frac{\bar{t}-\bar{s}}{\eta} + \sigma + \beta(\bar{t} - \bar{t}_\beta) + H_i\left(\frac{d(\bar{x}, \bar{y})}{\varepsilon} + \beta|\bar{x}| + \beta(|\bar{x} - \bar{x}_\beta|)\right) \le 0. \tag{28}$$

Since $\bar{s} > 0$ we know that $\psi_1(\bar{t}, \bar{s}, \bar{x}, \bar{y}) \ge \psi_1(\bar{t}, \bar{s} - \Delta t, \bar{x}, y)$ for a generic y and, by defining $\varphi(s, y) = -\big(\frac{(\bar{t}-s)^2}{2\eta} + \frac{d(\bar{x},y)^2}{2\varepsilon} + \frac{\beta}{2}|y|^2 + \frac{\beta}{2}|y - \bar{x}_\beta|^2 + \frac{\beta}{2}|s - \bar{s}_\beta|^2\big)$, this implies that, for a generic y,

$$w_\#(\bar{s}, \bar{y}) - \varphi(\bar{s}, \bar{y}) \le w_\#(\bar{s} - \Delta t, y) - \varphi(\bar{s} - \Delta t, y).$$

In particular, we have that for any $z \in J^{\Delta x}$

$$w_\#(\bar{s}, \bar{y}) - \varphi(\bar{s}, \bar{y}) \le w(\bar{s} - \Delta t, z) - \varphi(\bar{s} - \Delta t, z).$$

By the monotonicity of the scheme and the fact that the scheme is invariant by the addition of constants, adding $w_\#(\bar{s}, \bar{y}) - \varphi(\bar{s}, \bar{y})$ we get, for any $z \in J^{\Delta x}$,

$$w(\bar{s}, z) = S[\hat{w}(\bar{s} - \Delta t)](z) \ge S[\hat{\varphi}(\bar{s} - \Delta t)](z) + C.$$

By the monotonicity of the interpolation operator, this implies

$$w_\#(\bar{s}, \bar{y}) = \mathbb{I}[\hat{w}(\bar{s}, \cdot)](\bar{y}) \ge \mathbb{I}[S[\hat{\varphi}(\bar{s} - \Delta t)](\cdot)](\bar{y}) + w_\#(\bar{s}, \bar{y}) - \varphi(\bar{s}, \bar{y}).$$

Simplifying by $w_\#(\bar{s}, \bar{y})$, we obtain

$$-\sum_i \phi_i(\bar{y}) S[\hat{\varphi}(\bar{s} - \Delta t)](y_i) = -\mathbb{I}[S[\hat{\varphi}(\bar{s} - \Delta t)](\cdot)](\bar{y}) \ge -\varphi(\bar{s}, \bar{y}),$$

where ϕ_i are the basis functions of the interpolation operator. Adding and subtracting $\mathbb{I}[\hat{\varphi}(\bar{s}, \cdot)](\bar{y}) - \mathbb{I}[\hat{\varphi}(\bar{s} - \Delta t, \cdot)](\bar{y})$ and dividing by Δt , we get

$$\sum_i \phi_i(\bar{y}) \left(\frac{\varphi(\bar{s}-\Delta t, y_i) - S[\hat{\varphi}(\bar{s}-\Delta t)](y_i)}{\Delta t} + \frac{\varphi(\bar{s}, y_i) - \varphi(\bar{s}-\Delta t, y_i)}{\Delta t}\right) \ge \mathcal{O}\left(\frac{\Delta x^2}{\varepsilon}\right),$$

where we have used $\varphi_{xx} = \mathcal{O}(\frac{1}{\varepsilon})$ together with the properties of the interpolation operator. We observe that $\frac{\varphi(\bar{s},y_i) - \varphi(\bar{s}-\Delta t, y_i)}{\Delta t} = \varphi_s(\bar{s}, y_i) + \mathcal{O}(\Delta t/\eta)$; then, using the consistency definition (Definition 3.4), we obtain

$$\sum \phi_i(\bar{y}) \left(-\varphi_s(\bar{s}, y_i) + H_i(\varphi_x(\bar{s} - \Delta t, y_i))\right) \ge \mathcal{O}\left(\frac{\Delta t}{\eta} + \frac{\Delta x^2}{\varepsilon}\right) + \frac{\mathcal{E}(\Delta t, \Delta x)}{\varepsilon}.$$

By the regularity of φ and H (Lipschitz continuous) and the interpolation error for the Lipschitz function, there exists a positive constant K such that

$$\varphi_s(\bar{s}, \bar{y}) + H_i(\varphi_x(\bar{s} - \Delta t, \bar{y})) \ge -K\left(\frac{\Delta t}{\eta} + \frac{\Delta x^2}{\varepsilon}\right) + \frac{\mathcal{E}(\Delta t, \Delta x)}{\varepsilon}. \tag{29}$$

We subtract (29) from (28) and use the explicit form of φ, obtaining

$$\begin{aligned}\sigma + \beta(\bar{s} - \bar{s}_\beta) + \beta(\bar{t} - \bar{t}_\beta) + H_i\left(\frac{d(\bar{x}, \bar{y})}{\varepsilon} + \beta|\bar{x}| + \beta(|\bar{x} - \bar{x}_\beta|)\right)& \\ - H_i\left(\frac{d(\bar{x}, \bar{y})}{\varepsilon} - \beta|\bar{y}| - \beta(|\bar{y} - \bar{x}_\beta|)\right) \le K\left(\frac{\Delta t}{\eta} + \frac{\Delta x^2}{\varepsilon}\right) + \frac{\mathcal{E}(\Delta t, \Delta x)}{\varepsilon}&.\end{aligned}$$

Then, using that H_i is Lipschitz continuous and the basic estimates of Step 1, we arrive at

$$\sigma < K\sqrt{\beta} + K\left(\frac{\Delta t}{\eta} + \frac{\Delta x^2}{\varepsilon}\right) + \frac{\mathcal{E}(\Delta t, \Delta x)}{\varepsilon} =: \sigma^*. \tag{30}$$

Therefore, we have that for a $\sigma \geq \sigma^*$, at least one between $\bar{t}$ and $\bar{s}$ is equal to zero.

Step 3 *(conclusion).* If $\bar{t} = 0$ (a similar argument applies if $\bar{s} = 0$), we have

$$\psi_1(0, \bar{s}, \bar{x}, \bar{y}) \leq u_0(\overline{x}) - w_{\#}(\bar{s}, \overline{y}) \leq u_0(\overline{x}) - u_0(\overline{y}) + C\overline{s} + \mu_0 \leq K\varepsilon + K\eta + \mu_0.$$

Taking $\sigma = \sigma^*$, we obtain

$$\begin{aligned} &u(t,x) - w_{\#}(t,x) - \frac{\beta}{2}\left(|x|^2 + |y|^2 + |x - \bar{x}_\beta|^2 + |y - \bar{x}_\beta|^2 + |t - \bar{t}_\beta|^2 + |s - \bar{s}_\beta|^2\right) \\ &- \left(K\sqrt{\beta} + K\left(\frac{\Delta t}{\eta} + \frac{\Delta x^2}{\varepsilon}\right) + \frac{\mathcal{E}(\Delta t, \Delta x)}{\varepsilon}\right) T \leq K\varepsilon + K\eta + \mu_0, \end{aligned}$$

where, sending $\beta \to 0$ and choosing $\varepsilon = \eta = \sqrt{\Delta t}$, we get the desired estimate.

Case 2: $\overline{x}_\beta = 0$.

Step 1 *(basic estimates).* This step is identical to the previous case. In addition, we observe that, assuming

$$\sigma > K\sqrt{\beta} + K\left(\frac{\mathcal{E}(\Delta t, \Delta x)}{\varepsilon} + \frac{\Delta t}{\varepsilon} + \frac{\Delta x}{\varepsilon}\right) \tag{31}$$

(which is compatible with $\sigma > \sigma^*$), then there exists an $\bar{A} \in \mathbb{R}$ such that

$$\frac{\bar{s}_\beta - \bar{t}_\beta}{\eta} - K\left(\frac{\mathcal{E}(\Delta t, \Delta x)}{\varepsilon} + \frac{\Delta t}{\varepsilon}\right) + K\sqrt{\beta} > \bar{A} > \frac{\bar{s}_\beta - \bar{t}_\beta}{\eta} - \sigma + K\sqrt{\beta}. \tag{32}$$

Using the fact that $(t,x) \to \psi(t, \bar{s}_\beta, x)$ has a maximum in $(\bar{t}_\beta, \bar{x}_\beta)$ and that u is a subsolution, we get

$$\frac{\bar{t}_\beta - \bar{s}_\beta}{\eta} + \sigma + F_{-L_0}\left(\partial_x \varphi(\bar{t}_\beta, 0)\right) \leq 0, \tag{33}$$

with $\varphi(t,x) = w_{\#}(\bar{s}_\beta, x) + \frac{(t - \bar{s}_\beta)^2}{2\eta} + \beta|x|^2 + \sigma t$, and from (33) and (32),

$$\bar{A} > F_{-L_0}\left(\partial_x \varphi(\bar{t}_\beta, 0)\right). \tag{34}$$

We use (34), the definition of F_{-L_0}, and the coercivity of the Hamiltonians to obtain the existence of values λ_i such that

$$H_i(\lambda_i) = H_i^+(\lambda_i) = \bar{A} \tag{35}$$

(cf. Figure 2), which will be useful in the remaining part of the proof.

Now we move on to identify the right test function to treat this case. We duplicate the space variable differently than in Case 1. We consider, for $\varepsilon \in (0,1)$,

$$\begin{aligned} \psi_2(t,s,x,y) = u(t,x) - w_{\#}(s,y) - \frac{(t-s)^2}{2\eta} - \frac{d(x,y)^2}{2\varepsilon} - \frac{\beta}{2}(|x|^2 + |y|^2) \\ - \sigma t - (h(x) + h(y)) - \frac{\beta}{2}(t - \bar{t}_\beta)^2 - \frac{\beta}{2}(s - \bar{s}_\beta)^2 \\ \text{for } (t,s,x,y) \in [0,T) \times \{t_n : n = 0, \ldots, N_T\} \times J \times J, \end{aligned}$$

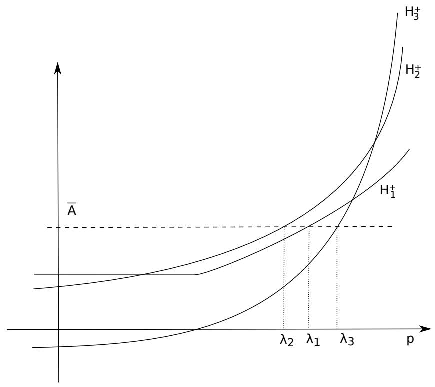

FIG. 2. *An example of H_i^+ functions.*

where $h(x) = \lambda_i x$ if $x \in J_i$ and the λ_i are defined in (35).

We denote by $(\bar{t}, \bar{s}, \bar{x}, \bar{y})$ the maximum point of ψ_2 (we keep the same notation as the previous case, but they are possibly different points). We remark that $(\bar{t}, \bar{s}, \bar{x}, \bar{y}) \to (\bar{t}_\beta, \bar{s}_\beta, \bar{x}_\beta, \bar{x}_\beta)$ as $\varepsilon \to 0$.

Step 2 *(viscosity inequalities).* We claim that for σ large enough, the supremum of ψ_1 is achieved for $\bar{t} = 0$ or $\bar{s} = 0$. We prove the assertion by contradiction. Suppose $\bar{t} > 0$ and $\bar{s} > 0$. We can have different scenarios: If $\bar{x}$ and $\bar{y}$ belong to the same arc (junction point excluded), the case is contained in Case 1. If instead $\bar{x} \in J_i \setminus \{0\}$, $\bar{y} \in J_j$ ($\bar{x}$ and $\bar{y}$ belong to different arcs), we can repeat the same argument to obtain (28) with the test function ψ_2. We have

$$\frac{\bar{t} - \bar{s}}{\eta} + \beta(\bar{t} - \bar{t}_\beta) + \sigma + H_i\left(\frac{d(\bar{x}, \bar{y})}{\varepsilon} + 2\beta|\bar{x}| + \lambda_i\right) \leq 0.$$

Observing that the argument inside the Hamiltonian is bigger than λ_i, we use (35) and arrive at

$$0 \geq \frac{\bar{t} - \bar{s}}{\eta} + \beta(\bar{t} - \bar{t}_\beta) + \sigma + H_i^+(\lambda_i) = \frac{\bar{t}_\beta - \bar{s}_\beta}{\eta} + \sigma + \bar{A} + K\sqrt{\beta},$$

which contradicts (32). Then this case cannot occur.

We go on to the last case to consider $\bar{x} = 0$, $\bar{y} \in J_i \setminus \{0\}$. First of all, we notice that the *basic estimates* (24)–(25) are still valid for $(\bar{t}, \bar{s}, \bar{x}, \bar{y})$ maximum points of ψ_2 since the added terms $h(x)$, $h(y)$ are easily included in the other linear elements of the estimates.

In this case, the difficulty comes from comparing two Hamiltonians evaluated, respectively, on the junction point and on one arc. Using the subsolution property with the test function ψ_2, we have as the first equation

$$\frac{\bar{t} - \bar{s}}{\eta} + \beta(\bar{t} - \bar{t}_\beta) + \sigma + F_{-L_0}\left(\frac{-|\bar{y}|}{\varepsilon} + \lambda_i\right) \leq 0, \tag{36}$$

where

$$F_{-L_0}\left(\frac{-|\bar{y}|}{\varepsilon} + \lambda_i\right) = \max\left(-L_0, \max_j\left(H_j^-\left(\frac{-|\bar{y}|}{\varepsilon} + \lambda_i\right)\right)\right).$$

From the definition of F_A it is also valid that

$$\frac{\bar{t}-\bar{s}}{\eta} + \beta(\bar{t}-\bar{t}_\beta) + \sigma + H_i^-\left(\frac{-|\bar{y}|}{\varepsilon} + \lambda_i\right) \leq 0. \tag{37}$$

Since $\bar{y} \in J_i \setminus \{0\}$, with the same argument as that used to obtain (29) (but for the test function ψ_2) and using the consistency result, we have

$$\frac{\bar{t}-\bar{s}}{\eta} + \beta(\bar{s}-\bar{s}_\beta) + H_i\left(\frac{-|\bar{y}|}{\varepsilon} - 2\beta\bar{y} + \lambda_i\right) \geq K\left(\frac{\mathcal{E}(\Delta t,\Delta x)}{\varepsilon} + \frac{\Delta t}{\eta} + \frac{\Delta x^2}{\varepsilon}\right).$$

Now recalling that $H^+(\lambda_i) = \bar{A}$,

$$\begin{aligned}
&\frac{\bar{t}-\bar{s}}{\eta} + \beta(\bar{s}-\bar{s}_\beta) + H_i^+\left(\frac{-|\bar{y}|}{\varepsilon} - 2\beta\bar{y} + \lambda_i\right) - K\left(\frac{\mathcal{E}(\Delta t,\Delta x)}{\varepsilon} + \frac{\Delta t}{\eta} + \frac{\Delta x^2}{\varepsilon}\right)\\
&\qquad \leq \frac{\bar{t}-\bar{s}}{\eta} + \beta(\bar{s}-\bar{s}_\beta) + H_i^+(\lambda_i) - K\left(\frac{\mathcal{E}(\Delta t,\Delta x)}{\varepsilon} + \frac{\Delta t}{\eta} + \frac{\Delta x^2}{\varepsilon}\right)\\
&\qquad\qquad \leq \frac{\bar{t}_\beta-\bar{s}_\beta}{\eta} + K\sqrt{\beta} + \bar{A} - K\left(\frac{\mathcal{E}(\Delta t,\Delta x)}{\varepsilon} + \frac{\Delta t}{\eta} + \frac{\Delta x^2}{\varepsilon}\right) < 0
\end{aligned}$$

for ε small enough, where we used $\beta(\bar{s}-\bar{s}_\beta) \leq K\sqrt{\beta}$ (basic estimates). We can claim that

$$\frac{\bar{t}-\bar{s}}{\eta} + \beta(\bar{s}-\bar{s}_\beta) + H_i^-\left(\frac{-|\bar{y}|}{\varepsilon} - 2\beta\bar{y} + \lambda_i\right) \geq K\left(\frac{\mathcal{E}(\Delta t,\Delta x)}{\varepsilon} + \frac{\Delta t}{\eta} + \frac{\Delta x^2}{\varepsilon}\right). \tag{38}$$

Finally, we subtract (38) from (36), obtaining the desired estimate on σ:

$$\sigma \leq K\sqrt{\beta} + K\left(\frac{\mathcal{E}(\Delta t,\Delta x)}{\varepsilon} + \frac{\Delta t}{\varepsilon} + \frac{\Delta x}{\varepsilon}\right) := \sigma^*. \tag{39}$$

In this case we obtain a contradiction with (31): since, assuming $\sigma > \sigma^*$, at least one between $\bar{t}$ and $\bar{s}$ is equal to zero.

Step 3 *(conclusion).* We obtain the same estimate as in Case 1.

It just remains to prove the general case (for which we do not assume that $u^0(x) \geq w_\#(0,x)$ for all $x \in J^{\Delta x}$). Remarking that $\bar{u} = u + \mu_1$ with $\mu_1 = \sup_{x\in J^{\Delta x}}(w_\#(0,x) - u_0(x))$ is a solution of the same equation of u but satisfying $\bar{u}(0,x) \geq w_\#(0,x)$ for all $x \in J^{\Delta x}$, we deduce that $\bar{u}$ satisfies

$$\begin{aligned}
&\sup_{(t,x)\in\mathcal{G}^\Delta}(u(t,x) + \mu_1 - w(t,x))\\
&\qquad\qquad \leq C\left(\frac{\mathcal{E}(\Delta t,\Delta x)}{\sqrt{\Delta t}} + \sqrt{\Delta t}\right) + \sup_{x\in J^{\Delta x}}|u_0(x) + \mu_1 - w(0,x)|,
\end{aligned}$$

which implies (23), ending the proof of the theorem. ∎

5. Numerical tests.

In this section, we present some numerical simulations to show the features and the convergence properties of the scheme proposed. In the first two tests assumptions (A1)–(A4) are verified, while in the last test (A4) does not hold.

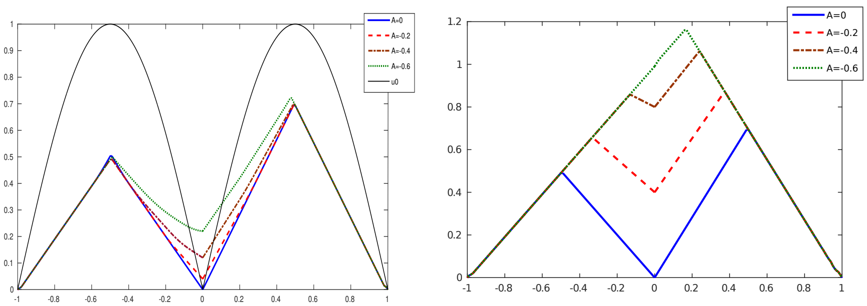

FIG. 3. *Initial condition and numerical solution at time* $t = 0.2$ *(left) and at time* $T = 2$ *(right), computed with parameter* $L_0 = 0, 0.2, 0.4, 0.6$.

Test 1. We consider a basic network composed of two edges connecting the nodes $(-1, 0)$ and $(1, 0)$ with a junction in 0. This case can be seen as a 1D problem in $\Gamma = \Gamma_1 \cup \Gamma_2 = [-1, 0] \cup [0, 1] = [-1, 1]$ with a discontinuity on the Hamiltonian at the origin. Despite its simplicity, this test helps us to understand the effect of the flux limiter contained in the operator F_A.

We consider the following Hamiltonian on Γ:

$$H(x,p) = \begin{cases} \frac{p^2}{2} - \frac{1}{2}, & x \in \Gamma_1, \\ \frac{p^2}{2} - 1, & x \in \Gamma_2. \end{cases}$$

This example was used as a benchmark also in [20]. Using the Legendre transform, we rewrite (5) as

$$H(x,p) = \begin{cases} \max\limits_{\alpha \in \mathbb{R}} \left(\alpha_1 p - \dfrac{\alpha_1^2}{2} \right) - \dfrac{1}{2}, & x \in \Gamma_1, \\ \max\limits_{\alpha \in \mathbb{R}} \left(\alpha_2 p - \dfrac{\alpha_2^2}{2} \right) - 1, & x \in \Gamma_2, \end{cases}$$

where we can deduce $L_1(\alpha_1) = (\alpha_1^2 + 1)/2$ and $L_2(\alpha_2) = (\alpha_1^2 + 2)/2$. We choose as initial condition $u_0(x) = \sin(\pi|x|)$, and we impose Dirichlet boundary conditions $u(t, -1) = u(t, 1) = 0$. The Dirichlet boundary conditions are implemented numerically by truncating the characteristics that cross the boundary, as in [15].

In Figure 3, we show the numerical solution at time $t = 0.2$ and $T = 2$ computed with parameter $L_0 = 0, 0.2, 0.4, 0.6$. We can observe that the asymmetry of the Hamiltonian with respect to the origin induces an asymmetric behavior on the solution. We can also highlight how the choice of parameter L_0 influences *globally* the value function of the problem. In fact, when $L_0 = 0$ the optimal control in $x = 0$ is simply $\alpha_0 = 0$, which corresponds to a zero cost, and since $u_0(0) = 0$, the solution $u(t, 0) = 0$ for each $t \in [0, T]$. This explains the choice of the name *flux limiter* for L_0: in this case the parameter blocks the passage of information between the two arcs which could be solved separately. In the case of $L_0 > 0$ the situation is different: the control $\alpha_0 = 0$ *does not correspond* to a null cost. A trajectory, which remains on the junction point, entails a cost. Furthermore, we observe that for values of $|L_0|$ sufficiently large, the behavior of the solution does not change anymore with respect to L_0. This happens because remaining on the junction point is no longer a convenient choice, i.e., the transition condition (6) is reached only by one nonincreasing

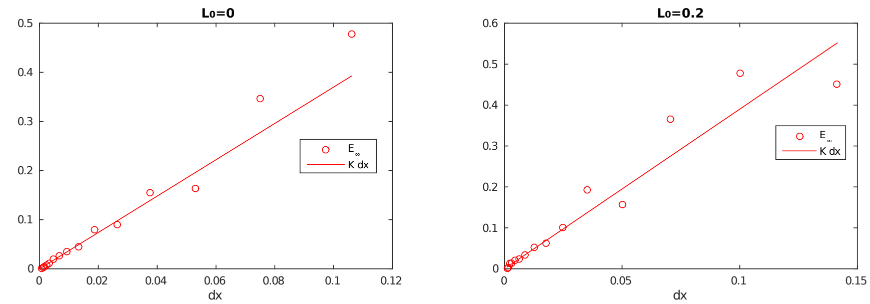

FIG. 4. *Graphic of* E_∞^Δ *with respect the space step, together with the line* $K\Delta x$*. Left:* $L_0 = 0$*, with* $K = 3.7$*. Right:* $L_0 = 0.2$*, with* $K = 3.9$*.*

Hamiltonian. Therefore, the flux limiter is not active anymore.

In Figure 4, we show the convergence rates in the case of $L_0 = 0$ and $L_0 = 0.2$. In the absence of an analytic exact solution, we compare the approximated solution $w(T, x)$ with an approximation $u(T, x)$ obtained on a very fine grid with $\Delta x = 10^{-4}$ and $\Delta t = 2.5\Delta x$. The error is evaluated with respect to the uniform discrete norm defined by

$$E_\infty^\Delta := \max_{x \in J^{\Delta x}} (|w(T, x) - u(T, x)|). \tag{40}$$

The error E_∞^Δ as a function of Δx is represented in Figure 4 for the choice $L_0 = 0$ (left) and $L_0 = 0.2$ (right), with the final time T and the time step fixed at $T = 0.2$ and $\Delta t = 2.5\Delta x$. We observe in the case $L_0 = 0$ a linear decay of the E_∞^Δ error; in particular, the E_∞^Δ errors fit with a linear regression curve of ratio $K_1 = 3.7$. We also observe the same convergence order in the case $L_0 = 0.2$, with almost the same ratio $K_1 = 3.9$. In this case, since Theorem 4.2 holds, their rate of convergence is 1 independently of the choice of Δt, so large time steps are allowed (cf. [15], where a similar property is discussed for Euclidean domains).

Test 2. We consider a simple junction network composed of three edges connecting the nodes $(0, 1)$, $(-1, -1)$, $(1, -1)$ with the junction point placed at $(0, 0)$. We denote by J_1 the edge connecting $(0, 1)$ to $(0, 0)$ and by J_2, J_3 the edges connecting $(0, 0)$ to $(1, -1)$ and $(-1, -1)$, respectively. The cost functions $L_i, i = 0, 1, 2, 3$, are defined as

$$L_i(\alpha_i) = \begin{cases} \frac{\alpha_i^2}{2} + 1 & \text{if } i = 1, 3, \\ \frac{\alpha_i^2}{2} + 2 & \text{if } i = 2, \\ 2 & \text{if } i = 0 \text{ and } \alpha_0 = 0, \\ +\infty & \text{if } i = 0 \text{ and } \alpha_0 \neq 0. \end{cases}$$

We impose Dirichlet boundary conditions on the boundary nodes:

$$u(t, x) = \begin{cases} 0 & \text{if } x = \{(-1, -1), (1, -1)\}, \\ \sqrt{2} + 1 & \text{if } x = \{(0, 1)\}. \end{cases}$$

The initial value u_0 is chosen as the restriction of $1 + x_2$ on J, where we denote $(x_1, x_2) = x$. In Figure 5, we show the color map of the initial condition and of the numerical solution at time $t = 0.5, 1, 1.5$, projected on the state coordinate plane. It is possible to observe that the initial datum u_0 (Figure 5, top left) quickly evolves to

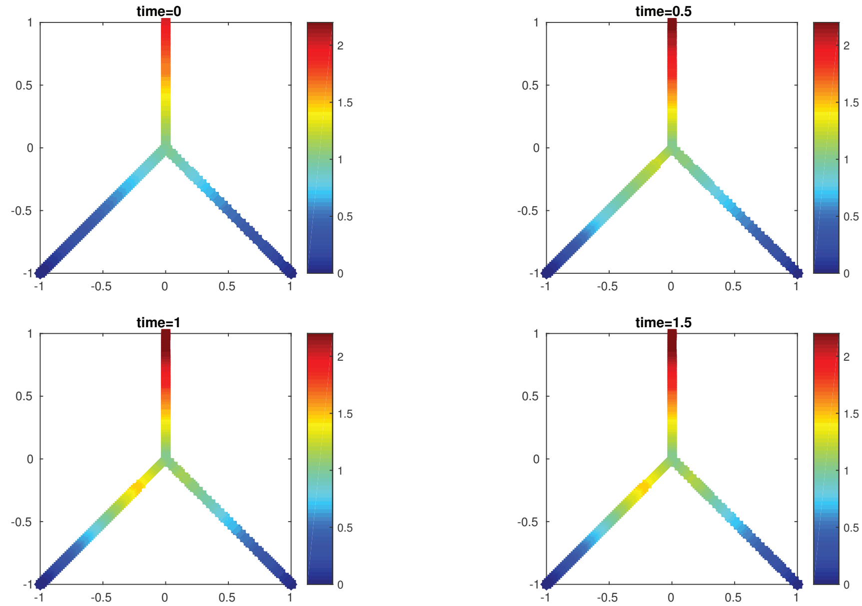

FIG. 5. *Projection on the state coordinate plane of the initial condition (top left), numerical solution at time $t = 0.5$ (top right), $t = 1$ (bottom left), and $t = 1.5$ (bottom right).*

the stationary solution (Figure 5, bottom right), which represents a weighted distance from the boundary points, with exit costs equal to the boundary values.

We compare the approximate solution at $T = 1.5$ with the exact solution of the corresponding stationary problem; this makes sense since the approximate solution has already reached the steady state. The exact steady state solution is

$$u(x) = \begin{cases} \sqrt{2} + x_2 & \text{if } x \in J_1, \\ \min\left(2\sqrt{(x_1-1)^2 + (x_2+1)^2}, \sqrt{2} + 2\sqrt{x_1^2 + x_2^2}\right) & \text{if } x \in J_2, \\ \sqrt{(x_1+1)^2 + (x_2+1)^2} & \text{if } x \in J_3. \end{cases} \tag{41}$$

In Figure 6, we show the behavior of the error (40) for various values of Δx, setting $\Delta t = 2.5\Delta x$. We observe as in the first test a linear decay of the E_∞^Δ, allowing large time steps.

Test 3. We conclude this section by treating a more complex network. We consider a network formed by 4-junctions and 8-arcs, defined in $(x_1, x_2) \in \mathbb{R}^2$ and connecting the points $V_1 = (-2, 0)$, $V_2 = (-1, 0)$, $V_3 = (0, 2)$, $V_4 = (0, 1)$, $V_5 = (2, 0)$, $V_6 = (1, 0)$, $V_7 = (0, -2)$, $V_8 = (0, -1)$. We define the edges $J_1 = \overline{V_1V_2}$, $J_2 = \overline{V_2V_4}$, $J_3 = \overline{V_2V_8}$, $J_4 = \overline{V_3V_4}$, $J_5 = \overline{V_4V_6}$, $J_6 = \overline{V_5V_6}$, $J_7 = \overline{V_6V_8}$, $J_8 = \overline{V_7V_8}$. We choose the costs on each arc as

$$L_i(x, \alpha_i) = \begin{cases} \frac{\alpha_i^2}{2} + (2 + x_1)^2 & \text{for } i = 1, 2, 3, 4, \\ \frac{\alpha_i^2}{2} + (-2 + x_2)^2 & \text{for } i = 5, 6, 7, 8. \end{cases}$$

We set the costs on the junction points V_i with $i = 2, 4, 8$ (when the relative control α_0 is null) equal to $L(V_i, \alpha) = 4$, and on V_6 equal to $L(V_6, \alpha) = 0.4$. We impose zero Dirichlet boundary conditions on the boundary points V_i with $i = 1, 3, 5, 7$.

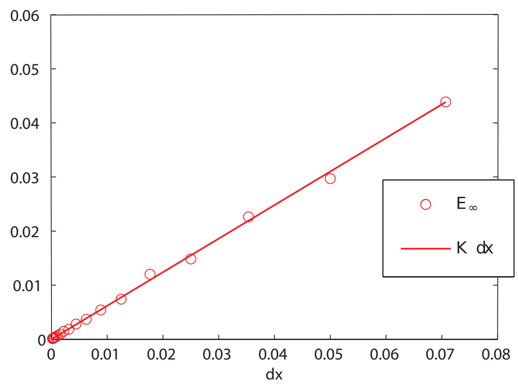

FIG. 6. *Graphic of* E_∞^Δ *with respect to the space step, together with the line* $K\Delta x$ *with* $K = 6.5$.

We choose as discretization steps $\Delta x = 0.013$, $\Delta t = 0.065$ and as the initial condition $u_0 = 2-(x_1^2+x_2^2)/2$. We observe that the CFL condition (13) is not verified, but we have convergence since Theorem 4.1 holds (but the estimate Theorem 4.3 does not hold).

In Figure 7, we show the initial value function and its evolution at different times. We observe that the value function starts from a symmetric configuration, and it loses its symmetry since the costs are not constant on the various edges of the network and they also depend on x. We highlight also that an optimal trajectory does not stop at junction points since it would run into a cost considerably higher than elsewhere in the network. The only exception is in V_6, where the waiting cost is 0.4, and it is where the value function assumes a local minimum. Clearly, for $t \to +\infty$ such a minimum tends to disappear since the cost is positive. In Figure 8, we also compare the value function at $t = 1.4$ with the value function computed using same data except for the cost in V_6, which is set to $L(V_6, \alpha) = 4$. We observe that since in this case the cost to stop in V_6 is higher, the local minimum disappears.

Choosing an effective Courant number is not a trivial task. In particular, large Courant numbers can considerably increase the complexity of the scheme near the junctions. In contrast, the opposite can produce low accuracy for long time approximations due to numerical diffusion effects.

In Table 1, we show the CPU times of a code implementing the proposed method for Test 3. The code is written in MATLAB R2018a and runs on a MacBook Pro 2017, 2.5 GHz, Intel Core i7. The variable **nc** represents the number of points used to compute the minimum by comparison, with respect to the time s_0, in (3).

The second and third columns show that the minimum complexity is reached when $\Delta t = \Delta x/8$, which corresponds to the case when the discrete characteristics do not cross the junctions. Choosing a time step lower than $\Delta x/8$ is no longer convenient in terms of computational time. In the last column, we consider the case when $L(V_i, \alpha) = 4$ for $i = 2, 4, 6, 8$. This choice corresponds to the case in which the characteristics never stop at the junctions. Therefore the minimization with respect to s_0 is not necessary, and **nc** is set to 1. In this case, the minimum complexity is reached using the largest time step.

In conclusion, Table 1 shows that large time steps may imply higher complexity due to the cost of exploring all the arcs and computing the minimum of the waiting time s_0 on the junctions. However, in semi-Lagrangian schemes, large time steps cor-

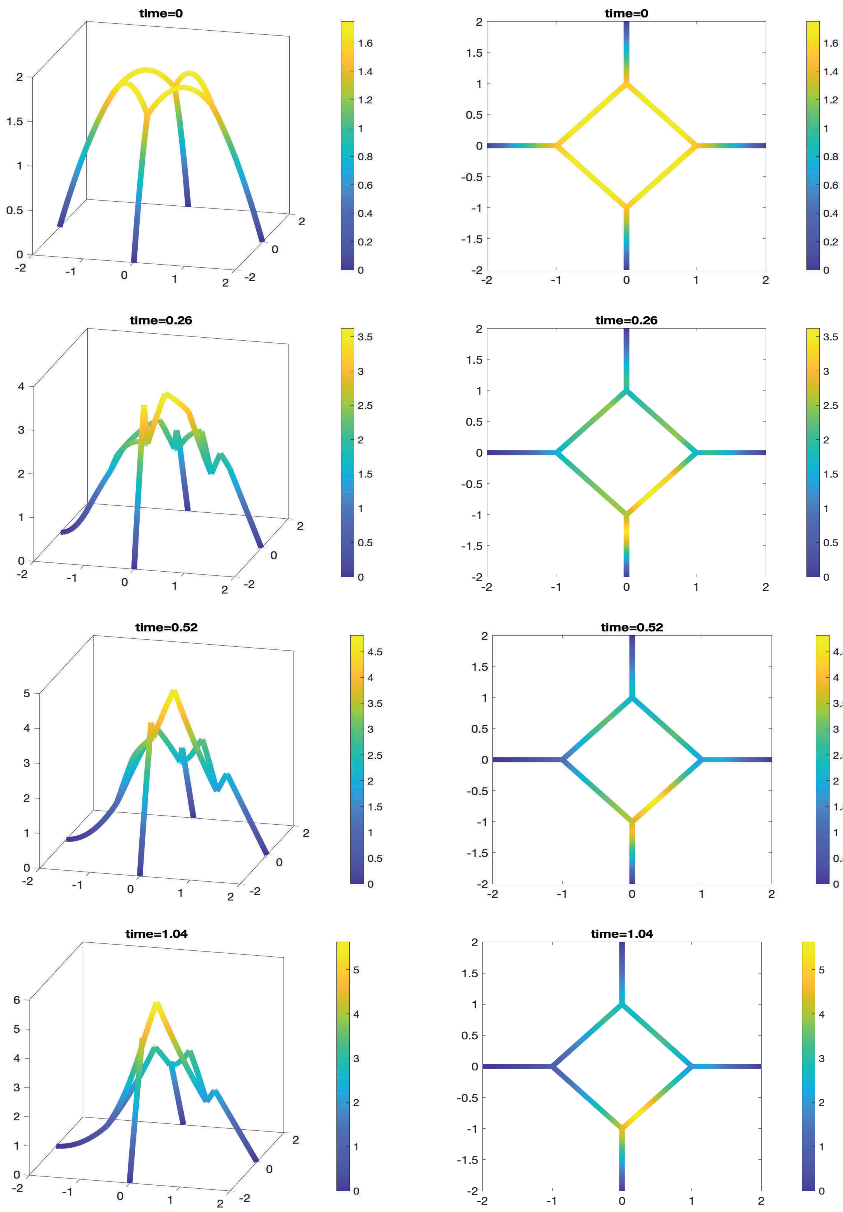

FIG. 7. *Evolution of the value function at times* $t = 0, 0.26, 0.52, 1.04$.

respond to less diffusive numerical approximations. Still, very large time steps can imply a low accuracy in the approximation of the characteristics (when the characteristics are not straight lines). The best choice, in terms of accuracy, may consist in choosing a time step which optimizes the consistency error proved in Proposition 3.5.

Appendix A. Definition of viscosity solution. Let us introduce the class of test functions. For $T > 0$, set $J_T = (0, T) \times J$. We define the class of test functions

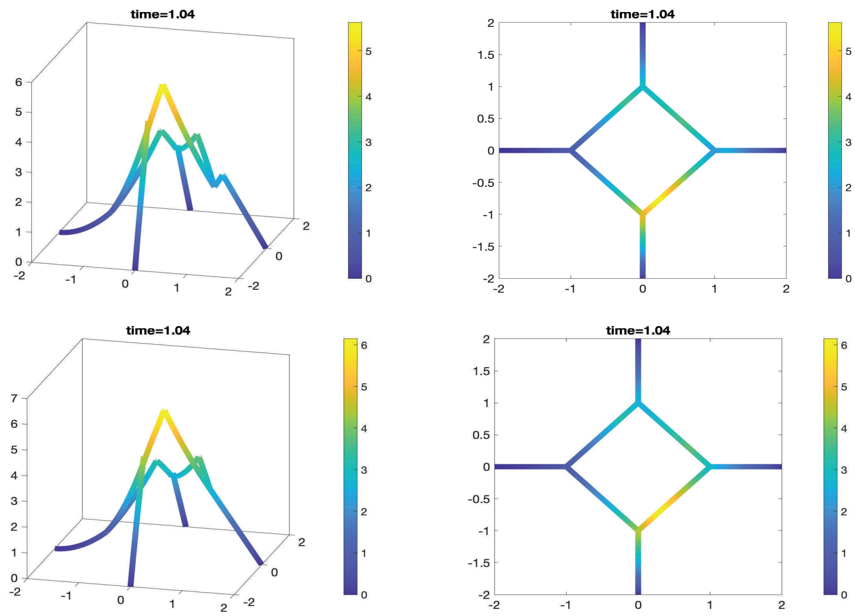

FIG. 8. *Comparison between the value functions varying the cost in* V_6*,* $L(V_6, \alpha) = 0.4$ *(above),* $L(V_6, \alpha) = 4$ *(below) at time* $t = 1.04$*.*

TABLE 1
CPU times in seconds for Test 3*, computed with* $\Delta x = 0.013$*.*

	$L(V_6,\alpha)=0.4$		$L(V_6,\alpha)=4$
Δt	$\mathbf{nc}=3$	$\mathbf{nc}=6$	$\mathbf{nc}=1$
$4\Delta x$	10.4	19.4	3.7
$2\Delta x$	9.8	19.0	3.7
Δx	9.6	18.0	3.9
$\Delta x/2$	8.8	15.6	4.1
$\Delta x/4$	8.3	13.2	5.2
$\Delta x/8$	7.4	7.5	7.2
$\Delta x/16$	14.7	14.8	14.4
$\Delta x/32$	29.5	30.0	28.5

on J_T and on J as

$$C^k(J_T) = \{\varphi \in C(J_T),\, \varphi \in C^k((0,T) \times J_i \;\forall i = 1, \dots, N)\},$$
$$C^k(J) = \{\varphi \in C(J),\, \varphi \in C^k(J_i) \;\forall i = 1, \dots, N\}.$$

We recall also the definition of upper and lower semicontinuous envelopes u^* and u_* of a (locally bounded) function u defined on $[0,T) \times J$,

$$u^*(t,x) = \limsup_{(s,y)\to(t,x)} u(s,y) \quad \text{and} \quad u_*(t,x) = \liminf_{(s,y)\to(t,x)} u(s,y).$$

We say that a test function φ touches a function u from below (resp., from above) at (t,x) if $u - \varphi$ reaches a local minimum (resp., maximum) at (t,x).

DEFINITION A.1 (flux-limited solutions). *Assume that the Hamiltonian satisfies some standard hypotheses of regularity, convexity, and coercivity and let* $u : [0,T) \times J \to \mathbb{R}$*.*

(i) *We say that u is a flux-limited subsolution (resp., flux-limited supersolution) of* (5) *in* $(0,T)\times J$ *if for all test functions* $\varphi \in C^1(J_T)$ *touching* u^* *from above (resp.,* u_* *from below) at* $(t_0,x_0)\in J_T$*, we have*

$$\begin{array}{lll} \varphi_t(t_0,x_0)+H_i(x_0,\varphi_x(t_0,x_0))\le 0 & (\textit{resp.}, \ge 0) & \textit{if } x_0\in J_i,\\ \varphi_t(t_0,x_0)+F_A(\varphi_x(t_0,x_0))\le 0 & (\textit{resp.}, \ge 0) & \textit{if } x_0=0. \end{array} \tag{42}$$

(ii) *We say that u is a flux-limited subsolution (resp., flux-limited supersolution) of* (5) *on* $[0,T)\times J$ *if additionally*

$$u^*(0,x)\le u_0(x) \quad (\textit{resp.},\ u_*(0,x)\ge u_0(x)) \ \ \forall x\in J. \tag{43}$$

(iii) *We say that u is a flux-limited solution if u is both a flux-limited subsolution and a flux-limited supersolution.*

In [21], an equivalent definition of viscosity solutions for (5) is proved. We use this equivalent definition in particular in the definition of the consistency in section 3. In the following theorem, we adapt this result for Hamiltonians depending on x. The proof is a straightforward adaptation of the one in [21].

THEOREM A.2 (equivalent definition for subsolutions/supersolutions). *Let* $\bar H^0 = \max_j \min_p H_j(0,p)$ *and consider* $A\in[\bar H^0,+\infty)$*. Given solutions* $p_i^A\in\mathbb{R}$ *of*

$$H_i\left(0,p_i^A\right)=H^+\left(0,p_i^A\right)=A, \tag{44}$$

let us fix any time-independent test function $\phi^0(x)$ *satisfying, for* $i=1,\dots,N$,

$$\partial_i\phi^0(0)=p_i^A.$$

Given a function $u:(0,T)\times J\to\mathbb{R}$*, the following properties hold true:*

(i) *If u is an upper semicontinuous subsolution of* (5) *with* $A=H_0$*, for* $x\neq 0$*, satisfying*

$$u(t,0)=\limsup_{(s,y)\to(t,0),\ y\in J_i^*} u(s,y), \tag{45}$$

then u is an H_0-flux limited subsolution.

(ii) *Given* $A>H_0$ *and* $t_0\in(0,T)$*, if u is an upper semicontinuous subsolution of* (5) *for* $x\neq 0$*, satisfying* (45)*, and if for any test function φ touching u from above at* $(t_0,0)$ *with*

$$\varphi(t,x)=\psi(t)+\phi^0(x), \tag{46}$$

for some $\psi\in C^2((0,+\infty))$*, we have*

$$\varphi_t+F_A(\varphi_x)\le 0 \quad \textit{at } (t_0,0),$$

then u is an A-flux limited subsolution at $(t_0,0)$.

(iii) *Given* $t_0\in(0,T)$*, if u is a lower semicontinuous supersolution of* (5) *for* $x\neq 0$*, and if for any test function φ satisfying* (46) *touching u from above at* $(t_0,0)$ *we have*

$$\varphi_t+F_A(\varphi_x)\ge 0 \quad \textit{at } (t_0,0),$$

then u is an A-flux limited supersolution at $(t_0,0)$.

Remark. In fact this theorem shows that, at the junction, it is sufficient to test with a test function having the form (46).

Appendix B. Proof of Proposition 3.2.

Proof. In this proof, we denote by C a constant that depends only on L_i and that may change line to line and with L_f the Lipschitz constant of a generic function f.

Let us just assume that $x, y \in J_i \cap J^{\Delta x}$. The latter is not restrictive since if $x \in J_j \cap J^{\Delta x}, y \in J_i \cap J^{\Delta x}$ with $j \neq i$, we come back to the case of a comparison between points belonging to the same arc, writing

$$|w(t_n, x) - w(t_n, y)| \leq |w(t_n, x) - w(t_n, 0)| + |w(t_n, 0) - w(t_n, y)|.$$

We denote by $\bar{\alpha}_i$ the optimal control of $S[w^{n-1}](y)$ associated to the i-arc. We consider three different cases:

(1) $\bar{\alpha}_i < |y|/\Delta t$ with $y \neq 0$. In this case, we consider α_i such that $x - \Delta t \alpha_i e_i = y - \Delta t \bar{\alpha}_i e_i$. This means that

$$|\alpha_i - \bar{\alpha}_i| = \frac{|x - y|}{\Delta t}. \tag{47}$$

Using the suboptimal control α_i for $S[\hat{w}^{n-1}](x)$ yields

$$\begin{aligned} w(t_n, x) - w(t_n, y) \leq\ & \mathbb{I}[\hat{w}^{n-1}]\left(-\left(x - \alpha_i \Delta t e_i\right)\right) + \Delta t L_i(x, \alpha_i) \\ & - \mathbb{I}[\hat{w}^{n-1}]\left(-\left(y - \bar{\alpha}_i \Delta t e_i\right)\right) - \Delta t L_i(y, \bar{\alpha}_i) \leq \Delta t L_{L_i} |\alpha_i - \bar{\alpha}_i| \leq L_{L_i}|x - y|. \end{aligned}$$

(2) $0 < \frac{|y|}{\Delta t} \leq \bar{\alpha}_i$. In this case the discrete trajectory starting from y passes through the junction. We denote by $(\bar{\alpha}_i, \bar{s}_0, \bar{\alpha}_0, \bar{j}, \bar{\alpha}_{\bar{j}})$ the optimal control associated with $S[\hat{w}^{n-1}](y)$. We distinguish two subcases:

(2.i) $x = 0$. In this case, we choose the suboptimal control $(\bar{s}_0 + \frac{|y|}{\bar{\alpha}_i}, \bar{\alpha}_0, \bar{j}, \bar{\alpha}_{\bar{j}})$ (if $\bar{\alpha}_0 \neq 0$, we replace it by 0 in order to stay in the origin), obtaining

$$\begin{aligned} w(t_n, x) - w(t_n, y) \leq\ & \mathbb{I}[\hat{w}^{n-1}]\left(-\left(\Delta t - \bar{s}_0 - \frac{|y|}{\bar{\alpha}_i}\right)\bar{\alpha}_{\bar{j}} e_{\bar{j}}\right) + \left(\bar{s}_0 + \frac{|y|}{\bar{\alpha}_i}\right) L_0(\bar{\alpha}_0) \\ & + \left(\Delta t - \bar{s}_0 - \frac{|y|}{\bar{\alpha}_i}\right) L_{\bar{j}}(0, \bar{\alpha}_{\bar{j}}) - \mathbb{I}[\hat{w}^{n-1}]\left(-\left(\Delta t - \bar{s}_0 - \frac{|y|}{\bar{\alpha}_i}\right)\bar{\alpha}_{\bar{j}} e_{\bar{j}}\right) \\ & - \left(\Delta t - \bar{s}_0 - \frac{|y|}{\bar{\alpha}_i}\right) L_{\bar{j}}(0, \bar{\alpha}_{\bar{j}}) - \bar{s}_0 L_0(\bar{\alpha}_0) - \frac{|y|}{\bar{\alpha}_i} L_i(y, \bar{\alpha}_i) \leq \frac{|y|}{\bar{\alpha}_i}\left(L_0(\bar{\alpha}_0) - L_i(y, \bar{\alpha}_i)\right). \end{aligned}$$

If $\bar{\alpha}_i \geq 1$, using that L_i is Lipschitz continuous, we get that there exists a constant C (depending only on $L_i(y, 0)$ and the Lipschitz constant of L_i) such that

$$\frac{|L_i(y, \bar{\alpha}_i)|}{|\bar{\alpha}_i|} \leq C.$$

Injecting the estimate above into (8) and using that $L_0(0)$ is bounded, we deduce

$$w(t_n, x) - w(t_n, y) \leq C|y| = Cd(x, y).$$

If $\bar{\alpha}_i \leq 1$, since $L_0(0)$ and $L_i(\bar{\alpha}_i)$ are bounded, there exists a constant C such that

$$w(t_n, x) - w(t_n, y) \leq C\frac{|y|}{\bar{\alpha}_i} \leq C\Delta t.$$

We finally get that in all the cases, $w(t_n,x) - w(t_n,y) \le C\left(\Delta t + d(x,y)\right)$.

(2.ii) $|x|>0$. In this case, we choose α_i such that $\frac{|x|}{\alpha_i} = \frac{|y|}{\bar\alpha_i}$. This implies in particular

$$x - \frac{|x|}{\alpha_i}\alpha_i e_i = y - \frac{|x|}{\alpha_i}\bar\alpha_i e_i,$$

and so $\frac{|x|}{\alpha_i}|\alpha_i - \bar\alpha_i| = |x-y| = d(x,y)$. Using the suboptimal control $(\alpha_i, \bar s_0, \bar\alpha_0, \bar j, \bar\alpha_{\bar j})$ for $S[w^{n-1}](x)$, we have

$$\begin{aligned}
&w(t_n,x) - w(t_n,y)\\
&\quad \le \mathbb{I}[\hat w^{n-1}]\left(-\left(\Delta t - \bar s_0 - \frac{|x|}{\alpha_i}\right)\bar\alpha_{\bar j} e_{\bar j}\right) + \left(\Delta t - \bar s_0 - \frac{|x|}{\alpha_i}\right) L_{\bar j}(0,\bar\alpha_{\bar j})\\
&\qquad + \bar s_0 L_0(\bar\alpha_0) + \frac{|x|}{\alpha_i} L_i(x,\alpha_i) - \mathbb{I}[\hat w^{n-1}]\left(-\left(\Delta t - \bar s_0 - \frac{|y|}{\bar\alpha_i}\right)\bar\alpha_{\bar j} e_{\bar j}\right)\\
&\qquad - \left(\Delta t - \bar s_0 - \frac{|y|}{\bar\alpha_i}\right) L_{\bar j}(0,\bar\alpha_{\bar j}) - \bar s_0 L_0(\bar\alpha_0) - \frac{|y|}{\bar\alpha_i} L_i(y,\bar\alpha_i)\\
&\qquad\qquad \le L_{L_i}\frac{|x|}{\alpha_i}|\alpha_i - \bar\alpha_i| \le L_{L_i} d(x,y).
\end{aligned}$$

(3) $y=0$. We denote by $(\bar s_0, \bar\alpha_0, \bar j, \bar\alpha_{\bar j})$ the optimal control associated to the operator $S[w^{n-1}](y)$. We distinguish two subcases again:

(3.i) $\bar s_0 = \Delta t$. We choose $\alpha_i \ge \max(1, \frac{|x|}{\Delta t})$ and the suboptimal control $(\alpha_i, \bar s_0 - \frac{|x|}{\alpha_i}, \bar\alpha_0)$ for $S[w^{n-1}](x)$, obtaining

$$\begin{aligned}
&w(t_n,x) - w(t_n,y)\\
&\quad \le \mathbb{I}[\hat w^{n-1}](0) + \left(\bar s_0 - \frac{|x|}{\alpha_i}\right) L_0(\bar\alpha_0) + \frac{|x|}{\alpha_i} L_i(x,\alpha_i) - \mathbb{I}[\hat w^{n-1}](0) - \bar s_0 L_0(\bar\alpha_0)\\
&\qquad\qquad \le \frac{|x|}{\alpha_i}\left(L_i(x,\alpha_i) - L_0(\bar\alpha_0)\right) \le L_{L_i} d(x,y).
\end{aligned}$$

Using that L_i is Lipschitz continuous, we get that there exists a constant C (depending only on $L_i(0), L_0(0)$ and on the Lipschitz constant of L_i) such that

$$\frac{|L_i(x,\bar\alpha_i)| + |L_0(\bar\alpha_0)|}{|\bar\alpha_i|} \le C.$$

This implies that $w(t_{n+1},x) - w(t_{n+1},y) \le C|x| = Cd(x,y)$.

(3.ii) $\bar s_0 < \Delta t$. We choose $\alpha_i \ge \max(1, |\bar\alpha_{\bar j}|)$ such that

$$\frac{|x|}{\alpha_i} \le \frac{\Delta t - \bar s_0}{2} \quad \text{and} \quad \frac{\Delta t - \bar s_0}{\Delta t - \bar s_0 - \frac{|x|}{\alpha_i}}|\bar\alpha_{\bar j}| \le \alpha_i. \tag{48}$$

We also set $\alpha_{\bar j} = \frac{\Delta t - \bar s_0}{\Delta t - \bar s_0 - \frac{|x|}{\alpha_i}}\bar\alpha_{\bar j}$, which satisfies in particular $\alpha_i \ge |\alpha_{\bar j}|$. Taking the suboptimal control $(\alpha_i, \bar{s}_0, \bar{\alpha}_0, \bar{j}, \alpha_{\bar{j}})$ for $S[w^{n-1}](x)$, we get

$$
\begin{aligned}
(49) \quad w(t_n,x) - w(t_n,y) &\le \mathbb{I}[\hat{w}^{n-1}]\left(-\left(\Delta t - \bar{s}_0 - \frac{|x|}{\alpha_i}\right)\alpha_{\bar{j}} e_{\bar{j}}\right) \\
&+ \left(\Delta t - \bar{s}_0 - \frac{|x|}{\alpha_i}\right) L_{\bar{j}}(0,\alpha_{\bar{j}}) + \bar{s}_0 L_0(\bar{\alpha}_0) + \frac{|x|}{\alpha_i} L_i(x,\alpha_i) \\
&- \mathbb{I}[\hat{w}^{n-1}]\left(-(\Delta t - \bar{s}_0)\,\bar{\alpha}_{\bar{j}} e_{\bar{j}}\right) - (\Delta t - \bar{s}_0)\, L_{\bar{j}}(0,\bar{\alpha}_{\bar{j}}) - \bar{s}_0 L_0(\bar{\alpha}_0) \\
&\le \frac{|x|}{\alpha_i}\left(L_i(x,\alpha_i) - L_{\bar{j}}(0,\alpha_{\bar{j}})\right) + (\Delta t - \bar{s}_0)\left(L_{\bar{j}}(0,\alpha_{\bar{j}}) - L_{\bar{j}}(0,\bar{\alpha}_{\bar{j}})\right) \\
&\le \frac{|x|}{\alpha_i}\left(L_i(x,\alpha_i) - L_{\bar{j}}(0,\alpha_{\bar{j}})\right) + (\Delta t - \bar{s}_0) L_{L_{\bar{j}}} |\alpha_{\bar{j}} - \bar{\alpha}_{\bar{j}}|.
\end{aligned}
$$

Using that $\alpha_i \ge 1$, we get that $\frac{|L_i(x,\alpha_i)}{\alpha_i} \le C$. In the same way (using that $\alpha_i \ge \alpha_{\bar{j}}$)

$$
\frac{|L_{\bar{j}}(0,\alpha_{\bar{j}})|}{\alpha_i} \le \frac{1}{\alpha_i}\left(L_{\bar{j}}(0,0) + L_{L_{\bar{j}}}|\alpha_{\bar{j}}|\right) \le L_{\bar{j}}(0,0) + L_{L_{\bar{j}}}\frac{|\alpha_{\bar{j}}|}{\alpha_i} \le C.
$$

Finally, using the definition of $\alpha_{\bar{j}}$, we observe that

$$
(\Delta t - \bar{s}_0)|\alpha_{\bar{j}} - \bar{\alpha}_{\bar{j}}| = |x|\frac{|\alpha_{\bar{j}}|}{\alpha_i} \le |x|.
$$

Injecting these estimates into (49), we obtain $w(t_n,x) - w(t_n,y) \le C|x| = Cd(x,y)$. ∎

Appendix C. Proof of Theorem 4.2.

Proof. The proof is made by induction assuming that for $n \ge 1$

$$
(50) \qquad |w^{n-1}(x) - u(t_{n-1},x)| \le (n-1)C\Delta x \quad \forall x \in J^{\Delta x}.
$$

Note that the thesis is clearly satisfied for $n = 1$. We then want to show that

$$
|w^n(x) - u(t_n,x)| \le nC\Delta x \quad \forall x \in J^{\Delta x}.
$$

From Proposition 2.1, we know that

$$
(51) \qquad u(t_n,x) := \inf_{y\in J} \inf_{(X(.),\alpha(.))\in\Gamma^{t_n,x}_{t_{n-1},y}} \left\{ u(t_{n-1},y) + \int_{t_{n-1}}^{t_n} L(X(\tau),\alpha(\tau))d\tau \right\},
$$

where we use the notation $L(X(\tau),\alpha(\tau)) \equiv L_i(\alpha_i)$ if $X(\tau) \in J_i$ (see (A4)).

We denote $\bar{\alpha} = (\bar{\alpha}_0, \bar{\alpha}_1, \ldots, \bar{\alpha}_N)$ and by $\bar{s}_0$ the optimal argument of $S[w^{n-1}](x)$, and we treat only the case where $x \in J_i\backslash\{0\}$ and $|x|/\bar{\alpha}_i < \Delta t$ (this corresponds to the more difficult case in which the optimal trajectory crosses the junction). We also denote by $\bar{X}(t)$ (with $t \in [t_{n-1}, t_n]$) the trajectory obtained applying the control $\bar{\alpha}$. Clearly such a trajectory belongs to $\Gamma^{t_n,x}_{t_{n-1},\bar{X}(t_{n-1})}$ with $\bar{X}(t_{n-1}) = \left(\Delta t - s_0 - \frac{|x|}{\bar{\alpha}_i}\right)e_j$ and

$$
(52) \qquad \begin{cases}
\bar{X}(t) \in J_i & \text{for } t \in \left[t_n - \frac{|x|}{\bar{\alpha}_i}, t_n\right), \\
\bar{X}(t) = 0 & \text{for } t \in \left[t_n - \frac{|x|}{\bar{\alpha}_i} - \bar{s}_0, t_n - \frac{|x|}{\bar{\alpha}_i}\right), \\
\bar{X}(t) \in J_j & \text{for } t \in \left[t_{n-1}, t_n - \frac{|x|}{\bar{\alpha}_i} - \bar{s}_0\right].
\end{cases}
$$

Therefore,

$$
\begin{aligned}
&u(t_n,x)-w^n(x)\\
&=\inf_{y\in J}\inf_{(X,\alpha)\in\Gamma_{t_{n-1},y}^{t_n,x}}\left[u(t_{n-1},y)+\int_{t_{n-1}}^{t_n}L(X(\tau),\alpha(\tau))d\tau\right]-S[w^n](x)\\
&\le u(t_{n-1},X(t_{n-1}))+\int_{t_{n-1}}^{t_n}L(X(\tau),\alpha(\tau))d\tau-S[w^n](x)\\
&\le u(t_{n-1},X(t_{n-1}))-\mathbb{I}[w^{n-1}](\bar X(t_{n-1}))|\\
&+\int_0^{\Delta t-\frac{|x|}{\bar\alpha_i}-\bar s_0}L(\bar X(\tau),\bar\alpha(\tau))d\tau-\left(\Delta t-\bar s_0-\frac{|x|}{\bar\alpha_i}\right)L_j(\bar\alpha_j)\\
&+\int_{\Delta t-\bar s_0-\frac{|x|}{\bar\alpha_i}}^{\Delta t-\frac{|x|}{\bar\alpha_i}}L(\bar X(\tau),\bar\alpha(\tau))d\tau-s_0L_0(\bar\alpha_0)+\int_{\Delta t-\frac{|x|}{\bar\alpha_i}}^{\Delta t}L(\bar X(\tau),\bar\alpha(\tau))d\tau-\frac{|x|}{\bar\alpha_i}L_i(\bar\alpha_i).
\end{aligned}
$$

Since $L(.,\alpha)$ is constant in time, the cost terms cancel. Moreover, using standard interpolation operator properties and (50), we observe that

$$
\begin{aligned}
&u(t_{n-1},X(t_{n-1}))-\mathbb{I}[w^{n-1}](\bar X(t_{n-1}))|\\
&=u(t_{n-1},X(t_{n-1}))-\mathbb{I}[u(t_{n-1},\cdot)](X(t_{n-1}))+(n-1)C\Delta x\\
&\quad+\mathbb{I}[u(t_{n-1},\cdot)-(n-1)C\Delta x](X(t_{n-1}))+\mathbb{I}[w^{n-1}](\bar X(t_{n-1}))\le nC\Delta x,
\end{aligned}
$$

and consequently

$$u(t_n,x)-w^n(x)\le nC\Delta x.$$

For the inverse inequality, we invert the whole argument. An additional difficulty arises in choosing the good control for the $S[w^{n-1}]$ term. We proceed considering a continuous optimal control $\bar\alpha(\cdot)$ for $u(t_n,x)$ in (51). Without loss of generality we assume that the associated trajectory $\bar X(t)$ is such that

$$
\left\{\begin{array}{ll}
\bar X(t)\in J_i & \text{for } t\in(\bar t_2,t_n]\,,\\
\bar X(t)=0 & \text{for } t\in(\bar t_1,\bar t_2]\,,\\
\bar X(t)\in J_j & \text{for } t\in[t_{n-1},\bar t_1]\,.
\end{array}\right. \tag{53}
$$

Indeed, we can exclude that an optimal trajectory passes in another arc or touches multiple times the junction point thanks to the convexity of the functions L. In fact, in such cases, it would be necessary for an optimal trajectory to pass twice for the same point, i.e., $X(\tilde t_1)=\tilde x$ and $X(\tilde t_2)=\tilde x$, with $X(t)\ne\tilde x$ for $t\in(\tilde t_1,\tilde t_2)$. This means that since $\dot X(t)=\bar\alpha(t)$, we have that

$$\int_{\tilde t_1}^{\tilde t_2}\bar\alpha(\tau)d\tau=X(\tilde t_1)-X(\tilde t_2)=0.$$

Then the average control on $[\tilde t_1,\tilde t_2]$ is zero. Using the strict convexity and Jensen's inequality, we find that the optimal control $\bar\alpha$ should be zero. This contradicts the definition of X.

We can now build a discrete control and an associated trajectory $(\hat\alpha,\hat X)$ for $S[\hat\varphi](x)$ such that

$$\hat\alpha_i=\frac{|x|}{t_n-\bar t_2}=\frac{1}{t_n-\bar t_2}\int_{\bar t_2}^{t_n}\bar\alpha_i(\tau)d\tau,\quad \hat s_0=\bar t_2-\bar t_1,$$

$$\hat{\alpha}_j = \frac{1}{\bar{t}_1 - t_{n-1}} \int_{t_{n-1}}^{\bar{t}_1} \bar{\alpha}_j(\tau) d\tau.$$

Then by construction $\hat{X}(t_{n-1}) = \bar{X}(t_{n-1})$ and

$$S[w^{n-1}](x) - u(t_n, x) = S[w^{n-1}](x) - u(t_{n-1}, y) + \int_{t_{n-1}}^{t_n} L(\bar{X}(\tau), \bar{\alpha}(\tau)) d\tau$$

$$\leq \mathbb{I}[w^{n-1}](\hat{X}(t_{n-1})) - u(t_{n-1}, \bar{X}(t_{n-1})) + (\Delta t - \bar{t}_2)\, L_i(\hat{\alpha}_i) - \int_{\bar{t}_2}^{\Delta t} L(\bar{X}(\tau), \bar{\alpha}(\tau)) d\tau$$

$$+ \left((\bar{t}_2 - \bar{t}_1) L_0 - \int_{\bar{t}_1}^{\bar{t}_2} L(\bar{X}(\tau), \bar{\alpha}(\tau)) d\tau \right) + \left(\bar{t}_1 L_j(\hat{\alpha}_j) - \int_{t_{n-1}}^{\bar{t}_1} L(\bar{X}(\tau), \bar{\alpha}(\tau)) d\tau \right).$$

Using Jensen's inequality, knowing that the L-functions are convex, we get

$$\begin{aligned}
&\bar{t}_1 L_j(\hat{\alpha}_j) - \int_{t_{n-1}}^{\bar{t}_1} L(\bar{X}(\tau), \bar{\alpha}(\tau)) d\tau \\
&= \bar{t}_1 L_j \left(\frac{1}{\bar{t}_1 - t_{n-1}} \int_{t_{n-1}}^{\bar{t}_1} \bar{\alpha}_j(\tau) d\tau \right) - \int_{t_{n-1}}^{\bar{t}_1} L_j(\bar{\alpha}_j(\tau)) d\tau \\
&\leq \int_{t_{n-1}}^{\bar{t}_1} L_j(\bar{\alpha}_j(\tau)) d\tau - \int_{t_{n-1}}^{\bar{t}_1} L_j(\bar{\alpha}_j(\tau)) d\tau = 0.
\end{aligned}$$

The two other cost terms can be treated in a similar way. Using that

$$\begin{aligned}
\mathbb{I}[w^{n-1}](\hat{X}(t_{n-1})) - u(t_{n-1}, \bar{X}(t_{n-1})) &\leq \mathbb{I}[w^{n-1}](\hat{X}(t_{n-1})) \\
&\quad - \mathbb{I}[u(t_{n-1}, \cdot) + (n-1)C\Delta x](\bar{X}(t_{n-1})) + \mathbb{I}[u(t_{n-1}, \cdot)](\bar{X}(t_{n-1})) \\
&\quad + (n-1)C\Delta x - u(t_{n-1}, \bar{X}(t_{n-1})) \leq nC\Delta x
\end{aligned}$$

for the basic properties of the interpolation operator and (50), we can claim that

$$w^n(x) - u(t_n, x) \leq nC\Delta x,$$

and this concludes the proof. ∎

REFERENCES

[1] Y. Achdou, F. Camilli, A. Cutrì, and N. Tchou, *Hamilton–Jacobi equations constrained on networks*, NoDEA Nonlinear Differential Equations Appl., 20 (2013), pp. 413–445.

[2] B. Andreianov, K. H. Karlsen, and N. H. Risebro, *A theory of L^1-dissipative solvers for scalar conservation laws with discontinuous flux*, Arch. Ration. Mech. Anal., 201 (2011), pp. 27–86.

[3] M. Bardi and I. C. Dolcetta, *Optimal Control and Viscosity Solutions of Hamilton-Jacobi-Bellman Equations*, Birkhäuser Boston, Boston, 1996.

[4] G. Barles, A. Briani, and E. Chasseigne, *A Bellman approach for two-domains optimal control problems in $\mathbb{R}^N$*, ESAIM Control Optim. Calc. Var., 19 (2013), pp. 710–739.

[5] G. Barles, A. Briani, E. Chasseigne, and C. Imbert, *Flux-limited and classical viscosity solutions for regional control problems*, ESAIM Control Optim. Calc. Var., 24 (2018), pp. 1881–1906.

[6] G. Barles and P. E. Souganidis, *Convergence of approximation schemes for fully nonlinear second order equations*, Asymptot. Anal., 4 (1991), pp. 271–283.

[7] F. Camilli, E. Carlini, and C. Marchi, *A flame propagation model on a network with application to a blocking problem*, Discrete Contin. Dynam. Syst. Ser. S, 11 (2018), pp. 825–843.

[8] F. Camilli, A. Festa, and D. Schieborn, *An approximation scheme for a Hamilton–Jacobi equation defined on a network*, Appl. Numer. Math., 73 (2013), pp. 33–47.

[9] F. Camilli, A. Festa, and S. Tozza, *A discrete Hughes' model for pedestrian flow on graphs*, Netw. Heterog. Media, 12 (2017), pp. 93–112.

[10] F. Camilli and C. Marchi, *A comparison among various notions of viscosity solution for Hamilton–Jacobi equations on networks*, J. Math. Anal. Appl., 407 (2013), pp. 112–118.

[11] G. Costeseque, J.-P. Lebacque, and R. Monneau, *A convergent scheme for Hamilton–Jacobi equations on a junction: Application to traffic*, Numer. Math., 129 (2015), pp. 405–447.

[12] M. G. Crandall, and P.-L. Lions, *Two approximations of solutions of Hamilton-Jacobi equations*, Math. Comp., 43 (1984), pp 1–19.

[13] E. Cristiani, B. Piccoli, and A. Tosin, *Multiscale modeling of granular flows with application to crowd dynamics*, Multiscale Model. Simul., 9 (2011), pp. 155–182, https://doi.org/10.1137/100797515.

[14] F. Charles, B. Després, and M. Mehrenberger, *Enhanced convergence estimates for semi-Lagrangian schemes: Application to the Vlasov–Poisson equation*, SIAM J. Numer. Anal., 51 (2013), pp. 840–863, https://doi.org/10.1137/110851511.

[15] M. Falcone and R. Ferretti, *Semi-Lagrangian Approximation Schemes for Linear and Hamilton–Jacobi Equations*, SIAM, 2014, https://doi.org/10.1137/1.9781611973051.

[16] N. Forcadel and W. Salazar, *Homogenization of second order discrete model and application to traffic flow*, Differential Integral Equations, 28 (2015), pp. 1039–1068.

[17] M. Garavello, K. Han, and B. Piccoli, *Models for Vehicular Traffic on Networks*, AIMS Ser. Appl. Math. 9, American Institute of Mathematical Sciences (AIMS), Springfield, MO, 2016.

[18] M. Garavello, R. Natalini, B. Piccoli, and A. Terracina, *Conservation laws with discontinuous flux*, Netw. Heterog. Media, 2 (2007), pp. 159–179.

[19] M. Garavello and B. Piccoli, *Traffic Flow on Networks*, Vol. 1, American Institute of Mathematical Sciences (AIMS), Springfield, MO, 2006.

[20] J. Guerand and M. Koumaiha, *Error estimates for finite difference schemes associated with Hamilton-Jacobi equations on a junction*, Numer. Math., 142 (2019), pp 525–575.

[21] C. Imbert and R. Monneau, *Flux-limited solutions for quasi-convex Hamilton-Jacobi equations on networks*, Ann. Sci. Éc. Norm. Supér. (4), 50 (2017), pp. 357–448.

[22] C. Imbert, R. Monneau, and H. Zidani, *A Hamilton-Jacobi approach to junction problems and application to traffic flows*, ESAIM Control Optim. Calc. Var., 19 (2013), pp. 129–166.

[23] J. P. Lebacque, *The Godunov scheme and what it means for first order traffic flow models*, in Proceedings of the 13th International Symposium on Transportation and Traffic Theory, J. B. Lesort, ed., Elsevier, 1996, pp. 647–677, https://trid.trb.org/view/481282.

[24] P.-L. Lions and P. Souganidis, *Well-posedness for multi-dimensional junction problems with Kirchoff-type conditions*, Atti Accad. Naz. Lincei Rend. Lincei Mat. Appl., 28 (2017), pp. 807–816.

[25] G. F. Newell, *A simplified theory of kinematic waves in highway traffic, Part* I: *General theory*, Transp. Res. Part B Methodol., 27 (1993), pp. 281–287.

[26] A. Quarteroni, R. Sacco, and F. Saleri, *Numerical Mathematics*, Texts Appl. Math. 37, Springer-Verlag, Berlin, Heidelberg, 2007.

[27] D. Schieborn and F. Camilli, *Viscosity solutions of eikonal equations on topological networks*, Calc. Var. Partial Differential Equations, 46 (2013), pp. 671–686.

[28] M. Treiber and A. Kesting, *Traffic Flow Dynamics: Data, Models and Simulation*, Springer-Verlag, Berlin, Heidelberg, 2013.