MATHEMATICS OF COMPUTATION
https://doi.org/10.1090/mcom/3437
Article electronically published on September 10, 2019

STOCHASTIC AND VARIATIONAL APPROACH TO FINITE DIFFERENCE APPROXIMATION OF HAMILTON-JACOBI EQUATIONS

KOHEI SOGA

Abstract. Previously, the author presented a stochastic and variational approach to the Lax-Friedrichs finite difference scheme applied to hyperbolic scalar conservation laws and the corresponding Hamilton-Jacobi equations with convex and superlinear Hamiltonians in the one-dimensional periodic setting, showing new results on the stability and convergence of the scheme [Soga, Math. Comp. **84** (2015), 629–651]. In the current paper, we extend these results to the higher dimensional setting. Our framework with a deterministic scheme provides approximation of viscosity solutions of Hamilton-Jacobi equations, their spatial derivatives and the backward characteristic curves at the same time, within an arbitrary time interval. The proof is based on stochastic calculus of variations with random walks, a priori boundedness of minimizers of the variational problems that verifies a CFL type stability condition, and the law of large numbers for random walks under the hyperbolic scaling limit. Convergence of approximation and the rate of convergence are obtained in terms of probability theory. The idea is reminiscent of the stochastic and variational approach to the vanishing viscosity method introduced in [Fleming, J. Differ. Eqs **5** (1969) 515–530].

1. Introduction

We consider finite difference approximation to viscosity solutions of initial value problems for the Hamilton-Jacobi equations

$$
\begin{cases} v_t(x,t)+H(x,t,v_x(x,t))=h \quad \text{in } \mathbb{R}^d\times(0,T], \\ v(x,0)=v^0(x) \quad \text{on } \mathbb{R}^d, \quad v^0\in Lip_r(\mathbb{R}^d),\ |v^0|\le R, \end{cases} \tag{1.1}
$$

where $v_t:=\frac{\partial v}{\partial t}$, $v_x:=(\frac{\partial v}{\partial x^1},\dots,\frac{\partial v}{\partial x^d})$, $d\ge 1$, h is a given constant, and $Lip_r(\mathbb{R}^d)$ denotes the family of Lipschitz functions $\mathbb{R}^d\to\mathbb{R}$ with Lipschitz constants bounded by $r>0$. Initial data is assumed to be bounded by $R>0$. We arbitrarily fix the constants T,r,R. The function H is assumed to satisfy the following (H1)–(H5):

(H1) $H(x,t,p):\mathbb{R}^d\times\mathbb{R}\times\mathbb{R}^d\to\mathbb{R}$, C^2.
(H2) $H_{pp}(x,t,p)$ is positive definite in $\mathbb{R}^d\times\mathbb{R}\times\mathbb{R}^d$.

Received by the editor March 22, 2018, and, in revised form, February 12, 2019, March 16, 2019, and July 8, 2019.

2010 *Mathematics Subject Classification.* Primary 65M06, 35F21, 49L25, 60G50.

Key words and phrases. Finite difference scheme, Hamilton-Jacobi equation, viscosity solution, calculus of variations, random walk, law of large numbers.

The main part of this work was done, when the author belonged to Unité de mathématiques pures et appliquées, CNRS UMR 5669 & École Normale Supérieure de Lyon, being supported by ANR-12-BS01-0020 WKBHJ as a researcher for academic year 2014–2015, hosted by Albert Fathi. The author was partially supported by JSPS Grant-in-aid for Young Scientists (B) #15K21369.

(H3) H is uniformly superlinear with respect to p, i.e., for each $a \geq 0$ there exists $b_1(a) \in \mathbb{R}$ such that $H(x,t,p) \geq a\|p\| + b_1(a)$ in $\mathbb{R}^d \times \mathbb{R} \times \mathbb{R}^d$.

(H4) $H, H_{x^i}, H_{p^i}, H_{x^i x^j}, H_{x^i p^j}, H_{p^i p^j}$ are uniformly bounded on $\mathbb{R}^d \times \mathbb{R} \times K$ for each compact set $K \subset \mathbb{R}^d$ for $i, j = 1, \ldots, d$.

(H5) For the Legendre transform of $H(x,t,\cdot)$, denoted by L, there exists $\alpha > 0$ such that $|L_{x^j}| \leq \alpha(1 + |L|)$ in $\mathbb{R}^d \times \mathbb{R} \times \mathbb{R}^d$ for $j = 1, \ldots, d$.

Here, $\|x\| := \sqrt{\sum_{1 \leq j \leq d} (x^j)^2}$ and $x \cdot y := \sum_{1 \leq j \leq d} x^j y^j$ for $x, y \in \mathbb{R}^d$. Note that, due to (H1)–(H3), the function $L(x,t,\xi) : \mathbb{R}^d \times \mathbb{R} \times \mathbb{R}^d \to \mathbb{R}$ in (H5) is well-defined and is given by

$$L(x,t,\xi) = \sup_{p \in \mathbb{R}^d} \{p \cdot \xi - H(x,t,p)\}.$$

We will see properties of L in Section 3.

These problems arise in many fields such as theories of optimal control, dynamical systems, and so on. Our motivation mainly comes from weak KAM theory. Weak KAM theory connects viscosity solutions of Hamilton-Jacobi equations and Hamiltonian/Lagrangian dynamics, where the central objects are viscosity solutions, their spatial derivatives, and the characteristic curves. In numerical analysis of weak KAM theory, it is important to develop a method that is able to approximate all of these objects at the same time. See [6], [9], [10], [11] for recent developments in the numerical analysis of weak KAM theory based on such a method in one-dimensional problems with Tonelli Hamiltonians, i.e., H is periodic in (x,t) with (H1)–(H3).

In the one-dimensional case, (1.1) is equivalent to the scalar conservation law

$$\begin{cases} u_t + H(x,t,u)_x = 0 & \text{in } \mathbb{R} \times (0,T], \\ u(x,0) = u^0(x) & \text{on } \mathbb{R}, \quad u^0 \in L^\infty(\mathbb{R}). \end{cases} \tag{1.2}$$

If $u^0 = v^0_x$, the viscosity solution v or entropy solution u is derived from the other and they satisfy the relation $u = v_x$. Therefore, approximation of u implies approximation of v, but the other way around is not necessarily true. In the pioneering work [7] on finite difference approximation of (1.2) with a wide class of functions H, stability and convergence of the Lax-Friedrichs finite difference scheme are proved within a restricted time interval based on a functional analytic approach. The restriction depends on the growth rate of H for $|p| \to \infty$. In [9], the author announced a stochastic and variational approach to the Lax-Friedrichs scheme, where the discretized equation of (1.2) with the Lax-Friedrichs scheme is converted to a discretized equation of (1.1), and stability and convergence are proved within an arbitrary time interval. Furthermore, this framework guarantees approximation of all of entropy solutions, viscosity solutions, and their characteristic curves at the same time. Application of these results to weak KAM theory is found in [10], [11]. The key point of the stochastic and variational approach is to characterize the so-called *numerical viscosity* of the discretized equation by space-time inhomogeneous random walks. Then, we obtain a stochastic Lax-Oleinik type operator with random walks for the discretized Hamilton-Jacobi equation. Convergence of the stochastic Lax-Oleinik type operator to the exact one for (1.1) is proved through the law of large numbers for random walks. The idea is very much reminiscent of the stochastic and variational approach to the vanishing viscosity method with the Brownian motions developed in [5].

The purpose of this paper is to generalize the results in [9] to the higher-dimensional problems. Here, we do not restrict ourselves to the case with Tonelli Hamiltonians, also because such a non-compact problem arises in different contexts. We introduce a simple (deterministic) scheme that is a direct generalization of the one-dimensional Lax-Friedrichs scheme on a staggered grid. Then, we formulate a stochastic Lax-Oleinik type operator with space-time inhomogeneous random walks on a grid in $\mathbb{R}^d$. Convergence of approximation is proved through the law of large numbers for random walks. *We will have convergent approximation of the viscosity solutions, their spatial derivatives, and the characteristic curves at the same time within an arbitrary time interval.* In the case of $d \geq 2$, there is no equivalence between Hamilton-Jacobi equations and scaler conservation laws. Therefore, we discretize the Hamilton-Jacobi equation so that both difference viscosity solutions and their discrete spatial derivatives can be controllable. Our approach yields new results on stability of the scheme and its convergence to the exact viscosity solution of (1.1) with the rate $O(\sqrt{\Delta x})$, from which convergent approximation of its derivatives and characteristic curves is derived. Here, $\Delta x, \Delta t > 0$ are spatial, temporal discretization parameters, respectively. In our proof, this rate of convergence comes from the hyperbolic scaling limit (i.e., $\Delta x, \Delta t \to 0$ with $\Delta t/\Delta x = O(1)$) of random walks. The result of the present paper would be a basic tool for numerical analysis of Hamilton-Jacobi equations including weak KAM theory in the higher-dimensional setting.

There are many results on finite difference approximation of the viscosity solution to (1.1). We refer to the pioneering work [3] and its generalization [12], where convergence of a class of finite difference schemes with the rate $O(\sqrt{\Delta x})$ is proved in an abstract setting, under the assumption that schemes are monotone (verification of this assumption is necessary for each scheme). In the recent work [1], a numerical scheme for Hamilton-Jacobi equations with separable Hamiltonians, i.e., functions H of the form $H(x,t,p) = f(p) + g(x,t)$, is developed based on a direct discretization of Lax-Oleinik type operators. As far as the author knows, in the literature, the functions H with (H1)–(H5) are not covered; furthermore, convergence of approximation is proved only for viscosity solutions. The main difficulty of finite difference approximation in (1.1) with our requirement is to verify a priori boundedness of the discrete derivatives of difference solutions (this yields monotonicity of the scheme). Verification of the a priori boundedness is even harder in one-dimensional cases if $H(x,t,p)$ is convex and superlinear with respect to p and is not of a separable form like $H(x,t,p) = f(p)+g(x,t)$. This difficulty is overcome by means of our stochastic and variational framework, where we may effectively use a priori boundedness of minimizers of the stochastic Lax-Oleinik type operator.

Although we do not demonstrate any numerical experiment in this paper, we give a short comment on possible implementation of our scheme. Note that our discrete problem (2.1) is an explicit purely deterministic recurrence formula; it is solvable by the four arithmetic operations WITHOUT any stochastic and variational treatment (stochastic and variational argument is necessary for theoretical analysis of (2.1)); its computational cost is more or less the same as the standard Lax-Friedrichs scheme applied to a scaler conservation law in $\mathbb{R}^d$; only in approximation of characteristic curves, one needs to deal with the random walk derived from the minimizing control explicitly given in (2) of Theorem 2.1; due to a finite speed of propagation, (2.1) is solvable locally in space (see Section 2). Some numerical experiments related to (2.1) with $d = 1$ and the periodic boundary condition are done in [6]. In our framework, hyperbolic scaling $\Delta t/\Delta x = O(1)$ is crucial. It is interesting to note that the scheme developed in [1] accepts $\Delta x, \Delta t \to 0$ with $\Delta x/\Delta t \to 0$, say $\Delta t = \sqrt{\Delta x}$, which is not hyperbolic scaling nor diffusive scaling. Such scaling would imply that, if one only needs approximation of viscosity solutions, the scheme in [1] could be quicker than ours in actual computation. Nevertheless, this might not be true, because schemes derived from direct discretization of Lax-Oleinik type operators require an algorithm of taking "infimum" on a computer, which is not the case with our scheme. It is an open problem to compare these schemes in terms of the CPU time.

2. Scheme and result

Let $\Delta x > 0$ and $\Delta t > 0$ be discretization parameters for space and time, respectively. Set $\Delta := (\Delta x, \Delta t)$, $|\Delta| := \max\{\Delta x, \Delta t\}$, $G_{even} := \{m\Delta x \,|\, m = (m^1, \dots, m^d) \in \mathbb{Z}^d, m^1 + \cdots + m^d = even\}$, $G_{odd} := \{m\Delta x \,|\, m \in \mathbb{Z}^d, m^1 + \cdots + m^d = odd\}$, $t_k := k\Delta t$ for $k \in \mathbb{N} \cup \{0\}$ and

$$\mathcal{G} := \bigcup_{k \geq 0} \Big\{ (G_{even} \times \{t_{2k}\}) \cup (G_{odd} \times \{t_{2k+1}\}) \Big\},$$
$$\tilde{\mathcal{G}} := \bigcup_{k \geq 0} \Big\{ (G_{odd} \times \{t_{2k}\}) \cup (G_{even} \times \{t_{2k+1}\}) \Big\}.$$

Figure 1 shows the two-dimensional G_{even} by the symbol $\circ$ and G_{odd} by $\bullet$. Note that $\mathcal{G} \cup \tilde{\mathcal{G}} = (\Delta x\mathbb{Z})^d \times (\Delta t \mathbb{Z}_{\geq 0})$, where $\Delta x\mathbb{Z} := \{i\Delta x \,|\, i \in \mathbb{Z}\}$. Each point of $(\Delta x\mathbb{Z})^d \times (\Delta t\mathbb{Z}_{\geq 0})$ is denoted by $(x_m, t_k) = (x^1_{m^1}, \dots, x^d_{m^d}, t_k)$ with $m = (m^1, \dots, m^d) \in \mathbb{Z}^d$. We sometimes use the notation $(x_m, t_k), (x_{m+\mathbf{1}}, t_{k+1})$ to indicate points of $\mathcal{G}$

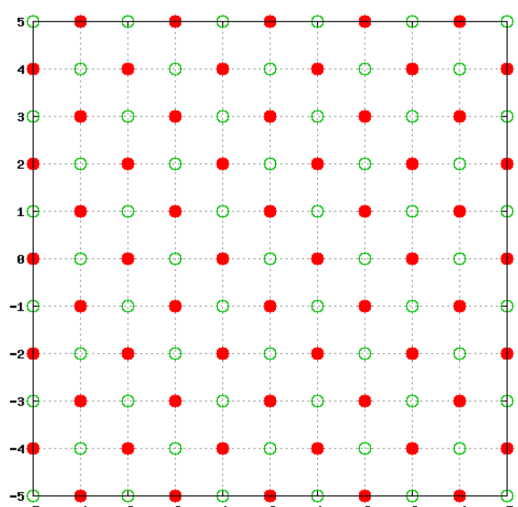

Figure 1

and $(x_{m+\mathbf{1}},t_k),(x_m,t_{k+1})$ to indicate points of $\tilde{\mathcal{G}}$ with $\mathbf{1}:=(1,0,\dots,0)\in\mathbb{Z}^d$. For $(x,t)\in\mathcal{G}\cup\tilde{\mathcal{G}}$, the notation $m(x)$, $k(t)$ denotes the index of x, t, respectively. For $t\geq 0$, $k(t)$ denotes the integer such that $t\in[k(t)\Delta t,k(t)\Delta t+\Delta t)$. Let $\{e_1,\dots,e_d\}$ be the standard basis of $\mathbb{R}^d$. Set

$$B:=\{\pm e_1,\dots,\pm e_d\}.$$

Let $v=v^k_{m+\mathbf{1}}$ denote the function $\tilde{\mathcal{G}}\ni(x_{m+\mathbf{1}},t_k)\mapsto v^k_{m+\mathbf{1}}\in\mathbb{R}$. Introduce the spatial difference derivatives of v that are defined at each point $(x_m,t_k)\in\mathcal{G}$ as

$$(D_{x^j}v)^k_m:=\frac{v^k_{m+e_j}-v^k_{m-e_j}}{2\Delta x},\quad (D_xv)^k_m:=\big((D_{x^1}v)^k_m,\dots,(D_{x^d}v)^k_m\big).$$

Introduce the temporal difference derivative of v as

$$(D_tv)^{k+1}_m:=\left(v^{k+1}_m-\frac{1}{2d}\sum_{\omega\in B}v^k_{m+\omega}\right)\frac{1}{\Delta t}.$$

Discretize (1.1) as

$$\text{(2.1)}\qquad \begin{cases} (D_tv)^{k+1}_m+H(x_m,t_k,(D_xv)^k_m)=h & \text{in } \tilde{\mathcal{G}}|_{0\leq k\leq k(T)},\\ v^0_{m+\mathbf{1}} \quad \text{is given on } G_{odd}, \end{cases}$$

where the initial data $v^0_{m+\mathbf{1}}$ is given as

$$v^0_{m+\mathbf{1}}:=v^0(x_{m+\mathbf{1}}).$$

Note that, in (2.1), v^{k+1}_m is unknown determined by $v^k_{m+\mathbf{1}}$ as a recursion. If $d=1$, the difference equation in (2.1) is exactly the same as the scheme studied in [9].

Evolution of (2.1) can be intuitively seen as Figure 2: the value v^{k+1}_m is determined by the values of the grid points • contained in the pyramid in the figure, where the pyramid grows up to $k=0$; the aspect ratio of the pyramid is determined by $\lambda:=\Delta t/\Delta x$, which is of $O(1)$ for hyperbolic scaling limit, i.e., $\Delta x,\Delta t\to 0$ with $0<\lambda_0\leq\lambda<\lambda_1$ (a finite speed of propagation). The main discovery in this paper is that the contribution of the value at each grid point • within the pyramid to the value v^{k+1}_m can be characterized by a probability measure that is given by a random walk wandering within the pyramid; the probability measure concentrates on a curve as $\Delta x,\Delta t\to 0$ under hyperbolic scaling (the law of large numbers); the curve is a characteristic curve of (1.1).

In order to formulate the above statement, we introduce space-time inhomogeneous random walks on $\tilde{\mathcal{G}}$, which correspond to characteristic curves of (1.1). For each point $(x_n,t_{l+1})\in\tilde{\mathcal{G}}$, we consider the backward random walks γ within $[t_{l'},t_{l+1}]$ which start from x_n at t_{l+1} and move by $\omega\Delta x$, $\omega\in B$, in each backward time step Δt:

$$\gamma=\{\gamma^k\}_{k=l',\dots,l+1},\quad \gamma^{l+1}=x_n,\quad \gamma^k=\gamma^{k+1}+\omega\Delta x.$$

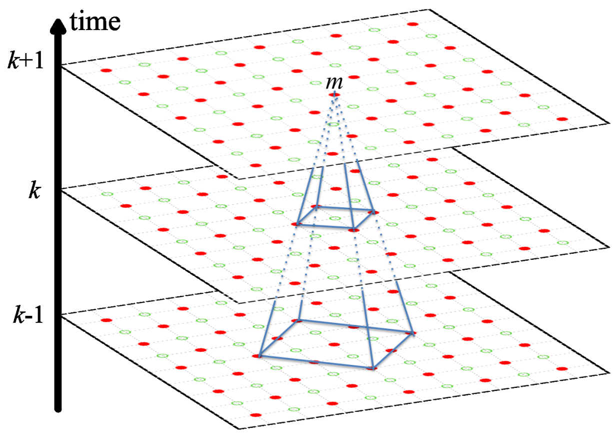

FIGURE 2

More precisely, we set the following objects for each $(x_n, t_{l+1}) \in \tilde{\mathcal{G}}$ and $l' \le l$:

$X_n^{l+1,k} := \{x_{m+\mathbf{1}} \mid\ (x_{m+\mathbf{1}}, t_k) \in \tilde{\mathcal{G}},\ \max_{1\le j\le d} |x^j_{m^j+\mathbf{1}^j} - x^j_{n^j}| \le (l+1-k)\Delta x\}$

(the set of all reachable points at time k, or the cross-section of the pyramid at time k),

$G_n^{l+1,l'} := \bigcup_{l'\le k\le l+1} \big(X_n^{l+1,k} \times \{t_k\}\big) \subset \tilde{\mathcal{G}}$

(the set of points within the pyramid for $l' \le k \le l+1$),

$\xi : G_n^{l+1,l'+1} \ni (x_m, t_{k+1}) \mapsto \xi_m^{k+1} \in [-(d\lambda)^{-1}, (d\lambda)^{-1}]^d, \quad \lambda := \Delta t/\Delta x,$

$\rho : G_n^{l+1,l'+1} \times B \ni (x_m, t_{k+1}; \omega) \mapsto \rho_m^{k+1}(\omega) := \left\{\frac{1}{2d} - \frac{\lambda}{2}(\omega\cdot\xi_m^{k+1})\right\} \in [0,1],$

$\gamma : \{l', l'+1, \dots, l+1\} \ni k \mapsto \gamma^k \in X_n^{l+1,k}$ with $\gamma^{l+1} = x_n$,

$\gamma^k = \gamma^{k+1} + \omega\Delta x,\ \omega \in B,$

$\Omega_n^{l+1,l'}$: the family of the above γ,

where ξ and ρ are not defined at l'. We may regard $\rho_m^{k+1}(\omega)$, $\omega \in B$, as the transition probability from (x_m, t_{k+1}) to $(x_m + \omega\Delta x, t_k)$, because we have

$$\sum_{\omega\in B} \rho_m^{k+1}(\omega) = \sum_{i=1}^{d} (\rho_m^{k+1}(e_i) + \rho_m^{k+1}(-e_i)) = 1.$$

We control transition of random walks by ξ, which is a kind of velocity field on $G_n^{l+1,l'}$. We define the density of each path $\gamma \in \Omega_n^{l+1,l'}$ as

$$\mu_n^{l+1,l'}(\gamma) := \prod_{l'\le k\le l} \rho^{k+1}_{m(\gamma^{k+1})}(\omega^{k+1}),$$

where $\omega^{k+1} := (\gamma^k - \gamma^{k+1})/\Delta x$. For each ξ, the density $\mu_n^{l+1,l'}(\cdot) = \mu_n^{l+1,l'}(\cdot;\xi)$ yields a probability measure of $\Omega_n^{l+1,l'}$, i.e.,

$$prob(A) = \sum_{\gamma \in A} \mu_n^{l+1,l'}(\gamma;\xi) \quad \text{for } A \subset \Omega_n^{l+1,l'}.$$

The expectation with respect to this probability measure is denoted by $E_{\mu_n^{l+1,l'}(\cdot;\xi)}[\cdot]$, i.e., for a random variable $f : \Omega_n^{l+1,l'} \to \mathbb{R}$,

$$E_{\mu_n^{l+1,l'}(\cdot;\xi)}[f(\gamma)] := \sum_{\gamma \in \Omega_n^{l+1,l'}} \mu_n^{l+1,l'}(\gamma;\xi) f(\gamma).$$

We remark that, since our transition probabilities are space-time inhomogeneous, the well-known law of large numbers and central limit theorem for random walks do not always hold. The author investigated the asymptotics of the probability measure of $\Omega_n^{l+1,l'}$ as $\Delta \to 0$ under hyperbolic scaling in the one-dimensional case [8]. We extend this investigation to the current multidimensional case in Section 3.

Let $v(x,t)$ be the viscosity solution of (1.1) (v uniquely exists as a Lipschitz function, which becomes locally semiconcave with a linear modulus in x for $t > 0$). It is well known that v satisfies for each $x \in \mathbb{R}^d$ and $t > 0$,

$$v(x,t) = \inf_{\gamma \in AC,\ \gamma(t)=x} \left\{ \int_0^t L(\gamma(s), s, \gamma'(s))ds + v^0(\gamma(0)) \right\} + ht,$$

where AC is the family of absolutely continuous curves $\gamma : [0,t] \to \mathbb{R}^d$ ($v^0 \mapsto v(\cdot,t)$ is called a Lax-Oleinik type operator for (1.1)). Due to Tonelli's theory, there exists a minimizing curve γ^* for $v(x,t)$, which is a C^2-solution of the Euler-Lagrange equation generated by L. We say that a point $(x,t) \in \mathbb{R}^d \times \mathbb{R}$ is regular if $v_x(x,t)$ exists. The viscosity solution v of (1.1) is Lipschitz and hence it is differentiable a.e. If (x,t) is regular, the minimizing curve γ^* for $v(x,t)$ is unique. Let (x,t) be regular and let γ^* be the minimizing curve for $v(x,t)$. Then, we have

$$v_x(x,t) = \int_0^t L_x(\gamma^*(s), s, \gamma^{*\prime}(s))ds + v_x^0(\gamma^*(0)),$$

where v^0 is supposed to be locally semiconcave; otherwise, $v_x^0(\gamma^*(0))$ must be replaced by $L_\xi(\gamma^*(0), 0, \gamma^{*\prime}(0))$. We refer to [2], [4] for more on viscosity solutions and calculus of variations.

Now we state the main results. The first theorem shows a stochastic and variational representation of the difference solution to (2.1) (a Lax-Oleinik type operator for (2.1)).

Theorem 2.1. *There exists $\lambda_1 > 0$ (depending on T, r, and R, but independent of Δ) for which the following statements hold for any small $\Delta = (\Delta x, \Delta t)$ with $\lambda = \Delta t/\Delta x < \lambda_1$:*

(1) *For each n and l with $0 < l+1 \le k(T)$ such that $(x_n, t_{l+1}) \in \tilde{\mathcal{G}}$, the expectation*

$$E_n^{l+1}(\xi) := E_{\mu_n^{l+1,0}(\cdot;\xi)}\Big[\sum_{0<k\le l+1} L(\gamma^k, t_{k-1}, \xi^k_{m(\gamma^k)})\Delta t + v^0_{m(\gamma^0)} \Big] + ht_{l+1}$$

with respect to the probability measure of $\Omega_n^{l+1,0}$ *has the infimum within all controls* $\xi : G_n^{l+1,1} \to [-(d\lambda)^{-1}, (d\lambda^{-1})]^d$. *There exists the unique minimizing control* ξ^* *to attain the infimum, which satisfies* $|\xi^{*j}| \le (d\lambda_1)^{-1} < (d\lambda)^{-1}$ *for all* $1 \le j \le d$.

(2) *Define the function* v *on* $\tilde{\mathcal{G}}|_{0\le k\le k(T)}$ *as* $v(x_m, t_{k+1}) := \inf_\xi E_m^{k+1}(\xi)$ *and* $v(x_{m+\mathbf{1}}, t_0) := v^0_{m+\mathbf{1}}$. *Then, the minimizing control* ξ^* *for* $\inf_\xi E_n^{l+1}(\xi)$ *satisfies*

$$\xi^{*k+1}_{\ m} = H_p(x_m, t_k, (D_x v)^k_m) \ \ (\Leftrightarrow \ (D_x v)^k_m = L_\xi(x_m, t_k, \xi^{*k+1}_{\ m})).$$

In particular, $(D_x v)^k_m$ *is uniformly bounded on* $\tilde{\mathcal{G}}|_{0\le k\le k(T)}$ *independently from* Δ *(a CFL-type condition).*

(3) $v = v^k_{m+\mathbf{1}}$ *defined in* (2) *is the solution of* (2.1).

(4) *For each* $\omega \in B$, *let* $\xi^*(\omega) : G^{l+1,1}_{n+\mathbf{1}+\omega} \to [-(d\lambda)^{-1}, (d\lambda^{-1})]^d$ *be the minimizing control for* $\inf_\xi E^{l+1}_{n+\mathbf{1}+\omega}(\xi)$. *Let* $\gamma \in \Omega^{l+1,0}_{n+\mathbf{1}+\omega}$ *be the minimizing random walk generated by* $\xi^*(\omega)$. *Then, we have for* $j = 1, \dots, d$,

$$\begin{aligned}(D_{x^j} v)^{l+1}_{n+\mathbf{1}} \le E_{\mu^{l+1,0}_{n+\mathbf{1}-e_j}(\cdot;\xi^*(-e_j))}\Big[&\sum_{0<k\le l+1} L_{x^i}(\gamma^k, t_{k-1}, \xi^{*k}_{m(\gamma^k)}(-e_j))\Delta t \\ &+ (D_{x^j} v)^0_{m(\gamma^0)+e_j}\Big] + \theta\Delta x,\\ (D_{x^j} v)^{l+1}_{n+\mathbf{1}} \ge E_{\mu^{l+1,0}_{n+\mathbf{1}+e_j}(\cdot;\xi^*(e_j))}\Big[&\sum_{0<k\le l+1} L_{x^i}(\gamma^k, t_{k-1}, \xi^{*k}_{m(\gamma^k)}(e_j))\Delta t \\ &+ (D_{x^j} v)^0_{m(\gamma^0)-e_j}\Big] - \theta\Delta x,\end{aligned}$$

where $\theta > 0$ *is independent of* Δ.

We define the interpolation $v_\Delta(x,t) : \mathbb{R}^d \times [0, t_{k(T)+1}) \to \mathbb{R}$ of the difference solution $v^k_{m+\mathbf{1}}$ to (2.1), $u^j_\Delta(x,t) : \mathbb{R}^d \times [0, t_{k(T)+1}) \to \mathbb{R}^d$ of its difference partial derivative $(D_{x^j} v)^k_m$, and $\gamma_\Delta(s) : [0,t] \to \mathbb{R}^d$ of each sample path $\gamma \in \Omega^{l+1,0}_n$ with $t \in [t_{l+1}, t_{l+2})$: Let $\tilde{B}$ be the set of all $\tilde\omega := \{\tilde\omega_1, \dots, \tilde\omega_d\}$ such that $\tilde\omega_i = e_i$ or $-e_i$; Introduce

$$\Theta(x) := \bigcup_{\tilde\omega\in\tilde B} \Big\{ x + \Delta x \sum_{i=1}^d \theta_i \tilde\omega_i \,|\, \theta_i \ge 0,\ \sum_{i=1}^d \theta_i < 1 \text{ if } \tilde\omega_d = e_d,\ \sum_{i=1}^d \theta_i \le 1 \text{ if } \tilde\omega_d = -e_d \Big\}$$

(if $d = 2$, $\Theta(x)$ is the square whose vertexes are $x \pm \Delta x e_1, x \pm \Delta x e_2$); Define

$$\begin{aligned} v_\Delta(x,t) &:= v^k_{m+\mathbf{1}} && \text{for } x \in \Theta(x_{m+\mathbf{1}}),\ t \in [t_k, t_{k+1}),\\ u^j_\Delta(x,t) &:= (D_{x^j} v)^k_m && \text{for } x \in \Theta(x_m),\ t \in [t_k, t_{k+1}),\\ \gamma_\Delta(s) &:= \begin{cases} x_n & \text{for } s \in [t_{l+1}, t],\\ \gamma^k + \frac{\gamma^{k+1}-\gamma^k}{\Delta t}(s - t_k) & \text{for } s \in [t_k, t_{k+1}]. \end{cases} \end{aligned}$$

Note that the probability measure of a random walk induces a probability measure

$$\sum_{\gamma\in\Omega^{l+1,0}_n} \mu^{l+1,0}_n(\gamma;\xi)\delta_{\gamma_\Delta}(\cdot)$$

on $C^0([0,t];\mathbb{R}^d)$, where δ_{γ_Δ} is the δ-measure supported by γ_Δ. The next theorem shows convergence of approximation.

Theorem 2.2. *Take the limit $\Delta \to 0$ under hyperbolic scaling, namely, $\Delta \to 0$ with $0 < \lambda_0 \le \lambda = \Delta t/\Delta x < \lambda_1$, where λ_1 is the one mentioned in Theorem 2.1 and λ_0 is a constant. Then, the following statements hold:*

(1) *Let v be the viscosity solution of (1.1) and let v_Δ be the interpolation of the solution of (2.1). Then, there exists $\beta > 0$ independent of Δ and v^0 for which we have*

$$\sup_{\mathbb{R}^d\times[0,T]} |v_\Delta - v| \le \beta\sqrt{\Delta x} \quad \text{as } \Delta \to 0.$$

(2) *Let $(x,t) \in \mathbb{R}^d \times (0,T]$ be regular and $\gamma^* : [0,t] \to \mathbb{R}^d$ be the minimizing curve for $v(x,t)$. Let $(x_n, t_{l+1}) \in \tilde{\mathcal{G}}|_{0\le k\le k(T)}$ be a point such that $t \in [t_{l+1}, t_{l+2})$ and $x \in \Theta(x_n)$. Let γ_Δ be the interpolation of the random walk $\gamma \in \Omega_n^{l+1,0}$ generated by the minimizing control for v_n^{l+1}. Then, we have for any $\varepsilon > 0$,*

$$Prob(\{\gamma \in \Omega_n^{l+1,0} \mid \|\gamma_\Delta - \gamma^*\|_{C^0([0,t])} \le \varepsilon\}) \to 1 \quad \text{as } \Delta \to 0.$$

In other words, the probability measure of minimizing random walk converges weakly to the δ-measure supported by γ^*. In particular, the average of γ_Δ converges uniformly to γ^* on $[0,t]$.

(3) *Suppose, in addition, that v^0 is locally semiconcave with a linear modulus. Let u^j_Δ be the interpolation of the difference partial derivatives of the solution to (2.1). Then, for each regular point $(x,t) \in \mathbb{R}^d \times (0,T]$, we have*

$$u^j_\Delta(x,t) \to v_{x^j}(x,t) \quad \text{as } \Delta \to 0 \quad (j = 1,\dots,d).$$

In particular, u^j_Δ converges to v_{x^j} pointwise a.e., and hence, for each compact set $K \subset \mathbb{R}^d$, $u^j_\Delta(\cdot,t)$ converges to $v_{x^j}(\cdot,t)$ in $L^1(K)$; u^j_Δ converges uniformly to v_{x^j} on $(K \times [0,T]) \setminus \mathcal{O}$, where $\mathcal{O}$ is a neighborhood of the set of points of singularity of v_x with arbitrarily small measure.

Remark 2.3. If $d = 1$, we do not need the assumption of semiconcavity in (3). Due to technical difficulty, it is not clear how to remove the assumption of semiconcavity for $d \ge 2$.

Although Remark 2.3 is already announced in the work [9] of the author, we refer to more details later.

3. Proof of result

First, we state some properties of $L(x,t,\xi)$.

Lemma 3.1. *Let $H(x,t,p)$ satisfy (H1)–(H4) and let $L(x,t,\xi)$ be the Legendre transform of $H(x,t,\cdot)$. Then, L satisfies the following (L1)–(L4):*

(L1) *$L(x,t,\xi) : \mathbb{R}^d \times \mathbb{R} \times \mathbb{R}^d \to \mathbb{R}$, C^2.*

(L2) *$L_{\xi\xi}(x,t,\xi)$ is positive definite in $\mathbb{R}^d \times \mathbb{R} \times \mathbb{R}^d$.*

(L3) *L is uniformly superlinear with respect to ξ, i.e., for each $a \ge 0$ there exists $b_2(a) \in \mathbb{R}$ such that $L(x,t,\xi) \ge a\|\xi\| + b_2(a)$ on $\mathbb{R}^d \times \mathbb{R} \times \mathbb{R}^d$.*

(L4) *$L, L_{x^i}, L_{\xi^i}, L_{x^ix^j}, L_{x^i\xi^j}, L_{\xi^i\xi^j}$ are uniformly bounded on $\mathbb{R}^d\times\mathbb{R}\times K$ for each compact set $K \subset \mathbb{R}^d$ for $i,j = 1,\dots,d$.*

Proof. Minor variation of the reasoning in Chapter 1 of [4] yields the assertion. For reader's convenience, we give a proof. Fix $(x,t,\xi) \in \mathbb{R}^d \times \mathbb{R} \times \mathbb{R}^d$ arbitrarily. Take $a > \|\xi\|$. By (H3), we have $\xi \cdot p - H(x,t,p) \le \xi \cdot p - (a\|p\| + b_1(a)) \le$

$(\|\xi\|-a)\|p\|-b_1(a)\to-\infty$ as $\|p\|\to+\infty$. Hence, $\sup_{p\in\mathbb{R}^d}\{\xi\cdot p-H(x,t,p)\}$ is achieved in a bounded ball of $\mathbb{R}^d$, namely, for each ξ, there exists $p^*\in\mathbb{R}^d$ such that $L(x,t,\xi)=\xi\cdot p^*-H(x,t,p^*)$ and $H_p(x,t,p^*)=\xi$. Since $H_p(x,t,p^*)=\xi$ is invertible with respect to p^* due to (H2), we have the C^1-map $p^*=p(x,t,\xi)$ such that $H_p(x,t,p(x,t,\xi))\equiv\xi$. Therefore, L is of C^1. Furthermore, direct computation yields

$$\begin{aligned}
L_\xi(x,t,\xi)&=p(x,t,\xi)\in C^1,\\
L_x(x,t,\xi)&=-H_x(x,t,p(x,t,\xi))\in C^1,\\
L_t(x,t,\xi)&=-H_t(x,t,p(x,t,\xi))\in C^1,\\
L_{\xi\xi}(x,t,\xi)&=p_\xi(x,t,\xi)=H_{pp}(x,t,p(x,t,\xi))^{-1},\\
L_{xx}(x,t,\xi)&=-H_{xx}(x,t,p(x,t,\xi))\\
&\qquad+H_{xp}(x,t,p(x,t,\xi))H_{pp}(x,t,p(x,t,\xi))^{-1}H_{px}(x,t,p(x,t,\xi)),\\
L_{x\xi}(x,t,\xi)&=-H_{xp}(x,t,p(x,t,\xi))H_{pp}(x,t,p(x,t,\xi))^{-1}.
\end{aligned}$$

(L1) and (L2) are proved.

For each $a\ge 0$, define $b_2(a):=-\max_{x\in\mathbb{R}^d,t\in\mathbb{R},\|p\|=a}H(x,t,p)$. It holds that $L(x,t,\xi)\ge\xi\cdot p-H(x,t,p)$ for all x,t,ξ,p. For each $\xi\in\mathbb{R}^d$, take $p:=a\xi/\|\xi\|$. Then, we see that $L(x,t,\xi)\ge a\|\xi\|+b_2(a)$, which holds for all x,t,ξ. (L3) is proved.

Let K be a compact subset of $\mathbb{R}^d$ and let $a>\max_{\xi\in K}\|\xi\|$. Then, we have $H(x,t,p)\ge a\|p\|+b_1(a)$ for all x,t,p. Hence, $\xi\cdot p-H(x,t,p)\le\|\xi\|\|p\|-(a\|p\|+b_1(a))<-b_1(a)$ for all $\xi\in K$ and $(x,t,p)\in\mathbb{R}^d\times\mathbb{R}\times\mathbb{R}^d$. Therefore, with (L3), we obtain $b_2(a)\le L(x,t,\xi)=\xi\cdot p(x,t,\xi)-H(x,t,p(x,t,\xi))<-b_1(a)$ on $\mathbb{R}^d\times\mathbb{R}\times K$. Finally, we show the boundedness of $p(x,t,\xi)$ on $\mathbb{R}^d\times\mathbb{R}\times K$, which ends the proof of (L4) with the above expressions of the derivatives of L. Suppose that there exist $(x_j,t_j,\xi_j)\in\mathbb{R}^d\times\mathbb{R}\times K$ for which the norm of $p_j:=p(x_j,t_j,\xi_j)$ goes to $+\infty$ as $j\to\infty$. Then, it holds that $L(x_j,t_j,\xi_j)=\xi_j\cdot p_j-H(x_j,t_j,p_j)\le\|\xi_j\|\|p_j\|-(a\|p_j\|+b_1(a))=(\|\xi_j\|-a)\|p_j\|-b_1(a)\to-\infty$ as $j\to\infty$. However, L is bounded below by $b_2(a)$ on $\mathbb{R}^d\times\mathbb{R}\times K$, and we reach a contradiction. $\square$

Proof of Theorem 2.1. We recall and define the following constants:

$$\begin{aligned}
&\sup_{x\in\mathbb{R}^d}|v^0(x)|\le R,\qquad \operatorname{ess\,sup}_{x\in\mathbb{R}^d}\|v^0_x(x)\|_\infty\le r,\\
&\|L_x(x,t,\xi)\|_\infty\le\alpha(1+|L(x,t,\xi)|),\qquad L_*:=\min\{0,\inf_{x\in\mathbb{R}^d,t\in\mathbb{R},\xi\in\mathbb{R}^d}L(x,t,\xi)\},\\
&\alpha_1:=T\max_{x\in\mathbb{R}^d,t\in\mathbb{R}}|L(x,t,0)|+R,\quad \alpha_2:=\alpha\{\alpha_1+R+(1+2|L_*|)T\},\\
&\lambda_1:=(d\max_{x\in\mathbb{R}^d,t\in\mathbb{R},\|p\|_\infty\le 1+r+\alpha_2}\|H_p(x,t,p)\|_\infty)^{-1},\\
&\theta:=T\max_{x\in\mathbb{R}^d,t\in\mathbb{R},\|\xi\|_\infty\le(d\lambda_1)^{-1},i,j}|L_{x^ix^j}(x,t,\xi)|,
\end{aligned}$$

where $\|x\|_\infty:=\max_{1\le j\le d}|x^j|$ for $x\in\mathbb{R}^d$. Let $\Delta=(\Delta x,\Delta t)$ be such that $\Delta x\theta\le 1$ and $\lambda=\Delta t/\Delta x<\lambda_1$.

Since L and v^0 are bounded below, there exists the infimum of $E^{l+1}_n(\xi)$ with respect to $\xi:G^{l+1,1}_n\to[-(d\lambda)^{-1},(d\lambda)^{-1}]^d$ for each n and $0\le l\le k(T)$. Since $G^{l+1,1}_n$ consists of a finite number of points, the compactness of a finite-dimensional Euclidean space implies that there exists a minimizing control ξ^* that attains the infimum.

We have the following equality for each $0 < k \le l$:

$$E_n^{l+1}(\xi) = \sum_{\gamma\in\Omega_n^{l+1,k}} \mu_n^{l+1,k}(\gamma;\xi|_{G_n^{l+1,k+1}})\Big(\sum_{k<k'\le l+1} L(\gamma^{k'},t_{k'-1},\xi^{k'}_{m(\gamma^{k'})})\Delta t + E^k_{m(\gamma^k)}(\xi|_{G^{k,1}_{m(\gamma^k)}})\Big) + h(t_{l+1}-t_k). \tag{3.1}$$

In order to check (3.1), set $\tilde\gamma := \gamma|_{k\le k'\le l+1} \in \Omega_n^{l+1,k}$ and $\hat\gamma := \gamma|_{0\le k'\le k} \in \Omega^{k,0}_{m(\tilde\gamma^k)}$ for each $\gamma \in \Omega_n^{l+1,0}$. Then, we have $\mu_n^{l+1,0}(\gamma) = \mu_n^{l+1,k}(\tilde\gamma)\mu^{k,0}_{m(\tilde\gamma^k)}(\hat\gamma)$. Hence, it follows from the definition of the random walk that we have for each $0 < k \le l$,

$$\begin{aligned}
E_n^{l+1}(\xi) &= \sum_{\gamma\in\Omega_n^{l+1,0}} \mu_n^{l+1,0}(\gamma;\xi)\Big(\sum_{0<k\le l+1} L(\gamma^k,t_{k-1},\xi^k_{m(\gamma^k)})\Delta t + v^0_{m(\gamma^0)}\Big) + ht_{l+1}\\
&= \sum_{\tilde\gamma\in G_n^{l+1,k}} \mu_n^{l+1,k}(\tilde\gamma;\xi|_{G_n^{l+1,k+1}})\Big\{\sum_{\hat\gamma\in G^{k,0}_{m(\tilde\gamma^k)}} \mu^{k,0}_{m(\tilde\gamma^k)}(\hat\gamma;\xi|_{G^{k,1}_{m(\tilde\gamma^k)}})\\
&\qquad \times\Big(\sum_{k<k'\le l+1} L(\tilde\gamma^{k'},t_{k'-1},\xi^{k'}_{m(\tilde\gamma^{k'})})\Delta t + \sum_{0<k'\le k} L(\hat\gamma^{k'},t_{k'-1},\xi^{k'}_{m(\hat\gamma^{k'})})\Delta t\\
&\qquad + v^0_{m(\hat\gamma^0)}\Big)\Big\} + ht_k + h(t_{l+1}-t_k),\\
&= \sum_{\tilde\gamma\in G_n^{l+1,k}} \mu_n^{l+1,k}(\tilde\gamma;\xi|_{G_n^{l+1,k+1}})\Big[\sum_{k<k'\le l+1} L(\tilde\gamma^{k'},t_{k'-1},\xi^{k'}_{m(\tilde\gamma^{k'})})\Delta t\\
&\qquad + \Big\{\sum_{\hat\gamma\in G^{k,0}_{m(\tilde\gamma^k)}} \mu^{k,0}_{m(\tilde\gamma^k)}(\hat\gamma;\xi|_{G^{k,1}_{m(\tilde\gamma^k)}})\Big(\sum_{0<k'\le k} L(\hat\gamma^{k'},t_{k'-1},\xi^{k'}_{m(\hat\gamma^{k'})})\Delta t\\
&\quad + v^0_{m(\hat\gamma^0)}\Big) + ht_k\Big\}\Big] + h(t_{l+1}-t_k),
\end{aligned}$$

which implies (3.1).

We observe that for each n and any ξ,

$$\begin{aligned}
E_n^1(\xi) &= L(x_n,t_0,\xi_n^1)\Delta t + \sum_{\omega\in B}\rho_n^1(\omega)v^0_{n+\omega} + ht_1\\
&= L(x_n,t_0,\xi_n^1)\Delta t + \frac{1}{2d}\sum_{\omega\in B} v^0_{n+\omega} - \frac{\lambda}{2}\sum_{\omega\in B}(\omega\cdot\xi_n^1)v^0_{n+\omega} + h\Delta t\\
&= L(x_n,t_0,\xi_n^1)\Delta t + \frac{1}{2d}\sum_{\omega\in B} v^0_{n+\omega}\\
&\qquad - \frac{\Delta t}{2\Delta x}\sum_{i=1}^d \xi_n^{1\,i}(v^0_{n+e_i} - v^0_{n-e_i}) + h\Delta t\\
&= L(x_n,t_0,\xi_n^1)\Delta t + \frac{1}{2d}\sum_{\omega\in B} v^0_{n+\omega} - \Delta t\xi_n^1\cdot(D_xv)^0_n + h\Delta t\\
&= -(\xi_n^1\cdot(D_xv)^0_n - L(x_n,t_0,\xi_n^1))\Delta t + \frac{1}{2d}\sum_{\omega\in B} v^0_{n+\omega} + h\Delta t\\
&\ge -H(x_n,t_0,(D_xv)^0_n)\Delta t + \frac{1}{2d}\sum_{\omega\in B} v^0_{n+\omega} + h\Delta t,
\end{aligned} \tag{3.2}$$

where the last inequality becomes an equality if and only if ξ is given as

$$\xi^1_n = H_p(x_n, t_0, (D_x v)^0_n).$$

Hence, the minimizing control ξ^* for $v^1_n = \inf_\xi E^1_n(\xi)$ satisfies

$$\xi^{*1}_{\ n} = H_p(x_n, t_0, (D_x v)^0_n), \quad \|\xi^*\|_\infty \le (d\lambda_1)^{-1} \tag{3.3}$$

with

$$\begin{aligned} v^1_n &= -H(x_n, t_0, (D_x v)^0_n)\Delta t + \frac{1}{2d}\sum_{\omega\in B} v^0_{n+\omega} + h\Delta t \\ &\Leftrightarrow (D_t v)^1_n + H(x_n, t_0, (D_x v)^0_n) = h. \end{aligned} \tag{3.4}$$

We proceed by induction. Suppose the statement

(A)$_l$ For some $l \ge 0$, the minimizing control ξ^* for $v^{l+1}_n = \inf_\xi E^{l+1}_n(\xi)$ is uniquely obtained for each n as $\xi^{*k+1}_{\ m} = H_p(x_m, t_k, (D_x v)^k_m)$ on $G^{l+1,1}_n$ with $\|\xi^{*k}_{\ m+\mathbf{1}}\|_\infty \le (d\lambda_1)^{-1}$.

(A)$_{l=0}$ is true. In order to see that (A)$_{l+1}$ is also true, we first examine the bound of $(D_x v)^{l+1}_{n+\mathbf{1}}$ for each n. Let $\xi^*(\omega)$ denote the unique minimizer for $v^{l+1}_{n+\mathbf{1}+\omega}$, $\omega \in B$. The variational property yields for each e_j,

$$\begin{aligned} v^{l+1}_{n+\mathbf{1}-e_j} &= E_{\mu^{l+1,0}_{n+\mathbf{1}-e_j}(\cdot;\xi^*(-e_j))}\Big[\sum_{0<k\le l+1} L(\gamma^k, t_{k-1}, \xi^{*k}_{\ m(\gamma^k)}(-e_j))\Delta t + v^0_{m(\gamma^0)}\Big] \\ &\quad + ht_k, \\ v^{l+1}_{n+\mathbf{1}+e_j} &\le E_{\mu^{l+1,0}_{n+\mathbf{1}-e_j}(\cdot;\xi^*(-e_j))}\Big[\sum_{0<k\le l+1} L(\gamma^k + e_j\cdot 2\Delta x, t_{k-1}, \xi^{*k}_{\ m(\gamma^k)}(-e_j))\Delta t \\ &\quad + v^0_{m(\gamma^0+e_j\cdot 2\Delta x)}\Big] + ht_k. \end{aligned}$$

Hence, from $v^{l+1}_{n+\mathbf{1}+e_j} - v^{l+1}_{n+\mathbf{1}-e_j}$, we obtain

$$\begin{aligned}
(D_{x^j}v)^{l+1}_{n+\mathbf{1}} &\le E_{\mu^{l+1,0}_{n+\mathbf{1}-e_j}(\cdot;\xi^*(-e_j))}\Big[\sum_{0<k\le l+1} L_{x^j}(\gamma^k, t_{k-1}, \xi^{*k}_{\ m(\gamma^k)}(-e_j))\Delta t \\
&\quad + (D_{x^j}v)^0_{m(\gamma^0)+e_j}\Big] + \theta\Delta x \\
&\le E_{\mu^{l+1,0}_{n+\mathbf{1}-e_j}(\cdot;\xi^*(-e_j))}\Big[\sum_{0<k\le l+1} L_{x^j}(\gamma^k, t_{k-1}, \xi^{*k}_{\ m(\gamma^k)}(-e_j))\Delta t\Big] + r + \theta\Delta x \\
&\le E_{\mu^{l+1,0}_{n+\mathbf{1}-e_j}(\cdot;\xi^*(-e_j))}\Big[\sum_{0<k\le l+1} \alpha\{1 + |L(\gamma^k, t_{k-1}, \xi^{*k}_{\ m(\gamma^k)}(-e_j))|\}\Delta t\Big] + r + \theta\Delta x \\
&\le E_{\mu^{l+1,0}_{n+\mathbf{1}-e_j}(\cdot;\xi^*(-e_j))}\Big[\sum_{0<k\le l+1} \alpha\{1 + L(\gamma^k, t_{k-1}, \xi^{*k}_{\ m(\gamma^k)}(-e_j)) + 2|L_*|\}\Delta t\Big] \\
&\quad + r + \theta\Delta x \\
&= E_{\mu^{l+1,0}_{n+\mathbf{1}-e_j}(\cdot;\xi^*(-e_j))}\Big[\sum_{0<k\le l+1} \alpha\{1 + L(\gamma^k, t_{k-1}, \xi^{*k}_{\ m(\gamma^k)}(-e_j)) + 2|L_*|\}\Delta t \\
&\quad + \alpha\{v^0_{m(\gamma^0)} - v^0_{m(\gamma^0)}\}\Big] + r + \theta\Delta x \\
&\le \alpha E_{\mu^{l+1,0}_{n+\mathbf{1}-e_j}(\cdot;\xi^*(-e_j))}\Big[\sum_{0<k\le l+1} L(\gamma^k, t_{k-1}, \xi^{*k}_{\ m(\gamma^k)}(-e_j))\Delta t + v^0_{m(\gamma^0)}\Big] \\
&\quad + \alpha T + 2\alpha|L_*|T + \alpha R + r + \theta\Delta x \\
&\le \alpha E_{\mu^{l+1,0}_{n+\mathbf{1}-e_j}(\cdot;0)}\Big[\sum_{0<k\le l+1} L(\gamma^k, t_{k-1}, 0)\Delta t + v^0_{m(\gamma^0)}\Big] \\
&\quad + \alpha T + 2\alpha|L_*|T + \alpha R + r + \theta\Delta x \\
&\le \alpha\alpha_1 + \alpha T + 2\alpha|L_*|T + \alpha R + r + 1 = \alpha_2 + r + 1.
\end{aligned}$$

We can show $(D_{x^j}v)^{l+1}_{n+\mathbf{1}} \ge -\alpha_2 - r - 1$ in a similar way with the minimizing control for $v^{l+1}_{n+\mathbf{1}+e_j}$. Therefore, we obtain

$$\begin{aligned}
&\|(D_x v)^{l+1}_{n+\mathbf{1}}\|_\infty \le 1 + r + \alpha_2, \\
&\|H_p(x_{n+\mathbf{1}}, t_{l+1}, (D_x v)^{l+1}_{n+\mathbf{1}})\|_\infty \le (d\lambda_1)^{-1}. \qquad (3.5)
\end{aligned}$$

With (3.1) and similar calculation in (3.2), we observe that for each n and for any ξ,

$$\begin{aligned}
E^{l+2}_{n+\mathbf{1}}(\xi) &= L(x_{n+\mathbf{1}}, t_{l+1}, \xi^{l+2}_{n+\mathbf{1}})\Delta t + \sum_{\omega\in B} \rho^{l+2}_{n+\mathbf{1}}(\omega) E^{l+1}_{n+\mathbf{1}+\omega}(\xi|_{G^{l+1,1}_{n+\mathbf{1}+\omega}}) + h\Delta t \\
&\ge L(x_{n+\mathbf{1}}, t_{l+1}, \xi^{l+2}_{n+\mathbf{1}})\Delta t + \sum_{\omega\in B} \rho^{l+2}_{n+\mathbf{1}}(\omega) v^{l+1}_{n+\mathbf{1}+\omega} + h\Delta t \qquad (3.6) \\
&= L(x_{n+\mathbf{1}}, t_{l+1}, \xi^{l+2}_{n+\mathbf{1}})\Delta t + \frac{1}{2d}\sum_{\omega\in B} v^{l+1}_{n+\mathbf{1}+\omega} - \xi^{l+2}_{n+\mathbf{1}} \cdot (D_x v)^{l+1}_{n+\mathbf{1}}\Delta t + h\Delta t \\
&\ge -H(x_{n+\mathbf{1}}, t_{l+1}, (D_x v)^{l+1}_{n+\mathbf{1}})\Delta t + \frac{1}{2d}\sum_{\omega\in B} v^{l+1}_{n+\mathbf{1}+\omega} + h\Delta t. \qquad (3.7)
\end{aligned}$$

It follows from the assumption $(\mathrm{A})_l$ of induction and the properties of the Legendre transform that the above inequalities (3.6) and (3.7) become equalities, if $\xi=\xi^*$ is given as

$$\xi^{*k+1}_{\ m}=H_p(x_m,t_k,(D_xv)^k_m)\quad \text{on } G^{l+2,1}_{n+\mathbf{1}}, \tag{3.8}$$

which makes sense due to (3.5). Hence, we obtain a minimizing control ξ^* for $v^{l+2}_{n+\mathbf{1}}=\inf_\xi E^{l+2}_{n+\mathbf{1}}(\xi)$ with $\|\xi^*\|_\infty\le(d\lambda_1)^{-1}$ and

$$\begin{aligned} v^{l+2}_{n+\mathbf{1}}&=-H(x_{n+\mathbf{1}},t_{l+1},(D_xv)^{l+1}_{n+\mathbf{1}})\Delta t+\frac{1}{2d}\sum_{\omega\in B}v^{l+1}_{n+\mathbf{1}+\omega}+h\Delta t\\ &\Leftrightarrow (D_tv)^{l+2}_{n+\mathbf{1}}+H(x_{n+\mathbf{1}},t_{l+1},(D_xv)^{l+1}_{n+\mathbf{1}})=h. \end{aligned}\tag{3.9}$$

Let ξ be another minimizing control for $v^{l+2}_{n+\mathbf{1}}$ different from the above ξ^*. If $\xi^{l+2}_{n+\mathbf{1}}\neq\xi^{*l+2}_{\ n+\mathbf{1}}$, then (3.7) becomes a strict inequality, yielding the contradiction that $v^{l+2}_{n+\mathbf{1}}>v^{l+2}_{n+\mathbf{1}}$. If $\xi^{l+2}_{n+\mathbf{1}}=\xi^{*l+2}_{\ n+\mathbf{1}}$, then there necessarily exists $\omega\in B$ such that $\xi|_{G^{l+1,1}_{n+\mathbf{1}+\omega}}$ is not the minimizing control for $v^{l+1}_{n+\mathbf{1}+\omega}$, because of the assumption $(\mathrm{A})_l$ of our induction; Hence (3.6) becomes a strict inequality, yielding the same contradiction.

Thus, we conclude that the minimizing control ξ^* for $v^{l+2}_{n+\mathbf{1}}(\xi)=\inf_\xi E^{l+2}_{n+\mathbf{1}}(\xi)$ is uniquely obtained for each n as (3.8) with $\|\xi^*\|_\infty\le(d\lambda_1)^{-1}$. By induction, (1) is clear. (2) follows from (3.3) and (3.8). (3) follows from (3.4) and (3.9). The inequalities in (4) are obtained in the above calculation to derive (3.5). □

Next, we study the hyperbolic scaling limit of our random walks in order to prove Theorem 2.2. Let $\bar\gamma^k$ be the averaged path of $\gamma\in\Omega^{l+1,0}_n$ generated by a control ξ, i.e.,

$$\bar\gamma^k:=\sum_{\gamma\in\Omega^{l+1,0}_n}\mu^{l+1,0}_n(\gamma)\gamma^k,\quad 0\le k\le l+1,$$

where we use the short notation $\mu^{l+1,l'}_n(\cdot)$ instead of $\mu^{l+1,l'}_n(\cdot;\xi)$. We see that $\bar\gamma$ satisfies

$$\bar\gamma^k=\bar\gamma^{k+1}-\bar\xi^{k+1}\Delta t\qquad\text{with}\quad \bar\xi^k:=\sum_{\gamma\in\Omega^{l+1,0}_n}\mu^{l+1,0}_n(\gamma)\xi^k_{m(\gamma^k)}.$$

In fact, we have

$$\begin{aligned}
\bar\gamma^k &= \sum_{\gamma\in\Omega_n^{l+1,0}} \mu_n^{l+1,0}(\gamma)\gamma^k = \sum_{\gamma\in\Omega_n^{l+1,k}} \mu_n^{l+1,k}(\gamma)\gamma^k \\
&= \sum_{\gamma\in\Omega_n^{l+1,k+1}} \sum_{\omega\in B} \mu_n^{l+1,k+1}(\gamma)\rho^{k+1}_{m(\gamma^{k+1})}(\omega)(\gamma^{k+1}+\omega\Delta x) \\
&= \bar\gamma^{k+1} + \sum_{\gamma\in\Omega_n^{l+1,k+1}} \sum_{\omega\in B} \mu_n^{l+1,k+1}(\gamma)\Big(\frac{1}{2d} - \frac{\lambda}{2}\omega\cdot\xi^{k+1}_{m(\gamma^{k+1})}\Big)\omega\Delta x \\
&= \bar\gamma^{k+1} - \sum_{\gamma\in\Omega_n^{l+1,k+1}} \sum_{\omega\in B} \mu_n^{l+1,k+1}(\gamma)\big(\omega\cdot\xi^{k+1}_{m(\gamma^{k+1})}\big)\omega\frac{\lambda}{2}\Delta x \\
&= \bar\gamma^{k+1} - \frac12\sum_{\omega\in B}(\omega\cdot\bar\xi^{k+1})\omega\Delta t \\
&= \bar\gamma^{k+1} - \frac12\sum_{i=1}^{d}\{(e_i\cdot\bar\xi^{k+1})e_i + (-e_i\cdot\bar\xi^{k+1})(-e_i)\}\Delta t \\
&= \bar\gamma^{k+1} - \bar\xi^{k+1}\Delta t.
\end{aligned}$$

Let $\eta(\gamma)$ be a random variable defined for each $\gamma\in\Omega_n^{l+1,0}$ as

$$\eta^k(\gamma) = \eta^{k+1}(\gamma) - \xi^{k+1}_{m(\gamma^{k+1})}\Delta t,\ \ \eta^{l+1}(\gamma) = x_n.$$

Let $\eta_\Delta(\gamma)(\cdot):[0,t]\to\mathbb{R}^d$ denote the linear interpolation of $\eta(\gamma)$ with $t\in[t_{l+1},t_{l+2})$,

$$\eta_\Delta(\gamma)(s) := \begin{cases} x_n \quad \text{for } s\in[t_{l+1},t], \\ \eta^k(\gamma) + \frac{\eta^{k+1}(\gamma)-\eta^k(\gamma)}{\Delta t}(s-t_k) \quad \text{for } s\in[t_k,t_{k+1}]. \end{cases}$$

Define $\tilde\sigma_i^k$ and $\tilde\delta_i^k$ for $i=1,\dots,d$ as

$$\tilde\sigma_i^k := E_{\mu_n^{l+1,0}(\cdot;\xi)}[|(\eta^k(\gamma)-\gamma^k)^i|^2], \quad \tilde\delta_i^k := E_{\mu_n^{l+1,0}(\cdot;\xi)}[|(\eta^k(\gamma)-\gamma^k)^i|],$$

where $(\eta^k(\gamma)-\gamma^k)^i$ denotes the ith component of $\eta^k(\gamma)-\gamma^k$. The following lemma is a key to the convergence of our difference scheme:

Lemma 3.2. *For any control ξ, we have*

$$(\tilde\delta_i^k)^2 \le \tilde\sigma_i^k \le (t_{l+1}-t_k)\frac{\Delta x}{\lambda} \quad \textit{for} \quad 0\le k\le l+1.$$

Remark 3.3. $\tilde\sigma_i^k$ can be seen as a generalization of the standard variance. The standard variance is of $O(1)$ as $\Delta\to0$ under hyperbolic scaling in general for space-time inhomogeneous random walks. However, $\tilde\sigma_i^k$ and $\tilde\delta_i^k$ always tend to 0 for any control ξ as $\Delta\to0$ under hyperbolic scaling. In the homogeneous case (i.e., ξ is constant), $\tilde\sigma_i^k$ is equal to the standard variance. See also [8].

Proof of Lemma 3.2. We observe that

$$
\begin{aligned}
\tilde{\sigma}^k_i &= \sum_{\gamma\in\Omega_n^{l+1,k}} \mu_n^{l+1,k}(\gamma)|(\eta^k(\gamma)-\gamma^k)^i|^2\\
&= \sum_{\gamma\in\Omega_n^{l+1,k+1}} \mu_n^{l+1,k+1}(\gamma)\sum_{\omega\in B}\rho^{k+1}_{m(\gamma^{k+1})}(\omega)\\
&\times\big\{(\eta^{k+1}(\gamma)-\gamma^{k+1})^i-(\xi^{k+1}_{m(\gamma^{k+1})})^i\Delta t-\omega^i\Delta x\big\}^2\\
&= \tilde{\sigma}^{k+1}_i+\sum_{\gamma\in\Omega_n^{l+1,k+1}} \mu_n^{l+1,k+1}(\gamma)\Big[\sum_{\omega\in B}\rho^{k+1}_{m(\gamma^{k+1})}(\omega)\big\{\lambda(\xi^{k+1}_{m(\gamma^{k+1})})^i+\omega^i\big\}^2\Big]\Delta x^2\\
&\quad-\sum_{\gamma\in\Omega_n^{l+1,k+1}} \mu_n^{l+1,k+1}(\gamma)\Big\{2\lambda(\eta^{k+1}(\gamma)-\gamma^{k+1})^i(\xi^{k+1}_{m(\gamma^{k+1})})^i\\
&\quad+2(\eta^{k+1}(\gamma)-\gamma^{k+1})^i\sum_{\omega\in B}\rho^{k+1}_{m(\gamma^{k+1})}(\omega)\omega^i\Big\}\Delta x.
\end{aligned}
$$

Since

$$
\begin{aligned}
\sum_{\omega\in B}\rho^{k+1}_{m(\gamma^{k+1})}(\omega)\omega^i &= \sum_{j=1}^d\Big\{\Big(\frac{1}{2d}-\frac{\lambda}{2}e_j\cdot\xi^{k+1}_{m(\gamma^{k+1})}\Big)e^i_j\\
&\qquad+\Big(\frac{1}{2d}-\frac{\lambda}{2}(-e_j)\cdot\xi^{k+1}_{m(\gamma^{k+1})}\Big)(-e^i_j)\Big\}\\
&= -\lambda(\xi^{k+1}_{m(\gamma^{k+1})})^i,
\end{aligned}
$$

we obtain

$$
\begin{aligned}
\tilde{\sigma}^k_i &= \tilde{\sigma}^{k+1}_i+\sum_{\gamma\in\Omega_n^{l+1,k+1}} \mu_n^{l+1,k+1}(\gamma)\Big[\sum_{\omega\in B}\rho^{k+1}_{m(\gamma^{k+1})}(\omega)\big\{\omega^i+\lambda(\xi^{k+1}_{m(\gamma^{k+1})})^i\big\}^2\Big]\Delta x^2\\
&= \tilde{\sigma}^{k+1}_i+\sum_{\gamma\in\Omega_n^{l+1,k+1}} \mu_n^{l+1,k+1}(\gamma)\Big[\sum_{\omega\in B}\rho^{k+1}_{m(\gamma^{k+1})}(\omega)\big[(\omega^i)^2+2\lambda\omega^i(\xi^{k+1}_{m(\gamma^{k+1})})^i\\
&\quad+\{\lambda(\xi^{k+1}_{m(\gamma^{k+1})})^i\}^2\big]\Big]\Delta x^2\\
&\le \tilde{\sigma}^{k+1}_i+\sum_{\gamma\in\Omega_n^{l+1,k+1}} \mu_n^{l+1,k+1}(\gamma)\Big[1+\{\lambda(\xi^{k+1}_{m(\gamma^{k+1})})^i\}^2\\
&\quad+2\lambda(\xi^{k+1}_{m(\gamma^{k+1})})^i\sum_{\omega\in B}\rho^{k+1}_{m(\gamma^{k+1})}(\omega)\omega^i\Big]\Delta x^2\\
&= \tilde{\sigma}^{k+1}_i+\sum_{\gamma\in\Omega_n^{l+1,k+1}} \mu_n^{l+1,k+1}(\gamma)\Big[1-\{\lambda(\xi^{k+1}_{m(\gamma^{k+1})})^i\}^2\Big]\Delta x^2\\
&\le \tilde{\sigma}^{k+1}_i+\frac{\Delta x}{\lambda}\Delta t.
\end{aligned}
$$

Since $\tilde{\sigma}^{l+1}_i=0$, the assertion is proved. □

The rate $O(\sqrt{\Delta x})$ of convergence of our scheme comes from the asymptotics of $\tilde{\delta}^k_i$ as we will see below.

We observe the following facts on the viscosity solution v of (1.1).

Lemma 3.4. *Let $\gamma^* : [0,t] \to \mathbb{R}^d$ be a minimizing curve for $v(x,t)$.*

(1) *The following regularity properties hold:*

$$L^c_\xi(\gamma^*(\tau),\tau,\gamma^{*\prime}(\tau)) \in \partial^-_x v(\gamma^*(\tau),\tau) \text{ for } 0 \le \tau < t,$$
$$L^c_\xi(\gamma^*(\tau),\tau,\gamma^{*\prime}(\tau)) \in \partial^+_x v(\gamma^*(\tau),\tau) \text{ for } 0 < \tau \le t,$$

where $\partial^-_x v$ (resp., $\partial^+_x v$) denotes the subdifferential (resp., superdifferential). In particular $v_x(\gamma^*(\tau),\tau)$ exists for $0 < \tau < t$ and is equal to $L^c_\xi(\gamma^*(\tau),\tau,\gamma^{*\prime}(\tau))$.

(2) $|\gamma^{*\prime}(\tau)| \le (d\lambda_1)^{-1}$ *for $0 \le \tau \le t$, where λ_1 is the number given in the proof of Theorem 2.1.*

(3) *If (x,t) is regular, we have for any $0 \le \tau < t$*

$$v_x(x,t) = \int_\tau^t L^c_x(\gamma^*(s),s,\gamma^{*\prime}(s))ds + L^c_\xi(\gamma^*(\tau),\tau,\gamma^{*\prime}(\tau)).$$

If the initial data v^0 is semiconcave, the superdifferential of v^0 is not empty and $L^c_\xi(\gamma^*(0),0,\gamma^{*\prime}(0)) = v^0_x(\gamma^*(0))$.

This lemma is well known (a proof is given in the same manner as the proof of Lemma 3.2 in [9]).

Proof of Theorem 2.2. Hereafter, $\beta_1,\beta_2,\dots$ are constants independent of Δx, Δt, x_m, t_k, and v^0 in (1.1).

We prove (1). Since the solution v of (1.1) is Lipschitz ((1) and (2) of Lemma 3.4 imply a Lipschitz constant independent of v^0), it is enough to show $|v^{l+1}_n - v(x_n,t_{l+1})| \le \beta\sqrt{\Delta x}$. Let γ^* be a minimizing curve for $v(x_n,t_{l+1})$. Consider the control ξ defined on $G^{l+1,1}_n$ as

$$\xi(x_m,t_{k+1}) := \gamma^{*\prime}(t_{k+1})$$

and the random walk γ generated by ξ. Then, $\eta(\gamma)$ is independent of $\gamma \in \Omega^{l+1,0}_n$ and satisfies

$$\begin{aligned}
&\|\eta^k(\gamma) - \gamma^*(t_k)\|_\infty \le \beta_1\Delta x \quad \text{for all } 0 \le k \le l+1,\\
&\Big| \int_0^{t_{l+1}} L(\gamma^*(s),s,\gamma^{*\prime}(s))ds + v^0(\gamma^*(0))\\
&\qquad - \Big(\sum_{0<k\le l+1} L(\eta^k(\gamma),t_{k-1},\xi^k_{m(\gamma^k)})\Delta t + v^0(\eta^0(\gamma))\Big)\Big| \le \beta_2\Delta x.
\end{aligned}$$

It follows from Lemma 3.2 that

$$\begin{aligned}
v^{l+1}_n &\le E_{\mu(\cdot;\xi)}\Big[\sum_{0<k\le l+1} L(\gamma^k,t_{k-1},\xi^k_{m(\gamma^k)})\Delta t + v^0_{m(\gamma^0)}\Big] + ht_{l+1}\\
&\le E_{\mu(\cdot;\xi)}\Big[\sum_{0<k\le l+1} L(\gamma^k,t_{k-1},\xi^k_{m(\gamma^k)})\Delta t + v^0(\gamma^0)\Big] + ht_{l+1} + r\Delta x\\
&\le E_{\mu(\cdot;\xi)}\Big[\sum_{0<k\le l+1} L(\eta^k(\gamma),t_{k-1},\xi^k_{m(\gamma^k)})\Delta t + v^0(\eta^0(\gamma))\Big] + ht_{l+1} + \beta_3\sqrt{\Delta x}\\
&\le \int_0^{t_{l+1}} L(\gamma^*(s),s,\gamma^{*\prime}(s))ds + v^0(\gamma^*(0)) + ht_{l+1} + \beta_4\sqrt{\Delta x}.
\end{aligned}$$

Hence, we obtain

$$v^{l+1}_n - v(x_n,t_{l+1}) \le \beta_4\sqrt{\Delta x}.$$

Let ξ^* be the minimizing control for v_n^{l+1}. Consider the linear interpolation of $\eta(\gamma)$ within $[0,t_{l+1}]$. Then, we have

$$\eta_\Delta(\gamma)'(s)=\xi^{*k}_{\ m(\gamma^k)}\ \text{ for } s\in(t_{k-1},t_k),$$

$$v(x_n,t_{l+1})\le\int_0^{t_{l+1}}L(\eta_\Delta(s),s,\eta_\Delta(\gamma)'(s))dt+v^0(\eta_\Delta(\gamma)(0))+ht_{l+1}\ \text{ for } \gamma\in\Omega_n^{l+1,0},$$

$$v(x_n,t_{l+1})\le E_{\mu_n^{l+1,0}(\cdot;\xi^*)}\Big[\int_0^{t_{l+1}}L(\eta_\Delta(s),s,\eta_\Delta(\gamma)'(s))dt+v^0(\eta_\Delta(\gamma)(0))\Big]+ht_{l+1}.$$

It follows from Lemma 3.2 that

$$\begin{aligned}
v_n^{l+1}&=E_{\mu_n^{l+1,0}(\cdot;\xi^*)}\Big[\sum_{0<k\le l+1}L(\gamma^k,t_{k-1},\xi^{*k}_{\ m(\gamma^k)})\Delta t+v^0_{m(\gamma^0)}\Big]+ht_{l+1}\\
&\ge E_{\mu_n^{l+1,0}(\cdot;\xi^*)}\Big[\sum_{0<k\le l+1}L(\gamma^k,t_{k-1},\xi^{*k}_{\ m(\gamma^k)})\Delta t+v^0(\gamma^0)\Big]+ht_{l+1}-r\Delta x\\
&\ge E_{\mu_n^{l+1,0}(\cdot;\xi^*)}\Big[\sum_{0<k\le l+1}L(\eta^k(\gamma),t_{k-1},\xi^{*k}_{\ m(\gamma^k)})\Delta t+v^0(\eta^0(\gamma))\Big]\\
&\qquad+ht_{l+1}-\beta_5\sqrt{\Delta x}\\
&\ge E_{\mu_n^{l+1,0}(\cdot;\xi^*)}\Big[\int_0^{t_{l+1}}L(\eta_\Delta(\gamma)(s),s,\eta_\Delta(\gamma)'(s))ds+v^0(\eta_\Delta(\gamma)(0))\Big]\\
&\qquad+ht_{l+1}-\beta_6\sqrt{\Delta x}.
\end{aligned}$$

Therefore, we obtain

$$v_n^{l+1}-v(x_n,t_{l+1})\ge-\beta_6\sqrt{\Delta x}.$$

Thus, (1) is proved.

In order to prove (2), we prepare two lemmas.

Lemma 3.5. *Let γ^* be the unique minimizer for $v(x,t)$. Define the set Γ^ε with $\varepsilon>0$ and $b>0$ as the family of all Lipschitz curves $\nu:[0,t]\to\mathbb{R}^d$ such that*

$$\|\nu(t)-\gamma^*(t)\|_\infty\le\varepsilon,\quad \|\nu'(s)\|_\infty\le b \text{ for a.e. } s\in[0,t],$$

$$\int_0^t L(\nu(s),s,\nu'(s))ds+v^0(\nu(0))\le\int_0^t L(\gamma^*(s),s,\gamma^{*\prime}(s))ds+v^0(\gamma^*(0))+\varepsilon.$$

Then, we have as $\varepsilon\to0$,

$$\sup_{\nu\in\Gamma^\varepsilon}\|\nu-\gamma^*\|_{C^0([0,t])}\to0,\ \ \sup_{\nu\in\Gamma^\varepsilon}\|\nu'-\gamma^{*\prime}\|_{L^2([0,t])}\to0.$$

Here, $\|\nu\|_{C^0([0,t])}:=\sup_{s\in[0,t]}\|\nu(s)\|$ and $\|\nu\|_{L^2([0,t])}:=\{\int_0^t\|\nu(s)\|^2ds\}^{\frac12}$. Lemma 3.5 is proved in a similar way to the proof of Lemma 3.4 in [9].

Lemma 3.6. *Let $f:[0,t]\to\mathbb{R}$ be a Lipschitz function with a Lipschitz constant θ satisfying $f(t)=0$. Then, it holds that $\|f\|_{C^0([0,t])}\le\theta\|f\|_{L^2([0,t])}+\sqrt{\|f\|_{L^2([0,t])}}$.*

See Lemma 3.5 of [9] for a proof.

We prove (2). For each fixed $\varepsilon>0$, define the set

$$\Omega^\varepsilon_\Delta:=\{\gamma\in\Omega_n^{l+1,0}\mid\|\gamma_\Delta-\gamma^*\|_{C^0([0,t])}\le\varepsilon\}.$$

Our task is to prove that $prob(\Omega^\varepsilon_\Delta)\to1$ as $\Delta\to0$. We first obtain an estimate of $\|\gamma_\Delta-\gamma^*\|_{L^2([0,t])}$ and then convert it into that of $\|\gamma_\Delta-\gamma^*\|_{C^0([0,t])}$ using Lemma 3.6.

Observe that

$$(E_{\mu_n^{l+1,0}(\cdot;\xi^*)}[\|\gamma_\Delta-\gamma^*\|_{L^2([0,t])}])^2 \le E_{\mu_n^{l+1,0}(\cdot;\xi^*)}[\|\gamma_\Delta-\gamma^*\|^2_{L^2([0,t])}] \tag{3.10}$$
$$\le 2E_{\mu_n^{l+1,0}(\cdot;\xi^*)}[\|\gamma_\Delta-\eta_\Delta(\gamma)\|^2_{L^2([0,t])}]$$
$$+\,2E_{\mu_n^{l+1,0}(\cdot;\xi^*)}[\|\eta_\Delta(\gamma)-\gamma^*\|^2_{L^2([0,t])}],$$

where $E_{\mu(\cdot;\xi^*)}[\|\gamma_\Delta-\eta_\Delta(\gamma)\|^2_{L^2([0,t])}]$ tends to 0 as $\Delta\to 0$ due to Lemma 3.2. We show that $E_{\mu_n^{l+1,0}(\cdot;\xi^*)}[\|\eta_\Delta(\gamma)-\gamma^*\|^2_{L^2([0,t])}]$ also tends to 0 as $\Delta\to 0$. For this purpose, define the set

$$\tilde\Omega^\varepsilon_\Delta := \{\gamma\in\Omega_n^{l+1,0} \mid \|\eta_\Delta(\gamma)-\gamma^*\|_{C^0([0,t])}\le\varepsilon,\ \ \|\eta_\Delta(\gamma)'-\gamma^{*\prime}\|_{L^2([0,t])}\le\varepsilon\}$$

for each fixed $\varepsilon>0$, and show $prob(\tilde\Omega^\varepsilon_\Delta)\to 1$ as $\Delta\to 0$. It follows from Lemma 3.2 that

$$v_n^{l+1} = E_{\mu_n^{l+1,0}(\cdot;\xi^*)}\Big[\int_0^t L(\eta_\Delta(\gamma)(s),s,\eta_\Delta(\gamma)'(s))ds + v^0(\eta_\Delta(\gamma)(0))\Big]$$
$$+\,ht_{l+1}+O(\sqrt{\Delta x}).$$

By (1), we have

$$v_n^{l+1}-v(x,t) = O(\sqrt{\Delta x})$$
$$= E_{\mu_n^{l+1,0}(\cdot;\xi^*)}\Big[\int_0^t L(\eta_\Delta(\gamma)(s),s,\eta_\Delta(\gamma)'(s))ds + v^0(\eta_\Delta(\gamma)(0))$$
$$-\Big(\int_0^t L(\gamma^*(s),s,\gamma^{*\prime}(s))ds + v^0(\gamma^*(0))\Big)\Big] + h(t_{l+1}-t)$$
$$+\,O(\sqrt{\Delta x}).$$

Hence, we obtain

$$E_{\mu_n^{l+1,0}(\cdot;\xi^*)}\Big[\int_0^t L(\eta_\Delta(\gamma)(s),s,\eta_\Delta(\gamma)'(s))ds + v^0(\eta_\Delta(\gamma)(0)) \tag{3.11}$$
$$-\Big(\int_0^t L(\gamma^*(s),s,\gamma^{*\prime}(s))ds + v^0(\gamma^*(0))\Big)\Big] = O(\sqrt{\Delta x}).$$

Consider the set

$$\Omega^+ \quad := \quad \Big\{\gamma\in\Omega_n^{l+1,0}\,\Big|\, \int_0^t L(\eta_\Delta(\gamma)(s),s,\eta_\Delta(\gamma)'(s))ds + v^0(\eta_\Delta(\gamma)(0))$$
$$-\Big(\int_0^t L(\gamma^*(s),s,\gamma^{*\prime}(s))ds + v^0(\gamma^*(0))\Big) \ge \Delta x^{\frac14}\Big\}.$$

Since γ^* is a minimizing curve for $v(x,t)$, we have for each $\gamma\in\Omega_n^{l+1,0}$,

$$0 \quad\le\quad \int_0^t L(\eta_\Delta(\gamma)(s)+x-x_n,s,\eta_\Delta(\gamma)'(s))ds + v^0(\eta_\Delta(\gamma)(0)+x-x_n)$$
$$-\Big(\int_0^t L(\gamma^*(s),s,\gamma^{*\prime}(s))ds + v^0(\gamma^*(0))\Big)$$
$$\le\quad \int_0^t L(\eta_\Delta(\gamma)(s),s,\eta_\Delta(\gamma)'(s))ds + v^0(\eta_\Delta(\gamma)(0))$$
$$-\Big(\int_0^t L(\gamma^*(s),s,\gamma^{*\prime}(s))ds + v^0(\gamma^*(0))\Big) + \beta_7\Delta x.$$

Hence, noting that $\sum_{\Omega_n^{l+1,0}} = \sum_{\Omega^+} + \sum_{\Omega_n^{l+1,0}\setminus\Omega^+}$ in (3.11), we obtain

$$O(\sqrt{\Delta x}) \geq prob(\Omega^+)\Delta x^{\frac{1}{4}} + prob(\Omega_n^{l+1,0}\setminus\Omega^+)(-\beta_7\Delta x) \geq prob(\Omega^+)\Delta x^{\frac{1}{4}} - \beta_7\Delta x,$$

which yields $prob(\Omega^+) = O(\Delta x^{\frac{1}{4}})$. Since γ^* is the unique minimizing curve, it follows from Lemma 3.5 that $\Omega_n^{l+1,0}\setminus\Omega^+ \subset \tilde{\Omega}_\Delta^\varepsilon$ for $\Delta x \ll \varepsilon$, which means that $prob(\tilde{\Omega}_\Delta^\varepsilon) \to 1$ as $\Delta \to 0$. Therefore, we obtain the convergence $E_{\mu_n^{l+1,0}(\cdot;\xi^*)}[\|\eta_\Delta(\gamma) - \gamma^*\|^2_{L^2([0,t])}] \to 0$ as $\Delta \to 0$, which implies that the left-hand side of (3.10) tends to 0 as $\Delta \to 0$. From this, we see that for any $\varepsilon' > 0$ there exists $\delta(\varepsilon') > 0$ such that if $|\Delta| < \delta(\varepsilon')$ we have

$$E_{\mu_n^{l+1,0}(\cdot;\xi^*)}[\|\gamma_\Delta - (\gamma^* + \gamma_\Delta(t) - x)\|_{L^2([0,t])}] \leq \varepsilon'.$$

Define $\Omega^{++} := \{\gamma \in \Omega_n^{l+1,0} \mid \|\gamma_\Delta - (\gamma^* + \gamma_\Delta(t) - x)\|_{L^2([0,t])} \geq \sqrt{\varepsilon'}\}$. Then, we have $prob(\Omega^{++}) \leq \sqrt{\varepsilon'}$. By Lemma 3.6 (note that $\gamma_\Delta(t) - (\gamma^*(t) + \gamma_\Delta(t) - x) = 0$), we obtain $\|\gamma_\Delta - (\gamma^* + \gamma_\Delta(t) - x)\|_{C^0([0,t])} \leq O(\varepsilon'^{\frac{1}{4}})$ for all $\gamma \in \Omega_n^{l+1,0}\setminus\Omega^{++}$. If $\varepsilon' \ll \varepsilon^4$, we have $\|\gamma_\Delta - \gamma^*\|_{C^0([0,t])} \leq \varepsilon$ for all $\gamma \in \Omega_n^{l+1,0}\setminus\Omega^{++}$ with $|\Delta| < \delta(\varepsilon')$. Therefore, it holds that, for any $\varepsilon' > 0$ with $\varepsilon' \ll \varepsilon$, there exists $\delta(\varepsilon') > 0$ such that if $|\Delta| < \delta(\varepsilon')$ we have $\Omega_n^{l+1,0}\setminus\Omega^{++} \subset \Omega_\Delta^\varepsilon$ and $prob(\Omega_\Delta^\varepsilon) \geq 1 - \sqrt{\varepsilon'}$, which means $prob(\Omega_\Delta^\varepsilon) \to 1$ as $\Delta \to 0$.

Note that, if $x \in \mathbb{R}^d$ (a regular point), $m \in \mathbb{Z}^d$ are fixed and if $n \in \mathbb{Z}^d$ is taken so that $x \in \Theta(x_n)$ accordingly to Δ, the minimizing random walks for v_{n+m}^{l+1} converge to γ^* in the same way as the above, because $m\Delta x \to 0$ as $\Delta x \to 0$.

We prove (3). Let (x,t) be an arbitrary regular point with $t > 0$. Let γ^* be the unique minimizing curve for $v(x,t)$. We have for $j = 1,\ldots,d$,

$$v_{x^j}(x,t) = \int_0^t L_{x^j}(\gamma^*(s), s, \gamma^{*\prime}(s))ds + v^0_{x^j}(\gamma^*(0)). \tag{3.12}$$

Let $(x_{n+\mathbf{1}}, t_{l+1}) \in \mathcal{G}$ be a point such that $t \in [t_{l+1}, t_{l+2})$ and $x \in \Theta(x_{n+\mathbf{1}})$. Then, $u^j_\Delta(x,t) = (D_{x^j}v)^{l+1}_{n+\mathbf{1}}$. By (4) of Theorem 2.1, we have

$$\begin{aligned}
(D_{x^j}v)^{l+1}_{n+\mathbf{1}} &\leq E_{\mu^{l+1,0}_{n+\mathbf{1}-e_j}(\cdot;\xi^*(-e_j))}\Big[\sum_{0<k\leq l+1} L_{x^j}(\gamma^k, t_{k-1}, \xi^k_{m(\gamma^k)}(-e_j))\Delta t \\
&\quad + (D_{x^j}v)^0_{m(\gamma^0)+e_j}\Big] + \theta\Delta x \\
&\leq E_{\mu^{l+1,0}_{n+\mathbf{1}-e_j}(\cdot;\xi^*(-e_j))}\Big[\int_0^t L_{x^j}(\eta_\Delta(\gamma)(s), s, \eta_\Delta(\gamma)'(s))ds \\
&\quad + (D_{x^j}v)^0_{m(\gamma^0)+e_j}\Big] + \beta_8\sqrt{\Delta x}.
\end{aligned}$$

Hence, we have

$$\begin{aligned}
&(D_{x^j}v)^{l+1}_{n+\mathbf{1}} - v_{x^j}(x,t) \leq R_1 + R_2 + \beta_8\sqrt{\Delta x}, \\
&R_1 = E_{\mu^{l+1,0}_{n+\mathbf{1}-e_j}(\cdot;\xi^*(-e_j))}\Big[\int_0^t \big(L_{x^j}(\eta_\Delta(\gamma)(s), s, \eta_\Delta(\gamma)'(s)) \\
&\qquad\qquad - L_{x^j}(\gamma^*(s), s, \gamma^{*\prime}(s))\big)ds\Big], \\
&R_2 = E_{\mu^{l+1,0}_{n+\mathbf{1}-e_j}(\cdot;\xi^*(-e_j))}\Big[(D_{x^j}v)^0_{m(\gamma^0)+e_j} - v^0_{x^j}(\gamma^*(0))\Big].
\end{aligned}$$

We already know from the proof of (2) that, for each j and any $\varepsilon > 0$, we have

$$Prob(\{\gamma \in \Omega^{l+1,0}_{n+\mathbf{1}-e_j} \mid \|\eta_\Delta(\gamma) - \gamma^*\|_{C^0([0,t])} > \varepsilon\}) \to 0 \quad \text{as } \Delta \to 0,$$
$$Prob(\{\gamma \in \Omega^{l+1,0}_{n+\mathbf{1}-e_j} \mid \|\eta_\Delta(\gamma)' - \gamma^{*\prime}\|_{L^2([0,t])} > \varepsilon\}) \to 0 \quad \text{as } \Delta \to 0,$$

which yields $R_1 \to 0$ as $\Delta \to 0$. Since v^0 is semiconcave, v^0 is differentiable at $\gamma^*(0)$. Furthermore, due to Lemma 3.7 below, we have

$$\sup_{\{x\in\mathbb{R}^d \mid \ \|x-\gamma^*(0)\|_\infty \le \varepsilon\}} |v^0_{x^j}(x) - v^0_{x^j}(\gamma^*(0))| \to 0 \quad \text{as } \varepsilon \to 0,$$

where "sup" is taken with respect to x such that the partial derivative $v^0_{x^j}(x)$ exists (we do not ask the derivative of v^0). By (2), we have for any $\varepsilon > 0$,

$$Prob(\{\gamma \in \Omega^{l+1,0}_{n+\mathbf{1}-e_j} \mid \|\gamma^0 - \gamma^*(0)\|_\infty > \varepsilon\}) \to 0 \quad \text{as } \Delta \to 0.$$

Therefore, noting that the map $[0, 2\Delta x] \ni s \mapsto v^0(\gamma^0 + se_j)$ is Lipschitz and a.e. differentiable, we have

$$(D_{x^j}v)^0_{m(\gamma^0)+e_j} \quad = \quad \frac{v^0(\gamma^0 + 2\Delta x \cdot e_j) - v^0(\gamma^0)}{2\Delta x} = \frac{1}{2\Delta x}\int_0^{2\Delta x} v^0_{x^j}(\gamma^0 + se_j)ds,$$

which yields $R_2 \to 0$ as $\Delta \to 0$. Thus, we conclude that

$$\limsup_{\Delta\to 0}(u^j_\Delta(x,t) - v_{x^j}(x,t)) \le 0.$$

Similar reasoning yields

$$\liminf_{\Delta\to 0}(u^j_\Delta(x,t) - v_{x^j}(x,t)) \ge 0. \qquad \square$$

Lemma 3.7. *Let $v^0 \in Lip(\mathbb{R}^d)$ be locally semiconcave with a linear modulus. Let $x_* \in \mathbb{R}^d$ be a point of differentiability of v^0 with the derivative $v^0_x(x_*) = (v^0_{x^1}(x_*), \dots, v^0_{x^d}(x_*))$. Suppose that, for a sequence $\{x_j\}_{j\in\mathbb{N}}$, $x_j \to x_*$ as $j \to \infty$, there exists the partial derivative $v^0_{x^i}(x_j)$ for some i for all $j \in \mathbb{N}$. Then, we have $v^0_{x^i}(x_j) \to v^0_{x^i}(x_*)$ as $j \to \infty$.*

Proof. Let $\alpha > 0$ be a modulus of semiconcavity in a neighborhood A of x_*. Then, $w(x) := \alpha|x|^2 - v^0(x)$ is convex in A. Therefore, it is enough to prove the assertion for a convex function $w \in Lip(A)$ instead of v^0. Set $\tilde{w}(x) := w(x) - w(x_*) - w_x(x_*) \cdot (x - x_*)$, which is still Lipschitz, convex and differentiable at x_* with $\tilde{w}_x(x_*) = 0$. We proceed by contradiction. Suppose that $\tilde{w}_{x^i}(x_j)$ does not converge to 0 as $j \to \infty$. Since $\{\tilde{w}_{x^i}(x_j)\}_{j\in\mathbb{N}}$ is bounded, there exists a subsequence, still denoted by $\{\tilde{w}_{x^i}(x_j)\}_{j\in\mathbb{N}}$, that converges to a number $a \ne 0$ as $j \to \infty$. Since the map $(-1, 1) \ni s \mapsto \tilde{w}(x_j + se_i)$ is Lipschitz, it is a.e. differentiable. Hence, we have for $t \in (-1, 1)$,

$$\tilde{w}(x_j + te_i) = \tilde{w}(x_j) + \int_0^t \tilde{w}_{x^i}(x_j + se_i)ds.$$

Since $\tilde{w}$ is differentiable at x_*, we have

$$\tilde{w}(x) = \tilde{w}(x_*) + \tilde{w}_x(x_*) \cdot (x - x_*) + o(|x - x_*|) = o(|x - x_*|).$$

Therefore, we have

$$\tilde{w}(x_j + te_i) \quad = \quad o(|x_j + te_i - x_*|) = o(|x_j - x_*|) + \int_0^t \tilde{w}_{x^i}(x_j + se_i)ds.$$

Set $t_j := \mathrm{sign}(a)|x_j - x_*|$ and insert $t = t_j$. Then, we have

$$o(|x_j - x_*|) = \int_0^{t_j} \tilde{w}_{x^i}(x_j + se_i)ds.$$

Since the map $s \mapsto \tilde{w}(x_j + se_i)$ is convex by assumption, $\tilde{w}_{x^i}(x_j + se_i)$ is non-decreasing with respect to s. Therefore, for $0 < \varepsilon \ll |a|$, we have if $a > 0$,

$$a - \varepsilon \le \tilde{w}_{x^i}(x_j) \le \tilde{w}_{x^i}(x_j + se_i) \quad \text{for all } s \in [0, t_j] \text{ with sufficiently large } j$$

and if $a < 0$,

$$\tilde{w}_{x^i}(x_j + se_i) \le \tilde{w}_{x^i}(x_j) \le a + \varepsilon \quad \text{for all } s \in [t_j, 0] \text{ with sufficiently large } j.$$

Thus, we obtain

$$o(|x_j - x_*|) = \Big| \int_0^{t_j} \tilde{w}_{x^i}(x_j + se_i)ds \Big| \ge |a \mp \varepsilon||t_j| = |a \mp \varepsilon||x_j - x_*|,$$

which is a contradiction. □

We conclude this section with discussion on Remark 2.3. Currently, we fail to obtain (3) of Theorem 2.2 without semiconcavity of initial data, if $d \ge 2$. We will see the difficulty and how to overcome it for $d = 1$. Without semiconcavity, uncountably many minimizing curves $\gamma^*(s)$ meet at one point at $s = 0$ (so-called "rarefaction") and $v^0_x(\gamma^*(0))$ does not exist for such γ^*. In such a case, we may use the following formula: For any $\tau \in [0, t)$,

$$v_{x^j}(x,t) = \int_\tau^t L_{x^j}(\gamma^*(s), s, \gamma^{*\prime}(s))ds + L_{\xi^j}(\gamma^*(\tau), \tau, \gamma^{*\prime}(\tau)). \tag{3.13}$$

We have a discrete version of (3.13): We observe that for each e_j,

$$\begin{aligned} v^{l+1}_{n+\mathbf{1}-e_j} &= \sum_{\gamma \in \Omega^{l+1,0}_{n+\mathbf{1}-e_j}} \mu^{l+1,0}_{n+\mathbf{1}-e_j}(\gamma; \xi^*|_{G^{l+1,1}_{n+\mathbf{1}-e_j}}) \Big(\sum_{0<k\le l+1} L(\gamma^k, t_{k-1}, {\xi^*}^k_{m(\gamma^k)})\Delta t \\ &\quad + v^0(\gamma^0) \Big) + ht_{l+1} \\ &= \sum_{\gamma \in \Omega^{l+1,k(\tau)}_{n+\mathbf{1}-e_j}} \mu^{l+1,k(\tau)}_{n+\mathbf{1}-e_j}(\gamma; \xi^*|_{G^{l+1,k(\tau)+1}_{n+\mathbf{1}-e_j}}) \Big(\sum_{k(\tau)<k\le l+1} L(\gamma^k, t_{k-1}, {\xi^*}^k_{m(\gamma^k)})\Delta t \\ &\quad + v^{k(\tau)}_{m(\gamma^{k(\tau)})} \Big) + h(t_{l+1} - t_{k(\tau)}). \end{aligned}$$

For each e_j, define the control ζ on $G^{l+1,1}_{n+\mathbf{1}+e_j}$ as

$$\zeta(x_m, t_{k+1}) := \begin{cases} \xi^*(x_m - e_j \cdot 2\Delta x, t_{k+1}) & \text{for } k(\tau) < k+1 \le l+1, \\ \xi^*(x_m, t_{k+1}) & \text{for } 0 < k+1 \le k(\tau). \end{cases}$$

Then, we have

$$
\begin{aligned}
v^{l+1}_{n+\mathbf{1}+e_j} \le & \sum_{\gamma\in\Omega^{l+1,0}_{n+\mathbf{1}+e_j}} \mu^{l+1,0}_{n+\mathbf{1}+e_j}(\gamma;\zeta)\Big(\sum_{0<k\le l+1} L(\gamma^k,t_{k-1},\zeta^k_{m(\gamma^k)})\Delta t + v^0(\gamma^k)\Big) + ht_{l+1}\\
= & \sum_{\gamma\in\Omega^{l+1,k(\tau)}_{n+\mathbf{1}-e_j}} \mu^{l+1,k(\tau)}_{n+\mathbf{1}-e_j}(\gamma;\xi^*|_{G^{l+1,k(\tau)+1}_{n+\mathbf{1}-e_j}})\\
& \times\Big(\sum_{k(\tau)<k\le l+1} L(\gamma^k+e_j\cdot 2\Delta x, t_{k-1},{\xi^*}^k_{m(\gamma^k)})\Delta t + v^{k(\tau)}_{m(\gamma^{k(\tau)}+e_j\cdot 2\Delta x)}\Big)\\
& +h(t_{l+1}-t_{k(\tau)}).
\end{aligned}
$$

Hence, with Lemma 3.2, we obtain

$$
\begin{aligned}
(D_{x^j}v)^{l+1}_{n+\mathbf{1}} \quad = \quad & \frac{v^{l+1}_{n+\mathbf{1}+e_j}-v^{l+1}_{n+\mathbf{1}-e_j}}{2\Delta x}\\
\le \quad & \sum_{\gamma\in\Omega^{l+1,k(\tau)}_{n+\mathbf{1}-e_j}} \mu^{l+1,k(\tau)}_{n+\mathbf{1}-e_j}(\gamma;\xi^*|_{G^{l+1,k(\tau)+1}_{n+\mathbf{1}-e_j}})\\
& \times\Big(\sum_{k(\tau)<k\le l+1} L_{x^j}(\gamma^k,t_{k-1},{\xi^*}^k_{m(\gamma^k)})\Delta t + (D_{x^j}v)^{k(\tau)}_{m(\gamma^{k(\tau)}+e_j\Delta x)}\Big)\\
& +\beta_8\Delta x\\
\le \quad & \sum_{\gamma\in\Omega^{l+1,k(\tau)}_{n+\mathbf{1}-e_j}} \mu^{l+1,k(\tau)}_{n+\mathbf{1}-e_j}(\gamma;\xi^*|_{G^{l+1,k(\tau)+1}_{n+\mathbf{1}-e_j}})\\
& \times\Big(\int_\tau^t L_{x^j}(\eta_\Delta(\gamma)(s),s,\eta_\Delta(\gamma)'(s))ds + (D_{x^j}v)^{k(\tau)}_{m(\gamma^{k(\tau)}+e_j\Delta x)}\Big)\\
& +\beta_9\sqrt{\Delta x}.
\end{aligned}
$$

Now, we compare this inequality and (3.13). Since we already know about the convergence of the minimizing random walks γ and $\eta(\gamma)$ to γ^*, it is enough to estimate

$$
\begin{aligned}
(3.14)\quad & \sum_{\gamma\in\Omega^{l+1,k(\tau)}_{n+\mathbf{1}-e_j}} \mu^{l+1,k(\tau)}_{n+\mathbf{1}-e_j}(\gamma;\xi^*|_{G^{l+1,k(\tau)+1}_{n+\mathbf{1}-e_j}})\Big((D_{x^j}v)^{k(\tau)}_{m(\gamma^{k(\tau)}+e_j\Delta x)}\Big)\\
& \qquad -L_{\xi^j}(\gamma^*(\tau),\tau,{\gamma^*}'(\tau)))\\
& = \sum_{\gamma\in\Omega^{l+1,k(\tau)}_{n+\mathbf{1}-e_j}} \mu^{l+1,k(\tau)}_{n+\mathbf{1}-e_j}(\gamma;\xi^*|_{G^{l+1,k(\tau)+1}_{n+\mathbf{1}-e_j}})\\
& \qquad \times\Big(L_{\xi^j}(\gamma^{k(\tau)}+e_j\Delta x,t_{k(\tau)},{\xi^*}^{k(\tau)+1}_{m(\gamma^{k(\tau)}+e_j\Delta x)}) - L_{\xi^j}(\gamma^*(\tau),\tau,{\gamma^*}'(\tau))\Big),
\end{aligned}
$$

where we note that $(D_{x^j}v)^k_m = L_{\xi^j}(x_m,t_k,{\xi^*}^{k+1}_m)$. Unfortunately, ${\xi^*}^{k(\tau)+1}_{m(\gamma^{k(\tau)}+e_j\Delta x)}$ is NOT equal to ${\xi^*}^{k(\tau)+1}_{m(\gamma^{k(\tau)+1})}$, i.e., $\gamma^{k(\tau)+1}$ is not necessarily equal to $\gamma^{k(\tau)}+e_j\Delta x$, which prevents us from direct application of the L^2-convergence of $\eta_\Delta(\gamma)'(\cdot)$ to ${\gamma^*}'(\cdot)$.

If $d=1$, we manage to estimate (3.14) by means of the one-sided Lipschitz estimate of the discrete derivative, or the entropy condition, proved in Proposition 2.8 of [10] (this holds also in our whole space setting): There exists $M^k>0$

independent of Δ such that

$$\frac{(D_xv)^k_{m+2}-(D_xv)^k_m}{2\Delta x}\le M^k.$$

Let Ω_x be the set of all $\gamma\in\Omega_n^{l+1,k(\tau)+1}$ such that $\gamma^{k(\tau)+1}=x$. We have

$$\begin{aligned}
&\sum_{\gamma\in\Omega_n^{l+1,k(\tau)}}\mu_n^{l+1,k(\tau)}(\gamma;\xi^*|_{G_n^{l+1,k(\tau)+1}})\Big((D_xv)^{k(\tau)}_{m(\gamma^{k(\tau)}+\Delta x)}-(D_xv)^{k(\tau)}_{m(\gamma^{k(\tau)+1})}\Big)\\
&=\sum_{x\in X_n^{l+1,k(\tau)+1}}\sum_{\gamma\in\Omega_x}\mu_n^{l+1,k(\tau)+1}(\gamma;\xi^*|_{G_n^{l+1,k(\tau)+2}})\Big\{\rho^{k(\tau)+1}_{m(x)}(+1)\Big((D_xv)^{k(\tau)}_{m(x+\Delta x+\Delta x)}\\
&\quad-(D_xv)^{k(\tau)}_{m(x)}\Big)+\rho^{k(\tau)+1}_{m(x)}(-1)\Big((D_xv)^{k(\tau)}_{m(x-\Delta x+\Delta x)}-(D_xv)^{k(\tau)}_{m(x)}\Big)\Big\}\\
&\le M^{k(\tau)}\cdot 2\Delta x.
\end{aligned}$$

Hence, we obtain

$$\begin{aligned}
(3.14)|_{d=1}\ \le&\sum_{\gamma\in\Omega_n^{l+1,k(\tau)}}\mu_n^{l+1,k(\tau)}(\gamma;\xi^*|_{G_n^{l+1,k(\tau)+1}})(D_{x^j}v)^{k(\tau)}_{m(\gamma^{k(\tau)+1})}+M^{k(\tau)}\cdot 2\Delta x\\
&-L_\xi(\gamma^*(\tau),\tau,\gamma^{*\prime}(\tau))\\
=&\sum_{\gamma\in\Omega_n^{l+1,k(\tau)}}\mu_n^{l+1,k(\tau)}(\gamma;\xi^*|_{G_n^{l+1,k(\tau)+1}})\Big(L_\xi(\gamma^{k(\tau)+1},t_{k(\tau)},{\xi^*}^{k(\tau)+1}_{m(\gamma^{k(\tau)+1})})\\
&-L_\xi(\gamma^*(\tau),\tau,\gamma^{*\prime}(\tau))\Big)+M^{k(\tau)}\cdot 2\Delta x\\
=&\sum_{\gamma\in\Omega_n^{l+1,0}}\mu_n^{l+1,0}(\gamma;\xi^*|_{G_n^{l+1,1}})\Big(L_\xi(\gamma^{k(\tau)+1},t_{k(\tau)},{\xi^*}^{k(\tau)+1}_{m(\gamma^{k(\tau)+1})})\\
&-L_\xi(\gamma^*(\tau),\tau,\gamma^{*\prime}(\tau))\Big)+M^{k(\tau)}\cdot 2\Delta x.
\end{aligned}$$

According to the above proof of (2) of Theorem 2.2, for any $\varepsilon>0$, there exists $\delta>0$ such that if $|\Delta|<\delta$ we have

$$\begin{aligned}
&E_{\mu_n^{l+1,0}(\cdot;\xi^*|_{G_n^{l+1,1}})}[\|\eta_\Delta(\gamma)'-\gamma^{*\prime}\|_{L^1([0,t])}]\\
&\quad\le E_{\mu_n^{l+1,0}(\cdot;\xi^*|_{G_n^{l+1,1}})}[\sqrt{t}\|\eta_\Delta(\gamma)'-\gamma^{*\prime}\|_{L^2([0,t])}]<\varepsilon.
\end{aligned}$$

Therefore, there exists $\tau\in[t/2,t]$ such that

$$\begin{aligned}
&E_{\mu_n^{l+1,0}(\cdot;\xi^*|_{G_n^{l+1,1}})}[|\eta_\Delta(\gamma)'(\tau)-\gamma^{*\prime}(\tau)|]\\
&\quad=E_{\mu_n^{l+1,0}(\cdot;\xi^*|_{G_n^{l+1,1}})}[|{\xi^*}^{k(\tau)+1}_{m(\gamma^{k(\tau)+1})}-\gamma^{*\prime}(\tau)|]<2\varepsilon/t.
\end{aligned}$$

Since $M^{k(\tau)}$ with $\tau\ge t/2$ is bounded for $\Delta\to0$, we obtain $\limsup_{\Delta\to0}(3.14)|_{d=1}=0$ and $\limsup_{\Delta\to0}(u_\Delta(x,t)-v_x(x,t))=0$. Similarly, we obtain $\liminf_{\Delta\to0}(u_\Delta(x,t)-v_x(x,t))=0$.

Acknowledgment

The author thanks Norihisa Ikoma for a discussion on Lemma 3.7.

References

[1] A. Bouillard, E. Faou, and M. Zavidovique, *Fast weak-KAM integrators for separable Hamiltonian systems*, Math. Comp. **85** (2016), no. 297, 85–117, DOI 10.1090/mcom/2976. MR3404444

[2] P. Cannarsa and C. Sinestrari, *Semiconcave Functions, Hamilton-Jacobi Equations, and Optimal Control*, Progress in Nonlinear Differential Equations and their Applications, vol. 58, Birkhäuser Boston, Inc., Boston, MA, 2004. MR2041617

[3] M. G. Crandall and P.-L. Lions, *Two approximations of solutions of Hamilton-Jacobi equations*, Math. Comp. **43** (1984), no. 167, 1–19, DOI 10.2307/2007396. MR744921

[4] A. Fathi, *Weak KAM Theorem in Lagrangian Dynamics*, Cambridge University Press (2011).

[5] W. H. Fleming, *The Cauchy problem for a nonlinear first order partial differential equation*, J. Differential Equations **5** (1969), 515–530, DOI 10.1016/0022-0396(69)90091-6. MR0235269

[6] T. Nishida and K. Soga, *Difference approximation to Aubry-Mather sets of the forced Burgers equation*, Nonlinearity **25** (2012), no. 9, 2401–2422, DOI 10.1088/0951-7715/25/9/2401. MR2967111

[7] O. A. Oleĭnik, *Discontinuous solutions of non-linear differential equations*, Amer. Math. Soc. Transl. (2) **26** (1963), 95–172, DOI 10.1090/trans2/026/05. MR0151737

[8] K. Soga, *Space-time continuous limit of random walks with hyperbolic scaling*, Nonlinear Anal. **102** (2014), 264–271, DOI 10.1016/j.na.2014.02.012. MR3182814

[9] K. Soga, *Stochastic and variational approach to the Lax-Friedrichs scheme*, Math. Comp. **84** (2015), no. 292, 629–651, DOI 10.1090/S0025-5718-2014-02863-9. MR3290958

[10] K. Soga, *More on stochastic and variational approach to the Lax-Friedrichs scheme*, Math. Comp. **85** (2016), no. 301, 2161–2193, DOI 10.1090/mcom/3061. MR3511278

[11] K. Soga, *Selection problems of $\mathbb{Z}^2$-periodic entropy solutions and viscosity solutions*, Calc. Var. Partial Differential Equations **56** (2017), no. 4, Art. 119, 30, DOI 10.1007/s00526-017-1208-7. MR3672390

[12] P. E. Souganidis, *Approximation schemes for viscosity solutions of Hamilton-Jacobi equations*, J. Differential Equations **59** (1985), no. 1, 1–43, DOI 10.1016/0022-0396(85)90136-6. MR803085

Department of Mathematics, Faculty of Science and Technology, Keio University, 3-14-1 Hiyoshi, Kohoku-ku, Yokohama, 223-8522, Japan

Email address: soga@math.keio.ac.jp