SIAM J. CONTROL OPTIM.
Vol. 63, No. 1, pp. 375–401

POLICY ITERATION FOR THE DETERMINISTIC CONTROL PROBLEMS—A VISCOSITY APPROACH*

WENPIN TANG†, HUNG VINH TRAN‡, AND YUMING PAUL ZHANG§

Abstract. This paper is concerned with the convergence rate of policy iteration for (deterministic) optimal control problems in continuous time. To overcome the problem of ill-posedness due to lack of regularity, we consider a semidiscrete scheme by adding a viscosity term via finite differences in space. We prove that the policy iteration (PI) for the semidiscrete scheme converges exponentially fast and provide a bound on the error induced by the semidiscrete scheme. We also consider the discrete space-time scheme, where both space and time are discretized. The convergence rate of PI and the discretization error are studied.

Key words. finite differences, Hamilton–Jacobi–Bellman equations, optimal control, policy iteration

MSC codes. 49L25, 65M06, 93E20

DOI. 10.1137/24M1631602

1. Introduction.

Optimal control is ubiquitous in science and engineering with a variety of applications including aerospace engineering [6, 10], chemical engineering [38], economy [29], operations research [45, 48] and robotics [2, 15]. Dynamic programming (DP) has proved to be an efficient tool for solving multistage optimal control problems since its inception by Bellman [5]. In recent years, reinforcement learning (RL) has shown great success in resolving complex decision-making problems, notably AlphaGo [49] and humanoid tasks [22]. Policy iteration (PI), as a class of approximate or adaptive dynamic programming (ADP), is instrumental in many RL algorithms [51].

The idea of PI dates back to Howard [24] in a stochastic environment known as the Markov decision process (MDP). Subsequent works [7, 40, 41] explored PI for MDPs in discrete time and space; recently, [8, 36] considered PI for (deterministic) optimal control problems in discrete time and continuous space. In these works, PIs are proved to converge to the optimal control under suitable conditions on the model parameters. Furthermore, [42, 47] studied the convergence rate of PI for infinite horizon MDP. On the other hand, many real-world problems are modeled by dynamical systems evolving in continuous time, and it is known that DP for optimal control in continuous time and space entails the Hamilton–Jacobi–Bellman (HJB) partial differential equation

*Received by the editors January 16, 2024; accepted for publication (in revised form) October 27, 2024; published electronically January 24, 2025.
https://doi.org/10.1137/24M1631602
Funding: The work of the first author is supported by NSF grant DMS-2206038, the Columbia Innovation Hub grant, and the Tang Family Assistant Professorship. The work of the second author is partially supported by NSF CAREER grant DMS-1843320, a Simons Fellowship, and a Vilas Faculty Early-Career Investigator Award. The work of the third author is partially supported by Simons Foundation Travel Support MPS-TSM-00007305 and a start-up grant at Auburn University.
†Department of Industrial Engineering and Operations Research, Columbia University, New York, NY 10027 USA (wt2319@columbia.edu).
‡Department of Mathematics, University of Wisconsin Madison, Madison, WI 53706 USA (hung@math.wisc.edu).
§Department of Mathematics and Statistics, Auburn University, Auburn, AL 36849 USA (yzhangpaul@auburn.edu).

(PDE). Despite its importance, PI for optimal control problems in continuous time and space has mostly been studied in the linear quadratic setting [32, 55] or those with a specific structure that allows solvability to some extent [1]. It was not until recently that the general space-time problems were considered in [34] under a fixed point assumption. For the stochastic control problems, [31, 44] showed that PI converges exponentially fast in the case where controls are only exercised on the drift term of the state process. Similar results were derived for the corresponding entropy-regularized problems [26, 53]. Recently, PI for mean field games was considered in [11, 13, 14]. We also mention that, in a closely related direction, [9, 57] studied value iteration for optimal control problems. References [28, 37] proposed differential dynamic programming. It relies on a quadratic approximation to the value function, which requires the second-order property of the model parameters. See [35, 56] for recent progress on the theory and applications of ADP for optimal control and RL.

In this paper, we study the convergence rate of PI for optimal control problems in continuous time and their discretization under general conditions on the model parameters. We will assume that the cost function, the control, and the vector field that controls the system's state are all uniformly bounded and Lipschitz continuous. However, some of our results hold under more general assumptions (see Remark 3.4). Note that the convergence analysis in [1, 32, 55] relies on the specific structure of the problem, while [34] assumed that the HJB operator enjoys a fixed point or a contraction property, which is hard to verify. None of these works quantified the convergence of PI to the optimal control. Moreover, PI for continuous-time control problems may even be ill-posed due to lack of regularity. Our idea is to introduce a viscosity term "$h\Delta^h$" in the policy evaluation, where h is the mesh size and Δ^h is the discrete Laplacian in space. We call it a *semidiscrete scheme*. Essentially, the viscosity term is of order 1, which ensures that the finite difference scheme is monotone. A monotone scheme is commonly desirable for numerical implementation, so the addition of the finite difference viscosity term is natural. On the other hand, the viscosity term in the semidiscrete scheme mimics the vanishing viscosity approximation to first-order PDEs [20], which forces PI to converge exponentially fast (Theorem 3.1 and Theorem 3.3), as for the stochastic control problems. We also prove that the discrepancy between the optimal control problem and its semidiscrete scheme is of order $\sqrt{h}$ as $h \to 0$ (Theorem 3.5). If further assuming the cost function and the vector field to be uniformly bounded in $W^{2,\infty}$ in space, then the policy in PI converges almost everywhere (Theorem 3.7). Furthermore, we consider the time-discretization, called a *discrete space-time scheme*. The same results hold for PI for the discrete space-time scheme (Theorem 4.1 and Theorem 4.2). Our results echo recent work [23], which asserts that noise enhances the convergence of finite horizon RL algorithms. In our setting, noise corresponds to the viscosity term, and the importance of a finite horizon is seen from various bounds with exponential dependence in time. Our analysis relies on PDE techniques (which are also useful in analyzing vanishing viscosity approximations for mean field games [52]) and may carry over to the study of differential games in solving Hamilton–Jacobi–Bellman–Issacs (HJBI) equations.

To the best of our knowledge, the exponential convergence results in Theorems 3.1, 3.3, and 4.1 are new in the literature and they are essentially optimal. For the quantitative convergence of the solutions to the semidiscrete scheme and the discrete space-time scheme to these of the continuous equations in Theorems 3.5 and 4.2, we follow the approach of Crandall and Lions [16]. Note that [16] does not deal with PI and approximated optimal policies.

The rest of the paper is organized as follows. In section 2, we provide background and present the semidiscrete and the discrete space-time schemes. In section 3, we study the semidiscrete scheme, and in section 4, we analyze the discrete space-time scheme. We provide further PDE perspectives in section 5. We conclude with section 7.

2. Setup and preliminary results. In this section, we present the semidiscrete and the discrete space-time schemes. Consider a system whose state is governed by the ordinary differential equation

$$\frac{dx(t)}{dt} = f(t, x(t), \alpha(t)), \tag{2.1}$$

where, for $0 \le t \le T$, $x(t) \in \mathbb{R}^d$ is the system state, $\alpha(t) \in A \subset \mathbb{R}^m$ is the control or policy, and $f : [0,T] \times \mathbb{R}^d \times A \to \mathbb{R}^d$ is Lipschitz continuous. Here, A is a given compact subset of $\mathbb{R}^m$. The objective is

$$J(t,x,\alpha) := \int_t^T c(s, x(s), \alpha(s))\, ds + q(x(T)) \quad \text{given } x(t) = x, \tag{2.2}$$

and the goal is to minimize this objective function. Denote by

$$v_*(t,x) := \inf_{\alpha \in \mathcal{A}_t} J(t,x,\alpha), \tag{2.3}$$

where $\mathcal{A}_t$ is the standard admissible policy defined as $\mathcal{A}_t = \{\alpha : [t,T] \to A : \alpha \text{ is measurable}\}$. It is known that, under suitable conditions on $c(\cdot)$ and $q(\cdot)$ (see [21, Chapter 2] or [54, Chapter 2]), v_* defined by (2.3) is the viscosity solution to

$$\begin{cases} \partial_t v(t,x) + H(t,x,\nabla v(t,x)) = 0 & \text{in } (0,T) \times \mathbb{R}^d, \\ v(T,x) = q(x) & \text{on } \mathbb{R}^d, \end{cases} \tag{2.4}$$

where the Hamiltonian $H : [0,T] \times \mathbb{R}^d \times \mathbb{R}^d \to \mathbb{R}$ is given by

$$H(t,x,p) := \inf_{a \in A} \left[c(t,x,a) + p \cdot f(t,x,a) \right].$$

We assume the above infimum is achieved at a unique $a \in A$. Denote by

$$\alpha(t,x,p) := \operatorname*{arg\,min}_{a \in A} \left[c(t,x,a) + p \cdot f(t,x,a) \right]. \tag{2.5}$$

The optimal policy is given by

$$\alpha_*(t,x) = \alpha(t,x,\nabla v_*(t,x)). \tag{2.6}$$

We impose the following assumptions.

(A1) $c(\cdot,\cdot,\cdot), f(\cdot,\cdot,\cdot), q(\cdot)$ are uniformly bounded and Lipschitz continuous in all of their dependencies.

(A2) $\alpha(\cdot,\cdot,\cdot)$, the unique solution to (2.5), is uniformly Lipschitz continuous in all of its dependencies on $[0,T] \times \mathbb{R}^d \times A$.

Condition (A2) is restrictive, which is required to ensure the well-posedness and regularity properties of the PI algorithm. It is hard to relax this condition because the control α appears directly in PI.

PI is an ADP that alternates between policy evaluation to get the value function with the current control and policy improvement to optimize the value function. More precisely, for $n = 0, 1, \ldots$, the iterative procedure is as follows:

- Given $\alpha_n(t,x)$, solve the linear PDE

$$\begin{cases} \partial_t v_n(t,x) + c(t,x,\alpha_n(t,x)) + \nabla v_n(t,x) \cdot f(t,x,\alpha_n(t,x)) = 0 & \text{in } (0,T)\times\mathbb{R}^d, \\ v_n(T,x) = q(x) & \text{on } \mathbb{R}^d. \end{cases} \tag{2.7}$$

- Set

$$\alpha_{n+1}(t,x) = \alpha(t,x,\nabla v_n(t,x)) = \underset{a\in A}{\arg\min}\left[c(t,x,a) + \nabla v_n(t,x)\cdot f(t,x,a)\right]. \tag{2.8}$$

The key is to understand how the sequence $\{v_n\}$ approximates the optimal value v_* and how $\{\alpha_n\}$ approximates the optimal policy α_*.

On the other hand, it is not clear whether the PI scheme (2.7) and (2.8) is well-posed. Intuitively, to make sense of α_{n+1}, we need v_n to be Lipschitz continuous, for which we then need α_n to be Lipschitz. This, in turn, requires ∇v_{n-1} to be Lipschitz. After iterations, we need v_0 to be smooth, which is not generally true.

Throughout the paper, we denote by $\mathbb{N}$ the set of all positive natural numbers and $\mathbb{Z}$ the set of all integers. For any $h>0$, we write $h\mathbb{Z}^d := \{hz \,|\, z\in\mathbb{Z}^d\}$. Let $\mathbb{R}^d$ be the Euclidean space of dimension d and $|\cdot|$ the Euclidean distance. For $R>0$, by B_R we mean the ball in $\mathbb{R}^d$ of radius R and centered at the origin. For a vector field $f:[0,T]\times\mathbb{R}^d\times A\to\mathbb{R}^d$, we denote its infinity norm by $\|f\|_\infty$. For a function $g:[0,T]\times\mathbb{R}^d\to\mathbb{R}$, the spatial gradient is denoted as $\nabla g(t,x) = \nabla_x g(t,x)$, and the partial derivative with respect to time is denoted as $\partial_t g(t,x)$.

We write C as various universal constants that only depend on d and the constants in (A1) and (A2) unless otherwise stated. Specifically, since T,h are not universal constants, we keep track of the dependence on T,h in most estimates. The constants C might vary from one line to another. By C_X or $C(X)$, we mean a constant that depends on universal constants and X.

2.1. Semidiscrete schemes. For $T>0$, $h\in(0,1)$, $N\geq\max\{1,\|f\|_\infty/2\}$, and a given Lipschitz continuous function $\alpha_0:\mathbb{R}\times\mathbb{R}^d\to A$, we solve for $n=0,1,\ldots$:

$$\begin{cases} \partial_t v_n^h(t,x) + c(t,x,\alpha_n(t,x)) + \nabla^h v_n^h(t,x)\cdot f(t,x,\alpha_n(t,x)) & \\ \qquad\qquad = -Nh\Delta^h v_n^h(t,x) & \text{in } (0,T)\times\mathbb{R}^d, \\ v_n^h(T,x) = q(x) & \text{on } \mathbb{R}^d. \end{cases} \tag{2.9}$$

Then, set

$$\alpha_{n+1}(t,x) = \alpha(t,x,\nabla^h v_n^h(t,x)) \qquad \text{in } (0,T)\times\mathbb{R}^d. \tag{2.10}$$

Here, for any $\varphi:\mathbb{R}^d\to\mathbb{R}$ and $h\in\mathbb{R}\setminus\{0\}$, we use the notations

$$\nabla^h\varphi(x) := \left(\frac{\varphi(x+he_1)-\varphi(x-he_1)}{2h},\ldots,\frac{\varphi(x+he_d)-\varphi(x-he_d)}{2h}\right),$$

$$\Delta^h\varphi(x) := \sum_{i=1}^d \frac{\varphi(x+he_i)-2\varphi(x)+\varphi(x-he_i)}{h^2}.$$

Later, we will also write $D^h\varphi(x) := \left(\frac{\varphi(x+he_1)-\varphi(x)}{h},\ldots,\frac{\varphi(x+he_d)-\varphi(x)}{h}\right)$. It is clear that

$$\nabla^h\varphi(x) = \frac{1}{2}\left(D^h\varphi(x) + D^{-h}\varphi(x)\right). \tag{2.11}$$

The assumption $N \geq \|f\|_\infty/2$ guarantees that the numerical Hamiltonian is monotone, and, as a consequence of this, the following comparison principle holds (see, e.g., [16, 39, 54]).

LEMMA 2.1. *Let v_0^h and $\tilde{v}_0^h$ be, respectively, a bounded continuous super- and subsolution to* (2.9) *with $n=0$ and satisfy $\tilde{v}_0^h \leq v_0^h$ at $t=T$. Then, $\tilde{v}_0^h \leq v_0^h$ in $[0,T]\times\mathbb{R}^d$. Here, by a supersolution (resp., subsolution), we mean that it satisfies* (2.9) *with the first equality replaced by $\leq$ (resp., $\geq$) and the second equality replaced by $\geq$ (resp., $\leq$).*

First, we show that the scheme (2.9) and (2.10) is well-posed.

PROPOSITION 2.2. *Assume* (A1) *and* (A2) *and that $N \geq \max\{1, \|f\|_\infty/2\}$. Then, the iterative process* (2.9) *and* (2.10) *is well-defined; that is, there are Lipschitz continuous functions v_n^h, α_n satisfying* (2.9) *and* (2.10), *and v_n^h are uniformly bounded for all $n \geq 0$ and $h > 0$.*

Proof. Since α_0 is Lipschitz continuous, the unique solvability of (2.9) for $n=0$ follows from [33, Theorem 2.4]. If one can show that v_0^h is uniformly bounded and Lipschitz continuous with Lipschitz constant C_h, then α_1 is Lipschitz continuous with Lipschitz constant C_h'/h for some $C_h'>0$ by the assumption that α is Lipschitz and

$$\begin{aligned}
&\left|\nabla^h v_0^h(t,x) - \nabla^h v_0^h(s,y)\right| \\
&\quad \leq \left|\left(\frac{v_0^h(t,x+he_1) - v_0^h(s,y+he_1) - v_0^h(t,x-he_1) + v_0^h(s,y-he_1)}{2h}, \dots\right)\right| \\
&\quad \leq h^{-1} C_h(|x-y| + |t-s|).
\end{aligned}$$

From the same argument, we obtain a unique bounded and Lipschitz solution v_1^h. The existence of solutions then follows from iterations.

First, we prove the boundedness of v_0^h. Since $c(\cdot,\cdot,\cdot), q(\cdot)$ are uniformly bounded, we have that $\pm[\|q\|_\infty + \|c\|_\infty(T-t)]$ are a supersolution and a subsolution to (2.9) with $n=0$, respectively. Hence, by Lemma 2.1,

$$-\|q\|_\infty - \|c\|_\infty(T-t) \leq v_0^h(t,x) \leq \|q\|_\infty + \|c\|_\infty(T-t)$$

for all $(t,x)\in[0,T]\times\mathbb{R}^d$. The same bound holds for all v_n^h by this argument.

Next, we show that v_0^h is Lipschitz continuous with Lipschitz constant independent of h when $T=T_0$ is sufficiently small depending only on the Lipschitz norms of c, f and α_0 presented in assumptions (A1) and (A2). The general result for any $T>0$ follows immediately by iterations and shifting in time on $[0,T_0]$, $[T_0, 2T_0]$, $\dots$, to $[kT_0, (k+1)T_0]$, where $kT_0 < T \leq (k+1)T_0$ for some $k\in\mathbb{N}$. For simplicity of notation, write

$$G(t,x,p) := c(t,x,\alpha_0(t,x)) + p\cdot f(t,x,\alpha_0(t,x)).$$

Then, for $M := 2\|\nabla q\|_\infty + 1$, define

$$\tilde{G}(t,x,p) := \begin{cases} G(t,x,p) & \text{if } |p|\leq M, \\ G(t,x,Mp/|p|) & \text{if } |p| > M. \end{cases}$$

It follows from (A1) and the Lipschitz continuity of α_0 that G is Lipschitz continuous in (t,x) with Lipschitz constant $C(1+|p|)$. Thus, also using that $\|f\|_\infty \leq 2N$, we get that, for all t,x,p,

$$|\tilde{G}_t(t,x,p)|, |\tilde{G}_x(t,x,p)| \leq C(1+M), \quad |\tilde{G}_p(t,x,p)| \leq 2N, \tag{2.12}$$

where C only depends on the Lipschitz norms of c, f, and α_0.

Now, let $\tilde{v}^h$ be the solution to

$$\begin{cases} \partial_t \tilde{v}^h(t,x) + \tilde{G}(t,x,\nabla^h \tilde{v}^h(t,x)) = -Nh\Delta^h \tilde{v}^h(t,x), \\ \tilde{v}^h(T,x) = q(x). \end{cases}$$

The goal is to show that $\tilde{v}^h$ is Lipschitz continuous and $\tilde{v}^h = v_0^h$ in $[0,T] \times \mathbb{R}^d$.

It follows from the equation of $\tilde{v}^h$ that $p_s(t,x) := \frac{\tilde{v}^h(t,x+se) - \tilde{v}^h(t,x)}{s}$ for any $e \in \mathbb{S}^{d-1}$ and that $s \in (0,1)$ satisfies

(2.13)
$$\begin{cases} \partial_t p_s(t,x) + G_1(t,x) + G_2(t,x) \cdot \nabla^h p_s(t,x) = -Nh\Delta^h p_s(t,x) & \text{in } (0,T) \times \mathbb{R}^d, \\ p_s(T,x) = \dfrac{q(x+se) - q(x)}{s} & \text{on } \mathbb{R}^d, \end{cases}$$

where

$$G_1(t,x) := \frac{1}{s} \int_0^s \tilde{G}_x\left(t, x+ze, \nabla^h \tilde{v}^h(t,x+se)\right) \cdot e\, dz,$$
$$G_2(t,x) := \int_0^1 \tilde{G}_p\left(t,x,\nabla^h \tilde{v}^h(t,x) + z(\nabla^h \tilde{v}^h(t,x+se) - \nabla^h \tilde{v}^h(t,x))\right) dz.$$

It is clear from (2.12) that $|G_1| \le C(1+M)$ and $|G_2| \le 2N$. This yields that the comparison principle for (2.13) holds. Thus, by comparing p_s with $\pm(\|\nabla q\|_\infty + C(1+M)(T-t))$, we obtain $|p_s(t,x)| \le \|\nabla q\|_\infty + C(1+M)(T-t)$. Sending $s \to 0$ yields that, for some C depending only on (A1), $|\nabla_e \tilde{v}^h(t,x)| \le \|\nabla q\|_\infty + C(1+M)(T-t)$. Thus, if $t \le T \le (2C)^{-1}$, we have that $\tilde{v}^h(t,x)$ is Lipschitz and

$$\sup_{(t,x)\in[0,T]\times\mathbb{R}^d} |\nabla \tilde{v}^h(t,x)| \le \|\nabla q\|_\infty + 1/2 + M/2 = M.$$

From the definition of ∇^h, we get $\sup_{(t,x)\in[0,T]\times\mathbb{R}^d} |\nabla^h \tilde{v}^h(t,x)| \le M$. Hence, $\tilde{v}^h$ is a solution to (2.9) for $n = 0$. The uniqueness of the solution to (2.9) yields that $v_0^h \equiv \tilde{v}^h$. So, we obtain the uniform Lipschitz continuity of v_0^h in space with the Lipschitz constant of the form $C\exp(CT)$. The Lipschitz regularity in time follows from the equation. ∎

We point out that the Lipschitz constant of v_n^h may depend on both n and h for $n \ge 1$. Another consequence of the comparison principle is that the functions v_n^h are monotone decreasing in n.

PROPOSITION 2.3. *Under the assumptions of Proposition* 2.2, *we have that, for all* $n \ge 0$,

$$v_{n+1}^h(t,x) \le v_n^h(t,x) \qquad \textit{for all } (t,x) \in [0,T] \times \mathbb{R}^d.$$

Proof. By the definition of α_n,

$$\begin{aligned} &c(t,x,\alpha_{n+1}(t,x)) + \nabla^h v_n^h(t,x) \cdot f(t,x,\alpha_{n+1}(t,x)) \\ &\qquad \le c(t,x,\alpha_n(t,x)) + \nabla^h v_n^h(t,x) \cdot f(t,x,\alpha_n(t,x)). \end{aligned}$$

Thus, $v_n^h = v_n^h(t,x)$ is a supersolution to (2.9) with subscripts $n+1$ because it satisfies

$$\partial_t v_n^h + c(t,x,\alpha_{n+1}(t,x)) + \nabla^h v_n^h \cdot f(t,x,\alpha_{n+1}(t,x)) \le -Nh\Delta^h v_n^h \quad \text{in } (0,T) \times \mathbb{R}^d.$$

Therefore, the comparison principle (Lemma 2.1) yields that $v_{n+1}^h \le v_n^h$ in $[0,T]\times\mathbb{R}^d$ for each $n\ge 0$. ∎

Since v_n^h is uniformly bounded for all $n\ge 0$, the monotonicity property of v_n^h in n from Proposition 2.3 yields that v_n^h converges locally uniformly as $n\to\infty$. We denote the limit as v^h. Then, by the stability property of viscosity solutions, v^h solves

$$\begin{cases} \partial_t v^h(t,x) + H(t,x,\nabla^h v^h(t,x)) = -Nh\Delta^h v^h(t,x) & \text{in } (0,T)\times\mathbb{R}^d, \\ v^h(T,x) = q(x) & \text{on } \mathbb{R}^d, \end{cases} \tag{2.14}$$

where

$$\begin{aligned} H(t,x,p) &:= c(t,x,\alpha(t,x,p)) + p\cdot f(t,x,\alpha(t,x,p)) \\ &= \min_{a\in A}\left[c(t,x,a) + p\cdot f(t,x,a)\right]. \end{aligned} \tag{2.15}$$

Since $\alpha(t,x,p)$ is assumed to be uniformly Lipschitz continuous in all of its dependencies, there exists $C>0$ such that, for all $(t,x,p)\in[0,T]\times\mathbb{R}^d\times\mathbb{R}^d$,

$$|H_t(t,x,p)|,\, |H_x(t,x,p)| \le C(1+|p|), \quad |H_p(t,x,p)|\le C. \tag{2.16}$$

The same proof of uniform boundedness and Lipschitz continuity for v_0^h in Proposition 2.2 can show that v^h is uniformly bounded and Lipschitz continuous and that the estimates are uniform in $h>0$. In Lemma 2.4, we also consider the unique solution v to (2.4), as one expects that it is the limit of v^h as $h\to 0$. We will prove this fact in Theorem 3.5.

Lemma 2.4. *Under the assumptions of Proposition* 2.2, *let v_0^h, v^h and v be, respectively, solutions to* (2.9) *(for $n=0$)*, (2.14)*, and* (2.4). *Then, in $[0,T]\times\mathbb{R}^d$, v_0^h, v^h and v are bounded by $C(1+T)$ and are Lipschitz continuous with Lipschitz constant $C\exp(CT)$ for some universal constant $C>0$.*

For a general class of first-order Hamilton–Jacobi (continuous) equations, we refer to [3, 4] for the regularity results.

2.2. Discrete space-time schemes. Now, we consider the scheme that is discrete in both space and time. Let $\tau, h\in(0,1)$ and N such that

$$\max\{1, \|f\|_\infty/2\} \le N \le h/(2d\tau). \tag{2.17}$$

Assuming that $T/\tau\in\mathbb{N}$, we denote

$$\mathbb{N}_T^\tau := \{0,\tau,2\tau,\ldots,T\}, \quad \mathbb{Z}_h^d := h\mathbb{Z}^d,$$

$$\Omega_T^{\tau,h} := \mathbb{N}_T^\tau\times\mathbb{Z}_h^d, \quad \text{and} \quad \Omega_T' := (\mathbb{N}_T^\tau\backslash\{0\})\times\mathbb{Z}_h^d.$$

Given a Lipschitz continuous function $\alpha_0(t,x)$, let $V_n^{\tau,h}:\Omega_T^{\tau,h}\to\mathbb{R}$ be defined iteratively for $n=0,1,\ldots$ as follows:

$$\begin{cases} \partial_t^\tau V_n^{\tau,h}(t,x) + c(t,x,\alpha_n(t,x)) + \nabla^h V_n^{\tau,h}(t,x)\cdot f(t,x,\alpha_n(t,x)) & \\ \qquad\qquad = -Nh\Delta^h V_n^{\tau,h}(t,x) & \text{in } \Omega_T', \\ V_n^{\tau,h}(T,x) = q(x) & \text{on } \mathbb{Z}_h^d \end{cases} \tag{2.18}$$

with

$$\alpha_{n+1}(t,x) := \alpha(t,x,\nabla^h V_n^{\tau,h}(t,x)) \qquad \text{in } \Omega_T'. \tag{2.19}$$

Here, we used the notation $\partial_t^\tau V_n^{\tau,h}(t,x) := \frac{V_n^{\tau,h}(t,x) - V_n^{\tau,h}(t-\tau,x)}{\tau}$.

We also consider the following equation:

$$
(2.20)\qquad \begin{cases} \partial_t^\tau V^{\tau,h}(t,x) + H(t,x,\nabla^h V^{\tau,h}(t,x)) = -Nh\Delta^h V^{\tau,h}(t,x) & \text{in } \Omega'_T, \\ V^{\tau,h}(T,x) = q(x) & \text{on } \mathbb{Z}_h^d \end{cases}
$$

where H is given by (2.15). The goal is to show that $V_n^{\tau,h}$ converges to $V^{\tau,h}$ as $n\to\infty$ and $V^{\tau,h}$ converges to v as $\tau,h\to 0$, where v is given by (2.4).

We will use the following operator. For each $t\in\mathbb{N}_T^\tau$, let $\mathcal{F}_t : L^\infty(\mathbb{Z}_h^d)\to L^\infty(\mathbb{Z}_h^d)$ be defined as

$$
(2.21)\qquad \mathcal{F}_t(U)(x) := U(x) + \tau H(t,x,\nabla^h U(x)) + Nh\tau\Delta^h U(x).
$$

Then, the equation in (2.20) can be rewritten as $V_n^{\tau,h}(t-\tau,x) = \mathcal{F}_t(V_n^{\tau,h}(t,\cdot))(x)$. We need

$$
\max\{1, \|H_p\|_\infty/2\} \le N \le h/(2d\tau)
$$

(which corresponds to (2.17) as $\|H_p\|_\infty = \|f\|_\infty$) to guarantee a monotonicity property of the operator $\mathcal{F}_t$. That is, for all $t\in\mathbb{N}_T^\tau$ and $U,V\in L^\infty(\mathbb{Z}_h^d)$ satisfying $U\le V$, we have that $\mathcal{F}_t(U)\le\mathcal{F}_t(V)$; see, e.g., [16, 54]. It is easy to see that the same holds if we replace $H(t,x,p)$ by $c(t,x,\alpha_n(t,x)) + p\cdot f(t,x,\alpha_n(t,x))$ because $\|f\|_\infty \le 2N$.

The monotonicity property is important because it immediately implies the comparison principle of (2.20) and the scheme (2.18) and (2.19), in the sense that is similar to Lemma 2.1. As a consequence of this, one can show the following properties.

PROPOSITION 2.5. *Assume* (A1) *and* (A2) *and* (2.17). *Then, in $\Omega_T^{\tau,h}$, the solutions $V_n^{\tau,h}, V^{\tau,h}$ are bounded by $C(1+T)$ and are Lipschitz continuous with Lipschitz constant $C\exp(CT)$ for some universal constant $C>0$. Moreover, for all $n\ge 0$, we have that $V_{n+1}^{\tau,h}(t,x)\le V_n^{\tau,h}(t,x)$ for all $(t,x)\in\Omega_T^{\tau,h}$.*

The proof of Proposition 2.5 is similar to those of Propositions 2.2 and 2.3 and Lemma 2.4, and hence, we skip it.

3. Analysis of semidiscrete schemes.

3.1. Convergence of PI. We show that, for each fixed $h\in(0,1)$, $v_n^h\to v^h$ as $n\to\infty$ exponentially fast in an L^2_{loc} norm. We will assume $T\ge 1$ for convenience.

THEOREM 3.1. *Assume* (A1) *and* (A2) *and $N\ge 1$. Let v_n^h and v^h be, respectively, continuous solutions to* (2.9) *and* (2.14). *Then, there exists a universal constant $C>0$ such that, for all $n\ge 1$, $R\ge 1$, and $t\in[0,T]$, we have that*

$$
\int_{B_R} \left|v_n^h(t,x) - v^h(t,x)\right|^2 dx \le \frac{h}{2^{n+1}} \int_t^T \int_{\mathbb{R}^d} \exp\left[C(1+\|\nabla^h v^h\|_\infty^2)(s-t)/h\right] \times \left|D^h(v_0^h(s,x) - v^h(s,x))\right|^2 \min\left\{1, e^{-|x|+R+1}\right\} dxds.
$$

In particular, we have that

$$
\sup_{t\in[0,T]} \int_{B_R} \left|v_n^h(t,x) - v^h(t,x)\right|^2 dx \le C2^{-n}\exp\left[C\exp(CT)/h\right] R^d.
$$

Proof. In this proof, we write $v_n := v_n^h$ and $v := v^h$ and assume $T \geq 1$ for simplicity. For any fixed $R \geq 1$, let $\varphi = \varphi_R : [0,\infty) \to (0,1]$ be C^1 and satisfy

$$\begin{aligned} &\varphi(r) = 1 \quad \text{ on } [0,R], \qquad \varphi(r) = e^{-r+R} \quad \text{ on } [R+1,\infty), \\ &-\varphi'(r) \in [0, 4\varphi(r)] \quad \text{ for all } r > 0. \end{aligned} \tag{3.1}$$

It is clear that such φ exists. Later, we write $\varphi(x) := \varphi(|x|)$ for $x \in \mathbb{R}^d$.

Next, for some $A > 0$ to be determined, set

$$E_{t,n} := \frac{1}{2} e^{At} \int_{\mathbb{R}^d} |v_n(t,x) - v(t,x)|^2 \varphi(x)\, dx, \tag{3.2}$$

which is finite since v_n, v are uniformly bounded. Direct computation yields that

$$\frac{d}{dt} E_{t,n} = A E_{t,n} + e^{At} \underbrace{\int_{\mathbb{R}^d} (v_n(t,x) - v(t,x))\,(\partial_t v_n(t,x) - \partial_t v(t,x))\,\varphi(x)\,dx}_{=:X_{t,n}}. \tag{3.3}$$

Recall from (2.15) that $H(t,x,p) = c(t,x,\alpha(t,x,p)) + \nabla v \cdot f(t,x,\alpha(t,x,p))$. Below, we write

$$\begin{aligned} c := c(t,x,\alpha(t,x,\nabla^h v(t,x))) \quad &\text{and} \quad f := f(t,x,\alpha(t,x,\nabla^h v(t,x))), \\ c_n := c(t,x,\alpha_n(t,x)) \quad &\text{and} \quad f_n := f(t,x,\alpha_n(t,x))) \end{aligned}$$

for simplicity. We will also drop (t,x) from the notations of $v(t,x)$ and $v_n(t,x)$ and (x) from $\varphi(x)$ when there is no confusion. Direct computation yields that

$$\begin{aligned} &\int_{\mathbb{R}^d} (\Delta^h v) v \varphi\, dx \\ &\quad = -\int_{\mathbb{R}^d} |D^h v|^2 \varphi\, dx + \frac{1}{h^2} \sum_{i=1}^d \int_{\mathbb{R}^d} v(t,x+he_i)(v(t,x+he_i) - v(t,x))\varphi(x)\, dx \\ &\qquad - \frac{1}{h^2} \sum_{i=1}^d \int_{\mathbb{R}^d} v(t,x)(v(t,x) - v(t,x-he_i))\varphi(x)\, dx \\ &\quad = -\int_{\mathbb{R}^d} |D^h v|^2 \varphi\, dx - \int_{\mathbb{R}^d} v D^{-h} v \cdot D^{-h} \varphi\, dx, \end{aligned}$$

where the last equality was obtained by a change of variable. We then deduce from the equation that

$$\begin{aligned} X_{t,n} &= -\int_{\mathbb{R}^d} (v_n - v)(\nabla^h v_n \cdot f_n + c_n + Nh\Delta^h v_n - \nabla^h v \cdot f - c - Nh\Delta^h v)\varphi\, dx \\ &\geq Nh \int_{\mathbb{R}^d} |D^h(v_n - v)|^2 \varphi\, dx - Nh \int_{\mathbb{R}^d} |v_n - v||D^{-h}(v_n - v)||D^{-h}\varphi|\, dx \\ &\quad - \int_{\mathbb{R}^d} |v_n - v| \left(|\nabla^h(v_n - v)||f_n| + |f_n - f||\nabla^h v| + |c_n - c| \right) \varphi\, dx. \end{aligned} \tag{3.4}$$

Due to (3.1), $|D^{-h}\varphi(x)| \leq C\varphi(x)$ for some constant $C > 0$. Also, using $\|f\|_\infty < \infty$ and (2.11), we have $|\nabla^h(v_n - v)||f_n| \leq C(|D^h(v_n - v)| + |D^{-h}(v_n - v)|)$. Since v is Lipschitz continuous, $|\nabla^h v| \leq M$ for some $M \geq 1$. So, by (2.10) and the uniform Lipschitz continuity of f, c, and α, we have that, for some $C > 0$,

$$|f_n - f||\nabla^h v| + |c_n - c| \leq CM(|D^h(v_{n-1} - v)| + |D^{-h}(v_{n-1} - v)|). \tag{3.5}$$

With all these, if denoting

$$G^h_{t,n} := \int_{\mathbb{R}^d} |D^h(v_n(t,x) - v(t,x))|^2 \varphi(x)\, dx,$$

it follows from (3.4) that, for some $C > 0$,

$$\begin{aligned} X_{t,n} \ge NhG^h_{t,n} &- C\int_{\mathbb{R}^d} |v_n - v|(|D^h(v_n - v)| + |D^{-h}(v_n - v)|)\varphi\, dx \\ &- CM\int_{\mathbb{R}^d} |v_n - v|(|D^h(v_{n-1} - v)| + |D^{-h}(v_{n-1} - v)|)\varphi\, dx. \end{aligned}$$

Denote $w_n(t,x) := v_n(t,x) - v(t,x)$. Since $\varphi(x - he_i) \le (1 + Ch)\varphi(x)$ by the choice of φ, there exists $C > 0$ such that

$$\begin{aligned} G^{-h}_{t,n} &= \int_{\mathbb{R}^d} \sum_{i=1}^d h^{-2} |w_n(t,x) - w_n(t, x - he_i)|^2 \varphi(x)\, dx \\ &\le (1 + Ch)\int_{\mathbb{R}^d} \sum_{i=1}^d h^{-2}|w_n(t,x) - w_n(t,x - he_i)|^2 \varphi(x - he_i)\, dx \\ &= (1 + Ch)G^h_{t,n}. \end{aligned} \tag{3.6}$$

Then, using (3.2) and Young's inequality, we get, for some universal $C > 0$ and any $\sigma_1, \sigma_2 > 0$, that

$$\begin{aligned} X_{t,n} \ge NhG^h_{t,n} &- \frac{\sigma_1}{2 + Ch}\int_{\mathbb{R}^d} (|D^h(v_n - v)|^2 + |D^{-h}(v_n - v)|^2)\varphi\, dx \\ &- \frac{\sigma_2}{2 + Ch}\int_{\mathbb{R}^d} (|D^h(v_{n-1} - v)|^2 + |D^{-h}(v_{n-1} - v)|^2)\varphi\, dx \\ &- C(2 + Ch)(\sigma_1^{-1} + M^2\sigma_2^{-1})\int_{\mathbb{R}^d} |v_n - v|^2 \varphi\, dx \\ \ge (Nh - \sigma_1)&G^h_{t,n} - \sigma_2 G^h_{t,n-1} - C(\sigma_1^{-1} + M^2\sigma_2^{-1})e^{-At}E_{t,n}. \end{aligned}$$

Using this and $E_{T,n} = 0$ and integrating (3.3) over $[t, T]$, we obtain, for some universal $C > 0$,

$$\begin{aligned} -E_{t,n} &\ge (A - C\sigma_1^{-1} - CM^2\sigma_2^{-1})\int_t^T E_{s,n}\, ds \\ &\quad + (Nh - \sigma_1)\int_t^T e^{As} G^h_{s,n}\, ds - \sigma_2 \int_t^T e^{As} G^h_{s,n-1}\, ds. \end{aligned} \tag{3.7}$$

Now, taking $\sigma_1 := h/2$, $\sigma_2 := h/4$ and $A := 6CM^2/h$, then (3.7) and $N \ge 1$ yield that

$$\int_t^T e^{As} G^h_{s,n}\, ds \le \frac{1}{2}\int_t^T e^{As} G^h_{s,n-1}\, ds \le \cdots \le 2^{-n}\int_t^T e^{As} G^h_{s,0}\, ds.$$

With this, (3.7) also shows that $E_{t,n} \le \frac{h}{4}\int_t^T e^{As} G^h_{s,n-1}\, ds \le \frac{h}{2^{n+1}}\int_t^T e^{As} G^h_{s,0}\, ds$. Therefore, for all $n \ge 0$ and $t \in [0, T]$, we obtain that

$$\int_{B_R} |v_n(t,x) - v(t,x)|^2\, dx \le \frac{h}{2^{n+1}}\int_t^T \int_{\mathbb{R}^d} e^{A(s-t)}|D^h(v_0(s,x) - v(s,x))|^2 \varphi(x)\, dxds,$$

which, combined with Lemma 2.4, concludes the proof. ∎

Remark 3.2. In the proof of Theorem 3.1, we only used the following: uniform Lipschitz continuity of f, c, and α and uniform boundedness of f and $|\nabla^h v^h|$. In particular, the solutions v_n^h and v^h are allowed to have certain growth at $x=\infty$, and the comparison principle is not needed.

By Theorem 3.1, we immediately have the convergence of the policies.

THEOREM 3.3. *Assume* (A1) *and* (A2) *and that* $N \geq 1$. *Then, there exists a universal constant* $C>0$ *such that, for all* $n \geq 0$ *and* $R \geq 1$, *we have that*

$$\sup_{t\in[0,T]} \int_{B_R} |\alpha(t,x,\nabla^h v_n^h(t,x)) - \alpha(t,x,\nabla^h v^h(t,x))|^2 dx \leq C 2^{-n} \exp\left[\frac{C}{h}\exp(CT)\right] R^d.$$

Proof. Since α is Lipschitz continuous,

$$\begin{aligned}&\int_{B_R} \left|\alpha(t,x,\nabla^h v_n^h(t,x)) - \alpha(t,x,\nabla^h v^h(t,x))\right|^2 dx \\ &\quad\leq \frac{C}{h^2}\sum_{i=1}^d \int_{B_R} \left|v_n^h(t,x+he_i) - v^h(t,x+he_i) - v_n^h(t,x-he_i) + v^h(t,x-he_i)\right|^2 dx.\end{aligned}$$

We can then conclude the proof from Theorem 3.1. ∎

Remark 3.4. Here, we consider the problem where f is linear in x and c and q are quadratic in x (again, we assume that the control set A is compact). To compute the values and the optimal policy on $[0,T]\times B_R$ (then $(t,x)\in[0,T]\times B_R$), by (2.1), we have that $|x(t)| \leq CRe^{CT}$ for some $C>0$. Thus, by (2.2) and (2.3), we only need the information of c, f, and q (and hence, H) for $|x| \leq C'Re^{C'T}$ for some $C'>0$. We can then perform a cut-off of c, f, and q for $|x| \geq 2C'Re^{C'T}$ so that c, f, and q are globally bounded and the value function and the optimal policy remain the same on $[0,T]\times B_R$. This shows that the boundedness conditions we impose are not restrictive.

Of course, this argument does not work if we need to study the problem globally, but Theorems 3.1 and 3.3 deal exactly with this bounded setting.

3.2. Convergence of v^h as $h \to 0$. Let v^h and v be, respectively, solutions to (2.14) and (2.4). We show that $|v^h - v| \leq C_T\sqrt{h}$, where the rate is sharp (we refer to a simple example given in [18]). We also point out that, for a semi-Lagrangian scheme (which preserves the optimization structure), it is possible to obtain a first-order estimate $O(h)$ if the discretized solution is semiconcave; see [17, 46]. However, our scheme is based on finite difference, and it is unclear whether or not v^h is semiconcave. Along this line, our Theorem 3.8 provides a weak semiconcavity result for v^h.

THEOREM 3.5. *Assume* (A1) *and* (A2) *and that* $N \geq \max\{1, \|f\|_\infty/2\}$. *Then, there exists a universal constant* $C>0$ *such that*

$$\sup_{(t,x)\in[0,T]\times\mathbb{R}^d} |v(t,x) - v^h(t,x)| \leq C(1+T)(1+\|\nabla v\|_\infty)\sqrt{h}.$$

In particular, we have that $\sup_{(t,x)\in[0,T]\times\mathbb{R}^d} |v(t,x)-v^h(t,x)| \leq C\exp(CT)\sqrt{h}$.

Remark 3.6. This rate was obtained in [19, 33] for a large class of parabolic Bellman equations with Lipschitz coefficients. We apply a different argument—the classical doubling variable method that is used in [16], in which a discrete space-time homogeneous Hamilton–Jacobi equation is discussed. This argument allows us to obtain the same sharp estimate for the scheme (2.18), while it seems that the method in [19, 33] cannot (see Remark 4.3). See also [12] for a different proof of this convergence rate via the nonlinear adjoint method.

Proof. We assume that $T \geq 1$. Suppose, for some $(t_0, x_0) \in [0,T] \times \mathbb{R}^d$, that

$$8\sigma := v(t_0, x_0) - v^h(t_0, x_0) \geq \frac{1}{2} \sup_{(t,x)\in[0,T]\times\mathbb{R}^d} \left[v(t,x) - v^h(t,x)\right] > 0. \tag{3.8}$$

Below, we will show that $\sigma \leq CT(1 + \|\nabla v\|_\infty)\sqrt{h}$.

Consider a smooth function $g : \mathbb{R}^{d+1} \to [0,1]$ such that

(g1) $g(t,x) = 1 - t^2 - |x|^2$ if $t^2 + |x|^2 < 1/2$,

(g2) $0 \leq g(t,x) \leq 1/2$ if $t^2 + |x|^2 > 1/2$ and $g(t,x) = 0$ if $t^2 + |x|^2 > 1$.

For $\varepsilon > 0$, denote $g_\varepsilon(t,x) := g(t/\varepsilon, x/\varepsilon)$ and

$$L := \sup\left\{v(t,x), -v^h(t,x) : (t,x) \in [0,T] \times \mathbb{R}^d\right\} + 1 \geq 1.$$

By Lemma 2.4, $\sigma \leq L \leq CT$ for some universal constant $C > 0$. Next, for $\phi(x) := (1 + |x|^2)^{1/2}$ and $R \geq |x_0| + T$, we define $\Phi^h : [0,T]^2 \times \mathbb{R}^{2d} \to \mathbb{R}$ by

$$\begin{aligned}\Phi^h(t,s,x,y) &:= v(t,x) - v^h(s,y) - \frac{\sigma}{T}(2T - t - s)\\ &\quad - \frac{\sigma}{R}(\phi(x) + \phi(y)) + (8L + 2\sigma) g_\varepsilon(t-s, x-y).\end{aligned}$$

Since v, v^h are bounded and continuous, there exists $(t_1, s_1, x_1, y_1) \in [0,T]^2 \times \mathbb{R}^{2d}$ such that

$$\Phi^h(t_1, s_1, x_1, y_1) = \max_{[0,T]^2 \times \mathbb{R}^{2d}} \Phi^h(t,s,x,y). \tag{3.9}$$

Due to $\phi(x_0) \leq R$, by (3.8),

$$\Phi^h(t_1, s_1, x_1, y_1) \geq \Phi^h(t_0, t_0, x_0, x_0) \geq 8L + 6\sigma. \tag{3.10}$$

Since $\max\{v(t_1, x_1), -v^h(s_1, y_1)\} \leq L$, we deduce that $\Phi^h(t_1, s_1, x_1, y_1) \leq 2L + (8L + 2\sigma) g_\varepsilon(t_1 - s_1, x_1 - y_1)$, which, together with (3.10), implies that $g_\varepsilon(t_1 - s_1, x_1 - y_1) \geq 3/4$. Then, by (g1), we get that, for some $C > 0$,

$$g_\varepsilon(t-s, x-y) = 1 - \varepsilon^{-2}|t-s|^2 - \varepsilon^{-2}|x-y|^2 \tag{3.11}$$

whenever $|t - t_1|, |s - s_1|, |x - x_1|, |y - y_1| \leq \varepsilon/C$.

Now, by (3.9), the mapping

$$(t,x) \mapsto v(t,x) + \frac{\sigma}{T} t - \frac{\sigma}{R}\phi(x) + (8L + 2\sigma) g_\varepsilon(t - s_1, x - y_1) \tag{3.12}$$

is maximized at $(t,x) = (t_1, x_1)$. Together with the fact that v is Lipschitz continuous (taking $M := 1 + \|\nabla v\|_\infty$) and $|\nabla \phi| \leq 1$, we find that $|\nabla_x g_\varepsilon(t_1 - s_1, x_1 - y_1)| \leq (M + \sigma R^{-1})(8L + 2\sigma)^{-1}$ and $|\partial_t g_\varepsilon(t_1 - s_1, x_1 - y_1)| \leq (M + \sigma T^{-1})(8L + 2\sigma)^{-1}$. By (3.11), $\sigma \leq L \leq CT$, and $R \geq T$, these yield that

$$|x_1 - y_1| \leq C\varepsilon^2 (M + \sigma R^{-1})(L + \sigma)^{-1} \leq C\varepsilon^2 M L^{-1} \tag{3.13}$$

and

$$|t_1 - s_1| \leq C\varepsilon^2 (M + \sigma T^{-1})(L + \sigma)^{-1} \leq C\varepsilon^2 M L^{-1}. \tag{3.14}$$

Now, we first assume that $t_1, s_1 < T$. In view of (3.12), we apply the viscosity solution test for v to get that

$$
\begin{aligned}
&-\frac{\sigma}{T} - (8L+2\sigma)\,\partial_t g_\varepsilon(t_1-s_1, x_1-y_1) \\
&\quad + H\left(t_1, x_1, \frac{\sigma}{R}\nabla\phi(x_1) - (8L+2\sigma)\nabla_x g_\varepsilon(t_1-s_1, x_1-y_1))\right) \geq 0.
\end{aligned} \tag{3.15}
$$

Similarly, since $(s,y) \to v^h(s,y) - \frac{\sigma}{T}s + \frac{\sigma}{R}\phi(y) - (8L+2\sigma)g_\varepsilon(t_1-s, x_1-y)$ is minimized at (s_1, y_1), the comparison principle yields

$$
\begin{aligned}
&\frac{\sigma}{T} - (8L+2\sigma)\,\partial_t g_\varepsilon(t_1-s_1, x_1-y_1) \\
&\quad + H\left(s_1, y_1, -\frac{\sigma}{R}\nabla^h\phi(y_1) - (8L+2\sigma)\nabla_x^h g_\varepsilon(t_1-s_1, x_1-y_1)\right) \\
&\quad - Nh\Delta^h\left[\frac{\sigma}{R}\phi(y_1) - (8L+2\sigma)g_\varepsilon(t_1-s_1, x_1-y_1))\right] \leq 0.
\end{aligned}
$$

Thus, we get that

$$
\begin{aligned}
\frac{2\sigma}{T} &\leq H\left(t_1, x_1, \frac{\sigma}{R}\nabla\phi(x_1) - (8L+2\sigma)\nabla_x g_\varepsilon(t_1-s_1, x_1-y_1))\right) \\
&\quad - H\left(s_1, y_1, -\frac{\sigma}{R}\nabla\phi(y_1) - (8L+2\sigma)\nabla_x^h g_\varepsilon(t_1-s_1, x_1-y_1)\right) \\
&\quad + Nh\Delta^h\left[\frac{\sigma}{R}\phi(y_1) - (8L+2\sigma)g_\varepsilon(t_1-s_1, x_1-y_1))\right].
\end{aligned} \tag{3.16}
$$

It follows from (3.11) that, for $h \ll \varepsilon$, we have, at point (t_1-s_1, x_1-y_1), that

$$
\nabla_x^h g_\varepsilon = \nabla_x g_\varepsilon = 2\varepsilon^{-2}(x_1-y_1), \quad \Delta^h g_\varepsilon = -2d\varepsilon^{-2}. \tag{3.17}
$$

Due to $|\nabla\phi| \leq 1$ and $\Delta^h\phi \leq C$, we get that

$$
Nh\Delta^h\left[\frac{\sigma}{R}\phi(y_1) - (8L+2\sigma)g_\varepsilon(t_1-s_1, x_1-y_1))\right] \leq CL\varepsilon^{-2}h. \tag{3.18}
$$

Using (3.16)–(3.18) and the regularity of H (see (2.16)), we obtain, for some universal C, that

$$
\begin{aligned}
2\sigma T^{-1} &\leq C\sigma R^{-1}(|\nabla\phi(x_1)| + |\nabla\phi(y_1)|) + CL\varepsilon^{-2}h \\
&\quad + C(|t_1-s_1| + |x_1-y_1|)\,[1 + (8L+2\sigma)|\nabla_x g_\varepsilon(t_1-s_1, x_1-y_1)|] \\
&\leq C\sigma R^{-1} + CL\varepsilon^{-2}h + C(|t_1-s_1| + |x_1-y_1|)\left(1 + L\varepsilon^{-2}|x_1-y_1|\right),
\end{aligned}
$$

which, by (3.13) and (3.14), yields that $\sigma T^{-1} \leq C\sigma R^{-1} + CL\varepsilon^{-2}h + C\varepsilon^2 M^2 L^{-1}$. Now, we take $\varepsilon := M^{-1/2}L^{1/2}h^{1/4}$ and pass $R \to \infty$. Then, when h is sufficiently small, we obtain that $\sigma \leq CTM\sqrt{h}$ for some universal $C > 0$. This finishes the proof of the upper bound of $\sup_{[0,T]\times\mathbb{R}^d}(v - v^h)$ in the case when $t_1, s_1 < T$.

Next, suppose that one of t_1 and s_1 is equal to T. We only prove this for the case when $t_1 = T$. By (3.10) and the definition of Φ^h,

$$
8L + 6\sigma \leq v(t_1, x_1) - v^h(s_1, y_1) + (8L+2\sigma)g_\varepsilon(t_1-s_1, x_1-y_1).
$$

It follows from the proof of Lemma 2.4 that v^h is Lipschitz continuous with unit Lipschitz constant when $|T-t| \leq C$. Note that $\varepsilon^2 ML^{-1} \leq C$. Hence, (3.13) and (3.14) yield that

$$
\begin{aligned}
8L + 6\sigma &\leq |v(T, x_1) - q(y_1)| + |q(y_1) - v^h(s_1, y_1)| + (8L+2\sigma)g_\varepsilon(T-s_1, x_1-y_1) \\
&\leq C(|x_1-y_1| + |T-s_1|) + 8L + 2\sigma \leq C\varepsilon^2 ML^{-1} + 8L + 2\sigma.
\end{aligned}
$$

This yields that $\sigma \leq C\sqrt{h}$ for some universal $C > 0$.

Finally, the upper bound estimate for $\sup_{[0,T]\times B_R}(v^h - v)$ follows by using the same argument as the above. Applying Lemma 2.4 permits us to conclude. ∎

3.3. Almost everywhere convergence of the policy. It was proved in [27] that, under suitable assumptions, the solution v to (2.4) is semiconcave in space. From this, we are able to derive the almost everywhere convergence of the policies.

We say that a function $g:\mathbb{R}^d \to \mathbb{R}$ is uniformly semiconcave if there exists $C>0$ such that, for all $x,y\in\mathbb{R}^d$, we have that $g(x+y)+g(x-y)-2g(x)\le C|y|^2$. If g is uniformly bounded and Lipschitz continuous, and both $\pm g$ are uniformly semiconcave, then g is bounded in $W^{2,\infty}(\mathbb{R}^d)$, and we denote by

$$\|g\|_{W^{2,\infty}} = \|g\|_\infty + \|\nabla g\|_\infty + \|\nabla^2 g\|_\infty.$$

We make the following assumption:

(A3) $q(\cdot)$ is uniformly semiconcave, and $c(t,\cdot,a), f(t,\cdot,a)$ are bounded in $W^{2,\infty}(\mathbb{R}^d)$ uniformly in $t\in[0,T]$ and $a\in A$.

THEOREM 3.7. *Under the assumptions of Theorem* 3.5*, further assume* (A3)*. Then, $v(t,\cdot)$ is uniformly semiconcave for all $t\in[0,T]$. Moreover, for each $t\in[0,T]$ we have, for a.e. $x\in\mathbb{R}^d$, that*

$$\alpha(t_h,x_h,\nabla^h v^h(t_h,x_h)) \to \alpha(t,x,\nabla v(t,x)) \qquad \text{as } h\to 0,$$

where $[0,T]\times\mathbb{R}^d \ni (t_h,x_h)\to(t,x)$ as $h\to 0$.

We next show a weak type of semiconcavity of v^h.

THEOREM 3.8. *Under the assumptions of Theorem* 3.7*, there exists $C>0$ (also depending on* (A3)*) such that, for all $h\in(0,1)$, $t\in[0,T]$, and $x,y\in\mathbb{R}^d$,*

$$v^h(t,x+y)+v^h(t,x-y)-2v^h(t,x)\le C\exp(CT)\,(|y|^2+\sqrt{h}).$$

The proofs of the two theorems are similar to those of Theorem 4.4 and Theorem 4.5, and we choose to write the full details down there (as it is slightly more complicated there).

4. Analysis of discrete space-time schemes.

4.1. Convergence of PI. The parallel result of Theorem 3.1 on the convergence of $V_n^{\tau,h}\to V^{\tau,h}$ holds the same (see Figure 1 for a numerical illustration). However, the proof is more involved due to the discretization in the time direction. In it, we will emphasize the difference.

THEOREM 4.1. *Assume* (A1) *and* (A2) *and that $N\ge 1$. Let $V_n := V_n^{\tau,h}$ and $V := V^{\tau,h}$ be, respectively, continuous solutions to* (2.18) *and* (2.20)*. Then, there exists a universal constant $C>0$ such that, if $C(1+\|\nabla^h V\|_\infty^2)\tau\le h$, we have, for all $n\ge 1$, $R\ge 1$, and $t\in\mathbb{N}_T^\tau$,*

$$\sum_{x\in\mathbb{Z}_h^d,|x|\le R} |V_n(t,x)-V(t,x)|^2 \le \frac{h\tau}{2^{n+1}} \sum_{t\le s\in\mathbb{N}_T^\tau}\sum_{x\in\mathbb{Z}_h^d} \exp\left[\frac{C}{h}\exp(1+\|\nabla^h V\|_\infty^2)(s-t)\right]|D^h(V_0(s,x)-V(s,x))|^2 \min\left\{1,e^{-|x|+R+1}\right\}.$$

In particular, we have that

$$\max_{t\in\mathbb{N}_T^\tau}\sum_{x\in\mathbb{Z}_h^d,|x|\le R} |V_n(t,x)-V(t,x)|^2 \le C2^{-n}\exp\left[C\exp(CT)/h\right]R^d,$$

$$\max_{t\in\mathbb{N}_T^\tau}\sum_{x\in\mathbb{Z}_h^d,|x|\le R}|\alpha(t,x,\nabla^h V_n(t,x))-\alpha(t,x,\nabla^h V(t,x))|^2\le\frac{C}{2^n}\exp\left[\frac{C}{h}\exp(CT)\right]R^d.$$

Proof. Assume that $T\ge 1$ for simplicity. Let $\varphi=\varphi_R:[0,\infty)\to[0,1]$ be C^1 and satisfy (3.1), and let $A:=CT^2/h$ for some $C>0$ to be determined. Then, for $t\in\mathbb{N}_T^\tau$, set

$$E_{t,n}:=\frac{1}{2}e^{At}\sum_{x\in\mathbb{Z}_h^d}|V_n(t,x)-V(t,x)|^2\varphi(|x|),$$

which is finite. Direct computation yields that

$$\begin{aligned}
&\frac{E_{t,n}-E_{t-\tau,n}}{\tau}\\
&\quad\ge Ae^{-A\tau}E_{t,n}+\frac{1}{2}e^{A(t-\tau)}\sum_{x\in\mathbb{Z}_h^d}(V_n(t,x)+V_n(t-\tau,x)-V(t,x)-V(t-\tau,x))\\
&\qquad\times(\partial_t^\tau V_n(t,x)-\partial_t^\tau V(t,x))\varphi(|x|)\\
&\quad=Ae^{-A\tau}E_{t,n}+e^{A(t-\tau)}\sum_{x\in\mathbb{Z}_h^d}(V_n(t,x)-V(t,x))(\partial_t^\tau V_n(t,x)-\partial_t^\tau V(t,x))\varphi(|x|)\\
&\qquad-\frac{\tau}{2}e^{A(t-\tau)}\sum_{x\in\mathbb{Z}_h^d}\left|\partial_t^\tau V_n(t,x)-\partial_t^\tau V(t,x)\right|^2\varphi(|x|)\\
&\quad=:Ae^{-A\tau}E_{t,n}+e^{A(t-\tau)}X_{t,n}-\frac{\tau}{2}e^{A(t-\tau)}Y_{t,n}.
\end{aligned}\tag{4.1}$$

First, we consider the term $Y_{t,n}$ (which does not appear in the semidiscretization problem in Theorem 3.1). Similarly as before, for simplicity of notations, we write that

$$\begin{aligned}
\alpha:=\alpha(t,x,\nabla^h V(t,x)),\quad \alpha_n:=\alpha(t,x,\nabla^h V_{n-1}(t,x)),\\
c_n:=c(t,x,\alpha_n(t,x)),\quad\text{and}\quad f_n:=f(t,x,\alpha_n(t,x))).
\end{aligned}$$

We will also drop (t,x) from the notations of $V(t,x),V_n(t,x)$, and $(|x|)$ from $\varphi(|x|)$. It follows from (2.18) and (2.20) that

$$Y_{t,n}=\sum_{x\in\mathbb{Z}_h^d}\left|c_n+\nabla^h V_n\cdot f_n+Nh\Delta^h V_n-H(t,x,\nabla^h V)-Nh\Delta^h V\right|^2\varphi(|x|).$$

Recall that $H(t,x,\nabla^h V)=c(t,x,\alpha)+f(t,x,\alpha)\cdot\nabla^h V$ and $|\nabla^h V|\le M$ for some $M\ge 1$. So, the regularity assumptions and (2.11) yield that

$$\begin{aligned}
Y_{t,n}\le C\sum_{x\in\mathbb{Z}_h^d}\big(&M^2|D^h V_{n-1}-D^h V|^2+M^2|D^{-h}V_{n-1}-D^{-h}V|^2\\
&+|D^h V_n-D^h V|^2+|D^{-h}V_n-D^{-h}V|^2\big)\varphi(|x|)\le C\left(M^2G_{t,n-1}^h+G_{t,n}^h\right),
\end{aligned}$$

where, in the last inequality, we used the notation $G_{t,n}^h:=\sum_{x\in\mathbb{Z}_h^d}|D^h V_n(t,x)-D^h V(t,x)|^2\varphi(|x|)$ and (3.6) with the above-defined $G_{t,n}^h$ (which clearly holds the same).

Next, we consider the term $X_{t,n}$. Note that, for any $v \in L^\infty(\mathbb{Z}_h^d)$,

$$\sum_{x\in\mathbb{Z}_h^d} \Delta^h v\, v\varphi = -\sum_{x\in\mathbb{Z}_h^d} |D^h v|^2\varphi - \sum_{x\in\mathbb{Z}_h^d} vD^{-h}v\cdot D^{-h}\varphi.$$

So, similarly as before (also using (3.1), (3.6), the uniform Lipschitz assumptions, and Young's inequality), we have, for some universal $C>0$ and for any $\sigma_1,\sigma_2>0$,

$$\begin{aligned} X_{t,n} &\geq NhG^h_{t,n} - C\sum_{x\in\mathbb{Z}_h^d} |V_n - V|(|D^h(V_n-V)| + |D^{-h}(V_n-V)|)\varphi \\ &\quad - CM\sum_{x\in\mathbb{Z}_h^d} |V_n - V|(|D^h(V_{n-1}-V)| + |D^{-h}(V_{n-1}-V)|)\varphi \\ &\geq (Nh-\sigma_1)G^h_{t,n} - \sigma_2 G^h_{t,n-1} - C(\sigma_1^{-1} + M^2\sigma_2^{-1})e^{-At}E_{t,n}. \end{aligned}$$

Since $E_{T,n}\equiv 0$, putting the above together and summing up (4.1) with respect to t yields that

$$\begin{aligned} &-E_{t,n}/\tau \\ &\quad \geq (Nh-\sigma_1 - C\tau)\sum_{t+\tau\leq s\in\mathbb{N}_T^\tau} e^{A(s-\tau)}G^h_{s,n} - (\sigma_2 + CM^2\tau)\sum_{t+\tau\leq s\in\mathbb{N}_T^\tau} e^{A(s-\tau)}G^h_{s,n-1} \\ &\qquad + (A - C\sigma_1^{-1} - CM^2\sigma_2^{-1})e^{-A\tau}\sum_{t+\tau\leq s\in\mathbb{N}_T^\tau} E^h_{s,n} \end{aligned} \tag{4.2}$$

for some universal constant $C>0$.

Finally, we take $\sigma_1 := h/4$, $\sigma_2 := h/8$, $A := 12CM^2/h$. Then, if $\tau \leq h/(8CM^2)$, (4.2) yields that

$$\sum_{t+\tau\leq s\in\mathbb{N}_T^\tau} e^{A(s-\tau)}G^h_{s,n} \leq \frac{1}{2}\sum_{t+\tau\leq s\in\mathbb{N}_T^\tau} e^{A(s-\tau)}G^h_{s,n-1} \leq \cdots \leq 2^{-n}\sum_{t+\tau\leq s\in\mathbb{N}_T^\tau} e^{A(s-\tau)}G^h_{s,0},$$

and then, $E_{t,n} \leq \frac{h\tau}{4}\sum_{t+\tau\leq s\in\mathbb{N}_T^\tau} e^{A(s-\tau)}G^h_{s,n-1} \leq \frac{h\tau}{2^{n+1}}\sum_{t+\tau\leq s\in\mathbb{N}_T^\tau} e^{A(s-\tau)}G^h_{s,0}$. This, together with Proposition 2.5, concludes the proof of the first claim as before.

The second claim follows similarly as in Theorem 3.3. ∎

By shifting the solutions, we obtain uniform pointwise exponential convergence of $V_n^{\tau,h}$ to $V^{\tau,h}$ and $\alpha(\cdot,\cdot,\nabla^h V_n^{\tau,h}(\cdot,\cdot))$ to $\alpha(\cdot,\cdot,\nabla^h V^{\tau,h}(\cdot,\cdot))$ as $n\to\infty$ in $\Omega_T^{\tau,h}$.

4.2. Convergence of $V^{\tau,h}$ as $\tau,h\to 0$. Let $V^{\tau,h}$ and v be, respectively, solutions to (2.20) and (2.4). Theorem 4.2 proves that the difference between $V^{\tau,h}$ and v is at most of order $\sqrt{h}$. The argument follows the idea of [16, Theorem 1], which considered the discrete space-time scheme for the homogeneous Hamilton–Jacobi equation $v_t + H(Dv) = 0$.

THEOREM 4.2. *Assume* (A1) *and* (A2) *and* (2.17). *Then, there exists a universal* $C>0$ *such that*

$$\sup_{(t,x)\in\Omega_T^{\tau,h}} |v(t,x) - V^{\tau,h}(t,x)| \leq C(1+T)(1+\|\nabla v\|_\infty)\sqrt{h}.$$

In particular, we have that $\sup_{(t,x)\in\Omega_T^{\tau,h}} |v(t,x) - V^{\tau,h}(t,x)| \leq C\exp(CT)\sqrt{h}$.

Remark 4.3. It was shown in [18, 19, 33] that

$$\sup_{(t,x)\in\Omega_T^{\tau,h}} |v(t,x) - V^{\tau,h}(t,x)| \le C(\tau^{1/4} + h^{1/2}) \quad \text{for some } C = C(T) > 0,$$

where v solves a general degenerate parabolic Bellman equation and $V^{\tau,h}$ is its space-time finite difference approximation. For the first-order equations, our Theorem 4.2 obtains a better convergence rate of $C(\tau^{1/2} + h^{1/2})$.

Proof. Assume that $T \ge 1$, and suppose, for some $(t_0, x_0) \in \Omega_T^{\tau,h}$, that

$$8\sigma := v(t_0,x_0) - V^{\tau,h}(t_0,x_0) \ge \frac{1}{2} \sup_{(t,x)\in\Omega_T^{\tau,h}} \left[v(t,x) - V^{\tau,h}(t,x)\right] > 0. \tag{4.3}$$

Let $D_{T,\tau,h} := [0,T] \times \mathbb{N}_T^\tau \times \mathbb{R}^d \times \mathbb{Z}_h^d$ and

$$L := \sup\left\{v(t,x), -V^{\tau,h}(t,x) \,:\, (t,x) \in \Omega_T^{\tau,h}\right\} + 1.$$

Then, $\sigma \le L \le CT$ for some universal constant $C > 0$. Moreover, let R, g, and g_ε with $\varepsilon \in (0,1)$, and ϕ be from the proof of Theorem 3.5, and define $\Phi^h : D_{T,\tau,h} \to \mathbb{R}$ by

$$\begin{aligned}\Phi^h(t,s,x,y) := v(t,x) - V^{\tau,h}(s,y) - \frac{\sigma}{T}(2T - t - s)\\ - \frac{\sigma}{R}(\phi(x) + \phi(y)) + (8L + 2\sigma) g_\varepsilon(t - s, x - y).\end{aligned}$$

Suppose that

$$\Phi^h(t_1,s_1,x_1,y_1) = \max_{D_{T,\tau,h}} \Phi^h(t,s,x,y). \tag{4.4}$$

It is clear that (3.10)–(3.14) hold the same. By (3.14), if $\tau \ll \varepsilon^2 M/L$, we get that

$$|t_1 - s_1 - \tau| \le C\varepsilon^2 M/L \quad \text{with } M = 1 + \|\nabla v\|_\infty. \tag{4.5}$$

First, assume that $t_1, s_1 < T$. The viscosity solution test for v shows (3.15) by (3.12). Next, since $\Omega_T^{\tau,h} \ni (s,y) \to V^{\tau,h}(s,y) - \frac{\sigma}{T}s + \frac{\sigma}{R}\phi(y) - (8L+2\sigma) g_\varepsilon(t_1 - s, x_1 - y)$ is minimized at (s_1, y_1), then, for all $(s,y) \in \Omega_T^{\tau,h}$,

$$\begin{aligned}V^{\tau,h}(s,y) \ge V^{\tau,h}(s_1,y_1) - \frac{\sigma}{T}(s_1 - s) + \frac{\sigma}{R}(\phi(y_1) - \phi(y))\\ - (8L + 2\sigma)\left[g_\varepsilon(t_1 - s_1, x_1 - y_1) - g_\varepsilon(t_1 - s, x_1 - y)\right] =: \tilde{V}(s,y).\end{aligned}$$

Recall that $s_1 + \tau \le T$ and that $\mathcal{F}_t$ from (2.21) satisfies the monotonicity property. We obtain

$$V^{\tau,h}(s_1,y_1) = \mathcal{F}_{s_1+\tau}(V^{\tau,h}(s_1+\tau,\cdot))(y_1) \ge \mathcal{F}_{s_1+\tau}(\tilde{V}(s_1+\tau,\cdot))(y_1),$$

which gives

$$\begin{aligned}0 \ge \frac{\sigma}{T} - (8L + 2\sigma)\, \partial_t^\tau g_\varepsilon(t_1 - s_1, x_1 - y_1)\\ + H\left(s_1 + \tau, y_1, -\frac{\sigma}{R}\nabla^h\phi(y_1) - (8L+2\sigma)\nabla_x^h g_\varepsilon(t_1 - s_1 - \tau, x_1 - y_1)\right)\\ - Nh\Delta^h\left[\frac{\sigma}{R}\phi(y_1) - (8L+2\sigma) g_\varepsilon(t_1 - s_1 - \tau, x_1 - y_1))\right].\end{aligned} \tag{4.6}$$

By (3.11), if $\tau, h \ll \varepsilon^2$,

$$|\partial_t^\tau g_\varepsilon(t_1 - s_1, x_1 - y_1) - \partial_t g_\varepsilon(t_1 - s_1, x_1 - y_1)| \le C\varepsilon^{-2}\tau, \tag{4.7}$$

$$\nabla_x^h g_\varepsilon(t_1 - s_1 - \tau, x_1 - y_1) = \nabla_x g_\varepsilon(t_1 - s_1, x_1 - y_1) = 2\varepsilon^{-2}(x_1 - y_1). \tag{4.8}$$

Combining (4.6) with (3.15) and using (4.7) and (4.8) yields that

$$\begin{aligned}\frac{2\sigma}{T} \le{}& H\left(t_1, x_1, \frac{\sigma}{R}\nabla\phi(x_1) - (8L + 2\sigma)2\varepsilon^{-2}(x_1 - y_1)\right)\\ &- H\left(s_1 + \tau, y_1, -\frac{\sigma}{R}\nabla\phi(y_1) - (8L + 2\sigma)2\varepsilon^{-2}(x_1 - y_1)\right)\\ &+ Nh\Delta^h\left[\frac{\sigma}{R}\phi(y_1) - (8L + 2\sigma)g_\varepsilon(t_1 - s_1 - \tau, x_1 - y_1))\right] + CL\varepsilon^{-2}\tau.\end{aligned} \tag{4.9}$$

The definitions of ϕ and g_ε show (3.18). Then, applying (2.16) and (3.18) into (4.9), if $(\tau \le)\, h \ll \varepsilon^2$, we deduce, for some $C > 0$, that

$$\begin{aligned}\sigma T^{-1} &\le C\sigma R^{-1}(|\nabla\phi(x_1)| + |\nabla\phi(y_1)|) + CL\varepsilon^{-2}h + CL\varepsilon^{-2}\tau\\ &\quad + C(|t_1 - s_1 - \tau| + |x_1 - y_1|)\left[1 + (8L + 2\sigma)2\varepsilon^{-2}|x_1 - y_1|\right]\\ &\le C\sigma R^{-1} + CL\varepsilon^{-2}h + C\varepsilon^2 M^2 L^{-1},\end{aligned} \tag{4.10}$$

where, in the second inequality, we also used (3.13) and (4.5).

Now, we take $\varepsilon := M^{-1/2}L^{1/2}h^{1/4}$ and send $R \to \infty$. It is clear that $\tau \ll \varepsilon^2 M/L$ is satisfied when h is small. We obtain from (4.10) that $\sigma \le CTM\sqrt{h}$, which finishes the proof of the upper bound of $\sup_{\Omega_T^{\tau,h}}(v - V^{\tau,h})$ in the case when $t_1, s_1 < T$.

Next, if at least one of t_1 and s_1 is equal to T, the argument of Theorem 3.5 applies the same, except that we need to use Proposition 2.5 in place of Lemma 2.4. Finally, the proof for the upper bound of $\sup_{\Omega_T^{\tau,h}}(V^{\tau,h} - v)$ is the same. ∎

4.3. Almost everywhere convergence of the policy. We show the almost everywhere convergence of the policy and some semiconcavity properties of the solution.

THEOREM 4.4. *Under the assumptions of Theorem* 4.2*, further assume* (A3)*. Then, v is uniformly semiconcave for all $t \in [0, T]$. Moreover, for each $t \in [0, T]$, we have, for a.e. $x \in \mathbb{R}^d$, that*

$$\alpha(t_h, x_h, \nabla^h V^{\tau_h,h}(t_h, x_h)) \to \alpha(t, x, \nabla v(t, x)) \quad \text{as } h \to 0,$$

where $\Omega_T^{\tau_h,h} \ni (t_h, x_h) \to (t, x)$ as $h \to 0$ and τ_h satisfies $0 < 2N\tau_h \le h$.

Proof. The semiconcavity of $v(t, \cdot)$ follows from [27].

For the second claim, it suffices to prove that, for a fixed $t \in [0, T)$ and a.e. $x \in \mathbb{R}^d$,

$$\nabla^h V^{\tau_h,h}(t_h, x_h) \to \nabla v(t, x) \quad \text{as } h \to 0. \tag{4.11}$$

For any function $g : \mathbb{R}^d \to \mathbb{R}$, we denote by $D^+g(x)$ the set of subdifferentials of g:

$$D^+g(x) := \left\{p \in \mathbb{R}^d \mid \limsup_{y \to x} \frac{g(y) - g(x) - p \cdot (y - x)}{|y - x|} \le 0\right\}.$$

Due to $v(t, \cdot)$ being semiconcave, $D^+v(t, x)$ is nonempty for all $x \in \mathbb{R}^d$.

Because $v(t, \cdot)$ is Lipschitz continuous, $\nabla_x v(t, x)$ exists for a.e. $x \in \mathbb{R}^d$. We fix one such x. Since $V^{\tau_h,h}$ are Lipschitz continuous uniformly in h, after passing to a subsequence of $h \to 0$, we can assume that $\nabla^h V^{\tau_h,h}(t_h, x_h) \to p$ for some $p \in \mathbb{R}^d$. Since $V^{\tau_h,h}(t_h, x_h) \to v(t,x)$ as $h \to 0$, the stability of the subdifferential yields that $p \in D^+ v(t,x)$, while, because $\nabla_x v(t,x)$ exists, we get that $p = \nabla_x v(t,x)$. Note that this is for any convergent subsequence of $\nabla^h V^{\tau_h,h}(t_h, x_h)$, and so, we obtain (4.11). ◻

In Theorem 4.5, we show a weak type of semiconcavity of $V^{\tau,h}(t,\cdot)$. We use the "doubling variable" method; see, e.g., [27].

THEOREM 4.5. *Under the assumptions of Theorem* 4.4, *there exists* $C > 0$ *(also depending on* (A3)*) such that, for all* $t \in \mathbb{N}_T^\tau$ *and* $x, y \in \mathbb{Z}_h^d$,

$$V^{\tau,h}(t, x+y) + V^{\tau,h}(t, x-y) - 2V^{\tau,h}(t,x) \leq C \exp(CT)\,(|y|^2 + \sqrt{h}).$$

Proof. It suffices to show that there exist $C_T, C_T' > 0$ depending on the assumptions such that

$$V^{\tau,h}(t,x) + V^{\tau,h}(t,z) - 2V^{\tau,h}(t,y) \leq C_T \left(|x-y|^2 + |z-y|^2 + |x+z-2y|\right) + C_T'\sqrt{h} \tag{4.12}$$

for all $t \in \mathbb{N}_T^\tau$ and $x, y, z \in \mathbb{Z}_h^d$. By the assumption on q, the inequality holds for $t = T$ with $C_T = \|q\|_{W^{2,\infty}} =: C_0$ and $C_T' = 0$.

Suppose, for contradiction, that (4.12) fails. Then, we have that, for some $C_1 \geq 1$ to be determined and some $C \geq 2$,

$$\begin{aligned} &V^{\tau,h}(t,x) + V^{\tau,h}(t,z) - 2V^{\tau,h}(t,y) \\ &\quad - 2C_0 e^{C_1(T-t)} \left(|x-y|^4 + |z-y|^4 + |x+z-2y|^2\right)^{1/2} \geq C e^{C_1(T-t)}\sqrt{h} \end{aligned} \tag{4.13}$$

for some $(t,x,y,z) = (t',x',y',z') \in \mathbb{N}_T^\tau \times \mathbb{Z}_h^d$. Since $V^{\tau,h}(t,\cdot)$ is Lipschitz continuous (with Lipschitz constant bounded by $C\exp(C(T-t))$ by Proposition 2.5 with a shift in time), after enlarging the constant C in (4.13) and replacing y' by $y' + \sqrt{h}$ if necessary, we can assume that

$$|x' + z' - 2y'| \geq \sqrt{h}. \tag{4.14}$$

We denote $\psi(x,y,z) := |x-y|^4 + |z-y|^4 + |x+z-2y|^2$, and by (4.14),

$$\delta := \psi(x',y',z')^{1/2} \geq \sqrt{h}.$$

Then, for all $\varepsilon > 0$ sufficiently small, we obtain from (4.13) that

$$\begin{aligned} &\Phi(t,x,y,z) \\ &\quad := e^{C_1 t}\left(V^{\tau,h}(t,x) + V^{\tau,h}(t,z) - 2V^{\tau,h}(t,y)\right) - C_0 e^{C_1 T}\left(\delta + \delta^{-1}\psi(x,y,z)\right) - \varepsilon|y|^2 \end{aligned}$$

satisfies $\Phi(t',x',y',z') \geq e^{C_1 T}\sqrt{h}$. With the positive ε-term, Φ obtains its positive maximum that is at least $e^{C_1 T}\sqrt{h}$ in $\Omega_T^{\tau,h}$ at some point $(t_0,x_0,y_0,z_0) \in \mathbb{N}_T^\tau \times \mathbb{Z}_h^d$, where (t_0,x_0,y_0,z_0) depends on ε and δ. It is clear that $t_0 \leq T - \tau$ by the choice of C_0. Moreover, for

$$\gamma_0 := \delta + \delta^{-1}\psi(x_0,y_0,z_0),$$

we have that

$$V^{\tau,h}(t_0,x_0) + V^{\tau,h}(t_0,z_0) - 2V^{\tau,h}(t_0,y_0) \geq C_0 e^{C_1(T-t_0)}\gamma_0 + e^{C_1(T-t_0)}\sqrt{h}. \tag{4.15}$$

Due to the uniform boundedness of $V^{\tau,h}$, by further taking ε to be small enough depending on C, T, and h, it is easy to get $\varepsilon|y_0| \le h$.

Now, since $\Omega_T^{\tau,h} \ni (t,x) \to e^{C_1 t}V^{\tau,h}(t,x) - C_0 e^{C_1 T}\delta^{-1}\left(|x-y_0|^4 + |x+z_0-2y_0|^2\right)$ is maximized at (t_0,x_0), we get, for all $(t,x) \in \Omega_T^{\tau,h}$, that

$$\begin{aligned} V^{\tau,h}(t,x) &\le e^{C_1(t_0-t)}V^{\tau,h}(t_0,x_0) + C_0 e^{C_1(T-t)}\delta^{-1}\left(|x-y_0|^4 + |x+z_0-2y_0|^2\right) \\ &\quad - C_0 e^{C_1(T-t_0)}\delta^{-1}\left(|x_0-y_0|^4 + |x_0+z_0-2y_0|^2\right) =: \tilde V(t,x). \end{aligned}$$

Due to the equation and the monotonicity property of $\mathcal{F}_t$ (as in (2.21)), $V^{\tau,h}(t_0,x_0) = \mathcal{F}_{t_0+\tau}(V^{\tau,h}(t_0+\tau,\cdot))(x_0) \le \mathcal{F}_{t_0+\tau}(\tilde V(t_0+\tau,\cdot))(x_0)$. By direct computation,

$$\nabla_x^h(|x-y_0|^4 + |x+z_0-2y_0|^2) = 4(|x-y_0|^2+h^2)(x-y_0) + 2(x+z_0-2y_0),$$

$$\Delta_x^h(|x-y_0|^4 + |x+z_0-2y_0|^2) = (8+4d)|x-y_0|^2 + 2dh^2 + 2d.$$

We then get that

$$\begin{aligned} &\frac{(1-e^{-C_1\tau})}{\tau}V^{\tau,h}(t_0,x_0) \le H\left(t_0+\tau, x_0, \nabla_x^h \tilde V(t_0+\tau,x_0)\right) + Nh\Delta_x^h \tilde V(t_0+\tau,x_0) \\ &\quad \le H\left(t_0+\tau, x_0, 2C_{T,\delta}(q_{x_0}+p_0)\right) + CC_{T,\delta}h(|x_0-y_0|^2+1), \end{aligned} \tag{4.16}$$

where

$$q_{x_0} := 2(|x_0-y_0|^2+h^2)(x_0-y_0),$$

$$C_{T,\delta} := C_0 e^{C_1(T-t_0-\tau)}/\delta, \quad \text{and} \quad p_0 := x_0+z_0-2y_0. \tag{4.17}$$

Similarly, since $\Omega_T^{\tau,h} \ni (t,z) \to e^{C_1 t}V^{\tau,h}(t,z) - C_0 e^{C_1 T}\delta^{-1}(|z-y_0|^4 + |x_0+z-2y_0|^2)$ is maximized at (t_0,z_0), we get that

$$\frac{(1-e^{-C_1\tau})}{\tau}V^{\tau,h}(t_0,z_0) \le H\left(t_0+\tau, z_0, 2C_{T,\delta}(q_{z_0}+p_0)\right) + CC_{T,\delta}h(|z_0-y_0|^2+1), \tag{4.18}$$

where $q_{z_0} := 2(|z_0-y_0|^2+h^2)(z_0-y_0)$.

Next, note that $\Omega_T^{\tau,h} \ni (t,y) \to 2e^{C_1 t}V^{\tau,h}(t,y) + C_0 e^{C_1 T}\delta^{-1}\psi(x_0,y,z_0) + \varepsilon|y|^2$ is minimized at (t_0,y_0). Hence, we get that $V^{\tau,h}(t_0,y_0) \ge \mathcal{F}_{t_0+\tau}(\hat V(t_0+\tau,\cdot))(y_0)$, where

$$\begin{aligned} \hat V(t,y) &:= e^{C_1(t_0-t)}V^{\tau,h}(t_0,y_0) - (\varepsilon/2)|y|^2 + (\varepsilon/2)|y_0|^2 \\ &\quad - (C_0/2)e^{C_1(T-t)}\delta^{-1}\psi(x_0,y,z_0) + (C_0/2)e^{C_1(T-t_0)}\delta^{-1}\psi(x_0,y_0,z_0). \end{aligned}$$

From this, we obtain that

$$\begin{aligned} &-\frac{(1-e^{-C_1\tau})}{\tau}V^{\tau,h}(t_0,y_0) \\ &\le -H(t_0+\tau, y_0, 2C_{T,\delta}(q_{y_0}+p_0) - \varepsilon y_0) - Nh\Delta_y^h \hat V(t_0+\tau,y_0) \\ &\le -H(t_0+\tau, y_0, 2C_{T,\delta}(q_{y_0}+p_0) - \varepsilon y_0) + CC_{T,\delta}h(|x_0-y_0|^2 + |z_0-y_0|^2+1) + Ch\varepsilon, \end{aligned}$$

where

$$q_{y_0} := (|x_0-y_0|^2+h^2)(x_0-y_0) + (|z_0-y_0|^2+h^2)(z_0-y_0)$$

and $C_{T,\delta}$ and p_0 are given in (4.17). Using $|H_p| \le C$ and $\varepsilon |y_0| \le h$ yields that

$$-\frac{(1-e^{-C_1\tau})}{\tau} V^{\tau,h}(t_0,y_0) \le -H\left(t_0+\tau, y_0, 2C_{T,\delta}(q_{y_0}+p_0)\right) + CC_{T,\delta}h(|x_0-y_0|^2+|z_0-y_0|^2+1)+Ch. \tag{4.19}$$

Now, let $\alpha \in \mathcal{A}$ be such that

$$H\left(t_0+\tau, y_0, 2C_{T,\delta}(q_{y_0}+p_0)\right) = c(t_0+\tau,y_0,\alpha) + 2C_{T,\delta} f(t_0+\tau,y_0,\alpha)\cdot(q_{y_0}+p_0).$$

By (2.15), denoting $c_\alpha(\cdot) := c(t_0+\tau,\cdot,\alpha)$ and $f_\alpha(\cdot) := f(t_0+\tau,\cdot,\alpha)$, we have that

$$\begin{aligned}
&H\left(t_0+\tau, x_0, 2C_{T,\delta}(q_{x_0}+p_0)\right) + H\left(t_0+\tau, z_0, 2C_{T,\delta}(q_{z_0}+p_0)\right)\\
&\quad - 2H\left(t_0+\tau, y_0, 2C_{T,\delta}(q_{y_0}+p_0)\right)\\
&\le c_\alpha(x_0)+c_\alpha(z_0)-2c_\alpha(y_0)+2C_{T,\delta}\left[f_\alpha(x_0)\cdot(q_{x_0}+p_0)\right.\\
&\quad \left. + f_\alpha(z_0)\cdot(q_{x_0}+p_0) - 2f_\alpha(y_0)\cdot(q_{y_0}+p_0)\right]\\
&= c_\alpha(x_0)+c_\alpha(z_0)-2c_\alpha(y_0)+2C_{T,\delta}\left[(f_\alpha(x_0)-f_\alpha(y_0))\cdot q_{x_0}\right.\\
&\quad \left. +(f_\alpha(z_0)-f_\alpha(y_0))\cdot q_{z_0} + (f_\alpha(x_0)+f_\alpha(z_0)-2f_\alpha(y_0))\cdot p_0\right]\\
&\le \|c_\alpha\|_{W^{2,\infty}}(|x_0-y_0|^2+|z_0-y_0|^2+|x_0+z_0-2y_0|)\\
&\quad + 2C_{T,\delta}\|f_\alpha\|_{\mathrm{Lip}}(|x_0-y_0||q_{x_0}|+|z_0-y_0||q_{z_0}|)\\
&\quad + 2C_{T,\delta}\|f_\alpha\|_{W^{2,\infty}}(|x_0-y_0|^2+|z_0-y_0|^2+|x_0+z_0-2y_0|)|x_0+z_0-2y_0|,
\end{aligned} \tag{4.20}$$

where we used $2q_{y_0} = q_{x_0}+q_{z_0}$ and that, for any $x,y,z \in \mathbb{R}^d$ and $g \in W^{2,\infty}(\mathbb{R}^d)$, $|g(x)+g(z)-2g(y)| \le \|g\|_{W^{2,\infty}}(|x-y|^2+|z-y|^2+|x+z-2y|)$. By Young's inequality, we get that

$$|x_0-y_0||q_{x_0}|+|z_0-y_0||q_{z_0}| \le 2|x_0-y_0|^4+2|z_0-y_0|^4+h^4.$$

Also, using the definitions of $C_{T,\delta}$ and ψ, we get the left-hand side of (4.20)

$$\le Ce^{C_1(T-t_0)}(\delta+\delta^{-1}\psi(x_0,y_0,z_0)+h^4) = Ce^{C_1(T-t_0)}\gamma_0 + Ce^{C_1(T-t_0)}h^4/\delta$$

with $C>0$ only depending on $\|q\|_{W^{2,\infty}}$, $\|c_\alpha\|_{W^{2,\infty}}$, and $\|f_\alpha\|_{W^{2,\infty}}$.

Now, summing up (4.16), (4.18), and (4.19) twice, we get that

$$\begin{aligned}
&\frac{(1-e^{-C_1\tau})}{\tau}\left[V^{\tau,h}(t_0,x_0)+V^{\tau,h}(t_0,z_0)-2V^{\tau,h}(t_0,y_0)\right]\\
&\le Ce^{C_1(T-t_0)}\gamma_0 + \frac{C}{\delta}e^{C_1(T-t_0)}h^4 + CC_{T,\delta}h(|x_0-y_0|^2+|z_0-y_0|^2+1)+Ch\\
&\le Ce^{C_1(T-t_0)}\gamma_0 + Ce^{C_1(T-t_0)}\delta^{-1}(|x_0-y_0|^4+|z_0-y_0|^4) + Ce^{C_1(T-t_0)}\sqrt{h}\\
&\le Ce^{C_1(T-t_0)}\gamma_0 + Ce^{C_1(T-t_0)}\sqrt{h},
\end{aligned}$$

where we used $\delta \ge \sqrt{h}$ in the second inequality. Finally, this and (4.15) yield that

$$C_1(C_0e^{C_1(T-t_0)}\gamma_0 + e^{C_1(T-t_0)}\sqrt{h}) \le Ce^{C_1(T-t_0)}\gamma_0 + Ce^{C_1(T-t_0)}\sqrt{h},$$

with $C>0$ depending only on d, N, and the regularity assumptions of q, c, f. Thus, if C_1 is sufficiently large depending only on the assumptions, we get a contradiction, which finishes the proof of (4.12), which finishes the proof. ∎

5. Generalization: A PDE perspective. In this section, we consider PI for HJB equations with a general Hamiltonian. For convenient use of the Legendre transform, we write the system in the forward-in-time setting. It is easy to carry over to the backward-in-time setting.

Suppose that $\mathcal{H} : [0,T] \times \mathbb{R}^d \times \mathbb{R}^d \to \mathbb{R}$ is continuous such that $\mathcal{H}(t,x,p)$ is convex in p. Let $\mathcal{L}(t,x,\mu)$ be the Legendre transform of $\mathcal{H}$; that is,

$$\mathcal{L}(t,x,\mu) := \sup_{p\in\mathbb{R}^d} [p\cdot\mu - \mathcal{H}(t,x,p)] \qquad \text{for } (t,x,\mu)\in[0,T]\times\mathbb{R}^d\times\mathbb{R}^d.$$

We always have the inequality $\mathcal{L}(t,x,\mu)+\mathcal{H}(t,x,p)\geq p\cdot\mu$, with the equality holding if and only if $\mu = \nabla_p\mathcal{H}(t,x,p)$ and if and only if $p=\nabla_\mu\mathcal{L}(t,x,\mu)$.

The HJB equation is

$$\begin{cases} \partial_t v(t,x) + \mathcal{H}(t,x,\nabla v(t,x)) = 0 & \text{in } (0,T)\times\mathbb{R}^d, \\ v(0,x) = q(x) & \text{on } \mathbb{R}^d. \end{cases} \tag{5.1}$$

Under some assumptions (see [3, 54]), it is a classical result that v is uniformly Lipschitz continuous if q is Lipschitz continuous. So, we can assume that

$$\|\nabla v\|_{L^\infty([0,T]\times\mathbb{R}^d)} \leq M \quad \text{for some } M>0. \tag{5.2}$$

Now, we take

$$m_1 := \min_{\substack{|p|=2M,\\ t\in[0,T],x\in\mathbb{R}^d}} \mathcal{H}(t,x,p) \quad \text{and} \quad m_2 \geq \max_{\substack{|p|=3M,\\ t\in[0,T],x\in\mathbb{R}^d}} [\mathcal{H}(t,x,p)-m_1]/M,$$

and we can assume that $m_2 \geq 2$. Then, define

$$\tilde{\mathcal{H}}(t,x,p) := \begin{cases} \mathcal{H}(t,x,p) & \text{if } |p|\leq 2M, \\ \max\{\mathcal{H}(t,x,p), m_1 + m_2(|p|-2M)\} & \text{if } 2M<|p|\leq 3M, \\ m_1+m_2(|p|-2M) & \text{if } |p|>3M. \end{cases}$$

It is not hard to verify that $\tilde{\mathcal{H}}$ is continuous in all its dependencies and is convex in p. Due to (5.2), v is also a solution of (5.1) with $\mathcal{H}$ replaced by $\tilde{\mathcal{H}}$. Moreover, for $N := m_2/2 \geq 1$, we have that

$$|\tilde{\mathcal{H}}_p(t,x,p)| \leq 2N \quad \text{in } [0,T]\times\mathbb{R}^d\times\mathbb{R}^d. \tag{5.3}$$

We define $\tilde{\mathcal{L}}$ as the Legendre transform of $\tilde{\mathcal{H}}$. Since the goal is to approximate v, it suffices to study $\tilde{\mathcal{H}}$ and $\tilde{\mathcal{L}}$ instead of $\mathcal{H}$ and $\mathcal{L}$. From now on, with a slight abuse of notation, we write $\mathcal{H}$ and $\mathcal{L}$ as $\tilde{\mathcal{H}}$ and $\tilde{\mathcal{L}}$, respectively.

With the modified operators, we can consider the semidiscretization. For $h>0$,

$$\begin{cases} \partial_t v^h(t,x) + \mathcal{H}(t,x,\nabla^h v^h(t,x)) = Nh\Delta^h v^h(t,x) & \text{in } (0,T)\times\mathbb{R}^d, \\ v^h(0,x) = q(x) & \text{on } \mathbb{R}^d. \end{cases} \tag{5.4}$$

As before, $N\geq \|\nabla_p\mathcal{H}\|_\infty/2$ guarantees that the finite difference scheme is monotone. We also assume that there exists $C>0$ such that, for all $t,x,p\in[0,T]\times\mathbb{R}^d\times\mathbb{R}^d$,

$$|\mathcal{H}_t(t,x,p)|, |\mathcal{H}_x(t,x,p)| \leq C(1+|p|), \quad |\mathcal{H}(t,x,0)|\leq C. \tag{5.5}$$

We can replace $C(1+|p|)$ by just C for the modified operator. We will not discuss the space-time discretization of (5.1) since it is similar.

From the above discussions, we note the PDE in (5.1) can be rewritten as

$$\partial_t v(t,x) + \sup_{\mu\in\mathbb{R}^d} \{\nabla v(t,x)\cdot\mu - \mathcal{L}(t,x,\mu)\} = 0,$$

and the supremum is achieved when $\mu(t,x) = \nabla_p \mathcal{H}(t,x,\nabla v(t,x))$. Therefore, we give the following iteration scheme for (5.4). Fixing small $h>0$, we start with a uniformly bounded and Lipschitz continuous function $v_0^h(t,x)$ and then iteratively compute v_n^h as follows. For $n\geq 1$, let $v_n^h = v_n^h(t,x)$ be the solution to

$$\begin{cases} \partial_t v_n^h + \mu_{n-1}^h(t,x)\cdot\nabla^h v_n^h - \mathcal{L}(t,x,\mu_{n-1}^h(t,x)) = Nh\Delta^h v_n^h & \text{in } (0,T)\times\mathbb{R}^d, \\ v_n^h(0,x) = q(x) & \text{on } \mathbb{R}^d, \end{cases} \tag{5.6}$$

where we denoted $\mu_n^h(t,x) := \nabla_p \mathcal{H}(t,x,\nabla^h v_n^h(t,x))$. Note that $\mathcal{L}(t,x,\mu_n^h(t,x))$ is finite due to $\mu_n^h \leq 2N$. Essentially, v_n^h solves a linearized equation of (5.4).

Let v_n^h (for each $n\geq 1$ with given v_0^h), v^h, and v be, respectively, Lipschitz continuous solutions to (5.6), (5.4), and (5.1). We have the following monotonicity property.

PROPOSITION 5.1. *Suppose that* $N\geq \max\{1, \|\nabla_p\mathcal{H}\|_\infty/2\}$ *and that* $\mathcal{H}(t,x,p)$ *is convex in* p *and satisfies* (5.3) *and* (5.5). *Let* q *and* v_0^h *be uniformly bounded and Lipschitz continuous for all* $h>0$. *Then, the solutions* v_n^h *are uniformly bounded for all* $n\geq 1$ *and* $h>0$. *Moreover, we have, for all* $n\geq 0$, *that*

$$v_{n+1}^h(t,x) \leq v_n^h(t,x) \qquad \text{for all } (t,x)\in[0,T]\times\mathbb{R}^d.$$

In the current setting, the terms $\mu_n^h(t,x)$, $-\mathcal{L}(t,x,\mu)$, and $\nabla_p\mathcal{H}(t,x,p)$ correspond to $\alpha_n^h(t,x)$, $c(t,x,a)$, and $\alpha(t,x,p)$ in section 2, respectively. The proof of Proposition 5.1 then is identical to that of Proposition 2.3 after converting the problem to a backward-in-time setting by considering $w_n^h(t,x) = v_n^h(T-t,x)$.

Additionally, we have the following convergence results. The proof of Theorem 5.2 follows those of Theorems 3.1 and 3.3, and the proof of Theorem 5.3 is analogous to those of Theorems 3.7 and 3.8.

THEOREM 5.2. *Under the assumptions of Proposition* 5.1, *for all* $R\geq 1$, *there exists a constant* C *depending only on* T *and the assumptions such that we have, for all* $t\in[0,T]$, *that*

$$\int_{B_R} \left|v_n^h(t,x) - v^h(t,x)\right|^2 dx \leq C2^{-n}he^{Ct/h}R^d,$$

$$\int_{B_R} \left|\nabla_p\mathcal{H}(t,x,\nabla^h v_n^h(t,x)) - \nabla_p\mathcal{H}(t,x,\nabla^h v^h(t,x))\right|^2 dx \leq C2^{-n}e^{Ct/h}R^d/h.$$

Moreover, we have that $\sup_{(t,x)\in[0,T]\times\mathbb{R}^d} |v^h(t,x) - v(t,x)| \leq C\sqrt{h}$.

Next, let $\mathcal{H}$ take the form $\mathcal{H}(t,x,p) := \sup_{a\in A}[c(t,x,a) + p\cdot f(t,x,a)]$, where A is some set, $c: [0,T]\times\mathbb{R}^d\times A\to\mathbb{R}$, and $f: [0,T]\times\mathbb{R}^d\times A\to\mathbb{R}^d$.

THEOREM 5.3. *Under the assumptions of Theorem* 5.2, *assume that* $c(t,\cdot,a)$, $f(t,\cdot,a)$ *are bounded in* $W^{2,\infty}(\mathbb{R}^d)$ *uniformly for all* $t\in[0,T]$ *and* $a\in A$. *Then, for each* $t\in[0,T]$, *we have that, for a.e.* $x\in\mathbb{R}^d$,

$$\alpha(t_h, x_h, \nabla^h v^h(t_h,x_h)) \to \alpha(t,x,\nabla v(t,x)) \qquad \text{as } h\to 0,$$

where $[0,T]\times\mathbb{R}^d \ni (t_h,x_h)\to(t,x)$ *as* $h\to 0$.

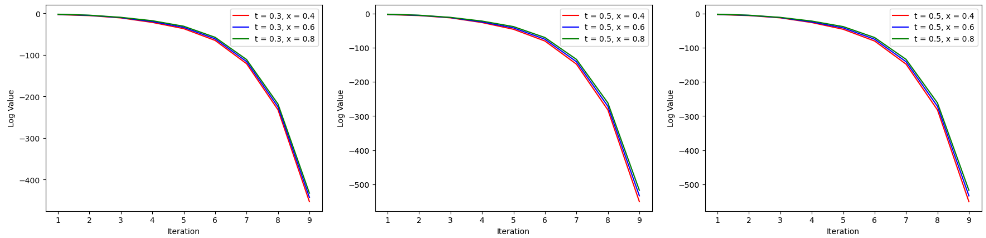

FIG. 1. *Convergence of PI* (2.18) *and* (2.19) *for* $f(t,x,a)=a$, $c(t,x,a)=\frac{1}{2}a^2$, *and* $q\equiv 0$ *with* $A=[-2,2]$, $\tau=0.025$, $h=0.1$, *and* $N=2$.

Moreover, there exists $C>0$ *depending only on the assumptions such that, for all* $h\in(0,1)$, $t\in[0,T]$, *and* $x,y\in\mathbb{R}^d$, $v^h(t,x+y)+v^h(t,x-y)-2v^h(t,x)\le C\exp(CT)(|y|^2+\sqrt{h})$.

6. Numerical experiments. In this section, we provide numerical experiments to illustrate the convergence of PI. We take $f(t,x,a)=a$, $c(t,x,a)=\frac{1}{2}a^2$, and $q\equiv 0$ so that $v_*\equiv 0$. Figure 1 (the semilog plot) shows the exponential convergence of PI, corroborating Theorem 4.1. For the vanishing viscosity approximations (Theorem 4.2), we refer to [43] for numerical illustration.

7. Conclusion. In this paper, we study the convergence rate of PI for optimal control problems in continuous time. To overcome the problem of ill-posedness, we consider a semidiscrete scheme by adding a viscosity term using finite differences. We prove that PI for the semidiscrete scheme converges exponentially fast and provide a bound on the discrepancy between the semidiscrete scheme and the optimal control. We also study the discrete space-time scheme, where both space and time are discretized.

Several future directions are related to PI. First, we plan to relax condition (A2) by considering

$$\alpha(t,x,p)\in\operatorname*{arg\,min}_{a\in A}\left[c(t,x,a)+p\cdot f(t,x,a)\right]$$

to replace (2.5). One possible approach is to use the minimization property of α rather than its precise value in the proof. Note, however, that without the Lipschitz continuity of α, we might not have the exponential convergence of the approximate optimal policies in Theorem 3.3. Second, we have only proved a weak form of semiconcavity for v^h in Theorem 3.8. It remains an open question whether v^h is semiconcave.

Several future directions are related to the nondiscretized PI. It remains unclear under which conditions on the model parameters the PI (2.7) and (2.8) is well-defined and converges exponentially fast. For instance, for $f(t,x,a)=a$, $c(t,x,a)=\frac{1}{2}|a|^2$, and $q\equiv 0$, the HJB equation is $\partial_t v-\frac{1}{2}|\nabla v|^2=0$ and $v(T,x)=0$, which has the solution $v_*\equiv 0$. On the other hand, PI yields that $v_n(t,x)=c_n(t)x^2$ with $c_1(t)=\frac{1}{2}$ for a suitable initialization. It is easy to check that $c_n(t)\le 2^{-n}$ for $n\ge 1$, and thus, we get the exponential convergence of v_n to v_* on any compact set. Moreover, it is also interesting to adapt PI to the differential game setting and design efficient numerical schemes (see, e.g., [25]). We refer to [30, 50] for the use of PI to solve numerically fully nonlinear HJB and HJBI equations.

REFERENCES

[1] M. Abu-Khalaf and F. L. Lewis, *Nearly optimal control laws for nonlinear systems with saturating actuators using a neural network HJB approach*, Automatica, 41 (2005), pp. 779–791, https://doi.org/10.1016/j.automatica.2004.11.034.

[2] S. Bansal, M. Chen, S. Herbert, and C. J. Tomlin, *Hamilton–Jacobi reachability: A brief overview and recent advances*, in 2017 IEEE 56th Annual Conference on Decision and Control (CDC), IEEE, 2017, pp. 2242–2253, https://doi.org/10.1109/CDC.2017.8263977.

[3] G. Barles, *Regularity results for first order Hamilton–Jacobi equations*, Differential Integral Equations, 3 (1990), pp. 103–125, https://doi.org/10.57262/die/1371586188.

[4] G. Barles, *Uniqueness and regularity results for first-order Hamilton–Jacobi equations*, Indiana Univ. Math. J., 39 (1990), pp. 443–466, https://doi.org/10.1512/iumj.1990.39.39024.

[5] R. Bellman, *Dynamic Programming*, Princeton University Press, Princeton, NJ, 1957.

[6] J. Z. Ben-Asher, *Optimal Control Theory with Aerospace Applications*, AIAA Education Series, AIAA, Reston, VA, 2010.

[7] D. P. Bertsekas, *Dynamic Programming and Optimal Control Vol.* II, Approximate Dynamic Programming, 4th ed., Athena Scientific, Nashua, NH, 2012.

[8] D. P. Bertsekas, *Value and policy iterations in optimal control and adaptive dynamic programming*, IEEE Trans. Neural Netw. Learn. Syst., 28 (2017), pp. 500–509, https://doi.org/10.1109/TNNLS.2015.2503980.

[9] T. Bian and Z.-P. Jiang, *Value iteration and adaptive dynamic programming for data-driven adaptive optimal control design*, Automatica, 71 (2016), pp. 348–360, https://doi.org/10.1016/j.automatica.2016.05.003.

[10] L. Blackmore, B. Açikmeşe, and D. P. Scharf, *Minimum-landing-error powered-descent guidance for Mars landing using convex optimization*, J. Guid. Control Dyn, 33 (2010), pp. 1161–1171, https://doi.org/10.2514/1.47202.

[11] S. Cacace, F. Camilli, and A. Goffi, *A policy iteration method for mean field games*, ESAIM Control Optim. Calc. Var., 27 (2021), 85, https://doi.org/10.1051/cocv/2021081.

[12] F. Cagnetti, D. Gomes, and H. V. Tran, *Convergence of a semi-discretization scheme for the Hamilton–Jacobi equation: A new approach with the adjoint method*, Appl. Numer. Math., 73 (2013), pp. 2–15, https://doi.org/10.1016/j.apnum.2013.05.004.

[13] F. Camilli, *A policy iteration method for mean field games*, IFAC-PapersOnLine, 55 (2022), pp. 406–411, https://doi.org/10.1016/j.ifacol.2022.11.087.

[14] F. Camilli and Q. Tang, *Rates of convergence for the policy iteration method for mean field games systems*, J. Math. Anal. Appl., 512 (2022), 126138, https://doi.org/10.1016/j.jmaa.2022.126138.

[15] M. Chen and C. J. Tomlin, *Hamilton–Jacobi reachability: Some recent theoretical advances and applications in unmanned airspace management*, Annu. Rev. Control Robot. Auton. Syst., 1 (2018), pp. 333–358, https://doi.org/10.1146/annurev-control-060117-104941.

[16] M. G. Crandall and P.-L. Lions, *Two approximations of solutions of Hamilton–Jacobi equations*, Math. Comp., 43 (1984), pp. 1–19, https://doi.org/10.1090/S0025-5718-1984-0744921-8.

[17] I. C. Dolcetta and H. Ishii, *Approximate solutions of the Bellman equation of deterministic control theory*, Appl. Math. Optim., 11 (1984), pp. 161–181, https://doi.org/10.1007/BF01442176.

[18] H. Dong and N. V. Krylov, *Rate of convergence of finite-difference approximations for degenerate linear parabolic equations with C^1 and C^2 coefficients*, Electron. J. Differential Equations, 2005 (2005), 102, https://digital.library.txst.edu/items/6ede96ff-a602-4f4a-802a-429bc2de67a1.

[19] H. Dong and N. V. Krylov, *The rate of convergence of finite-difference approximations for parabolic Bellman equations with Lipschitz coefficients in cylindrical domains*, Appl. Math. Optim., 56 (2007), pp. 37–66, https://doi.org/10.1007/s00245-007-0879-4.

[20] L. C. Evans, *On solving certain nonlinear partial differential equations by accretive operator methods*, Israel J. Math., 36 (1980), pp. 225–247, https://doi.org/10.1007/BF02762047.

[21] W. H. Fleming and H. M. Soner, *Controlled Markov Processes and Viscosity Solutions*, 2nd ed., Stoch. Model. Appl. Probab. 25, Springer, Cham, 2006.

[22] T. Haarnoja, A. Zhou, P. Abbeel, and S. Levine, *Soft actor-critic: Off-policy maximum entropy deep reinforcement learning with a stochastic actor*, in Proceedings of the 35th International Conference on Machine Learning (ICML), PMLR, 2018, pp. 1861–1870.

[23] B. Hambly, R. Xu, and H. Yang, *Policy gradient methods for the noisy linear quadratic regulator over a finite horizon*, SIAM J. Control Optim., 59 (2021), pp. 3359–3391, https://doi.org/10.1137/20M1382386.

[24] R. A. Howard, *Dynamic Programming and Markov Processes*, Technology Press of M.I.T., Wiley, Hoboken, NJ, 1960.

[25] K. Huang, X. Di, Q. Du, and X. Chen, *A game-theoretic framework for autonomous vehicles velocity control: Bridging microscopic differential games and macroscopic mean field games*, Discrete Contin. Dyn. Syst. Ser. B, 25 (2020), pp. 4869–4903, https://doi.org/10.3934/dcdsb.2020131.

[26] Y.-J. Huang, Z. Wang, and Z. Zhou, *Convergence of Policy Improvement for Entropy-Regularized Stochastic Control Problems*, preprint, arXiv:2209.07059, 2022.

[27] H. Ishii and P.-L. Lions, *Viscosity solutions of fully nonlinear second-order elliptic partial differential equations*, J. Differential Equations, 83 (1990), pp. 26–78, https://doi.org/10.1016/0022-0396(90)90068-Z.

[28] D. H. Jacobson and D. Q. Mayne, *Differential Dynamic Programming*, Mod. Analytic Computat. Methods Sci. Math. 24, Elsevier, New York, 1970.

[29] M. I. Kamien and N. L. Schwartz, *Dynamic Optimization: The Calculus of Variations and Optimal Control in Economics and Management*, 2nd ed., Adv. Textb. Economics 31, Elsevier, Amsterdam, 1991.

[30] E. L. Kawecki and T. Sprekeler, *Discontinuous Galerkin and C^0-IP finite element approximation of periodic Hamilton–Jacobi–Bellman–Isaacs problems with application to numerical homogenization*, ESAIM Math. Model. Numer. Anal., 56 (2022), pp. 679–704, https://doi.org/10.1051/m2an/2022017.

[31] B. Kerimkulov, D. Šiška, and L. Szpruch, *Exponential convergence and stability of Howard's policy improvement algorithm for controlled diffusions*, SIAM J. Control Optim., 58 (2020), pp. 1314–1340, https://doi.org/10.1137/19M1236758.

[32] D. Kleinman, *On an iterative technique for Riccati equation computations*, IEEE Trans. Automat. Control, 13 (1968), pp. 114–115, https://doi.org/10.1109/TAC.1968.1098829.

[33] N. V. Krylov, *The rate of convergence of finite-difference approximations for Bellman equations with Lipschitz coefficients*, Appl. Math. Optim., 52 (2005), pp. 365–399, https://doi.org/10.1007/s00245-005-0832-3.

[34] J. Lee and R. S. Sutton, *Policy iterations for reinforcement learning problems in continuous time and space—fundamental theory and methods*, Automatica, 126 (2021), 109421, https://doi.org/10.1016/j.automatica.2020.109421.

[35] F. L. Lewis and D. Vrabie, *Reinforcement learning and adaptive dynamic programming for feedback control*, IEEE Circuits Syst. Mag., 9 (2009), pp. 32–50, https://doi.org/10.1109/MCAS.2009.933854.

[36] Y. Li, K. H. Johansson, J. Mårtensson, and D. P. Bertsekas, *Data-driven rollout for deterministic optimal control*, in 2021 60th IEEE Conference on Decision and Control (CDC), IEEE, 2021, pp. 2169–2176, https://doi.org/10.1109/CDC45484.2021.9683499.

[37] D. Q. Mayne, *A second-order gradient method for determining optimal trajectories of non-linear discrete-time systems*, Int. J. Control, 3 (1966), pp. 85–95, https://doi.org/10.1080/00207176608921369.

[38] E. Nolasco, V. S. Vassiliadis, W. Kähm, S. D. Adloor, R. Al Ismaili, R. Conejeros, T. Espaas, N. Gangadharan, V. Mappas, F. Scott, and Q. Zhang, *Optimal control in chemical engineering: Past, present and future*, Comput. Chem. Eng., 155 (2021), 107528, https://doi.org/10.1016/j.compchemeng.2021.107528.

[39] S. Osher and C.-W. Shu, *High-order essentially nonoscillatory schemes for Hamilton–Jacobi equations*, SIAM J. Numer. Anal., 28 (1991), pp. 907–922, https://doi.org/10.1137/0728049.

[40] W. B. Powell, *Approximate Dynamic Programming: Solving the Curses of Dimensionality*, Wiley Ser. Prob. Stat., Wiley, Hoboken, NJ, 2007.

[41] M. L. Puterman, *Markov Decision Processes: Discrete Stochastic Dynamic Programming*, Wiley Ser. Probab. Math. Stat., Wiley, Hoboken, NJ, 1994.

[42] M. L. Puterman and S. L. Brumelle, *On the convergence of policy iteration in stationary dynamic programming*, Math. Oper. Res., 4 (1979), pp. 60–69, https://doi.org/10.1287/moor.4.1.60.

[43] J. Qian, T. Sprekeler, H. V. Tran, and Y. Yu, *Optimal Rate of Convergence in Periodic Homogenization of Viscous Hamilton–Jacobi Equations*, preprint, arXiv:2402.03091, 2024.

[44] C. Reisinger, W. Stockinger, and Y. Zhang, *Linear convergence of a policy gradient method for some finite horizon continuous time control problems*, SIAM J. Control Optim., 61 (2023), pp. 3526–3558, https://doi.org/10.1137/22M1492180.

[45] I. M. Ross, M. Karpenko, and R. J. Proulx, *A nonsmooth calculus for solving some graph-theoretic control problems*, IFAC-PapersOnLine, 49 (2016), pp. 462–467, https://doi.org/10.1016/j.ifacol.2016.10.208.

[46] L. Saluzzi, A. Alla, and M. Falcone, *Error estimates for a tree structure algorithm solving finite horizon control problems*, ESAIM Control Optim. Calc. Var., 28 (2022), 69, https://doi.org/10.1051/cocv/2022067.

[47] M. S. Santos and J. Rust, *Convergence properties of policy iteration*, SIAM J. Control Optim., 42 (2004), pp. 2094–2115, https://doi.org/10.1137/S0363012902399824.

[48] S. Sethi and X. Y. Zhou, *Stochastic dynamic job shops and hierarchical production planning*, IEEE Trans. Automat. Control, 39 (1994), pp. 2061–2076, https://doi.org/10.1109/9.328820.

[49] D. Silver, J. Schrittwieser, K. Simonyan, I. Antonoglou, A. Huang, A. Guez, T. Hubert, L. Baker, M. Lai, A. Bolton, et al., *Mastering the game of Go without human knowledge*, Nature, 550 (2017), pp. 354–359, https://doi.org/10.1038/nature24270.

[50] I. Smears and E. Süli, *Discontinuous Galerkin finite element approximation of Hamilton–Jacobi–Bellman equations with Cordes coefficients*, SIAM J. Numer. Anal., 52 (2014), pp. 993–1016, https://doi.org/10.1137/130909536.

[51] R. S. Sutton and A. G. Barto, *Reinforcement learning: An introduction*, in Adaptive Computation and Machine Learning, 2nd ed., MIT Press, Cambridge, MA, 2018.

[52] W. Tang and Y. P. Zhang, *The Convergence Rate of Vanishing Viscosity Approximations for Mean Field Games*, preprint, arXiv:2303.14560, 2023.

[53] W. Tang and X. Y. Zhou, *Regret of exploratory policy improvement and q-learning*, preprint, arXiv:2411.01302, 2024.

[54] H. V. Tran, *Hamilton–Jacobi Equations—Theory and Applications*, Grad. Stud. Math. 213, American Mathematical Society, Providence, RI, 2021.

[55] D. Vrabie, O. Pastravanu, M. Abu-Khalaf, and F. L. Lewis, *Adaptive optimal control for continuous-time linear systems based on policy iteration*, Automatica, 45 (2009), pp. 477–484, https://doi.org/10.1016/j.automatica.2008.08.017.

[56] D. Wang, H. He, and D. Liu, *Adaptive critic nonlinear robust control: A survey*, IEEE Trans. Cybernet., 47 (2017), pp. 3429–3451, https://doi.org/10.1109/TCYB.2017.2712188.

[57] Q. Wei, D. Liu, and H. Lin, *Value iteration adaptive dynamic programming for optimal control of discrete-time nonlinear systems*, IEEE Trans. Cybernet., 46 (2015), pp. 840–853, https://doi.org/10.1109/TCYB.2015.2492242.