arXiv:2301.00419v1 [math.OC] 1 Jan 2023

确定性控制问题的策略迭代——一种黏性方法

WENPIN TANG、HUNG VINH TRAN、YUMING PAUL ZHANG

Abstract. 本文研究连续时间（确定性）最优控制问题中策略迭代的收敛速率。为克服因缺乏正则性而导致的适定性问题，我们通过空间有限差分添加黏性项，构造半离散格式。我们证明，该半离散格式下的策略迭代以指数速度收敛，并给出半离散格式所引入误差的界。我们还研究了空间和时间均离散化的时空离散格式，并分析了策略迭代的收敛速率及离散误差。

1. 引言

最优控制广泛应用于科学和工程领域，包括航空航天工程 [6, 10], 化学工程 [31], 经济学 [24], 运筹学[36, 37] 以及机器人学 [2, 12]。自贝尔曼提出以来，动态规划（DP）已被证明是求解多阶段最优控制问题的有效工具 [5]。近年来，强化学习（RL）在解决复杂决策问题方面取得了巨大成功，阿尔法围棋 [38] 和人形机器人任务 [18]便是典型例子。作为近似或自适应动态规划（ADP）的一类，策略迭代（PI）在许多强化学习算法中发挥着重要作用[40]。

策略迭代（PI）的思想可追溯至 Howard [20]，当时研究的是一种称为马尔可夫决策过程（MDP）的随机环境。后续研究 [7, 33, 34] 探讨了离散时间和离散空间 MDP 的策略迭代；近期， [8, 30] 则研究了离散时间、连续空间下（确定性）最优控制问题的策略迭代。这些研究证明，在模型参数满足适当条件时，策略迭代能够收敛至最优控制。另一方面，许多现实问题都通过连续时间演化的动力系统建模；众所周知，连续时间和连续空间下的最优控制动态规划（DP）涉及哈密顿–雅可比–贝尔曼（HJB）偏微分方程（PDE）。尽管策略迭代十分重要，直到最近，连续时间和连续空间下最优控制问题的策略迭代才得到严格研究。 [1, 43] 证明了连续时间线性二次最优控制问题中策略迭代的收敛性；在不动点假设下，更一般的情形由 [28] 解决。对于随机控制问题，[26, 35] 表明，当控制仅作用于状态过程的漂移项时，策略迭代呈指数速度收敛。对应熵正则化问题也得出了类似结果 [22, 41]。此外，我们还提及一个密切相关的研究方向： [9, 45]

Date: 2023年1月3日。*Key words and phrases.*有限差分、哈密顿-雅可比-贝尔曼方程、最优控制、策略迭代。W. Tang 的工作得到美国国家科学基金会项目 DMS-2113779 和 DMS-2206038，以及哥伦比亚大学启动经费的支持。H. Tran 的工作部分得到美国国家科学基金会 CAREER 项目 DMS-1843320、西蒙斯学者奖和维拉斯学院青年研究者奖的支持。1

研究了最优控制问题的价值迭代。关于最优控制和强化学习（RL）中近似动态规划（ADP）理论与应用的最新进展，参见 [29, 44] 。

本文研究了在模型参数满足相当一般条件的情况下，连续时间最优控制问题及其离散化中策略迭代的收敛速率。需要指出的是， [1, 43] 中的收敛分析依赖于问题的线性二次结构，而 [28] 则假设哈密顿-雅可比-贝尔曼算子具有不动点或压缩性质，但这一条件很难验证。这些工作均未对策略迭代收敛到最优控制的过程进行量化。此外，由于缺乏正则性，连续时间控制问题的策略迭代甚至可能是不适定的。我们的思路是在策略评估中引入黏性项“$h\Delta^h$”，其中h为网格尺寸， Δ^h 为空间中的离散拉普拉斯算子。我们将其称为半离散格式。从本质上说，黏性项的阶为 1，这保证了有限差分格式是单调的。事实上，单调格式通常有利于数值实现，因此加入有限差分黏性项是很自然的。另一方面，半离散格式中的黏性项模拟了一阶偏微分方程的消失黏性近似 [16], ，从而使策略迭代以指数速度收敛（定理 3.1），与随机控制问题的情形相同。我们还证明，最优控制问题与其半离散格式之间的差异为 $\sqrt{h}$ ，当 $h \to 0$ 时成立（定理 3.3）。此外，我们还考虑时间离散化，称之为时空离散格式。对于时空离散格式中的策略迭代，同样的结果也成立（定理 4.1 和定理 4.2）。我们的结果与近期的研究[19], 相呼应；该研究指出，噪声能够促进有限时域强化学习算法的收敛。在我们的框架中，噪声对应于黏性项，而有限时域的重要性则体现在各种对时间呈指数依赖的界中。我们的分析依赖于偏微分方程技术，也可能适用于通过求解哈密顿-雅可比-贝尔曼-艾萨克斯（HJBI）方程研究微分博弈。

本文其余部分安排如下。第 2 节介绍背景，并给出半离散格式和时空离散格式。第 3 节研究半离散格式，第 4 节分析时空离散格式。第 5 节进一步从偏微分方程的角度进行探讨。最后在第 6 节作总结。

2. Setup and preliminary results

本节介绍半离散格式和时空离散格式。考虑一个状态由以下常微分方程控制的系统：

$$\frac{dx_t}{dt} = f(t, x_t, \alpha_t), \tag{2.1}$$

其中，对于 $0 \le t \le T$， $x_t \in \mathbb{R}^d$ 是系统状态， $\alpha_t \in A \subset \mathbb{R}^m$ 是控制或策略。这里， A 是 $\mathbb{R}^m$的一个给定紧子集。目标为

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

目标是使该目标函数最小化。记

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

其中， $\mathcal{A}_t$是定义为 $\mathcal{A}_t = \{\alpha \colon [t, T] \to A \colon \ \alpha$可测的$\}$的标准可容许策略。在对 $c(\cdot)$ 和 $q(\cdot)$ 的适当条件下（参见 [17, 第 2] 章或 [42, 第

2])，由 (2.3) 定义的 v_* 是以下方程的黏性解

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

其中，哈密顿量 H 由 $H(t,x,p) := \inf_{a\in A}\left[c(t,x,a) + p\cdot f(t,x,a)\right]$给出。最优策略为

$$\alpha_*(t,x) = \alpha(t,x,\nabla v_*), \tag{2.5}$$

其中

$$\alpha(t,x,p) := \arg\min_{a\in A}\left[c(t,x,a) + p\cdot f(t,x,a)\right], \tag{2.6}$$

为简便起见，本文假设其始终唯一。

策略迭代是一种近似动态规划方法，它通过交替进行策略评估（利用当前控制求得价值函数）和策略改进（优化价值函数）来实现。更确切地说，对于 $n = 0, 1, \cdots$，迭代过程如下：

- 给定 $\alpha_n(t,x)$，求解偏微分方程

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

- 设

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

关键在于理解序列 $\{v_n\}$ 如何逼近最优值 v_*，以及 $\{\alpha_n\}$ 如何逼近最优策略 α_*。

另一方面，尚不清楚策略迭代格式 (2.7)–(2.8) 是否适定。直观来说，为了使 α_{n+1} 有意义，我们需要 v_n 是利普希茨连续的；为此，我们又需要 α_n 是利普希茨连续的。这进而要求 ∇v_{n-1} 是利普希茨连续的。经过多次迭代，这意味着我们需要 v_0 是光滑的，而这通常并不成立。

2.1. 半离散格式。对于 $T \geq 1$、$h \in (0,1)$、$N \geq \max\{1, \|f\|_\infty/2\}$ 以及给定的连续函数 α_0：$\mathbb{R}\times\mathbb{R}^d \to A$，我们求解 $n = 0, 1, \cdots$

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

然后令

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

这里，对于任意的 $\varphi\colon\ \mathbb{R}^d \to \mathbb{R}$ 和 $h \in \mathbb{R}\setminus\{0\}$，我们采用以下记号

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

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

之后我们还将写作 $D^h\varphi(x) := \left(\frac{\varphi(x+he_1)-\varphi(x)}{h}, \cdots, \frac{\varphi(x+he_d)-\varphi(x)}{h}\right)$。显然，

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

假设 $N \geq \|f\|_\infty/2$ 保证数值哈密顿量是单调的，因此，下述比较原理成立（例如，参见 [13, 32, 42]）。

引理 2.1. 设 v_0^h 和 $\tilde{v}_0^h$ 分别是方程 (2.9) 的有界连续上解和下解，其中 $n = 0$，并且在 $t = T$ 处满足 $\tilde{v}_0^h \leq v_0^h$。则 $\tilde{v}_0^h \leq v_0^h$在 $[0, T] \times \mathbb{R}^d$中成立。这里，上解（分别地，下解）是指满足方程 (2.9)，但将第一个等式替换为 $\leq$（分别地，$\geq$）的函数。

首先，我们证明格式 (2.9)–(2.10) 是适定的。我们需要以下假设：

(A1) $c(\cdot,\cdot,\cdot)$ $f(\cdot,\cdot,\cdot)$ $q(\cdot)$ 关于其所有自变量均一致有界且利普希茨连续。 - 赖关系。(A2) $\alpha(\cdot,\cdot,\cdot)$ 和 $\alpha_0(\cdot,\cdot)$ 关于其所有自变量均一致利普希茨连续。

在本文中，除非另有说明，我们用 C 表示只依赖于 d、N 以及 (A1)–(A2) 中常数的各种通用常数。具体而言，由于 T 不是通用常数，我们在大多数估计中会追踪对 T 的依赖。常数 C 可能在不同的行中取不同的值。C_X 或 $C(X)$ 表示依赖于通用常数和 X 的常数。

命题 2.2. 假设 (A1)–(A2) 且 $N \geq \max\{1, \|f\|_\infty/2\}$。那么迭代过程 (2.9)–(2.10) 定义良好，也就是说，存在满足 (2.9)–(2.10) 的利普希茨连续函数 v_n^h、α_n，并且对于所有 $n \geq 0$ 和 $h > 0$，v_n^h 都一致有界。

证明. 由于 α_0 是利普希茨连续的，根据 [27, 定理 2.4]，方程 (2.9) 关于 $n = 0$ 唯一可解。若能证明 v_0^h 一致有界且利普希茨连续，其 Lipschitz 常数为 C_h，那么根据 α 是利普希茨连续的假设，α_1 就是利普希茨连续的，其 Lipschitz 常数为 C_h/h。通过相同的论证，我们得到一个唯一的有界利普希茨解 v_1^h。解的存在性随后由迭代得到。

首先证明 v_0^h 的有界性。由于 $c(\cdot,\cdot,\cdot)$ $q(\cdot)$ 一致有界，因此 $\pm[\|q\|_\infty + \|c\|_\infty(T-t)]$ 分别是方程 (2.9) 在 $n = 0$ 下的上解和下解。因此

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

对所有 $(t,x) \in [0,T] \times \mathbb{R}^d$ 均成立。实际上，通过这一论证，对所有 v_n^h，同样的界也成立。

接下来证明，当 T 足够小时，v_0^h 是利普希茨连续的，且其 Lipschitz常数与 h 无关。这里，T 只需依赖于假设 (A1)，而不依赖于 q。一般情形的结论可通过迭代立即得到。为简化记号，记

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

那么对于 $M := 2\|\nabla q\|_\infty + 1$，定义

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

由 (A1) 和 N 的选取可得

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

现在令 $\tilde{v}^h$ 为下式的解

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

目标是证明 $\tilde{v}^h$ 是利普希茨连续的，并且 $\tilde{v}^h = v_0^h$ 在 $[0,T]\times\mathbb{R}^d$ 中。

注意，对于任意 $e \in \mathbb{S}^{d-1}$ 和 $s \in (0,1)$，$p_s := \frac{\tilde{v}^h(t,x+se)-\tilde{v}^h(t,x)}{s}$ 满足

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

其中

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

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

由 (2.12) 显然可知，$|G_1| \le C(1+M)$ 和 $|G_2| \le 2N$。这说明 (2.13) 的比较原理成立。因此，通过将 p_s 与 $\pm(\|\nabla q\|_\infty + C(1+M)(T-t))$ 进行比较，我们得到 $|p_s(t,x)| \le \|\nabla q\|_\infty + C(1+M)(T-t)$。令 $s\to 0$ 趋于，可得对于某个仅依赖于 (A1) 的 C，$|\nabla_e \tilde{v}^h(t,x)| \le \|\nabla q\|_\infty + C(1+M)(T-t)$。因此，如果 $t \le T \le (2C)^{-1}$，则 $\tilde{v}^h(t,x)$ 是利普希茨的，并且

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

根据 ∇^h 的定义，我们得到 $\sup_{(t,x)\in[0,T]\times\mathbb{R}^d}|\nabla^h\tilde{v}^h(t,x)| \le M$. 因此，对于 $n=0$，$\tilde{v}^h$ 是方程 (2.9) 的一个解。方程 (2.9) 的解的唯一性表明 $v_0^h \equiv \tilde{v}^h$。因此，我们得到 v_0^h 关于空间的一致利普希茨连续性，其 Lipschitz 常数具有 $C\exp(CT)$ 的形式。时间上的 Lipschitz 正则性可由方程推出。□

我们指出，当 $n \ge 1$ 时，v_n^h 的 Lipschitz常数可能同时依赖于 n 和 h 。比较原理的另一个推论是，函数v_n^h 关于 n单调递减。

命题2.3. 在命题 2.2 的假设下，对所有 $n \ge 0$,

$$v_{n+1}^h \le v_n^h \qquad in\ [0,T]\times\mathbb{R}^d.$$

证明. 根据 α_n 的定义,

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

因此，v_n^h 是带有下标 $n+1$ 的方程 (2.9) 的上解，因为它满足

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

因此，由比较原理可得，对每个 $n\ge 0$，$v_n^h \le v_{n+1}^h$ 属于 $[0,T]\times\mathbb{R}^d$。 □

由于对于所有 $n \geq 0$, v_n^h 都是一致有界的，由单调性可知，当 $n \to \infty$ 时， v_n^h 局部一致收敛。将其极限记为 v^h。那么，根据黏性解的稳定性， v^h 是下列方程的解

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

其中

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

由于假设 $\alpha(t,x,p)$ 关于其所有自变量均一致利普希茨连续，因此存在 $C>0$，使得对所有 $(t,x,p) \in [0,T]\times\mathbb{R}^d\times\mathbb{R}^d$,

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

此外，可以预期 v^h 在 $h\to 0$ 时收敛，且极限 v 应为方程 (2.4) 的唯一解。

命题 2.2 已证明 v_0^h 一致有界，且关于 $[0,T]\times\mathbb{R}^d$ 利普希茨连续。由同样的证明可知，对于所有 $h>0$, v^h 和 v 也一致有界且利普希茨连续。

引理 2.4. 在命题 2.2 的假设下，设 v_0^h、v^h 和 v 分别为方程 (2.9)（对于 $n=0$）、(2.14) 和 (2.4) 的解。那么在 $[0,T]\times\mathbb{R}^d$ 中， v_0^h 、 v^h 和 v 均由 $C(1+T)$ 有界，且为利普希茨连续，其 Lipschitz常数为 $C\exp(CT)$，其中 $C>0$ 为某个普适常数。

关于一类一般的一阶 Hamilton-Jacobi（连续）方程，其正则性结果请参见 [3, 4]。

2.2. 离散时空格式。现在我们考虑一种在空间和时间上均离散的格式。设 $\tau, h \in (0,1)$，且 N满足

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

假设 $T/\tau \in \mathbb{N}^+$，记

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

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

给定一个利普希茨连续函数 $\alpha_0(t,x)$，令 $V^{\tau,h}_n : \Omega_T^{\tau,h} \to \mathbb{R}$,通过迭代定义如下，其中$n=0, 1, \cdots$:

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

with

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

这里我们使用记号 $\partial_t^\tau V^{\tau,h}_n(t,x) := \frac{V^{\tau,h}_n(t,x) - V^{\tau,h}_n(t-\tau,x)}{\tau}$。

我们还考虑以下方程

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

其中， H 由 (2.15) 给出。目标是证明，当 $n \to \infty$ 时， $V^{\tau,\,n\,h}$ 收敛到 $V^{\tau,\,h}$；并且当 τ, $h \to 0$ 时，$V^{\tau,\,h}$ 收敛到 v，其中 v 由 (2.4) 给出。

与之前类似，我们记作

$$c := c(t,x,\alpha(t,x,\nabla^h V^{\tau,h})) \quad \text{and} \quad f := f(t,x,\alpha(t,x,\nabla^h V^{\tau,h}))$$
$$c_n := c(t,x,\alpha_n(t,x)) \quad \text{and} \quad f_n := f(t,x,\alpha_n(t,x)))$$

为简便起见，在不会引起歧义时。

我们将使用以下算子。对于每个 $t \in \mathbb{N}_T^\tau$，令 $\mathcal{F}_t\colon L^\infty(\mathbb{Z}_h^d) \to L^\infty(\mathbb{Z}_h^d)$ 定义为

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

那么，(2.20) 中的方程可改写为 $V^{\tau,\,n\,h}(t-\tau,x) = \mathcal{F}_t(V^{\tau,\,n\,h}(t,\cdot))(x)$。我们需要

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

（当 $\|H_p\|_\infty = \|f\|_\infty$ 时，这对应于 (2.17)）以保证算子 $\mathcal{F}_t$ 的单调性。也就是说，对于所有满足 $U \le V$ 的 $t \in \mathbb{N}_T^\tau$ 和 $U, V \in L^\infty(\mathbb{Z}_h^d)$，都有 $\mathcal{F}_t(U) \le \mathcal{F}_t(V)$，例如参见 [13, 42]。容易看出，如果将 $H(t,x,p)$ 替换为 $c_n(t,x) + p \cdot f_n(t,x)$，同样的结论仍然成立。

单调性很重要，因为它立即推出 (2.20) 的比较原理以及格式 (2.18)–(2.19)，其意义与引理 2.1 类似。因此，可以证明以下性质。

命题 2.5. 假设 (A1)–(A2) 和 (2.17) 成立。那么在 $\Omega_T^{\tau\,h}$，中，解 $V^{\tau,\,n\,h}, V^{\tau,\,h}$ 由 $C(1+T)$ 有界，并且是利普希茨连续的，其 Lipschitz常数为 $C\exp(CT)$，其中 $C > 0$ 为某个普适常数。此外，对所有 $n \ge 0$，在 $\Omega_T^{\tau\,h}$，中都有 $V^{\tau,\;\;h}_{\;\;n+1} \le V^{\tau,\,n\,h}$。

命题 2.5 的证明与命题 2.2、2.3 及引理 2.4 的证明类似，因此略去。

3. Analysis of semi-discrete schemes

3.1. PI 的收敛性。我们证明，对于每个固定的 $h \in (0,1)$， $v_n^h \to v^h$ 在 L^2_{loc} 范数下以 $n \to \infty$ 指数速度收敛。

定理 3.1。假设 (A1)–(A2) 以及 $N \ge 1$。令 v_n^h 和 v^h 分别为方程 (2.9) 和 (2.14) 的连续解。则存在一个普适常数 $C > 0$，使得对于所有 $n \ge 1$、 $R \ge 1$ 和 $t \in [0,T]$，都有

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

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

证明。此处记 $v_n := v_n^h$ 和 $v := v^h$，并为简便起见假设 $T \geq 1$。对于任意固定的 $R \geq 1$，令 $\varphi = \varphi_R\colon [0,\infty) \to (0,1]$ 为 C^1 且满足

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

显然，这样的 φ 存在。

接下来，取待定的某个 $A \geq 1$，令

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

由于 v_n、v 一致有界，该量是有限的。直接计算得到

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

回顾（2.15），$H(t,x,\nabla^h v) = c(t,x,\alpha(t,x,\nabla^h v)) + \nabla v \cdot f(t,x,\alpha(t,x,\nabla^h v))$。为简便起见，记$c := c(t,x,\alpha(t,x,\nabla^h v))$，$f := f(t,x,\alpha(t,x,\nabla^h v))$，$c_n := c(t,x,\alpha_n(t,x))$以及 $f_n := f(t,x,\alpha_n(t,x))$。于是，利用

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

由该方程可得

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

由 (3.1)，$|D^{-h}\varphi(|x|)| \leq C\varphi(|x|)$，其中某个常数为 $C > 0$。此外，利用 $\|f\|_\infty < \infty$和 (2.11)，我们有 $|\nabla^h(v_n - v)||f_n| \leq C(|D^h(v_n - v)| + |D^{-h}(v_n - v)|)$。由于 v 是利普希茨连续的，存在某个 $M \geq 1$，使得 $|\nabla^h v| \leq M$。因此，由 (2.10) 以及f、c 和 α的一致利普希茨连续性可知，存在某个 $C > 0$，

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

综合以上，若记

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

由 (3.4) 可知，存在某个 $C>0$,

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

根据对 φ 的选择，存在 $C>0$ 使得

$$G^{-h}_{t,n} \leq (1 + Ch)G^h_{t,n} \tag{3.6}$$

然后，利用 (3.2) 和杨氏不等式，对于任意 σ_1、$\sigma_2 > 0$，可得

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

利用这一点以及 $E_{T,\ n} = 0$，并在 $[t,T]$ 上对 (3.3) 积分，可得对于某个通用的 $C>0$,

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

现在取 $\sigma_1 := h/2$、$\sigma_2 := h/4$ 和 $A := 6CM^2/h$，则 (3.7) 和 $N \geq 1$ 给出

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

由此，(3.7) 还表明 $E_{t,\ n} \leq \frac{h}{4}\int_t^T e^{As}G^h_{s,\ n-1}\, ds \leq \frac{h}{2^{n+1}}\int_t^T e^{As}G^h_{s,\ 0}\, ds$。因此，对于所有 $n \geq 0$ 和 $t \in [0,T]$，我们得到

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

结合引理 2.4，即可完成证明。 □

注 3.1。在定理 3.1 的证明中，我们只用到了以下条件：f、c 和 α 的一致利普希茨连续，以及 f 和 $|\nabla^h v^h|$ 的一致有界性。特别地，允许解 v^h_n 和 v^h 在 $x = \infty$ 处具有一定的增长性，并且不需要比较原理。

根据定理 3.1，我们立即得到策略的收敛性。

定理 3.2. 假设 (A1)–(A2) 和 $N \geq 1$。那么存在一个普适常数 $C>0$，使得对于所有 $n \geq 0$ 和 $R \geq 1$，我们有

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

证明. 由于 α 是利普希茨连续的,

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

由定理 3.1，我们即可得出证明。 □

3.2. 当 $h\to 0$ 时 v^h 的收敛。设 v^h 和 v 分别为方程 (2.14) 和 (2.4) 的解。我们证明 $|v^h-v|\leq C_T\sqrt{h}$，其收敛速率是尖锐的（参见 [14] 中给出的一个简单例子）。

定理 3.3. 假设 (A1)–(A2) 且 $N\geq \max\{1,\|f\|_\infty/2\}$。那么存在一个普适常数 $C>0$，使得

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

特别地，我们有 $\sup_{(t,x)\in[0,T]\times\mathbb{R}^d}|v(t,x)-v^h(t,x)|\leq C\exp(CT)\sqrt{h}$。

注 3.2. 对于一大类具有利普希茨系数的抛物型贝尔曼方程，这一收敛速率已在 [15, 27] 中得到。我们采用了不同的论证方法——经典的变量加倍法；[13] 中使用该方法讨论了一个离散的时空齐次哈密顿–雅可比方程。该论证使我们能够为格式 (2.18) 得到同样尖锐的估计，而 [15, 27] 中的方法似乎无法做到这一点（见注 4.1）。另请参见 [11]，其中通过非线性伴随方法给出了这一收敛速率的另一种证明。

证明。设 $T\geq 1$。假设对于某个 $(t_0,x_0)\in[0,T]\times\mathbb{R}^d$，使得

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

下面我们将证明 $\sigma\leq CT(1+\|\nabla v\|_\infty)\sqrt{h}$。

考虑一个光滑函数 $g\colon\mathbb{R}^{d+1}\to[0,1]$，使得

(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 $t^2+|x|^2>1$.

对于 $\varepsilon>0$，记 $g_\varepsilon(t,x):=g(t/\varepsilon,x/\varepsilon)$，并且

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

由引理 2.4，$\sigma\leq L\leq CT$，其中 $C>0$ 是某个普适常数。接下来，对于 $\phi(x):=(1+|x|^2)^{1/2}$ 和 $R\geq|x_0|+T$，我们定义 $\Phi^h\colon[0,T]^2\times\mathbb{R}^{2d}\to\mathbb{R}$ 为

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

由于 v、v^h 有界，存在 $(t_1,s_1,x_1,y_1)\in[0,T]^2\times\mathbb{R}^{2d}$，使得

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

由于 $\phi(x_0) \le R$，根据 (3.8)，

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

由于 $\max\{v(t_1, x_1) - v^h(s_1, y_1)\} \le L$，我们推得 $\Phi^h(t_1, s_1, x_1, y_1) \le 2L + (8L + 2\sigma) g_\varepsilon(t_1 - s_1, x_1 - y_1)$，结合 (3.10) 可知 $g_\varepsilon(t_1 - s_1, x_1 - y_1) \ge 3/4$。于是根据 (g1)，可知存在某个 $C > 0$，

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

只要 $|t - t_1|, |s - s_1|, |x - x_1|, |y - y_1| \le \varepsilon/C$。

现在，根据 (3.9)，映射

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

在 $(t, x) = (t_1, x_1)$ 处取得最大值。结合 v 是利普希茨连续的这一事实（取 $M := 1 + \|\nabla v\|_\infty$）以及 $|\nabla \phi| \le 1$，我们得到 $|\nabla_x g_\varepsilon(t_1 - s_1, x_1 - y_1)| \le (M + \sigma R^{-1})(8L + 2\sigma)^{-1}$，并且 $|\partial_t g_\varepsilon(t_1 - s_1, x_1 - y_1)| \le (M + \sigma T^{-1})(8L + 2\sigma)^{-1}$。由 (3.11)，$\sigma \le L \le CT$ 和 $R \ge T$，可得

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

and

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

现在，我们首先假设 $t_1, s_1 < T$。根据 (3.12)，对 v 应用粘性解检验，得到

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

类似地，由于 $(s, y) \to v^h(s, y) - \frac{\sigma}{T}s + \frac{\sigma}{R}\phi(y) - (8L + 2\sigma) g_\varepsilon(t_1 - s, x_1 - y)$ 在 (s_1, y_1) 处取到最小值，由比较原理可得

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

因此得到

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

由 (3.11) 可知，对于 $h \ll \varepsilon$，在点 $(t_1 - s_1, x_1 - y_1)$ 处有，

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

由于 $|\nabla\phi| \le 1$ 和 $\Delta^h\phi \le C$，我们得到

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

利用 (3.16)–(3.18) 以及 H 的正则性（见 (2.16)），我们得到某个通用的 C,

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

由 (3.13) 和 (3.14) 可得 $\sigma T^{-1}\le C\sigma R^{-1}+CL\varepsilon^{-2}h+C\varepsilon^2M^2L^{-1}$。现在取 $\varepsilon := M^{-1/2}L^{1/2}h^{1/4}$并令 $R\to\infty$ 趋于无穷。于是，当 h 足够小时，我们得到 $\sigma\le CTM\sqrt{h}$，其中某个通用常数为 $C>0$。这就完成了在 t_1、$s_1<T$ 情形下对 $\sup_{[0,T]\times\mathbb{R}^d}(v-v^h)$ 上界的证明。

接下来，假设 t_1 和 s_1 中有一个等于 T。我们仅证明 $t_1=T$ 的情形。由 (3.10) 及 Φ^h 的定义,

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

由引理 2.4 的证明可知，当 $|T-t|\le C$ 时，v^h 是利普希茨连续的，且 Lipschitz常数为 1。注意到 $\varepsilon^2ML^{-1}\le C$。因此，由 (3.13) 和 (3.14) 可得

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

由此得到 $\sigma\le C\sqrt{h}$，其中 $C>0$为某个普适常数。

最后，利用上述相同的论证可得 $\sup_{[0,T]\times B_R}(v^h-v)$ 的上界估计。应用引理 2.4 即可得出结论。□

3.3. 策略的几乎处处收敛。文献 [23] 证明了，方程 (2.4) 的解 v 在空间上是半凹的。由此，我们可以推导出策略的几乎处处收敛。

下文中，若存在 $C>0$，使得对所有 x、$y\in\mathbb{R}^d$，均有 $g(x+y)+g(x-y)-2g(x)\le C|y|^2$，则称函数 $g:\mathbb{R}^d\to\mathbb{R}$ 一致半凹。若 g 一致有界且利普希茨连续，并且 $\pm g$ 均一致半凹，则 g 在 $W^{2,\infty}(\mathbb{R}^d)$ 中有界。我们作如下假设:

(A3) $q(\cdot)$ 一致半凹，且 $c(t,\cdot,a)$ $f(t,\cdot,a)$ 在 $W^{2,\infty}(\mathbb{R}^d)$ 中关于 $t\in[0,T]$ 和 $a\in A$ 一致有界。

定理 3.4. 在定理 3.3 的假设下，进一步假设 (A3) 成立。则对所有 $t\in[0,T]$，$v(t,\cdot)$ 一致半凹。此外，对每个 $t\in[0,T]$，对于几乎处处的 $x\in\mathbb{R}^d$，有:

$$\alpha(t_h,x_h,\nabla^h v^h(t_h,x_h))\to\alpha(t,x,\nabla v(t,x))\quad as\ h\to0$$

其中，当 $h\to0$ 时，$[0,T]\times\mathbb{R}^d\ni(t_h,x_h)\to(t,x)$。

接下来，我们证明 v^h 的一种较弱的半凹性。

定理 3.5. 在定理 3.4 的假设下，存在 $C>0$（也依赖于 (A3)），使得对所有 $h\in(0,1)$、$t\in[0,T]$ 和 $x,y\in\mathbb{R}^d$,

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

这两个定理的证明与定理 4.3 和定理 4.4 的证明类似，我们选择在后文完整写出证明细节（因为后文的证明略微复杂一些）。

4. Analysis of discrete space-time schemes

4.1. PI 的收敛性。定理 3.1 关于 $V^\tau,\ _n{}^h \to V^{\tau,\ h}$ 收敛性的对应结果同样成立。不过，由于时间方向上的离散化，证明更加复杂。在证明中，我们将强调其中的差异。

定理 4.1. 假设 (A1)–(A2) 和 $N \geq 1$。设 $V_n := V^{\tau}{}_{,n}{}^h$ 和 $V := V^{\tau,h}$ 分别为 (2.18) 和 (2.20) 的连续解。则存在一个普适常数 $C > 0$，使得若 $C(1 + \|\nabla^h V\|_\infty^2)\tau \leq h$，则对所有 $n \geq 1$、$R \geq 1$ 和 $t \in \mathbb{N}_T^\tau$，有

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

特别地，我们有

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

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

证明. 假设 $T \geq 1$。设 $\varphi = \varphi_R : [0,\infty) \to [0,1]$ 为 C^1 且满足 (3.1)，并设 $A := CT^2/h$，其中 $C > 0$ 待定。则对 $t \in \mathbb{N}_T^\tau$，令

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

它是有限的。直接计算可得

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

首先，我们考虑项 $Y_{t,n}$（该项不会出现在定理 3.1 的半离散化问题中）。由方程 (2.18) 和 (2.20) 可得

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

回想一下，$\alpha=\alpha(t,x,\nabla^h V)$、$H(t,x,\nabla^h V)=c(t,x,\alpha)+f(t,x,\alpha)\cdot\nabla^h V$，且 $|\nabla^h V|\le M$对某个 $M\ge 1$成立。由于 $\alpha_n=\alpha(t,x,\nabla^h V_{n-1})$，正则性假设和 (2.11) 给出

$$Y_{t,n}\le C\sum_{x\in\mathbb{Z}_h^d}\Big(M^2|D^hV_{n-1}-D^hV|^2+M^2|D^{-h}V_{n-1}-D^{-h}V|^2+$$
$$|D^hV_n-D^hV|^2+|D^{-h}V_n-D^{-h}V|^2\Big)\varphi(|x|)\le C\left(M^2G^h_{t,n-1}+G^h_{t,n}\right),$$

其中，在最后一个不等式中，我们使用了记号 $G^h_{t,\,n}:=\sum_{x\in\mathbb{Z}_h^d}|D^hV_n(t,x)-D^hV(t,x)|^2\varphi(|x|)$。以及上述定义的$G^h_{t,\,n}$（显然同样成立）所对应的方程 (3.6)。

接下来，我们考虑项 $X_{t,\,n}$。注意，对于任意 $v\in L^\infty(\mathbb{Z}_h^d)$，$\sum_{x\in\mathbb{Z}_h^d}\Delta^h v\,v\varphi=-\sum_{x\in\mathbb{Z}_h^d}|D^hv|^2\varphi-\sum_{x\in\mathbb{Z}_h^d}vD^{-h}v\cdot D^{-h}\varphi$。因此，与前面类似（同时使用方程 (3.1)、(3.6)、一致利普希茨假设和杨氏不等式），对于任意 σ_1、$\sigma_2>0$，我们有

$$\begin{aligned}X_{t,n}&\ge NhG^h_{t,n}-C\sum_{x\in\mathbb{Z}_h^d}|V_n-V|(|D^h(V_n-V)|+|D^{-h}(V_n-V)|)\varphi\\&\quad-CM\sum_{x\in\mathbb{Z}_h^d}|V_n-V|(|D^h(V_{n-1}-V)|+|D^{-h}(V_{n-1}-V)|)\varphi\\&\ge(Nh-\sigma_1)G^h_{t,n}-\sigma_2G^h_{t,n-1}-C(\sigma_1^{-1}+M^2\sigma_2^{-1})e^{-At}E_{t,n}.\end{aligned}$$

由于 $E_{T,\,n}\equiv 0$，将上述结果结合起来，并对 t 求和 (4.1)，得到 d

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

对于某个普适常数 $C>0$。

最后我们取 $\sigma_1:=h/4$，$\sigma_2:=h/8$，$A:=12CM^2/h$。那么，如果 $\tau\le h/(8CM^2)$，由 (4.2) 可得

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

然后 $E_{t,\,n}\le\frac{h\tau}{4}\sum_{t+\tau\le s\in\mathbb{N}_T^\tau}e^{A(s-\tau)}G^h_{s,\,n-1}\le\frac{h\tau}{2^{n+1}}\sum_{t+\tau\le s\in\mathbb{N}_T^\tau}e^{A(s-\tau)}G^h_{s,\,0}$。结合命题 2.5，即可像之前一样完成第一个断言的证明。

第二个断言与定理 3.2 类似可得。 □

通过平移这些解，我们实际上得到了 $V^{\tau,\,n\,h}$ 到 $V^{\tau,\,h}$ 以及 $\alpha(\cdot,\cdot,\nabla^h V^{\tau,\,n\,h})$ 到 $\alpha(\cdot,\cdot,\nabla^h V^{\tau,\,h})$ 的一致逐点指数收敛，当 $n\to\infty$ 时，在 $\Omega_T^{\tau\,h,}$ 中。

4.2. $V^{\tau,h}$ 的收敛性（当 $\tau,h\to 0$ 时）。设 $V^{\tau,h}$ 和 v 分别是方程 (2.20) 和 (2.4) 的解。以下定理证明，$V^{\tau,h}$ 与 v 之差至多为 $\sqrt{h}$ 阶。该证明沿用了 [13, 定理 1], 的思路；该定理研究了齐次哈密顿-雅可比方程 $v_t+H(Dv)=0$的时空离散格式。

定理 4.2。假设 (A1)–(A2) 和 (2.17) 成立。则存在一个普适的 $C>0$，使得

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

特别地，我们有 $\sup_{(t,x)\in\Omega_T^{\tau,h}} |v(t,x)-V^{\tau,h}(t,x)| \le C\exp(CT)\sqrt{h}$。

注 4.1. 文献 [14, 15, 27] 中证明了

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

其中 v 是一个一般的退化抛物型 Bellman 方程的解，而 $V^{\tau,h}$ 是其时空有限差分近似。对于一阶方程，我们的定理 4.2 得到了更好的收敛速率。

证明。假设 $T\ge 1$。并设对某个 $(t_0,x_0)\in\Omega_T^{\tau,h}$，使得

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

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

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

那么，对某个普适常数 $C>0$，有 $\sigma\le L\le CT$。此外，令 R、g 和 g_ε 满足 $\varepsilon\in(0,1)$，并令 ϕ 来自定理 3.3 的证明，再定义 $\Phi^h: D_{T,\tau,h}\to\mathbb{R}$ 如下

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

假设

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

显然，(3.10)–(3.14) 仍然成立。由 (3.14)，若 $\tau\ll\varepsilon^2M/L$，则得到

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

首先，假设 $t_1,s_1<T$。对 v 的粘性解检验根据 (3.12) 得出 (3.15)。接下来，由于 $\Omega_T^{\tau,h}\ni(s,y)\to V^{\tau,h}(s,y)-\frac{\sigma}{T}s+\frac{\sigma}{R}\phi(y)-(8L+2\sigma)g_\varepsilon(t_1-s,x_1-y)$ 在 (s_1,y_1) 处取最小值，因此对所有 $(s,y)\in\Omega_T^{\tau,h}$，

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

回顾一下，(2.21) 中的 $s_1+\tau\le T$ 和 $\mathcal{F}_t$ 满足单调性。我们得到

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

从而得到

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

根据 (3.11)，如果 τ、$h \ll \varepsilon^2$,

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

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

结合 (4.6) 和 (3.15)，并利用 (4.7) 和 (4.8)，可得

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

ϕ 和 g_ε 的定义表明 (3.18)。然后，将 (2.16) 和 (3.18) 代入 (4.9)，若 ($\tau \leq$) $h \ll \varepsilon^2$，则可推出存在某个 $C>0$，使得

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

其中，在第二个不等式中我们还使用了 (3.13) 和 (4.5)。

现在取 $\varepsilon := M^{-1/2}L^{1/2}h^{1/4}$，并令 $R\to\infty$。显然，当 h 足够小时，$\tau \ll \varepsilon^2 M/L$ 得以满足。由 (4.10) 可得 $\sigma \leq CTM\sqrt{h}$，从而完成了在 t_1、$s_1 < T$ 情形下对 $\sup_{\Omega_T^{\tau,h}}(v - V^{\tau,h})$ 上界的证明。

接下来，如果 t_1 和 s_1 中至少有一个等于 T，则定理 3.3 的论证同样适用，只是需要用命题 2.5 替代引理 2.4。最后，关于 $\sup_{\Omega_T^{\tau,h}}(V^{\tau,h} - v)$ 的上界的证明也相同。□

4.3. 策略的几乎处处收敛。我们证明策略几乎处处收敛，并说明解具有某种半凹性。

定理 4.3. 在定理 4.2 的假设下，进一步假设 (A3) 成立。则对所有 $t\in[0,T]$，v 都是一致半凹的。此外，对每个 $t\in[0,T]$，我们有，对几乎处处的 $x\in\mathbb{R}^d$,

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

其中 $\Omega_T^{\tau_h,h} \ni (t_h,x_h) \to (t,x)$，当 $h\to 0$ 时，并且 τ_h 满足 $0 < 2N\tau_h \leq h$。

证明。$v(t,\cdot)$ 的半凹性由 [23] 得出。

对于第二个论断，只需证明对于固定的 $t\in[0,T)$，以及几乎处处的 $x\in\mathbb{R}^d$,

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

对于任意函数 $g:\mathbb{R}^d\to\mathbb{R}$，记 $D^+g(x)$ 为 g 的次微分集合：

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

由于 $v(t,\cdot)$ 是半凹的，因此对所有 $x\in\mathbb{R}^d$，$D^+v(t,x)$ 非空。

由于 $v(t,\cdot)$ 是利普希茨连续的，$\nabla_x v(t,x)$ 对几乎处处的 $x\in\mathbb{R}^d$ 存在。我们固定这样一个 x。由于 $V^{\tau_h,h}$ 关于 h 一致利普希茨连续，在取 $h\to 0$ 的一个子列后，我们可以假设对某个 $p\in\mathbb{R}^d$，$\nabla^h V^{\tau_h,h}(t_h,x_h)\to p$。由于当 $h\to 0$ 时，$V^{\tau_h,h}(t_h,x_h)\to v(t,x)$，次微分稳定性推出 $p\in D^+v(t,x)$。另一方面，由于 $\nabla_x v(t,x)$ 存在，我们得到 $p=\nabla_x v(t,x)$。注意，这对 $\nabla^h V^{\tau_h}$ 的任意收敛子列以及 $^h(t_h,x_h)$ 都成立，因此我们得到 (4.11)。□

下面，我们证明 $V^{\tau,h}(t,\cdot)$ 的一种弱形式的半凹性。我们采用“倍增变量”方法，参见例如 [23]。

定理 4.4. 在定理 4.3 的假设下，存在 $C>0$（也依赖于 (A3)），使得对所有 $t\in\mathbb{N}_T^\tau$ 和 $x,y\in\mathbb{Z}_h^d$，

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

证明. 只需证明存在依赖于这些假设的 $C_T,C_T'>0$，使得

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

对所有 $t\in\mathbb{N}_T^\tau$ 和 $x,y,z\in\mathbb{Z}_h^d$ 均成立。根据关于 q 的假设，当 $t=T$ 满足 $C_T=\|q\|_{W^{2,\infty}}=:C_0$，且 $C_T'=0$ 时，该不等式成立。

反设 (4.12) 不成立。则对于某个待定的 $C_1\ge 1$ 以及某个 $C\ge 2$，有

$$\begin{aligned}&V^{\tau,h}(t,x)+V^{\tau,h}(t,z)-2V^{\tau,h}(t,y)\\&\quad-2C_0e^{C_1(T-t)}\left(|x-y|^4+|z-y|^4+|x+z-2y|^2\right)^{1/2}\ge Ce^{C_1(T-t)}\sqrt{h}\end{aligned}\quad(4.13)$$

对于某个 $(t,x,y,z)=(t',x',y',z')\in\mathbb{N}_T^\tau\times\mathbb{Z}_h^d$。由于 $V^{\tau,h}(t,\cdot)$ 是利普希茨连续的（根据命题 2.5，并将时间平移后，其 Lipschitz常数有界于 $C\exp(C(T-t))$），在必要时增大 (4.13) 中的常数 C 后，我们可以假设

$$|x'+z'-2y'|\ge\sqrt{h}.\quad(4.14)$$

记 $\psi(x,y,z):=|x-y|^4+|z-y|^4+|x+z-2y|^2$，由 (4.14) 得，$\delta:=\psi(x',y',z')^{1/2}\ge\sqrt{h}$。于是，对所有充分小的 $\varepsilon>0$ sufficiently small，我们由 (4.13) 得到

$$\Phi(t,x,y,z):=e^{C_1t}\left(V^{\tau,h}(t,x)+V^{\tau,h}(t,z)-2V^{\tau,h}(t,y)\right)-C_0e^{C_1T}\left(\delta+\delta^{-1}\psi(x,y,z)\right)-\varepsilon|y|^2$$

满足 $\Phi(t',x',y',z')\ge e^{C_1T}\sqrt{h}$。由于存在正的 ε-项，Φ 在 $\Omega_T^{\tau,h}$ 内取得一个至少为 $e^{C_1T}\sqrt{h}$ 的正最大值，该最大值在某点 $(t_0,x_0,y_0,z_0)\in\mathbb{N}_T^\tau\times\mathbb{Z}_h^d$ 处取得，其中 (t_0,x_0,y_0,z_0) 取决于 ε 和 δ。根据 C_0 的选取，显然有 $t_0\le T-\tau$。此外，对于 $\gamma_0:=\delta+\delta^{-1}\psi(x_0,y_0,z_0)$，有

$$V^{\tau,h}(t_0,x_0)+V^{\tau,h}(t_0,z_0)-2V^{\tau,h}(t_0,y_0)\ge C_0e^{C_1(T-t_0)}\gamma_0+e^{C_1(T-t_0)}\sqrt{h}.\quad(4.15)$$

由于 $V^{\tau,h}$ 一致有界，通过进一步令 ε 足够小（其取值取决于 C、T 和 h），容易得到 $\varepsilon|y_0|\le h$。

现在，由于 $\Omega_T^{\tau,h} \ni (t,x) \to e^{C_1 t} V^{\tau,h}(t,x) - C_0 e^{C_1 T} \delta^{-1} \left(|x-y_0|^4 + |x+z_0-2y_0|^2 \right)$ 在 (t_0, x_0) 处取最大值，因此对所有 $(t,x) \in \Omega_T^{\tau,h}$，都有

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

由于方程以及 $\mathcal{F}_t$（定义见 (2.21)）的单调性，$V^{\tau,h}(t_0,x_0) = \mathcal{F}_{t_0+\tau}(V^{\tau,h}(t_0+\tau,\cdot))(x_0) \le \mathcal{F}_{t_0+\tau}(\tilde{V}(t_0+\tau,\cdot))(x_0)$。通过直接计算，

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

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

于是得到

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

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

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

类似地，由于 $\Omega_T^{\tau,h} \ni (t,z) \to e^{C_1 t} V^{\tau,h}(t,z) - C_0 e^{C_1 T} \delta^{-1} (|z-y_0|^4 + |x_0+z-2y_0|^2)$ 在 (t_0, z_0) 处取最大值，我们得到

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

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

接下来，注意到 $\Omega_T^{\tau,h} \ni (t,y) \to 2e^{C_1 t} V^{\tau,h}(t,y) + C_0 e^{C_1 T} \delta^{-1} \psi(x_0,y,z_0)$ 在 (t_0,y_0) 处取最小值。因此得到 $V^{\tau,h}(t_0,y_0) \ge \mathcal{F}_{t_0+\tau}(\hat{V}(t_0+\tau,\cdot))(y_0)$，其中

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

由此可得

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

其中 $q_{y_0} := (|x_0-y_0|^2 + h^2)(x_0-y_0) + (|z_0-y_0|^2 + h^2)(z_0-y_0)$，且 $C_{T,\delta}$ 和 p_0 在 (4.17) 中给出。利用 $|H_p| \le C$ 和 $\varepsilon|y_0| \le h$ 可得

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

现在令 $\alpha \in \mathcal{A}$ 满足

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

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

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

其中我们用到了 $2q_{y_0} = q_{x_0}+q_{z_0}$，并且对于任意 x、y、$z\in\mathbb{R}^d$ 和 $g\in W^2,\infty(\mathbb{R}^d)$，$|g(x)+g(z)-2g(y)|\le\|g\|_{W^2,\infty}(|x-y|^2+|z-y|^2+|x+z-2y|)$。由杨氏不等式可得 $|x_0-y_0||q_{x_0}|+|z_0-y_0||q_{z_0}|\le 2|x_0-y_0|^4+2|z_0-y_0|^4+h^4$。此外，利用 C_T、δ 和 ψ 的定义，可得 (4.20) 左端为 $\le Ce^{C_1(T-t_0)}(\delta+\delta^{-1}\psi(x_0,y_0,z_0)) = Ce^{C_1(T-t_0)}\gamma_0 + Ce^{C_1(T-t_0)}h^4/\delta$，其中 $C>0$ 仅依赖于 $\|q\|_{W^2,\infty}$、$\|c_\alpha\|_{W^2,\infty}$ 和 $\|f_\alpha\|_{W^2,\infty}$。

现在将 (4.16)、(4.18) 和 (4.19) 的两倍相加，得到

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

在第二个不等式中，我们使用了 $\delta\ge\sqrt{h}$。最后，由此及 (4.15) 可得

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

其中 $C>0$ 仅取决于 d、N 以及 q、c、f 的正则性假设。因此，若 C_1 足够大（仅取决于这些假设），便会得到矛盾，从而完成 (4.12) 的证明，也就完成了证明。□

5. Generalization: a PDE perspective

本节中，我们讨论具有一般哈密顿量的 HJB 方程的策略迭代。为便于使用 Legendre变换，我们将系统写成正向时间形式。将其转化为反向时间形式也很容易。

假设 $\mathcal{H}:\ [0,T]\times\mathbb{R}^d\times\mathbb{R}^d\to\mathbb{R}$ 是连续的，且 $\mathcal{H}(t,\ x,\ p)$ 关于 p是凸的。令 $\mathcal{L}(t,\ x,\ \mu)$ 为 $\mathcal{H}$的 Legendre变换，即

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

我们总有以下不等式 $\mathcal{L}(t,x,\mu)+\mathcal{H}(t,x,p)\ge p\cdot\mu$，且当且仅当 $\mu=\nabla_p\mathcal{H}(t,x,p)$ 时等号成立，并且当且仅当 $p=\nabla_\mu\mathcal{L}(t,x,\mu)$。

HJB 方程为

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

在某些假设下（见 [3, 42]），v 是利普希茨连续的，这是一个经典结果。因此，我们可以假设

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

现在我们取 $m_1 := \min_{|p|=2M,\ t\in[0,T],\ x\in\mathbb{R}^d} \mathcal{H}(t,x,p)$，以及 $m_2 \ge \max_{|p|=3M,\ t\in[0,T],\ x\in\mathbb{R}^d} [\mathcal{H}(t,x,p) - m_1]/M$，并且可以假设 $m_2 \ge 2$。然后定义

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

不难验证，$\mathcal{H}$ 关于其所有变量都是连续的，并且关于 p是凸的。由于 (5.2)，将 $\mathcal{H}$ 替换为 $\mathcal{H}$后，v 也是 (5.1) 的一个解。此外，对于 $N := m_2/2 \ge 1$，我们有

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

我们将 $\mathcal{L}$ 定义为 $\mathcal{H}$的 Legendre变换。由于目标是逼近 v，因此只需研究 $\mathcal{H}$ 和 $\mathcal{L}$，而不是 $\mathcal{H}$ 和 $\mathcal{L}$。从现在起，略微滥用记号，我们分别将 $\mathcal{H}$ 和 $\mathcal{L}$ 记作 $\mathcal{H}$ 和 $\mathcal{L}$。

使用修改后的算子，我们可以考虑半离散化。对于 $h>0$,

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

与之前一样，$N \ge \|\nabla_p \mathcal{H}\|_\infty/2$ 保证有限差分格式是单调的。我们还假设存在 $C>0$，使得对于所有 t、x、$p\in[0,T]\times\mathbb{R}^d\times\mathbb{R}^d$,

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

实际上，对于修改后的算子，我们可以将 $C(1+|p|)$ 替换为仅有的 C。由于类似，我们不讨论（5.1）的时空离散化。

现在我们给出（5.4）的迭代格式。固定一个较小的 $h>0$，从一个一致有界且利普希茨连续的函数 $v_0^h(t,x)$ 开始，然后按以下方式迭代计算 v_n^h。对于 $n\ge 1$，令 $v_n^h(t,x)$ 为下列方程的解：

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

其中，我们记 $\mu_n^h(t,x) := \nabla_p\mathcal{H}(t,x,\nabla^h v_n^h)$。注意，由于 $\mu_n^h \le 2N$，$\mathcal{L}(t,x,\mu_n^h)$ 是有限的。本质上，v_n^h 求解的是（5.4）的一个线性化方程。

令 v_n^h（对每个具有给定 v_0^h 的 $n\ge 1$），v^h 和 v 分别为方程 (5.6)、(5.4) 和 (5.1) 的利普希茨连续解。我们有如下单调性。

命题 5.1. 假设 $N \geq \max\{1, \|\nabla_p \mathcal{H}\|_\infty/2\}$，且 $\mathcal{H}(t,x,p)$ 关于 p 是凸的，并满足 (5.3) 和 (5.5)。设 q 和 v_0^h 对所有 $h > 0$ 均一致有界且利普希茨连续。那么，解 v_n^h 对所有 $n \geq 1$ 和 $h > 0$ 均一致有界。此外，对所有 $n \geq 0$,

$$v_{n+1}^h \leq v_n^h \quad in\ [0,T] \times \mathbb{R}^d.$$

我们还有以下收敛结果。

定理 5.2。在命题 5.1 的假设下，对于所有 $R \geq 1$，存在一个仅依赖于 T 和这些假设的常数 C，使得对于所有 $t \in [0,T]$，我们有

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

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

此外，我们有 $\sup_{(t,x)\in[0,T]\times\mathbb{R}^d} |v^h(t,x) - v(t,x)| \leq C\sqrt{h}$。

接下来，令 $\mathcal{H}$ 具有如下形式：$\mathcal{H}(t,x,p) := \sup_{a\in A} [c(t,x,a) + p \cdot f(t,x,a)]$,，其中 A 是某个集合，c：$[0,T] \times \mathbb{R}^d \times A \to \mathbb{R}$ 且 f：$[0,T] \times \mathbb{R}^d \times A \to \mathbb{R}^d$。

定理 5.3。在定理 5.2 的假设下，假设 $c(t,\cdot,a)$ $f(t,\cdot,a)$ 在 $W^{2,\infty}(\mathbb{R}^d)$ 中一致地有界，对所有 $t \in [0,T]$ 和 $a \in A$ 均成立。则对每个 $t \in [0,T]$，对几乎处处的 $x \in \mathbb{R}^d$，我们有

$$\alpha(t_h, x_h, \nabla^h v^h(t_h,x_h)) \to \alpha(t,x,\nabla v(t,x)) \quad as\ h \to 0,$$

当 $h \to 0$ 时，$[0,T] \times \mathbb{R}^d \ni (t_h, x_h) \to (t,x)$。

此外，存在一个仅取决于这些假设的 $C > 0$，使得对于所有 $h \in (0,1)$、$t \in [0,T]$ 和 $x, y \in \mathbb{R}^d$, $v^h(t,x+y) + v^h(t,x-y) - 2v^h(t,x) \leq C\exp(CT)(|y|^2 + \sqrt{h})$。

6. Conclusion

本文研究了连续时间最优控制问题中策略迭代的收敛速率。为克服适定性问题，我们通过采用有限差分添加黏性项，考虑一种半离散格式。我们证明，该半离散格式的策略迭代快速指数收敛，并给出了半离散格式与最优控制之间差异的界。我们还研究了时空离散格式，其中空间和时间均进行了离散化。

本工作有几个可拓展的方向。首先，模型参数满足什么条件时，策略迭代 (2.7)–(2.8) 会指数快速收敛？例如，对于 $f(t,x,a) = a, c(t,x,a) = \frac{1}{2}|a|^2$ 和 $q \equiv 0$，HJB 方程为 $\partial_t v - \frac{1}{2}|\nabla v|^2 = 0$ 且 $v(T,x) = 0$，其解为 $v_* \equiv 0$。另一方面，在适当的初始化下，策略迭代得到 $v_n(t,x) = c_n(t)x^2$，其中 $c_1(t) = \frac{1}{2}$。容易验证，对于 $n \geq 1$，$c_n(t) \leq 2^{-n}$，因此我们得到 v_n 在任意紧集上指数收敛到 v_*。然而，目前尚不清楚应对模型参数施加哪些恰当条件，才能使策略迭代指数快速收敛。将策略迭代应用于微分博弈，并设计高效的数值格式（例如参见 [21]），也很值得研究。关于使用策略迭代数值求解完全非线性 HJB 和 HJBI 方程的工作，可参见 [25, 39]。

References

[1] M. Abu-Khalaf 和 F. L. Lewis。利用神经网络哈密顿-雅可比-贝尔曼方法，为带饱和执行器的非线性系统设计近似最优控制律。 *Automatica*，41(5):779–791，2005。[2] S. Bansal、M. Chen、S. Herbert 和 C. J. Tomlin。哈密顿-雅可比可达性：简要概述与最新进展。载于 *56th Annual Conference on Decision and Control (CDC)*，第 2242–2253 页，2017。[3] G. Barles。一阶哈密顿-雅可比方程的正则性结果。 *Differential Integral Equations*，3 1 103 125，1990 ( ): – 。[4] G. Barles。一阶哈密顿-雅可比方程的唯一性与正则性结果。

Indiana Univ. Math.，( ): – 。[5] R. *J.* 39 2 443 466, 1990

Bellman。 *Dynamic programming*。普林斯顿大学出版社，1957。[6] J. Z. Ben-Asher。 *Optimal Control Theory with Aerospace Applications*。AIAA教育系列。AIAA，2010。[7] D. P. Bertsekas。 *Dynamic programming and optimal control. Vol. II. Approximate dynamic programming*。Athena Scientific，第四版，2012。[8] D. P. Bertsekas。最优控制与自适应动态规划中的价值迭代与策略迭代。 *IEEE Trans. Neural Netw. Learn. Syst.* 28 3 500 509, 2017，( ): – 。[9] T.

Bian 和 Z.-P. Jiang。用于数据驱动自适应最优控制设计的价值迭代与自适应动态规划。 *Automatica*，71:348–360，2016。[10] L. Blackmore、B. A.cikme.se 和 D. P. Scharf。利用凸优化实现火星着陆的最小着陆误差动力下降制导。 *J. Guid. Control Dyn*，33(4):1161–1171，2010。[11] F. Cagnetti、D. Gomes 和 H. V. Tran。哈密顿-雅可比方程半离散化格式的收敛性：伴随方法的新途径。 *Appl. Numer. Math.*，73:2–15，2013。[12] M. Chen 和 C. J. Tomlin。哈密顿-雅可比可达性：无人空域管理领域的一些最新理论进展与应用。 *Annu. Rev. Control Robot. Auton. Syst.*，1(1):333–358，2018。[13] M. G. Crandall 和 P.-L. Lions。哈密顿-雅可比方程解的两种近似。

Math.，( ): – 。[14] *Comp.* 43 167 1 19, 1984

H. Dong 和 N. V. Krylov。具有 C^1 和 C^2 系数的退化线性抛物方程有限差分近似的收敛速度。 *Electron. J. Differential Equations*，第 102 号，25 页，2005。[15] H. Dong 和 N. V. Krylov。柱状区域中具有利普希茨系数的抛物型贝尔曼方程有限差分近似的收敛速度。 *Appl. Math. Optim.*，56(1):37–66，2007。[16] L. C. Evans。利用增生算子方法求解某些非线性偏微分方程。 *Israel J. Math.* 36 3 4 225 247, 1980，(-): – 。[17]

W. H. Fleming 和 H. M. Soner。 *Controlled Markov processes and viscosity solutions* 25 *Stochastic*，卷（ *Modelling and Applied Probability* 2006 *Controlled Markov processes and viscosity solutions* 25 *Stochastic*）。施普林格，第二版， *Modelling and Applied Probability* 2006 *Controlled Markov processes and viscosity solutions* 25 *Stochastic*。[18] T. Haarnoja、A. Zhou、P. Abbeel 和 S. Levine。软演员-评论家算法：采用随机演员的离策略最大熵深度强化学习。载于 *35th International Conference on Machine Learning (ICML)* 1861 1870, 2018，第 – 页。[19]

B. Hambly、R. Xu 和 H. Yang。有限时域内含噪线性二次调节器的策略梯度方法。 *SIAM J. Control Optim.*，59(5):3359–3391，2021。[20] R. A. Howard。 *Dynamic programming and Markov processes*。麻省理工学院技术出版社；John Wiley & Sons, Inc.，1960。[21] K. Huang 、 X. Di 、 Q. Du 和 X. Chen。自动驾驶车辆速度控制的博弈论框架：衔接微观微分博弈与宏观平均场博弈。

Discrete Contin. Dyn. Syst. Ser.，( ): – 。[22] *B* 25 12 4869 4903, 2020

Y.-J. Huang、Z. Wang 和 Z. Zhou。熵正则化随机控制问题中策略改进的收敛性。2022。arXiv:2209.07059。[23] H. Ishii 和 P.-L. Lions。完全非线性二阶椭圆偏微分方程的黏性解。 *J. Differential Equations*，83(1):26–78，1990。[24] M. I. Kamien 和 N. L. Schwartz。 *Dynamic Optimization: The Calculus of Variations and Optimal Control in Economics and Management* 31 *Advanced Textbooks in Economics*，卷（ *Dynamic Optimization: The Calculus of Variations and Optimal Control in Economics and Management* 31 *Advanced Textbooks in Economics*）。爱思唯尔，第二版，1991。[25] E. L. Kawecki 和 T. Sprekeler。间断伽辽金与 C^0-IP 有限元方法对周期性哈密顿-雅可比-贝尔曼-艾萨克斯问题的近似及其在数值均匀化中的应用。

ESAIM Math.，( ): – 。 *Model. Numer. Anal.* 56 2 679 704, 2022

[26] B. Kerimkulov、D. ˇSiˇska 和 L. Szpruch。受控扩散过程的 Howard 策略改进算法的指数收敛性与稳定性。 *SIAM J. Control Optim.*，58(3):1314–1340，2020。[27] N. V. Krylov。具有 Lipschitz 系数的贝尔曼方程有限差分近似的收敛速率。 *Appl. Math. Optim.*，52(3):365–399，2005。[28] J. Lee 和 R. S. Sutton。连续时间与空间中强化学习问题的策略迭代——基础理论与方法。 *Automatica*，126:论文编号 109421，15，2021。[29] F. L. Lewis 和 D. Vrabie。用于反馈控制的强化学习与自适应动态规划。*IEEE Circuits Syst. Mag.*，9(3):32–50，2009。[30] Y. Li、K. H. Johansson、J. M ̊artensson 和 D. P. Bertsekas。确定性最优控制的数据驱动滚动优化，2021。arXiv:2105.03116。[31] E. Nolasco、V. S. Vassiliadis、W. K¨ahm、S. D. Adloor、R. Al Ismaili、R. Conejeros、T. Espaas、N. Gangadha-ran、V. Mappas 和 F. Scott。化学工程中的最优控制：过去、现在与未来。 *Comput.*，：。 *Chem. Eng.* 155 107528, 2021
[32] S. Osher 和 C.-W. Shu。哈密顿-雅可比方程的高阶本质无振荡格式。 *SIAM J. Numer. Anal.* 28 4 907 922, 1991 ，( ): – 。[33] W. B. Powell。 *Approximate dynamic programming: solving the curses of dimensionality*。威利概率与统计丛书。John Wiley & Sons, Inc，2007。[34] M. L. Puterman。 *Markov decision processes: discrete stochastic dynamic programming*。威利概率与数理统计丛书。约翰·威立父子公司，1994。[35] C. Reisinger、W. Stockinger 和 Y. Zhang。有限时域连续时间随机控制问题的策略梯度方法线性收敛，2022。arXiv:2203.11758。[36] I. M. Ross、M. Karpenko 和 R. J. Proulx。求解若干图论控制问题的非光滑微积分。 *IFAC-PapersOnLine*，49(18):462–467，2016。[37] S. Sethi 和 X. Y. Zhou。随机动态作业车间与分层生产计划。 *IEEE Trans.*，( ): *Automat. Control* 39 10 2061 2076, 1994
– 。[38] D. Silver、J. Schrittwieser、K. Simonyan、I. Antonoglou、A. Huang、A. Guez、T. Hubert、L. Baker、M. Lai 和 A. Bolton。无需人类知识掌握围棋。 *Nature*，550(7676):354–359，2017。[39] I. S mears 和 E. S¨uli。具有 Cordes系数的哈密顿-雅可比-贝尔曼方程的间断伽辽金有限元近似。 *SIAM J. Numer. Anal.*，52(2):993–1016，2014。[40] R. S. Sutton 和 A. G. Barto。 *Reinforcement learning: an introduction*。自适应计算与机器学习丛书。麻省理工学院出版社，第二版，2018。[41] W. Tang 和 X. Y. Zhou。连续时间 *Q*学习的遗憾。 2023+。准备中。[42] H. V. TRAN。 *Hamilton-Jacobi equations—theory and applications* 213 *Graduate Studies in Math-* ，第 *ematics* 2021 卷。美国数学会，。[43] D. Vrabie、O. Pastravanu、M. Abu-Khalaf 和 F. L. Lewis。基于策略迭代的连续时间线性系统自适应最优控制。 *Automatica*，45(2):477–484，2009。[44] D. Wang、H. He 和 D. Liu。自适应评论家非线性鲁棒控制：综述。 *IEEE Trans. Cybern.*，47 10 3429 3451，2017 ( ): – 。[45] Q. Wei、D. Liu 和 H. Lin。离散时间非线性系统最优控制的价值迭代自适应动态规划。 *IEEE Trans. Cybern.*，46(3):840–853，2015。

(W. TANG) Department of Industrial Engineering and Operations Research, Columbia University, S.W. Mudd Building, 500 W 120th St, New York, NY 10027
Email address: wt2319@columbia.edu

（H. V. Tran） Department of Mathematics, University of Wisconsin Madison, Van Vleck Hall, 480Lincoln Drive, Madison, WI 53706*Email address*: hung@math.wisc.edu

（Y. P. ZHANG） Department of Mathematics and Statistics, Auburn University, Parker Hall, 221Roosevelt Concourse, Auburn, AL 36849*Email address*: yzhangpaul@auburn.edu