环境异质性下采用线性函数逼近的联邦时序差分学习

Han Wang *hw2786@columbia.edu*
哥伦比亚大学电气工程系

Aritra Mitra *amitra2@ncsu.edu*
电气与计算机工程系 北卡罗来纳州立大学

Hamed Hassani *hassani@seas.upenn.edu*
电气与系统工程系 宾夕法尼亚大学

George J. Pappas *pappasg@seas.upenn.edu*
宾夕法尼亚大学电气与系统工程系

James Anderson *james.anderson@columbia.edu*
哥伦比亚大学电气工程系

已在 OpenReview 上评审： *https://openreview.net/forum?id=hdQspgyFrk*

arXiv:2302.02212v2 [cs.LG] 1 Jul 2024

摘要

我们通过研究策略评估问题，开启了对环境异质性下联邦强化学习的探索。我们的设定涉及 N 个智能体，它们与共享相同状态与动作空间、但奖励函数和状态转移核各不相同的环境交互。假设智能体可以通过中央服务器进行通信，我们提出以下问题：交换信息能否加快共同策略的评估过程？为回答这一问题，我们首次对采用线性函数逼近的联邦时序差分（TD）学习算法进行了全面的有限时间分析，同时考虑了马尔可夫采样、智能体所处环境的异质性，以及为节省通信而进行的多次本地更新。我们的分析关键依赖于若干新颖的方法：（i）推导时序差分固定点的扰动界，并刻画其与智能体底层马尔可夫决策过程（MDP）异质性的关系；（ii）引入虚拟马尔可夫决策过程，以紧密近似联邦时序差分算法的动态过程；（iii）利用虚拟马尔可夫决策过程，明确建立与联邦优化之间的联系。综合这些方法，我们证明，在低异质性情形下，交换模型估计可使收敛速度随智能体数量实现线性加速。我们的理论贡献意义重大，因为这是多智能体／联邦强化学习领域首个此类成果，可与异构联邦优化领域大量类似的研究成果相互补充。

1 引言

在广受欢迎的联邦学习（FL）范式中（Konečn`等人，2016年；McMahan 等人，2017年），一组智能体旨在找到一个能够解释其集体观测的共同统计模型。合作的动机在于：如果生成智能体观测的底层分布“相似”，那么每个智能体最终学到的模型就可能比仅使用自身

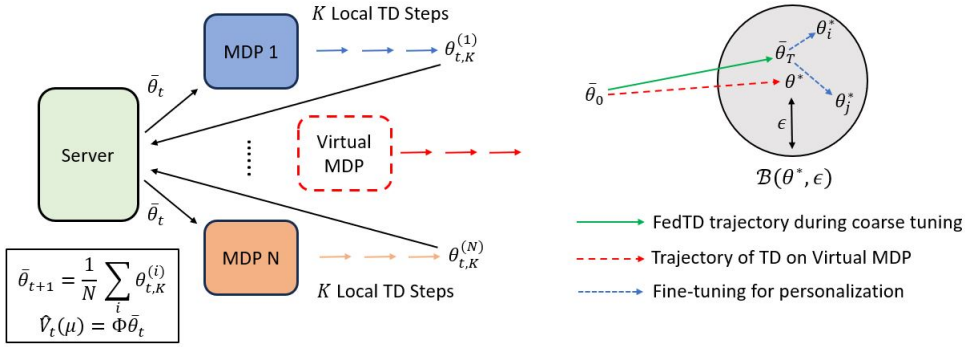


图 1：（**左**） FedTD(0) 的工作原理示意图。每个智能体在自己的马尔可夫决策过程中执行 K 本地时序差分更新步，并将更新后的模型发送到服务器。虚拟马尔可夫决策过程用于近似 FedTD(0) 的动态。服务器上的全局模型 θ_t 用于构建与策略 μ 相关的价值函数的线性参数化近似。（**右**） FedTD(0) 帮助每个智能体快速收敛到以虚拟马尔可夫决策过程的最优参数 θ^* 为中心的球 $\mathcal{B}(\theta^*, \epsilon)$ 内。此处，ϵ 体现了各智能体马尔可夫决策过程的异质性。使用 FedTD(0) 的输出 θ_T 后，每个智能体 i 便可基于自己的数据进行微调，从而精确收敛到自身的最优参数 θ_i^*。

数据时学到的模型“更好”。这一思想已通过经典的联邦学习算法 FedAvg （及其诸多变体）形式化：智能体通过中央服务器通信并共享本地模型，同时对原始数据保密。为提高通信效率——这是联邦学习中的一项关键考量——智能体会在连续的通信轮次之间执行多次本地模型更新。大量文献分析了 FedAvg 的性能，重点关注统计异质性，即由智能体底层数据分布的差异所导致的异质性（Sahu 等人，2018年；Khaled 等人，2019年；2020年；Li 等人，2019年；Koloskova 等人，2020年；Woodworth 等人，2020b；Malinovskiy 等人，2020年；Pathak 与 Wainwright，2020 年；Wang 等人，2020 年；Karimireddy 等人，2020b ；Acar 等人，2021 年；Gorbunov 等人，2021年；Mitra 等人，2021年；Mishchenko 等人，2022年）。值得注意的是，上述研究聚焦于在分布式优化框架下建模的监督学习问题。然而，对于多个智能体与可能各不相同的环境进行交互时的序贯决策，异质性所产生的影响几乎无人知晓。这正是我们希望通过本研究填补的空白。

近期的综述论文（Qi 等人，2021年）介绍了一种联邦强化学习（FRL）框架，该框架将联邦学习（FL）的一些关键理念融入强化学习（RL）；本文详细讨论了联邦强化学习在机器人学（Liu 等人，2019年）、自动驾驶（Chen 等人，2015年）和边缘计算（Wang 等人，2019年）中的应用。由于强化学习算法通常需要大量样本才能达到可接受的精度，联邦强化学习旨在通过利用多个智能体与相似环境交互所获得的信息来提高样本效率。重要的是，与标准联邦学习一样，联邦强化学习框架要求智能体对其原始数据（例如奖励、状态和动作）保密，并遵守严格的通信约束。

本工作的动机与范围。尽管联邦强化学习是一个很有前景的想法，但在现实中，不同智能体最终与完全相同的环境交互的情况并不多见。遗憾的是，几乎所有多智能体强化学习（MARL）和联邦强化学习研究都默认了这一假设（Doan 等人，2019年；Liu 与 Olshevsky，2021a；Khodadadian 等人，2022年；Shen 等人，2023年）。有鉴于这一略显不切实际却普遍存在的假设，**本文的主要动机是构建一个系统的理论框架，以探究混合来自不同马尔可夫过程的信息时会产生怎样的预期结果**。这一问题具有基础性意义；尽管我们从联邦强化学习的视角出发来探讨它[1]，，但同样可以将其与随机控制和估计问题联系起来：在这些问题中，人们试图融合由具有噪声输入的非同一动力系统生成的数据（Wang 等人，2022b；Guo 等人，2023年；Xin 等人，2023年）。

[1]正如统计异质性是联邦学习中的一项重大挑战，环境异质性也被确定为联邦强化学习中的一项关键开放性挑战（Qi 等人，2021年）。

为了以严谨的方式开启对联邦强化学习中异质性的研究，我们聚焦于最简单的强化学习问题，即策略评估。我们的设定涉及 N 个智能体，每个智能体都与一个建模为马尔可夫决策过程的环境进行交互。各智能体的马尔可夫决策过程具有相同的状态与动作空间，但奖励函数和状态转移核各不相同，从而体现环境异质性。每个智能体都致力于计算与共同策略 μ相对应的折扣累积奖励（价值函数）。值得注意的是，由 μ 所诱导的价值函数可能因环境而异。这引出了我们要探讨的核心问题：智能体能否利用来自可能不同的马尔可夫决策过程的信息，加快自身价值函数的学习过程？正如我们接下来将要说明的，即使对于策略评估而言，这个问题也并非易于回答；因此，我们选择以策略评估作为研究起点。尽管如此，近期研究表明，只需对策略评估中的时序差分学习分析作少许修改（Srikant 与 Ying，2019年），便可分析用于控制的Q学习（Chen 等人，2019年）。因此，我们认为，本文的研究成果也可以适当地扩展到Q学习等控制算法。

上述联邦强化学习设置的一个典型应用是自动驾驶系统：不同地理位置的车辆共享能够体现其学习经验的本地模型，以训练一个可从所有车辆的集体探索数据中获益的共享模型。车辆（智能体）执行的操作基本相同（例如转向、制动、加速等），但可能处于不同的环境中（例如道路和天气状况、路线、驾驶法规等）。

1.1 我们的贡献

我们研究时序差分（TD）学习算法的联邦版本 TD(0)（Sutton，1988年）。我们将该算法的结构称为 FedTD(0)，其流程如下。在每次迭代中，每个智能体根据策略 μ执行一个动作、观测奖励，并根据其自身马尔可夫决策过程（MDP）转移到新状态。随后，它使用 TD(0)和线性函数逼近来更新一个用于逼近自身价值函数的本地模型。为以通信高效的方式利用其他智能体的数据，每个智能体会定期与中央服务器同步，并在两次同步之间执行多次本地模型更新——如图1所示。值得注意的是，与联邦学习（FL）一样，智能体只交换模型，从不交换其个人观测。我们对环境异质性下的 FedTD(0)进行了全面分析，并作出以下贡献：

1. **异质性对 TD(0) 不动点的影响。** 为理解 FedTD(0) 的行为，我们首先提出以下问题：在这些马尔可夫决策过程上，转移核和奖励函数的异质性会如何体现为 *TD(0)*（采用线性函数逼近）的长期行为差异？ 定理 1 通过刻画扰动马尔可夫决策过程会如何扰动该马尔可夫决策过程的 TD(0) 不动点，给出了答案。为得到这一结果，我们结合了马尔可夫链和线性方程扰动理论的结果。定理 1 首次为 TD(0) 不动点建立了扰动结果，并补充了强化学习文献中类似的研究成果，例如 模拟引理，该引理由 Kearns & Singh (2002) 提出。因此，定理 1 也可作为强化学习中具有独立价值的工具。

2. **虚拟 MDP 框架。** 在联邦学习算法中，例如 FedAvg，智能体损失函数的负梯度均值驱动 FedAvg 的迭代序列趋向于全局损失函数的最小化点。在我们的设定中，并不存在这样的全局损失函数。 那么，对不同马尔可夫决策过程的 *TD(0)* 更新方向取平均后，最终会得到什么？ 为回答这个问题，我们在第 3.1 节构建了一个虚拟马尔可夫决策过程，并刻画了该虚拟马尔可夫决策过程的若干重要性质，以辅助后续分析。在此过程中，我们推导出一个简单却关键的结果（命题 1），涉及与非周期且不可约马尔可夫链相关联的马尔可夫矩阵的凸组合。这个结果似乎是新的，也可能具有独立的研究价值。

3. **马尔可夫采样与异质性下的线性加速。** 我们最重要的贡献是首次对一种联邦强化学习算法进行了分析，该算法 *FedTD(0)*，同时考虑了线性函数逼近、马尔可夫采样、多次本地更新和异质性。 在定理 2中，我们证明，经过 T 个通信轮次，每轮进行 K 次本地模型更新~步后， FedTD(0) 保证以 $O(1/NKT)$ 的速率收敛至每个智能体最优参数的一个邻域。该邻域的大小取决于各智能体

的马尔可夫决策过程之间的异质性程度。这一结果的关键意义在于，在低异质性情形下，每个智能体都能通过协作实现 N 倍的收敛线性加速，并快速收敛至自身最优参数附近。我们可以将其视为一个“粗调阶段”。与联邦学习中的常见做法一致（Collins 等人，2022年），每个智能体都可以使用 FedTD(0) 的解，再基于自身数据进行微调（个性化）。图 1 直观地展示了这一过程。定理 2 意义重大，因为这是联邦强化学习领域的首个结果，补充了大量考虑异质性影响的联邦优化研究（Sahu 等人，2018年；Khaled 等人，2019年；2020年；Li 等人，2019年；Koloskova 等人，2020年；Woodworth 等人，2020b；Malinovskiy 等人，2020年；Pathak 与 Wainwright，2020年；Wang 等人，2020年；Karimireddy 等人，2020b；Acar 等人，2021年；Gorbunov 等人，2021年；Mitra 等人，2021年；Mishchenko 等人，2022年）。

4. **新型证明框架**。 人们可能会认为，定理 2 的证明不过是将标准的 FedAvg 分析与时序差分学习的分析简单结合起来。下面简要说明实际情况并非如此，并将更详细的解释留到第 5.1 节。首先，在集中式时序差分分析（Bhandari 等人，2018年；Srikant 与 Ying，2019年）以及针对多智能体强化学习/联邦强化学习的现有分析中（Doan 等人，2019年；Liu 与 Olshevsky，2021a；Khodadadian 等人，2022年），在马尔可夫决策过程相同的情况下，更新规则的动态对应于单个马尔可夫决策过程。而在我们的设定中， FedTD(0) 的动态甚至可能不对应任何马尔可夫决策过程！因此，相较于现有的强化学习分析，我们需要新的工具。其次，现有的联邦学习分析本质上是分布式优化证明，而联邦时序差分学习并不对应于最小化任何固定损失函数。此外，与联邦学习中的独立同分布数据模型不同，每个智能体在 FedTD(0) 中观测到的数据元组均来自同一条马尔可夫轨迹。由此产生复杂的时间相关性，即使在单智能体情形下也难以处理。因此，我们也无法直接采用联邦学习的证明方法。为此，我们提出了一种新的分析框架，论证 FedTD(0) 的动态可以通过虚拟马尔可夫决策过程上的 TD(0) 动态来近似，误差项则刻画了各智能体马尔可夫决策过程的异质性。仔细追踪该误差项随时间的传播情况，便能反映异质性的影响；而要在马尔可夫采样和本地步数条件下确立线性加速，则需要付出更多努力。

5. **异质性引入的偏差**。 定理2 中的收敛结果包含一个由异质性导致的偏差项，即使将步长任意减小，该偏差项也无法消除。 这样的偏差项是否不可避免？ 我们在定理3中研究了 FedTD(0) 的“稳态”确定性版本，以探讨这个问题。即使在这个简单情形下，我们也证明，依赖于异质性的一种自然度量的偏差项，仍会不可避免地 出现在 FedTD(0) 的长期动力学中。该结果进一步揭示了异质性对联邦强化学习的影响。

1.2 相关工作

接下来，我们将讨论与本文最相关的几条研究脉络。

1. TD学习算法的有限时间分析。Tsitsiklis 和 Van Roy（1997）在其开创性论文中，运用随机逼近理论的工具，对带有价值函数逼近的时序差分（TD）学习算法（Sutton，1988年；Sutton 等人，1998年）进行了渐近收敛分析。几年后，Korda 与 La（2015年）的研究给出了TD学习的有限时间收敛速率。然而，Narayanan 与 Szepesvári（2017年）指出，Korda 与 La（2015年）中的证明存在一些问题。在第5节所述的独立同分布观测模型下，Dalal等人（2018年）和Lakshminarayanan 与 Szepesvári（2017年）成功解决了Korda 与 La（2015年）中的这些问题。尽管如此，对于具有挑战性的马尔可夫设置（本文所研究的情形），非渐近收敛分析仍然难以实现，直到Bhandari 等人（2018年）的研究出现。Bhandari 等人（2018年）巧妙地揭示了TD学习动态与梯度下降之间的一些联系，而Srikant 与 Ying（2019年）则提出了一种运用Stein方法的替代证明技术。Liu 与 Olshevsky（2021b）还给出了一种有趣的解释：他们认为，稳态时序差分方向相当于对一个经过适当选择的函数进行“梯度分裂”。最近，Mitra（2024年）基于一种新颖的归纳证明技术，为采用线性函数逼近的TD学习以及更一般的非线性压缩随机逼近方案给出了简洁的证明。尽管上述所有

研究均为策略评估任务提供了上界；关于极小极大下界，我们请读者参阅Khamaru 等人（2021年）。

2. 多智能体与联邦强化学习。Doan 等人（2019年）和 Liu 与 Olshevsky（2021a）分析了点对点网络中采用线性函数逼近的多智能体时序差分学习。这两种方法都未考虑本地步数或马尔可夫采样。Shen 等人（2023年）研究了异步演员-评论家算法的并行版本，并证明了线性加速结果，不过这一结果是在独立同分布采样假设下得出的。最近，Khodadadian 等人（2022年）和 Dal Fabbro 等人（2023年）研究了马尔可夫采样对联邦时序差分学习的影响。然而，上述论文均考虑所有智能体的马尔可夫决策过程相同的同质设定。相比之下，我们的工作必须应对这样一项挑战：理解混合来自不同马尔可夫决策过程的时序差分更新方向所产生的长期影响。据我们所知，只有 Jin 等人（2022年）和 Xie 和 Song（2023年）这两篇论文对联邦强化学习中的异质性进行了理论分析。然而，他们的分析仅限于简单得多、未采用函数逼近的表格设定。具体而言，Xie 和 Song（2023年）的工作仅给出了渐近结果，即没有提供有限时间收敛速率。此外，与我们的工作不同，Jin 等人（2022年）和 Xie 和 Song（2023年）都没有给出明确的线性加速结果。总之，我们率先为函数逼近、环境异质性和马尔可夫采样下的联邦强化学习建立了有限时间理论。在不同设定下，Zhang 等人（2024年）提出了 `FEDSARSA`算法来解决在策略联邦强化学习问题，而 Wang 等人（2023年）提出了 `FedLQR` 来解决联邦控制设计问题。关于联邦学习相关工作的更详细介绍见附录。

2 模型与问题表述

我们考虑由元组 $\mathrm{M} = \mathrm{S}, \mathrm{A}, \mathrm{R}, \mathrm{P}$ γ定义的马尔可夫决策过程（MDP）（Sutton 等人，1998年），其中 S 是大小为 n 的有限状态空间，A 是有限动作空间，P 是一组依赖于动作的马尔可夫转移核，R 是奖励函数，$\gamma \in (0, 1)$ 是折扣因子。我们考虑评估给定策略 μ 的价值函数 V_μ 的问题，其中 μ: S $\to$ A。策略 μ 诱导出一个由转移矩阵 P_μ 和奖励函数 R_μ 表征的马尔可夫奖励过程（MRP）。在策略 μ 作用于初始状态 s 时，$P_\mu(s, s')$ 表示从状态 s 转移到状态 s' 的概率，而$R_\mu(s)$ 是期望即时奖励。从初始状态 s 出发执行策略 μ 所获得的折扣期望累积奖励为：

$$V_\mu(s) = \mathbb{E}\left[\sum_{t=0}^{\infty} \gamma^t R_\mu(s_t) | s_0 = s\right],$$

其中 s_t 是马尔可夫链在时刻 t的状态。根据 Tsitsiklis& Van Roy (1997年)，我们知道 V_μ 是策略特定的贝尔曼算子 T_μ 的 $\mathbb{R}^n \to \mathbb{R}^n$不动点，即 $T_\mu V_\mu = V_\mu$，其中对于任意 $V \in \mathbb{R}^n$，

$$(T_\mu V)(s) = R_\mu(s) + \gamma \sum_{s' \in \mathcal{S}} P_\mu(s, s') V(s'), \ \forall s \in \mathcal{S}.$$

采用线性函数逼近的时序差分学习。我们考虑状态数量非常大的情形，因此直接计算价值函数 V_μ 在实际中不可行。为缓解维数灾难，一种常见方法（Sutton 等人，1998年）是对价值函数 V_μ采用低维线性函数逼近。设 $\{\Phi_k\}_{k=1}^d$ 为 d 中一组线性独立的基向量，$\Phi \in \mathbb{R}^{n \times d}$ 为以这些基向量为列的矩阵，即 Φ 的第 k 列为 Φ_k。那么，V_μ 在 $\{\Phi_k\}_{k=1}^d$ 张成空间中的参数化近似 $\hat{} V_\theta$ 可表示为 $\hat{} V_\theta = \Phi\theta$，其中 $\theta \in \mathbb{R}^d$ 为待学习的参数向量。值得注意的是，这种方法具有可处理性，因为 $d \ll n$。我们将 Φ 的第 s 行记为$\phi(s) \in \mathbb{R}^d$，并称其为对应于状态 s 的固定特征向量。我们记 $\hat{} V_\theta(s) = \phi(s)^\top \theta$，并作出标准假设（Bhandari 等人，2018年）：$\|\phi(s)\|^2 \le 1, \forall s \in \mathrm{S}$。

目标是在$\{\Phi_k\}_{k=1}^d$的张成空间中，找到 V_μ 的最佳线性近似。更准确地说，我们要寻找一个参数向量 θ^*，使 $\hat{} V_\theta$ 与 V_μ 之间的距离（按适当的意义）最小。当底层

马尔可夫决策过程未知时，实现这一目标最常用的方法之一是经典的 TD(0)算法。 TD(0)从初始猜测 $\theta_0 \in \mathbb{R}^d$ 开始。随后，在第 t次迭代中，执行给定策略 μ后，会观测到一个新的数据元组 $O_t = (s_t,\ r_t = R_\mu(s_t)\ s_{t+1})$，其中包含当前状态、即时奖励和下一状态。将 TD(0)的更新方向定义为 $g_t(\theta_t) \triangleq \left(r_t + \gamma\phi(s_{t+1})^\top \theta_t - \phi(s_t)^\top \theta_t\right)\phi(s_t)$。使用步长 $\alpha_t \in (0,\ 1)$，随后将参数 θ_t 更新为

$$\theta_{t+1} = \theta_t + \alpha_t g_t(\theta_t).$$

在一些温和的技术假设下，Tsitsiklis& Van Roy（1997年）证明了， TD（0）中的迭代序列几乎必然渐近收敛于向量 θ^*，其中 θ^* 是投影贝尔曼方程 $\Pi_D T_\mu(\Phi\theta^*) = \Phi\theta^*$的唯一解。这里， D 是一个对角矩阵，其对角元素由马尔可夫矩阵 π 的平稳分布的各元素给出 P_μ。此外， Π_D（$\cdot$）是关于内积$\langle\cdot,\ \cdot\rangle_D$的投影算子，它将向量投影到由$\{\phi_k\}_{k=1}^d$ 张成的子空间上。[2]

目标。我们研究一个多智能体强化学习问题，其中智能体与具有相似但不完全相同的马尔可夫决策过程进行交互，这些马尔可夫决策过程共享相同的状态与动作空间。所有智能体都希望评估同一个策略。我们的目标是理解：智能体能否通过利用其他智能体的数据，以更高的样本效率评估其自身马尔可夫决策过程的价值函数？现有研究监督学习和经验风险最小化中统计异质性的联邦学习分析，无法回答这一问题，因为我们的问题并不涉及最小化静态损失函数。因此，我们提出的上述问题并不简单，需要若干新的思想和工具。在下一节中，我们将通过系统地实现以下目标，开始构建这些工具。

目标 1。 正式定义智能体的马尔可夫决策过程中的模型异质性。

目标 2。 刻画在智能体的马尔可夫决策过程中运行 TD(0)算法时，这种模型异质性如何导致其不动点 出现差异。

目标 3。 引入虚拟马尔可夫决策过程这一概念，它对于推理结合非同一马尔可夫决策过程信息的算法的长期行为至关重要。

3 异构联邦强化学习

我们考虑一个联邦强化学习设置，其中包含 N 个智能体，它们与可能各不相同的环境进行交互。智能体 i 的环境由以下马尔可夫决策过程刻画：$\mathcal{M}^{(i)} = \langle \mathcal{S}, \mathcal{A}, \mathcal{R}^{(i)}, \mathcal{P}^{(i)}, \gamma\rangle$。虽然所有智能体共享相同的状态与动作空间，但其环境的奖励函数和状态转移核可能不同。我们关注一个策略评估问题：所有智能体都希望评估一个共同策略 μ，该策略诱导出由元组 $\{P_\mu^{(i)}, R_\mu^{(i)}\}_{i\in[N]}$ 刻画的马尔可夫奖励过程。[3] 智能体 i 旨在为其自身价值函数 $V_\mu^{(i)}$寻找一种线性参数化近似。显然，智能体 i 无需与任何其他智能体交互，只需运行 TD(0) 即可做到这一点。然而，我们提出的关键问题关乎**辅助信息的价值**：利用其他智能体的数据，相较于单独行动，它能否以更少的样本达到理想的近似水平？自然地，上述问题的答案取决于智能体马尔可夫决策过程的异质程度。因此，我们引入以下定义。

假设*1*。（马尔可夫核异质性）存在一个 $\epsilon > 0$，使得对于所有智能体 i、 $j \in [N]$，均有 $|P(i)(s, s') - P(j)(s, s')| \le \epsilon|P(i)(s, s')|$、 $\forall s$、 $s' \in \mathcal{S}$。这里，对于每个 $i \in [N]$， $P(i)(s, s')$ 表示矩阵 $P^{(i)}$ 的第（s， s'）个元素。

假设2。 （奖励异质性）存在一个 $\epsilon_1 > 0$，使得对于所有 i、 $j \in [N]$，均有 $\|R^{(i)} - R^{(j)}\| \le \epsilon_1$。

显然， ϵ 和 ϵ_1 的值越小，越能体现智能体的马尔可夫决策过程之间的相似性。假设所有智能体都可以通过中央服务器进行通信。通过这种通信，标准联邦学习任务旨在找到一个共同的

[2]我们将使用 $\|\cdot\|_D$ 表示由矩阵 D 诱导的范数，并使用 $\|\cdot\|$ 表示向量的标准欧几里得范数以及矩阵的 ℓ_2 诱导范数。[3]下文中，我们将省略 $P^{(i)}$ 和 $R^{(i)}$ 对策略 μ 的依赖关系。

模型，使其能够“拟合”所有智能体的数据。类似地，我们的目标是找到一个共同的参数 θ，使得 $\hat{V}_\theta = \Phi\theta$ 能够近似每个 $V^{(i)}_\mu$，$i \in [N]$。这个共同 θ 的作用是快速（即利用所有智能体的样本）提供一个粗略模型，供智能体用作热启动，然后再根据个人数据进行微调。这里存在一种天然的权衡：联邦化有助于更快地收敛到一个粗略模型，但如果智能体的马尔可夫决策过程差异很大，这样的模型可能无法准确地刻画任何智能体的价值函数。那么，更多数据究竟有益还是有害？

异质性对时序差分固定点的影响。 要回答上述问题，我们需要仔细理解针对马尔可夫决策过程的结构异质性假设（即假设1 和2）如何体现为 TD (0) 在这些马尔可夫决策过程上的长期动力学差异。由于长期动力学与不动点密切相关，我们首先着手刻画不同马尔可夫决策过程之间 TD (0) 不动点的“接近程度”。为此，我们作出以下标准假设。

假设 3。 对于每个 $i \in [N]$，由策略 μ 诱导、对应于状态转移矩阵 $P^{(i)}$ 的马尔可夫链是非周期且不可约的。

上述假设意味着每个 $i \in [N]$ 都存在唯一的平稳分布 $\pi^{(i)}$；令 $D^{(i)}$ 为一个对角矩阵，其对角线上的元素为 $\pi^{(i)}$。对于每个智能体 i，我们用 θ^*_i 表示投影贝尔曼方程 $\Pi_{D^{(i)}} T^{(i)}_\mu(\Phi\theta^*_i) = \Phi\theta^*_i$ 的解，该方程对应于智能体 i。换言之，θ^*_i 是 $V^{(i)}_\mu$ 在 $\{\phi_k\}_{k=1}^d$ 张成空间中的最佳线性近似。根据第2节，我们知道，TD(0) 在智能体 i 的马尔可夫奖励过程上的迭代序列几乎必然渐近收敛到 θ^*_i。我们的目标是将差距 $\|\theta^*_i - \theta^*_j\|$ 表示为假设1和假设2中出现的异质性参数 ϵ 和 ϵ_1 的函数。我们将利用的关键观察是，对于每个 $i \in [N]$，θ^*_i 是线性方程 $A_i\theta^*_i = b_i$ 的唯一解，其中 $\bar{A}_i = \Phi^\top D^{(i)}(\Phi - \gamma P^{(i)}\Phi)$ 且 $\bar{b}_i = \Phi^\top D^{(i)} R^{(i)}$.。对于智能体 $j \neq i$，将 $\bar{A}_j$ 和 $\bar{b}_j$ 视为 A_i 和 b_i 的扰动版本后，我们便可以借助线性方程扰动理论（Horn & Johnson, 2012a，第5.8章）的结果来界定 $\|\theta^*_i - \theta^*_j\|$。为此，我们首先回顾马尔可夫链扰动理论（O’cinneide, 1993）中的一项结果，该结果表明，在假设1下，对于任意一对 i、$j \in [N]$，其平稳分布 $\pi^{(i)}$ 和 $\pi^{(j)}$ 都彼此接近。

引理1。 *(平稳分布的扰动界)* 假设 *1* 成立。则对于任意一对智能体 $i, j \in [N]$，平稳分布 $\pi^{(i)}$ 和 $\pi^{(j)}$ 满足：

$$\|\pi^{(i)} - \pi^{(j)}\|_1 \leq 2(n-1)\epsilon + \mathcal{O}(\epsilon^2). \tag{1}$$

现在我们将使用上述结果来界定 $\|A_i - A_j\|$ 和 $\|b_i - b_j\|$.。为陈述我们的结果，我们作出标准假设：对于每个 $i \in [N]$，都有 $|R^{(i)}(s)| \leq R_{\max}, \forall s \in \mathcal{S}$，即奖励一致有界。在 (Tsitsiklis & Van Roy, 1997年) 中，已证明 $-A_i$ 是一个负定矩阵；因此，存在 $\exists \delta_1 > 0$ 使得 $\|A_i\| \geq \delta_1, \forall i \in [N]$。我们还假设存在 $\exists \delta_2 > 0$ 使得 $\|b_i\| \geq \delta_2, \forall i \in [N]$。在下面给出的第一个技术结果中，我们将给出 TD 不动点的扰动界。

定理1. *(TD(0) 不动点的扰动界)* 对所有 i，$j \in [N]$，我们有：

1. $\|\bar{A}_i - \bar{A}_j\| \leq A(\epsilon) \triangleq \gamma\sqrt{n}\epsilon + (1+\gamma)\left(2(n-1)\epsilon + \mathcal{O}(\epsilon^2)\right)$.
2. $\|\bar{b}_i - \bar{b}_j\| \leq b(\epsilon, \epsilon_1) \triangleq R_{\max}\left(2(n-1)\epsilon + \mathcal{O}(\epsilon^2)\right) + \mathcal{O}(\epsilon_1)$。

3. 假设 $\exists H > 0$ *s.t.* $\|\theta^*_i\| \leq H, \forall i \in [N]$。令 $\kappa(A_i)$ 为 A_i 的条件数。则：

$$\|\theta^*_i - \theta^*_j\| \leq \Gamma(\epsilon, \epsilon_1) \triangleq \max_{i \in [N]} \left\{ \frac{\kappa(\bar{A}_i) H}{1 - \kappa(\bar{A}_i)\frac{A(\epsilon)}{\delta_1}} \left(\frac{A(\epsilon)}{\delta_1} + \frac{b(\epsilon, \epsilon_1)}{\delta_2} \right) \right\}.$$

讨论。 定理1 揭示了马尔可夫决策过程的奖励和转移核中的异质性，如何从固定点视角映射为 TD在此类马尔可夫决策过程上的极限行为差异。它将这样一种直觉形式化：如果异质性程度（由 ϵ 和 ϵ_1 所刻画）较小，那么.....也

智能体各自的马尔可夫决策过程的 TD（0）极限点之间的差距也较小。这一结果具有新颖性，并补充了强化学习文献中的类似扰动结果，例如模拟引理（Kearns & Singh，2002）。[4]

接下来，我们将介绍虚拟马尔可夫决策过程这一关键概念，并基于定理 1 阐述该虚拟马尔可夫决策过程的性质与各智能体各自马尔可夫决策过程的性质之间的关系。

3.1 虚拟马尔可夫决策过程

在标准的联邦学习设定中，目标通常是最小化全局损失函数 $f(x) = (1/N)\sum_{i\in[N]} f_i(x)$，它由 N 个智能体的局部损失函数构成；此处，$f_i(x)$ 是智能体 i 的局部损失函数。在联邦学习中，由于智能体的损失函数存在异质性，因此会产生“漂移”效应（Charles & Konečn`y, 2020; Karimireddy 等人, 2020b）：每个智能体 i 的局部迭代点会朝着 $f_i(x)$ 的最小化点漂移。然而，当异质性适中时，智能体的迭代序列的平均值会收敛到 $f(x)$ 的最小化点。为了针对联邦强化学习建立类似的理论，我们需要先回答：当我们对来自不同马尔可夫决策过程的 TD(0) 个更新方向取平均时，平均后的 TD(0) 个更新方向会将我们引向何处？我们引入虚拟马尔可夫决策过程这一概念，正是为了回答这个问题。为了对能够捕捉各智能体各自环境“平均值”的虚拟环境进行建模，我们构造一个马尔可夫决策过程 $\bar{\mathrm{M}} = \mathrm{S}, \mathrm{A}, \mathrm{R}, \mathrm{P}\ \gamma$，其中 $\bar{\mathrm{P}} = (1/N)\sum_{i=1}^{N}\mathrm{P}^{(i)}$，且 $\bar{\mathrm{R}} = (1/N)\sum_{i=1}^{N}\mathrm{R}^{(i)}$。请注意，虚拟马尔可夫决策过程是我们专为分析而构造的虚拟马尔可夫决策过程，通常可能与任何一个智能体的马尔可夫决策过程都不一致。

虚拟马尔可夫决策过程的性质。当应用于$\mathcal{M}$时，我们要评估的策略 μ会诱导出一个由元组 $\{\bar{P}、\bar{R}\}$刻画的虚拟马尔可夫奖励过程。显然，$\bar{P} = (1/N)\sum_{i=1}^{N} P^{(i)}$，并且 $\bar{R} = (1/N)\sum_{i=1}^{N} R^{(i)}$。下面的结果展示了虚拟马尔可夫奖励过程如何继承个体马尔可夫奖励过程的某些基本性质；该结果具有相当的普适性，也可能具有独立的研究价值。

命题 1。（马尔可夫矩阵的凸组合）设 $\{P^{(i)}\}_{i=1}^{N}$ 是与马尔可夫链相关联的一组马尔可夫矩阵，这些马尔可夫链具有相同的状态，并且均为非周期且不可约。于是，对于满足 $w_i \geq 0$、$\forall i \in [N]$ 和 $\sum_{i\in[N]} w_i = 1$ 的任意一组权重 $\{w_i\}_{i=1}^{N}$，与矩阵 $\sum_{i\in[N]} w_i P^{(i)}$ 对应的马尔可夫链同样是非周期且不可约的。

上述结果立即表明，与 P 相对应的马尔可夫链是非周期且不可约的。因此，该马尔可夫链存在唯一的平稳分布 $\bar{\pi}$；令 D 为相应的对角矩阵。与之前一样，定义 $\bar{A} \triangleq \Phi^\top \bar{D}(\Phi - \gamma\bar{P}\Phi)$，$\bar{b} \triangleq \Phi^\top \bar{D}\bar{R}$并用 θ^* 表示方程 $\bar{A}\theta^* = \bar{b}$ 的解。我们的下一个结果是定理1的推论，并刻画了对于每个θ_i^*而言，θ^*与 $i \in [N]$之间的差距。

命题 2。*（虚拟马尔可夫奖励过程与个体马尔可夫奖励过程“接近”）* 固定任意 $i \in [N]$。使用与定理 1 中相同的定义，我们有 $\|\bar{A}_i - \bar{A}\| \leq A(\epsilon)$，$\|\bar{b}_i - \bar{b}\| \leq b(\epsilon, \epsilon_1)$ 以及 $\|\theta_i^* - \theta^*\| \leq \Gamma(\epsilon, \epsilon_1)$。

我们将在后文论证，联邦时序差分算法（将在第 4 节介绍）会收敛到一个以虚拟马尔可夫奖励过程的 TD(0) 不动点 θ^* 为中心的球内。因此，命题 2 尤为重要，因为它告诉我们，在低异质性情形下，收敛到 θ^*附近，也就意味着收敛到每个智能体 θ_i^* 的最优参数 i附近。这说明了研究 FedTD(0) 在虚拟马尔可夫奖励过程上的收敛行为是合理的。定义 $\Sigma_v \triangleq \Phi^\top \bar{D}\Phi$。这个矩阵的最小特征值最终将决定所提出算法的收敛速率。本节最后给出一个结果，表明该特征值有一个严格大于零的下界。

命题 3。对于虚拟马尔可夫奖励过程，有 $\lambda_{\max}(\Sigma_v) \leq 1$，且存在 $\exists\ \bar{\omega} > 0$ 使得 $\lambda_{\min}(\Sigma_v) \geq \bar{\omega}$.

4 联邦 TD 算法

本节将介绍 FedTD(0) 算法（见算法 1）。FedTD(0) 的目标是生成一个模型 θ，使得 $\hat{} V_\theta$ 能够良好地近似每个智能体 i的价值函数 $V_\mu^{(i)}$，对应于

[4]模拟引理指出，如果两个具有相同状态空间和动作空间的马尔可夫决策过程相似，那么由它们上的共同策略导出的价值函数也相似。

算法 1 FedTD(0) 的描述

1: **输入：** 策略 μ、局部步长 α_l、全局步长 $\alpha_g^{(t)}$ 取决于通信轮次 t
2: **初始化：** $\bar{\theta}_0 = \theta_0$ 和 $s_{0,0}^{(i)} = s_0, \forall i \in [N]$
3: **对于**每轮 $t = 0, \ldots, T-1$ **执行**
4: **对于**每个智能体 $i \in [N]$ **执行**
5: **对于** $k = 0, \ldots, K-1$ **使用**初始模型 $\theta^{(i)} = \bar{\theta}_t$
6: 智能体 i执行 $\mu(s_{t,k}^{(i)})$，观察到 $O_{t,k}^{(i)} = (s_{t,k}^{(i)}, r_{t,k}^{(i)}, s_{t,k+1}^{(i)})$，并更新本地模型：
$\theta_{t,k+1}^{(i)} = \theta_{t,k}^{(i)} + \alpha_l g_i(\theta_{t,k}^{(i)})$其中 $g_i(\theta_{t,k}^{(i)}) \triangleq \left(r_{t,k}^{(i)} + \gamma\phi(s_{t,k+1}^{(i)})^\top \theta_t^{(i)} - \phi(s_{t,k}^{(i)})^\top \theta_t^{(i)}\right)\phi(s_{t,k}^{(i)})$
7: **结束循环**
8: 智能体 i 发送 $\Delta_t^{(i)} = \theta_{t,K}^{(i)} - \theta_t$ 返回服务器
9: **结束循环**
10: 服务器广播以下全局模型：$\bar{\theta}_{t+1} = \Pi_{2,\mathcal{H}}(\bar{\theta}_t + (\alpha_g^{(t)}/N) \sum_{i\in[N]} \Delta_t^{(i)})$
11: **结束循环**

策略 μ。与标准联邦学习算法以及在多智能体强化学习/联邦强化学习（同质设定）中使用的算法（Doan 等人，2019；Khodadadian 等人，2022）一致，智能体会对其原始观测（即奖励、状态和动作）保密，仅交换本地模型。在每一轮 t 中，每个智能体 $i \in [N]$ 从一个共同的全局模型 θ_t 开始，并使用其本地数据执行 K 次本地更新，形式如下：在每次本地迭代 k中，智能体 i 执行动作 $\mu(s_{t,k}^{(i)})$ 并观测到一个数据元组 $O_{t,k}^{(i)}$，该元组基于其自身的马尔可夫奖励过程（MRP），即 $\{P^{(i)}, R^{(i)}\}$；这里需要指出的是，各智能体的观测彼此独立。智能体 i 随后利用其数据元组，沿着第 6 行中的方向 $g_i(\theta_{t,k}^{(i)})$ 更新其本地模型 $\theta_{t,k}^{(i)}$。由于每个智能体都希望从其他智能体获取的样本中受益，因此会通过服务器进行间歇性通信。然而，这种通信需要受到限制，因为通信效率是联邦学习中的一个关键问题。因此，智能体仅每隔 K 个时间步才将其本地模型的差异 $\Delta_t^{(i)}$ 上传至服务器。服务器对这些模型差异取平均，并进行投影，以构建全局模型 θ_{t+1}，随后将其广播给所有智能体（第 10 行）。这里，我们用 $\Pi_{2,\mathcal{H}}(\cdot)$ 表示到凸紧子集 $\mathbf{H} \subset \mathbb{R}^d$ 上的标准欧几里得投影，并假设该子集包含每个 $\theta_i^*, i \in [N]$，以及 θ^*。此投影步骤可确保全局模型不会发散，这在随机逼近（Borkar，2009）和强化学习（Bhandari 等人，2018；Doan 等人，2019）中很常见。随后，每个智能体从该全局模型出发，继续进行本地更新。

我们注意到，FedTD(0) 的结构与 FedAvg （及其许多变体）相似：智能体使用各自的数据，在彼此隔离的情况下执行多次本地模型更新（以节省通信开销），并定期通过服务器进行同步。然而，标准联邦学习算法与 FedTD(0) 的动态特性存在显著差异，因此要为后者推导有限时间收敛结果颇具挑战。在下一节分析 FedTD(0) 时，我们将说明这些挑战的本质，并讨论如何克服它们。

5 主要结果与分析

为了陈述关于 FedTD(0) 的主要收敛结果，我们需要引入几个对象。首先，令 H 表示算法*1*第 10 行中集合H的半径。此外，定义 $G \triangleq R_{\max} + 2H$ 和 $\nu \triangleq (1-\gamma)\bar{\omega}$，其中 $\bar{\omega}$与命题3中的定义相同。在分析中，我们将利用有限状态、非周期且不可约马尔可夫链的几何混合性质（Levin & Peres，2017年）。具体而言，在假设3下，对每个$i \in [N]$，存在某个 $m_i \geq 1$和 $\rho_i \in (0, 1)$，使得对所有 $t \geq 0$和 $0 \leq k \leq K-1$:

$$d_{TV}\left(\mathbb{P}\left(s_{t,k}^{(i)} = \cdot \mid s_{0,0}^{(i)} = s\right), \pi^{(i)}\right) \leq m_i \rho_i^{tK+k}, \forall s \in \mathcal{S}.$$

这里，我们用 $d_{TV}(P,\ Q)$表示概率测度 P和Q.之间的全变差距离。对于任意 $\bar{\epsilon} > 0$，将 $P_{(i)}$的混合时间定义为 $\tau_i^{\text{mix}}(\bar{\epsilon})\ \triangleq \min\{t \in \mathbb{N}_0 \mid m_i \rho_i^t \leq \bar{\epsilon}\}$。最后，令$\tau(\bar{\epsilon})\ = \max_{i\in[N]} \tau_i^{\text{mix}}(\bar{\epsilon})$表示混合最慢的马尔可夫链对应的混合时间。正如所料，且如下文的主要结果所形式化的，正是这条混合最慢的马尔可夫链决定了 FedTD(0)收敛速率中的某些项。

定理 2. *(主要结果)* 存在一个递减的全局步长序列 $\{\alpha_g^{(t)}\}$、一个固定的本地步长 α_l，以及一组凸权重，使得全局模型 $\{\theta_t\}$ 的凸组合 θ_T 对每个智能体 $i \in [N]$ 在经过 T 轮次后满足以下条件：

$$\mathbb{E}\left[\left\|V_{\tilde{\theta}_T} - V_{\theta_i^*}\right\|_{\bar{D}}^2\right] \leq \tilde{\mathcal{O}}\left(\frac{\tau^2 G^2 + K^2}{K^2 T^2} + \frac{c_{quad}(\tau)}{\nu^2 NKT} + \frac{c_{lin}(\tau)}{\nu^4 KT^2} + Q(\epsilon, \epsilon_1)\right), \tag{2}$$

其中，$\tau = \lceil \tau^{\text{mix}}(\alpha_T^2) \rceil_K$，$\alpha_T = K\alpha_l \alpha_g^{(T)}$，而 $c_{quad}(\tau)$ 和 $c_{lin}(\tau)$ 分别是关于 τ 的二次函数和线性函数。此外，$B(\epsilon, \epsilon_1) = H\left(\sqrt{n}\epsilon + 2(n-1)\epsilon + \mathcal{O}(\epsilon^2) + O(\epsilon_1)\right)$，$\Gamma(\epsilon, \epsilon_1)$ 的定义与定理1中相同，而 $Q(\epsilon, \epsilon_1) = \tilde{O}(B(\epsilon,\epsilon_1)_\nu^G + \Gamma^2(\epsilon, \epsilon_1))$。

上述结果的证明见附录 I。现在我们讨论其含义。

讨论。为理解定理 2，我们先指出，方程 (2) 中的项 $Q(\epsilon, \epsilon_1)$ 刻画了异质性的影响；稍后我们将对此项进行讨论。当 $T \gg N$ 时，方程 (2) 中前三项里占主导地位的是 $c_{\text{quad}}(\tau)/(\nu^2 NKT)$。为说明这一项的紧确性，我们注意到，在集中式设置中（即当 $N = 1$ 时），给定 KT 个样本，TD(0) 的收敛速率为 $O(1/(\nu^2 KT))$（Bhandari 等人，2018）。因此，我们的分析表明，每个智能体只需在 KT 次迭代中进行 T 次通信，i 就能实现相对于智能体数量的线性加速。在低异质性情形下，即当 $Q(\epsilon, \epsilon_1)$ 较小时，我们注意到，通过结合来自不同马尔可夫决策过程的数据，FedTD(0) 能够保证快速收敛到一个对每个智能体的价值函数都具有良好近似的模型；这里所说的快速，是指相较于完全不进行通信时每个智能体所能达到的收敛速率，实现 N 倍加速。因此，只需少量通信，FedTD(0) 就能快速为每个智能体提供一个良好的模型，随后智能体可对其进行微调，以实现个性化。定理 2 意义重大，因为它是多智能体强化学习/联邦强化学习在异质环境下的首个此类结果，并补充了异构联邦优化领域中大量类似的结果（Sahu 等人，2018；Khaled 等人，2019；2020；Li 等人，2019；Koloskova 等人，2020；Woodworth 等人，2020b；Malinovskiy 等人，2020；Pathak 与 Wainwright，2020；Wang 等人，2020；Karimireddy 等人，2020b；Acar 等人，2021；Gorbunov 等人，2021；Mitra 等人，2021；Mishchenko 等人，2022）。

当所有马尔可夫决策过程都相同时，$Q(\epsilon, \epsilon_1) = 0$。但当马尔可夫决策过程各不相同时，我们是否还应期待出现这样的项？为了进一步理解异质性的影响，只需消除我们设定中的所有随机性。因此，假设我们将算法 1 中每个智能体 i 的随机 TD(0) 方向 $g_i(\theta_{t,k}^{(i)},)$替换为其稳态确定性版本 $\bar{g}_i(\theta_{t,k}^{(i)},) = \bar{b}_i - \bar{A}_i \theta_{t,k}^{(i)}, ,$ 其中 $\bar{A}_i$ 和 $\bar{b}_i$ 与第 3 节中的定义相同。我们将由此得到的确定性算法称为均值路径 FedTD(0)。为简洁起见，我们省略投影步骤。在下一个结果中，我们利用稳态 TD(0)方向的仿射性质，刻画异质性对 FedTD(0)极限行为的影响。

定理3。 *(异质性偏差)* 假设 $N = 2$ 和 $K = 1$。令步长 $\alpha = \alpha_l \alpha_g^{(t)}$满足以下条件：$I - \alpha\hat{A}$ 是*Schur*稳定的，其中 $\hat{A} = \left(\bar{A}_1 + \bar{A}_2\right)/2$。定义$e_{i,t} \triangleq \bar{\theta}_t - \theta_i^*, i \in \{1,2\}$。那么，均值路径 *FedTD(0)*的输出满足：

$$\lim_{t\to\infty} e_{1,t} = \frac{1}{2}\hat{A}^{-1}\bar{A}_2(\theta_1^* - \theta_2^*); \lim_{t\to\infty} e_{2,t} = \frac{1}{2}\hat{A}^{-1}\bar{A}_1(\theta_2^* - \theta_1^*). \tag{3}$$

讨论： 对于定理3中描述的情形，均值路径 FedTD(0) 的更新遵循确定性递推 $\bar{\theta}_{t+1} = (I - \alpha\hat{A})\bar{\theta}_t + \alpha\hat{b}$，其中 $\hat{b} = (1/2)(\bar{b}_1 + \bar{b}_2)$。这是一个离散时间线性时不变系统（LTI）。当且仅当状态转移矩阵 $(I - \alpha\hat{A})$ 是 Schur 稳定时，该系统的动力学才稳定，这证明了在定理 α 中选择 3的合理性。这一结果传达的主要信息是，均值路径 FedTD(0) 的极限点与 θ_i^* 以下两者中任一者的最优参数之间的差距

MRP 取决于各 MRP 的最优参数之间的差异，而这自然是两个 MRP 异质性的指标。此外，该项不依赖于步长 α，即无论将 α 缩小到多小，都无法消除异质性导致的偏差。与此观察一致，注意方程 (2) 中的 $Q(\epsilon, \epsilon_1)$ 也与步长无关。上述讨论有助于理解为何定理 2 中会出现 $Q(\epsilon, \epsilon_1)$ 形式的项。值得注意的是，即使本地步数仅为 1，也就是智能体在每个时间步都与服务器通信，方程 (3) 中的偏差项仍然存在。这与标准联邦学习设置有一个关键区别：在标准设置中，只有当本地步数 K 严格大于 1 时，异质性的影响才会显现（Charles & Konečnỳ, 2021年；Karimireddy 等人，2020b；Mitra 等人，2021年）。

5.1 主要技术挑战与分析中新颖要素概述

挑战。我们总结定理 2 分析中遇到的主要技术挑战。首先， FedTD(0) 更新方向可能并不对应于任何马尔可夫决策过程的 TD(0) 更新方向。这一挑战是我们所研究设定独有的，在集中式 TD(0) 分析中（Bhandari 等人，2018；Srikant 与 Ying，2019）以及针对同质马尔可夫决策过程的现有多智能体强化学习/联邦强化学习分析中（Doan 等人，2019；Khodadadian 等人，2022）均未出现。其次，与处理每个智能体独立同分布观测的标准联邦学习分析不同，我们的设定更为复杂，因为每个智能体的数据均由马尔可夫链生成。此外，对于每个智能体 i，参数序列 $\{\theta_{t\,k}^{(i)},\}$ 与数据元组 $\{O_{t\,k}^{(i)},\}$ 之间存在错综复杂的耦合关系。第三， FedTD(0) 中的同步步骤会在任一给定智能体的本地参数与所有其他智能体的过往观测之间产生复杂的统计依赖关系。第四，在寻求建立关于智能体数量 N 的线性加速性质（即 $O(1/NKT)$ 类型的速率）时，控制梯度偏差（$1/NK$）$\sum_{i=1}^{N}\sum_{k=0}^{K-1}\left(g_i(\theta_{t\,k}^{(i)},,O_{t\,k}^{(i)},)-\bar{g}_i(\theta_{t\,k}^{(i)},)\right)$ 和梯度范数 $\mathbb{E}\|(1/NK)\sum_{i=1}^{N}\sum_{k=0}^{K-1}g_i(\theta_{t\,k}^{(i)},)\|^2$需要非常精细的分析。特别是，像集中式分析中那样（Bhandari 等人，2018）直接利用投影半径界定各项，并不能得到线性加速性质。最后，在上述不同随机变量强耦合的情况下，我们还需要控制环境异质性所导致的“客户端漂移”效应。

定理 2 的证明概要。我们的首个关键创新是在第 3 节的结果基础上，证明平均路径（稳态） FedTD(0) 更新方向$(1/N)\sum_{i=1}^{N}\bar{g}_i(\theta)$ 与我们在第 TD(0) 节构造的虚拟马尔可夫奖励过程的平均路径更新方向 $\bar{g}(\theta)=b-A\theta$ “接近”。这里，3.1 中定义了 b, A。形式化地，我们有以下结果。3.1

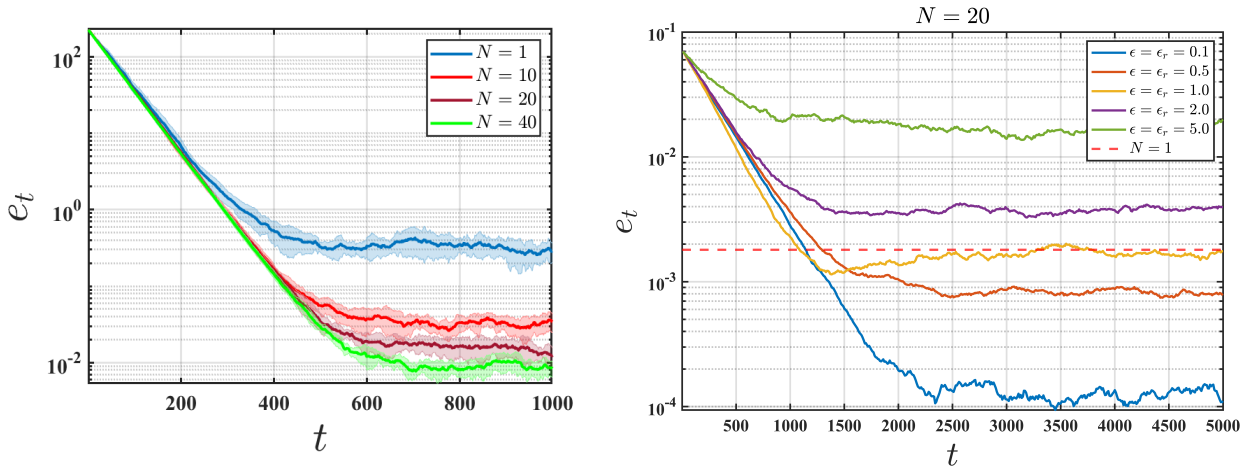
引理2。 *(稳态伪梯度异质性)* 对于每个 $\theta\in\mathcal{H}$，有：

$$\left\|\bar{g}(\theta)-\frac{1}{N}\sum_{i=1}^{N}\bar{g}_i(\theta)\right\|\leq B(\epsilon,\epsilon_1), \tag{4}$$

其中 $B(\epsilon,\epsilon_1)$ 与定理 2 中相同，且 $\bar{g}(\theta)$ 是虚拟马尔可夫奖励过程的稳态 *TD(0)* 方向 .

根据 Bhandari 等人（2018），我们知道 $\bar{g}(\theta)$ 类似于一个伪梯度，指向虚拟马尔可夫奖励过程的最优模型 θ^*。根据命题 2，我们知道 θ^* 接近于 $\theta_i^*, \forall i\in[N]$；引理2告诉我们，至少在稳态下， FedTD(0) 的迭代序列将收敛到每个智能体最优模型的一个邻域内，该邻域的大小取决于异质性程度。虽然这有助于建立直觉，但引理2所传达的所有有价值见解仅涉及FedTD(0)的稳态动力学，也就是说，我们提到的所有统计挑战仍有待解决。具体而言，如前所述，我们不能直接使用来自 Bhandari 等人（2018）集中式分析的形如 $\mathbb{E}\left[\|(1/NK)\sum_{i=1}^{N}\sum_{k=0}^{K-1}g_i(\theta_{t\,k}^{(i)},)\|^2\right]=O(G^2)$ 的投影界，因为本地模型可能不属于集合 H。此外，这样做会掩盖线性加速效应。我们通过将每个智能体 i 的随机 TD 方向分解为 $g_i(\theta_{t\,k}^{(i)},)=b_i(O_{t\,k}^{(i)},)-A_i(O_{t\,k}^{(i)},)\theta_{t\,k}^{(i)},.$ 来克服这一困难。由于 $A_i(O_{t\,k}^{(i)},)$ 和 $b_i(O_{t\,k}^{(i)},)$ 仅取决于马尔可夫链的随机性，且 $O_{t\,k}^{(i)}$ 和 $O_{t\,k}^{(j)}$ 相互独立，我们可以证明 $(1/NK)\sum_{i=1}^{N}\sum_{k=0}^{K-1}A_i(O_{t\,k}^{(i)},)$ 和 $(1/NK)\sum_{i=1}^{N}\sum_{k=0}^{K-1}b_i(O_{t\,k}^{(i)},)$ 的方差会按 NK缩小（忽略高阶项）。此外，为了考虑到 $A_i(O_{t\,k}^{(i)},)$ 和 $b_i(O_{t\,k}^{(i)},)$ 在不同智能体间存在差异，我们采用

引理 2。仔细整合这些部分，便可得到定理2中的最终速率。详细分析以及一些模拟结果见附录。



(a) 线性加速效果的模拟　　(b) 异质性水平影响的模拟

图2：马尔可夫采样下 FedTD(0) 的性能。(a) 不同智能体数量 N 下 FedTD(0) 的性能。第一个智能体的马尔可夫决策过程 $\mathcal{M}^{(1)}$ 是随机生成的，其状态空间大小为 $n = 100$。其余马尔可夫决策过程是对 $\mathcal{M}^{(1)}$ 的扰动，异质性水平分别为 $\epsilon = 0.05$ 和 $\epsilon_1 = 0.1$。我们使用运行误差 $e_t = \|\bar{\theta}_t - \theta_1^*\|^2$ 来评估收敛情况。(b) 固定智能体数量为 $N = 20$ 时，不同异质性水平下 FedTD(0) 的性能。与理论结果一致，增大 N 会降低误差，而增大异质性水平会增大 FedTD(0) 收敛到的球的半径。两幅图中的本地步数均设为 $K = 10$ 。

6 结论

在这项工作中，我们研究了环境异质性下的联邦强化学习问题，并探讨了以下问题：智能体能否利用与可能不同的马尔可夫决策过程交互的其他智能体所提供的信息，加快学习自身价值函数的过程？为回答这个问题，我们研究了采用线性函数逼近的联邦 TD(0) 算法的收敛性；在不同环境下的 N 个智能体协作评估同一个策略。与现有工作的主要区别在于：(i) 提出了环境异质性的新定义；(ii) 刻画了异质性对 TD(0) 不动点的影响；(iii) 引入虚拟马尔可夫决策过程来分析 FedTD(0) 算法的长期行为；(iv) 利用虚拟马尔可夫决策过程，明确建立了联邦强化学习与联邦监督学习/优化之间的联系。基于这些要素，我们证明，若智能体所处环境之间的环境异质性较小，那么在独立同分布和马尔可夫设置下，以及进行多次本地更新时， FedTD(0) 均可实现线性加速。

这项工作有一些有趣的扩展方向。首先，研究除 TD(0) 算法之外其他强化学习算法的联邦变体是很自然的。其次，探究传统联邦学习优化文献中使用的个性化技术能否用于解决联邦强化学习问题，也很有意义。我们能否不再学习一个通用的价值函数/策略，而是设计在高异质性情形下可能表现更好的个性化价值函数/策略？我们将这一有趣问题的探索留作未来工作。

致谢

Anderson 和 Wang 获得美国国家科学基金会通过 EPCN 项目提供的部分资助，资助编号为 2144634 和 2231350。

Hamed Hassani 获得规模化学习赋能优化研究所（TILOS）的资助，资助编号为 NSF-CCF-2112665。

参考文献

Durmus Alp Emre Acar、Yue Zhao、Ramon Matas Navarro、Matthew Mattina、Paul N Whatmough 和 Venkatesh Saligrama。基于动态正则化的联邦学习。*arXiv* 预印本 *arXiv:2111.04263*，2021年。

Jalaj Bhandari、Daniel Russo 和 Raghav Singal。采用线性函数逼近的时序差分学习的有限时间分析。载于学习理论会议，第 1691–1692 页。PMLR，2018年。

Vivek S Borkar。随机逼近：动力系统视角，第 48 卷。施普林格，2009年。

Zachary Charles 和 Jakub Konečn`y。本地更新方法中学习率的重要性不容小觑。*arXiv* 预印本 *arXiv:2007.00878*，2020年。

Zachary Charles 和 Jakub Konečn`y。《联邦学习与元学习中的收敛与准确率权衡》。载于国际人工智能与统计会议，第 2575–2583 页。PMLR，2021年。

Chenyi Chen、Ari Seff、Alain Kornhauser 和 Jianxiong Xiao。《DeepDriving：学习自动驾驶中直接感知的可供性》。载于*IEEE* 国际计算机视觉会议论文集，第 2722–2730 页，2015年。

Zaiwei Chen、Sheng Zhang、Thinh T Doan、Siva Theja Maguluri 和 John-Paul Clarke。《线性函数逼近下 Q学习的性能：稳定性与有限时间分析》。*arXiv* 预印本 *arXiv:1905.11425*，第 4 页，2019年。

Liam Collins、Hamed Hassani、Aryan Mokhtari 和 Sanjay Shakkottai。《利用共享表示实现个性化联邦学习》。载于国际机器学习会议，第 2089–2099 页。PMLR，2021年。

Liam Collins、Hamed Hassani、Aryan Mokhtari 和 Sanjay Shakkottai。FedAvg 与微调：本地更新促成表征学习。*arXiv* 预印本 *arXiv:2205.13692*，2022年。

Nicolò Dal Fabbro、Aritra Mitra 和 George J Pappas。有限速率擦除信道上的联邦时序差分学习：马尔可夫采样下的线性加速。*IEEE*控制系统快报，2023年。

Gal Dalal、Balázs Szörényi、Gugan Thoppe 和 Shie Mannor。函数逼近下 TD(0) 的有限样本分析。载于 *AAAI*人工智能会议论文集，第 32 卷，2018年。Yuyang Deng、Mohammad Mahdi Kamani 和 Mehrdad Mahdavi。自适应个性化联邦学习。*arXiv* 预印本 *arXiv:2003.13461*，2020年。Thinh Doan、Siva Maguluri 和 Justin Romberg。多智能体强化学习中线性函数逼近下分布式 TD(0) 的有限时间分析。载于 国际机器学习会议，第 1626–1635 页。PMLR，2019年。Alireza Fallah、Aryan Mokhtari 和 Asuman Ozdaglar。个性化联邦学习：一种元学习方法。*arXiv* 预印本 *arXiv:2002.07948*，2020年。Georg Frobenius、Ferdinand Georg Frobenius、Ferdinand Georg Frobenius、Ferdinand Georg Frobenius 和德国数学家。关于由非负元素构成的矩阵。1912年。Avishek Ghosh、Jichan Chung、Dong Yin 和 Kannan Ramchandran。一种高效的聚类联邦学习框架。神经信息处理系统进展，33:19586–19597，2020年。Eduard Gorbunov、Filip Hanzely 和 Peter Richtárik。本地随机梯度下降：统一理论与高效新方法。载于 国际人工智能与统计会议，第 3556–3564 页。PMLR，2021年。

Taosha Guo、Abed AlRahman Al Makdah、Vishaal Krishnan 和 Fabio Pasqualetti。线性二次高斯控制的模仿与迁移学习。*arXiv* 预印本 *arXiv:2303.09002*，2023年。Farzin Haddadpour 和 Mehrdad Mahdavi。联邦学习中局部下降方法的收敛性。*arXiv* 预印本 *arXiv:1910.14425*，2019年。Farzin Haddadpour、Mohammad Mahdi Kamani、Mehrdad Mahdavi 和 Viveck Cadambe。周期平均的本地随机梯度下降：更紧致的分析与自适应同步。载于神经信息处理系统进展，第 11082–11094 页，2019年。Filip Hanzely、Slavomír Hanzely、Samuel Horváth 和 Peter Richtárik。个性化联邦学习的下界与最优算法。神经信息处理系统进展，33:2304–2315，2020年。Roger A Horn 和 Charles R Johnson。矩阵分析。剑桥大学出版社，2012a。Roger A Horn 和 Charles R Johnson。矩阵分析。剑桥大学出版社，2012b。Hao Jin、Yang Peng、Wenhao Yang、Shusen Wang 和 Zhihua Zhang。环境异质性下的联邦强化学习。载于国际人工智能与统计会议，第 18–37 页。PMLR，2022年。Sai Praneeth Karimireddy、Satyen Kale、Mehryar Mohri、Sashank Reddi、Sebastian Stich 和 Ananda Theertha Suresh。Scaffold：联邦学习的随机控制平均。载于 国际机器学习会议，第 5132–5143 页。PMLR，2020a。Sai Praneeth Karimireddy、Satyen Kale、Mehryar Mohri、Sashank Reddi、Sebastian Stich 和 Ananda Theertha Suresh。Scaffold：联邦学习的随机控制平均。载于 国际机器学习会议，第 5132–5143 页。PMLR，2020b。Michael Kearns 和 Satinder Singh。多项式时间内近乎最优的强化学习。机器学习，49(2):209–232，2002。Ahmed Khaled、Konstantin Mishchenko 和 Peter Richtárik。异质数据上的局部梯度下降的首次分析。*arXiv* 预印本 *arXiv:1909.04715*，2019年。Ahmed Khaled、Konstantin Mishchenko 和 Peter Richtárik。相同及异质数据上本地随机梯度下降的更紧致理论。载于国际人工智能与统计会议，第 4519–4529 页。PMLR，2020年。Koulik Khamaru、Ashwin Pananjady、Feng Ruan、Martin J Wainwright 和 Michael I Jordan。时序差分学习是否最优？一种实例依赖分析。*SIAM*数据科学数学杂志，3(4):1013–1040，2021年。Sajad Khodadadian、Pranay Sharma、Gauri Joshi 和 Siva Theja Maguluri。联邦强化学习：马尔可夫采样下的线性加速。载于 国际机器学习会议，第 10997–11057 页。PMLR，2022年。Anastasia Koloskova、Nicolas Loizou、Sadra Boreiri、Martin Jaggi 和 Sebastian U Stich。变化拓扑与本地更新下去中心化随机梯度下降的统一理论。*arXiv* 预印本 *arXiv:2003.10422*，2020年。Jakub Konečn`y、H Brendan McMahan、Daniel Ramage 和 Peter Richtárik。联邦优化：面向端侧智能的分布式机器学习。*arXiv* 预印本 *arXiv:1610.02527*，2016年。Nathaniel Korda 和 Prashanth La。函数逼近下的 TD(0)：集中界与一种具有指数收敛性的中心化变体。载于国际机器学习会议，第 626–634 页。PMLR，2015年。Yassine Laguel、Krishna Pillutla、Jerôme Malick 和 Zaid Harchaoui。面向异构设备联邦学习的超分位数方法。载于 *2021*年第*55*届年度信息科学与系统会议（*CISS*），第 1–6 页。IEEE，2021年。

Chandrashekar Lakshminarayanan 和 Csaba Szepesvári。线性随机近似：常数步长与迭代平均。*arXiv*预印本 *arXiv:1709.04073*，2017年。David A Levin 和 Yuval Peres。马尔可夫链与混合时间，第107卷。美国数学学会，2017年。Xiang Li、Kaixuan Huang、Wenhao Yang、Shusen Wang 和 Zhihua Zhang。关于非独立同分布数据上 FedAvg 的收敛性。*arXiv*预印本 *arXiv:1907.02189*，2019年。Boyi Liu、Lujia Wang 和 Ming Liu。终身联邦强化学习：一种用于云机器人系统导航的学习架构。*IEEE*机器人与自动化快报，4(4):4555–4562，2019年。Rui Liu 和 Alex Olshevsky。几乎无需通信的分布式 TD(0)。*arXiv*预印本 *arXiv:2104.07855*，2021a。Rui Liu 和 Alex Olshevsky。作为梯度分裂的时序差分学习。载于国际机器学习会议，第6905–6913页。PMLR，2021b。Grigory Malinovskiy、Dmitry Kovalev、Elnur Gasanov、Laurent Condat 和 Peter Richtarik。从本地随机梯度下降到联邦学习的本地不动点方法。载于国际机器学习会议，第6692–6701页。PMLR，2020年。Brendan McMahan、Eider Moore、Daniel Ramage、Seth Hampson 和 Blaise Aguera y Arcas。基于去中心化数据的深度网络通信高效学习。载于人工智能与统计学，第1273–1282页。PMLR，2017年。Konstantin Mishchenko、Grigory Malinovsky、Sebastian Stich 和 Peter Richtárik。ProxSkip：是的！本地梯度步可证明能加速通信！终于实现了！*arXiv*预印本 *arXiv:2202.09357*，2022年。Aritra Mitra。线性函数逼近下时序差分学习的简明有限时间分析。*arXiv*预印本 *arXiv:2403.02476*，2024年。Aritra Mitra、Rayana Jaafar、George J Pappas 和 Hamed Hassani。联邦学习中的线性收敛：应对客户端异质性与稀疏梯度。神经信息处理系统进展，34：14606–14619，2021年。C Narayanan 和 Csaba Szepesvári。函数逼近下时序差分学习的有限时间界：某些“最新”结果存在的问题。技术报告，技术报告，2017年。Colm Art O’cinneide。马尔可夫链的逐项扰动理论与误差分析。数值数学，65(1):109–120，1993年。Reese Pathak 和 Martin J Wainwright。FedSplit：一种用于快速联邦优化的算法框架。*arXiv*预印本 *arXiv:2005.05238*，2020年。Hossein Pishro-Nik。概率、统计学与随机过程导论。2016年。Jiaju Qi、Qihao Zhou、Lei Lei 和 Kan Zheng。联邦强化学习：技术、应用与开放挑战。*arXiv*预印本 *arXiv:2108.11887*，2021年。Amirhossein Reisizadeh、Aryan Mokhtari、Hamed Hassani、Ali Jadbabaie 和 Ramtin Pedarsani。Fed-paq：一种采用周期平均与量化的通信高效联邦学习方法。载于国际人工智能与统计会议，第2021–2031页。PMLR，2020年。Anit Kumar Sahu、Tian Li、Maziar Sanjabi、Manzil Zaheer、Ameet Talwalkar 和 Virginia Smith。异构网络中联邦优化的收敛性。*arXiv*预印本 *arXiv:1812.06127*，3，2018年。

Felix Sattler、Klaus-Robert Müller 和 Wojciech Samek。聚类联邦学习：隐私约束下与模型无关的分布式多任务优化。*IEEE*神经网络与学习系统汇刊，32(8):3710–3722，2020年。Han Shen、Kaiqing Zhang、Mingyi Hong 和 Tianyi Chen。迈向对异步优势演员-评论家的理解：收敛性与线性加速。*IEEE*信号处理汇刊，2023年。Artin Spiridonoff、Alex Olshevsky 和 Ioannis Ch Paschalidis。本地随机梯度下降的通信开销仅取决于工作节点数量。*arXiv*预印本 *arXiv:2006.02582*，2020年。Rayadurgam Srikant 和 Lei Ying。线性随机近似与时序差分学习的有限时间误差界。载于学习理论会议，第2803–2830页。PMLR，2019年。Sebastian U Stich。本地随机梯度下降收敛迅速且通信开销低。*arXiv*预印本 *arXiv:1805.09767*，2018年。Lili Su、Jiaming Xu 和 Pengkun Yang。混合回归中联邦学习的全局收敛性。*arXiv*预印本 *arXiv:2206.07279*，2022年。Richard S Sutton。通过时序差分方法学习预测。机器学习，3(1):9–44，1988年。Richard S Sutton、Andrew G Barto 等。强化学习导论。1998年。Canh T Dinh、Nguyen Tran 和 Josh Nguyen。基于 Moreau 包络的个性化联邦学习。神经信息处理系统进展，33:21394–21405，2020年。Alysa Ziying Tan、Han Yu、Lizhen Cui 和 Qiang Yang。迈向个性化联邦学习。*IEEE*神经网络与学习系统汇刊，2022年。John N Tsitsiklis 和 Benjamin Van Roy。函数逼近下时序差分学习的分析。载于*IEEE*自动控制汇刊，1997年。Han Wang、Siddartha Marella 和 James Anderson。FedADMM：一种允许部分参与的联邦原始-对偶算法。载于*2022*年 *IEEE* 第*61*届决策与控制会议 *(CDC)*，第287–294页。IEEE，2022a。Han Wang、Leonardo F Toso 和 James Anderson。FedSysID：一种样本高效系统辨识的联邦方法。*arXiv*预印本 *arXiv:2211.14393*，2022b。Han Wang、Leonardo F Toso、Aritra Mitra 和 James Anderson。异构动力系统的无模型学习：一种联邦线性二次调节方法。*arXiv*预印本 *arXiv:2308.11743*，2023年。Jianyu Wang 和 Gauri Joshi。协同随机梯度下降：通信高效随机梯度下降算法设计与分析的统一框架。*arXiv*预印本 *arXiv:1808.07576*，2018年。Jianyu Wang、Qinghua Liu、Hao Liang、Gauri Joshi 和 H Vincent Poor。应对异构联邦优化中的目标不一致问题。神经信息处理系统进展，33，2020年。Xiaofei Wang、Yiwen Han、Chenyang Wang、Qiyang Zhao、Xu Chen 和 Min Chen。边缘内人工智能：通过联邦学习实现移动边缘计算、缓存与通信的智能化。*IEEE*网络，33(5):156–165，2019年。Blake Woodworth、Kumar Kshitij Patel、Sebastian U Stich、Zhen Dai、Brian Bullins、H Brendan McMahan、Ohad Shamir 和 Nathan Srebro。本地随机梯度下降优于小批量随机梯度下降吗？*arXiv*预印本 *arXiv:2002.07839*，2020a。Blake E Woodworth、Kumar Kshitij Patel 和 Nati Srebro。异构分布式学习中的小批量随机梯度下降与本地随机梯度下降。神经信息处理系统进展，33:6281–6292，2020b。

Zhijie Xie 和 Shenghui Song。Fedkl：通过惩罚 KL 散度应对联邦强化学习中的数据异质性。*IEEE*通信领域精选期刊，41(4):1227–1242，2023年。

Lei Xin、Lintao Ye、George Chiu 和 Shreyas Sundaram。利用相似系统的数据学习动力系统。*arXiv*预印本 *arXiv:2302.04344*，2023年。

Chenyu Zhang、Han Wang、Aritra Mitra 和 James Anderson。同策略异质联邦强化学习的有限时间分析。国际学习表征会议，2024年。

目录

1 引言 1

1.1 我们的贡献 .3 1.2 相关工作 .4 **2 模型与问题表述** 5 **3 异构联邦强化学习** 6 3.1 虚拟马尔可夫决策过程 .8 **4 联邦时序差分算法** 8 **5 主要结果与分析** 9 5.1 主要技术挑战及分析中新颖要素概述 . . . . . 11 **6 结论** 12 **A 补充文献综述** 20 **B 扰动界 TD(0) 不动点** 21 B.1 定理1的证明.21 **C 虚拟马尔可夫决策过程的性质** 22

C.1 命题 1 的证明 .22 C.2 命题 2 的证明 .22 C.3 命题 3 的证明 .22

D 伪梯度异质性：引理2的证明 23

E 独立同分布和马尔可夫设置中使用的辅助结果 23

F 符号 26

G 预热：对 分析 FedTD 在独立同分布采样下 27

G.1 定理4的辅助引理.27 G.1.1 方差缩减.27 G.1.2 每轮进展.28 G.1.3 漂移项分析.30 G.1.4 参数选择.32

G.2 定理4的证明.33

H 异质性偏差：定理3的证明 35

I 马尔可夫设置的证明 36

I.1 概述 .36I.2 定理 2 的辅助引理 .36I.2.1 分解形式 .36I.2.2 方差缩减.38I.2.3 界定 $\mathbb{E}\left[\left\|\bar{\theta}_t - \bar{\theta}_{t-\tau}\right\|^2\right]$. 40I.2.4 漂移项分析. .46I.2.5 每轮进展 .48I.2.6 参数选择 .56I.3 定理 2 的证明. .58

J 模拟结果 60

J.1 独立同分布设定下的模拟结果 .60J.2 马尔可夫设置下的模拟结果 .61J.3 马尔可夫设置下异质性水平影响的模拟。 62

A 补充文献综述

联邦学习算法。关于联邦学习算法发展的文献浩如烟海；因此，我们在此仅介绍其中一些最相关、最具代表性的研究。最常用的联邦学习算法 FedAvg 最初由 McMahan 等人（2017年）提出。随后，多项研究对 FedAvg 进行了详细的理论分析，既包括所有客户端都最小化同一目标函数的同质情形（Stich，2018年；Wang 与 Joshi，2018年；Spiridonoff 等人，2020年；Reisizadeh 等人，2020年；Haddadpour 等人，2019年；Woodworth 等人，2020a），也包括更具挑战性的异质情形（Khaled 等人，2019年；2020年；Haddadpour 与 Mahdavi，2019年；Li 等人，2019年；Koloskova 等人，2020年）。在后一种情形下，人们很快发现，FedAvg 会受到“客户端漂移”效应的影响，从而损害其收敛性能（Charles 与 Konečn`y，2020年；2021年）。

此后，人们投入了大量精力，通过多种技术途径改善 FedAvg 的收敛保证：FedProx 中的近端方法（Sahu 等人，2018年）；FedSplit中的算子分裂（Pathak 与 Wainwright，2020年）；Scaffold 中的方差缩减（Karimireddy 等人，2020a）和 S-Local-SVRG中的方差缩减（Gorbunov 等人，2021年）；FedLin 中的梯度跟踪（Mitra 等人，2021年）；Acar 等人（2021年）提出的动态正则化；以及 FedADMM 中的交替方向乘子法（Wang 等人，2022a）。尽管这些方法在多个方面改进了 FedAvg，但它们都未能为在任意统计异质性下执行多次本地更新提供任何理论依据。就在最近，Mishchenko 等人（2022年）提出了ProxSkip 算法，并证明即使存在任意异质性，该算法也确实可以通过本地步数节省通信开销。

解决联邦学习中统计分布异质性的其他方法包括个性化（Deng 等人，2020年；Fallah 等人，2020年；T Dinh 等人，2020年；Hanzely 等人，2020年；Tan 等人，2022年）、聚类（Ghosh 等人，2020年；Sattler 等人，2020年；Su 等人，2022年）、表征学习（Collins 等人，2021年）以及使用分位数（Laguel 等人，2021年）。

B TD(0) 不动点的扰动界

B.1 定理 1 的证明

在本节中，我们将证明定理 1中给出的 TD(0)固定点的扰动界。首先，我们注意到：

$$\begin{aligned}
\|\bar{A}_i - \bar{A}_j\| &= \|\Phi^\top D^{(i)}(\Phi - \gamma P^{(i)}\Phi) - \Phi^\top D^{(j)}(\Phi - \gamma P^{(j)}\Phi)\| \\
&\le \|\Phi^\top D^{(i)}(\Phi - \gamma P^{(i)}\Phi) - \Phi^\top D^{(i)}(\Phi - \gamma P^{(j)}\Phi) + \\
&\qquad \Phi^\top D^{(i)}(\Phi - \gamma P^{(j)}\Phi) - \Phi^\top D^{(j)}(\Phi - \gamma P^{(j)}\Phi)\| \\
&\le \|\Phi^\top D^{(i)}(\Phi - \gamma P^{(i)}\Phi) - \Phi^\top D^{(i)}(\Phi - \gamma P^{(j)}\Phi)\| + \\
&\qquad \|\Phi^\top D^{(i)}(\Phi - \gamma P^{(j)}\Phi) - \Phi^\top D^{(j)}(\Phi - \gamma P^{(j)}\Phi)\| \\
&\overset{(a)}{\le} \gamma\|\Phi\|^2\|D^{(i)}\|\|P^{(i)} - P^{(j)}\| + \|\Phi\|^2\|D^{(i)} - D^{(j)}\|\|(I - \gamma P^{(j)})\| \\
&\overset{(b)}{\le} \gamma\sqrt{n}\epsilon + (1+\gamma)[2(n-1)\epsilon + \mathcal{O}(\epsilon^2)],
\end{aligned} \tag{5}$$

其中，(a) 根据三角不等式得出。(b) 中的第一项利用了这样一个事实： $\|\Phi\| \le 1$、 $\|D^{(i)}\| \le 1$，以及

$$\|P^{(i)} - P^{(j)}\| \le \sqrt{n}\|P^{(i)} - P^{(j)}\|_\infty \le \epsilon\sqrt{n}\|P^{(i)}\|_\infty = \epsilon\sqrt{n},$$

其中，在第二个不等式中我们使用了假设 1 。(b) 中的第二项利用了以下事实：$\|I - \gamma P^{(j)}\| \le 1 + \gamma$、$\|D^{(i)} - D^{(j)}\| \le \|D^{(i)} - D^{(j)}\|_1 \le \|\pi^{(i)} - \pi^{(j)}\|_1$，以及引理 1。

接下来，我们对其进行界定

$$\begin{aligned}
\|\bar{b}_i - \bar{b}_j\| &= \|\Phi D^{(i)}R^{(i)} - \Phi D^{(j)}R^{(j)}\| \\
&\le \|\Phi D^{(i)}R^{(i)} - \Phi D^{(i)}R^{(j)}\| + \|\Phi D^{(i)}R^{(j)} - \Phi D^{(j)}R^{(j)}\| \\
&\le \|\Phi\|\|D^{(i)}\|\|R^{(i)} - R^{(j)}\| + \|\Phi\|\|D^{(i)} - D^{(j)}\|\|R^{(j)}\| \\
&\le \epsilon_1 + R_{\max}\left(2(n-1)\epsilon + \mathcal{O}(\epsilon^2)\right),
\end{aligned} \tag{6}$$

其中，我们在最后一个不等式中使用了假设2，并采用与上文推导 $\|A_i - A_j\|$相同的方法。

现在，我们可以推导不动点之间的差距：

$$\frac{\|\theta_i^* - \theta_j^*\|}{\|\theta_i^*\|} \le \frac{\kappa(\bar{A}_i)}{1 - \kappa(\bar{A}_i)\frac{\|\bar{A}_i - \bar{A}_j\|}{\|\bar{A}_i\|}}\left(\frac{\|\bar{A}_i - \bar{A}_j\|}{\|\bar{A}_i\|} + \frac{\|\bar{b}_i - \bar{b}_j\|}{\|\bar{b}_i\|}\right). \tag{7}$$

这里，我们利用了（Horn & Johnson，2012b）第5节5.8中的线性方程扰动理论。最后，对于任意 $\|\theta_i^*\| \le H$，我们有

$$\|\theta_i^* - \theta_j^*\| \le \Gamma(\epsilon, \epsilon_1) \triangleq \frac{\kappa(\bar{A}_i)H}{1 - \kappa(\bar{A}_i)\frac{A(\epsilon)}{\delta_1}}\left(\frac{A(\epsilon)}{\delta_1} + \frac{b(\epsilon, \epsilon_1)}{\delta_2}\right),$$

其中我们使用了 fact，其中 δ_1 和 δ_2 是分别为 $\|A_i\|$ 和 $\|b_i\|$提供下界的正常数。

C 虚拟马尔可夫决策过程的性质

C.1 命题 1 的证明

在证明该命题之前，我们先给出以下来自（Pishro-Nik，2016年）的事实：马尔可夫矩阵 P不可约且非周期，当且仅当存在一个正整数 k，使得矩阵P^k 的每个元素都严格为正，即 $P^k_{s,\,s'} > 0$，对所有 $s, s' \in S$ 均成立。

对于每个马尔可夫矩阵 $P_{(i)}$，根据上述事实以及论文中的假设 3，我们知道存在这样的整数 k_i。然后我们定义集合 $J = \{i \in [N] | w_i > 0\}$。由于 $\sum_{i=1}^{N} w_i = 1$，并且 $w_i \geq 0$ 对所有 $i \in [N]$均成立，因此我们知道 J 是一个非空集合。如果定义 $k = \min^{i\in[J]}\{k_i\}$ 以及 $j = \arg\min^{i\in[J]}\{k_i\}$，则有：

$$\left(\sum_{i\in[N]} w_i P^{(i)}\right)^{\bar{k}} = \underbrace{w_j^{\bar{k}}\left(P^{(j)}\right)^{\bar{k}}}_{\text{positive}} + \underbrace{\cdots\cdots}_{\text{nonnegative}}, \tag{8}$$

其中，$w_j^{\bar{k}}\left(P^{(j)}\right)^{\bar{k}}$ 的每个元素都严格为正，而求和中的其他矩阵均为非负矩阵。因此，我们可以得出结论：与马尔可夫矩阵 $\sum_{i\in[N]} w_i P^{(i)}$ 相关联的马尔可夫链也是不可约且非周期的。

C.2 命题 2 的证明

遵循与定理 1 中类似的论证，我们对 $\|A_i - A\|$ 进行界定：

$$\begin{aligned}\|\bar{A}_i - \bar{A}\| &= \|\Phi^\top D^{(i)}(\Phi - \gamma P^{(i)}\Phi) - \Phi^\top \bar{D}(\Phi - \gamma\bar{P}\Phi)\| \\ &\overset{(a)}{\leq} \gamma\|\Phi\|^2\|D^{(i)}\|\|P^{(i)} - \bar{P}\| + \|\Phi\|^2\|D^{(i)} - \bar{D}\|\|(I - \gamma\bar{P})\| \\ &\overset{(b)}{\leq} \gamma\sqrt{n}\epsilon + (1+\gamma)[2(n-1)\epsilon + \mathcal{O}(\epsilon^2)] = A(\epsilon),\end{aligned} \tag{9}$$

其中，不等式 (a) 的推导与方程 (5) 中 (a) 的推理相同，(b) 使用了与方程 (5) 中 (b) 相同的事实，并且 $\|P^{(i)} - \bar{P}\| \leq \frac{1}{N}\sum_{j=1}^{N}\|P^{(i)} - P^{(j)}\| \leq \epsilon\sqrt{n}$ 以及 $\|D^{(i)} - \bar{D}\| \leq 2(n-1)\epsilon + \mathcal{O}(\epsilon^2)$。

基于上述事实：(i) $\|\bar{R}\| \leq \frac{1}{N}\sum_{i=1}^{N}\|R^{(i)}\| \leq R_{\max}$，(ii) $\|R^{(i)} - \bar{R}\| \leq \frac{1}{N}\sum_{j=1}^{N}\|R^{(i)} - R^{(j)}\| \leq \epsilon_1$以及 (iii) $\|D^{(i)} - \bar{D}\| \leq 2(n-1)\epsilon + \mathcal{O}(\epsilon^2)$，我们通过证明 $\|\bar{b}_i - \bar{b}\| \leq b(\epsilon,\ \epsilon_1)$ 来完成证明。为此，我们遵循方程 (6) 中的相同步骤，并通过与方程 (7) 相同的分析，对 $\|\theta_i^* - \theta^*\|$ 的界进行证明。

C.3 命题3的证明

由于虚拟马尔可夫决策过程是各智能体的马尔可夫决策过程的平均，即$\bar{P} = \frac{1}{N}\sum_{i=1}^{N} P^{(i)}$，根据命题 1，虚拟马尔可夫链是不可约且非周期的。对称半正定矩阵的最大特征值是凸函数。因此，我们有 $\lambda_{\max}(\Phi^\top\bar{D}\Phi) \leq \sum_{s\in\mathcal{S}}\bar{\pi}(s)\lambda_{\max}\left(\phi(s)\phi(s)^\top\right) \leq \sum_{s\in\mathcal{S}}\bar{\pi}(s) = 1$.

为了证明存在 $\omega > 0$ 使得 $\lambda_{\min}(\Phi^\top\bar{D}\Phi) \geq \omega > 0$，我们将证明 $\Phi^\top\bar{D}\Phi$ 是一个正定矩阵。由于 Φ 列满秩，这相当于证明 D 是一个正定矩阵。根据 D的定义，证明 D 正定等价于论证平稳分布向量 $\bar{\pi}$ 的每个元素都严格为正；这里，$\bar{\pi}^\top\bar{P} = \bar{\pi}$. 为此，根据命题 1，我们知道，与 P 相关联的马尔可夫链是非周期且不可约的。根据佩龙-弗罗贝尼乌斯定理（Frobenius 等人，1912年），我们可以得出结论，$\bar{\pi}$ 的每个元素确实都严格为正。如果我们选择$\omega = \min_{s\in\mathcal{S}}\{\bar{\pi}(s)\} > 0$，则有 $\lambda_{\min}(\Phi^\top\bar{D}\Phi) \geq \omega > 0$.

D 伪梯度异质性：引理2 的证明

对于每个 $\theta \in \mathcal{H}$，我们有：

$$
\begin{aligned}
\Big\|\bar{g}(\theta) - \frac{1}{N}\sum_{i=1}^{N}\bar{g}_i(\theta)\Big\| &= \Big\|\Phi^T\bar{D}(\bar{T}_\mu\Phi\theta - \Phi\theta) - \frac{1}{N}\Big(\sum_{i=1}^{N}\Phi^T D^{(i)}(T_\mu^{(i)}\Phi\theta - \Phi\theta)\Big)\Big\| \\
&\overset{(a)}{\leq} \frac{1}{N}\sum_{i=1}^{N}\Big\|\Phi^T\bar{D}(\bar{T}_\mu\Phi\theta - \Phi\theta) - \Phi^T D^{(i)}(T_\mu^{(i)}\Phi\theta - \Phi\theta)\Big\| \\
&\overset{(b)}{\leq} \frac{1}{N}\sum_{i=1}^{N}\Big\|\bar{D}\Big[\frac{1}{N}\sum_{j=1}^{N}R^{(j)} + \gamma\frac{1}{N}\sum_{j=1}^{N}P^{(j)}\Phi\theta - \Phi\theta\Big] - D^{(i)}(T_\mu^{(i)}\Phi\theta - \Phi\theta)\Big\| \\
&\leq \frac{1}{N}\sum_{i=1}^{N}\Big\|\bar{D}\Big[\frac{1}{N}\sum_{j=1}^{N}R^{(j)} + \gamma\frac{1}{N}\sum_{j=1}^{N}P^{(j)}\Phi\theta - \Phi\theta\Big] - \bar{D}(T_\mu^{(i)}\Phi\theta - \Phi\theta) \\
&\quad + \bar{D}(T_\mu^{(i)}\Phi\theta - \Phi\theta) - D^{(i)}(T_\mu^{(i)}\Phi\theta - \Phi\theta)\Big\| \\
&\overset{(c)}{\leq} \frac{1}{N}\sum_{i=1}^{N}\Big\|\bar{D}\Big[\frac{1}{N}\sum_{j=1}^{N}R^{(j)} + \gamma\frac{1}{N}\sum_{j=1}^{N}P^{(j)}\Phi\theta - \Phi\theta\Big] - \bar{D}(T_\mu^{(i)}\Phi\theta - \Phi\theta)\Big\| \\
&+ \frac{1}{N}\sum_{i=1}^{N}\Big\|\bar{D}(T_\mu^{(i)}\Phi\theta - \Phi\theta) - D^{(i)}(T_\mu^{(i)}\Phi\theta - \Phi\theta)\Big\| \\
&\leq \frac{1}{N}\sum_{i=1}^{N}\Big\|\bar{D}\Big\|\Big\|\frac{1}{N}\sum_{j=1}^{N}R^{(j)} - R^{(i)}\Big\| + \gamma\Big\|\frac{1}{N}\sum_{j=1}^{N}P^{(j)} - P^{(i)}\Big\|\Big\|\Phi\theta\Big\| \\
&+ \frac{1}{N}\sum_{i=1}^{N}\Big\|\bar{D} - D^{(i)}\Big\|\Big\|T_\mu^{(i)}\Phi\theta - \Phi\theta\Big\| \\
&\overset{(d)}{\leq} \frac{1}{N}\sum_{i=1}^{N}\Big\|\frac{1}{N}\sum_{j=1}^{N}R^{(j)} - R^{(i)}\Big\|_2 + \gamma\Big\|\frac{1}{N}\sum_{j=1}^{N}P^{(j)} - P^{(i)}\Big\|\Big\|\Phi\theta\Big\| \\
&+ \frac{1}{N}\sum_{i=1}^{N}\Big\|\bar{D} - D^{(i)}\Big\|\Big\|T_\mu^{(i)}\Phi\theta - \Phi\theta\Big\| \\
&\overset{(e)}{\leq} \Big[\epsilon_1 + \gamma\sqrt{n}\epsilon\|\Phi\theta\| + \Big[2(n-1)\epsilon + \mathcal{O}(\epsilon^2)\Big]\|\Phi\theta\| \\
&\leq H\Big[\mathcal{O}(\epsilon_1) + \gamma\sqrt{n}\epsilon + 2(n-1)\epsilon + \mathcal{O}(\epsilon^2)\Big] = B(\epsilon, \epsilon_1). \qquad (10)
\end{aligned}
$$

不等式 (a) 和 (c) 由三角不等式得出，(b) 是由于 $\|\Phi\| \leq 1$；(d) 是由于以下事实$\|\bar{D}\| \leq 1$；而 (e) 则利用以下事实：(i) $\|R^{(i)} - \bar{R}\| \leq \epsilon_1$；(ii) $\|P^{(i)} - P^{(j)}\| \leq \sqrt{n}\|P^{(i)} - P^{(j)}\|_\infty \leq \epsilon\sqrt{n}\|P_{(i)}\|_\infty = \epsilon\sqrt{n}$，这又可由定理 1的证明得出；(iii) $\|D^{(i)} - \bar{D}\| \leq 2(n-1)\epsilon + \mathcal{O}(\epsilon^2)$，这又可由定理 1 或方程 (5) 的证明得出；以及 (iv) $\|\theta\| \leq H$ 对于任意 $\theta \in \mathcal{H}$。

E 独立同分布和马尔可夫设置中使用的辅助结果

在本附录中，我们反复使用以下不等式（通常不作明确说明）：

- 给定任意两个向量 x，$y \in \mathbb{R}^d$，对于任意 $\beta > 0$，有

$$
\|x + y\|^2 \leq (1 + \beta)\|x\|^2 + \left(1 + \frac{1}{\beta}\right)\|y\|^2. \qquad (11)
$$

- 给定任意两个向量 x， $y \in \mathbb{R}^d$，对于任意 $\beta > 0$，有

$$\langle x, y \rangle \leq \frac{\beta}{2}\|x\|^2 + \frac{1}{2\beta}\|y\|^2. \tag{12}$$

  这个不等式称为杨氏不等式。

- 给定 m 向量 x_1，…，$x_m \in \mathbb{R}^d$，下面是詹森不等式的一个简单应用：

$$\left\|\sum_{i=1}^{m} x_i\right\|^2 \leq m\sum_{i=1}^{m}\|x_i\|^2. \tag{13}$$

我们证明虚拟马尔可夫决策过程的以下结论。

引理3. 对于任意 θ_1， $\theta_2 \in \mathbb{R}^d$，

$$(\theta_2 - \theta_1)^\top [\bar{g}(\theta_1) - \bar{g}(\theta_2)] \geq (1-\gamma)\left\|\hat{V}_{\theta_1} - \hat{V}_{\theta_2}\right\|_{\bar{D}}^2. \tag{14}$$

证明。 考虑一个平稳状态序列，其随机初始状态为 $s \sim \bar{\pi}$，后续状态为 s'；在给定 s的条件下，后续状态从 $\bar{P}(\cdot \mid s)$中抽取。定义 $\phi \triangleq \phi(s)$和 $\phi' \triangleq \phi(s')$。定义 $\chi_1 \triangleq \hat{V}_{\theta_2}(s) - \hat{V}_{\theta_1}(s) = (\theta_2 - \theta_1)^\top \phi$ 以及 $\chi_2 \triangleq \hat{V}_{\theta_2}(s') - \hat{V}_{\theta_1}(s') = (\theta_2 - \theta_1)^\top \phi'$。根据平稳性， χ_1 和 χ_2 是两个边缘分布相同的相关随机变量。根据定义，$\mathbb{E}[\chi_1^2] = \mathbb{E}[\chi_2^2] = \left\|\hat{V}_{\theta_2} - \hat{V}_{\theta_2}\right\|_{\bar{D}}^2$，因为 s， s' 是从 $\bar{\pi}$中抽取的。并且我们有，

$$\bar{g}(\theta_1) - \bar{g}(\theta_2) = \mathbb{E}\left[\phi\left(\gamma\phi' - \phi\right)^\top(\theta_1 - \theta_2)\right] = \mathbb{E}\left[\phi\left(\chi_1 - \gamma\chi_2\right)\right].$$

因此，

$$\begin{aligned}(\theta_2 - \theta_1)^\top [\bar{g}(\theta_1) - \bar{g}(\theta_2)] &= \mathbb{E}\left[\chi_1\left(\chi_1 - \gamma\chi_2\right)\right] \\ &= \mathbb{E}\left[\chi_1^2\right] - \gamma\mathbb{E}\left[\chi_1\chi_2\right] \\ &\geq (1-\gamma)\mathbb{E}\left[\chi_1^2\right] \\ &= (1-\gamma)\left\|\hat{V}_{\theta_2} - \hat{V}_{\theta_2}\right\|_{\bar{D}}^2,\end{aligned}$$

其中我们使用柯西-施瓦茨不等式得出 $\mathbb{E}[\chi_1\chi_2] \leq \sqrt{\mathbb{E}[\chi_1^2]}\sqrt{\mathbb{E}[\chi_2^2]} = \mathbb{E}[\chi_1^2]$。 □

引理4. 对于任意 θ_1、 $\theta_2 \in \mathbb{R}^d$，有

$$\|\bar{g}(\theta_1) - \bar{g}(\theta_2)\| \leq 2\left\|\hat{V}_{\theta_1} - \hat{V}_{\theta_2}\right\|_{\bar{D}}. \tag{15}$$

证明。根据引理3的分析，我们有

$$\begin{aligned}\|\bar{g}(\theta_1) - \bar{g}(\theta_2)\| &= \|\mathbb{E}\left[\phi\left(\chi_1 - \gamma\chi_2\right)\right]\| \\ &\leq \sqrt{\mathbb{E}\left[\|\phi\|^2\right]}\sqrt{\mathbb{E}\left[\left(\chi_1 - \gamma\chi_2\right)^2\right]} \\ &\leq \sqrt{\mathbb{E}\left[\chi_1^2\right]} + \gamma\sqrt{\mathbb{E}\left[\chi_2^2\right]} \\ &= (1+\gamma)\sqrt{\mathbb{E}\left[\chi_1^2\right]},\end{aligned}$$

其中，第二个不等式由 $\|\phi\| \leq 1$ 得出，最后一个等式由 $\mathbb{E}\left[\chi_1^2\right] = \mathbb{E}\left[\chi_2^2\right]$ 得出。利用 $\mathbb{E}\left[\chi_1^2\right] = \left\|\hat{V}_{\theta_2} - \hat{V}_{\theta_2}\right\|_{\bar{D}}^2$ 和 $1 + \gamma \leq 2$ 这一事实，即可完成证明。 □

利用该引理，接下来我们证明稳态 TD(0) 更新方向 $\bar{g}$ 和 $\bar{g}_i$ 是 2-利普希茨的 .

引理 5. *(2-稳态 TD(0) 更新方向的利普希茨性)* 对任意 $\theta_1, \theta_2 \in \mathbb{R}^d$，有

$$\|\bar{g}(\theta_1) - \bar{g}(\theta_2)\| \leq 2\|\theta_1 - \theta_2\|. \tag{16}$$

对于每个智能体 $i \in [N]$，有

$$\|\bar{g}_i(\theta_1) - \bar{g}_i(\theta_2)\| \leq 2\|\theta_1 - \theta_2\|. \tag{17}$$

证明。 根据引理 4，我们可以轻松得出，虚拟马尔可夫决策过程的稳态 TD(0) 更新方向 $\bar{g}$ 是 2-利普希茨的，即

$$\|\bar{g}(\theta_1) - \bar{g}(\theta_2)\| \leq 2\|\theta_1 - \theta_2\|, \tag{18}$$

根据 $\lambda_{\max}(\Phi^\top \bar{D}\Phi) \leq 1$这一事实。我们可以用相同的推理证明方程 (17)，因为$\|\bar{g}_i(\theta_1) - \bar{g}_i(\theta_2)\| \leq 2\left\|\hat{V}_{\theta_1} - \hat{V}_{\theta_2}\right\|_{D_i}$ 对每个 $i \in [N]$ 均成立，这一点可见于 Bhandari 等人 (2018年)。 □

接下来，我们证明引理 5中关于每个智能体的随机 TD(0) 更新方向的利普希茨性质的类似结论 i。

引理 6. *(2-利普希茨连续性随机 TD(0) 更新方向)* 对于任意 θ_1、$\theta_2 \in \mathbb{R}^d$ 和 $i \in [N]$，我们有 e

$$\|g_i(\theta_1) - g_i(\theta_2)\| \leq 2\|\theta_1 - \theta_2\|.$$

证明。在本证明中，我们将用到这样一个事实：在第 t 个通信轮次的第 k 次本地更新时，智能体 i 的随机 TD(0) 更新方向是参数 θ 的仿射函数。具体而言，我们有$g_i(\theta) = b_i(O_{t\,k}^{(i)}) - A_i(O_{t\,k}^{(i)})\theta$，其中 $A_i(O_{t\,k}^{(i)}) = \phi(s_{t\,k}^{(i)})(\phi^\top(s_{t\,k}^{(i)}) - \gamma\phi^\top(s_{t\,k+1}^{(i)}))$ 且 $b_i(O_{t\,k}^{(i)}) = r(s_{t\,k}^{(i)})\phi(s_{t\,k}^{(i)})$。因此，我们有

$$\begin{aligned}
\|g_i(\theta_1) - g_i(\theta_2)\| &= \left\|A_i(O_{t,k}^{(i)})(\theta_1 - \theta_2)\right\| \\
&\leq \left\|A_i(O_{t,k}^{(i)})\right\|\|\theta_1 - \theta_2\| \\
&\leq \left(\left\|\phi(s_{t,k}^i)\right\|^2 + \gamma\left\|\phi(s_{t,k}^i)\right\|\left\|\phi(s_{t,k+1}^i)\right\|\right)\|\theta_1 - \theta_2\| \\
&\leq 2\|\theta_1 - \theta_2\|,
\end{aligned}$$

其中，最后一步用到了 $\|\phi(s)\| \leq 1$，$\forall s \in \mathcal{S}$。 □

F 记号

在后续分析中，我们用 $\mathcal{F}_k^t$ 表示滤过，它涵盖了截至第 k 个本地步、轮次 t内的所有随机性。我们还用 $\mathcal{F}^t$ 表示滤过，它涵盖了截至轮次 $t-1$结束时的所有随机性。略微滥用记号，$\mathcal{F}_{-1}^t$ 应解释为$\mathcal{F}^t$。根据对 FedTD(0) 的描述，显然，对于每个 $i \in [N]$， $\theta_{t\,k}^{(i)}$. 是$\mathcal{F}_{k-1}^t$-可测的，而 $\bar{\theta}_t$ 是$\mathcal{F}^t$-可测的。此外，我们用 $\mathbb{E}_t$ 表示以截至轮次 $t-1$结束时的所有随机性为条件的期望。

为简便起见，我们定义 $\delta_t = \frac{1}{NK}\sum_{i=1}^{N}\sum_{k=0}^{K-1}\left\|\theta_t^{(i)},_k - \bar{\theta}_t\right\|$ 和 $\Delta_t = \frac{1}{NK}\sum_{i=1}^{N}\sum_{k=0}^{K-1}\left\|\theta_t^{(i)},_k - \bar{\theta}_t\right\|^2$。后一个项称为漂移项。注意，根据詹森不等式，对于所有 t ，$(\delta_t)^2 \leq \Delta_t$ 均成立。除非另有说明， $\|\cdot\|$ 表示欧几里得范数。

步长：在本文中，我们会遇到三种步长：局部步长 α_l、全局步长α_g和有效步长 α。我们的一些结果将依赖于随通信轮次 t变化而衰减的有效步长；我们将用 $\{\alpha_t\}$表示这样一个衰减的有效步长序列。局部步长 α_ℓ 始终保持不变，而有效步长的衰减将通过使服务器处的全局步长随通信轮次变化而衰减来实现。因此，我们将用$\{\alpha_g^{(t)}\}$ 表示服务器处衰减的全局步长序列。接下来，除非下标中另有说明，证明中出现的所有步长均指有效步长。

G 热身： FedTD 在独立同分布采样下的分析

为隔离异质性的影响并深入理解我们的主要证明思路，本节将分析一个更简单的独立同分布设定。具体而言，我们假设对于每个智能体 $i \in [N]$，数据元组$\{O_{t\,k}^{(i)},\}$ 均从马尔可夫矩阵 $\pi^{(i)}$ 的平稳分布 $P^{(i)}$中独立同分布地采样。这种独立同分布假设在强化学习算法的有限时间分析中很常见（Dalal等人，2018年；Bhandari 等人，2018年；Doan 等人，2019年）。接下来，对于固定的 θ 以及每个 $i \in [N]$，定义 $\bar{g}_i(\theta) \triangleq \mathbb{E}_{O^{(i)}}[g_i(\theta)]$ 作为期望的 TD（0）次迭代点处的更新方向 θ 当马尔可夫元组 $O_{t\,k}^{(i)}$ 达到其平稳分布 $\pi^{(i)}$。我们采用以下标准的有界方差假设（Bhandari 等人，2018年）；联邦学习分析中也采用了类似假设。

假设4。 $\mathbb{E}\|g_i(\theta) - \bar{g}_i(\theta)\|^2 \le \sigma^2$ 对所有智能体 $i \in [N]$、每一轮次 t 及本地更新 k均成立，并且$\forall\theta$.

令 H 表示集合$\mathcal{H}$的半径。此外，定义 $G \triangleq R_{\max} + 2H$ 和 $\nu = (1-\gamma)\bar{\omega}$，其中 $\bar{\omega}$ 与命题3中的定义相同。以下给出 FedTD（0）在独立同分布设定下的收敛结果。

定理4。 *(独立同分布设定)* 存在一个递减的全局步长序列 $\{\alpha_g^{(t)}\}$、一个固定的本地步长 α_l，以及一组凸权重，使得全局模型 $\{\theta_t\}$ 的一个凸组合 θ_T 在经过 T 轮次后，对每个 $i \in [N]$ 都满足以下条件：

$$\mathbb{E}\left\|V_{\tilde{\theta}_T} - V_{\theta_i^*}\right\|_{\bar{D}}^2 \le \tilde{\mathcal{O}}\Big(\frac{G^2}{K^2T^2} + \frac{\sigma^2}{\nu^2 NKT} + \frac{\sigma^2}{\nu^4 KT^2} + Q(\epsilon,\epsilon_1)\Big), \tag{19}$$

其中 $Q(\epsilon,\epsilon_1) = \tilde{\mathcal{O}}\big(\frac{B(\epsilon,\epsilon_1)}{\nu}\big)^G + \Gamma^2(\epsilon,\epsilon_1))$，$B(\epsilon,\epsilon_1) = H\big(\sqrt{n}\epsilon + 2(n-1)\epsilon + \mathcal{O}(\epsilon^2) + \mathcal{O}(\epsilon_1)\big)$，且 $\Gamma(\epsilon,\epsilon_1)$ 的定义见定理*1*。

接下来，我们将对上述结果进行详细的收敛分析。

定理4的辅助引理 G.1

G.1.1 方差缩减

引理 7。 *(独立同分布设定下的方差缩减)*。在独立同分布设定下，根据假设*4*，在每一轮次 t，我们有 $\mathbb{E}\left\|\frac{1}{NK}\sum_{i=1}^{N}\sum_{k=0}^{K-1}\left[g_i(\theta_{t\,k}^{(i)},) - \bar{g}_i(\theta_{t\,k}^{(i)},)\right]\right\|^2 \le \frac{\sigma^2}{NK}$。 ——

证明。定义 $Y_{(i)\,t,k} \triangleq g_i(O_{t\,k}^{(i)},,\theta_{t\,k}^{(i)},) - \bar{g}_i(\theta_{t\,k}^{(i)},)$。由于 $\{O_{t\,k}^{(i)},\}$ 随时间从其平稳分布 $\pi^{(i)}$ 中独立同分布地抽取，因此我们有 $\mathbb{E}[Y_{(i}{}_{k}] = \mathbb{E}\left[\mathbb{E}[Y_{)\,t,(i}{}_{k} \mid \theta_{t}^{(i)})\,t,,k}]\right] = 0.$ 如前所述，对于每个 $i \in [N]$, $\theta_{t\,k}^{(i)}$, 都是 F_{k-1}^{t}-可测的。若以 F_{k-1}^{t} 为条件，则可知 $\theta_{t\,k}^{(i)}$ 和 $\theta_{t\,k}^{(j)}$ 是确定的，而 $Y_{(i)\,t,k}$ 和 $Y_{(j)\,t,k}$ 中唯一的随机性来自 $O_{t\,k}^{(i)}$ 和 $O_{t\,k}^{(j)},,$ 它们相互独立。因此，在以 F_{k-1}^{t} 为条件时， $Y_{(i)\,t,k}$ 和 $Y_{(j)\,t,k}$ 相互独立。

对于每个 $i \ne j \in [N]$，我们有

$$\mathbb{E}\left[\left\langle Y_{t,k}^{(i)}, Y_{t,k}^{(j)}\right\rangle\right] = \mathbb{E}\left[\mathbb{E}\left[\left\langle Y_{t,k}^{(i)}, Y_{t,k}^{(j)}\right\rangle \mid \mathcal{F}_{k-1}^{t}\right]\right] \overset{(a)}{=} \mathbb{E}\left[\left\langle \mathbb{E}[Y_{t,k}^{(i)} \mid \mathcal{F}_{k-1}^{t}], \mathbb{E}[Y_{t,k}^{(j)} \mid \mathcal{F}_{k-1}^{t}]\right\rangle\right] = 0, \tag{20}$$

其中，(a) 是由以下事实得出的： $Y_{(i)\,t,\,k}$ 和 $Y_{(j)\,t,\,k}$ 在给定 $\mathcal{F}_{k-1}^{t}$. 的条件下相互独立。对于每个 $k < l$ 和 $i,\ j \in [N]$,

$$\mathbb{E}\left[\left\langle Y_{t,k}^{(i)}, Y_{t,l}^{(j)}\right\rangle\right] = \mathbb{E}\left[\mathbb{E}\left[\left\langle Y_{t,k}^{(i)}, Y_{t,l}^{(j)}\right\rangle \,\middle|\, \mathcal{F}_{l-1}^{t}\right]\right] = \mathbb{E}\left[\left\langle Y_{t,k}^{(i)}, \mathbb{E}[Y_{t,l}^{(j)} \mid \mathcal{F}_{l-1}^{t}]\right\rangle\right] = 0. \tag{21}$$

那么，

$$\mathbb{E}\left\|\frac{1}{NK}\sum_{i=1}^{N}\sum_{k=0}^{K-1}\left[g_i(\theta_{t,k}^{(i)}) - \bar{g}_i(\theta_{t,k}^{(i)})\right]\right\|^2$$

$$
\begin{aligned}
&= \mathbb{E}\left\|\frac{1}{NK}\sum_{i=1}^{N}\sum_{k=0}^{K-1}Y_{t,k}^{(i)}\right\|^2 \\
&= \frac{1}{N^2K^2}\sum_{i=1}^{N}\sum_{k=0}^{K-1}\mathbb{E}\|Y_{t,k}^{(i)}\|^2 + \frac{2}{N^2K^2}\underbrace{\sum_{i<j}\sum_{k=0}^{K-1}\mathbb{E}[\langle Y_{t,k}^{(i)}, Y_{t,k}^{(j)}\rangle]}_{0} \\
&+ \frac{2}{N^2K^2}\sum_{i,j=1}^{N}\sum_{k<l}\underbrace{\mathbb{E}[\langle Y_{t,k}^{(i)}, Y_{t,l}^{(j)}\rangle]}_{0} \\
&\leq \frac{\sigma^2}{NK},
\end{aligned}
$$

其中第二个方程　　等式源于方程 (20)) 和方程 (21)，最后一个不等式源于假设 4。 □

G.1.2 每轮进展

首先，我们在以下引理中刻画每次迭代时误差的下降。

引理8. （每轮进展）。若局部步长 α_l 满足 $\alpha_l \leq \frac{(1-\gamma\,\bar{\omega})}{48K}$，则对于任意全局步长 α_g，*FedTD(0)* 的更新满足

$$
\begin{aligned}
\mathbb{E}\|\bar{\theta}_{t+1}-\theta^*\|^2 \leq (1+\zeta_1)\mathbb{E}\left\|\bar{\theta}_t-\theta^*\right\|^2 + 2\alpha\mathbb{E}\langle\bar{g}(\bar{\theta}_t),\bar{\theta}_t-\theta^*\rangle + 6\alpha^2\mathbb{E}\left\|\bar{g}(\bar{\theta}_t)\right\|^2 \\
+ 4\alpha^2\left(\frac{1}{\zeta_1}+6\right)\mathbb{E}[\Delta_t] + \frac{2\alpha^2\sigma^2}{NK} + 2\alpha B(\epsilon,\epsilon_1)G + 6\alpha^2B^2(\epsilon,\epsilon_1),
\end{aligned}
\tag{22}
$$

其中 ζ_1 为任意正数，α 为有效步长，即 $\alpha = K\alpha_l\alpha_g$。

证明。

$$
\begin{aligned}
&\mathbb{E}\|\bar{\theta}_{t+1}-\theta^*\|^2 \\
&= \mathbb{E}\left\|\Pi_{2,\mathcal{H}}\left(\bar{\theta}_t + \frac{\alpha}{NK}\sum_{i=1}^{N}\sum_{k=0}^{K-1}g_i(\theta_{t,k}^{(i)}) - \theta^*\right)\right\|^2 \quad \text{(updating rule)} \\
&\leq \mathbb{E}\left\|\bar{\theta}_t + \frac{\alpha}{NK}\sum_{i=1}^{N}\sum_{k=0}^{K-1}g_i(\theta_{t,k}^{(i)}) - \theta^*\right\|^2 \quad \text{(projection is non-expansive)} \\
&= \mathbb{E}\left\|\bar{\theta}_t-\theta^*\right\|^2 + 2\mathbb{E}\langle\frac{\alpha}{NK}\sum_{i=1}^{N}\sum_{k=0}^{K-1}g_i(\theta_{t,k}^{(i)}),\bar{\theta}_t-\theta^*\rangle + \mathbb{E}\left\|\frac{\alpha}{NK}\sum_{i=1}^{N}\sum_{k=0}^{K-1}g_i(\theta_{t,k}^{(i)})\right\|^2 \\
&= \mathbb{E}\left\|\bar{\theta}_t-\theta^*\right\|^2 + \frac{2\alpha}{NK}\underbrace{\sum_{i=1}^{N}\sum_{k=0}^{K-1}\mathbb{E}\langle g_i(\theta_{t,k}^{(i)}) - \bar{g}_i(\theta_{t,k}^{(i)}),\bar{\theta}_t-\theta^*\rangle}_{C_1=0} \\
&+ \frac{2\alpha}{NK}\sum_{i=1}^{N}\sum_{k=0}^{K-1}\mathbb{E}\langle\bar{g}_i(\theta_{t,k}^{(i)}),\bar{\theta}_t-\theta^*\rangle + \mathbb{E}\left\|\frac{\alpha}{NK}\sum_{i=1}^{N}\sum_{k=0}^{K-1}g_i(\theta_{t,k}^{(i)})\right\|^2 \\
&= \mathbb{E}\left\|\bar{\theta}_t-\theta^*\right\|^2 + \frac{2\alpha}{NK}\sum_{i=1}^{N}\sum_{k=0}^{K-1}\mathbb{E}\langle\bar{g}_i(\theta_{t,k}^{(i)}),\bar{\theta}_t-\theta^*\rangle + \mathbb{E}\left\|\frac{\alpha}{NK}\sum_{i=1}^{N}\sum_{k=0}^{K-1}g_i(\theta_{t,k}^{(i)})\right\|^2 \\
&\leq \mathbb{E}\left\|\bar{\theta}_t-\theta^*\right\|^2 + \frac{2\alpha}{NK}\sum_{i=1}^{N}\sum_{k=0}^{K-1}\mathbb{E}\langle\bar{g}_i(\theta_{t,k}^{(i)}),\bar{\theta}_t-\theta^*\rangle \\
&+ 2\mathbb{E}\left\|\frac{\alpha}{NK}\sum_{i=1}^{N}\sum_{k=0}^{K-1}\left[g_i(\theta_{t,k}^{(i)}) - \bar{g}_i(\theta_{t,k}^{(i)})\right]\right\|^2 + 2\mathbb{E}\left\|\frac{\alpha}{NK}\sum_{i=1}^{N}\bar{g}_i(\theta_{t,k}^{(i)})\right\|^2 \quad \text{(Young's inequality (12))}
\end{aligned}
$$

$$
\begin{aligned}
&\overset{(a)}{\leq} \mathbb{E}\left\|\bar{\theta}_t-\theta^*\right\|^2+\frac{2\alpha}{NK}\sum_{i=1}^{N}\sum_{k=0}^{K-1}\mathbb{E}\langle\bar{g}_i(\theta_{t,k}^{(i)}),\bar{\theta}_t-\theta^*\rangle+\frac{2\sigma^2}{NK}+2\mathbb{E}\left\|\frac{\alpha}{NK}\sum_{i=1}^{N}\bar{g}_i(\theta_{t,k}^{(i)})\right\|^2\\
&=\mathbb{E}\left\|\bar{\theta}_t-\theta^*\right\|^2+\frac{2\alpha}{NK}\sum_{i=1}^{N}\sum_{k=0}^{K-1}\mathbb{E}\langle\bar{g}_i(\theta_{t,k}^{(i)})-\bar{g}_i(\bar{\theta}_t)+\bar{g}_i(\bar{\theta}_t)-\bar{g}(\bar{\theta}_t)+\bar{g}(\bar{\theta}_t),\bar{\theta}_t-\theta^*\rangle\\
&+2\mathbb{E}\left\|\frac{\alpha}{NK}\sum_{i=1}^{N}\sum_{k=0}^{K-1}\bar{g}_i(\theta_{t,k}^{(i)})\right\|^2+\frac{2\alpha^2\sigma^2}{NK}\\
&\leq\mathbb{E}\left\|\bar{\theta}_t-\theta^*\right\|^2+\frac{2\alpha}{NK}\sum_{i=1}^{N}\sum_{k=0}^{K-1}\mathbb{E}\langle\bar{g}_i(\theta_{t,k}^{(i)})-\bar{g}_i(\bar{\theta}_t),\bar{\theta}_t-\theta^*\rangle+\frac{2\alpha}{N}\sum_{i=1}^{N}\mathbb{E}\langle\bar{g}_i(\bar{\theta}_t)-\bar{g}(\bar{\theta}_t),\bar{\theta}_t-\theta^*\rangle\\
&+2\alpha\mathbb{E}\langle\bar{g}(\bar{\theta}_t),\bar{\theta}_t-\theta^*\rangle+2\mathbb{E}\left\|\frac{\alpha}{NK}\sum_{i=1}^{N}\sum_{k=0}^{K-1}\bar{g}_i(\theta_{t,k}^{(i)})\right\|^2+\frac{2\alpha^2\sigma^2}{NK}\\
&\leq(1+\zeta_1)\mathbb{E}\left\|\bar{\theta}_t-\theta^*\right\|^2+\frac{1}{\zeta_1}\mathbb{E}\left\|\frac{\alpha}{NK}\sum_{i=1}^{N}\sum_{k=0}^{K-1}\left[\bar{g}_i(\theta_{t,k}^{(i)})-\bar{g}_i(\bar{\theta}_t)\right]\right\|^2+2\alpha B(\epsilon,\epsilon_1)G\\
&+2\alpha\mathbb{E}\langle\bar{g}(\bar{\theta}_t),\bar{\theta}_t-\theta^*\rangle+2\mathbb{E}\left\|\frac{\alpha}{NK}\sum_{i=1}^{N}\sum_{k=0}^{K-1}\bar{g}_i(\theta_{t,k}^{(i)})\right\|^2+\frac{2\alpha^2\sigma^2}{NK}\quad\text{(Eq (12) and Lemma 2)}\\
&\leq(1+\zeta_1)\mathbb{E}\left\|\bar{\theta}_t-\theta^*\right\|^2+\frac{4\alpha^2}{\zeta_1 NK}\sum_{i=1}^{N}\sum_{k=0}^{K-1}\mathbb{E}\left\|\theta_{t,k}^{(i)}-\bar{\theta}_t\right\|^2+2\alpha B(\epsilon,\epsilon_1)G\\
&+2\alpha\mathbb{E}\langle\bar{g}(\bar{\theta}_t),\bar{\theta}_t-\theta^*\rangle+2\mathbb{E}\left\|\frac{\alpha}{NK}\sum_{i=1}^{N}\sum_{k=0}^{K-1}\bar{g}_i(\theta_{t,k}^{(i)})\right\|^2+\frac{2\alpha^2\sigma^2}{NK}\quad\text{(2-Lipschitz of }\bar{g}_i\text{ in Lemma 5)}\\
&\leq(1+\zeta_1)\mathbb{E}\left\|\bar{\theta}_t-\theta^*\right\|^2+\frac{4\alpha^2}{\zeta_1}\mathbb{E}[\Delta_t]+\frac{2\alpha^2\sigma^2}{NK}+2\alpha B(\epsilon,\epsilon_1)G\\
&+2\alpha\mathbb{E}\langle\bar{g}(\bar{\theta}_t),\bar{\theta}_t-\theta^*\rangle+2\mathbb{E}\left\|\frac{\alpha}{NK}\sum_{i=1}^{N}\sum_{k=0}^{K-1}\left[\bar{g}_i(\theta_{t,k}^{(i)})-\bar{g}_i(\bar{\theta}_t)+\bar{g}_i(\bar{\theta}_t)-\bar{g}(\bar{\theta}_t)+\bar{g}(\bar{\theta}_t)\right]\right\|^2\\
&\leq(1+\zeta_1)\mathbb{E}\left\|\bar{\theta}_t-\theta^*\right\|^2+\frac{4\alpha^2}{\zeta_1}\mathbb{E}[\Delta_t]+\frac{2\alpha^2\sigma^2}{NK}+2\alpha B(\epsilon,\epsilon_1)G\\
&+2\alpha\mathbb{E}\langle\bar{g}(\bar{\theta}_t),\bar{\theta}_t-\theta^*\rangle+6\mathbb{E}\left\|\frac{\alpha}{NK}\sum_{i=1}^{N}\sum_{k=0}^{K-1}\bar{g}_i(\theta_{t,k}^{(i)})-\bar{g}_i(\bar{\theta}_t)\right\|^2\\
&+6\mathbb{E}\left\|\frac{\alpha}{N}\sum_{i=1}^{N}\left[\bar{g}_i(\bar{\theta}_t)-\bar{g}(\bar{\theta}_t)\right]\right\|^2+6\mathbb{E}\left\|\alpha\bar{g}(\bar{\theta}_t)\right\|^2\quad\text{(Eq (12) and Lemma 2)}\\
&\leq(1+\zeta_1)\mathbb{E}\left\|\bar{\theta}_t-\theta^*\right\|^2+\frac{4\alpha^2}{\zeta_1}\mathbb{E}[\Delta_t]+\frac{2\alpha^2\sigma^2}{NK}+2\alpha B(\epsilon,\epsilon_1)G\\
&+2\alpha\mathbb{E}\langle\bar{g}(\bar{\theta}_t),\bar{\theta}_t-\theta^*\rangle+24\alpha^2\mathbb{E}[\Delta_t]\quad\text{(2-Lipschitz of }\bar{g}_i\text{)}\\
&+6\alpha^2B^2(\epsilon,\epsilon_1)+6\alpha^2\mathbb{E}\left\|\bar{g}(\bar{\theta}_t)\right\|^2\quad\text{(Eq (12))}\\
&=(1+\zeta_1)\mathbb{E}\left\|\bar{\theta}_t-\theta^*\right\|^2+2\alpha\mathbb{E}\langle\bar{g}(\bar{\theta}_t),\bar{\theta}_t-\theta^*\rangle+6\alpha^2\mathbb{E}\left\|\bar{g}(\bar{\theta}_t)\right\|^2\\
&+4\alpha^2\left(\frac{1}{\zeta_1}+6\right)\mathbb{E}[\Delta_t]+\frac{2\alpha^2\sigma^2}{NK}+2\alpha B(\epsilon,\epsilon_1)G+6\alpha^2B^2(\epsilon,\epsilon_1),
\end{aligned}
\tag{23}
$$

其中，(a) is due to Lemma 7. Furthermore, the reason why $\mathcal{C}_1=0$ is as follows:

$$
\mathcal{C}_1=\sum_{i=1}^{N}\sum_{k=0}^{K-1}\mathbb{E}\langle g_i(\theta_{t,k}^{(i)})-\bar{g}_i(\theta_{t,k}^{(i)}),\bar{\theta}_t-\theta^*\rangle
$$

$$
\begin{aligned}
&= \sum_{i=1}^{N}\sum_{k=0}^{K-2}\mathbb{E}\langle g_i(\theta_{t,k}^{(i)}) - \bar{g}_i(\theta_{t,k}^{(i)}), \bar{\theta}_t - \theta^*\rangle + \\
&\qquad \sum_{i=1}^{N}\mathbb{E}\langle g_i(\theta_{t,K-1}^{(i)}) - \bar{g}_i(\theta_{t,K-1}^{(i)}), \bar{\theta}_t - \theta^*\rangle \\
&= \sum_{i=1}^{N}\sum_{k=0}^{K-2}\mathbb{E}\langle g_i(\theta_{t,k}^{(i)}) - \bar{g}_i(\theta_{t,k}^{(i)}), \bar{\theta}_t - \theta^*\rangle + \\
&\qquad \sum_{i=1}^{N}\mathbb{E}\left[\mathbb{E}\left[\langle g_i(\theta_{t,K-1}^{(i)}) - \bar{g}_i(\theta_{t,K-1}^{(i)}), \bar{\theta}_t - \theta^*\rangle \mid \mathcal{F}_{K-1}^t\right]\right] \\
&= \sum_{i=1}^{N}\sum_{k=0}^{K-2}\mathbb{E}\langle g_i(\theta_{t,k}^{(i)}) - \bar{g}_i(\theta_{t,k}^{(i)}), \bar{\theta}_t - \theta^*\rangle + \\
&\qquad \sum_{i=1}^{N}\mathbb{E}\left[\left\langle \bar{\theta}_t - \theta^*, \underbrace{\mathbb{E}\left[g_i(\theta_{t,k}^{(i)}) - \bar{g}_i(\theta_{t,k}^{(i)}) \mid \mathcal{F}_{K-1}^t\right]}_{0}\right\rangle\right] \\
&= \sum_{i=1}^{N}\sum_{k=0}^{K-2}\mathbb{E}\langle g_i(\theta_{t,k}^{(i)}) - \bar{g}_i(\theta_{t,k}^{(i)}), \bar{\theta}_t - \theta^*\rangle.
\end{aligned}
$$

我们可以通过依次对$\mathcal{F}_{K-2}^t$、$\cdots$、$\mathcal{F}_1^t$、$\mathcal{F}_0^t$进行条件化，重复这一过程。 □

G.1.3 漂移项分析

现在我们来界定漂移项 Δ_t。

Lemma 9. （有界客户端漂移）第 t 轮的漂移项 Δ_t 可以被界定为 a s

$$
\mathbb{E}[\Delta_t] = \frac{1}{NK}\sum_{i=1}^{N}\sum_{k=0}^{K-1}\mathbb{E}\left\|\theta_{t,k}^{(i)} - \bar{\theta}_t\right\|^2 \le 27(\sigma^2 + 3KB^2(\epsilon,\epsilon_1) + 2KG^2)\frac{\alpha^2}{K\alpha_g^2}, \tag{24}
$$

若固定的局部步长 α_l 满足 $\alpha_l \le$最小值 $\frac{(1-\gamma\ \bar{\omega}\ )}{48K}$. ———

证明。

$$
\begin{aligned}
&\mathbb{E}\left\|\theta_{t,k}^{(i)} - \bar{\theta}_t\right\|^2 \\
&= \mathbb{E}\left\|\theta_{t,k-1}^{(i)} + \alpha_l g_i(\theta_{t,k-1}^{(i)}) - \bar{\theta}_t\right\|^2 \quad \text{(updating rule)} \\
&= \mathbb{E}\left\|\theta_{t,k-1}^{(i)} + \alpha_l \bar{g}_i(\theta_{t,k-1}^{(i)}) - \bar{\theta}_t + \alpha_l\left(g_i(\theta_{t,k-1}^{(i)}) - \bar{g}_i(\theta_{t,k-1}^{(i)})\right)\right\|^2 \\
&= \mathbb{E}\left\|\theta_{t,k-1}^{(i)} + \alpha_l \bar{g}_i(\theta_{t,k-1}^{(i)}) - \bar{\theta}_t\right\|^2 + \alpha_l^2\mathbb{E}\left\|g_i(\theta_{t,k-1}^{(i)}) - \bar{g}_i(\theta_{t,k-1}^{(i)})\right\|^2 \\
&+ 2\alpha_l \underbrace{\mathbb{E}\left[\mathbb{E}\left\langle g_i(\theta_{t,k-1}^{(i)}) - \bar{g}_i(\theta_{t,k-1}^{(i)}), \theta_{t,k-1}^{(i)} + \alpha_l \bar{g}_i(\theta_{t,k-1}^{(i)}) - \bar{\theta}_t \;\middle|\; \mathcal{F}_{k-1}^t\right\rangle\right]}_{\mathcal{C}_2=0} \\
&\overset{(a)}{\le} (1+\zeta_2)\mathbb{E}\left\|\theta_{t,k-1}^{(i)} + \alpha_l \bar{g}(\theta_{t,k-1}^{(i)}) - \bar{\theta}_t\right\|^2 + (1+\frac{1}{\zeta_2})\alpha_l^2\mathbb{E}\left\|\bar{g}(\theta_{t,k-1}^{(i)}) - \bar{g}_i(\theta_{t,k-1}^{(i)})\right\|^2 \\
&+ \alpha_l^2\mathbb{E}\left\|g_i(\theta_{t,k-1}^{(i)}) - \bar{g}_i(\theta_{t,k-1}^{(i)})\right\|^2 \\
&\overset{(b)}{\le} (1+\zeta_2)(1+\zeta_3)\mathbb{E}\left\|\theta_{t,k-1}^{(i)} + \alpha_l \bar{g}(\theta_{t,k-1}^{(i)}) - \bar{\theta}_t - \alpha_l \bar{g}(\bar{\theta}_t)\right\|^2
\end{aligned}
$$

$$
\begin{aligned}
&+(1+\zeta_2)(1+\frac{1}{\zeta_3})\alpha_l^2\mathbb{E}\left\|\bar{g}(\bar{\theta}_t)\right\|^2 \\
&+(1+\frac{1}{\zeta_2})\alpha_l^2\mathbb{E}\left\|\bar{g}(\theta_{t,k-1}^{(i)})-\bar{g}(\bar{\theta}_t)+\bar{g}(\bar{\theta}_t)-\bar{g}_i(\bar{\theta}_t)+\bar{g}_i(\bar{\theta}_t)-\bar{g}_i(\theta_{t,k-1}^{(i)})\right\|^2+\alpha_l^2\sigma^2 \\
&\overset{(c)}{\leq}(1+\zeta_2)(1+\zeta_3)\mathbb{E}\left\|\theta_{t,k-1}^{(i)}+\alpha_l\bar{g}(\theta_{t,k-1}^{(i)})-\bar{\theta}_t-\alpha_l\bar{g}(\bar{\theta}_t)\right\|^2+(1+\zeta_2)(1+\frac{1}{\zeta_3})\alpha_l^2\mathbb{E}\left\|\bar{g}(\bar{\theta}_t)\right\|^2 \\
&+3(1+\frac{1}{\zeta_2})\alpha_l^2\mathbb{E}\left\|\bar{g}(\theta_{t,k-1}^{(i)})-\bar{g}(\bar{\theta}_t)\right\|^2+3(1+\frac{1}{\zeta_2})\alpha_l^2\mathbb{E}\left\|\bar{g}(\bar{\theta}_t)-\bar{g}_i(\bar{\theta}_t)\right\|^2 \\
&+3(1+\frac{1}{\zeta_2})\alpha_l^2\mathbb{E}\left\|\bar{g}_i(\bar{\theta}_t)-\bar{g}_i(\theta_{t,k-1}^{(i)})\right\|^2+\alpha_l^2\sigma^2 \\
&\overset{(d)}{\leq}(1+\zeta_2)(1+\zeta_3)\left[1-(2\alpha_l(1-\gamma)-4\alpha_l^2)\bar{\omega}\right]\mathbb{E}\left\|\theta_{t,k-1}^{(i)}-\bar{\theta}_t\right\|^2 \\
&+(1+\zeta_2)(1+\frac{1}{\zeta_3})\alpha_l^2\mathbb{E}\left\|\bar{g}(\bar{\theta}_t)\right\|^2 \\
&+12(1+\frac{1}{\zeta_2})\alpha_l^2\mathbb{E}\left\|\theta_{t,k-1}^{(i)}-\bar{\theta}_t\right\|^2+3(1+\frac{1}{\zeta_3})\alpha_l^2B^2(\epsilon,\epsilon_1) \\
&+12(1+\frac{1}{\zeta_2})\alpha_l^2\mathbb{E}\left\|\theta_{t,k-1}^{(i)}-\bar{\theta}_t\right\|^2+\alpha_l^2\sigma^2 \\
&=(1+\zeta_2)(1+\zeta_3)\left[1-(2\alpha_l(1-\gamma)-4\alpha_l^2)\bar{\omega}+\frac{24(1+\frac{1}{\zeta_3})\alpha_l^2}{(1+\zeta_2)(1+\zeta_3)}\right]\mathbb{E}\left\|\theta_{t,k-1}^{(i)}-\bar{\theta}_t\right\|^2 \\
&+\underbrace{(1+\zeta_2)(1+\frac{1}{\zeta_3})\alpha_l^2\mathbb{E}\left\|\bar{g}(\bar{\theta}_t)\right\|^2+3(1+\frac{1}{\zeta_3})\alpha_l^2B^2(\epsilon,\epsilon_1)+\alpha_l^2\sigma^2}_{\mathcal{D}_1},
\end{aligned}
$$

其中，(a) 中我们使用了方程（11）中的不等式，且常数 ζ_2 为任意正数；(b) 中我们使用了假设 4，并采用了与方程（11）相同的推理，且常数 ζ_3为任意正数；(c) 中我们使用了方程（13）中的不等式来界定第三项；(d) 中我们使用引理 3 和引理 4 来界定第一项，使用 2-利普希茨性质来界定第三项和第五项，并使用梯度异质性界 $\bar{g}$， $\bar{g}_i$ （即引理 5）

根据引理 2 中的第四项。若定义 $\zeta_4 \triangleq (1+\zeta_2)(1+\zeta_3)$ $\left[1-(2\alpha_l(1-\gamma)-4\alpha_l^2)\bar{\omega}+\frac{24(1+\frac{1}{\zeta_3})^2}{\alpha(1+\zeta_2)(1+\zeta_3)}\right]$
并如上所述定义 $\mathcal{D}_1$ ，我们有

$$
\mathbb{E}\left\|\theta_{t,k}^{(i)}-\bar{\theta}_t\right\|^2\leq\zeta_4\mathbb{E}\left\|\theta_{t,k-1}^{(i)}-\bar{\theta}_t\right\|^2+\mathcal{D}_1. \tag{25}
$$

接下来，我们令 $\zeta_2=\zeta_3=\frac{1}{K-1}$， $K\geq 2$，并选择局部步长 α_l 使其满足

$$
\frac{\alpha_l(1-\gamma)\bar{\omega}}{2}\geq 4\alpha_l^2\bar{\omega}\ \ \&\ \ \frac{\alpha_l(1-\gamma)\bar{\omega}}{2}\geq\frac{24(1+\frac{1}{\zeta_3})\alpha_l^2}{(1+\zeta_2)(1+\zeta_3)},
$$

使得 $\left[1-(2\alpha_l(1-\gamma)-4\alpha_l^2)\bar{\omega}+\frac{24(1+\frac{1}{\zeta_2})\alpha_l^2}{(1+\zeta_2)(1+\zeta_3)}\right]\leq 1-\alpha_l(1-\gamma)\bar{\omega}$。当 $\alpha_l\leq$最小$(\frac{1-\gamma}{48K}\frac{\bar{\omega}}{})$方程 (25) 变为

$$
\mathbb{E}\left\|\theta_{t,k}^{(i)}-\bar{\theta}_t\right\|^2\leq(1+\frac{3}{K-1})\left[1-\alpha_l(1-\gamma)\bar{\omega}\right]\mathbb{E}\left\|\theta_{t,k-1}^{(i)}-\bar{\theta}_t\right\|^2+\mathcal{D}_1.
$$

将上述递推展开，并使用 $\theta_r^{(i)}{}_{,0}=\bar{\theta}_t$，可得

$$
\begin{aligned}
\mathbb{E}\left\|\theta_{t,k}^{(i)}-\bar{\theta}_t\right\|^2&\leq\sum_{s=0}^{k-1}\mathcal{D}_1\left\{\Pi_{j=s+1}^{k-1}(1+\frac{3}{K-1})\left[1-\alpha(1-\gamma)\bar{\omega}\right]\right\} \\
&\overset{(e)}{\leq}\sum_{s=0}^{k-1}\left[\alpha_l^2\sigma^2+3K\alpha_l^2B^2(\epsilon,\epsilon_1),+2\alpha_l^2K\mathbb{E}\left\|\bar{g}(\bar{\theta}_t)\right\|^2\right]
\end{aligned}
$$

$$
\begin{aligned}
&\times \Pi_{j=s+1}^{k-1}(1+\frac{3}{K-1})[1-\alpha_l(1-\gamma)\bar{\omega}] \\
\leq &\sum_{s=0}^{k-1}\left[\alpha_l^2\sigma^2+3\alpha_l^2KB^2(\epsilon,\epsilon_1)+2\alpha_l^2K\mathbb{E}\left\|\bar{g}(\bar{\theta}_t)\right\|^2\right] \\
&\times (1+\frac{3}{K-1})^{K-1}\Pi_{j=s+1}^{k-1}[1-\alpha_l(1-\gamma)\bar{\omega}] \\
\overset{(f)}{\leq} &27(\sigma^2+3KB^2(\epsilon,\epsilon_1)+2K\mathbb{E}\left\|\bar{g}(\bar{\theta}_t)\right\|^2)\sum_{s=0}^{k-1}\alpha_l^2\times\underbrace{\Pi_{j=s+1}^{k-1}[1-\alpha(1-\gamma)\bar{\omega}]}_{\leq 1} \\
\leq &27(\sigma^2+3KB^2(\epsilon,\epsilon_1)+2KG^2)K\alpha_l^2 \quad \text{(constant local step-size)}
\end{aligned}
$$

其中我们用到了以下事实：对于 (e)，$(1+\zeta_2)(1+\frac{1}{\zeta_3})\leq 2K$；对于 ~~$(f)$~~，$(1+\frac{3}{K-1})^{K-1}\leq 27$。将 $\alpha_l=\frac{\alpha}{K\alpha_g}$ 代入即可完成证明。 —— □

将方程 (24) 代入方程 (22，可得

$$
\begin{aligned}
&\mathbb{E}\left\|\bar{\theta}_{t+1}-\theta^*\right\|^2 \qquad (26)\\
\leq &(1+\zeta_1)\mathbb{E}\left\|\bar{\theta}_r-\theta^*\right\|^2+2\alpha\mathbb{E}\langle\bar{g}(\bar{\theta}_r),\bar{\theta}_r-\theta^*\rangle+6\alpha^2\mathbb{E}\left\|\bar{g}(\bar{\theta}_r)\right\|^2 \\
&+108\frac{\alpha^4}{K\alpha_g^2}(6+\frac{1}{\zeta_1})(\sigma^2+3KB^2(\epsilon,\epsilon_1)+2KG^2)+\frac{2\alpha^2\sigma^2}{NK}+2\alpha B(\epsilon,\epsilon_1)G+6\alpha^2B^2(\epsilon,\epsilon_1). \qquad (27)
\end{aligned}
$$

G.1.4 参数选择

引理 10。定义 $(1-\gamma)\bar{\omega}$。若选择任意有效步长 $\alpha=K\alpha_g\alpha_l<\frac{(1-\gamma)\bar{\omega}}{96}$、任意局部步长 ~~$\alpha_l\leq$~~ 最小值 $(\frac{1-\gamma\bar{\omega}}{48K})$，并选择常数 ~~$\zeta_1=$~~ $\alpha\nu$，则 *FedTD(0)* 的更新满足

$$
\begin{aligned}
\nu_1\mathbb{E}\left\|V_{\bar{\theta}_t}-V_{\theta^*}\right\|_{\bar{D}}^2\leq&(\frac{1}{\alpha}-\nu_1)\mathbb{E}\left\|\bar{\theta}_t-\theta^*\right\|^2-\frac{1}{\alpha}\mathbb{E}\left\|\bar{\theta}_{t+1}-\theta^*\right\|^2+\underbrace{\frac{2\alpha\sigma^2}{NK}}_{O(\alpha^1)} \\
&+\underbrace{\frac{1080\alpha^2}{K\alpha_g^2\nu}(\sigma^2+3KB^2(\epsilon,\epsilon_1)+2KG^2)}_{O(\alpha^2)}+\underbrace{2B(\epsilon,\epsilon_1)G+6\alpha B^2(\epsilon,\epsilon_1)}_{heterogeneity\ term}, \qquad (28)
\end{aligned}
$$

其中 $\nu_1=\frac{\nu}{4}=\frac{(1-\gamma)\bar{\omega}}{4}$。

证明。根据方程（26）以及 $\zeta_1=\alpha\nu$，我们知道

$$
\begin{aligned}
&\mathbb{E}\left\|\bar{\theta}_{t+1}-\theta^*\right\|^2 \\
\leq &(1+\zeta_1)\mathbb{E}\left\|\bar{\theta}_t-\theta^*\right\|^2+2\alpha\mathbb{E}\langle\bar{g}(\bar{\theta}_t),\bar{\theta}_t-\theta^*\rangle+6\alpha^2\mathbb{E}\left\|\bar{g}(\bar{\theta}_r)\right\|^2 \\
&+108\frac{\alpha^4}{K\alpha_g^2}(6+\frac{1}{\zeta_1})(\sigma^2+3KB^2(\epsilon,\epsilon_1)+2KG^2)+\frac{2\alpha^2\sigma^2}{NK}+2\alpha B(\epsilon,\epsilon_1)G+6\alpha^2B^2(\epsilon,\epsilon_1) \\
\leq &(1+\alpha\nu-2\alpha\nu)\mathbb{E}\left\|\bar{\theta}_t-\theta^*\right\|^2+24\alpha^2\mathbb{E}\left\|V_{\bar{\theta}_t}-V_{\theta^*}\right\|_{\bar{D}}^2+\frac{2\alpha^2\sigma^2}{NK} \quad \text{(Lemma 3 and 4)} \\
&+108\frac{\alpha^4}{K\alpha_g^2}(6+\frac{1}{\alpha\nu})(\sigma^2+3KB^2(\epsilon,\epsilon_1)+2KG^2)+2\alpha B(\epsilon,\epsilon_1)G+6\alpha^2B^2(\epsilon,\epsilon_1) \\
\leq &(1-\frac{\alpha\nu}{2})\mathbb{E}\left\|\bar{\theta}_t-\theta^*\right\|^2-\frac{\alpha\nu}{2}\mathbb{E}\left\|\bar{\theta}_t-\theta^*\right\|^2+24\alpha^2\mathbb{E}\left\|V_{\bar{\theta}_t}-V_{\theta^*}\right\|_{\bar{D}}^2+\frac{2\alpha^2\sigma^2}{NK}
\end{aligned}
$$

$$
\begin{aligned}
&+108\frac{\alpha^4}{K\alpha_g^2}(6+\frac{1}{\alpha\nu})(\sigma^2+3KB^2(\epsilon,\epsilon_1)+2KG^2)+2\alpha B(\epsilon,\epsilon_1)G+6\alpha^2B^2(\epsilon,\epsilon_1)\\
&\overset{(a)}{\leq}(1-\frac{\alpha\nu}{2})\mathbb{E}\left\|\bar{\theta}_t-\theta^*\right\|^2-\frac{\alpha\nu}{2}\mathbb{E}\left\|V_{\bar{\theta}_t}-V_{\theta^*}\right\|_{\bar{D}}^2+\frac{\alpha\nu}{4}\mathbb{E}\left\|V_{\bar{\theta}_t}-V_{\theta^*}\right\|_{\bar{D}}^2+\frac{2\alpha^2\sigma^2}{NK}\\
&+108\frac{\alpha^4}{K\alpha_g^2}(6+\frac{1}{\alpha\nu})(\sigma^2+3KB^2(\epsilon,\epsilon_1)+2KG^2)+2\alpha B(\epsilon,\epsilon_1)G+6\alpha^2B^2(\epsilon,\epsilon_1),
\end{aligned}
$$

其中 (a) 来自 $\lambda_{\max}(\Phi^T\bar{D}\Phi)\leq 1$ 和 $24\alpha^2\leq 24\alpha\frac{(1-\gamma)\bar{w}}{96}=\frac{\alpha\nu}{4}$。将 $\mathbb{E}\left\|V_{\bar{\theta}_t}-V_{\theta^*}\right\|_{\bar{D}}^2$（位于 (a) 右侧）移至上述不等式左侧，得到：

$$
\begin{aligned}
\frac{\alpha\nu}{4}\mathbb{E}\left\|V_{\bar{\theta}_t}-V_{\theta^*}\right\|_{\bar{D}}^2&\leq(1-\frac{\alpha\nu}{2})\mathbb{E}\left\|\bar{\theta}_t-\theta^*\right\|^2-\mathbb{E}\left\|\bar{\theta}_{t+1}-\theta^*\right\|^2+\frac{2\alpha^2\sigma^2}{NK}\\
&+108(\frac{6\alpha^4}{K\alpha_g^2}+\frac{\alpha^3}{K\alpha_g^2\nu})(\sigma^2+3KB^2(\epsilon,\epsilon_1)+2KG^2)+2\alpha B(\epsilon,\epsilon_1)G+6\alpha^2B^2(\epsilon,\epsilon_1).
\end{aligned}
$$

将上面不等式两边除以 α，并将 ν 改为 ν_1，可得 e:

$$
\begin{aligned}
&\nu_1\mathbb{E}\left\|V_{\bar{\theta}_t}-V_{\theta^*}\right\|_{\bar{D}}^2\\
&\leq(\frac{1}{\alpha}-\nu_1)\mathbb{E}\left\|\bar{\theta}_t-\theta^*\right\|^2-\frac{1}{\alpha}\mathbb{E}\left\|\bar{\theta}_{t+1}-\theta^*\right\|^2+\frac{2\alpha\sigma^2}{NK}\\
&+108(\frac{6\alpha^3}{K\alpha_g^2}+\frac{4\alpha^2}{K\alpha_g^2\nu_1})(\sigma^2+3KB^2(\epsilon,\epsilon_1)+2KG^2)+2B(\epsilon,\epsilon_1)G+6\alpha B^2(\epsilon,\epsilon_1)\\
&\leq(\frac{1}{\alpha}-\nu_1)\mathbb{E}\left\|\bar{\theta}_t-\theta^*\right\|^2-\frac{1}{\alpha}\mathbb{E}\left\|\bar{\theta}_{t+1}-\theta^*\right\|^2+\underbrace{\frac{2\alpha\sigma^2}{NK}}_{O(\alpha^1)}\\
&+\underbrace{\frac{1080\alpha^2}{K\alpha_g^2\nu_1}(\sigma^2+3KB^2(\epsilon,\epsilon_1)+2KG^2)}_{O(\alpha^2)}+\underbrace{2B(\epsilon,\epsilon_1)G+6\alpha B^2(\epsilon,\epsilon_1)}_{\text{heterogeneity term}},
\end{aligned}
$$

其中，最后一个不等式用到了 $\alpha\leq 1$ 这一事实。 □

有了这些引理，我们现在可以证明定理 4，为清晰起见，我们在此重述该定理。

G.2 定理 4 的证明

给定固定的局部步长 $\alpha_l=\frac{1}{2}(1-\gamma\ \bar{\omega})$，递减的有效步长 $\alpha_t=\frac{8}{\nu(a+t+1)}=\frac{8}{(1-\gamma)\bar{\omega}(a+t+1)}$，递减的全局步长 $\alpha_g^{(t)}=\frac{\alpha_t}{K\alpha_l}$ 以及权重 $w_t=(a+t)$，我们有

$$
\mathbb{E}\left\|V_{\tilde{\theta}_T}-V_{\theta_i^*}\right\|_{\bar{D}}^2\leq\tilde{\mathcal{O}}\left(\frac{G^2}{K^2T^2}+\frac{\sigma^2}{\nu^4KT^2}+\frac{\sigma^2}{\nu^2NKT}+\frac{B(\epsilon,\epsilon_1)G}{\nu}+\Gamma^2(\epsilon,\epsilon_1)\right)\tag{29}
$$

对任意智能体 $i\in[N]$ 均成立。

证明。我们取 $a>0$ 的有效步长 $\alpha_t=\frac{8}{\nu(a+t+1)}=\frac{2}{\nu_1(a+t+1)}$。此外，我们定义权重 $w_t=(a+t)$，并定义

$$
\tilde{\theta}_T=\frac{1}{W}\sum_{t=1}^{T}w_t\bar{\theta}_t,
$$

其中 $W=\sum_{t=1}^{T}w_t\geq\frac{1}{2}T(a+T)$。由正定二次型的凸性（$\lambda_{\min}(\Phi^T\bar{D}\Phi)\geq\bar{\omega}>0$），可得

$$
\nu_1\mathbb{E}\left\|V_{\tilde{\theta}_T}-V_{\theta^*}\right\|_{\bar{D}}^2
$$

$$
\begin{aligned}
&\leq \frac{\nu_1}{W}\sum_{t=1}^{T}(a+t)\mathbb{E}\left\|V_{\bar{\theta}_t}-V_{\theta^*}\right\|_{\bar{D}}^2 \\
&\overset{(28)}{\leq} \frac{\nu_1(a+1)(a+2)G^2}{2W}+\frac{1}{W}\sum_{t=1}^{T}\left[\frac{2(a+t)\alpha_t}{NK}\sigma^2\right] \\
&+\frac{1}{W}\sum_{t=1}^{T}\left[\frac{1080(a+t)\alpha_t^2}{K\alpha_g^2\nu_1}(\sigma^2+3KB^2(\epsilon,\epsilon_1)+2KG^2)\right] \\
&+\frac{1}{W}\sum_{t=1}^{T}(a+t)\left[2B(\epsilon,\epsilon_1)G+6\alpha_t B^2(\epsilon,\epsilon_1)\right] \\
&\leq \frac{\nu_1(a+1)(a+2)G^2}{2W}+\frac{2\sigma^2}{NKW}\sum_{t=1}^{T}(a+t)\alpha_t \\
&+\frac{1080(\sigma^2+3KB^2(\epsilon,\epsilon_1)+2KG^2)}{K\alpha_g^2\nu_1 W}\sum_{t=1}^{T}(a+t)\alpha_t^2+2B(\epsilon,\epsilon_1)G+\frac{6B^2(\epsilon,\epsilon_1)}{W}\sum_{t=1}^{T}(a+t)\alpha_t \\
&\leq \frac{\nu_1(a+1)(a+2)G^2}{2W}+\frac{4\sigma^2}{\nu_1 NKW}\cdot T \\
&+\frac{4320(\sigma^2+3KB^2(\epsilon,\epsilon_1)+2KG^2)}{K\alpha_g^2\nu_1^3 W}\cdot(1+\log(a+T))+2B(\epsilon,\epsilon_1)G+\frac{12B^2(\epsilon,\epsilon_1)}{\nu_1 W}\cdot T,
\end{aligned}
$$

其中我们使用了 $\left\|V_{\bar{\theta}_0}-V_{\theta^*}\right\|_{\bar{D}}^2 \leq G^2$.。将两边同时除以 ν_1，把 ν_1 改为 ν，并使用 $W \geq {}^{T}{}_{(a+T)\,2}$，可得:

$$
\mathbb{E}\left\|V_{\tilde{\theta}_T}-V_{\theta^*}\right\|_{\bar{D}}^2 \leq \tilde{\mathcal{O}}\left(\frac{G^2}{K^2T^2}+\frac{\sigma^2}{\nu^4KT^2}+\frac{\sigma^2}{\nu^2NKT}+\frac{B(\epsilon,\epsilon_1)G}{\nu}\right).
$$

我们结合定理 $\mathbb{E}\left\|V_{\tilde{\theta}_T}-V_{\theta_i^*}\right\|_{\bar{D}}^2 \leq 2\mathbb{E}\left\|V_{\tilde{\theta}_T}-V_{\theta^*}\right\|_{\bar{D}}^2+2\mathbb{E}\left\|V_{\theta_i^*}-V_{\theta^*}\right\|_{\bar{D}}^2$中的第三点，利用以下不等式完成证明:

1。 □

H 异质性偏差：定理3 的证明

本节中，我们证明定理3。

定理3 的证明。由于 θ_1^* 和 θ_2^* 分别是智能体1 和2的 TD(0) 不动点，因此有$\theta_1^* = \bar{A}_1^{-1}\bar{b}_1$ 和$\theta_2^* = \bar{A}_2^{-1}\bar{b}_2$。具有 $k = 1$ 和 $\alpha = \alpha_g\alpha_l$ 的均值路径 FedTD(0) 满足：

$$
\begin{aligned}
&\bar{\theta}_{t+1} = \bar{\theta}_t + \alpha(-\hat{A}\bar{\theta}_t + \hat{b}) \\
\implies &\bar{\theta}_{t+1} - \theta_1^* = \bar{\theta}_t - \theta_1^* + \alpha(-\hat{A}(\bar{\theta}_t - \theta_1^* + \theta_1^*) + \hat{b}) \\
\implies &e_{1,t+1} = (I - \alpha\hat{A})e_{1,t} - \alpha\hat{A}\theta_1^* + \alpha\hat{b} \\
\implies &e_{1,t+1} = (I - \alpha\hat{A})e_{1,t} - \alpha\left(\frac{\bar{A}_1 + \bar{A}_2}{2}\right)\bar{A}_1^{-1}\bar{b}_1 + \alpha\frac{\bar{b}_1 + \bar{b}_2}{2} \\
\implies &e_{1,t+1} = (I - \alpha\hat{A})e_{1,t} - \alpha\frac{\bar{A}_2\bar{A}_1^{-1}\bar{b}_1}{2} + \alpha\frac{\bar{b}_2}{2} \\
\implies &e_{1,t+1} = (I - \alpha\hat{A})e_{1,t} - \frac{\alpha\bar{A}_2}{2}\left(\bar{A}_1^{-1}\bar{b}_1 - \bar{A}_2^{-1}\bar{b}_2\right) \\
\implies &e_{1,t+1} = \underbrace{(I - \alpha\hat{A})}_{\tilde{A}}e_{1,t} + \underbrace{\frac{\alpha\bar{A}_2}{2}(\theta_2^* - \theta_1^*)}_{\tilde{y}}.
\end{aligned} \tag{30}
$$

$\tilde{}$ $\tilde{}$现在注意到 $e_{1,\ t+1} = \mathcal{A}e_{1,\ t} + \mathcal{Y}$ 可以视为一个离散时间线性时不变（LTI）系统 $\tilde{}$ $\tilde{}$，其中 α 的选取满足：$\mathcal{A}$ 是 Schur稳定的，即 $|\lambda_{\max}(\mathcal{A})| < 1$。在第 t 次迭代时，我们有：

$$
e_{1,t} = \tilde{\mathcal{A}}^t e_{1,0} + \sum_{k=0}^{t-1}\tilde{\mathcal{A}}^k\tilde{\mathcal{Y}}.
$$

$\tilde{}$由于 $t \to \infty$，小增益定理告诉我们，因为 $\rho(\mathcal{A}) < 1$ （其中 $\rho(\cdot)$ 表示谱半径），$\sum_{k=0}^{t-1}$ $\tilde{}$$\mathcal{A}^k$ 存在，且由 $(I -$ $\tilde{}$$\mathcal{A})^{-1}$ 给出。因此我们可以得出

$$
\begin{aligned}
\lim_{t\to\infty} e_{1,t} &= (I - \tilde{\mathcal{A}})^{-1}\tilde{\mathcal{Y}} \\
&= \left(\alpha\hat{A}\right)^{-1}\frac{\alpha\bar{A}_2}{2}(\theta_1^* - \theta_2^*) \\
&= \frac{1}{2}\hat{A}^{-1}\bar{A}_2(\theta_1^* - \theta_2^*).
\end{aligned} \tag{31}
$$

$e_{2,\ t}$ 的极限表达式可通过相同的分析得到。

I 马尔可夫设定的证明

现在，我们转而证明本文的主要结果，即定理 2。

I.1 概述

如正文所述，分析中需要克服的主要障碍之一是，一般而言，$\mathbb{E}[(1/N)\sum_{i=1}^{N}(g_i(\theta_{t\,k}^{(i)},,O_{t\,k}^{(i)},)-\bar{g}_i(\theta_{t\,k}^{(i)},))]\neq 0$。为了证明能够实现线性加速，我们首先在 I.2.1 小节中将每个智能体 i 的随机 TD 方向分解为 $g_i(\theta_{t\,k}^{(i)},)=b_i(O_{t\,k}^{(i)},)-A_i(O_{t\,k}^{(i)},)\theta_{t\,k}^{(i)}$，并在 I.2.2 小节中证明 $(1/NK)\sum_{i=1}^{N}\sum_{k=0}^{K-1}A_i(O_{t\,k}^{(i)},)$ 和 $(1/NK)\sum_{i=1}^{N}\sum_{k=0}^{K-1}b_i(O_{t\,k}^{(i)},)$ 的方差缩小了 NK。为了使用 Bhandari 等人（2018）提出的耦合的信息论控制方法，将参数 $\theta_{t\,k}^{(i)}$ 与观测$O_{t\,k}^{(i)}$ 之间的随机性解耦，我们需要在 I.2.3 小节中界定 $\mathbb{E}\left[\left\|\bar{\theta}_t-\bar{\theta}_{t-\tau}\right\|^2\right]$。与独立同分布设定和传统联邦学习中的分析一样，我们分别在 I.2.4、I.2.5 和 I.2.6 小节中刻画漂移项、每次迭代的误差下降和参数选择。最后，我们在 I.3 小节中证明定理 2。

附加记号：在假设3下，对于每个马尔可夫决策过程 i，存在某个 $m_i\geq 1$ 和某个 $\rho_i\in(0,1)$，使得对于所有 $t\geq 0$ 和 $0\leq k\leq K-1$，都有

$$d_{TV}\left(\mathbb{P}\left(s_{t,k}^{(i)}=\cdot\mid s_{0,0}^{(i)}=s\right),\pi^{(i)}\right)\leq m_i\rho_i^{tK+k},\forall s\in\mathcal{S}.$$

此外，我们定义 $\rho=\max^{i\in[N]}\{\rho_i\}$，$m=\max^{i\in[N]}\{m_i\}$。

定理2的辅助引理

I.2.1 分解形式

首先 我们证明定理2 的第一步是重写智能体 i的 FedTD(0)处的更新方向，表示为 ：

$$g_i(\theta_{t,k}^{(i)})=-A_i(O_{t,k}^{(i)})\theta_{t,k}^{(i)}+b_i(O_{t,k}^{(i)})$$

其中 $A_i(O_{t\,k}^{(i)},)=\phi(s_{t\,k}^{(i)},)(\phi^\top(s_{t\,k}^{(i)},)-\gamma\phi^\top(s_{t}^{(i)},_{k+1}))$ 且 $b_i(O_{t\,k}^{(i)},)=r(s_{t\,k}^{(i)},)\phi(s_{t\,k}^{(i)},)$。注意，$\mathbb{E}[b_i(O_{t\,k}^{(i)},)]$ 的稳态值不等于 0。为方便起见，我们对 g_i 进行适当的中心化处理，将其改写为:

$$g_i(\theta_{t,k}^{(i)})=-A_i(O_{t,k}^{(i)})(\theta_{t,k}^{(i)}-\theta_i^*)+\underbrace{b_i(O_{t,k}^{(i)})-A_i(O_{t,k}^{(i)})\theta_i^*}_{Z_i(O_{t,k}^{(i)})}. \tag{32}$$

定义 $Z_i(O_{t\,k}^{(i)},)\triangleq b_i(O_{t\,k}^{(i)},)-A_i(O_{t\,k}^{(i)},)\theta_i^*$。由于 $\bar{g}_i(\theta)\triangleq\mathbb{E}_{O_{t}^{(i)},_{k}\sim\pi^{(i)}}[g_i(\theta)],$，我们有:

$$\bar{g}_i(\theta_{t,k}^{(i)})=-\bar{A}_i(\theta_{t,k}^{(i)}-\theta_i^*). \tag{33}$$

其中 $\bar{A}_i=\Phi^\top D^{(i)}(\Phi-\gamma P^{(i)}\Phi)$。注意，$\mathbb{E}_{O_{t}^{(i)},_{k}\sim\pi^{(i)}}\left[Z_i(O_{t\,k}^{(i)},)\right]$ 等于 0。考虑到上述定义，我们得到以下引理:

引理 11. *(一致范数界)* 存在一些常数 c_1、c_2、$c_3\geq 0$，使得对所有 $i\in[N].$，$\left\|A_i\left(O_{t,k}^{(i)}\right)\right\|\leq c_1:=1+\gamma$，$\left\|\bar{A}_i\right\|\leq c_2:=1+\gamma$以及 $\left\|Z_i\left(O_{t,k}^{(i)}\right)\right\|\leq c_3:=R_{\max}+c_1H$均成立

证明. 根据定义以及 $\|\phi(s)\| \leq 1$ 这一事实，我们有

$$\left\|A_i\left(O_{t,k}^{(i)}\right)\right\| = \left\|\phi(s_{t,k}^{(i)})(\phi^\top(s_{t,k}^{(i)}) - \gamma\phi^\top(s_{t,k+1}^{(i)}))\right\| \leq \left\|\phi(s_{t,k}^{(i)})\right\|\left\|\phi^\top(s_{t,k}^{(i)}) - \gamma\phi^\top(s_{t,k+1}^{(i)})\right\| \leq 1+\gamma.$$

同样地，利用对于任意 $s \in \mathcal{S}$，均有 $r(s) \leq R_{\max}$ 这一事实，我们应用相同的推理可得

$$\left\|\bar{A}_i\right\| \leq 1+\gamma \;\;\&\;\; \left\|Z_i\left(O_{t,k}^{(i)}\right)\right\| \leq R_{\max} + c_1 H$$

. □

引理 12。 存在一些常数 L_1、$L_2 \geq 0$，使得

$$\begin{aligned}
&\left\|\bar{A}_i - \mathbb{E}\left[A_i\left(O_{t_2,k_2}^{(i)}\right) \mid \mathcal{F}_{k_1}^{t_1}\right]\right\| \leq L_1\rho^{(t_2-t_1)K+k_2-k_1},\\
&\left\|\bar{A}_i - \mathbb{E}_{t_1}\left[A_i\left(O_{t_2,k_2}^{(i)}\right)\right]\right\| \leq L_1\rho^{(t_2-t_1)K+k_2},\\
&\left\|\mathbb{E}\left[Z_i\left(O_{t_2,k_2}^{(i)}\right) \mid \mathcal{F}_{k_1}^{t_1}\right]\right\| \leq L_2\rho^{(t_2-t_1)K+k_2-k_1},\\
&\left\|\mathbb{E}_{t_1}\left[Z_i\left(O_{t_2,k_2}^{(i)}\right)\right]\right\| \leq L_2\rho^{(t_2-t_1)K+k_2}
\end{aligned}$$

对任意 $i \in [N]$、$0 \leq k_1$、$k_2 \leq K-1$ 和 $t_2 \geq t_1 \geq 0$ 均成立。

证明。我们有：

$$\begin{aligned}
\left\|\mathbb{E}\left[Z_i\left(O_{t_2,k_2}^{(i)}\right) \mid \mathcal{F}_{k_1}^{t_1}\right]\right\| &= \left\|\mathbb{E}\left[Z_i\left(O_{t_2,k_2}^{(i)}\right) \mid \mathcal{F}_{k_1}^{t_1}\right] - \mathbb{E}_{O_{t_2,k_2}^{(i)}\sim\pi^{(i)}}\left[Z_i\left(O_{t_2,k_2}^{(i)}\right) \mid \mathcal{F}_{k_1}^{t_1}\right]\right\|\\
&= \Bigg\| \sum_{s_{t_2,k_2}^{(i)}, s_{t_2+1,k_2+1}^{(i)}} \Big(\pi^{(i)}(s_{t_2,k_2}^{(i)})P(s_{t_2+1,k_2+1}^{(i)} \mid s_{t_2,k_2}^{(i)})\\
&\quad - P(s_{t_2,k_2}^{(i)} = \cdot \mid s_{t_1,k_1}^{(i)})P(s_{t_2+1,k_2+1}^{(i)} \mid s_{t_2,k_2}^{(i)})\Big) Z_i(O_{t_2,k_2}^{(i)})\Bigg\|\\
&\leq \sum_{s_{t_2,k_2}^{(i)}} \left|\pi^{(i)}(s_{t_2,k_2}^{(i)}) - P(s_{t_2,k_2}^{(i)} = \cdot \mid s_{t_1,k_1}^{(i)})\right| \left\|Z_i(O_{t_2,k_2}^{(i)})\right\|\\
&\overset{(a)}{\leq} \sum_{s_{t_2,k_2}^{(i)}} \left|\pi^{(i)}(s_{t_2,k_2}^{(i)}) - P(s_{t_2,k_2}^{(i)} = \cdot \mid s_{t_1,k_1}^{(i)})\right| (R_{\max} + c_1 H)\\
&= 2(R_{\max} + c_1 H) d_{TV}\left(\mathbb{P}\left(s_{t_2,k_2}^{(i)} = \cdot \mid s_{t_1,k_1}^{(i)} = s\right), \pi^{(i)}\right)\\
&\leq 2(R_{\max} + c_1 H) m_i \rho_i^{(t_2-t_1)K+k_2-k_1}
\end{aligned}$$

其中 (a) 由引理 11 得出，最后一步根据假设 3得出。通过选取$L_2 \triangleq \max^{i\in[N]}\{2(R_{\max} + c_1 H)m_i\} = 2c_3 m$，即可完成证明。并采用相同的分析来界定：

$$\begin{aligned}
\left\|\bar{A}_i - \mathbb{E}\left[A_i\left(O_{t_2,k_2}^{(i)}\right) \mid \mathcal{F}_{k_1}^{t_1}\right]\right\| &= \left\|\|\mathbb{E}\left[A_i\left(O_{t_2,k_2}^{(i)}\right) \mid \mathcal{F}_{k_1}^{t_1}\right] - \mathbb{E}_{O_{t_2,k_2}^{(i)}\sim\pi^{(i)}}\left[A_i\left(O_{t_2,k_2}^{(i)}\right) \mid \mathcal{F}_{k_1}^{t_1}\right]\right\|\\
&= \Bigg\| \sum_{s_{t_2,k_2}^{(i)}, s_{t_2+1,k_2+1}^{(i)}} \Big(\pi^{(i)}(s_{t_2,k_2}^{(i)})P(s_{t_2+1,k_2+1}^{(i)} \mid s_{t_2,k_2}^{(i)})\\
&\quad - P(s_{t_2,k_2}^{(i)} = \cdot \mid s_{t_1,k_1}^{(i)})P(s_{t_2+1,k_2+1}^{(i)} \mid s_{t_2,k_2}^{(i)})\Big) A_i(O_{t_2,k_2}^{(i)})\Bigg\|\\
&\leq \sum_{s_{t_2,k_2}^{(i)}} \left|\pi^{(i)}(s_{t_2,k_2}^{(i)}) - P(s_{t_2,k_2}^{(i)} = \cdot \mid s_{t_1,k_1}^{(i)})\right| \left\|A_i(O_{t_2,k_2}^{(i)})\right\|
\end{aligned}$$

$$\overset{(b)}{\leq} 2c_1 d_{TV}\left(\mathbb{P}\left(s^{(i)}_{t_2,k_2}=\cdot \mid s^{(i)}_{t_1,k_1}=s\right),\pi^{(i)}\right)$$
$$\leq 2c_1 m_i \rho_i^{(t_2-t_1)K+k_2-k_1}$$

通过选取 $L_1 \triangleq \max^{i\in[N]}\{2c_1 m_i\} = 2c_1 m$. 即可完成证明。我们采用相同的推理来证明其余三个不等式。

□

I.2.2 方差缩减

现在我们可以介绍马尔可夫设定下的方差缩减引理。下面的引理给出了独立同分布设定下方差缩减引理的一个对应形式。7 基于轨迹在各智能体之间相互独立这一假设，不难理解，$(1/NK)\sum_{i=1}^{N}\sum_{k=0}^{K-1} A_i(O^{(i)}_{t\,k,})$ 和 $(1/NK)\sum_{i=1}^{N}\sum_{k=0}^{K-1} b_i(O^{(i)}_{t\,k,})$ 的方差可以按智能体数量 N 缩减。然而，方差能否按 K （局部迭代的数量）缩减则并不明显，因为每个智能体 $O^{(i)}_{t\,k_1,}$ 的观测与 $O^{(i)}_{t\,,\,k_2}$ 在不同局部步数 k_1，k_2 上存在相关性。由于马尔可夫链具有几何混合性质，经过混合时间后，$O^{(i)}_{t\,,\,k_1}$ 与 $O^{(i)}_{t\,,\,k_2}$ 之间的相关性将按几何速率衰减。基于这一事实，我们证明，$(1/NK)\sum_{i=1}^{N}\sum_{k=0}^{K-1} A_i(O^{(i)}_{t\,k,})$ 和 $(1/NK)\sum_{i=1}^{N}\sum_{k=0}^{K-1} b_i(O^{(i)}_{t\,k,})$ 的方差会按 NK 缩减，同时还会产生一个依赖于混合时间 τ 的高阶加性项。形式化表述如下：

引理13。（马尔可夫设置下的方差缩减）对于任意 $0<\tau<t$，存在 $d_1, d_2>0$，使得：

$$\mathbb{E}_{t-\tau}\left[\left\|\frac{1}{NK}\sum_{i=1}^{N}\sum_{k=0}^{K-1}\left[A_i(O^{(i)}_{t,k})-\bar{A}_i\right]\right\|\right] \leq \frac{d_1}{\sqrt{NK}}+2L_1\rho^{\tau K}, \tag{34}$$

$$\mathbb{E}_{t-\tau}\left[\left\|\frac{1}{NK}\sum_{i=1}^{N}\sum_{k=0}^{K-1}\left[A_i(O^{(i)}_{t,k})-\bar{A}_i\right]\right\|^2\right] \leq \frac{d_1^2}{NK}+4L_1^2\rho^{2\tau K}, \quad and \tag{35}$$

$$\mathbb{E}_{t-\tau}\left[\left\|\frac{1}{NK}\sum_{i=1}^{N}\sum_{k=0}^{K-1} Z_i(O^{(i)}_{t,k})\right\|\right] \leq \frac{d_2}{\sqrt{NK}}+2L_2\rho^{\tau K}, \tag{36}$$

$$\mathbb{E}_{t-\tau}\left[\left\|\frac{1}{NK}\sum_{i=1}^{N}\sum_{k=0}^{K-1} Z_i(O^{(i)}_{t,k})\right\|^2\right] \leq \frac{d_2^2}{NK}+4L_2^2\rho^{2\tau K}, \tag{37}$$

where $d_1 \triangleq \sqrt{(c_1+c_2)^2+\frac{2(c_1+c_2)L_1\rho}{1-\rho}}$ *and* $d_2 \triangleq \sqrt{c_3^2+\frac{2c_3L_2\rho}{1-\rho}}$.

证明。

$$\mathbb{E}_{t-\tau}\left[\left\|\frac{1}{NK}\sum_{i=1}^{N}\sum_{k=0}^{K-1} Z_i(O^{(i)}_{t,k})\right\|\right]$$
$$= \mathbb{E}_{t-\tau}\left[\sqrt{\left(\frac{1}{NK}\sum_{i=1}^{N}\sum_{k=0}^{K-1} Z_i(O^{(i)}_{t,k})\right)^{\top}\left(\frac{1}{NK}\sum_{i=1}^{N}\sum_{k=0}^{K-1} Z_i(O^{(i)}_{t,k})\right)}\right]$$
$$\leq \sqrt{\mathbb{E}_{t-\tau}\left[\left(\frac{1}{NK}\sum_{i=1}^{N}\sum_{k=0}^{K-1} Z_i(O^{(i)}_{t,k})\right)^{\top}\left(\frac{1}{NK}\sum_{i=1}^{N}\sum_{k=0}^{K-1} Z_i(O^{(i)}_{t,k})\right)\right]}$$

（平方根函数的凹性和詹森不等式）

$$= \left\{ \mathbb{E}_{t-\tau} \left[\frac{1}{N^2K^2} \sum_{i=1}^{N} \sum_{k=0}^{K-1} Z_i(O_{t,k}^{(i)})^\top Z_i(O_{t,k}^{(i)}) + \underbrace{\frac{2}{N^2K^2} \sum_{i=1}^{N} \sum_{k<l} Z_i(O_{t,k}^{(i)})^\top Z_i(O_{t,l}^{(i)})}_{T_1} \right. \right.$$

$$\left. \left. + \underbrace{\frac{2}{N^2K^2} \sum_{i<j} \sum_{k=0}^{K-1} Z_i(O_{t,k}^{(i)})^\top Z_j(O_{t,k}^{(j)})}_{T_2} + \underbrace{\frac{2}{N^2K^2} \sum_{i<j} \sum_{k<l} Z_i(O_{t,k}^{(i)})^\top Z_j(O_{t,l}^{(j)})}_{T_3} \right] \right\}^{\frac{1}{2}} \tag{38}$$

其中， T_1 还可以进一步界定为:

$$\begin{aligned}
\mathbb{E}_{t-\tau}[T_1] &= \mathbb{E}_{t-\tau} \left[\frac{2}{N^2K^2} \sum_{i=1}^{N} \sum_{k<l} Z_i(O_{t,k}^{(i)})^\top Z_i(O_{t,l}^{(i)}) \right] \\
&= \mathbb{E}_{t-\tau} \left[\frac{2}{N^2K^2} \sum_{i=1}^{N} \sum_{k<l} Z_i(O_{t,k}^{(i)})^\top \mathbb{E}\left[Z_i(O_{t,l}^{(i)}) \mid \mathcal{F}_k^t \right] \right] \\
&\le \mathbb{E}_{t-\tau} \left[\frac{2}{N^2K^2} \sum_{i=1}^{N} \sum_{k<l} \left\| Z_i(O_{t,k}^{(i)}) \right\| \left\| \mathbb{E}\left[Z_i(O_{t,l}^{(i)}) \mid \mathcal{F}_k^t \right] \right\| \right] \\
&\qquad \text{(Cauchy–Schwarz inequality)} \\
&\le \mathbb{E}_{t-\tau} \left[\frac{2}{N^2K^2} \sum_{i=1}^{N} \sum_{k<l} c_3 L_2 \rho^{(l-k)} \right] \quad \text{(Lemma 11 and 12)} \\
&\le \mathbb{E}_{t-\tau} \left[\frac{2}{N^2K^2} \sum_{i=1}^{N} \sum_{k=0}^{K-1} \sum_{m=1}^{\infty} c_3 L_2 \rho^{m} \right] \\
&= \frac{2c_3 L_2 NK}{N^2K^2} \frac{\rho}{1-\rho} = \frac{2c_3 L_2 \rho}{NK(1-\rho)}.
\end{aligned}$$

并且， T_2 可以界定为:

$$\begin{aligned}
\mathbb{E}_{t-\tau}[T_2] &= \frac{2}{N^2K^2} \sum_{i<j} \sum_{k=0}^{K-1} \mathbb{E}_{t-\tau}\left[Z_i(O_{t,k}^{(i)}) \right]^\top \mathbb{E}_{t-\tau}\left[Z_j(O_{t,k}^{(j)}) \right] \ (O_{t,k}^{(i)} \text{ and } O_{t,k}^{(j)} \text{ are independent}) \\
&\le \frac{2}{N^2K^2} \sum_{i<j} \sum_{k=0}^{K-1} L_2^2 \rho^{2\tau K + 2k} \quad \text{(Lemma 12)} \\
&\le \frac{2}{K} L_2^2 \rho^{2\tau K}.
\end{aligned}$$

同时， T_3 可以界定为:

$$\begin{aligned}
\mathbb{E}_{t-\tau}[T_3] &= \frac{2}{N^2K^2} \sum_{i<j} \sum_{k<l} \mathbb{E}_{t-\tau}\left[Z_i(O_{t,k}^{(i)}) \right]^\top \mathbb{E}_{t-\tau}\left[Z_j(O_{t,l}^{(j)}) \right] \ (O_{t,k}^{(i)} \text{ and } O_{t,l}^{(j)} \text{ are independent}) \\
&\le \frac{2}{N^2K^2} \sum_{i<j} \sum_{k<l} L_2^2 \rho^{2\tau K + k + l} \quad \text{(Lemma 12)} \\
&\le 2L_2^2 \rho^{2\tau K}
\end{aligned}$$

将 T_1、 T_2 和 T_3 的上界代入方程（38），得:

$$\mathbb{E}_{t-\tau} \left[\left\| \frac{1}{NK} \sum_{i=1}^{N} \sum_{k=0}^{K-1} Z_i(O_{t,k}^{(i)}) \right\| \right]$$

$$
\begin{aligned}
&\leq \sqrt{\frac{1}{N^2K^2}\sum_{i=1}^{N}\sum_{k=0}^{K-1}\mathbb{E}_{t-\tau}\left[Z_i(O_{t,k}^{(i)})^\top Z_i(O_{t,k}^{(i)})\right] + \frac{2c_3L_2\rho}{NK(1-\rho)} + \frac{2}{K}L_2^2\rho^{2\tau K} + 2L_2^2\rho^{2\tau K}} \\
&\overset{(a)}{\leq} \sqrt{\frac{NK}{N^2K^2}c_3^2 + \frac{2c_3L\rho}{NK(1-\rho)} + \frac{2}{K}L_2^2\rho^{2\tau K} + 2L_2^2\rho^{2\tau K}} \\
&\leq \sqrt{\frac{1}{NK}\left(c_3^2 + \frac{2c_3L_2\rho}{1-\rho}\right) + 4L_2^2\rho^{2\tau K}} \quad (K \geq 1) \\
&\leq \sqrt{\frac{1}{NK}\left(c_3^2 + \frac{2c_3L_2\rho}{1-\rho}\right)} + \sqrt{4L_2^2\rho^{2\tau K}} = \sqrt{\frac{1}{NK}\left(c_3^2 + \frac{2c_3L_2\rho}{1-\rho}\right)} + 2L_2\rho^{\tau K}.
\end{aligned}
$$

其中 (a) 用到了 $\left\|Z_i\left(O_{t,k}^{(i)}\right)\right\| \leq c_3$ 这一事实，该事实在引理 11中提及。其他不等式的证明采用相同的推理过程。

□

I.2.3 界定 $\mathbb{E}\left[\left\|\bar{\theta}_t - \bar{\theta}_{t-\tau}\right\|^2\right]$

引理14. （界定 $\|\theta_t - \theta_{t-\tau}\|^2$）考虑 $\tau = \lceil \tau^{\text{mix}}(\alpha_T^2)_K \rceil$，并选择有效步长

$$
\alpha \leq \min\left\{\frac{1}{30c_4(\tau+1)}, \frac{1}{96c_4^2\tau}, 1\right\}
$$

其中 $c_4 = 3c_1$。对于任意 $t \geq 2\tau$，我们有如下界：

$$
\begin{aligned}
\mathbb{E}_{t-2\tau}\left[\left\|\bar{\theta}_t - \bar{\theta}_{t-\tau}\right\|^2\right] \leq\ & 8\alpha^2\tau^2c_4^2\mathbb{E}_{t-2\tau}\left[\left\|\bar{\theta}_t - \theta^*\right\|^2\right] + 14\alpha^2\tau^2\frac{d_2^2}{NK} + \frac{52L_2^2\alpha^4\tau}{1-\rho^2} \\
&+ 4\alpha^2c_4^2\tau\sum_{s=0}^{\tau}E_{t-2\tau}[\Delta_{t-s}] + 3200\alpha^2c_4^2c_1^2\tau^3\Gamma^2(\epsilon,\epsilon_1) + 4\alpha^2c_1^2\tau^2\Gamma^2(\epsilon,\epsilon_1). \qquad (39)
\end{aligned}
$$

证明。对于任意 $l \geq 2\tau$，我们有

$$
\begin{aligned}
&\left\|\bar{\theta}_{l+1} - \bar{\theta}_l\right\|^2 \\
&= \left\|\Pi_{2,\mathcal{H}}\left(\bar{\theta}_l + \frac{\alpha}{NK}\sum_{i=1}^{N}\sum_{k=0}^{K-1}g_i(\theta_{l,k}^{(i)})\right) - \bar{\theta}_l\right\|^2 \\
&\leq \left\|\bar{\theta}_l + \frac{\alpha}{NK}\sum_{i=1}^{N}\sum_{k=0}^{K-1}g_i(\theta_{l,k}^{(i)}) - \bar{\theta}_l\right\|^2 \\
&= \alpha^2\left\|\frac{1}{NK}\sum_{i=1}^{N}\sum_{k=0}^{K-1}\left[-A_i(O_{l,k}^{(i)})\left(\theta_{l,k}^{(i)} - \theta_i^*\right) + Z_i(O_{l,k}^{(i)})\right]\right\|^2 \\
&\leq 2\alpha^2\left\|\frac{1}{NK}\sum_{i=1}^{N}\sum_{k=0}^{K-1}\left[-A_i(O_{l,k}^{(i)})\left(\theta_{l,k}^{(i)} - \theta^*\right) + Z_i(O_{l,k}^{(i)})\right]\right\|^2 \\
&\qquad + 2\alpha^2\left\|\frac{1}{NK}\sum_{i=1}^{N}\sum_{k=0}^{K-1}\left[-A_i(O_{l,k}^{(i)})\left(\theta^* - \theta_i^*\right)\right]\right\|^2 \\
&\overset{(a)}{\leq} 2\alpha^2\left\|\frac{1}{NK}\sum_{i=1}^{N}\sum_{k=0}^{K-1}\left[-A_i(O_{l,k}^{(i)})\left(\theta_{l,k}^{(i)} - \theta^*\right) + Z_i(O_{l,k}^{(i)})\right]\right\|^2 + 2\alpha^2c_1^2\Gamma^2(\epsilon,\epsilon_1)
\end{aligned}
$$

$$
\begin{aligned}
&= 6\alpha^2\Big\|\frac{1}{NK}\sum_{i=1}^{N}\sum_{k=0}^{K-1}A_i(O_{l,k}^{(i)})\Big(\theta_{l,k}^{(i)}-\bar{\theta}_l\Big)\Big\|^2+6\alpha^2\Big\|\frac{1}{NK}\sum_{i=1}^{N}\sum_{k=0}^{K-1}A_i(O_{l,k}^{(i)})\Big(\bar{\theta}_l-\theta^*\Big)\Big\|^2\\
&+6\alpha^2\Big\|\frac{1}{NK}\sum_{i=1}^{N}\sum_{k=0}^{K-1}Z_i(O_{l,k}^{(i)})\Big\|^2+2\alpha^2c_1^2\Gamma^2(\epsilon,\epsilon_1)\\
&\le 6\alpha^2\left(\frac{c_1}{NK}\sum_{i=1}^{N}\sum_{k=0}^{K-1}\Big\|\theta_{l,k}^{(i)}-\bar{\theta}_l\Big\|\right)^2\\
&\qquad+6\alpha^2c_1^2\Big\|\bar{\theta}_l-\theta^*\Big\|^2+6\alpha^2\Big\|\frac{1}{NK}\sum_{i=1}^{N}\sum_{k=0}^{K-1}Z_i(O_{l,k}^{(i)})\Big\|^2+2\alpha^2c_1^2\Gamma^2(\epsilon,\epsilon_1),
\end{aligned}
\tag{40}
$$

其中 (a) 来自定理 1 中不动点距离的上界，以及 $\left\|A_i\left(O_{t,k}^{(i)}\right)\right\|\le c_1$引理 11中的事实。对上述不等式两边开平方，得到:

$$
\begin{aligned}
&\Big\|\bar{\theta}_{l+1}-\bar{\theta}_l\Big\|\\
&\le 3\sqrt{\alpha^2\left(\frac{c_1}{NK}\sum_{i=1}^{N}\sum_{k=0}^{K-1}\Big\|\theta_{l,k}^{(i)}-\bar{\theta}_l\Big\|\right)^2}+3\sqrt{\alpha^2c_1^2\Big\|\bar{\theta}_l-\theta^*\Big\|^2}\\
&+3\sqrt{\alpha^2\Big\|\frac{1}{NK}\sum_{i=1}^{N}\sum_{k=0}^{K-1}Z_i(O_{l,k}^{(i)})\Big\|^2}+\sqrt{2\alpha^2c_1^2\Gamma^2(\epsilon,\epsilon_1)}\\
&\le \frac{3\alpha c_1}{NK}\sum_{i=1}^{N}\sum_{k=0}^{K-1}\Big\|\theta_{l,k}^{(i)}-\bar{\theta}_l\Big\|+3\alpha c_1\Big\|\bar{\theta}_l-\theta^*\Big\|+3\alpha\Big\|\frac{1}{NK}\sum_{i=1}^{N}\sum_{k=0}^{K-1}Z_i(O_{l,k}^{(i)})\Big\|+2\alpha c_1\Gamma(\epsilon,\epsilon_1).
\end{aligned}
\tag{41}
$$

利用 $\Big\|\bar{\theta}_{l+1}-\theta^*\Big\|\le\Big\|\bar{\theta}_l-\theta^*\Big\|+\Big\|\bar{\theta}_{l+1}-\bar{\theta}_l\Big\|$这一事实，我们得到:

$$
\begin{aligned}
\Big\|\bar{\theta}_{l+1}-\theta^*\Big\|\le(1+3\alpha c_1)\Big\|\bar{\theta}_l-\theta^*\Big\|+\frac{3\alpha c_1}{NK}\sum_{i=1}^{N}\sum_{k=0}^{K-1}\Big\|\theta_{l,k}^{(i)}-\bar{\theta}_l\Big\|\\
+3\alpha\Big\|\frac{1}{NK}\sum_{i=1}^{N}\sum_{k=0}^{K-1}Z_i(O_{l,k}^{(i)})\Big\|+2\alpha c_1\Gamma(\epsilon,\epsilon_1).
\end{aligned}
\tag{42}
$$

为简单起见，我们定义 $c_4\triangleq 3c_1$ 和 $\delta_l\triangleq\frac{1}{NK}\sum_{i=1}^{N}\sum_{k=0}^{K-1}\Big\|\theta_{l,k}^{(i)}-\bar{\theta}_l\Big\|$。对方程（42）两边平方，可得:

$$
\begin{aligned}
&\Big\|\bar{\theta}_{l+1}-\theta^*\Big\|^2\\
&\le(1+\alpha c_4)^2\Big\|\bar{\theta}_l-\theta^*\Big\|^2+\alpha^2c_4^2\delta_l^2+9\alpha^2\Big\|\frac{1}{NK}\sum_{i=1}^{N}\sum_{k=0}^{K-1}Z_i(O_{l,k}^{(i)})\Big\|^2+4\alpha^2c_1^2\Gamma^2(\epsilon,\epsilon_1)\\
&\underbrace{+6\alpha(1+\alpha c_4)\Big\|\bar{\theta}_l-\theta^*\Big\|\Big\|\frac{1}{NK}\sum_{i=1}^{N}\sum_{k=0}^{K-1}Z_i(O_{l,k}^{(i)})\Big\|}_{H_1}+\underbrace{2\alpha c_4(1+\alpha c_4)\Big\|\bar{\theta}_l-\theta^*\Big\|\delta_l}_{H_2}\\
&\underbrace{+6\alpha^2c_4\delta_l\Big\|\frac{1}{NK}\sum_{i=1}^{N}\sum_{k=0}^{K-1}Z_i(O_{l,k}^{(i)})\Big\|}_{H_3}+\underbrace{4\alpha^2c_1c_4\delta_l\Gamma(\epsilon,\epsilon_1)}_{H_4}
\end{aligned}
$$

$$+\underbrace{4\alpha c_1(1+\alpha c_4)\left\|\bar{\theta}_l-\theta^*\right\|\Gamma(\epsilon,\epsilon_1)}_{H_5}+\underbrace{12\alpha^2 c_1\left\|\frac{1}{NK}\sum_{i=1}^{N}\sum_{k=0}^{K-1}Z_i(O_{l,k}^{(i)})\right\|\Gamma(\epsilon,\epsilon_1)}_{H_6}. \tag{43}$$

我们可以进一步将 H_1 界定为:

$$\begin{aligned}
H_1 &= 6\alpha(1+\alpha c_4)\left\|\bar{\theta}_l-\theta^*\right\|\left\|\frac{1}{NK}\sum_{i=1}^{N}\sum_{k=0}^{K-1}Z_i(O_{l,k}^{(i)})\right\| \\
&= 2\sqrt{3\alpha(1+\alpha c_4)}\left\|\bar{\theta}_l-\theta^*\right\|\cdot\sqrt{3\alpha(1+\alpha c_4)}\left\|\frac{1}{NK}\sum_{i=1}^{N}\sum_{k=0}^{K-1}Z_i(O_{l,k}^{(i)})\right\| \\
&\le 3\alpha(1+\alpha c_4)\left\|\bar{\theta}_l-\theta^*\right\|^2+3\alpha(1+\alpha c_4)\left\|\frac{1}{NK}\sum_{i=1}^{N}\sum_{k=0}^{K-1}Z_i(O_{l,k}^{(i)})\right\|^2 \\
&\le 6\alpha\left\|\bar{\theta}_l-\theta^*\right\|^2+6\alpha\left\|\frac{1}{NK}\sum_{i=1}^{N}\sum_{k=0}^{K-1}Z_i(O_{l,k}^{(i)})\right\|^2.
\end{aligned} \tag{44}$$

其中，最后一个不等式用到了事实 $1+\alpha c_4\le 2$。类似地，我们可以对 H_2 给出上界:

$$H_2=2\alpha c_4(1+\alpha c_4)\left\|\bar{\theta}_l-\theta^*\right\|\delta_l\le 2\alpha\left\|\bar{\theta}_l-\theta^*\right\|^2+2\alpha c_4^2\delta_l^2. \tag{45}$$

并且我们可以对 H_3 给出上界:

$$H_3=6\alpha^2c_4\delta_l\left\|\frac{1}{NK}\sum_{i=1}^{N}\sum_{k=0}^{K-1}Z_i(O_{l,k}^{(i)})\right\|\le 3\alpha^2\left\|\frac{1}{NK}\sum_{i=1}^{N}\sum_{k=0}^{K-1}Z_i(O_{l,k}^{(i)})\right\|^2+3\alpha^2c_4^2\delta_l^2. \tag{46}$$

对于 H_4, H_5, H_6，我们有:

$$H_4=4\alpha^2c_1c_4\delta_l\Gamma(\epsilon,\epsilon_1)\le 2\alpha^2c_4^2\delta_l^2+2\alpha^2c_1^2\Gamma^2(\epsilon,\epsilon_1),$$

$$H_5=4\alpha c_1(1+\alpha c_4)\left\|\bar{\theta}_l-\theta^*\right\|\Gamma(\epsilon,\epsilon_1)\le 4\alpha\left\|\bar{\theta}_l-\theta^*\right\|^2+4\alpha c_1^2\Gamma^2(\epsilon,\epsilon_1),$$

$$H_6=12\alpha^2c_1\left\|\frac{1}{NK}\sum_{i=1}^{N}\sum_{k=0}^{K-1}Z_i(O_{l,k}^{(i)})\right\|\Gamma(\epsilon,\epsilon_1)\le 6\alpha^2\left\|\frac{1}{NK}\sum_{i=1}^{N}\sum_{k=0}^{K-1}Z_i(O_{l,k}^{(i)})\right\|^2+6\alpha^2c_1^2\Gamma^2(\epsilon,\epsilon_1),$$

将 H_1, H_2, ..., H_6的上界代入方程（43），并注意到$(1+\alpha c_4)^2\le 1+3\alpha c_4$，因为$\alpha c_4\le 1$，我们有:

$$\begin{aligned}
&\left\|\bar{\theta}_{l+1}-\theta^*\right\|^2 \\
&\le(1+\alpha(3c_4+12))\left\|\bar{\theta}_l-\theta^*\right\|^2+(6\alpha^2+2\alpha)c_4^2\delta_l^2 \\
&+(18\alpha^2+6\alpha)\left\|\frac{1}{NK}\sum_{i=1}^{N}\sum_{k=0}^{K-1}Z_i(O_{l,k}^{(i)})\right\|^2+(12\alpha^2+4\alpha)c_1^2\Gamma^2(\epsilon,\epsilon_1) \\
&\le(1+\alpha h_1)\left\|\bar{\theta}_l-\theta^*\right\|^2+8\alpha c_4^2\delta_l^2+24\alpha\left\|\frac{1}{NK}\sum_{i=1}^{N}\sum_{k=0}^{K-1}Z_i(O_{l,k}^{(i)})\right\|^2+16\alpha c_1^2\Gamma^2(\epsilon,\epsilon_1),
\end{aligned} \tag{47}$$

其中，为简洁起见，我们记为 $h_1\triangleq 3c_4+12$。对于任意 $t-\tau\le l\le t$，在上述不等式两边以 $\mathcal{F}_{t-2\tau}$ 为条件，可得:

$$\mathbb{E}_{t-2\tau}\left\|\bar{\theta}_{l+1}-\theta^*\right\|^2$$

$$
\begin{aligned}
&\leq (1+\alpha h_1)\,\mathbb{E}_{t-2\tau}\left\|\bar{\theta}_l-\theta^*\right\|^2+24\alpha\mathbb{E}_{t-2\tau}\left\|\frac{1}{NK}\sum_{i=1}^{N}\sum_{k=0}^{K-1}Z_i(O_{l,k}^{(i)})\right\|^2\\
&\quad+8\alpha c_4^2\mathbb{E}_{t-2\tau}\left[\delta_l^2\right]+\alpha M_3(\epsilon,\epsilon_1)\\
&\leq (1+\alpha h_1)\,\mathbb{E}_{t-2\tau}\left\|\bar{\theta}_l-\theta^*\right\|^2+24\alpha\left[\frac{d_2^2}{NK}+4L_2^2\rho^{2(l-t+2\tau)K}\right]\quad\text{(Lemma 13)}\\
&+8\alpha c_4^2\mathbb{E}_{t-2\tau}\left[\delta_l^2\right]+\alpha M_3(\epsilon,\epsilon_1)\\
&\overset{(a)}{\leq}(1+\alpha h_1)\,\mathbb{E}_{t-2\tau}\left\|\bar{\theta}_l-\theta^*\right\|^2+24\alpha\left[\frac{d_2^2}{NK}+4L_2^2\alpha^2\rho^{2(l-t+\tau)K}\right]\\
&+8\alpha c_4^2\mathbb{E}_{t-2\tau}\left[\delta_l^2\right]+\alpha M_3(\epsilon,\epsilon_1)\\
&\leq (1+\alpha h_1)\,\mathbb{E}_{t-2\tau}\left\|\bar{\theta}_l-\theta^*\right\|^2+\alpha c_t(l)+8\alpha c_4^2\mathbb{E}_{t-2\tau}\left[\delta_l^2\right]+\alpha M_3(\epsilon,\epsilon_1),
\end{aligned}
\tag{48}
$$

其中，为简洁起见，我们记 $M_3(\epsilon,\epsilon_1)\triangleq 16c_1^2\Gamma^2(\epsilon,\epsilon_1)$ 和 $c_t(l)=24\left[\frac{d_2^2}{NK}+4L_2^2\alpha^2\rho^{2(l-t+\tau)K}\right]$。不等式$(a)$是由 $\rho^{2\tau K}\leq\alpha_T^4\leq\alpha_t^2$ 得出的。接下来的步骤中，我们尝试将 $\mathbb{E}_{t-2\tau}\left\|\bar{\theta}_{l+1}-\theta^*\right\|^2$ 映射为 $\mathbb{E}_{t-2\tau}\left\|\bar{\theta}_{t-\tau}-\theta^*\right\|^2$，对于任意 $t-\tau\leq l\leq t$。递归应用方程（48），可得:

$$
\begin{aligned}
&\mathbb{E}_{t-2\tau}\left\|\bar{\theta}_{l+1}-\theta^*\right\|^2\\
&\leq (1+\alpha h_1)^{l+1-t+\tau}\,\mathbb{E}_{t-2\tau}\left\|\bar{\theta}_{t-\tau}-\theta^*\right\|^2+\alpha\sum_{k=t-\tau}^{l}(1+\alpha h_1)^{l-k}\left(c_t(k)+M_3(\epsilon,\epsilon_1)\right)\\
&+8\alpha c_4^2\mathbb{E}_{t-2\tau}\left[\sum_{k=t-\tau}^{l}(1+\alpha h_1)^{l-k}\,\delta_k^2\right]\\
&\overset{(b)}{\leq}(1+\alpha h_1)^{\tau+1}\,\mathbb{E}_{t-2\tau}\left\|\bar{\theta}_{t-\tau}-\theta^*\right\|^2+\underbrace{\alpha\sum_{k=t-\tau}^{t}(1+\alpha h_1)^{l-k}\left(c_t(k)+M_3(\epsilon,\epsilon_1)\right)}_{H_7}\\
&+8\alpha c_4^2\mathbb{E}_{t-2\tau}\underbrace{\left[\sum_{k=t-\tau}^{t}(1+\alpha h_1)^{l-k}\,\delta_k^2\right]}_{H_8}
\end{aligned}
\tag{49}
$$

其中，(b) 是由 $l\leq t$ 导致的。对于 H_7，我们有:

$$
\begin{aligned}
H_7&\leq\sum_{k=t-\tau}^{t}(1+\alpha h_1)^{t-k}\left(c_t(k)+M_3(\epsilon,\epsilon_1)\right)\quad(l\leq t)\\
&=\sum_{k'=0}^{\tau}(1+\alpha h_1)^{\tau-k'}\left(c_t(k'+t-\tau)+M_3(\epsilon,\epsilon_1)\right)\\
&\quad(\text{ changing index } k \text{ into } k' \text{ with } k'=k+\tau-t)\\
&\leq 24\sum_{k'=0}^{\tau}(1+\alpha h_1)^{\tau-k'}\left[\frac{d_2^2}{NK}+4L_2^2\alpha^2\rho^{2k'K}+M_3(\epsilon,\epsilon_1)\right]\\
&\quad(\text{Substituting the definition of } c_t(k') \text{ inside})\\
&=24\left[\left(\frac{d_2^2}{NK}+M_3(\epsilon,\epsilon_1)\right)\frac{(1+\alpha h_1)^{\tau+1}-1}{\alpha h_1}+4L_2^2\alpha^2(1+\alpha h_1)^{\tau}\sum_{k'=0}^{\tau}\left(\frac{\rho^{2K}}{1+\alpha h_1}\right)^{k'}\right]\\
&\leq 24\left[\left(\frac{d_2^2}{NK}+M_3(\epsilon,\epsilon_1)\right)\frac{(1+\alpha h_1)^{\tau+1}-1}{\alpha h_1}+4L_2^2\alpha^2(1+\alpha h_1)^{\tau}\sum_{k'=0}^{\tau}\rho^{2k'K}\right](1+\alpha h_1\geq 1)
\end{aligned}
$$

$$\leq 24\left[\left(\frac{d_2^2}{NK}+M_3(\epsilon,\epsilon_1)\right)\frac{(1+\alpha h_1)^{\tau+1}-1}{\alpha h_1}+4L_2^2\alpha^2\left(1+\alpha h_1\right)^{\tau}\frac{1}{1-\rho^2}\right].$$

这里我们沿用（Khodadadian 等人，2022年）中的分析。注意，对于 $x \leq \frac{2}{\tau}$，我们有 $(1+x)^{\tau+1} \leq 1+2x(\tau+1)$。如果 $\alpha \leq \frac{1}{4h_1\tau} \leq \frac{1}{h_1\tau^2}$ 且 $\alpha \leq \frac{1}{2h_1(\tau+1)}$，则有 $(1+\alpha h_1)^{\tau+1} \leq 1+2\alpha h_1(\tau+1) \leq 2$ 以及 $(1+\alpha h_1)^{\tau} \leq 1+2\alpha h_1\tau \leq 1+1/2 \leq 2$。因此，我们有

$$H_7 \leq 24\left[\left(\frac{d_2^2}{NK}+M_3(\epsilon,\epsilon_1)\right)2(\tau+1)+\frac{8L_2^2\alpha^2}{1-\rho^2}\right].$$

我们采用类似的分析来推导 H_8 的上界：

$$H_8=\sum_{k=0}^{\tau}(1+\alpha h_1)^{\tau-k}\delta_{t-\tau+k}^2 \leq \sum_{k=0}^{\tau}(1+\alpha h_1)^{\tau}\delta_{t-\tau+k}^2 \leq \sum_{k=0}^{\tau}(1+2\alpha h_1\tau)\delta_{t-\tau+k}^2 \leq 2\sum_{k=0}^{\tau}\delta_{t-k}^2.$$

将 H_7 和 H_8 的上界代入方程（49），可得：

$$\begin{aligned}\mathbb{E}_{t-2\tau}\left\|\bar{\theta}_{l+1}-\theta^*\right\|^2 \leq 2\mathbb{E}_{t-2\tau}\left\|\bar{\theta}_{t-\tau}-\theta^*\right\|^2+24\alpha\left[\left(\frac{d_2^2}{NK}+M_3(\epsilon,\epsilon_1)\right)2(\tau+1)+\frac{8L_2^2\alpha^2}{1-\rho^2}\right]\\ +16\alpha c_4^2\sum_{k=0}^{\tau}\mathbb{E}_{t-2\tau}[\delta_{t-k}^2].\end{aligned}$$

接下来，很容易推导出 $\mathbb{E}_{t-2\tau}\left\|\bar{\theta}_l-\theta^*\right\|^2$ 的上界：

$$\begin{aligned}\mathbb{E}_{t-2\tau}\left\|\bar{\theta}_l-\theta^*\right\|^2 \leq 2\mathbb{E}_{t-2\tau}\left\|\bar{\theta}_{t-\tau}-\theta^*\right\|^2+24\alpha\left[\left(\frac{d_2^2}{NK}+M_3(\epsilon,\epsilon_1)\right)4\tau+\frac{8L_2^2\alpha^2}{1-\rho^2}\right]\\ +16\alpha c_4^2\sum_{k=0}^{\tau}\mathbb{E}_{t-2\tau}[\delta_{t-k}^2].\end{aligned} \tag{50}$$

此外，根据三角不等式，可得：

$$\begin{aligned}&\left\|\bar{\theta}_t-\bar{\theta}_{t-\tau}\right\|^2\\ &\leq\left(\sum_{s=t-\tau}^{t-1}\left\|\bar{\theta}_{s+1}-\bar{\theta}_s\right\|\right)^2 \leq \tau\sum_{s=t-\tau}^{t-1}\left\|\bar{\theta}_{s+1}-\bar{\theta}_s\right\|^2\\ &\leq\tau\sum_{s=t-\tau}^{t-1}\left[\alpha^2c_4^2\left\|\bar{\theta}_s-\theta^*\right\|^2+\alpha^2c_4^2\delta_s^2+6\alpha^2\left\|\frac{1}{NK}\sum_{i=1}^{N}\sum_{k=0}^{K-1}Z_i(O_{s,k}^{(i)})\right\|^2+2\alpha^2c_1^2\Gamma^2(\epsilon,\epsilon_1)\right]\end{aligned}$$

其中最后一个不等式 性是由方程 (40) 得出的，其中 $c_4=3c_1$。对两边取期望，可得 ：

$$\begin{aligned}&\mathbb{E}_{t-2\tau}\left\|\bar{\theta}_t-\bar{\theta}_{t-\tau}\right\|^2\\ &\leq\tau\sum_{s=t-\tau}^{t-1}\left[\alpha^2c_4^2\mathbb{E}_{t-2\tau}\left\|\bar{\theta}_s-\theta^*\right\|^2+\alpha^2c_4^2\delta_s^2\right.\\ &\quad\left.+6\alpha^2\mathbb{E}_{t-2\tau}\left\|\frac{1}{NK}\sum_{i=1}^{N}\sum_{k=0}^{K-1}Z_i(O_{s,k}^{(i)})\right\|^2+2\alpha^2c_1^2\Gamma^2(\epsilon,\epsilon_1)\right]\\ &\leq\tau\alpha^2c_4^2\sum_{s=t-\tau}^{t-1}\left[2\mathbb{E}_{t-2\tau}\left\|\bar{\theta}_{t-\tau}-\theta^*\right\|^2\right.\end{aligned}$$

$$
\begin{aligned}
&+ 24\alpha\left[\left(\frac{d_2^2}{NK}+M_3(\epsilon,\epsilon_1)\right)4\tau+\frac{8L_2^2\alpha^2}{1-\rho^2}\right]+16\alpha c_4^2\sum_{k=0}^{\tau}\mathbb{E}_{t-2\tau}[\delta_{t-k}^2]\Big] \text{ (Eq (50))}\\
&+6\alpha^2\tau\sum_{s=t-\tau}^{t-1}\left(\frac{d_2^2}{NK}+4L_2^2\rho^{2(s-t+2\tau)K}\right)\\
&+\alpha^2c_4^2\tau\sum_{s=t-\tau}^{t-1}\mathbb{E}_{t-2\tau}[\delta_s^2]+2\alpha^2c_1^2\tau^2\Gamma^2(\epsilon,\epsilon_1)\quad\text{(Lemma 13)}\\
&\overset{(a)}{\leq}\tau^2\alpha^2c_4^2\left[2\mathbb{E}_{t-2\tau}\left\|\bar{\theta}_{t-\tau}-\theta^*\right\|^2+96\left(\frac{d_2^2}{NK}\alpha\tau+\frac{2L_2^2\alpha^3}{1-\rho^2}\right)\right]+6\alpha^2\tau\left[\frac{d_2^2}{NK}\tau+\frac{4L_2^2\alpha^2}{1-\rho^{2K}}\right]\\
&+\alpha^2c_4^2\tau(1+16\alpha\tau c_4^2)\sum_{s=0}^{\tau}E_{t-2\tau}[\delta_{t-s}^2]+96\alpha^2c_4^2\tau^3M_3(\epsilon,\epsilon_1)+2\alpha^2c_1^2\tau^2\Gamma^2(\epsilon,\epsilon_1)\\
&\overset{(b)}{\leq}2\tau^2\alpha^2c_4^2\mathbb{E}_{t-2\tau}\left\|\bar{\theta}_{t-\tau}-\theta^*\right\|^2+\frac{d_2^2}{NK}\alpha^2\tau^2\left(96\alpha\tau c_4^2+6\right)+\frac{12L_2^2\alpha^4\tau}{1-\rho^2}\left(16\alpha c_4^2\tau+2\right)\\
&+\alpha^2c_4^2\tau(1+16\alpha\tau c_4^2)\sum_{s=0}^{\tau}E_{t-2\tau}[\Delta_{t-s}]+96\alpha^2c_4^2\tau^3M_3(\epsilon,\epsilon_1)+2\alpha^2c_1^2\tau^2\Gamma^2(\epsilon,\epsilon_1)
\end{aligned}\tag{51}
$$

其中我们用到了以下事实：$\rho^{2\tau K}\leq\alpha^2$ 对于 (a) 和 (b) 成立，并且对于所有$t\geq 0$，由詹森不等式可知 $\delta_t^2\leq\Delta_t$ 在最后一个不等式中成立。我们选择 α 使得 $96\alpha\tau c_4^2+6\leq 7$、$16\alpha c_4^2\tau+2\leq\frac{13}{6}$ 和 $1+16\alpha\tau c_4^2\leq 2$，当以下条件成立时，这一结论成立

$$
\alpha\leq\min\left\{\frac{1}{96\tau c_4^2},\frac{1}{96c_4^2\tau},\frac{1}{16\tau c_4^2},1\right\}.
$$

根据 $\|\bar{\theta}_{t-\tau}-\theta^*\|^2\leq 2\|\bar{\theta}_t-\bar{\theta}_{t-\tau}\|^2+2\|\bar{\theta}_t-\theta^*\|^2$ 这一事实以及对 α 的要求，我们有

$$
\begin{aligned}
2\alpha^2\tau^2c_4^2\mathbb{E}_{t-2\tau}\|\bar{\theta}_{t-\tau}-\theta^*\|^2&\leq 4\alpha^2\tau^2c_4^2\mathbb{E}_{t-2\tau}\|\bar{\theta}_t-\bar{\theta}_{t-\tau}\|^2+4\alpha^2\tau^2c_4^2\mathbb{E}_{t-2\tau}\|\bar{\theta}_t-\theta^*\|^2\\
&\leq 0.5\mathbb{E}_{t-2\tau}\|\bar{\theta}_t-\bar{\theta}_{t-\tau}\|^2+4\alpha^2\tau^2c_4^2\mathbb{E}_{t-2\tau}\|\bar{\theta}_t-\theta^*\|^2\quad(4\alpha^2\tau^2c_4^2\leq 0.5)\\
&\overset{(a)}{\leq}\tau^2\alpha^2c_4^2\mathbb{E}_{t-2\tau}\left\|\bar{\theta}_{t-\tau}-\theta^*\right\|^2+\frac{7d_2^2}{2NK}\alpha^2\tau^2+\frac{13L_2^2\alpha^4\tau}{(1-\rho^2)}\\
&+\alpha^2c_4^2\tau\sum_{s=0}^{\tau}E_{t-2\tau}[\Delta_{t-s}]+48\alpha^2c_4^2\tau^3M_3(\epsilon,\epsilon_1)+\alpha^2c_1^2\tau^2\Gamma^2(\epsilon,\epsilon_1)\\
&+4\alpha^2\tau^2c_4^2\mathbb{E}_{t-2\tau}\|\bar{\theta}_t-\theta^*\|^2
\end{aligned}\tag{52}
$$

其中，(a) 是由方程 (51) 以及 α 的取值所致。通过重新排列各项，将 $\tau^2\alpha^2c_4^2\mathbb{E}_{t-2\tau}\left\|\bar{\theta}_{t-\tau}-\theta^*\right\|^2$ 相关项合并，可得：

$$
\begin{aligned}
\alpha^2\tau^2c_4^2\mathbb{E}_{t-2\tau}\|\bar{\theta}_{t-\tau}-\theta^*\|^2&\leq\frac{7d_2^2}{2NK}\alpha^2\tau^2+\frac{13L_2^2\alpha^4\tau}{(1-\rho^2)}\\
&+\alpha^2c_4^2\tau\sum_{s=0}^{\tau}E_{t-2\tau}[\Delta_{t-s}]+48\alpha^2c_4^2\tau^3M_3(\epsilon,\epsilon_1)+\alpha^2c_1^2\tau^2\Gamma^2(\epsilon,\epsilon_1)\\
&+4\alpha^2\tau^2c_4^2\mathbb{E}_{t-2\tau}\|\bar{\theta}_t-\theta^*\|^2
\end{aligned}\tag{53}
$$

将此不等式代入方程 (51) 以及 $M_3(\epsilon,\ \epsilon_1)$ 的定义，即完成证明。注意，我们要求有效步长

$$
\alpha\leq\min\left\{\frac{1}{4h_1\tau},\frac{1}{2h_1(\tau+1)},\frac{1}{96c_4^2\tau},1\right\}
$$

在此证明中，当 $\alpha\leq\min\left\{\frac{1}{30c_4{}^1(\tau+1)}\ \frac{1}{96c_4^2\tau},\ 1\right\}$ 时成立，因为 $c_4=3c_1\geq 1$。

□

I.2.4 漂移项分析。

现在我们对漂移项作如下界定:

引理 15。 （有界客户端漂移）若 $\alpha_l \leq \frac{1}{2\sqrt{2}c_1 (K-1)}$，则漂移项满足

$$\mathbb{E}[\Delta_t] = \frac{1}{NK}\sum_{i=1}^{N}\sum_{k=0}^{K-1}\mathbb{E}\left\|\theta_{t,k}^{(i)} - \bar{\theta}_t\right\|^2 \leq \frac{4\alpha^2}{K\alpha_g^2}\left[c_3^2 + \frac{2c_3L_2\rho}{1-\rho} + 8c_1^2(K-1)H^2\right]. \tag{54}$$

证明。

$$\begin{aligned}
&\frac{1}{NK}\sum_{i=1}^{N}\sum_{k=0}^{K-1}\mathbb{E}\left\|\theta_{t,k}^{(i)} - \bar{\theta}_t\right\|^2 = \frac{1}{NK}\sum_{i=1}^{N}\sum_{k=0}^{K-1}\mathbb{E}\left\|\bar{\theta}_t + \alpha_l\sum_{s=0}^{k-1}g_i(\theta_{t,s}^{(i)}) - \bar{\theta}_t\right\|^2 \\
&= \alpha_l^2\frac{1}{NK}\sum_{i=1}^{N}\sum_{k=0}^{K-1}\mathbb{E}\left\|\sum_{s=0}^{k-1} -A_i(O_{t,s}^{(i)})\left(\theta_{t,s}^{(i)} - \theta_i^*\right) + Z_i(O_{t,s}^{(i)})\right\|^2 \\
&\leq 2\alpha_l^2\frac{1}{NK}\sum_{i=1}^{N}\sum_{k=0}^{K-1}\mathbb{E}\left\|\sum_{s=0}^{k-1} -A_i(O_{t,s}^{(i)})\left(\theta_{t,s}^{(i)} - \theta_i^*\right)\right\|^2 + 2\alpha_l^2\frac{1}{NK}\sum_{i=1}^{N}\sum_{k=0}^{K-1}\mathbb{E}\left\|\sum_{s=0}^{k-1}Z_i(O_{t,s}^{(i)})\right\|^2 \\
&\leq 2\alpha_l^2\frac{1}{NK}\sum_{i=1}^{N}\sum_{k=0}^{K-1}k\sum_{s=0}^{k-1}\mathbb{E}\left\|A_i(O_{t,s}^{(i)})\left(\theta_{t,s}^{(i)} - \theta_i^*\right)\right\|^2 + 2\alpha_l^2\frac{1}{NK}\sum_{i=1}^{N}\sum_{k=0}^{K-1}\sum_{s=0}^{k-1}\mathbb{E}\left\|Z_i(O_{t,s}^{(i)})\right\|^2 \\
&+ 2\alpha_l^2\frac{1}{NK}\sum_{i=1}^{N}\sum_{k=0}^{K-1}\sum_{\substack{s,s'=0\\ s\neq s'}}^{k-1}\mathbb{E}\left\langle Z_i(O_{t,s}^{(i)}), Z_i(O_{t,s'}^{(i)})\right\rangle \\
&\leq 2\alpha_l^2\frac{1}{NK}\sum_{i=1}^{N}\sum_{k=0}^{K-1}kc_1^2\sum_{s=0}^{k-1}\mathbb{E}\left\|\theta_{t,s}^{(i)} - \theta_i^*\right\|^2 + 2\alpha_l^2\frac{1}{NK}\sum_{i=1}^{N}\sum_{k=0}^{K-1}kc_3^2 \quad \text{(Lemma 11)} \\
&+ 2\alpha_l^2\frac{1}{NK}\sum_{i=1}^{N}\sum_{k=0}^{K-1}\sum_{\substack{s,s'=0\\ s\neq s'}}^{k-1}\mathbb{E}\left[\mathbb{E}\left\langle Z_i(O_{t,s}^{(i)}), Z_i(O_{t,s'}^{(i)})\right\rangle \mid \mathcal{F}_s^t\right] \\
&\leq 2\alpha_l^2\frac{1}{NK}\sum_{i=1}^{N}\sum_{k=0}^{K-1}kc_1^2\sum_{s=0}^{k-1}\mathbb{E}\left\|\theta_{t,s}^{(i)} - \theta_i^*\right\|^2 + 2\alpha_l^2\frac{1}{NK}\sum_{i=1}^{N}\sum_{k=0}^{K-1}kc_3^2 \\
&+ 2\alpha_l^2\frac{1}{NK}\sum_{i=1}^{N}\sum_{k=0}^{K-1}\sum_{\substack{s,s'=0\\ s\neq s'}}^{k-1}\mathbb{E}\left[\left\langle Z_i(O_{t,s}^{(i)}), \mathbb{E}\left[Z_i(O_{t,s'}^{(i)}) \mid \mathcal{F}_s^t\right]\right\rangle\right] \\
&\leq 2\alpha_l^2\frac{1}{NK}\sum_{i=1}^{N}\sum_{k=0}^{K-1}kc_1^2\sum_{s=0}^{k-1}\mathbb{E}\left\|\theta_{t,s}^{(i)} - \theta_i^*\right\|^2 + 2\alpha_l^2\frac{1}{NK}\sum_{i=1}^{N}\sum_{k=0}^{K-1}kc_3^2 \\
&+ 4\alpha_l^2\frac{1}{NK}\sum_{i=1}^{N}\sum_{k=0}^{K-1}\sum_{\substack{s,s'=0\\ s<s'}}^{k-1}\mathbb{E}\left[\left\|Z_i(O_{t,s}^{(i)})\right\|\left\|\mathbb{E}\left[Z_i(O_{t,s'}^{(i)}) \mid \mathcal{F}_s^t\right]\right\|\right] \\
&\leq 4\alpha_l^2\frac{1}{NK}\sum_{i=1}^{N}\sum_{k=0}^{K-1}kc_1^2\sum_{s=0}^{k-1}\mathbb{E}\left\|\theta_{t,s}^{(i)} - \bar{\theta}_t\right\|^2 + 4\alpha_l^2\frac{1}{NK}\sum_{i=1}^{N}\sum_{k=0}^{K-1}kc_1^2\sum_{s=0}^{k-1}\mathbb{E}\left\|\bar{\theta}_t - \theta_i^*\right\|^2 \quad \text{(Eq (11))}
\end{aligned}$$

$$
\begin{aligned}
&+ 2\alpha_l^2 \frac{1}{NK} \sum_{i=1}^{N} \sum_{k=0}^{K-1} kc_3^2 + 4\alpha_l^2 \frac{1}{NK} \sum_{i=1}^{N} \sum_{k=0}^{K-1} \sum_{\substack{s,s'=0 \\ s<s'}}^{k-1} c_3 L_2 \rho^{s'-s} \quad \text{(Lemma 12)} \\
&\leq 4\alpha_l^2 \frac{1}{NK} \sum_{i=1}^{N} \sum_{k=0}^{K-1} kc_1^2 \sum_{s=0}^{k-1} \mathbb{E} \left\| \theta_{t,s}^{(i)} - \bar{\theta}_t \right\|^2 + 4\alpha_l^2 \frac{1}{NK} \sum_{i=1}^{N} \sum_{k=0}^{K-1} 4kc_1^2 (K-1) H^2 \\
&+ 2\alpha_l^2 \frac{1}{NK} \sum_{i=1}^{N} \sum_{k=0}^{K-1} kc_3^2 + 4\alpha_l^2 \frac{1}{NK} \sum_{i=1}^{N} \sum_{k=0}^{K-1} \underbrace{\sum_{\substack{s,s'=0 \\ s<s'}}^{k-1} c_3 L_2 \rho^{s'-s}}_{\mathcal{M}_1}
\end{aligned} \tag{55}
$$

其中，我们在最后一个不等式中用到了 $\bar{\theta}_t, \theta_i^* \in \mathcal{H}$ 的性质，即 $\|\bar{\theta}_t\| \leq H^2$ 和 $\|\theta_i^*\| \leq H^2$。现在我们对$\mathcal{M}_1$ 进行界定:

$$
\sum_{\substack{s,s'=0 \\ s<s'}}^{k-1} c_3 L_2 \rho^{s'-s} = c_3 L_2 \sum_{s=0}^{k-1} \sum_{s'=s+1}^{k-1} \rho^{s'-s} = c_3 L_2 \sum_{s=0}^{k-1} \frac{\rho - \rho^{s-s'}}{1-\rho} \leq c_3 L_2 \frac{\rho k}{1-\rho} \tag{56}
$$

定义 $\mathcal{R}_K \triangleq \sum_{i=1}^{N} \sum_{k=0}^{K-1} \mathbb{E} \left\| \theta_t^{(i)}{}_{,\,k} - \bar{\theta}_t \right\|^2$ 并注意到 $\mathcal{R}_K$ 关于 K单调递增。根据此定义，将 $\mathcal{M}_1$ 的上界代入方程 (55)，可得:

$$
\begin{aligned}
\mathcal{R}_K &\leq 4\alpha_l^2 \sum_{i=1}^{N} \sum_{k=0}^{K-1} kc_1^2 \sum_{s=0}^{k-1} \mathbb{E} \left\| \theta_{t,s}^{(i)} - \bar{\theta}_t \right\|^2 + 4\alpha_l^2 \sum_{i=1}^{N} \sum_{k=0}^{K-1} 4kc_1^2 (K-1) H^2 \\
&\quad + 2\alpha_l^2 \sum_{i=1}^{N} \sum_{k=0}^{K-1} kc_3^2 + 4\alpha_l^2 \sum_{i=1}^{N} \sum_{k=0}^{K-1} c_3 L_2 \frac{\rho k}{1-\rho} \\
&\leq 2\alpha_l^2 (K-1) NK \left[c_3^2 + \frac{2c_3 L_2 \rho}{1-\rho} + 8c_1^2 (K-1) H^2 \right] \\
&\quad + 4\alpha_l^2 c_1^2 (K-1) \sum_{k=1}^{K-1} \underbrace{\sum_{i=1}^{N} \sum_{s=0}^{k-1} \mathbb{E} \left\| \theta_{t,s}^{(i)} - \bar{\theta}_t \right\|^2}_{\mathcal{R}_k} \\
&= 2\alpha_l^2 (K-1) NK \left[c_3^2 + \frac{2c_3 L_2 \rho}{1-\rho} + 8c_1^2 (K-1) H^2 \right] + 4\alpha_l^2 c_1^2 (K-1) \sum_{k=1}^{K-1} \mathcal{R}_k
\end{aligned} \tag{57}
$$

根据$\mathcal{R}_k$的单调性，可得

$$
\mathcal{R}_K \leq 2\alpha_l^2 (K-1) NK \left[c_3^2 + \frac{2c_3 L_2 \rho}{1-\rho} + 8c_1^2 (K-1) H^2 \right] + 4\alpha_l^2 c_1^2 (K-1)^2 \mathcal{R}_{K-1}
$$

令 α_l使得 $4\alpha_l^2 c_1^2 (K-1)^2 \leq \frac{1}{2}$，即 $\alpha_l \leq \frac{1}{2\sqrt{2}c_1 (K-1)}$，则有如下递推关系:

$$
\mathcal{R}_K \leq \frac{1}{2} \mathcal{R}_{K-1} + 2\alpha_l^2 (K-1) NK \left[c_3^2 + \frac{2c_3 L_2 \rho}{1-\rho} + 8c_1^2 (K-1) H^2 \right] \tag{58}
$$

对于所有 $k \in [K]$. 接下来，我们展开递推关系，回退 $K-1$ 步，并利用$\mathcal{R}_1 = 0$这一事实，可得:

$$
\begin{aligned}
\mathcal{R}_K &\leq \left\{ \sum_{l=1}^{\infty} \left(\frac{1}{2} \right)^l \right\} \left(2\alpha_l^2 (K-1) NK \left[c_3^2 + \frac{2c_3 L_2 \rho}{1-\rho} + 8c_1^2 (K-1) H^2 \right] \right) \\
&= 4\alpha_l^2 (K-1) NK \left[c_3^2 + \frac{2c_3 L_2 \rho}{1-\rho} + 8c_1^2 (K-1) H^2 \right]
\end{aligned} \tag{59}
$$

我们通过两边除以 NK并代入 $\alpha_l = \frac{\alpha}{K\alpha_g}$来完成证明。 —— □

I.2.5 每轮进度

引理16. *(每轮进展)*。如果局部步长 $\alpha_l \leq \frac{1}{2\sqrt{2}c_1(K-1)}$，且有效步长~~$\alpha = K\alpha_l\alpha_g$~~ 满足：

$$\alpha \leq \min\{\frac{\xi_1}{24(c_1+c_2)^2+24\xi_1^2+16}, 1, \frac{\xi_1(c_1+c_2)}{2L_1+8\tau^2c_4^2}, \frac{1}{30c_4(\tau+1)}, \frac{1}{96c_4^2\tau}, \mathcal{X}\},$$ [5]

其中

$$\mathcal{X} = \frac{2B(\epsilon,\epsilon_1)G + 3\xi_1(c_1+c_2)\Gamma^2(\epsilon,\epsilon_1)}{4B^2(\epsilon,\epsilon_1) + 24(c_1+c_2)^2\Gamma^2(\epsilon,\epsilon_1) + 2L_1\Gamma(\epsilon,\epsilon_1)G + 6400c_1^2c_4^2\tau^3\Gamma^2(\epsilon,\epsilon_1) + 8c_1^2\tau^2\Gamma^2(\epsilon,\epsilon_1)},$$

并选择 $\tau = \lceil \frac{\tau^{\text{mix}}(\alpha_T^2)}{K} \rceil$，则有，

$$\begin{aligned}
&\mathbb{E}_{t-2\tau}\|\bar{\theta}_{t+1}-\theta^*\|^2 \\
&\leq \underbrace{(1+32\alpha\xi_1(c_1+c_2))\,\mathbb{E}_{t-2\tau}\left\|\bar{\theta}_t-\theta^*\right\|^2 + 2\alpha\mathbb{E}_{t-2\tau}\left\langle \bar{g}(\bar{\theta}_t), \bar{\theta}_t-\theta^*\right\rangle + 4\alpha^2\mathbb{E}_{t-2\tau}\left\|\bar{g}(\bar{\theta}_t)\right\|^2}_{\textit{Expected progress for the virtual MDP}} \\
&+ \underbrace{\frac{9+28\tau^2}{NK}\alpha^2 d_2^2}_{\textit{Linear speedup}} + \underbrace{\alpha^3\left(36L_2^2 + \frac{108\tau}{1-\rho^2}L_2^2 + 4L_1G^2 + 2L_2G\right)}_{\textit{High order terms: } O(\alpha^3)} \\
&+ \underbrace{\frac{4\alpha^3}{K\alpha_g^2}(\frac{14}{\xi_1}+14\xi_1)(c_1+c_2)\left[c_3^2 + \frac{2c_3L_2\rho}{1-\rho} + 4c_1^2(K-1)H^2\right]}_{\textit{drift term}} \\
&+ \underbrace{4\alpha B(\epsilon,\epsilon_1)G + 6\alpha\xi_1(c_1+c_2)\Gamma^2(\epsilon,\epsilon_1)}_{\textit{heterogeneity term}}. \qquad (60)
\end{aligned}$$

其中，ξ_1 是任意普适正常数。

证明。根据更新规则以及投影算子具有非扩张性这一事实，我们有：

$$\begin{aligned}
&\mathbb{E}_{t-\tau}\left\|\bar{\theta}_{t+1}-\theta^*\right\|^2 \\
&= \mathbb{E}_{t-\tau}\left\|\Pi_{2,\mathcal{H}}\left(\bar{\theta}_t + \frac{\alpha}{NK}\sum_{i=1}^{N}\sum_{k=0}^{K-1} g_i(\theta_{t,k}^{(i)})\right) - \theta^*\right\|^2 \\
&\leq \mathbb{E}_{t-\tau}\left\|\bar{\theta}_t + \frac{\alpha}{NK}\sum_{i=1}^{N}\sum_{k=0}^{K-1} g_i(\theta_{t,k}^{(i)}) - \theta^*\right\|^2 \\
&= \mathbb{E}_{t-\tau}\left\|\bar{\theta}_t-\theta^*\right\|^2 + 2\mathbb{E}_{t-\tau}\left\langle \frac{\alpha}{NK}\sum_{i=1}^{N}\sum_{k=0}^{K-1}\bar{g}_i(\theta_{t,k}^{(i)}), \bar{\theta}_t-\theta^*\right\rangle \\
&\quad + 2\mathbb{E}_{t-\tau}\left\langle \frac{\alpha}{NK}\sum_{i=1}^{N}\sum_{k=0}^{K-1}\left[g_i(\theta_{t,k}^{(i)}) - \bar{g}_i(\theta_{t,k}^{(i)})\right], \bar{\theta}_t-\theta^*\right\rangle
\end{aligned}$$

[5] 由于 $\mathcal{X}$ 中的分母由异质性项构成，而该项非常小，因此会使 $\mathcal{X}$ 变大，这一要求很容易满足。总的来说，步长的可行集并非为空。

$$
\begin{aligned}
&+\alpha^2\mathbb{E}_{t-\tau}\Big\|\frac{1}{NK}\sum_{i=1}^{N}\sum_{k=0}^{K-1}g_i(\theta_{t,k}^{(i)})\Big\|^2\\
&\le \underbrace{\mathbb{E}_{t-\tau}\left\{\Big\|\bar{\theta}_t-\theta^*\Big\|^2+2\Big\langle\frac{\alpha}{N}\sum_{i=1}^{N}\bar{g}_i(\bar{\theta}_t),\bar{\theta}_t-\theta^*\Big\rangle+2\Big\langle\frac{\alpha}{NK}\sum_{i=1}^{N}\sum_{k=0}^{K-1}\bar{g}_i(\theta_{t,k}^{(i)})-\bar{g}_i(\bar{\theta}_t),\bar{\theta}_t-\theta^*\Big\rangle\right\}}_{\mathcal{B}_1}\\
&+2\alpha\mathbb{E}_{t-\tau}\underbrace{\Big\langle\frac{1}{NK}\sum_{i=1}^{N}\sum_{k=0}^{K-1}\big[g_i(\theta_{t,k}^{(i)})-\bar{g}_i(\theta_{t,k}^{(i)})\big],\bar{\theta}_t-\theta^*\Big\rangle}_{\mathcal{B}_2}+\alpha^2\mathbb{E}_{t-\tau}\underbrace{\Big\|\frac{1}{NK}\sum_{i=1}^{N}\sum_{k=0}^{K-1}g_i(\theta_{t,k}^{(i)})\Big\|^2}_{\mathcal{B}_3}
\end{aligned}\tag{61}
$$

现在，我们通过将梯度偏差项 $\mathcal{B}_2$ 分解为三个项，开始对其进行界定：

$$
\begin{aligned}
&\Big\langle\frac{1}{NK}\sum_{i=1}^{N}\sum_{k=0}^{K-1}\big[g_i(\theta_{t,k}^{(i)})-\bar{g}_i(\theta_{t,k}^{(i)})\big],\bar{\theta}_t-\theta^*\Big\rangle\\
=&\underbrace{\Big\langle\frac{1}{NK}\sum_{i=1}^{N}\sum_{k=0}^{K-1}\big[g_i(\theta_{t,k}^{(i)})-\bar{g}_i(\theta_{t,k}^{(i)})\big],\bar{\theta}_t-\bar{\theta}_{t-\tau}\Big\rangle}_{\mathcal{B}_{21}}\\
&+\underbrace{\Big\langle\frac{1}{NK}\sum_{i=1}^{N}\sum_{k=0}^{K-1}\big[g_i(\theta_{t,k}^{(i)})-g_i(\theta_{t-\tau,k}^{(i)})-\bar{g}_i(\theta_{t,k}^{(i)})+\bar{g}_i(\theta_{t-\tau,k}^{(i)})\big],\bar{\theta}_{t-\tau}-\theta^*\Big\rangle}_{\mathcal{B}_{22}}\\
&+\underbrace{\Big\langle\frac{1}{NK}\sum_{i=1}^{N}\sum_{k=0}^{K-1}\big[g_i(\theta_{t-\tau,k}^{(i)})-\bar{g}_i(\theta_{t-\tau,k}^{(i)})\big],\bar{\theta}_{t-\tau}-\theta^*\Big\rangle}_{\mathcal{B}_{23}}.
\end{aligned}\tag{62}
$$

接下来，我们将 $\mathbb{E}_{t-\tau}[\mathcal{B}_{21}]$ 界定为：

$$
\begin{aligned}
&\mathbb{E}_{t-\tau}\Big\langle\frac{1}{NK}\sum_{i=1}^{N}\sum_{k=0}^{K-1}\big[g_i(\theta_{t,k}^{(i)})-\bar{g}_i(\theta_{t,k}^{(i)})\big],\bar{\theta}_t-\bar{\theta}_{t-\tau}\Big\rangle\\
&\le\mathbb{E}_{t-\tau}\Big\|\frac{1}{NK}\sum_{i=1}^{N}\sum_{k=0}^{K-1}g_i(\theta_{t,k}^{(i)})-\bar{g}_i(\theta_{t,k}^{(i)})\Big\|\Big\|\bar{\theta}_t-\bar{\theta}_{t-\tau}\Big\|\\
&\overset{(a)}{=}\mathbb{E}_{t-\tau}\left[\Big\|\frac{1}{NK}\sum_{i=1}^{N}\sum_{k=0}^{K-1}(-A_i(O_{t,k}^{(i)})+\bar{A}_i)(\theta_{t,k}^{(i)}-\theta_i^*)+Z_i(O_{t,k}^{(i)})\Big\|\Big\|\bar{\theta}_t-\bar{\theta}_{t-\tau}\Big\|\right]\\
&\le\mathbb{E}_{t-\tau}\left[\Big\|\frac{1}{NK}\sum_{i=1}^{N}\sum_{k=0}^{K-1}(A_i(O_{t,k}^{(i)})-\bar{A}_i)(\theta_{t,k}^{(i)}-\theta_i^*)\Big\|\Big\|\bar{\theta}_t-\bar{\theta}_{t-\tau}\Big\|\right]\\
&\quad+\mathbb{E}_{t-\tau}\left[\Big\|\frac{1}{NK}\sum_{i=1}^{N}\sum_{k=0}^{K-1}Z_i(O_{t,k}^{(i)})\Big\|\Big\|\bar{\theta}_t-\bar{\theta}_{t-\tau}\Big\|\right]\\
&\le\frac{1}{NK}\sum_{i=1}^{N}\sum_{k=0}^{K-1}\mathbb{E}_{t-\tau}\left[\Big\|(A_i(O_{t,k}^{(i)})-\bar{A}_i)(\theta_{t,k}^{(i)}-\theta_i^*)\Big\|\Big\|\bar{\theta}_t-\bar{\theta}_{t-\tau}\Big\|\right]\\
&\quad+\frac{\alpha}{2}\mathbb{E}_{t-\tau}\left[\Big\|\frac{1}{NK}\sum_{i=1}^{N}\sum_{k=0}^{K-1}Z_i(O_{t,k}^{(i)})\Big\|^2\right]+\frac{1}{2\alpha}\mathbb{E}_{t-\tau}\left[\Big\|\bar{\theta}_t-\bar{\theta}_{t-\tau}\Big\|^2\right]\\
&\overset{(b)}{\le}\frac{(c_1+c_2)}{NK}\sum_{i=1}^{N}\sum_{k=0}^{K-1}\mathbb{E}_{t-\tau}\Big\|\theta_{t,k}^{(i)}-\theta_i^*\Big\|\Big\|\bar{\theta}_t-\bar{\theta}_{t-\tau}\Big\|
\end{aligned}
$$

$$
\begin{aligned}
&+\frac{\alpha}{2}\mathbb{E}_{t-\tau}\left\|\frac{1}{NK}\sum_{i=1}^{N}\sum_{k=0}^{K-1}Z_i(O_{t,k}^{(i)})\right\|^2+\frac{1}{2\alpha}\mathbb{E}_{t-\tau}\left\|\bar{\theta}_t-\bar{\theta}_{t-\tau}\right\|^2\\
&\leq\frac{\xi_1(c_1+c_2)}{2NK}\sum_{i=1}^{N}\sum_{k=0}^{K-1}\mathbb{E}_{t-\tau}\left\|\theta_{t,k}^{(i)}-\theta_i^*\right\|^2+\frac{(c_1+c_2)}{2\xi_1 NK}\sum_{i=1}^{N}\sum_{k=0}^{K-1}\mathbb{E}_{t-\tau}\left\|\bar{\theta}_t-\bar{\theta}_{t-\tau}\right\|^2\\
&\quad\text{(Young's inequality (12) with constant } \xi_1)\\
&+\frac{\alpha}{2}\mathbb{E}_{t-\tau}\left\|\frac{1}{NK}\sum_{i=1}^{N}\sum_{k=0}^{K-1}Z_i(O_{t,k}^{(i)})\right\|^2+\frac{1}{2\alpha}\mathbb{E}_{t-\tau}\left\|\bar{\theta}_t-\bar{\theta}_{t-\tau}\right\|^2\\
&\overset{(c)}{\leq}\frac{3\xi_1(c_1+c_2)}{2NK}\sum_{i=1}^{N}\sum_{k=0}^{K-1}\mathbb{E}_{t-\tau}\left\|\theta_{t,k}^{(i)}-\bar{\theta}_t\right\|^2+\frac{3\xi_1(c_1+c_2)}{2NK}\sum_{i=1}^{N}\sum_{k=0}^{K-1}\mathbb{E}_{t-\tau}\left\|\bar{\theta}_t-\theta^*\right\|^2\\
&+\frac{3\xi_1(c_1+c_2)}{2NK}\sum_{i=1}^{N}\sum_{k=0}^{K-1}\mathbb{E}_{t-\tau}\left\|\theta^*-\theta_i^*\right\|^2+\frac{(c_1+c_2)}{2\xi_1 NK}\sum_{i=1}^{N}\sum_{k=0}^{K-1}\mathbb{E}_{t-\tau}\left\|\bar{\theta}_t-\bar{\theta}_{t-\tau}\right\|^2\\
&+\frac{\alpha}{2}\mathbb{E}_{t-\tau}\left\|\frac{1}{NK}\sum_{i=1}^{N}\sum_{k=0}^{K-1}Z_i(O_{t,k}^{(i)})\right\|^2+\frac{1}{2\alpha}\mathbb{E}_{t-\tau}\left\|\bar{\theta}_t-\bar{\theta}_{t-\tau}\right\|^2\\
&=\frac{3\xi_1(c_1+c_2)}{2NK}\sum_{i=1}^{N}\sum_{k=0}^{K-1}\mathbb{E}_{t-\tau}\left\|\theta_{t,k}^{(i)}-\bar{\theta}_t\right\|^2+\frac{3\xi_1(c_1+c_2)}{2}\mathbb{E}_{t-\tau}\left\|\bar{\theta}_t-\theta^*\right\|^2\\
&+\frac{3\xi_1(c_1+c_2)}{2}\Gamma^2(\epsilon,\epsilon_1)+\frac{(c_1+c_2)}{2\xi_1}\mathbb{E}_{t-\tau}\left\|\bar{\theta}_t-\bar{\theta}_{t-\tau}\right\|^2\\
&+\frac{\alpha}{2}\mathbb{E}_{t-\tau}\left\|\frac{1}{NK}\sum_{i=1}^{N}\sum_{k=0}^{K-1}Z_i(O_{t,k}^{(i)})\right\|^2+\frac{1}{2\alpha}\mathbb{E}_{t-\tau}\left\|\bar{\theta}_t-\bar{\theta}_{t-\tau}\right\|^2\\
&\overset{(d)}{\leq}\frac{3\xi_1(c_1+c_2)}{2NK}\sum_{i=1}^{N}\sum_{k=0}^{K-1}\mathbb{E}_{t-\tau}\left\|\theta_{t,k}^{(i)}-\bar{\theta}_t\right\|^2+\frac{3\xi_1(c_1+c_2)}{2}\mathbb{E}_{t-\tau}\left\|\bar{\theta}_t-\theta^*\right\|^2\\
&+\frac{3\xi_1(c_1+c_2)}{2}\Gamma^2(\epsilon,\epsilon_1)+\frac{c_1+c_2}{2\xi_1}\mathbb{E}_{t-\tau}\left\|\bar{\theta}_t-\bar{\theta}_{t-\tau}\right\|^2\\
&+\frac{\alpha}{2}\left[\frac{d_2^2}{NK}+4L_2^2\rho^{2\tau K}\right]+\frac{1}{2\alpha}\mathbb{E}_{t-\tau}\left\|\bar{\theta}_t-\bar{\theta}_{t-\tau}\right\|^2\\
&=\frac{3\xi_1(c_1+c_2)}{2NK}\sum_{i=1}^{N}\sum_{k=0}^{K-1}\mathbb{E}_{t-\tau}\left\|\theta_{t,k}^{(i)}-\bar{\theta}_t\right\|^2+\frac{3\xi_1(c_1+c_2)}{2}\mathbb{E}_{t-\tau}\left\|\bar{\theta}_t-\theta^*\right\|^2\\
&+\frac{3\xi_1(c_1+c_2)}{2}\Gamma^2(\epsilon,\epsilon_1)+\left(\frac{c_1+c_2}{2\xi_1}+\frac{1}{2\alpha}\right)\mathbb{E}_{t-\tau}\left\|\bar{\theta}_t-\bar{\theta}_{t-\tau}\right\|^2+\frac{\alpha}{2}\left[\frac{d_2^2}{NK}+\underbrace{4L_2^2\rho^{2\tau K}}_{\leq 4L_2^2\alpha^2}\right],
\end{aligned}
\tag{63}
$$

其中，(*a*) 由 $g_i(\theta_{t,k}^{(i)},)=-A_i(O_{t,k}^{(i)},)(\theta_{t,k}^{(i)}-\theta_i^*)+Z_i(O_{t,k}^{(i)},)$ 得出，(*b*) 由引理 12（$A_i(O_{t,k}^{(i)},)$ 和 $\bar{A}_i$ 的上界）得出，(*c*) 由公式（13）得出，(*d*) 则由引理13得出。

并且我们将$\mathcal{B}_{22}$ 界定为:

$$
\begin{aligned}
\mathcal{B}_{22}&=\left\langle\frac{1}{NK}\sum_{i=1}^{N}\sum_{k=0}^{K-1}\left[g_i(\theta_{t,k}^{(i)})-g_i(\theta_{t-\tau,k}^{(i)})-\bar{g}_i(\theta_{t,k}^{(i)})+\bar{g}_i(\theta_{t-\tau,k}^{(i)})\right],\bar{\theta}_{t-\tau}-\theta^*\right\rangle\\
&\leq\frac{1}{NK}\sum_{i=1}^{N}\sum_{k=0}^{K-1}\left\|g_i(\theta_{t,k}^{(i)})-g_i(\theta_{t-\tau,k}^{(i)})-\bar{g}_i(\theta_{t,k}^{(i)})+\bar{g}_i(\theta_{t-\tau,k}^{(i)})\right\|\left\|\bar{\theta}_{t-\tau}-\theta^*\right\|\\
&\quad\text{(Cauchy-Schwarz inequality)}
\end{aligned}
$$

$$
\begin{aligned}
&\leq \frac{1}{NK}\sum_{i=1}^{N}\sum_{k=0}^{K-1}\left[\left\|g_i(\theta_{t,k}^{(i)})-g_i(\theta_{t-\tau,k}^{(i)})\right\|+\left\|\bar{g}_i(\theta_{t,k}^{(i)})-\bar{g}_i(\theta_{t-\tau,k}^{(i)})\right\|\right]\left\|\bar{\theta}_{t-\tau}-\theta^*\right\| \\
&\overset{(a)}{\leq} \frac{1}{NK}\sum_{i=1}^{N}\sum_{k=0}^{K-1}\left[2\left\|\theta_{t,k}^{(i)}-\theta_{t-\tau,k}^{(i)}\right\|+2\left\|\theta_{t,k}^{(i)}-\theta_{t-\tau,k}^{(i)}\right\|\right]\left\|\bar{\theta}_{t-\tau}-\theta^*\right\| \\
&\leq \frac{1}{NK}\sum_{i=1}^{N}\sum_{k=0}^{K-1}\left[4\left\|\theta_{t,k}^{(i)}-\bar{\theta}_t\right\|+4\left\|\bar{\theta}_t-\bar{\theta}_{t-\tau}\right\|+4\left\|\bar{\theta}_{t-\tau}-\theta_{t-\tau,k}^{(i)}\right\|\right]\left\|\bar{\theta}_{t-\tau}-\theta^*\right\| \\
&\qquad \text{(Triangle inequality)} \\
&\leq 4\delta_t\left\|\bar{\theta}_{t-\tau}-\theta^*\right\|+4\left\|\bar{\theta}_t-\bar{\theta}_{t-\tau}\right\|\left\|\bar{\theta}_{t-\tau}-\theta^*\right\|+4\delta_{t-\tau}\left\|\bar{\theta}_{t-\tau}-\theta^*\right\| \\
&\leq \frac{2}{\xi_2}\Delta_t+\frac{2}{\xi_2}\Delta_{t-\tau}+(2\xi_2+4\xi_2)\left\|\bar{\theta}_{t-\tau}-\theta^*\right\|^2+\frac{2}{\xi_2}\left\|\bar{\theta}_t-\bar{\theta}_{t-\tau}\right\|^2 \\
&\qquad \text{(Young's inequality (12) with constants } \xi_2 \text{ and } \delta_t^2\leq\Delta_t) \\
&\leq \frac{2}{\xi_2}\Delta_t+\frac{2}{\xi_2}\Delta_{t-\tau}+12\xi_2\left\|\bar{\theta}_t-\theta^*\right\|^2+(12\xi_2+\frac{2}{\xi_2})\left\|\bar{\theta}_t-\bar{\theta}_{t-\tau}\right\|^2 \quad \text{(Eq 13)}
\end{aligned} \tag{64}
$$

其中，(a) 是由稳态 $\bar{g}$ 的 2-Lipschitz 性质（即引理5）和随机方向 g_i （即引理 6）得出的，$\delta_t=\frac{1}{NK}\sum_{i=1}^{N}\sum_{k=0}^{K-1}\left\|\theta_t^{(i)},_k-\bar{\theta}_t\right\|$ 以及 $\Delta_t\triangleq\frac{1}{NK}\sum_{i=1}^{N}\sum_{k=0}^{K-1}\mathbb{E}\left\|\theta_t^{(i)},_k-\bar{\theta}_t\right\|^2$.

现在，我们将$\mathcal{B}_{23}$ 界定为:

$$
\begin{aligned}
&\mathbb{E}_{t-\tau}[\mathcal{B}_{23}] \\
&=\left\langle\frac{1}{NK}\sum_{i=1}^{N}\sum_{k=0}^{K-1}\mathbb{E}_{t-\tau}\left[g_i(\theta_{t-\tau,k}^{(i)})-\bar{g}_i(\theta_{t-\tau,k}^{(i)})\right],\bar{\theta}_{t-\tau}-\theta^*\right\rangle \\
&\leq \frac{1}{NK}\sum_{i=1}^{N}\sum_{k=0}^{K-1}\left\|\bar{\theta}_{t-\tau}-\theta^*\right\|\left\|\mathbb{E}_{t-\tau}\left[g_i(\theta_{t-\tau,k}^{(i)})-\bar{g}_i(\theta_{t-\tau,k}^{(i)})\right]\right\| \text{ (Cauchy-Schwarz inequality)} \\
&=\frac{1}{NK}\sum_{i=1}^{N}\sum_{k=0}^{K-1}\left\|\bar{\theta}_{t-\tau}-\theta^*\right\|\left\|\mathbb{E}_{t-\tau}\left[-A_i(O_{t,k}^{(i)})(\theta_{t-\tau,k}^{(i)}-\theta_i^*)+Z_i(O_{t,k}^{(i)})+\bar{A}_i(\theta_{t-\tau,k}^{(i)}-\theta_i^*)\right]\right\| \\
&\leq \frac{1}{NK}\sum_{i=1}^{N}\sum_{k=0}^{K-1}\left\|\bar{\theta}_{t-\tau}-\theta^*\right\|\left\{\left\|\mathbb{E}_{t-\tau}(A_i(O_{t,k}^{(i)})-\bar{A}_i)(\theta_{t-\tau,k}^{(i)}-\theta_i^*)\right\|+\left\|\mathbb{E}_{t-\tau}\left[Z_i(O_{t,k}^{(i)})\right]\right\|\right\} \\
&\overset{(a)}{\leq} \frac{1}{NK}\sum_{i=1}^{N}\sum_{k=0}^{K-1}\left\|\bar{\theta}_{t-\tau}-\theta^*\right\|\left\{L_1\rho^{\tau K+k}\left\|\theta_{t-\tau,k}^{(i)}-\theta_i^*\right\|+L_2\rho^{\tau K+k}\right\} \\
&\leq \frac{1}{NK}\sum_{i=1}^{N}\sum_{k=0}^{K-1}\left\|\bar{\theta}_{t-\tau}-\theta^*\right\| \\
&\times\left\{L_1\rho^{\tau K+k}\left[\left\|\theta_{t-\tau,k}^{(i)}-\bar{\theta}_{t-\tau}\right\|+\left\|\bar{\theta}_{t-\tau}-\theta^*\right\|+\left\|\theta^*-\theta_i^*\right\|\right]+L_2\rho^{\tau K+k}\right\} \\
&\overset{(b)}{\leq} \alpha^2L_1\left\|\bar{\theta}_{t-\tau}-\theta^*\right\|\delta_{t-\tau}+\alpha^2L_1\left\|\bar{\theta}_{t-\tau}-\theta^*\right\|^2+\alpha^2L_1\Gamma(\epsilon,\epsilon_1)G+\alpha^2L_2G \\
&\leq \alpha^2L_1\left\|\bar{\theta}_{t-\tau}-\theta^*\right\|^2+\alpha^2L_1\Delta_{t-\tau}+\alpha^2L_1\left\|\bar{\theta}_{t-\tau}-\theta^*\right\|^2+\alpha^2L_1\Gamma(\epsilon,\epsilon_1)G+\alpha^2L_2G \\
&\overset{(c)}{\leq} 2\alpha^2L_1G^2+\alpha^2L_2G+\alpha^2L_1\Delta_{t-\tau}+\alpha^2L_1\Gamma(\epsilon,\epsilon_1)G,
\end{aligned} \tag{65}
$$

其中，(a) 由引理 12 得出；(b) 由以下事实得出： $\bar{\theta}_{t-\tau},\theta^*\in\mathcal{H}$，其半径为 $H\leq\frac{G}{2}$，且 $\tau=\lceil\log_\rho(\alpha^2/K)\rceil$（即 $\rho^{\tau K}\leq\alpha^2$）；(c) 则由以下事实得出： $\bar{\theta}_{t-\tau},\theta^*\in\mathcal{H}$。那么，项 $\mathcal{B}_2$ 可被

界定为：

$$
\begin{aligned}
&\mathbb{E}_{t-\tau}[\mathcal{B}_2] \\
&= \mathbb{E}_{t-\tau}[\mathcal{B}_{21} + \mathcal{B}_{22} + \mathcal{B}_{23}] \\
&\leq \frac{3\xi_1(c_1+c_2)}{2}\mathbb{E}_{t-\tau}[\Delta_t] + \frac{3\xi_1(c_1+c_2)}{2}\mathbb{E}_{t-\tau}\left\|\bar{\theta}_t - \theta^*\right\|^2 + \frac{3\xi_1(c_1+c_2)}{2}\Gamma^2(\epsilon,\epsilon_1) \\
&+ \left(\frac{c_1+c_2}{2\xi_1} + \frac{1}{2\alpha}\right)\mathbb{E}_{t-\tau}\left\|\bar{\theta}_t - \bar{\theta}_{t-\tau}\right\|^2 + \frac{\alpha}{2}\left[\frac{d_2^2}{NK} + 4L_2^2\rho^{2\tau K}\right] \\
&+ \frac{2}{\xi_2}\mathbb{E}_{t-\tau}[\Delta_t] + \frac{2}{\xi_2}\mathbb{E}_{t-\tau}[\Delta_{t-\tau}] + 12\xi_2\mathbb{E}_{t-\tau}\left\|\bar{\theta}_t - \theta^*\right\|^2 + (12\xi_2 + \frac{2}{\xi_2})\mathbb{E}_{t-\tau}\left\|\bar{\theta}_t - \bar{\theta}_{t-\tau}\right\|^2 \\
&+ 2\alpha^2 L_1 G^2 + \alpha^2 L_2 G + \alpha^2 L_1 \mathbb{E}_{t-\tau}[\Delta_{t-\tau}] + \alpha^2 L_1 \Gamma(\epsilon,\epsilon_1)G \\
&\leq \left(\frac{3\xi_1(c_1+c_2)}{2} + 12\xi_2\right)\mathbb{E}_{t-\tau}\left\|\bar{\theta}_t - \theta^*\right\|^2 + \left(\frac{c_1+c_2}{2\xi_1} + \frac{1}{2\alpha} + 12\xi_2 + \frac{2}{\xi_2}\right)\mathbb{E}_{t-\tau}\left\|\bar{\theta}_t - \bar{\theta}_{t-\tau}\right\|^2 \\
&+ \left(\frac{3\xi_1(c_1+c_2)}{2} + \frac{2}{\xi_2}\right)\mathbb{E}_{t-\tau}[\Delta_t] + \left(\frac{2}{\xi_2} + \alpha^2 L_1\right)\Delta_{t-\tau} + \frac{\alpha}{2}\left[\frac{d_2^2}{NK} + 4L_2^2\alpha^2\right] \\
&+ 2\alpha^2 L_1 G^2 + \alpha^2 L_2 G + \frac{3\xi_1(c_1+c_2)}{2}\Gamma^2(\epsilon,\epsilon_1) + \alpha^2 L_1 \Gamma(\epsilon,\epsilon_1)G
\end{aligned} \tag{66}
$$

接下来，我们将 $\mathcal{B}_3$ 界定为：

$$
\begin{aligned}
&\mathbb{E}_{t-\tau}[\mathcal{B}_3] \\
&= \mathbb{E}_{t-\tau}\left\|\frac{1}{NK}\sum_{i=1}^{N}\sum_{k=0}^{K-1}\left[g_i(\theta_{t,k}^{(i)}) - \bar{g}_i(\theta_{t,k}^{(i)}) + \bar{g}_i(\theta_{t,k}^{(i)}) - \bar{g}_i(\bar{\theta}_t) + \bar{g}_i(\bar{\theta}_t) - \bar{g}(\bar{\theta}_t) + \bar{g}(\bar{\theta}_t)\right]\right\|^2 \\
&\leq 4\mathbb{E}_{t-\tau}\left\|\frac{1}{NK}\sum_{i=1}^{N}\sum_{k=0}^{K-1}\left(g_i(\theta_{t,k}^{(i)}) - \bar{g}_i(\theta_{t,k}^{(i)})\right)\right\|^2 + 4\mathbb{E}_{t-\tau}\left\|\frac{1}{NK}\sum_{i=1}^{N}\sum_{k=0}^{K-1}\left(\bar{g}_i(\theta_{t,k}^{(i)}) - \bar{g}_i(\bar{\theta}_t)\right)\right\|^2 \\
&+ 4\mathbb{E}_{t-\tau}\left\|\frac{1}{NK}\sum_{i=1}^{N}\sum_{k=0}^{K-1}\left(\bar{g}_i(\bar{\theta}_t) - \bar{g}(\bar{\theta}_t)\right)\right\|^2 + 4\mathbb{E}_{t-\tau}\left\|\frac{1}{NK}\sum_{i=1}^{N}\sum_{k=0}^{K-1}\bar{g}(\bar{\theta}_t)\right\|^2 \quad \text{(Eq 13)} \\
&= 4\mathbb{E}_{t-\tau}\left\|\frac{1}{NK}\sum_{i=1}^{N}\sum_{k=0}^{K-1}\left[\left(\bar{A}_i - A_i(O_{t,k}^{(i)})\right)(\theta_{t,k}^{(i)} - \theta_i^*) + Z_i(O_{t,k}^{(i)})\right]\right\|^2 + 4\mathbb{E}_{t-\tau}\left\|\bar{g}(\bar{\theta}_t)\right\|^2 \\
&+ 4\mathbb{E}_{t-\tau}\left\|\frac{1}{NK}\sum_{i=1}^{N}\sum_{k=0}^{K-1}\left(\bar{g}_i(\theta_{t,k}^{(i)}) - \bar{g}_i(\bar{\theta}_t)\right)\right\|^2 + 4\mathbb{E}_{t-\tau}\underbrace{\left\|\frac{1}{N}\sum_{i=1}^{N}\left(\bar{g}_i(\bar{\theta}_t) - \bar{g}(\bar{\theta}_t)\right)\right\|^2}_{\text{Lemma 2}} \\
&\overset{(a)}{\leq} 8\mathbb{E}_{t-\tau}\left\|\frac{1}{NK}\sum_{i=1}^{N}\sum_{k=0}^{K-1}\left(\bar{A}_i - A_i(O_{t,k}^{(i)})\right)(\theta_{t,k}^{(i)} - \theta_i^*)\right\|^2 + 8\mathbb{E}_{t-\tau}\left\|\frac{1}{NK}\sum_{i=1}^{N}\sum_{k=0}^{K-1}Z_i(O_{t,k}^{(i)})\right\|^2 \\
&+ 16\frac{1}{NK}\sum_{i=1}^{N}\sum_{k=0}^{K-1}\mathbb{E}_{t-\tau}\left\|\theta_{t,k}^{(i)} - \bar{\theta}_t\right\|^2 + 4B^2(\epsilon,\epsilon_1) + 4\mathbb{E}_{t-\tau}\left\|\bar{g}(\bar{\theta}_t)\right\|^2 \\
&\leq \frac{8}{NK}\sum_{i=1}^{N}\sum_{k=0}^{K-1}\mathbb{E}_{t-\tau}\left\|\bar{A}_i - A_i(O_{t,k}^{(i)})\right\|^2\left\|\theta_{t,k}^{(i)} - \theta_i^*\right\|^2 + 8\mathbb{E}_{t-\tau}\left\|\frac{1}{NK}\sum_{i=1}^{N}\sum_{k=0}^{K-1}Z_i(O_{t,k}^{(i)})\right\|^2 \\
&+ 16\mathbb{E}_{t-\tau}[\Delta_t] + 4B^2(\epsilon,\epsilon_1) + 4\mathbb{E}_{t-\tau}\left\|\bar{g}(\bar{\theta}_t)\right\|^2 \\
&\leq \frac{8(c_1+c_2)^2}{NK}\sum_{i=1}^{N}\sum_{k=0}^{K-1}\mathbb{E}_{t-\tau}\left\|\theta_{t,k}^{(i)} - \theta_i^*\right\|^2 + 8\mathbb{E}_{t-\tau}\left\|\frac{1}{NK}\sum_{i=1}^{N}\sum_{k=0}^{K-1}Z_i(O_{t,k}^{(i)})\right\|^2 \\
&+ 16\mathbb{E}_{t-\tau}[\Delta_t] + 4B^2(\epsilon,\epsilon_1) + 4\mathbb{E}_{t-\tau}\left\|\bar{g}(\bar{\theta}_t)\right\|^2
\end{aligned}
$$

$$
\begin{aligned}
&\leq \frac{24(c_1+c_2)^2}{NK}\sum_{i=1}^{N}\sum_{k=0}^{K-1}\mathbb{E}_{t-\tau}\left\|\theta_{t,k}^{(i)}-\bar{\theta}_t\right\|^2+\frac{24(c_1+c_2)^2}{NK}\sum_{i=1}^{N}\sum_{k=0}^{K-1}\mathbb{E}_{t-\tau}\left\|\bar{\theta}_t-\theta^*\right\|^2\\
&+\frac{24(c_1+c_2)^2}{NK}\sum_{i=1}^{N}\sum_{k=0}^{K-1}\mathbb{E}_{t-\tau}\left\|\theta_i^*-\theta^*\right\|^2+8\mathbb{E}_{t-\tau}\left\|\frac{1}{NK}\sum_{i=1}^{N}\sum_{k=0}^{K-1}Z_i(O_{t,k}^{(i)})\right\|^2\\
&+16\mathbb{E}_{t-\tau}[\Delta_t]+4B^2(\epsilon,\epsilon_1)+4\mathbb{E}_{t-\tau}\left\|\bar{g}(\bar{\theta}_t)\right\|^2\\
&=24(c_1+c_2)^2\mathbb{E}_{t-\tau}[\Delta_t]+24(c_1+c_2)^2\mathbb{E}_{t-\tau}\left\|\bar{\theta}_t-\theta^*\right\|^2+24(c_1+c_2)^2\Gamma^2(\epsilon,\epsilon_1)\\
&+8\mathbb{E}_{t-\tau}\left\|\frac{1}{NK}\sum_{i=1}^{N}\sum_{k=0}^{K-1}Z_i(O_{t,k}^{(i)})\right\|^2+16\mathbb{E}_{t-\tau}[\Delta_t]+4B^2(\epsilon,\epsilon_1)+4\mathbb{E}_{t-\tau}\left\|\bar{g}(\bar{\theta}_t)\right\|^2\\
&\overset{(b)}{\leq}(24(c_1+c_2)^2+16)\mathbb{E}_{t-\tau}[\Delta_t]+8(\frac{d_2^2}{NK}+\underbrace{4L_2^2\rho^{\tau K}}_{\leq 4L_2^2\alpha^2})+24(c_1+c_2)^2\mathbb{E}_{t-\tau}\left\|\bar{\theta}_t-\theta^*\right\|^2\\
&+4\mathbb{E}_{t-\tau}\left\|\bar{g}(\bar{\theta}_t)\right\|^2+4B^2(\epsilon,\epsilon_1)+24(c_1+c_2)^2\Gamma^2(\epsilon,\epsilon_1),
\end{aligned}
\tag{67}
$$

其中，(a) 源于 2-Lipschitz 性质 $\bar{g}_i$ (即引理 5) 以及梯度异质性（即引理 2)；(b) 源于引理 13。

接下来，我们将$\mathcal{B}_1$ 界定如下:

$$
\begin{aligned}
&\mathbb{E}_{t-\tau}[\mathcal{B}_1]\\
&=\mathbb{E}_{t-\tau}\left\|\bar{\theta}_t-\theta^*\right\|^2+2\mathbb{E}_{t-\tau}\left\langle\frac{\alpha}{N}\sum_{i=1}^{N}\bar{g}_i(\bar{\theta}_t),\bar{\theta}_t-\theta^*\right\rangle\\
&\quad+2\mathbb{E}_{t-\tau}\left\langle\frac{\alpha}{NK}\sum_{i=1}^{N}\sum_{k=0}^{K-1}\bar{g}_i(\theta_{t,k}^{(i)})-\bar{g}_i(\bar{\theta}_t),\bar{\theta}_t-\theta^*\right\rangle\\
&\leq\mathbb{E}_{t-\tau}\left\|\bar{\theta}_t-\theta^*\right\|^2+2\alpha\mathbb{E}_{t-\tau}\left\langle\frac{1}{N}\sum_{i=1}^{N}\bar{g}_i(\bar{\theta}_t)-\bar{g}(\bar{\theta}_t),\bar{\theta}_t-\theta^*\right\rangle+2\alpha\mathbb{E}_{t-\tau}\left\langle\bar{g}(\bar{\theta}_t),\bar{\theta}_t-\theta^*\right\rangle\\
&+2\alpha\mathbb{E}_{t-\tau}\left\langle\frac{1}{NK}\sum_{i=1}^{N}\sum_{k=0}^{K-1}\bar{g}_i(\theta_{t,k}^{(i)})-\bar{g}_i(\bar{\theta}_t),\bar{\theta}_t-\theta^*\right\rangle\\
&\leq\mathbb{E}_{t-\tau}\left\|\bar{\theta}_t-\theta^*\right\|^2+2\alpha\mathbb{E}_{t-\tau}\left\|\frac{1}{N}\sum_{i=1}^{N}\bar{g}_i(\bar{\theta}_t)-\bar{g}(\bar{\theta}_t)\right\|\left\|\bar{\theta}_t-\theta^*\right\|+2\alpha\mathbb{E}_{t-\tau}\left\langle\bar{g}(\bar{\theta}_t),\bar{\theta}_t-\theta^*\right\rangle\\
&+\frac{\alpha}{\xi_3}\mathbb{E}_{t-\tau}\left\|\frac{1}{NK}\sum_{i=1}^{N}\sum_{k=0}^{K-1}\left(\bar{g}_i(\theta_{t,k}^{(i)})-\bar{g}_i(\bar{\theta}_t)\right)\right\|^2+\alpha\xi_3\mathbb{E}_{t-\tau}\left\|\bar{\theta}_t-\theta^*\right\|^2\\
&\qquad\text{(Young's inequality Eq (12) with constant } \xi_3)\\
&\overset{(a)}{\leq}\mathbb{E}_{t-\tau}\left\|\bar{\theta}_t-\theta^*\right\|^2+2\alpha B(\epsilon,\epsilon_1)G+2\alpha\mathbb{E}_{t-\tau}\left\langle\bar{g}(\bar{\theta}_t),\bar{\theta}_t-\theta^*\right\rangle\\
&+\frac{\alpha}{\xi_3}\mathbb{E}_{t-\tau}\left\|\frac{1}{NK}\sum_{i=1}^{N}\sum_{k=0}^{K-1}\left(\bar{g}_i(\theta_{t,k}^{(i)})-\bar{g}_i(\bar{\theta}_t)\right)\right\|^2+\alpha\xi_3\mathbb{E}_{t-\tau}\left\|\bar{\theta}_t-\theta^*\right\|^2\\
&\overset{(b)}{\leq}\mathbb{E}_{t-\tau}\left\|\bar{\theta}_t-\theta^*\right\|^2+2\alpha B(\epsilon,\epsilon_1)G+2\alpha\mathbb{E}_{t-\tau}\left\langle\bar{g}(\bar{\theta}_t),\bar{\theta}_t-\theta^*\right\rangle\\
&\quad+\frac{4\alpha}{\xi_3}\mathbb{E}_{t-\tau}[\Delta_t]+\alpha\xi_3\mathbb{E}_{t-\tau}\left\|\bar{\theta}_t-\theta^*\right\|^2,
\end{aligned}
\tag{68}
$$

其中，(a) 是因为 θ_t，$\theta^*\in\mathcal{H}$ 以及梯度异质性；(b) 是由引理 $\bar{g}$ 中的函数 5 的 2-Lipschitz 性质导致的。

将方程 (68$_1$ 中 B 的上界、$\mathcal{B}_2$ 方程 (66中的上界，以及$\mathcal{B}_3$ 方程 (67中的上界代入方程 (61)，可得:

$$\begin{aligned}
\mathbb{E}_{t-\tau}\left\|\bar{\theta}_{t+1}-\theta^*\right\|^2 \leq\ & \mathbb{E}_{t-\tau}\left\|\bar{\theta}_t-\theta^*\right\|^2+2\alpha\mathbb{E}_{t-\tau}\left\langle \bar{g}(\bar{\theta}_t),\bar{\theta}_t-\theta^*\right\rangle+4\alpha^2\mathbb{E}_{t-\tau}\left\|\bar{g}(\bar{\theta}_t)\right\|^2\\
&+\left(\alpha\xi_3+\alpha(3\xi_1(c_1+c_2)+24\xi_2)+24\alpha^2(c_1+c_2)^2\right)\mathbb{E}_{t-\tau}\left\|\bar{\theta}_t-\theta^*\right\|^2\\
&+\alpha\left(\frac{c_1+c_2}{\xi_1}+\frac{1}{\alpha}+24\xi_2+\frac{4}{\xi_2}\right)\mathbb{E}_{t-\tau}\left\|\bar{\theta}_t-\bar{\theta}_{t-\tau}\right\|^2\\
&+\frac{9d_2^2}{NK}\alpha^2+36L_2^2\alpha^4+4\alpha^3L_1G^2+2\alpha^3L_2G+\left(\frac{4\alpha}{\xi_2}+2\alpha^3L_1\right)\Delta_{t-\tau}\\
&+\alpha\left(\frac{4}{\xi_3}+3\xi_1(c_1+c_2)+\frac{4}{\xi_2}+\alpha^2\left(24(c_1+c_2)^2+16\right)\right)\mathbb{E}_{t-\tau}[\Delta_t]\\
&+2\alpha B(\epsilon,\epsilon_1)G+4\alpha^2B^2(\epsilon,\epsilon_1)+24\alpha^2(c_1+c_2)^2\Gamma^2(\epsilon,\epsilon_1)\\
&+3\alpha\xi_1(c_1+c_2)\Gamma^2(\epsilon,\epsilon_1)+2\alpha^3L_1\Gamma(\epsilon,\epsilon_1)G
\end{aligned}\tag{69}$$

在 $\mathcal{F}_{t-2\tau}$ 的条件下，并利用引理 14 给出上界 $\mathbb{E}_{t-2\tau}\left\|\bar{\theta}_t-\bar{\theta}_{t-\tau}\right\|^2$，我们有:

$$\begin{aligned}
&\mathbb{E}_{t-2\tau}\left\|\bar{\theta}_{t+1}-\theta^*\right\|^2\\
&\leq \mathbb{E}_{t-2\tau}\left\|\bar{\theta}_t-\theta^*\right\|^2+2\alpha\mathbb{E}_{t-2\tau}\left\langle \bar{g}(\bar{\theta}_t),\bar{\theta}_t-\theta^*\right\rangle+4\alpha^2\mathbb{E}_{t-2\tau}\left\|\bar{g}(\bar{\theta}_t)\right\|^2\\
&+\underbrace{\left(\alpha\xi_3+\alpha(3\xi_1(c_1+c_2)+24\xi_2)+24\alpha^2(c_1+c_2)^2\right)}_{\mathcal{E}_1}\mathbb{E}_{t-2\tau}\left\|\bar{\theta}_t-\theta^*\right\|^2\\
&+\underbrace{\alpha\left(\frac{c_1+c_2}{\xi_1}+\frac{1}{\alpha}+24\xi_2+\frac{4}{\xi_2}\right)}_{\mathcal{E}_2}\left\{8\alpha^2\tau^2c_4^2\mathbb{E}_{t-2\tau}\left[\left\|\bar{\theta}_t-\theta^*\right\|^2\right]+14\alpha^2\tau^2\frac{d_2^2}{NK}+\frac{52L_2^2\alpha^4\tau}{1-\rho^2}\right.\\
&\left.+4\alpha^2c_4^2\tau\sum_{s=0}^{\tau}E_{t-2\tau}[\Delta_{t-s}]+3200\alpha^2c_1^2c_4^2\tau^3\Gamma^2(\epsilon,\epsilon_1)+4\alpha^2c_1^2\tau^2\Gamma^2(\epsilon,\epsilon_1)\right\}\\
&+\frac{9d_2^2}{NK}\alpha^2+36L_2^2\alpha^4+4\alpha^3L_1G^2+2\alpha^3L_2G+\left(\frac{4\alpha}{\xi_2}+2\alpha^3L_1\right)\mathbb{E}_{t-2\tau}[\Delta_{t-\tau}]\\
&+\underbrace{\alpha\left(\frac{4}{\xi_3}+3\xi_1(c_1+c_2)+\frac{4}{\xi_2}+\alpha^2\left(24(c_1+c_2)^2+16\right)\right)}_{\mathcal{E}_3}\mathbb{E}_{t-2\tau}[\Delta_t]\\
&+2\alpha B(\epsilon,\epsilon_1)G+4\alpha^2B^2(\epsilon,\epsilon_1)+24\alpha^2(c_1+c_2)^2\Gamma^2(\epsilon,\epsilon_1)\\
&+3\alpha\xi_1(c_1+c_2)\Gamma^2(\epsilon,\epsilon_1)+2\alpha^3L_1\Gamma(\epsilon,\epsilon_1)G
\end{aligned}\tag{70}$$

如果我们选择步长 α 使得 $\alpha\mathcal{E}_2=\alpha\left(\frac{c_1+c_2}{\xi_1}+\frac{1}{\alpha}+24\xi_2+\frac{4}{\xi_2}\right)\leq 2$, $\xi_1=\xi_2=\xi_3$, $\mathcal{E}_1=\alpha\xi_3+\alpha(3\xi_1(c_1+c_2)+24\xi_2)+24\alpha^2(c_1+c_2)^2\leq 28\alpha\xi_1(c_1+c_2)+24\alpha^2(c_1+c_2)^2\leq 30\alpha\xi_1(c_1+c_2)$ $(c_1,\ c_2>1)$ 且$\mathcal{E}_3=\frac{4}{\xi_3}+3\xi_1(c_1+c_2)+\frac{4}{\xi_2}+\alpha^2\left(24(c_1+c_2)^2+16\right)\leq(\frac{9}{\xi_1}+9\xi_1)(c_1+c_2)$，即，

$$\alpha\leq\frac{1}{\left(\frac{c_1+c_2}{\xi_1}+24\xi_2+\frac{4}{\xi_2}\right)}=\frac{\xi_1}{(c_1+c_2+24\xi_1^2+4)}$$

$$\alpha\leq\min\{\frac{\xi_1}{12(c_1+c_2)},1,\frac{(\frac{5}{\xi_1}+5\xi_1)(c_1+c_2)}{24(c_1+c_2)^2+16}\},$$

当 $\alpha\leq\min\{\frac{\xi_1}{24\ c_1+c_2\ ()^2+24\xi_1^2+16},1\}$ 时足以成立，那么我们有:

$$\mathbb{E}_{t-2\tau}\left\|\bar{\theta}_{t+1}-\theta^*\right\|^2$$

$$
\begin{aligned}
&\leq \mathbb{E}_{t-2\tau}\left\|\bar{\theta}_t-\theta^*\right\|^2+2\alpha\mathbb{E}_{t-2\tau}\left\langle\bar{g}(\bar{\theta}_t),\bar{\theta}_t-\theta^*\right\rangle+4\alpha^2\mathbb{E}_{t-2\tau}\left\|\bar{g}(\bar{\theta}_t)\right\|^2\\
&+30\alpha\xi_1(c_1+c_2)\mathbb{E}_{t-2\tau}\left\|\bar{\theta}_t-\theta^*\right\|^2\\
&+2\left\{8\alpha^2\tau^2c_4^2\mathbb{E}_{t-2\tau}\left[\left\|\bar{\theta}_t-\theta^*\right\|^2\right]+14\alpha^2\tau^2\frac{d_2^2}{NK}+\frac{52L_2^2\alpha^4\tau}{1-\rho^2}\right.\\
&\left.+4\alpha^2c_4^2\tau\sum_{s=0}^{\tau}E_{t-2\tau}[\Delta_{t-s}]+3200\alpha^2c_1^2c_4^2\tau^3\Gamma^2(\epsilon,\epsilon_1)+4\alpha^2c_1^2\tau^2\Gamma^2(\epsilon,\epsilon_1)\right\}\\
&+\frac{9d_2^2}{NK}\alpha^2+36L_2^2\alpha^4+4\alpha^3L_1G^2+2\alpha^3L_2G+\left(\frac{4\alpha}{\xi_2}+2\alpha^3L_1\right)\mathbb{E}_{t-2\tau}[\Delta_{t-\tau}]\\
&+\alpha\left(\frac{4}{\xi_3}+3\xi_1(c_1+c_2)+\frac{4}{\xi_2}+\alpha^2\left(24(c_1+c_2)^2+16\right)\right)\mathbb{E}_{t-2\tau}[\Delta_t]\\
&+2\alpha B(\epsilon,\epsilon_1)G+4\alpha^2B^2(\epsilon,\epsilon_1)+24\alpha^2(c_1+c_2)^2\Gamma^2(\epsilon,\epsilon_1)\\
&+3\alpha\xi_1(c_1+c_2)\Gamma^2(\epsilon,\epsilon_1)+2\alpha^3L_1\Gamma(\epsilon,\epsilon_1)G\\
&\leq \mathbb{E}_{t-2\tau}\left\|\bar{\theta}_t-\theta^*\right\|^2+2\alpha\mathbb{E}_{t-2\tau}\left\langle\bar{g}(\bar{\theta}_t),\bar{\theta}_t-\theta^*\right\rangle+4\alpha^2\mathbb{E}_{t-2\tau}\left\|\bar{g}(\bar{\theta}_t)\right\|^2\\
&+\left(30\alpha\xi_1(c_1+c_2)+16\alpha^2\tau^2c_4^2\right)\mathbb{E}_{t-2\tau}\left\|\bar{\theta}_t-\theta^*\right\|^2\\
&+\frac{9+28\tau^2}{NK}\alpha^2d_2^2+36\left(1+\frac{3\tau}{1-\rho^2}\right)L_2^2\alpha^4+4\alpha^3L_1G^2+2\alpha^3L_2G\\
&+\left(\frac{4\alpha}{\xi_1}+2\alpha^3L_1\right)\mathbb{E}_{t-2\tau}[\Delta_{t-\tau}]+\alpha(\frac{9}{\xi_1}+9\xi_1)(c_1+c_2)\mathbb{E}_{t-2\tau}[\Delta_t]+8\alpha^2c_4^2\tau\sum_{s=0}^{\tau}E_{t-2\tau}[\Delta_{t-s}]\\
&+2\alpha B(\epsilon,\epsilon_1)G+4\alpha^2B^2(\epsilon,\epsilon_1)+24\alpha^2(c_1+c_2)^2\Gamma^2(\epsilon,\epsilon_1)\\
&+3\alpha\xi_1(c_1+c_2)\Gamma^2(\epsilon,\epsilon_1)+2\alpha^3L_1\Gamma(\epsilon,\epsilon_1)G\\
&+6400\alpha^2c_1^2c_4^2\tau^3\Gamma^2(\epsilon,\epsilon_1)+8\alpha^2c_1^2\tau^2\Gamma^2(\epsilon,\epsilon_1)
\end{aligned}
\tag{71}
$$

如果我们选择步长 α，使高阶 $O(\alpha^2)$ 项被一阶项
$O\ \alpha\ ,\quad 4\alpha^2B^2\ \epsilon\ \epsilon_1\ +24\alpha^2\ c_1+c_2\ ^2\Gamma^2\ \epsilon\ \epsilon_1\ +2\alpha^3L_1\Gamma\ \epsilon\ \epsilon_1\ G+6400\alpha^2c_1^2c_4^2\tau^3\Gamma^2\ \epsilon\ \epsilon_1\ +8\alpha^2c_1^2\tau^2\Gamma^2\ \epsilon\ \epsilon_1\ \leq(\ )$ 所主导，即 $(,)\ (\ )\ (,)\ (,)\ (,)\ (,)\ (,)\ (\ )\ (,)$，即，
$2\alpha B\ \epsilon\ \epsilon_1\ G+3\alpha\xi_1\ c_1+c_2\ \Gamma^2\ \epsilon\ \epsilon_1\ ,$

$$
\alpha\leq\min\{\frac{2B(\epsilon,\epsilon_1)G+3\xi_1(c_1+c_2)\Gamma^2(\epsilon,\epsilon_1)}{4B^2(\epsilon,\epsilon_1)+24(c_1+c_2)^2\Gamma^2(\epsilon,\epsilon_1)+2L_1\Gamma(\epsilon,\epsilon_1)G+6400c_1^2c_4^2\tau^3\Gamma^2(\epsilon,\epsilon_1)+8c_1^2\tau^2\Gamma^2(\epsilon,\epsilon_1)},1\},
$$

我们有:

$$
\begin{aligned}
&\mathbb{E}_{t-2\tau}\left\|\bar{\theta}_{t+1}-\theta^*\right\|^2\\
&\leq \mathbb{E}_{t-2\tau}\left\|\bar{\theta}_t-\theta^*\right\|^2+2\alpha\mathbb{E}_{t-2\tau}\left\langle\bar{g}(\bar{\theta}_t),\bar{\theta}_t-\theta^*\right\rangle+4\alpha^2\mathbb{E}_{t-2\tau}\left\|\bar{g}(\bar{\theta}_t)\right\|^2\\
&+\left(30\alpha\xi_1(c_1+c_2)+16\alpha^2\tau^2c_4^2\right)\mathbb{E}_{t-2\tau}\left\|\bar{\theta}_t-\theta^*\right\|^2\\
&+\frac{9+28\tau^2}{NK}\alpha^2d_2^2+36\left(1+\frac{3\tau}{1-\rho^2}\right)L_2^2\alpha^4+4\alpha^3L_1G^2+2\alpha^3L_2G\\
&+\left(\frac{4\alpha}{\xi_1}+2\alpha^3L_1\right)\mathbb{E}_{t-2\tau}[\Delta_{t-\tau}]+\alpha(\frac{9}{\xi_1}+9\xi_1)(c_1+c_2)\mathbb{E}_{t-2\tau}[\Delta_t]+8\alpha^2c_4^2\tau\sum_{s=0}^{\tau}E_{t-2\tau}[\Delta_{t-s}]\\
&+4\alpha B(\epsilon,\epsilon_1)G+6\alpha\xi_1(c_1+c_2)\Gamma^2(\epsilon,\epsilon_1)
\end{aligned}
\tag{72}
$$

根据引理（15），我们得到 $\mathbb{E}_{t-2\tau}[\Delta_t]$、$\mathbb{E}_{t-2\tau}[\Delta_{t-\tau}]$ 和 $\tau\sum_{s=0}^{\tau}E_{t-2\tau}[\Delta_{t-s}]$.的上界。于是有:

$$
E_{t-2\tau}\left\|\bar{\theta}_{t+1}-\theta^*\right\|^2
$$

$$
\begin{aligned}
&\leq \mathbb{E}_{t-2\tau}\left\|\bar{\theta}_t-\theta^*\right\|^2+2\alpha\mathbb{E}_{t-2\tau}\left\langle\bar{g}(\bar{\theta}_t),\bar{\theta}_t-\theta^*\right\rangle+4\alpha^2\mathbb{E}_{t-2\tau}\left\|\bar{g}(\bar{\theta}_t)\right\|^2\\
&+\underbrace{\left(30\alpha\xi_1(c_1+c_2)+16\alpha^2\tau^2c_4^2\right)}_{\mathcal{E}_4}\mathbb{E}_{t-2\tau}\left\|\bar{\theta}_t-\theta^*\right\|^2\\
&+\frac{9+28\tau^2}{NK}\alpha^2d_2^2+\alpha^3\left(36L_2^2+\frac{108\tau}{1-\rho^2}L_2^2+4L_1G^2+2L_2G\right)\\
&+\frac{4\alpha^2}{K\alpha_g^2}\underbrace{\left(\frac{4\alpha}{\xi_1}+2\alpha^3L_1+\alpha(\frac{9}{\xi_1}+9\xi_1)(c_1+c_2)+8\alpha^2c_4^2\tau^2\right)}_{\mathcal{E}_5}\left[c_3^2+\frac{2c_3L_2\rho}{1-\rho}+4c_1^2(K-1)H^2\right]\\
&+4\alpha B(\epsilon,\epsilon_1)G+6\alpha\xi_1(c_1+c_2)\Gamma^2(\epsilon,\epsilon_1)
\end{aligned}
\tag{73}
$$

如果我们选择步长使得 $\mathcal{E}_4=30\alpha\xi_1(c_1+c_2)+16\alpha^2\tau^2c_4^2\leq 32\alpha\xi_1(c_1+c_2)$ 以及 $\mathcal{E}_5=\frac{4\alpha}{\xi_1}+2\alpha^3L_1+\alpha(\frac{9}{\xi_1}+9\xi_1)(c_1+c_2)+8\alpha^2c_4^2\tau^2\leq\alpha(\frac{14}{\xi_1}+14\xi_1)(c_1+c_2)$，即，

$$
\alpha\leq\min\{\frac{\xi_1(c_1+c_2)}{8\tau^2c_4^2},1,\frac{(\frac{1}{\xi_1}+\xi_1)(c_1+c_2)}{2L_1+8c_4^2\tau^2}\},
$$

这足以在 $\alpha\leq\frac{\xi_1}{(c_1+c_2)\,2L_1+8\tau^2c_4^2}$ 成立时满足，因此我们有：

$$
\begin{aligned}
E_{t-2\tau}\left\|\bar{\theta}_{t+1}-\theta^*\right\|^2&\leq\mathbb{E}_{t-2\tau}\left\|\bar{\theta}_t-\theta^*\right\|^2+2\alpha\mathbb{E}_{t-2\tau}\left\langle\bar{g}(\bar{\theta}_t),\bar{\theta}_t-\theta^*\right\rangle+4\alpha^2\mathbb{E}_{t-2\tau}\left\|\bar{g}(\bar{\theta}_t)\right\|^2\\
&+32\alpha\xi_1(c_1+c_2)\mathbb{E}_{t-2\tau}\left\|\bar{\theta}_t-\theta^*\right\|^2\\
&+\frac{9+28\tau^2}{NK}\alpha^2d_2^2+\alpha^3\left(36L_2^2+\frac{108\tau}{1-\rho^2}L_2^2+4L_1G^2+2L_2G\right)\\
&+\frac{4\alpha^3}{K\alpha_g^2}(\frac{14}{\xi_1}+14\xi_1)(c_1+c_2)\left[c_3^2+\frac{2c_3L_2\rho}{1-\rho}+8c_1^2(K-1)H^2\right]\\
&+4\alpha B(\epsilon,\epsilon_1)G+6\alpha\xi_1(c_1+c_2)\Gamma^2(\epsilon,\epsilon_1).
\end{aligned}
\tag{74}
$$

□

I.2.6 参数选择

根据引理 16，我们有：

$$
\begin{aligned}
&\mathbb{E}_{t-2\tau}\left\|\bar{\theta}_{t+1}-\theta^*\right\|^2\\
&\leq(1+32\alpha\xi_1(c_1+c_2))\,\mathbb{E}_{t-2\tau}\left\|\bar{\theta}_t-\theta^*\right\|^2+2\alpha\mathbb{E}_{t-2\tau}\left\langle\bar{g}(\bar{\theta}_t),\bar{\theta}_t-\theta^*\right\rangle+4\alpha^2\mathbb{E}_{t-2\tau}\left\|\bar{g}(\bar{\theta}_t)\right\|^2\\
&+\frac{9+28\tau^2}{NK}\alpha^2d_2^2+\alpha^3\left(36L_2^2+\frac{108\tau}{1-\rho^2}L_2^2+4L_1G^2+2L_2G\right)\\
&+\frac{4\alpha^3}{K\alpha_g^2}(\frac{14}{\xi_1}+14\xi_1)(c_1+c_2)\left[c_3^2+\frac{2c_3L_2\rho}{1-\rho}+8c_1^2(K-1)H^2\right]\\
&+4\alpha B(\epsilon,\epsilon_1)G+6\alpha\xi_1(c_1+c_2)\Gamma^2(\epsilon,\epsilon_1).
\end{aligned}
\tag{75}
$$

命题4. 如果 α 满足引理 16的要求，取 $\xi_1=\frac{1-\gamma\,\bar{\omega})()}{32\;c_1+c_2}$ 和 $\tau=\lceil\tau^{\text{mix}}(\alpha_T^2)_K\rceil$，则有：

$$
\nu_1\mathbb{E}_{t-2\tau}\left\|V_{\bar{\theta}_t}-V_{\theta^*}\right\|_{\bar{D}}^2\leq(\frac{1}{\alpha}-\nu_1)\mathbb{E}_{t-2\tau}\left\|\bar{\theta}_t-\theta^*\right\|^2-\frac{1}{\alpha}\mathbb{E}_{t-2\tau}\left\|\bar{\theta}_{t+1}-\theta^*\right\|^2+\frac{9+28\tau^2}{NK}\alpha d_2^2
$$

$$+\alpha^2\left(36L_2^2+\frac{108\tau}{1-\rho^2}L_2^2+4L_1G^2+2L_2G\right)$$
$$+\frac{\alpha^2c_6}{K}\left[c_3^2+\frac{2c_3L_2\rho}{1-\rho}+8c_1^2(K-1)H^2\right]+4B(\epsilon,\epsilon_1)G+\nu_1\Gamma^2(\epsilon,\epsilon_1) \tag{76}$$

其中 $\nu_1=\frac{\nu}{4}=\frac{(1-\gamma)\bar{\omega}}{4}$ 和 $c_6\triangleq\frac{4}{\alpha_g^2}(\frac{14}{\xi_1}+14\xi_1)(c_1+c_2)$。

证明。将 $\xi_1=\frac{(1-\gamma)\bar{\omega}}{32\ c_1+c_2}$ $(\ )$ $c_6\triangleq\frac{4}{\alpha_g^2}$ $(\frac{14}{\xi_1}+14\xi_1)$ (c_1+c_2) 和 $6\xi_1(c_1+c_2)\le\nu_1$代入方程（75），可得

$$\begin{aligned}
&\mathbb{E}_{t-2\tau}\left\|\bar{\theta}_{t+1}-\theta^*\right\|^2\le\mathbb{E}_{t-2\tau}\left\|\bar{\theta}_t-\theta^*\right\|^2+2\alpha\mathbb{E}_{t-2\tau}\left\langle\bar{g}(\bar{\theta}_t),\bar{\theta}_t-\theta^*\right\rangle+4\alpha^2\mathbb{E}_{t-2\tau}\left\|\bar{g}(\bar{\theta}_t)\right\|^2\\
&\quad+\alpha(1-\gamma)\bar{\omega}\mathbb{E}_{t-2\tau}\left\|\bar{\theta}_t-\theta^*\right\|^2\\
&\quad+\frac{9+28\tau^2}{NK}\alpha^2d_2^2+\alpha^3\left(36L_2^2+\frac{108\tau}{1-\rho^2}L_2^2+4L_1G^2+2L_2G\right)\\
&\quad+\frac{\alpha^3c_6}{K}\left[c_3^2+\frac{2c_3L_2\rho}{1-\rho}+8c_1^2(K-1)H^2\right]\\
&\quad+4\alpha B(\epsilon,\epsilon_1)G+\alpha\nu_1\Gamma^2(\epsilon,\epsilon_1)\\
&\overset{(a)}{\le}\mathbb{E}_{t-2\tau}\left\|\bar{\theta}_t-\theta^*\right\|^2-2\alpha(1-\gamma)\bar{\omega}\mathbb{E}_{t-2\tau}\left\|\bar{\theta}_t-\theta^*\right\|^2+16\alpha^2\mathbb{E}_{t-2\tau}\left\|V_{\bar{\theta}_t}-V_{\theta^*}\right\|_{\bar{D}}^2\\
&\quad+\alpha(1-\gamma)\bar{\omega}\mathbb{E}_{t-2\tau}\left\|\bar{\theta}_t-\theta^*\right\|^2\\
&\quad+\frac{9+28\tau^2}{NK}\alpha^2d_2^2+\alpha^3\left(36L_2^2+\frac{108\tau}{1-\rho^2}L_2^2+4L_1G^2+2L_2G\right)\\
&\quad+\frac{\alpha^3c_6}{K}\left[c_3^2+\frac{2c_3L_2\rho}{1-\rho}+8c_1^2(K-1)H^2\right]\\
&\quad+4\alpha B(\epsilon,\epsilon_1)G+\alpha\nu_1\Gamma^2(\epsilon,\epsilon_1)\\
&=\mathbb{E}_{t-2\tau}\left\|\bar{\theta}_t-\theta^*\right\|^2-\frac{\alpha(1-\gamma)\bar{\omega}}{2}\mathbb{E}_{t-2\tau}\left\|\bar{\theta}_t-\theta^*\right\|^2-\frac{\alpha(1-\gamma)\bar{\omega}}{2}\mathbb{E}_{t-2\tau}\left\|\bar{\theta}_t-\theta^*\right\|^2\\
&\quad+16\alpha^2\mathbb{E}_{t-2\tau}\left\|V_{\bar{\theta}_t}-V_{\theta^*}\right\|_{\bar{D}}^2+\frac{9+28\tau^2}{NK}\alpha^2d_2^2+\alpha^3\left(36L_2^2+\frac{108\tau}{1-\rho^2}L_2^2+4L_1G^2+2L_2G\right)\\
&\quad+\frac{\alpha^3c_6}{K}\left[c_3^2+\frac{2c_3L_2\rho}{1-\rho}+8c_1^2(K-1)H^2\right]\\
&\quad+4\alpha B(\epsilon,\epsilon_1)G+\alpha\nu_1\Gamma^2(\epsilon,\epsilon_1)\\
&\le\mathbb{E}_{t-2\tau}\left\|\bar{\theta}_t-\theta^*\right\|^2-\frac{\alpha(1-\gamma)\bar{\omega}}{2}\mathbb{E}_{t-2\tau}\left\|\bar{\theta}_t-\theta^*\right\|^2-\frac{\alpha(1-\gamma)\bar{\omega}}{2}\mathbb{E}_{t-2\tau}\left\|V_{\bar{\theta}_t}-V_{\theta^*}\right\|_{\bar{D}}^2\\
&\quad+16\alpha^2\mathbb{E}_{t-2\tau}\left\|V_{\bar{\theta}_t}-V_{\theta^*}\right\|_{\bar{D}}^2+\frac{9+28\tau^2}{NK}\alpha^2d_2^2+\alpha^3\left(36L_2^2+\frac{108\tau}{1-\rho^2}L_2^2+4L_1G^2+2L_2G\right)\\
&\quad+\frac{\alpha^3c_6}{K}\left[c_3^2+\frac{2c_3L_2\rho}{1-\rho}+8c_1^2(K-1)H^2\right]\\
&\quad+4\alpha B(\epsilon,\epsilon_1)G+\alpha\nu_1\Gamma^2(\epsilon,\epsilon_1)\\
&\overset{(b)}{\le}\mathbb{E}_{t-2\tau}\left\|\bar{\theta}_t-\theta^*\right\|^2-\frac{\alpha(1-\gamma)\bar{\omega}}{2}\mathbb{E}_{t-2\tau}\left\|\bar{\theta}_t-\theta^*\right\|^2-\frac{\alpha(1-\gamma)\bar{\omega}}{4}\mathbb{E}_{t-2\tau}\left\|V_{\bar{\theta}_t}-V_{\theta^*}\right\|_{\bar{D}}^2\\
&\quad+\frac{9+28\tau^2}{NK}\alpha^2d_2^2+\alpha^3\left(36L_2^2+\frac{108\tau}{1-\rho^2}L_2^2+4L_1G^2+2L_2G\right)\\
&\quad+\frac{\alpha^3c_6}{K}\left[c_3^2+\frac{2c_3L_2\rho}{1-\rho}+8c_1^2(K-1)H^2\right]\\
&\quad+4\alpha B(\epsilon,\epsilon_1)G+\alpha\nu_1\Gamma^2(\epsilon,\epsilon_1)
\end{aligned}$$

$$
\begin{aligned}
&\leq (1-2\alpha\nu_1)\mathbb{E}_{t-2\tau}\left\|\bar{\theta}_t-\theta^*\right\|^2-\alpha\nu_1\mathbb{E}_{t-2\tau}\left\|V_{\bar{\theta}_t}-V_{\theta^*}\right\|_{\bar{D}}^2+\frac{9+28\tau^2}{NK}\alpha^2d_2^2\\
&+\alpha^3\left(36L_2^2+\frac{108\tau}{1-\rho^2}L_2^2+4L_1G^2+2L_2G\right)\\
&+\frac{\alpha^3c_6}{K}\left[c_3^2+\frac{2c_3L_2\rho}{1-\rho}+8c_1^2(K-1)H^2\right]+4\alpha B(\epsilon,\epsilon_1)G+\alpha\nu_1\Gamma^2(\epsilon,\epsilon_1)
\end{aligned}
\tag{77}
$$

其中，（a）由引理 3 以及参数的选择得到；(b) 由 $16\alpha^2\leq\alpha(1-\gamma)\bar{\omega}$ 得到。重新整理各项并利用事实 $1-2\alpha\nu_1\leq1-\alpha\nu_1$，可得：

$$
\begin{aligned}
\alpha\nu_1\mathbb{E}_{t-2\tau}\left\|V_{\bar{\theta}_t}-V_{\theta^*}\right\|_{\bar{D}}^2&\leq(1-\alpha\nu_1)\mathbb{E}_{t-2\tau}\left\|\bar{\theta}_t-\theta^*\right\|^2-\mathbb{E}_{t-2\tau}\left\|\bar{\theta}_{t+1}-\theta^*\right\|^2+\frac{9+28\tau^2}{NK}\alpha^2d_2^2\\
&+\alpha^3\left(36L_2^2+\frac{108\tau}{1-\rho^2}L_2^2+4L_1G^2+2L_2G\right)\\
&+\frac{\alpha^3c_6}{K}\left[c_3^2+\frac{2c_3L_2\rho}{1-\rho}+8c_1^2(K-1)H^2\right]+4\alpha B(\epsilon,\epsilon_1)G+\alpha\nu_1\Gamma^2(\epsilon,\epsilon_1)
\end{aligned}
\tag{78}
$$

然后，两边同时除以 α，即可完成证明。 □

利用这些引理，我们现在可以证明定理 2。

I.3 定理 2的证明。

给定固定的局部步长 $\alpha_l\leq\frac{1}{4\sqrt{2}c_1(K-1)}$、递减的有效步长 $\alpha_t=\frac{8}{\nu(a+t+1)}=\frac{8}{(1-\gamma)\bar{\omega}(a+t+1)}$、递减的全局步长 $\alpha_g^{(t)}=\frac{\alpha_t}{K\alpha_l}$ 以及权重 $w_t=(a+t)$，我们有：

$$
\mathbb{E}\left\|V_{\bar{\theta}_T}-V_{\theta_i^*}\right\|_{\bar{D}}^2\leq\tilde{\mathcal{O}}\left(\frac{\tau^2G^2}{K^2T^2}+\frac{c_{\text{quad}}(\tau)}{\nu^2NKT}+\frac{c_{\text{lin}}(\tau)}{\nu^4KT^2}+\frac{B(\epsilon,\epsilon_1)G}{\nu}+\Gamma^2(\epsilon,\epsilon_1)\right)
\tag{79}
$$

证明。我们取步长 $\alpha_t=\frac{8}{\nu(a+t+1)}=\frac{8}{\nu_1(a+t+1)}$ 用于 $a>0$。此外，我们定义权重 $w_t=(a+t)$，并定义

$$
\tilde{\theta}_T=\frac{1}{W}\sum_{t=1}^{T}w_t\bar{\theta}_t
$$

其中 $W=\sum_{t=1}^{T}w_t\geq\frac{1}{2}T(a+T)$。根据正定二次型的凸性（$\lambda_{\min}(\Phi^T\bar{D}\Phi)\geq\bar{\omega}>0$），我们有

$$
\begin{aligned}
&\nu_1\mathbb{E}\left\|V_{\tilde{\theta}_T}-V_{\theta^*}\right\|_{\bar{D}}^2\\
&\leq\frac{\nu_1}{W}\sum_{t=1}^{T}(a+t)\mathbb{E}\left\|V_{\bar{\theta}_t}-V_{\theta^*}\right\|_{\bar{D}}^2\\
&\leq\frac{\nu_1}{W}\sum_{t=1}^{2\tau-1}(a+t)\mathbb{E}\left\|V_{\bar{\theta}_t}-V_{\theta^*}\right\|_{\bar{D}}^2+\frac{\nu_1}{W}\sum_{t=2\tau}^{T}(a+t)\mathbb{E}\left\|V_{\bar{\theta}_t}-V_{\theta^*}\right\|_{\bar{D}}^2\\
&\leq\nu_1\frac{(2\tau-1)(a+2\tau-1)G^2}{W}+\frac{\nu_1}{W}\sum_{t=2\tau}^{T}(a+t)\mathbb{E}\left\|V_{\bar{\theta}_t}-V_{\theta^*}\right\|_{\bar{D}}^2\\
&\overset{(76)}{\leq}\nu_1\frac{(2\tau-1)(a+2\tau-1)G^2}{W}+\frac{\nu_1(a+2\tau)(a+2\tau+1)G^2}{2W}
\end{aligned}
$$

$$
\begin{aligned}
&+\frac{1}{W}\sum_{t=2\tau}^{T}\left[\frac{(9+28\tau^2)d_2^2}{NK}(a+t)\alpha_t+(a+t)\alpha_t^2\left(36L_2^2+\frac{108\tau}{1-\rho^2}L_2^2+4L_1G^2+2L_2G\right)\right]\\
&+\frac{1}{W}\sum_{t=2\tau}^{T}\frac{(a+t)\alpha^2c_6}{K}\left[c_3^2+\frac{2c_3L_2\rho}{1-\rho}+8c_1^2(K-1)H^2\right]\\
&+\frac{1}{W}\sum_{t=2\tau}^{T}\left[4(a+t)B(\epsilon,\epsilon_1)G+(a+t)\nu_1\Gamma^2(\epsilon,\epsilon_1)\right] \qquad (80)
\end{aligned}
$$

其中 $\left\|V_{\bar{\theta}_{2\tau}}-V_{\theta^*}\right\|_{\bar{D}}^2\le G^2$. 我们知道 $\frac{1}{W}\sum_{t=2\tau}^{T}(a+t)\alpha_t^2\le\frac{1}{W}\sum_{t=1}^{T}(a+t)\nu_1^2\frac{4}{(a+t)^2}\le\frac{8}{\log(a+T)\nu_1^2T^2}$ 并且 $\frac{1}{W}\sum_{t=2\tau}^{T}(a+t)\alpha_t\le\frac{4}{\nu_1T}$. 将这些不等式代入方程 (80)，我们得到：

$$
\begin{aligned}
&\nu_1\mathbb{E}\left\|V_{\bar{\theta}}-V_{\theta^*}\right\|_{\bar{D}}^2\\
&\le\frac{3\nu_1(a+2\tau)(a+2\tau+1)G^2}{2W}+\frac{4(9+28\tau^2)d_2^2}{\nu_1NKT}\\
&+\frac{8\log(a+T)}{\nu_1^2T^2}\left(36L_2^2+\frac{108\tau}{1-\rho^2}L_2^2+4L_1G^2+2L_2G\right)\\
&+\frac{8c_6\log(a+T)}{\nu_1^2T^2K}\left[c_3^2+\frac{2c_3L_2\rho}{1-\rho}+8c_1^2(K-1)H^2\right]\\
&+4B(\epsilon,\epsilon_1)G+\nu_1\Gamma^2(\epsilon,\epsilon_1)\\
&=\frac{3\nu_1(a+2\tau)(a+2\tau+1)G^2}{2W}+\frac{4(9+28\tau^2)d_2^2}{\nu_1NKT}\\
&+\frac{8\log(a+T)}{\nu_1^2T^2K}\underbrace{\left[K\left(36L_2^2+\frac{108\tau}{1-\rho^2}L_2^2+4L_1G^2+2L_2G\right)+c_6\left(c_3^2+\frac{2c_3L_2\rho}{1-\rho}+8c_1^2(K-1)H^2\right)\right]}_{c_{\mathrm{lin}}(\tau)}\\
&+4B(\epsilon,\epsilon_1)G+\nu_1\Gamma^2(\epsilon,\epsilon_1) \qquad (81)
\end{aligned}
$$

其中，$c_{\mathrm{quad}}(\tau)=4d_2^2(9+28\tau^2)$。将等式两边除以 ν_1，把 ν_1 改写为 $\nu(\nu=(1-\gamma)\bar{\omega})$，并注意到 $c_6=\frac{4}{\alpha_g^2}(\frac{14}{\xi_1}+14\xi_1)(\overline{c_1}+c_2)=\mathcal{O}(\frac{1}{\nu})$，可得：

$$
\mathbb{E}\left\|V_{\tilde{\theta}_T}-V_{\theta^*}\right\|_{\bar{D}}^2\le\tilde{\mathcal{O}}\left(\frac{\tau^2G^2}{K^2T^2}+\frac{c_{\mathrm{quad}}(\tau)}{\nu^2NKT}+\frac{c_{\mathrm{lin}}(\tau)}{\nu^4KT^2}+\frac{B(\epsilon,\epsilon_1)G}{\nu}+\Gamma^2(\epsilon,\epsilon_1)\right). \qquad (82)
$$

我们通过使用不等式，$\mathbb{E}\left\|V_{\tilde{\theta}_T}-V_{\theta_i^*}\right\|_{\bar{D}}^2\le2\mathbb{E}\left\|V_{\tilde{\theta}_T}-V_{\theta^*}\right\|_{\bar{D}}^2+2\mathbb{E}\left\|V_{\theta_i^*}-V_{\theta^*}\right\|_{\bar{D}}^2$ 并结合定理1'>中的第三点来完成证明。 □

J 模拟结果

J.1 独立同分布设定的模拟结果

在本小节中，我们给出 FedTD(0) 在独立同分布采样设定下的数值结果，以验证定理4的理论结果。具体而言，首个智能体的马尔可夫决策过程$\mathcal{M}^{(1)}$ 是随机生成的，其状态空间大小为 $n = 100$。其余马尔可夫决策过程是对$\mathcal{M}^{(1)}$ 的扰动，异质性水平分别为$\epsilon = 0.1$ 和 $\epsilon_1 = 0.1$。本地步数设为 $K = 20$. 我们根据运行误差 $e_t = \|\bar{\theta}_t - \theta_1^*\|^2$评估收敛情况。每项实验运行10 次。图3展示了 10 次运行的均值和标准差。

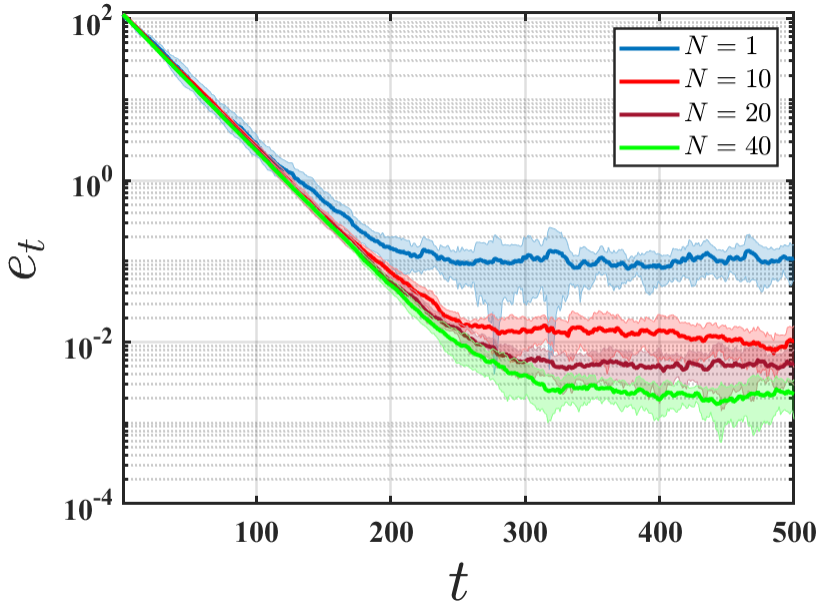


图 3：采用独立同分布采样时， FedTD(0) 在不同智能体数量 N 下的性能。实线表示均值，阴影区域表示十次运行的标准差。

如图 3所示， FedTD(0) 在所有 N取值下均收敛。较大的 N 会降低误差，这与定理 4中的理论分析一致。

J.2 马尔可夫设置下的模拟结果

本小节给出 FedTD(0) 在马尔可夫采样设置下的数值结果，以验证定理 2的理论结果。这里，所有马尔可夫决策过程均按照与独立同分布设置相同的方式生成，并将本地步数设为 $K = 20.$ 。其余参数均与小节 J.1 中的设置相同。

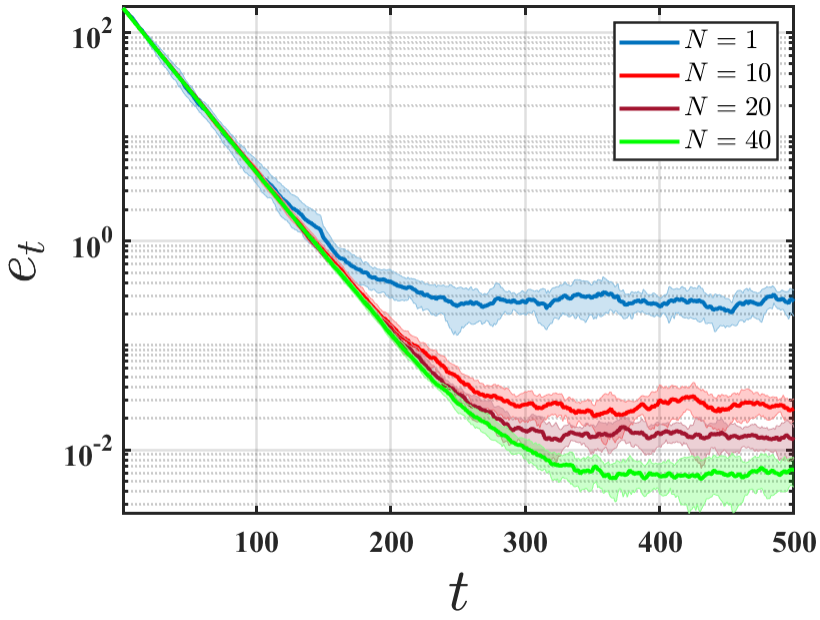
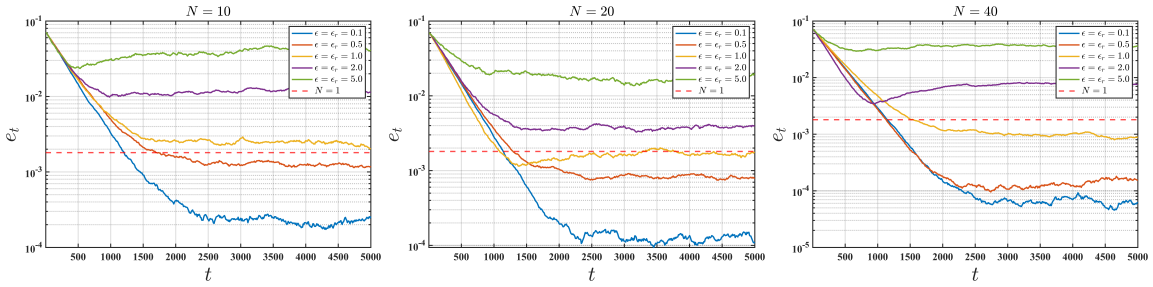


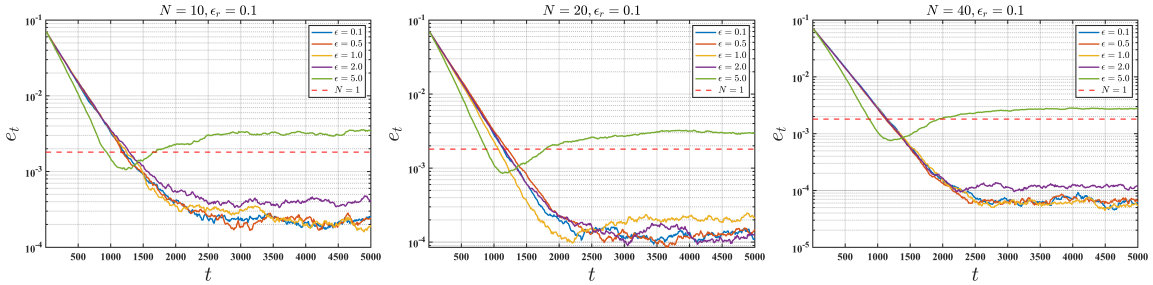
图4：采用马尔可夫采样时， FedTD(0) 随智能体数量 N 变化的性能。实线表示均值，阴影区域表示十次运行的标准差。

如图 4所示， FedTD(0) 对所有 N取值均收敛。增大 N 会减小误差，这与定理2中的理论分析一致。

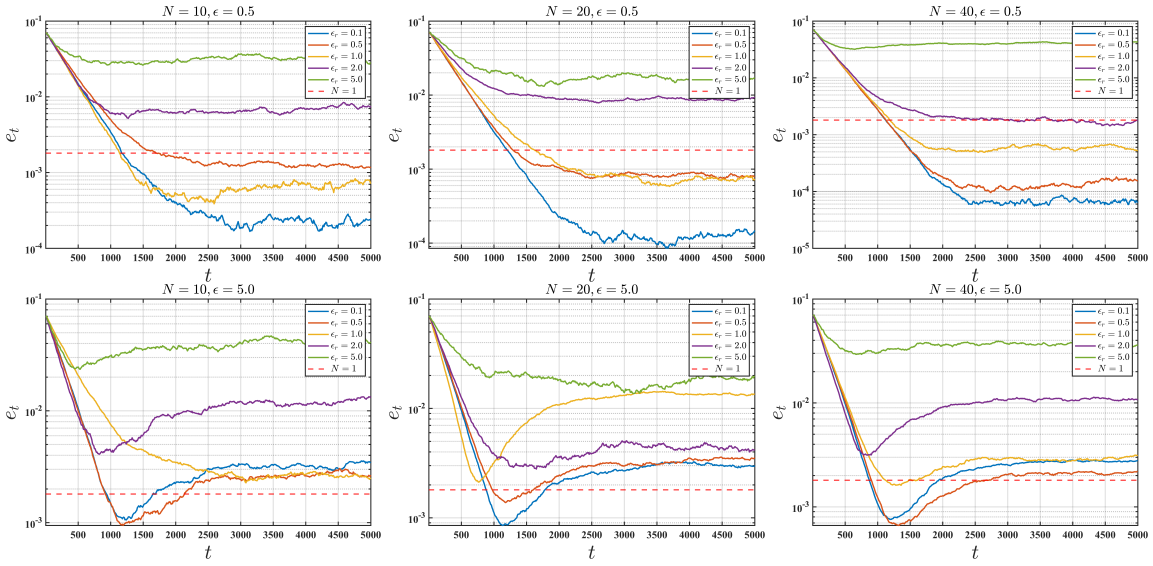
J.3 马尔可夫设定下异质性水平影响的仿真



(a) 不同奖励异质性和马尔可夫核异质性下异质性的影响。



(b) 固定奖励异质性 ϵ_r 并改变奖励异质性 ϵ时异质性的影响。



(c) 固定马尔可夫核异质性 ϵ 并改变奖励异质性 ϵ_r时异质性的影响。

如图 (a) ∼(c)所示，我们可以得出结论：异质性水平的提高会增大 FedTD(0) 收敛到的球的大小，这与定理 2 中的理论分析完全一致。