arXiv:2303.03390v1 [math.OC] 3 Mar 2023

用于离散时间高维随机最优控制问题贝尔曼方程的多项式运行时间非线性蒙特卡罗方法

Christian Beck[1], Arnulf Jentzen[2,3],
KonradKleinberg[4], 和 Thomas Kruse[5]

[1] 数学系，苏黎世联邦理工学院，瑞士；电子邮箱：christian.beck a○yahoo.de[2] 应用数学：分析与数值研究所，明斯特大学，德国；电子邮箱：ajentzen a○uni-muenster.de[3] 数据科学学院及深圳大数据研究院，香港中文大学（深圳），中国；电子邮箱：ajentzen a○cuhk.edu.ch[4] 数学与信息学系，伍珀塔尔大学，德国；电子邮箱：kleinberg a○uni-wuppertal.de[5] 数学与信息学系，伍珀塔尔大学，德国；电子邮箱：tkruse a○uni-wuppertal.de

2023年3月8日

摘要

离散时间随机最优控制 问题和马尔可夫决策过程 (MDP)分别是涉及不确定条件下序贯决策的问题的基础模型，因此构成了强化学习的理论基础。本文研究具有无限时间范围、有限控制集和一般状态空间的马尔可夫决策过程的数值逼近。我们的设定尤其涵盖离散时间马尔可夫过程的无限时间范围最优停止问题。求解马尔可夫决策过程的一个关键工具是贝尔曼方程，它刻画马尔可夫决策过程的价值函数，并确定最优控制策略。结合近期为求解某些非线性偏微分方程而提出的全历史递归多层皮卡逼近方法中的思想，以及 Q-学习中的思想，我们提出一类合适的非线性蒙特卡罗方法，并证明所提出的方法能够克服数值逼近贝尔曼方程解及相关离散时间随机最优控制问题时的维数灾难。

目录

1 引言 **2**

2 函数不动点方程解的存在性与唯一性 7

3 全历史递归多层不动点（MLFP）近似 10

3.1 MLFP 近似的数学描述.10 3.2 MLFP 近似的可测性与分布性质 . . . . . 11 3.3 MLFP 近似的可积性.13 3.4 MLFP 近似的递归误差界.15 3.5 MLFP 近似的非递归误差界.18

4 计算 复杂度分析：MLFP 近似 22

4.1 函数不动点方程的 MLFP 近似.22 4.2 最优控制问题贝尔曼方程的 MLFP 近似 . . . . 25 4.3 最优停止问题贝尔曼方程的 MLFP 近似 . . . 27

1 引言

强化学习 是机器学习领域的一个活跃研究方向，在经济学、工程学、金融、医疗保健、物流和机器人学等涉及序贯决策的许多领域中有着重要应用（例如，参见 Sutton & Barto [48], Bertsekas [11], 和 Bertsekas & Tistsiklis [13] 对该领域及其应用领域的综述）。尤其是，近年来，深度神经网络与强化学习的结合，即深度强化学习，在复杂决策问题中取得了显著成功（例如，参见 Li [37], François-Lavet 等 [22], 和 Arulkumaran 等 [2] 的综述文章）。强化学习的数学基础由随机最优控制 理论，尤其是马尔可夫决策过程 提供（例如，参见 Bertsekas & Shreve [12], Powell [41], 和 Puterman [42]）。

许多强化学习算法的基础，例如时序差分学习（参见，如 Sutton [47]）、 Q学习（参见，如 Watkins [51] 以及 Watkins & Dayan [52]）和 SARSA算法（参见，如 Rummery & Niranjan [44]），都是随机动态规划原理。该原理将给定状态下作出最优决策的问题，归结为求解一个特定的函数方程——所谓的贝尔曼方程。由于维数灾难，高维状态空间中贝尔曼方程的近似解是一个众所周知的难题（参见，如 Bellman [9], Novak & Wozniakowski [40, 第1], 章以及 Novak & Ritter [39]）。

在本文中，我们针对具有无限时间范围和有限控制集的马尔可夫决策过程（MDP），介绍非线性蒙特卡罗方法。这些方法具有多项式可处理性，即算法为近似计算贝尔曼方程的解所需的计算量，关于规定的近似精度的倒数 $1/\epsilon$ 以及底层状态空间的维数 $d \in \mathbb{N} = \{1, 2, 3, \ldots\}$ 至多呈多项式增长（$0, 1$] 。特别地，所提出的方法克服了贝尔曼方程解的数值逼近中的维数灾难。

更正式地说，我们考虑由可测空间 $(\mathbb{X}, \mathcal{X})$（通常为 $(\mathbb{R}^d, \mathcal{B}(\mathbb{R}^d))$，其中 $d \in \mathbb{N}$，但我们的框架也涵盖离散状态空间）、有限控制集 A、折扣因子 $\delta \in (0, 1)$、转移核 κ_a: $\mathbb{X} \times \mathcal{X} \to [0, 1], a \in A$，以及可测且有界的奖励函数 g: $\mathbb{X} \times A \to \mathbb{R}$ 所规定的马尔可夫决策过程。策略 $\mathcal{A}$ 的集合由所有可测

函数 $\alpha\colon \mathbb{X} \to A$ 组成（在适当的假设下，将范围限制为此类非随机平稳策略并不失一般性，参见例如 [12, 第 9] 章）。每个 $\alpha \in \mathcal{A}$ 都定义了一个滤过概率空间 $(\Omega, \mathcal{F}, (\mathcal{F}_k)_{k\in\mathbb{N}_0}, \mathbb{P})$ 上的马尔可夫过程 $X^\alpha = (X_k^\alpha)_{k\in\mathbb{N}_0}$，其状态空间为 $(\mathbb{X}, \mathcal{X})$，满足 $\mathbb{P}[X_{k+1}^\alpha \in B | \mathcal{F}_k] = \kappa_{\alpha(X_k^\alpha)}(B, X_k^\alpha)$，对所有 $k \in \mathbb{N}_0$, $B \in \mathcal{X}$ 均成立。策略 $\alpha \in \mathcal{A}$ 从 $x \in \mathbb{X}$ 开始时的期望收益为

$$J(x,\alpha) = \mathbb{E}\left[\sum_{k=0}^{\infty} \delta^k g(X_k^\alpha, \alpha(X_k^\alpha)) \middle| X_0^\alpha = x\right]. \tag{1}$$

价值函数 $v\colon \mathbb{X} \to \mathbb{R}$ 由 $v(x) = \sup_{\alpha\in\mathcal{A}} J(x,\alpha)$, $x \in \mathbb{X}$给出。动态规划原理确保在适当条件下，v 对所有 $x \in \mathbb{X}$ 均满足

$$v(x) = \max_{a\in A} \{g(x,a) + \delta\, \mathbb{E}[v(X^{x,a})]\}, \tag{2}$$

其中，对于所有 $x \in \mathbb{X}$, $a \in A$，随机变量 $X^{x,a}\colon \Omega \to \mathbb{X}$ 具有分布$\kappa_a(x,\cdot)$（例如，参见 [12, 第 9] 章的证明）。这类马尔可夫决策过程的一个重要子类是具有无限时间范围的离散时间最优停止问题。对于此类问题，控制集 A 由“停止”和“继续”两个元素组成；一旦采取“停止”控制，马尔可夫过程就会跳转到保持状态，且无法离开该状态，也不会产生进一步收益。

科学文献中已经提出并分析了大量形式为（2）的贝尔曼方程的数值逼近方法。贝尔曼方程的确定性数值逼近方法通常会遭受维数灾难。事实上，Chow & Tsitsiklis [17, 18] 证明，对于所有 $d \in \mathbb{N}$，当状态空间 $\mathbb{X}$ 为 d维单位立方体 $[0,1]^d$ 时，确定性数值逼近方法至少需要 $O(\epsilon^{-2d})$次计算操作，才能以精度 $\epsilon \in (0,1]$逼近贝尔曼方程的解。Rust [45] 证明，在某些假设下，通过采用随机算法可以克服这一维数灾难。Rust 的方法是在状态空间中随机抽取网格点，并在该随机网格上执行值迭代。该方法是基于模型的，因为它要求明确知道转移密度。在配套论文 [46]中，Rust 指出，在转移密度具有尖峰的情形下，其方法的收敛性可能会失效。此外，Kristensen 等人 [36] 报告称，随着状态维数增加，该方法的方差会显著上升。Bray [14] 从理论上证实了这一点：他证明，Rust 的方法只有在如下特殊情形下才能克服维数灾难，即该马尔可夫决策过程等价于另一个马尔可夫决策过程，并且其中除比例趋于零的一小部分状态变量外，其余状态变量均表现为历史无关的均匀随机变量。Broadie & Glasserman 在 [15] 和 [16]中设计了适用于最优停止问题的 Rust 方法变体。关于最优停止，文献中还包括各种其他随机算法。例如，对于基于回归的最优停止问题算法，可参见 [38], [49], 和 [50]；对于最优停止问题的基于对偶的算法，可参见 [43], [1], [26], [19], 和 [10]；对于最优停止问题的基于深度学习的算法，可参见 [7], [8],和 [25]。

在本文中，我们介绍了非线性蒙特卡洛算法。这些算法克服了贝尔曼方程逼近中的维数灾难，仅需非常弱的正则性假设，并且在模型无关的意义下运行，即不使用转移核 κ 的显式知识X^x，而只需能够获取独立的单步转移^a，$a \in A$ 对于任意动作 $x \in \mathbb{X}$和状态 Q的实现。我们的方法结合了-学习的思想与近期提出的全历史递归多层 *Picard*（MLP）逼近方法。事实证明，这些方法能够克服某些半线性偏微分方程（PDE）的数值逼近中的维数灾难（例如，参见 [3–6,20,21,24,29–33]）。 Q-学习

基于在（2）中交换期望与最大化运算顺序这一思想。形式上，Q-函数满足：对于所有 $x \in \mathbb{X}$、$a \in A$，有 $Q(x, a) = g(x, a) + \delta\,\mathbb{E}[v(X^{x,a})]$。结合（2）可推出，对于所有 $x \in \mathbb{X}$，有 $v(x) = \max_{a\in A} Q(x, a)$，因此，对于所有 $x \in \mathbb{X}$、$a \in A$，均有

$$Q(x,a) = g(x,a) + \delta\,\mathbb{E}\left[\max_{b\in A} Q(X^{x,a}, b)\right]. \tag{3}$$

在适当条件下，可以证明 Q 是该不动点方程的唯一解，并且不动点迭代序列 Q_n: $\mathbb{X} \times A \to \mathbb{R}$, $n \in \mathbb{N}_0$满足：对所有 $n \in \mathbb{N}$, $x \in \mathbb{X}$, $a \in A$，都有 $Q_0(x, a) = g(x, a)$，且

$$Q_n(x,a) = g(x,a) + \delta\,\mathbb{E}\left[\max_{b\in A} Q_{n-1}(X^{x,a}, b)\right] \tag{4}$$

收敛到 Q。接下来，我们采用多层蒙特卡洛近似中的一个核心思想，将迭代分解为多个层级，从而对所有 $n \in \mathbb{N}_0$, $x \in \mathbb{X}$, $a \in A$ 得到

$$Q_n(x,a) = g(x,a) + \delta\left[\sum_{l=0}^{n-1} \mathbb{E}\left[\left(\max_{b\in A} Q_l(X^{x,a}, b)\right) - \mathbb{I}^{\mathbb{N}}(l)\left(\max_{b\in A} Q_{l-1}(X^{x,a}, b)\right)\right]\right]. \tag{5}$$

在这一望远镜展开中，我们应用 Heinrich [27,28] 和 Giles [23] 的基本思想，在不同层级 $l \in \{1, \ldots, n\}$以不同精度用蒙特卡洛平均近似期望值。$(Q_l)_{l\in\mathbb{N}}$ 的收敛性保证了当 $l \in \{1, \ldots, n\}$ 较大时，Q_l 与 Q_{l-1} 之间的差异很小，因此对于较大的 $l \in \{1, \ldots, n\}$，我们使用较少的蒙特卡洛样本来近似期望值 $\mathbb{E}\,[(\max_{b\in A} Q_l(X^{x,a}, b)) - (\max_{b\in A} Q_{l-1}(X^{x,a}, b))]$，而对于较小的 $l \in \{1, \ldots, n\}$ 则使用较多样本。具体而言，我们固定 $M \in \mathbb{N}$，并使用 M^{n-l} 个独立蒙特卡洛样本来近似第 $l \in \{1, \ldots, n\}$ 层的期望值，由此得到下文式（6）中的全历史递归多层不动点（MLFP）近似方案。

为在本引言部分简要概述本文的贡献，现于下述结果——定理 1.1——中给出本文主要结果定理 4.1 的一个特例。定理 1.1 之后，我们将以文字解释定理 1.1 的陈述及其中出现的数学对象。

定理 *1.1*。 设 A 为一个非空集合，对于每个 $d \in \mathbb{N}$，令 g_d: $\mathbb{R}^d \times A \to \mathbb{R}$ 为$(\mathcal{B}(\mathbb{R}^d) \otimes 2^A)/\mathcal{B}(\mathbb{R})$-可测，假设$\sup_{d\in\mathbb{N}} \sup_{x\in\mathbb{R}^d} \sup_{a\in A} |g_d(x, a)| < \infty$，令$(\Omega, \mathcal{F}, \mathbb{P})$为一个概率空间，令 $\Theta = \cup_{n\in\mathbb{N}} \mathbb{Z}^n$，对于每个 $d \in \mathbb{N}$，令 $\mathcal{F}_d^\theta \subseteq \mathcal{F}$, $\theta \in \Theta$为 $\mathcal{F}$的独立子 σ-代数，对于每个 $d \in \mathbb{N}$，令 $X_d^\theta = (X^{\theta,d\ x\ a}, (\omega))_{(x,a,\omega)\in\mathbb{R}^d\times A\times\Omega}$: $\mathbb{R}^d \times A \times \Omega \to \mathbb{R}^d$, $\theta \in \Theta$为独立同分布随机场，并且对于所有 $d \in \mathbb{N}$, $\theta \in \Theta$满足 X_d^θ为$(\mathcal{B}(\mathbb{R}^d) \otimes 2^A \otimes \mathcal{F}_d^\theta)/\mathcal{B}(\mathbb{R}^d)$-可测，对于每个 $d \in \mathbb{N}$，令 $\delta_d \in [0, 1)$, $\mathcal{R}_d \in [0, \infty)$，令 $M \in \mathbb{N} \cap [\sup_{d\in\mathbb{N}}(4|A|^2(1 - \delta_d)^{-2}), \infty]$，对于每个$d \in \mathbb{N}$，令$\mathcal{Q}_{d,n}^\theta$: $\mathbb{R}^d \times A \times \Omega \to \mathbb{R}$, $n \in \mathbb{N}_0$, $\theta \in \Theta$，并满足对于所有 $n \in \mathbb{N}_0$, $\theta \in \Theta$, $x \in \mathbb{R}^d$, $a \in A$有

$$\begin{aligned}\mathcal{Q}_{d,n}^\theta(x,a) &= g_d(x,a)\\ &+ \sum_{l=0}^{n-1} \frac{\delta_d}{M^{n-l}} \sum_{i=1}^{M^{n-l}} \left[\max_{b\in A}\left\{\mathcal{Q}_{d,l}^{(\theta,l,i)}(X_d^{(\theta,l,i),x,a}, b)\right\} - \mathbb{I}^{\mathbb{N}}(l) \max_{b\in A}\left\{\mathcal{Q}_{d,\max\{l-1,0\}}^{(\theta,-l,i)}(X_d^{(\theta,l,i),x,a}, b)\right\}\right],\end{aligned} \tag{6}$$

并且对每个 $d \in \mathbb{N}$，令 $\mathcal{C}_{d,\,n} \in [0, \infty)$，$n \in \mathbb{N}_0$ 满足对所有 $n \in \mathbb{N}_0$，

$$\mathcal{C}_{d,n} \le \sum_{l=0}^{n-1} M^{n-l}\big(\mathcal{R}_d + \mathcal{C}_{d,l} + \mathcal{C}_{d,\max\{l-1,0\}}\,\mathbb{I}^{\mathbb{N}}(l)\big). \tag{7}$$

Then

(i) 对所有 $d \in \mathbb{N}$，都存在唯一的有界 $(\mathcal{B}(\mathbb{R}^d) \otimes 2^A)/\mathcal{B}(\mathbb{R})$-可测$Q_d\colon \ \mathbb{R}^d \times A \to \mathbb{R}$，其满足对所有 $x \in \mathbb{R}^d$、 $a \in A$，

$$Q_d(x,a) = g_d(x,a) + \delta_d \, \mathbb{E}\left[\max_{b \in A} Q_d(X_d^{0,x,a}, b)\right] \tag{8}$$

and

(ii) 存在 $N\colon (0,1] \to \mathbb{N}$ 以及 $c \in \mathbb{R}$，使得对所有 $d \in \mathbb{N}$, $\varepsilon \in (0,1]$，都有$\mathcal{C}_{d,N_\varepsilon} \leq c\mathcal{R}_d \varepsilon^{-c}$，并且

$$\sup_{x \in \mathbb{R}^d} \left(\mathbb{E}\left[\max_{a \in A} |Q_d(x,a) - \mathcal{Q}^0_{d,N_\varepsilon}(x,a)|^2\right]\right)^{1/2} \leq \varepsilon. \tag{9}$$

定理 1.1 可直接由下文第 4 节中的推论 4.2 得出。反过来，推论 4.2 是定理 4.1 的推论，而定理 4.1 是本文的主要结果。下面我们对上述定理 1.1 中出现的数学对象作一些说明。

在定理 1.1 中，我们在式 (6) 中为一个以状态空间维数 $d \in \mathbb{N}$ 为索引的马尔可夫决策过程序列引入了一个蒙特卡罗型近似算法。粗略地说，要构造式 (6) 中提出的蒙特卡罗型近似算法，我们需要足够多的独立随机变量，并且这些变量的索引集要足够大。定理 1.1 中引入的集合 $\Theta = \cup_{n \in \mathbb{N}} \mathbb{Z}^n$ 提供了这个足够大的索引集。定理 1.1 中的三元组 $(\Omega\ \mathcal{F}, \mathbb{P})$ 是定义这些随机变量的概率空间。在定理 1.1 中，对每个 $d \in \mathbb{N}$，我们都考虑一个状态空间为 $(\mathbb{R}^d, \mathcal{B}(\mathbb{R}^d))$ 的马尔可夫决策过程。我们假设该马尔可夫决策过程序列中的所有元素都具有相同的控制集 A。对于每个$d \in \mathbb{N}$, $x \in \mathbb{R}^d$, $a \in A$，状态空间为 $(\mathbb{R}^d, \mathcal{B}(\mathbb{R}^d))$ 的马尔可夫决策过程所对应的受控马尔可夫链的单步转移由随机变量 $X^{0,x\ a}_{0,d}, : \Omega \to \mathbb{R}^d$ 给出。用马尔可夫决策过程的术语来说，对于每个 $d \in \mathbb{N}$, $a \in A$，转移核 $\kappa_{d,a}\colon \mathbb{R}^d \times \mathcal{B}(\mathbb{R}^d) \to [0,1]$ 因而由 $X^{0,\cdot,\ a}_{0,d}$ 的分布确定。对于每个 $d \in \mathbb{N}$，定理 1.1 第一行中引入的函数 $g_d\colon \mathbb{R}^d \times A \to \mathbb{R}$ 是状态空间为 $(\mathbb{R}^d, \mathcal{B}(\mathbb{R}^d))$ 的马尔可夫决策过程的奖励函数。在定理 1.1 中，我们假设函数 g_d, $d \in \mathbb{N}$, 在 $d \in \mathbb{N}$ 上一致有界。对于每个$d \in \mathbb{N}$，定理 1.1 中引入的实数 $\delta_d \in [0,1)$ 是状态空间为 $(\mathbb{R}^d, \mathcal{B}(\mathbb{R}^d))$ 的马尔可夫决策过程的折扣因子。

定理 1.1 中的第 (i) 项确立了一个基本上众所周知的结果：在上述假设下，对于每个 $d \in \mathbb{N}$，与状态空间为 $(\mathbb{R}^d, \mathcal{B}(\mathbb{R}^d))$ 的马尔可夫决策过程相关的 Q-函数所对应的贝尔曼方程 (2) 都有唯一解 $Q_d\colon \mathbb{R}^d \times A \to \mathbb{R}$。

在定理 1.1 的 (6) 中，我们给出了所提出的 MLFP 近似方案，用于近似贝尔曼方程 (2) 的解。MLFP 近似 $\mathcal{Q}^\theta_{d,n}$, $d \in \mathbb{N}$,$n \in \mathbb{N}_0$, $\theta \in \Theta$ 由维数 $d \in \mathbb{N}$、不动点迭代的次数 $n \in \mathbb{N}_0$ 以及参数 $\theta \in \Theta$ 编号；对于 (6) 中不同位置出现的 MLFP 近似，参数 $\theta \in \Theta$ 各不相同。作为随机输入源，(6) 中提出的 MLFP 近似方案使用随机变量 $X^{\theta,x\ a}_{\theta,d}, : \Omega \to \mathbb{R}^d$, $\theta \in \Theta \setminus \{0\}$, $d \in \mathbb{N}$, $x \in \mathbb{R}^d$, $a \in A$。注意，对于每个 $\theta \in \Theta \setminus \{0\}$, $d \in \mathbb{N}$, $x \in \mathbb{R}^d$, $a \in A$，作为 (6) 中提出的 MLFP 近似方案的随机输入源所使用的随机变量 $X^{\theta,x\ a}_{\theta,d}, : \Omega \to \mathbb{R}^d$ 与用于构造贝尔曼方程 (2) 的随机变量 $X^{0,x\ a}_{0,d}, : \Omega \to \mathbb{R}^d$ 同分布。参数 $\theta \in \Theta$ 确保 (6) 中不同位置出现的 MLFP 近似相互独立，从而确保 (6) 可以通过递归函数调用来实现。定理 1.1 中的自然数 $M \in \mathbb{N} \cap [\sup_{d \in \mathbb{N}} (4|A|^2(1-\delta_d)^{-2})\ \infty]$ 决定了 (6) 中定义 MLFP 近似 $\mathcal{Q}^\theta_{d,n}$, $d \in \mathbb{N}$,$n \in \mathbb{N}_0$, $\theta \in \Theta$ 时所用的蒙特卡洛样本数。假设自然数 M 属于集合

$\mathbb{N} \cap [\sup_{d\in\mathbb{N}}(4|A|^2(1-\delta_d)^{-2}) \ \infty]$ 确保 $\sup_{d\in\mathbb{N}}(4|A|^2(1-\delta_d)^{-2}) < \infty$。这意味着控制集 A 是有限的，并且序列 $(\delta_d)_{d\in\mathbb{N}} \subseteq [0,1)$ 远离 1，即 $\sup_{d\in\mathbb{N}} \delta_d < 1$。注意，通过适当的对应，这一框架也涵盖了这样的情形：对于每个 $d \in \mathbb{N}$，都存在一个独立的控制集 A_d，使得 $\sup_{d\in\mathbb{N}} |A_d| < \infty$。

对于每个 $d \in \mathbb{N}$，定理 1.1 中的非负实数 $\mathcal{R}_d \in [0,\infty)$ 均视为计算任意随机变量 $X^{\theta,}_{d}{}^{x} : \Omega \to (\mathbb{R}^d)^A, \theta \in \Theta, x \in \mathbb{R}^d$ 的一个实现所需计算成本的上界。定理 1.1 的 (7) 中的实数 $\mathcal{C}_{d,}$ $_{n} \in [0,\infty), d \in \mathbb{N}, n \in \mathbb{N}_0$ 对应于 (6) 中 MLFP 近似方案的计算成本。更具体地说，对于每个 $d \in \mathbb{N}, n \in \mathbb{N}_0$，实数 $\mathcal{C}_{d,\ n} \in [0,\infty)$ 表示计算所有随机变量 $X^{\theta,}_{\ d}{}^{x} : \Omega \to (\mathbb{R}^d)^A, \theta \in \Theta, x \in \mathbb{R}^d$ 的实现所需计算成本的上界；这些实现是计算 $\mathcal{Q}^0_{d,\ n}(0,\cdot): \Omega \to \mathbb{R}^A$ 的一个实现所必需的。

定理 1.1 中的 (ii) 项证明，式 (2) 中贝尔曼方程的解可以通过式 (6) 中的 MLFP 近似方案来近似，其计算成本相对于指定近似精度的倒数 $1/\epsilon$ 至多呈多项式增长 $\epsilon \in (0,1]$，并且与计算任意随机变量 $X^{\theta}_{,d}{}^{x} : \Omega \to (\mathbb{R}^d)^A, \theta \in \Theta, x \in \mathbb{R}^d$ 的一次实现所需的计算成本 $\mathcal{R}_d$ 呈线性关系，其中 $d \in \mathbb{N}$ 是相关马尔可夫决策过程的状态空间维数。特别地，如果计算任意随机变量 $X^{\theta}_{,d}{}^{x} : \Omega \to (\mathbb{R}^d)^A, \theta \in \Theta, x \in \mathbb{R}^d$ 的一次实现所需的计算成本相对于相关马尔可夫决策过程的状态空间维数 $d \in \mathbb{N}$ 至多呈多项式增长（实际应用中往往如此），那么式 (6) 中的 MLFP 近似方案便克服了近似式 (2) 中贝尔曼方程解时的维数灾难。不过，我们想指出，如果 $\sup_{d\in\mathbb{N}} \delta_d$ 接近 1，定理 1.1 中 (ii) 项出现的常数 $c \in \mathbb{R}$ 可能变得任意大。因此，尽管式 (6) 中的非线性蒙特卡罗方法已被证明能够克服维数灾难，其计算成本仍可能高到不切实际。

下面我们还将对本文提出的定理 1.1 的推广和变体作一些说明。定理 1.1 考虑的是由欧几里得状态空间 ($d \in \mathbb{N}$，$\mathbb{R}^d(\mathbb{R}^d)$) 的维数 $\mathcal{B}$ 索引的 MDP 序列，而推论 4.2 考虑的是由一般索引集 $\mathfrak{D}$ 索引、具有一般可测状态空间 $(\mathbb{X}_d, \mathcal{X}_d)$ 的一族马尔可夫决策过程，$d \in \mathfrak{D}$。定理 1.1 考虑的是具有有界奖励函数的马尔可夫决策过程，而推论 4.2 允许奖励函数无界。定理 1.1 和推论 4.2 考虑的是马尔可夫决策过程的贝尔曼方程，而定理 4.1 考虑的是更一般的函数不动点方程（详情请参见定理 4.1 中的 (85)）。推论 4.3 证明，MLFP 近似方案的一种变体（详情请参见 (108)）能够克服维数灾难，用于近似最优停止问题的贝尔曼方程解（详情请参见 (110)）。

本文其余部分安排如下。在下文第 2 节中，我们将确立函数不动点方程解的存在性、唯一性和可积性。在下文第 3 节中，我们将介绍函数不动点方程解的 MLFP 近似，研究所引入的 MLFP 近似的可测性、分布和可积性，并为所考虑的函数不动点方程的精确解与所提出的 MLFP 近似之间的 L^2-距离建立递归上界，进而建立非递归上界。在第 4 节中，我们将第 2 节中确立的函数不动点方程解的存在性、唯一性和正则性，与第 3 节中建立的函数不动点方程 MLFP 近似的误差分析相结合，从而对函数不动点方程、马尔可夫决策过程（MDP）和最优停止问题的贝尔曼方程的 MLFP 近似进行计算复杂度分析。

2 函数不动点方程解的存在性与唯一性

定义 2.1. 设 $(\mathbb{X},\mathcal{X})$、$(\mathbb{Y},\mathcal{Y})$ 为非空可测空间，A 为非空集合，且 $\kappa_a\colon \mathbb{X}\times\mathcal{Y}\to[0,1], a\in A$ 满足：对所有 $a\in A, M\in\mathcal{Y}$，$\mathbb{X}\ni x\mapsto\kappa_a(x,M)\in[0,1]$ 是 $\mathcal{X}/\mathcal{B}\big([0,1]\big)$-可测的；并且对所有 $x\in\mathbb{X}, a\in A$，$\mathcal{Y}\ni M\mapsto\kappa_a(x,M)\in[0,1]$ 是 $(\mathbb{Y},\mathcal{Y})$ 上的概率测度。于是称 $(\kappa_a)_{a\in A}$ 是从 $(\mathbb{X},\mathcal{X})$ 到 $(\mathbb{Y},\mathcal{Y})$ 的一族随机核。

引理 *2.2.* 设 $c, L\in[0,\infty)$ 满足 $cL<1$，设 $(\mathbb{X},\mathcal{X})$ 是一个非空可测空间，设 A 是一个非空可数集，设 $\kappa=(\kappa_a)_{a\in A}$ 是从 $(\mathbb{X},\mathcal{X})$ 到 $(\mathbb{X},\mathcal{X})$ 的一族随机核，设 $\mathbb{R}^A=\{r\colon A\to\mathbb{R}\}$，设 $\widetilde{\mathcal{A}}=\bigotimes_{a\in A}\mathcal{B}(\mathbb{R})$，设 $f\colon \mathbb{X}\times\mathbb{R}^A\to\mathbb{R}$ 是 $(\mathcal{X}\otimes\widetilde{\mathcal{A}})/\mathcal{B}(\mathbb{R})$-可测的，设 $\mathfrak{w}\colon \mathbb{X}\to(0,\infty)^A$ 是 $\mathcal{X}/\bigotimes_{a\in A}\mathcal{B}((0,\infty))$-可测的，设

$$\mathfrak{W}=\left\{(u\colon\mathbb{X}\to\mathbb{R}^A)\colon u\ \textit{is}\ \mathcal{X}/\widetilde{\mathcal{A}}\textit{-measurable},\ \sup_{(x,a)\in\mathbb{X}\times A}\big[|\big(\mathfrak{w}(x)\big)(a)|^{-1}|\big(u(x)\big)(a)|\big]<\infty\right\},\quad(10)$$

假设对所有 $x\in\mathbb{X}$, $a\in A$, $r,s\in\mathbb{R}^A$，都有 $|f(x,r)-f(x,s)|\le L\sup_{b\in A}|r(b)-s(b)|$，$\int_{\mathbb{X}}\sup_{b\in A}\big(\mathfrak{w}(y)\big)(b)\kappa_a(x,dy)\le c\big(\mathfrak{w}(x)\big)(a)$，且 $\sup_{(t,b)\in\mathbb{X}\times A}[|\big(\mathfrak{w}(t)\big)(b)|^{-1}\int_{\mathbb{X}}|f(y,0)|\kappa_b(t,dy)]<\infty$。那么存在唯一的 $v\in\mathfrak{W}$，使得对所有 $x\in\mathbb{X}, a\in A$，都有

$$\int_{\mathbb{X}}|f(y,v(y))|\kappa_a(x,dy)<\infty\qquad and\qquad\big(v(x)\big)(a)=\int_{\mathbb{X}}f(y,v(y))\kappa_a(x,dy).\qquad(11)$$

引理 2.2 的证明。设 $\|\cdot\|\colon\ \mathfrak{W}\to[0,\infty)$ 对所有 $u\in\mathfrak{W}$ 满足 $\|u\|=\sup_{(x,a)\in\mathbb{X}\times A}\frac{|\big(u(x)\big)(a)|}{|\big(\mathfrak{w}(x)\big)(a)|}$。注意到 $\|\cdot\|$ 是 $\mathfrak{W}$ 上的范数，且 $(\mathfrak{W},\|\cdot\|)$ 是巴拿赫空间。假设对所有 $x\in\mathbb{X}$、r、$s\in\mathbb{R}^A$，均有 $|f(x,r)-f(x,s)|\le L\sup_{a\in A}|r(a)-s(a)|$，并且假设对所有 $x\in\mathbb{X}$、$a\in A$，均有 $\int_{\mathbb{X}}\sup_{b\in A}\big(\mathfrak{w}(y)\big)(b)\kappa_a(x,dy)\le c\big(\mathfrak{w}(x)\big)(a)$，则对所有 $u\in\mathfrak{W}$、$x\in\mathbb{X}$、$a\in A$，均有

$$\begin{aligned}
&\frac{1}{\big(\mathfrak{w}(x)\big)(a)}\left|\int_{\mathbb{X}}f(y,u(y))\kappa_a(x,dy)\right|\\
&\quad\le\frac{1}{\big(\mathfrak{w}(x)\big)(a)}\int_{\mathbb{X}}|f(y,u(y))-f(y,0)|\kappa_a(x,dy)+\frac{1}{\big(\mathfrak{w}(x)\big)(a)}\int_{\mathbb{X}}|f(y,0)|\kappa_a(x,dy)\\
&\quad\le\frac{L}{\big(\mathfrak{w}(x)\big)(a)}\int_{\mathbb{X}}\sup_{b\in A}|\big(u(y)\big)(b)|\kappa_a(x,dy)+\frac{1}{\big(\mathfrak{w}(x)\big)(a)}\int_{\mathbb{X}}|f(y,0)|\kappa_a(x,dy)\\
&\quad=\frac{L}{\big(\mathfrak{w}(x)\big)(a)}\int_{\mathbb{X}}\sup_{b\in A}\left\{\frac{|\big(u(y)\big)(b)|}{\big(\mathfrak{w}(y)\big)(b)}\big(\mathfrak{w}(y)\big)(b)\right\}\kappa_a(x,dy)+\frac{1}{\big(\mathfrak{w}(x)\big)(a)}\int_{\mathbb{X}}|f(y,0)|\kappa_a(x,dy)\\
&\quad\le\frac{L\|u\|}{\big(\mathfrak{w}(x)\big)(a)}\int_{\mathbb{X}}\sup_{b\in A}\big(\mathfrak{w}(y)\big)(b)\kappa_a(x,dy)+\frac{1}{\big(\mathfrak{w}(x)\big)(a)}\int_{\mathbb{X}}|f(y,0)|\kappa_a(x,dy)\\
&\quad\le cL\|u\|+\frac{1}{\big(\mathfrak{w}(x)\big)(a)}\int_{\mathbb{X}}|f(y,0)|\kappa_a(x,dy).
\end{aligned}\qquad(12)$$

假设 $\sup_{(x,a)\in\mathbb{X}\times A}\big[|\big(\mathfrak{w}(x)\big)(a)|^{-1}\int_{\mathbb{X}}|f(x,0)|\kappa_a(x,dy)\big]<\infty$ 表明，对于所有 $u\in\mathfrak{W}$，都有

$$\begin{aligned}
&\sup_{(x,a)\in\mathbb{X}\times A}\frac{1}{\big(\mathfrak{w}(x)\big)(a)}\left|\int_{\mathbb{X}}f(y,u(y))\kappa_a(x,dy)\right|\\
&\qquad\le cL\|u\|+\sup_{(x,a)\in\mathbb{X}\times A}\frac{1}{\big(\mathfrak{w}(x)\big)(a)}\int_{\mathbb{X}}|f(y,0)|\kappa_a(x,dy)<\infty.
\end{aligned}\qquad(13)$$

[35, 引理 14。20]确保对于所有 $u \in \mathfrak{W}$，映射 $\mathbb{X} \ni x \mapsto \big[A \ni a \mapsto \int_{\mathbb{X}} f(y, u(y))\kappa_a(x, dy) \in \mathbb{R}\big] \in \mathbb{R}^A$是 $\mathcal{X}/\widetilde{\mathcal{A}}$-可测的。令 $\Phi\colon\ \mathfrak{W} \to \mathfrak{W}$为满足对于所有 $u \in \mathfrak{W}$、$(x, a) \in \mathbb{X} \times A$ 的函数，即

$$[\Phi(u)](x)(a) = \int_{\mathbb{X}} f(y, u(y))\kappa_a(x, dy). \tag{14}$$

假设对所有 $x \in \mathbb{X}, r, s \in \mathbb{R}^A$，均有 $|f(x, r) - f(x, s)| \le L \sup_{a\in A} |r(a) - s(a)|$；假设 $\sup_{(x,a)\in\mathbb{X}\times A} \big[|\big(\mathfrak{w}(x)\big)(a)|^{-1} \int_{\mathbb{X}} \big|f(x, 0)\big|\kappa_a(x, dy)\big] < \infty$；以及假设对所有 $(x, a) \in \mathbb{X} \times A$，均有 $\int_{\mathbb{X}} \sup_{b\in A} \big(\mathfrak{w}(y)\big)(b)\kappa_a(x, dy) \le c\big(\mathfrak{w}(x)\big)(a)$，则可确保对所有 $u, v \in \mathfrak{W}, (x, a) \in \mathbb{X} \times A$，均有

$$\begin{aligned}
&\frac{1}{\big(\mathfrak{w}(x)\big)(a)}\Big|[\Phi(u)](x)(a) - [\Phi(v)](x)(a)\Big| \\
&\qquad \le \frac{1}{\big(\mathfrak{w}(x)\big)(a)} \int_{\mathbb{X}} \big|f(y, u(y)) - f(y, v(y))\big|\kappa_a(x, dy) \\
&\qquad \le \frac{L}{\big(\mathfrak{w}(x)\big)(a)} \int_{\mathbb{X}} \sup_{b\in A} |\big(u(y)\big)(b) - \big(v(y)\big)(b)|\kappa_a(x, dy) \\
&\qquad = \frac{L}{\big(\mathfrak{w}(x)\big)(a)} \int_{\mathbb{X}} \sup_{b\in A} \left\{ \frac{|\big(u(y)\big)(b) - \big(v(y)\big)(b)|}{\big(\mathfrak{w}(y)\big)(b)} \big(\mathfrak{w}(y)\big)(b) \right\} \kappa_a(x, dy) \\
&\qquad \le \frac{L\|u - v\|}{\big(\mathfrak{w}(x)\big)(a)} \int_{\mathbb{X}} \sup_{b\in A} \big(\mathfrak{w}(y)\big)(b)\kappa_a(x, dy) \\
&\qquad \le cL\|u - v\|.
\end{aligned} \tag{15}$$

由此可证明，对于所有的 $u, v \in \mathfrak{W}$，均有

$$\big\|\Phi(u) - \Phi(v)\big\| \le cL\|u - v\|. \tag{16}$$

假设 $cL < 1$，可知 Φ 是一个压缩映射。因此，巴拿赫不动点定理表明，存在唯一函数 $v \in \mathfrak{W}$，使得 $v = \Phi(v)$。由于 $v \in \mathfrak{W}$，可知对于所有的 $(x, a) \in \mathbb{X} \times A$，均有 $\int_{\mathbb{X}} |f(y, v(y))|\kappa_a(x, dy) < \infty$。引理 2.2 的证明至此完成。 □

推论 *2.3.* 设 $c, L \in [0, \infty)$ 满足 $cL < 1$，设 $(\mathbb{X}, \mathcal{X})$ 为非空可测空间，设 $(\Omega\ \mathcal{F}, \mathbb{P})$ 为概率空间，设 A 为非空可数集，设 $\mathbb{R}^A = \{r\colon\ A \to \mathbb{R}\}$，设$\widetilde{\mathcal{A}} = \bigotimes_{a\in A} \mathcal{B}(\mathbb{R})$，设 $\mathbb{X}^A = \{q\colon\ A \to \mathbb{X}\}$，设 $\widetilde{\mathcal{X}} = \bigotimes_{a\in A} \mathcal{X}$，设 $X = \big(X^{x,a}(\omega)\big)_{x\in\mathbb{X},\ a\in A,\ \omega\in\Omega}\colon\ \mathbb{X} \times \Omega \to \mathbb{X}^A$ 为 $(\mathcal{X} \otimes \mathcal{F})/\widetilde{\mathcal{X}}$-可测，设 $f\colon\ \mathbb{X} \times \mathbb{R}^A \to \mathbb{R}$ 为 $(\mathcal{X} \otimes \widetilde{\mathcal{A}})/\mathcal{B}(\mathbb{R})$- 可测，设 $\mathfrak{w}\colon\ \mathbb{X} \to (0, \infty)^A$ 为 $\mathcal{X}/\bigotimes_{a\in A} \mathcal{B}\big((0, \infty)\big)$- 可测，设

$$\mathfrak{W} = \big\{(u\colon \mathbb{X} \to \mathbb{R}^A) : u \text{ is } \mathcal{X}/\widetilde{\mathcal{A}}\text{-measurable}, \sup_{(x,a)\in\mathbb{X}\times A} \big[|\big(\mathfrak{w}(x)\big)(a)|^{-1}|\big(u(x)\big)(a)|\big] < \infty\big\},$$

假设对于所有 $(x, a) \in \mathbb{X} \times A$, $r, s \in \mathbb{R}^A$，都有 $|f(x, r) - f(x, s)| \le L \sup_{b\in A} |r(b) - s(b)|$, $\mathbb{E}[\sup_{b\in A} \big(\mathfrak{w}(X^{x,a})\big)(b)] \le c\big(\mathfrak{w}(x)\big)(a)$，且 $\sup_{(y,b)\in\mathbb{X}\times A} \big[|\big(\mathfrak{w}(y)\big)(b)|^{-1}\mathbb{E}[|f(X^{y,b}, 0)|]\big] < \infty$。那么存在一个唯一函数 $v \in \mathfrak{W}$，对所有 $(x, a) \in \mathbb{X} \times A$，满足

$$\mathbb{E}\big[|f(X^{x,a}, v(X^{x,a}))|\big] < \infty, \quad \text{and} \quad \big(v(x)\big)(a) = \mathbb{E}\big[f(X^{x,a}, v(X^{x,a}))\big]. \tag{17}$$

推论 2.3 的证明。令 $\kappa_a\colon \mathbb{X}\times\mathcal{X}\to[0,1]$, $a\in A$, 对所有 $(x,a)\in\mathbb{X}\times A$, $M\in\mathcal{X}$, 满足 $\kappa_a(x,M)=\mathbb{P}[X^{x,a}\in M]$。由于 X 是 $(\mathcal{X}\otimes\mathcal{F})/\mathcal{X}$-可测的，因此对所有 $(x,a)\in\mathbb{X}\times A$，都有 $X^{x,a}\colon\Omega\to\mathbb{X}$ 是 $\mathcal{F}/\mathcal{X}$-可测的。这意味着对所有 $(x,a)\in\mathbb{X}\times A$，$\mathcal{X}\ni M\mapsto\kappa_a(x,M)=\mathbb{P}[X^{x,a}\in M]\in[0,1]$ 是 $(\mathbb{X},\mathcal{X})$ 上的概率测度。由于 X 是 $(\mathcal{X}\otimes\mathcal{F})/\mathcal{X}$-可测的，且由 [35, 定理 14.16] 可得，对所有 $a\in A$, $M\in\mathcal{X}$，都有 $\mathbb{X}\ni x\mapsto\kappa_a(x,M)=\mathbb{P}[X^{x,a}\in M]$ 是 $\mathcal{X}/\mathcal{B}([0,1])$-可测的。因此，$(\kappa_a)_{a\in A}$ 是从 $(\mathbb{X},\mathcal{X})$ 到 $(\mathbb{X},\mathcal{X})$ 的一族随机核。假设对所有 $(x,a)\in\mathbb{X}\times A$，都有 $\mathbb{E}\big[\sup_{b\in A}\big(\mathfrak{w}(X^{x,a})\big)(b)\big]\le c\big(\mathfrak{w}(x)\big)(a)$，且 $\sup_{(y,b)\in\mathbb{X}\times A}\big[\big|\big(\mathfrak{w}(y)\big)(b)\big|^{-1}\mathbb{E}[|f(X^{y,b},0)|]\big]<\infty$，则可得对所有 $(x,a)\in\mathbb{X}\times A$，都有

$$\int_{\mathbb{X}}\sup_{b\in A}\big(\mathfrak{w}(y)\big)(b)\kappa_a(x,dy)=\int_{\mathbb{X}}\sup_{b\in A}\big(\mathfrak{w}(y)\big)(b)\big(X^{x,a}(\mathbb{P})\big)(dy)=\mathbb{E}[\sup_{b\in A}\big(\mathfrak{w}(X^{x,a})\big)(b)]\le c\big(\mathfrak{w}(x)\big)(a),\tag{18}$$

and

$$\begin{aligned}\sup_{(y,b)\in\mathbb{X}\times A}\frac{1}{\big(\mathfrak{w}(y)\big)(b)}\int_{\mathbb{X}}|f(t,0)|\kappa_b(y,dt)&=\sup_{(y,b)\in\mathbb{X}\times A}\frac{1}{\big(\mathfrak{w}(y)\big)(b)}\int_{\mathbb{X}}|f(t,0)|\big(X^{y,b}(\mathbb{P})\big)(dt)\\&=\sup_{(y,b)\in\mathbb{X}\times A}\frac{1}{\big(\mathfrak{w}(y)\big)(b)}\mathbb{E}\big[|f(X^{y,b},0)|\big]<\infty.\end{aligned}\tag{19}$$

引理 2.2 意味着存在唯一函数 $v\in\mathfrak{W}$，使得对所有 $(x,a)\in\mathbb{X}\times A$，均有

$$\mathbb{E}\Big[\big|f\big(X^{x,a},v(X^{x,a})\big)\big|\Big]=\int_{\mathbb{X}}|f(y,v(y))|\big(X^{x,a}(\mathbb{P})\big)(dy)=\int_{\mathbb{X}}|f(y,v(y))|\kappa_a(x,dy)<\infty,\tag{20}$$

and

$$\big(v(x)\big)(a)=\int_{\mathbb{X}}f(y,v(y))\kappa_a(x,dy)=\int_{\mathbb{X}}f(y,v(y))\big(X^{x,a}(\mathbb{P})\big)(dy)=\mathbb{E}\big[f\big(X^{x,a},v(X^{x,a})\big)\big].\tag{21}$$

因此，推论 2.3 的证明完成。 □

引理 *2.4.* 设 $c_f,c_{\mathfrak{w}},L\in[0,\infty)$ 且 $c_{\mathfrak{w}}L<1$，设 $(\mathbb{X},\mathcal{X})$ 为非空可测空间，设 $(\Omega,\mathcal{F},\mathbb{P})$ 为概率空间，设 A 为非空可数集，设 $\mathbb{R}^A=\{r\colon A\to\mathbb{R}\}$，设 $\widetilde{\mathcal{A}}=\bigotimes_{a\in A}\mathcal{B}(\mathbb{R})$，设 $\mathbb{X}^A=\{q\colon A\to\mathbb{X}\}$，设 $\widetilde{\mathcal{X}}=\bigotimes_{a\in A}\mathcal{X}$，设 $X=\big(X^{x,a}(\omega)\big)_{x\in\mathbb{X},\,a\in A,\,\omega\in\Omega}\colon\mathbb{X}\times\Omega\to\mathbb{X}^A$ 为 $(\mathcal{X}\otimes\mathcal{F})/\widetilde{\mathcal{X}}$-可测，设 $f\colon\mathbb{X}\times\mathbb{R}^A\to\mathbb{R}$ 为 $(\mathcal{X}\otimes\widetilde{\mathcal{A}})/\mathcal{B}(\mathbb{R})$-可测，设 $\mathfrak{w}\colon\mathbb{X}\to(0,\infty)^A$ 为 $\mathcal{X}/\bigotimes_{a\in A}\mathcal{B}((0,\infty))$-可测，设 $v\colon\mathbb{X}\to\mathbb{R}^A$ 为 $\mathcal{X}/\widetilde{\mathcal{A}}$-可测，假设对所有 $(x,a)\in\mathbb{X}\times A$, $r,s\in\mathbb{R}^A$，均有 $|f(x,r)-f(x,s)|\le L\sup_{b\in A}|r(b)-s(b)|$, $\sup_{(y,b)\in\mathbb{X}\times A}\big[\big|\big(\mathfrak{w}(y)\big)(b)\big|^{-1}\big|\big(v(y)\big)(b)\big|\big]<\infty$，$\big(\mathbb{E}\big[\sup_{b\in A}\big|\big(\mathfrak{w}(X^{x,a})\big)(b)\big|^2\big]\big)^{\frac12}\le c_{\mathfrak{w}}\big(\mathfrak{w}(x)\big)(a)$, $\big(\mathbb{E}\big[|f(X^{x,a},0)|^2\big]\big)^{\frac12}\le c_f\big(\mathfrak{w}(x)\big)(a)$, 且 $\big(v(x)\big)(a)=\mathbb{E}[f(X^{x,a},v(X^{x,a}))]$。则有

$$\sup_{(x,a)\in\mathbb{X}\times A}\frac{\big|\big(v(x)\big)(a)\big|}{\big(\mathfrak{w}(x)\big)(a)}\le\frac{c_f}{1-c_{\mathfrak{w}}L}\quad\text{and}\quad\sup_{(x,a)\in\mathbb{X}\times A}\frac{\big(\mathbb{E}\big[\sup_{b\in A}\big|\big(v(X^{x,a})\big)(b)\big|^2\big]\big)^{\frac12}}{\big(\mathfrak{w}(x)\big)(a)}\le\frac{c_fc_{\mathfrak{w}}}{1-c_{\mathfrak{w}}L}.\tag{22}$$

引理 *2.4* 的证明。由詹森不等式、三角不等式，以及对所有 $x\in\mathbb{X}$, $r,s\in\mathbb{R}^A$ 均有 $|f(x,r)-f(x,s)|\le L\sup_{a\in A}|r(a)-s(a)|$，可证明对所有

$(x,a)\in\mathbb{X}\times A$ 均有

$$\begin{aligned}\frac{|(v(x))(a)|}{(\mathfrak{w}(x))(a)}&=\frac{\left|\mathbb{E}\left[f(X^{x,a},v(X^{x,a}))\right]\right|}{(\mathfrak{w}(x))(a)}\leq\frac{\mathbb{E}\left[|f(X^{x,a},v(X^{x,a}))|^2\right]^{\frac{1}{2}}}{(\mathfrak{w}(x))(a)}\\&\leq L\frac{\mathbb{E}\left[\sup_{b\in A}|(v(X^{x,a}))(b)|^2\right]^{\frac{1}{2}}}{(\mathfrak{w}(x))(a)}+\frac{\mathbb{E}\left[|f(X^{x,a},0)|^2\right]^{\frac{1}{2}}}{(\mathfrak{w}(x))(a)}\\&\leq\frac{L}{(\mathfrak{w}(x))(a)}\mathbb{E}\left[\sup_{b\in A}\left\{\frac{|(v(X^{x,a}))(b)|}{(\mathfrak{w}(X^{x,a}))(b)}(\mathfrak{w}(X^{x,a}))(b)\right\}^2\right]^{\frac{1}{2}}+c_f\\&\leq\frac{L}{(\mathfrak{w}(x))(a)}\sup_{(y,b)\in\mathbb{X}\times A}\left\{\frac{|(v(y))(b)|}{(\mathfrak{w}(y))(b)}\right\}\mathbb{E}\left[\sup_{b\in A}\mathfrak{w}(X^{x,a})(b)^2\right]^{\frac{1}{2}}+c_f\\&\leq c_{\mathfrak{w}}L\sup_{(y,b)\in\mathbb{X}\times A}\left\{\frac{|(v(y))(b)|}{(\mathfrak{w}(y))(b)}\right\}+c_f.\end{aligned}\tag{23}$$

结合以上结果、假设 $\sup_{(x,a)\in\mathbb{X}\times A}[|(\mathfrak{w}(x))(a)|^{-1}|(v(x))(a)|]<\infty$，以及假设 $c_{\mathfrak{w}}L<1$，可得

$$\sup_{(x,a)\in\mathbb{X}\times A}\frac{|(v(x))(a)|}{(\mathfrak{w}(x))(a)}\leq\frac{c_f}{1-c_{\mathfrak{w}}L}.\tag{24}$$

由此以及假设对所有 $(x,a)\in\mathbb{X}\times A$ $\mathbb{E}\left[|(\mathfrak{w}(X^{x,a}))(b)|^2\right]^{\frac{1}{2}}\leq c_{\mathfrak{w}}(\mathfrak{w}(x))(a)$ 均有 $\sup_{(x,a)\in\mathbb{X}\times A}\frac{|(v(x))(a)|}{(\mathfrak{w}(x))(a)}<\infty$，可知对所有 (x,a) 均有

$$\frac{\left(\mathbb{E}\left[\sup_{b\in A}|(v(X^{x,a}))(b)|^2\right]\right)^{\frac{1}{2}}}{(\mathfrak{w}(x))(a)}\leq\left[\sup_{(y,b)\in\mathbb{X}\times A}\frac{|(v(y))(b)|}{(\mathfrak{w}(y))(b)}\right]\frac{\mathbb{E}\left[\sup_{b\in A}(\mathfrak{w}(X^{x,a}))(b)^2\right]^{\frac{1}{2}}}{(\mathfrak{w}(x))(a)}\leq\frac{c_fc_{\mathfrak{w}}}{1-c_{\mathfrak{w}}L}.\tag{25}$$

对 $\mathbb{X}\times A$ 取上确界可得

$$\sup_{(x,a)\in\mathbb{X}\times A}\frac{\left(\mathbb{E}\left[\sup_{b\in A}|(v(X^{x,a}))(b)|^2\right]\right)^{\frac{1}{2}}}{(\mathfrak{w}(x))(a)}\leq\frac{c_fc_{\mathfrak{w}}}{1-c_{\mathfrak{w}}L}.\tag{26}$$

引理 2.4 的证明至此完成。 □

3 全历史递归多层固定点（MLFP）近似

3.1 MLFP 近似的数学描述

设定 *3.1*。令 $M\in\mathbb{N}$，令 $\Theta=\bigcup_{n\in\mathbb{N}}\mathbb{Z}^n$，令 $(\mathbb{X},\mathcal{X})$ 为非空可测空间，令 $(\Omega,\mathcal{F},\mathbb{P})$ 为概率空间，令 A 为非空集合，令 $\mathbb{R}^A=\{r\colon A\to\mathbb{R}\}$，令 $\widetilde{\mathcal{A}}=\bigotimes_{a\in A}\mathcal{B}(\mathbb{R})$，令 $\|\cdot\|_\infty\colon\mathbb{R}^A\to[0,\infty]$ 满足对所有 $r\in\mathbb{R}^A$ 均有 $\|r\|_\infty=\sup_{a\in A}|r(a)|$，令 $\mathbb{X}^A=\{q\colon A\to\mathbb{X}\}$，令 $\widetilde{\mathcal{X}}=\bigotimes_{a\in A}\mathcal{X}$，令 $f\colon\mathbb{X}\times\mathbb{R}^A\to\mathbb{R}$ 为 $(\mathcal{X}\otimes\widetilde{\mathcal{A}})/\mathcal{B}(\mathbb{R})$-可测，令 $(\mathcal{F}^\theta)_{\theta\in\Theta}$，其中 $\mathcal{F}^\theta\subseteq\mathcal{F}$、$\theta\in\Theta$ 为相互独立的 σ-代数，令 $X^\theta=\left(X^\theta(x,a)(\omega)\right)_{x\in\mathbb{X},\,a\in A,\,\omega\in\Omega}\colon\mathbb{X}\times\Omega\to\mathbb{X}^A$，

$\theta \in \Theta$ 为独立同分布随机场，使得对所有 $\theta \in \Theta$ 均有 X^θ 是 $(\mathcal{X} \otimes \mathcal{F}^\theta)/\widetilde{\mathcal{X}}$-可测的，令 $V_n^\theta \colon \mathbb{X} \times \Omega \to \mathbb{R}^A$、$n \in \mathbb{N}_0$、$\theta \in \Theta$，满足对所有 $n \in \mathbb{N}_0$、$\theta \in \Theta$、$x \in \mathbb{X}$、$a \in A$ 均有

$$\begin{aligned}\big(V_n^\theta(x)\big)(a) = \sum_{l=0}^{n-1} \frac{1}{M^{n-l}} \sum_{i=1}^{M^{n-l}} \Big[& f\big(X^{(\theta,l,i),x,a}, V_l^{(\theta,l,i)}(X^{(\theta,l,i),x,a})\big) \\ & - \mathbb{I}^{\mathbb{N}}(l) f\big(X^{(\theta,l,i),x,a}, V_{\max\{l-1,0\}}^{(\theta,-l,i)}(X^{(\theta,l,i),x,a})\big)\Big]. \end{aligned} \tag{27}$$

3.2 MLFP 近似的可测性与分布性质

引理 3.2. 假设设定 *3.1*。对于所有 $n \in \mathbb{N}_0$, $\theta \in \Theta$， V_n^θ都是$(\mathcal{X} \otimes \sigma(\bigcup_{\eta \in \Theta} \mathcal{F}^{(\theta,\eta)}))/\widetilde{\mathcal{A}}$-可测的；并且对于所有 $n \in \mathbb{N}_0$, $\theta, \vartheta \in \Theta$，有

$$\mathbb{X} \times \Omega \ni (x,\omega) \mapsto \big[A \ni a \mapsto f\big(X^{\theta,x,a}(\omega), V_n^\vartheta(X^{\theta,x,a}(\omega),\omega)\big) \in \mathbb{R}\big] \in \mathbb{R}^A \tag{28}$$

是$(\mathcal{X} \otimes \sigma(\bigcup_{\eta \in \Theta} \mathcal{F}_{(\vartheta,\eta)} \cup \mathcal{F}^\theta))/\widetilde{\mathcal{A}}$-可测的。

引理 *3.2* 的证明。注意到（27）表明，对于所有 $\theta \in \Theta$, $x \in \mathbb{X}$, $a \in A$，都有$\big(V_0^\theta(x)\big)(a) = 0$。因此，对于所有 $\theta \in \Theta$， V_0^θ都是 $(\mathcal{X} \otimes \sigma\big(\bigcup_{\eta \in \Theta} \mathcal{F}^{(\theta,\eta)}\big))/\widetilde{\mathcal{A}}$-可测的。固定 $n \in \mathbb{N}$，并假设对于所有 $l \in \{0,1,\ldots,n-1\}$, $\theta \in \Theta$， V_l^θ都是 $(\mathcal{X} \otimes \sigma\big(\bigcup_{\eta \in \Theta} \mathcal{F}^{(\theta,\eta)}\big))/\widetilde{\mathcal{A}}$-可测的。注意，对于所有 $l \in \{0,1,\ldots,n-1\}$, $\theta \in \Theta$，有

$$\mathbb{X} \times \Omega \ni (x,\omega) \mapsto (x, V_l^\theta(x,\omega)) \in \mathbb{X} \times \mathbb{R}^A \tag{29}$$

是 $(\mathcal{X} \otimes \sigma\big(\bigcup_{\eta \in \Theta} \mathcal{F}^{(\theta,\eta)}\big))/(\mathcal{X} \otimes \widetilde{\mathcal{A}})$-可测的。由于对所有 $\theta \in \Theta$都有 X^θ是 $(\mathcal{X} \otimes \mathcal{F}^\theta)/\widetilde{\mathcal{X}}$-可测的，因此对所有 $\theta \in \Theta$, $a \in A$ 都有 $\mathbb{X} \times \Omega \ni (x,\omega) \mapsto X^{\theta,x,a}(\omega) \in \mathbb{X}$是 $(\mathcal{X} \otimes \mathcal{F}^\theta)/\mathcal{X}$-可测的。这意味着对所有 $\theta, \vartheta \in \Theta$, $a \in A$都有 $\mathbb{X} \times \Omega \ni (x,\omega) \mapsto (X^{\theta,x,a}(\omega)\ \omega) \in \mathbb{X} \times \Omega$ 是 $(\mathcal{X} \otimes \sigma\big(\bigcup_{\eta \in \Theta} \mathcal{F}^{(\vartheta,\eta)} \cup \mathcal{F}^\theta\big))/(\mathcal{X} \otimes \sigma\big(\bigcup_{\eta \in \Theta} \mathcal{F}^{(\vartheta,\eta)}\big))$-可测的。由此以及 (29) 可知，对所有 $l \in \{0,1,\ldots,n-1\}$, $\theta, \vartheta \in \Theta$, $a \in A$都有$\mathbb{X} \times \Omega \ni (x,\omega) \mapsto \big(X^{\theta,x,a}(\omega)\ V_l^\vartheta(X^{\theta,x,a}(\omega)\ \omega)\big) \in \mathbb{X} \times \mathbb{R}^A$是 $(\mathcal{X} \otimes \sigma\big(\bigcup_{\eta \in \Theta} \mathcal{F}^{(\vartheta,\eta)} \cup \mathcal{F}^\theta\big))/(\mathcal{X} \otimes \widetilde{\mathcal{A}})$-可测的。因此对所有 $l \in \{0,1,\ldots,n-1\}$, θ, $\vartheta \in \Theta$, $a \in A$都有 $\mathbb{X} \times \Omega \ni (x,\omega) \mapsto f\big(X^{\theta,x,a}(\omega)\ V_l^\vartheta(X^{\theta,x,a}(\omega)\ \omega)\big) \in \mathbb{R}$是 $(\mathcal{X} \otimes \sigma\big(\bigcup_{\eta \in \Theta} \mathcal{F}^{(\vartheta,\eta)} \cup \mathcal{F}^\theta\big))/\mathcal{B}(\mathbb{R})$-可测的。这意味着对所有 $l \in \{0,1,\ldots,n-1\}$, $\theta, \vartheta \in \Theta$都有

$$\mathbb{X} \times \Omega \ni (x,\omega) \mapsto \big[A \ni a \mapsto f\big(X^{\theta,x,a}(\omega), V_l^\vartheta(X^{\theta,x,a}(\omega),\omega)\big) \in \mathbb{R}\big] \in \mathbb{R}^A \tag{30}$$

是 $(\mathcal{X} \otimes \sigma\big(\bigcup_{\eta \in \Theta} \mathcal{F}^{(\vartheta,\eta)} \cup \mathcal{F}^\theta\big))/\widetilde{\mathcal{A}}$-可测。注意，对于所有 $l \in \{0,1,\ldots,n-1\}$, $i \in \mathbb{N}$, $\theta \in \Theta$, 均有 $\sigma\big(\bigcup_{\eta \in \Theta} \mathcal{F}^{(\theta,l,i,\eta)} \cup \mathcal{F}^{(\theta,l,i)}\big) \subseteq \sigma\big(\bigcup_{\eta \in \Theta} \mathcal{F}^{(\theta,\eta)}\big)$，且 $\sigma\big(\bigcup_{\eta \in \Theta} \mathcal{F}^{(\theta,-l,i,\eta)} \cup \mathcal{F}^{(\theta,-l,i)}\big) \subseteq \sigma\big(\bigcup_{\eta \in \Theta} \mathcal{F}^{(\theta,\eta)}\big)$。由此以及 (30) 可知，对于所有 $\theta \in \Theta$，均有 $V_n^\theta \colon \mathbb{X} \times \Omega \to \mathbb{R}^A$ 是 $(\mathcal{X} \otimes \sigma\big(\bigcup_{\eta \in \Theta} \mathcal{F}^{(\theta,\eta)}\big))/\widetilde{\mathcal{A}}$-可测。因此，归纳法证明了对于所有 $n \in \mathbb{N}_0$, $\theta \in \Theta$，均有 V_n^θ 是 $(\mathcal{X} \otimes \sigma\big(\bigcup_{\eta \in \Theta} \mathcal{F}^{(\theta,\eta)}\big))/\widetilde{\mathcal{A}}$-可测。由此以及 (30) 可知，对于所有 $n \in \mathbb{N}_0$,$\theta, \vartheta \in \Theta$，均有 $\mathbb{X} \times \Omega \ni (x,\omega) \mapsto \big[A \ni a \mapsto f\big(X^{\theta,x,a}(\omega)\ V_n^\vartheta(X^{\theta,x,a}(\omega)\ \omega)\big) \in \mathbb{R}\big] \in \mathbb{R}^A$ 是 $(\mathcal{X} \otimes \sigma(\bigcup_{\eta \in \Theta} \mathcal{F}^{(\vartheta,\eta)} \cup \mathcal{F}^\theta))/\widetilde{\mathcal{A}}$-可测。引理 3.2 的证明至此完成。 □

引理 3.3. 假设设定 *3.1* 成立。那么，对所有 $n \in \mathbb{N}_0$， V_n^θ, $\theta \in \Theta$ 是同分布随机场。

引理 3.3 的证明。对所有 $\theta \in \Theta$, $(x,a) \in \mathbb{X} \times A$，有 $\big(V_0^\theta(x)\big)(a) = 0$。因此，$V_0^\theta, \theta \in \Theta$ 是同分布随机场。固定 $n \in \mathbb{N}$，并假设对所有 $l \in \{0,1,\ldots,n-1\}$， $V_l^\theta, \theta \in \Theta$ 是同分布随机场。引理 3.2 和 [5, 引理 2.6] 保证对所有 $l \in \{0,1,\ldots,n-1\}$, $i \in \mathbb{N}$, $\theta \in \Theta$，有 $\mathbb{X} \times \Omega \ni (x,\omega) \mapsto (V_{\theta,l,i}$ $_l(x,\omega)$ $V_{\theta,-l,i})$ $\max_{\{l-1,0\}}(x,\omega)) \in \mathbb{R}^A \times \mathbb{R}^A$，且 $\mathbb{X} \times \Omega \ni (x,\omega) \mapsto (V_{0,l,i})$ $_l(x,\omega)$ $V_{0,-l,i})$ $\max_{\{l-1,0\}}$ $(x,\omega)) \in \mathbb{R}^A \times \mathbb{R}^A$ 是同分布随机场。这保证对所有 $l \in \{0,1,\ldots,n-1\}$, $i \in \mathbb{N}$ $\theta \in \Theta$ $\mathbb{X} \times \ni x\ \omega \mapsto f\ x\ V_l$ $x\ \omega$ $-_{\mathbb{I}^{\mathbb{N}}}$ $l\ f\ x\ V_{\{l-1,0\}}$ $x\ \omega \in$，有 $\mathbb{R}$ $\mathbb{X} \times \ni x\ \omega \mapsto f\ x\ V_l^{0\,l\,i}$ $x\ \omega$ $l\ f\ x\ V^{0\,-l\,i}$ $x\ \omega \in \mathbb{R}$ $\Omega(,)(,(_{\theta,l,i})(,))()(,(_{\theta,-l,i})\max(,))$，且 $\Omega(,)(,(,,)(,))^{-}\ \mathbb{I}^{\mathbb{N}}()(,(,,)\max_{\{l-1,0\}}(,))$ 是同分布随机场。结合这一点、对所有 $\theta \in \Theta$， X^θ 是 $(\mathcal{X} \otimes \mathcal{F}^\theta)/\widetilde{\mathcal{X}}$-可测这一假设，以及 [5, 引理 2.5]，可得对所有 $l \in \{0,1,\ldots,n-1\}$, $i \in \mathbb{N}$,$\theta \in \Theta$ ，有

$$\begin{aligned}\mathbb{X} \times A \times \Omega \ni (x,a,\omega) \mapsto\ & f\big(X^{(\theta,l,i),x,a}(\omega), V_l^{(\theta,l,i)}\big(X^{(\theta,l,i),x,a}(\omega),\omega\big)\big) \\ & - {}_{\mathbb{I}^{\mathbb{N}}}(l) f\big(X^{(\theta,l,i),x,a}(\omega), V^{(\theta,-l,i)}_{\max\{l-1,0\}}\big(X^{(\theta,l,i),x,a}(\omega),\omega\big)\big) \in \mathbb{R} \end{aligned} \tag{31}$$

and

$$\begin{aligned}\mathbb{X} \times A \times \Omega \ni (x,a,\omega) \mapsto\ & f\big(X^{(0,l,i),x,a}(\omega), V_l^{(0,l,i)}\big(X^{(0,l,i),x,a}(\omega),\omega\big)\big) \\ & - {}_{\mathbb{I}^{\mathbb{N}}}(l) f\big(X^{(0,l,i),x,a}(\omega), V^{(0,-l,i)}_{\max\{l-1,0\}}\big(X^{(0,l,i),x,a}(\omega),\omega\big)\big) \in \mathbb{R} \end{aligned} \tag{32}$$

是同分布随机场。这意味着，对于所有 $l \in \{0,1,\ldots,n-1\}$, $i \in \mathbb{N}$,$\theta \in \Theta$ ，都有

$$\begin{aligned}\mathbb{X} \times \Omega \ni (x,\omega) \mapsto \Big[A \ni a \mapsto\ & f\big(X^{(\theta,l,i),x,a}(\omega), V_l^{(\theta,l,i)}\big(X^{(\theta,l,i),x,a}(\omega),\omega\big)\big) \\ & - {}_{\mathbb{I}^{\mathbb{N}}}(l) f\big(X^{(\theta,l,i),x,a}(\omega), V^{(\theta,-l,i)}_{\max\{l-1,0\}}\big(X^{(\theta,l,i),x,a}(\omega),\omega\big)\big) \in \mathbb{R}\Big] \in \mathbb{R}^A \end{aligned} \tag{33}$$

and

$$\begin{aligned}\mathbb{X} \times \Omega \ni (x,\omega) \mapsto \Big[A \ni a \mapsto\ & f\big(X^{(0,l,i),x,a}(\omega), V_l^{(0,l,i)}\big(X^{(0,l,i),x,a}(\omega),\omega\big)\big) \\ & - {}_{\mathbb{I}^{\mathbb{N}}}(l) f\big(X^{(0,l,i),x,a}(\omega), V^{(0,-l,i)}_{\max\{l-1,0\}}\big(X^{(0,l,i),x,a}(\omega),\omega\big)\big) \in \mathbb{R}\Big] \in \mathbb{R}^A \end{aligned} \tag{34}$$

是同分布随机场。设 $g_k \colon (\mathbb{R}^A)^k \to \mathbb{R}^A$, $k \in \mathbb{N}$，满足对于所有 $k \in \mathbb{N}$, $r_1 \ldots$ $g_k\ r_1\ \ldots\ r_k$ g_k $,,r_k \in \mathbb{R}^A$，都有 $(,,) = \sum_{j=1}^k r_j$。注意，对于所有 $k \in \mathbb{N}$，都有 is$\widetilde{\mathcal{A}}^{\otimes k}/\widetilde{\mathcal{A}}$-可测。设 $U^\theta_{l,\,i} \colon \mathbb{X} \times \Omega \to \mathbb{R}^A$, $l \in \{0,1,\ldots,n-1\}$, $i \in \mathbb{N}$, $\theta \in \Theta$，满足对于所有$l \in \{0,1,\ldots,n-1\}$, $i \in \mathbb{N}$, $\theta \in \Theta$，都有 $(x,\omega) \in \mathbb{X} \times \Omega$

$$\begin{aligned} U^\theta_{l,i}(x,\omega) = \big[A \ni a \mapsto\ & f\big(X^{(\theta,l,i),x,a}(\omega), V_l^{(\theta,l,i)}\big(X^{(\theta,l,i),x,a}(\omega),\omega\big)\big) \\ & - {}_{\mathbb{I}^{\mathbb{N}}}(l) f\big(X^{(\theta,l,i),x,a}(\omega), V^{(\theta,-l,i)}_{\max\{l-1,0\}}\big(X^{(\theta,l,i),x,a}(\omega),\omega\big)\big) \in \mathbb{R}\big]. \end{aligned} \tag{35}$$

引理 3.2 表明，对所有 $l \in \{0,1,\ldots,n-1\}$, $i \in \mathbb{N}$, $\theta \in \Theta$，都有 $U^\theta_{l,i}$ 是 $(\mathcal{X} \otimes \sigma(\bigcup_{\eta\in\Theta} \mathcal{F}(_{\theta,l,i,\eta}) \cup \mathcal{F}^{(\theta,-l,i,\eta)} \cup \mathcal{F}(_{\theta,l,i})))/\widetilde{\mathcal{A}}$-可测的。注意，（27）保证对所有 $(x,\omega) \in \mathbb{X} \times \Omega$, $\theta \in \Theta$，都有 $V_n^\theta(x,\omega) = \sum_{l=0}^{n-1} \frac{1}{M^{n-l}} \sum_{i=1}^{M^{n-l}} U_l^{\theta,}{}_{i}(x,\omega)$。令 $\widetilde{M} = \sum_{j=1}^n M^j \in \mathbb{N}$，令$\backsim Y_n^\theta \colon \mathbb{X} \times \Omega \to (\mathbb{R}^A)^M$, $\theta \in \Theta$，使得对所有 $\theta \in \Theta$，

$$\begin{aligned} Y_n^\theta = \Big(& \frac{1}{M^n} U^\theta_{0,1}, \frac{1}{M^n} U^\theta_{0,2}, \ldots, \frac{1}{M^n} U^\theta_{0,M^n}, \frac{1}{M^{n-1}} U^\theta_{1,1}, \frac{1}{M^{n-1}} U^\theta_{1,2}, \ldots, \frac{1}{M^{n-1}} U^\theta_{1,M^{n-1}}, \ldots \\ & \ldots, \frac{1}{M} U^\theta_{n-1,1}, \frac{1}{M} U^\theta_{n-1,2}, \ldots, \frac{1}{M} U^\theta_{n-1,M} \Big). \end{aligned} \tag{36}$$

注意，对所有 $(x,\omega) \in \mathbb{X}\times\Omega$，$\theta\in\Theta$，都有

$$g_{\widetilde{M}}(Y_n^\theta(x,\omega)) = \sum_{l=0}^{n-1}\frac{1}{M^{n-l}}\sum_{i=1}^{M^{n-l}} U_{l,i}^\theta(x,\omega) = V_n^\theta(x,\omega). \tag{37}$$

对于所有 $l\in\{0,1,\dots,n-1\}$, $i\in\mathbb{N}$, $\theta\in\Theta$，均有 $U_{l,i}^\theta$ 是 $(\mathcal{X}\otimes\sigma(\bigcup_{\eta\in\Theta}\mathcal{F}_{(\theta,l,i,\eta)}\cup\mathcal{F}_{(\theta,-l,i,\eta)}\cup\mathcal{F}_{(\theta,l,i)}))/\widetilde{\mathcal{A}}$-可测；再结合 $(\mathcal{F}^\theta)_{\theta\in\Theta}$ 是相互独立的 σ-代数这一假设，以及引理 2.6]，可证明对于所有 $\theta\in\Theta$，Y_n^θ 和 Y_n^0 是同分布随机场。这意味着对于所有 $\theta\in\Theta$，V_n^θ 和 V_n^0 是同分布随机场。归纳法结合 V_0^θ, $\theta\in\Theta$ 是同分布随机场这一事实，可证明对于所有 $n\in\mathbb{N}_0$，均有 V_n^θ, $\theta\in\Theta$ 是同分布随机场。引理 3.3 的证明至此完成。

□

3.3 MLFP 近似的可积性

引理 *3.4.* 假设设定 *3.1*，假设 A 有限，令 $L\in[0,\infty)$，令 $\mathfrak{w}\colon \mathbb{X}\to(0,\infty)$ 为 $\mathcal{X}/\mathcal{B}((0,\infty))$-可测，假设对所有 $(x,a)\in\mathbb{X}\times A$, $r,s\in\mathbb{R}^A$，都有 $|f(x,r)-f(x,s)|\le L\max_{b\in A}|r(b)-s(b)|$，且 $\big(\mathbb{E}\big[|f(X^{0,x,a},0)|^2+|\mathfrak{w}(X^{0,x,a})|^2\big]\big)^{\frac{1}{2}}\le L\mathfrak{w}(x)$。则对所有 $n\in\mathbb{N}_0$，都有 $\sup_{x\in\mathbb{X}}\big(\mathbb{E}\big[\max_{a\in A}\big|\big(V_n^0(x)\big)(a)\big|^2\big]|\mathfrak{w}(x)|^{-2}\big)<\infty$。

引理 *3.4* 的证明。对于所有 $(x,a)\in\mathbb{X}\times A$, $\theta\in\Theta$，都有 $\big(V_0^\theta(x)\big)(a)=0$。因此有 $\sup_{x\in\mathbb{X}}(\mathbb{E}\big[\|V_0^0(x)\|_\infty^2\big]|\mathfrak{w}(x)|^{-2})<\infty$。固定 $n\in\mathbb{N}$，并假设对于所有 $l\in\{0,1,\dots,n-1\}$，都有 $\sup_{x\in\mathbb{X}}(\mathbb{E}\big[\|V_l^0(x)\|_\infty^2\big]|\mathfrak{w}(x)|^{-2})<\infty$。注意，对于所有 $m\in\mathbb{N}$，都有

$$\sup_{x\in\mathbb{X}}\frac{\mathbb{E}\big[\|V_m^0(x)\|_\infty^2\big]}{|\mathfrak{w}(x)|^2}\le\sup_{x\in\mathbb{X}}\sum_{a\in A}\frac{\mathbb{E}\big[|\big(V_m^0(x)\big)(a)|^2\big]}{|\mathfrak{w}(x)|^2}\le\sum_{a\in A}\sup_{x\in X}\frac{\mathbb{E}\big[|\big(V_m^0(x)\big)(a)|^2\big]}{|\mathfrak{w}(x)|^2}. \tag{38}$$

引理 3.3 和詹森不等式证明，对于所有 $m\in\mathbb{N}$，$(x,a)\in\mathbb{X}\times A$, $\theta\in\Theta$ 都有

$$\begin{aligned}
\mathbb{E}\big[|\big(V_m^0(x)\big)(a)|^2\big] &= \mathbb{E}\big[|\big(V_m^\theta(x)\big)(a)|^2\big]\\
&\le \sum_{l=0}^{m-1}\frac{m}{M^{m-l}}\sum_{i=1}^{M^{m-l}}\mathbb{E}\Big[\Big|f\big(X^{(\theta,l,i),x,a},V_l^{(\theta,l,i)}(X^{(\theta,l,i),x,a})\big)\\
&\qquad\qquad -\mathbb{I}^{\mathbb{N}}(l)f\big(X^{(\theta,l,i),x,a},V_{\max\{l-1,0\}}^{(\theta,-l,i)}(X^{(\theta,l,i),x,a})\big)\Big|^2\Big]\\
&\le \sum_{l=0}^{m-1}\frac{2m}{M^{m-l}}\sum_{i=1}^{M^{m-l}}\mathbb{E}\big[|f\big(X^{(\theta,l,i),x,a},V_l^{(\theta,l,i)}(X^{(\theta,l,i),x,a})\big)|^2\big]\\
&\qquad\qquad +\mathbb{E}\big[|f\big(X^{(\theta,l,i),x,a},V_{\max\{l-1,0\}}^{(\theta,-l,i)}(X^{(\theta,l,i),x,a})\big)|^2\big].
\end{aligned}\tag{39}$$

三角不等式、詹森不等式、对所有 $x\in\mathbb{X}$, $r,s\in\mathbb{R}^A$ 均满足 $|f(x,r)-f(x,s)|\le L\max_{a\in A}|r(a)-s(a)|$ 的假设，以及对所有 $(x,a)\in\mathbb{X}\times A$ 均满足 $\big(\mathbb{E}\big[|f(X^{0,x,a},0)|^2\big]\big)^{\frac{1}{2}}\le L\mathfrak{w}(x)$ 的假设，保证对所有 $m\in\mathbb{N}_0$，$(x,a)\in\mathbb{X}\times A$, $\theta,\vartheta\in\Theta$ 均有

$$\begin{aligned}
\mathbb{E}\big[|f\big(X^{\theta,x,a},V_m^\vartheta(X^{\theta,x,a})\big)|^2\big] &\le 2\mathbb{E}\big[|f\big(X^{\theta,x,a},V_m^\vartheta(X^{\theta,x,a})\big)-f(X^{\theta,x,a},0)|^2\big]+2\mathbb{E}\big[|f(X^{\theta,x,a},0)|^2\big]\\
&\le 2L^2\mathbb{E}\big[\max_{b\in A}|\big(V_m^\vartheta(X^{\theta,x,a})\big)(b)|^2\big]+2L^2|\mathfrak{w}(x)|^2.
\end{aligned}\tag{40}$$

引理 3.3 保证对所有 $m \in \mathbb{N}, \vartheta \in \Theta$， V_m^ϑ 和 V_m^0 同分布。结合这一点以及 A 有限的假设可知，对所有 $m \in \mathbb{N}, \vartheta \in \Theta$， $\mathbb{X} \times \Omega \ni (x, \omega) \mapsto \|V_m^\vartheta(x, w)\|_\infty^2 \in [0, \infty)$ 与 $\mathbb{X} \times \Omega \ni (x, \omega) \mapsto \|V_m^0(x, \omega)\|_\infty^2 \in [0, \infty)$ 同分布。这一点、引理 3.2 、对所有 $\theta \in \Theta$ 均满足 $X^{\theta\frown}$ 是 $(\mathcal{X} \otimes \mathcal{F}^\theta)/\mathcal{X}$-可测的假设、 σ-代数 $\mathcal{F}^\theta, \theta \in \Theta$ 相互独立的假设，以及 [31, 引理 2.2] 意味着对所有 $m \in \mathbb{N}_0$， $(x, a) \in \mathbb{X} \times A, \theta, \vartheta \in \Theta$ 且满足 $\theta \notin \bigcup_{\eta \in \Theta}\{(\vartheta, \eta)\}$ 时，均有

$$\begin{aligned}\mathbb{E}\big[\|V_m^\vartheta(X^{\theta,x,a})\|_\infty^2\big] &= \int_{\mathbb{X}} \mathbb{E}\big[\|V_m^\vartheta(y)\|_\infty^2\big]\big(X^{\theta,x,a}(\mathbb{P})\big)(dy) \\ &= \int_{\mathbb{X}} \mathbb{E}\big[\|V_m^0(y)\|_\infty^2\big]\big(X^{0,x,a}(\mathbb{P})\big)(dy). \end{aligned} \tag{41}$$

这、(40)、对于所有 $(x, a) \in \mathbb{X} \times A$ 均成立的假设，即 $(\mathbb{E}\big[|\mathfrak{w}(X^{0,\ x,\ a})|^2\big])^{\frac{1}{2}} \le L\mathfrak{w}(x)$，以及对于所有 $l \in \{0, 1, \ldots, n-1\}$ 均成立的假设，即 $\sup_{x \in \mathbb{X}}(\mathbb{E}\big[\|V_l^0(x)\|_\infty^2\big]|\mathfrak{w}(x)|^{-2}) < \infty$，证明了对于所有 $l \in \{0, 1, \ldots, n-1\}$， $(x, a) \in \mathbb{X} \times A, \theta, \vartheta \in \Theta$， 满足 $\theta \notin \bigcup_{\eta \in \Theta}\{(\vartheta, \eta)\}$ 的条件

$$\begin{aligned}\frac{\mathbb{E}\big[|f\big(X^{\theta,x,a}, V_l^\vartheta(X^{\theta,x,a})\big)|^2\big]}{|\mathfrak{w}(x)|^2} &\le \frac{2L^2}{|\mathfrak{w}(x)|^2}\int_{\mathbb{X}} \mathbb{E}\big[\|V_l^0(y)\|_\infty^2\big]\big(X^{0,x,a}(\mathbb{P})\big)(dy) + 2L^2 \\ &\le 2L^2\left(\sup_{y \in \mathbb{X}} \frac{\mathbb{E}\big[\|V_l^0(y)\|_\infty^2\big]}{|\mathfrak{w}(y)|^2}\right)\left(\frac{\mathbb{E}\big[|\mathfrak{w}(X^{0,x,a})|^2\big]}{|\mathfrak{w}(x)|^2}\right) + 2L^2 \\ &\le 2L^4\left(\sup_{y \in \mathbb{X}} \frac{\mathbb{E}\big[\|V_l^0(y)\|_\infty^2\big]}{|\mathfrak{w}(y)|^2}\right) + 2L^2. \end{aligned} \tag{42}$$

这确保了对于所有 $l \in \{0, 1, \ldots, n-1\}, \theta, \vartheta \in \Theta$， 满足 $\theta \notin \bigcup_{\eta \in \Theta}\{(\vartheta, \eta)\}$ 的条件

$$\sup_{(x,a) \in \mathbb{X} \times A} \frac{\mathbb{E}\big[|f\big(X^{\theta,x,a}, V_l^\vartheta(X^{\theta,x,a})\big)|^2\big]}{|\mathfrak{w}(x)|^2} \le 2L^4\left(\sup_{y \in \mathbb{X}} \frac{\mathbb{E}\big[\|V_l^0(y)\|_\infty^2\big]}{|\mathfrak{w}(y)|^2}\right) + 2L^2 < \infty. \tag{43}$$

这与（39）表明，对于所有 $a \in A$，都有 $\sup_{x \in \mathbb{X}}\big(\mathbb{E}\big[|\big(V_n^0(x)\big)(a)|^2\big]|\mathfrak{w}(x)|^{-2}\big) < \infty$。这、 A 为有限的假设以及（38）表明，$\sup_{x \in \mathbb{X}}\big(\mathbb{E}\big[\|V_n^0(x)\|_\infty^2\big]|\mathfrak{w}(x)|^{-2}\big) < \infty$。因此，通过归纳法可证明，对于所有 $n \in \mathbb{N}_0$，都有 $\sup_{x \in \mathbb{X}}\big(\mathbb{E}\big[\max_{a \in A}|\big(V_n^0(x)\big)(a)|^2\big]|\mathfrak{w}(x)|^{-2}\big) < \infty$。引理 3.4 的证明至此完成。 □

引理 *3.5.* 假设设定 *3.1*，假设 A 有限，令 $L \in [0, \infty)$，令 $\mathfrak{w}\colon \mathbb{X} \to (0, \infty)$ 为 $\mathcal{X}/\mathcal{B}((0, \infty))$-可测，假设对于所有 $(x, a) \in \mathbb{X} \times A, r, s \in \mathbb{R}^A$，均有 $|f(x, r) - f(x, s)| \le L\max_{b \in A}|r(b) - s(b)|$，$\big(\mathbb{E}\big[|f(X^{0,x,a}, 0)|^2 + |\mathfrak{w}(X^{0,x,a})|^2\big]\big)^{\frac{1}{2}} \le L\mathfrak{w}(x)$。那么对于所有 $n \in \mathbb{N}$，$(x, a) \in \mathbb{X} \times A, \theta \in \Theta$，均有 $\mathbb{E}\big[|f(X^{0,x,a}, V_{n-1}^0(X^{0,x,a}))| + |\big(V_n^\theta(x)\big)(a)|\big] < \infty$，并且 $\mathbb{E}\big[\big(V_n^\theta(x)\big)(a)\big] = \mathbb{E}\big[f(X^{0,x,a}, V_{n-1}^0(X^{0,x,a}))\big]$。

引理 3.5 的证明。假设对于所有 $x \in \mathbb{X}, r, s \in \mathbb{R}^A$，均有 $|f(x, r) - f(x, s)| \le L\max_{a \in A}|r(a) - s(a)|$，结合引理 3.3、引理 3.4、$0 \notin \bigcup_{\eta \in \Theta}\{(0, \eta)\}$这一事实、假设 $\big(\mathbb{E}\big[|f(X^{0,x,a}, 0)|^2\big]\big)^{\frac{1}{2}} \le L\mathfrak{w}(x)$，以及引理 2.2]，可得对于所有 $n \in \mathbb{N}$，$(x, a) \in \mathbb{X} \times A, \theta \in \Theta$，均有 $\mathbb{E}\big[|f(X^{0,x,a}, V_{n-1}^0(X^{0,x,a}))|^2 + |\big(V_n^\theta(x)\big)(a)|^2\big] < \infty$。这保证了对于所有 $n \in \mathbb{N}$，$(x, a) \in \mathbb{X} \times A, \theta \in \Theta$，均有 $\mathbb{E}\big[|f(X^{0,x,a}, V_{n-1}^0(X^{0,x,a}))| + |\big(V_n^\theta(x)\big)(a)|\big] < \infty$。由此，结合引理 3.2、假设 $X^\theta, \theta \in \Theta$ 是独立同分布随机场、引理 3.3，以及引理 2.2]，可得对于所有 $n \in \mathbb{N}_0$，$(x, a) \in \mathbb{X} \times A, \theta, \vartheta \in \Theta$，且满足 $\theta \notin \bigcup_{\eta \in \Theta}\{(\vartheta, \eta)\}$，均有

$$\mathbb{E}\big[f\big(X^{\theta,x,a}, V_n^\vartheta(X^{\theta,x,a})\big)\big] = \mathbb{E}\big[f\big(X^{0,x,a}, V_n^0(X^{0,x,a})\big)\big]. \tag{44}$$

由此可得，对于所有 $n \in \mathbb{N}$，$(x,a) \in \mathbb{X} \times A, \theta \in \Theta$，均有

$$\begin{aligned}\mathbb{E}\big[\big(V_n^\theta(x)\big)(a)\big] &= \sum_{l=0}^{n-1} \frac{1}{M^{n-l}} \sum_{i=1}^{M^{n-l}} \mathbb{E}\big[f\big(X^{(\theta,l,i),x,a}, V_l^{(\theta,l,i)}(X^{(\theta,l,i),x,a})\big)\big] \\ &\qquad - \mathbb{I}^{\mathbb{N}}(l)\mathbb{E}\big[f\big(X^{(\theta,l,i),x,a} V^{(\theta,-l,i)}_{\max\{l-1,0\}}(X^{(\theta,l,i)})\big)\big] \\ &= \sum_{l=0}^{n-1} \mathbb{E}\big[f\big(X^{0,x,a}, V_l^0(X^{0,x,a})\big)\big] - \mathbb{I}^{\mathbb{N}}(l)\mathbb{E}\big[f\big(X^{0,x,a}, V^0_{\max\{l-1,0\}}(X^{0,x,a})\big)\big] \\ &= \mathbb{E}\big[f\big(X^{0,x,a}, V^0_{n-1}(X^{0,x,a})\big)\big]. \end{aligned} \tag{45}$$

引理 3.5 的证明至此完成。 □

3.4 MLFP 近似的递归误差界

引理 *3.6*。假设设定 *3.1*，假设 A 为有限集，令 c、$L \in [0,\infty)$，令 $\mathfrak{w}\colon \mathbb{X} \to (0,\infty)$ 为 $\mathcal{X}/\mathcal{B}((0,\infty))$-可测，令 $v\colon \mathbb{X} \to \mathbb{R}^A$ 为 $\mathcal{X}/\widetilde{\mathcal{A}}$-可测，假设对于所有 $(x,a) \in \mathbb{X} \times A$、$r$、$s \in \mathbb{R}^A$，均有 $|f(x,r) - f(x,s)| \le L \max_{b\in A} |r(b) - s(b)|$、$\big(\mathbb{E}\big[|f(X^{0,x,a},0)|^2 + \|v(X^{0,x,a})\|_\infty^2 + |\mathfrak{w}(X^{0,x,a})|^2\big]\big)^{\frac{1}{2}} \le c\mathfrak{w}(x)$，且 $\big(v(x)\big)(a) = \mathbb{E}[f(X^{0,x,a}, v(X^{0,x,a}))]$。则对于所有 $n \in \mathbb{N}$、$(x,a) \in \mathbb{X} \times A$ $\theta \in \Theta$，有

$$\big|\big(v(x)\big)(a) - \mathbb{E}\big[\big(V_n^\theta(x)\big)(a)\big]\big| \le L\left(\mathbb{E}\Big[\big\|v(X^{0,x,a}) - V^0_{n-1}(X^{0,x,a})\big\|_\infty^2\Big]\right)^{\frac{1}{2}}. \tag{46}$$

引理 *3.6* 的证明。假设对于所有 ($x\ a$ $x\ \ a\ =$ $\mathbb{E}[f\ X^{0\,x\,a}\ v\ X^{0\,x\,a}\ ]$ 3 5, $x \in \mathbb{X}\ \ r\ s \in \mathbb{R}^A$,) $\in \mathbb{X} \times A$，均有 $(v())()$

(,, (,))，引理 。假设对于所有 , ,, 均有$|f(x,r) - f(x,s)| \le L\max_{a\in A}|r(a) - s(a)|$，由引理 3.4 和柯西-施瓦茨不等式可知，对于所有 $n \in \mathbb{N}, (x,a) \in \mathbb{X} \times A, \theta \in \Theta$，均有

$$\begin{aligned}\big|v(x)(a) - \mathbb{E}\big[V_n^\theta(x)(a)\big]\big| &\le \mathbb{E}\big[\big|f\big(X^{0,x,a}, v(X^{0,x,a})\big) - f\big(X^{0,x,a}, V^0_{n-1}(X^{0,x,a})\big)\big|\big] \\ &\le L\mathbb{E}\Big[\big\|v(X^{0,x,a}) - V^0_{n-1}(X^{0,x,a})\big\|_\infty\Big] \\ &\le L\Big(\mathbb{E}\Big[\big\|v(X^{0,x,a}) - V^0_{n-1}(X^{0,x,a})\big\|_\infty^2\Big]\Big)^{\frac{1}{2}}. \end{aligned} \tag{47}$$

至此，引理 3.6 的证明完成。 □

引理 *3.7*。假设设定 *3.1*，假设 A 有限，令 $c, L \in [0,\infty)$，令 $\mathfrak{w}\colon \mathbb{X} \to (0,\infty)$ 为 $\mathcal{X}/\mathcal{B}((0,\infty))$-可测，令 $v\colon \mathbb{X} \to \mathbb{R}^A$ 为 $\mathcal{X}/\widetilde{\mathcal{A}}$-可测，假设对所有 $(x,a) \in \mathbb{X} \times A$, $r, s \in \mathbb{R}^A$，都有 $|f(x,r) - f(x,s)| \le L\max_{b\in A}|r(b) - s(b)|$，$\big(\mathbb{E}\big[|f(X^{0,x,a},0)|^2 + \|v(X^{0,x,a})\|_\infty^2 + |\mathfrak{w}(X^{0,x,a})|^2\big]\big)^{\frac{1}{2}} \le c\mathfrak{w}(x)$，且 $\big(v(x)\big)(a) = \mathbb{E}\big[f(X^{0,x,a}, v(X^{0,x,a}))\big]$。则对所有 $n \in \mathbb{N}$，$(x,a) \in \mathbb{X} \times A$, $\theta \in \Theta$，有

$$\begin{aligned}&\big(\mathrm{Var}\big[\big(V_n^\theta(x)\big)(a)\big]\big)^{\frac{1}{2}} \\ &\le \frac{1}{\sqrt{M^n}}\Big(\mathbb{E}\big[|f(X^{0,x,a},0)|^2\big]\Big)^{\frac{1}{2}} + \mathbb{I}^{[2,\infty)}(n)\frac{L}{\sqrt{M}}\big(\mathbb{E}\big[\|V^0_{n-1}(X^{0,x,a}) - v(X^{0,x,a})\|_\infty^2\big]\big)^{\frac{1}{2}} \\ &\quad + \mathbb{I}^{[2,\infty)}(n)\frac{L\sqrt{M}}{\sqrt{M^n}}\big(\mathbb{E}\big[\|v(X^{0,x,a})\|_\infty^2\big]\big)^{\frac{1}{2}} + \sum_{l=1}^{n-2}\frac{L(1+\sqrt{M})}{\sqrt{M^{n-l}}}\big(\mathbb{E}\big[\|V_l^0(X^{0,x,a}) - v(X^{0,x,a})\|_\infty^2\big)^{\frac{1}{2}}. \end{aligned} \tag{48}$$

引理 *3.7* 的证明。引理 3.4 保证对所有 $n \in \mathbb{N}$，$(x,a) \in \mathbb{X}\times A, \theta \in \Theta$，都有 $\mathrm{Var}\left[\left(V_n^\theta(x)\right)(a)\right] < \infty$。引理 3.2 推出对所有 $l,i \in \mathbb{N}_0$，$(x,a) \in \mathbb{X} \times A, \theta \in \Theta$，都有

$$\begin{aligned}\Omega \ni \omega \mapsto f\big(X^{(\theta,l,i),x,a}(\omega), V_l^{(\theta,l,i)}(X^{(\theta,l,i),x,a}(\omega),\omega)\big) \\ - \mathbb{I}^{\mathbb{N}}(l) f\big(X^{(\theta,l,i),x,a}(\omega), V^{(\theta,-l,i)}_{\max\{l-1,0\}}(X^{(\theta,l,i),x,a}(\omega),\omega)\big) \in \mathbb{R}\end{aligned} \tag{49}$$

是 $\sigma\big(\bigcup_{\eta\in\Theta} \mathcal{F}_{(\theta,\ l,\ i,\ \eta)} \cup \mathcal{F}_{(\theta,\ -l,\ i,\ \eta)} \cup \mathcal{F}^{(\theta,l,i)}\big)/\mathcal{B}(\mathbb{R})$-可测。这表明，对所有 $n \in \mathbb{N}_0$，$(x,a) \in \mathbb{X} \times A$ 均成立：

$$\begin{aligned}\mathrm{Var}\left[\left(V_n^\theta(x)\right)(a)\right] = \sum_{l=0}^{n-1} \frac{1}{(M^{n-l})^2} \sum_{i=1}^{M^{n-l}} \mathrm{Var}\Big[f\big(X^{(\theta,l,i),x,a}, V_l^{(\theta,l,i)}(X^{(\theta,l,i),x,a})\big) \\ - \mathbb{I}^{\mathbb{N}}(l) f\big(X^{(\theta,l,i),x,a}, V^{(\theta,-l,i)}_{\max\{l-1,0\}}(X^{(\theta,l,i),x,a})\big)\Big].\end{aligned} \tag{50}$$

引理 3.2、引理 3.3 和 [5, 引理 2.6] 确保对所有 $l,i \in \mathbb{N}_0, \theta \in \Theta$，以下结论均成立：

$$\begin{aligned}&\mathbb{X}\times\Omega \ni (x,\omega) \mapsto \big(V_l^{(\theta,l,i)}(x,\omega), V^{(\theta,-l,i)}_{\max\{l-1,0\}}(x,\omega)\big) \in \mathbb{R}^A \times \mathbb{R}^A \quad \text{and} \\ &\mathbb{X}\times\Omega \ni (x,\omega) \mapsto \big(V_l^{0}(x,\omega), V^{1}_{\max\{l-1,0\}}(x,\omega)\big) \in \mathbb{R}^A \times \mathbb{R}^A.\end{aligned} \tag{51}$$

$\widetilde{\ }$是同分布随机场。结合这一点以及 f 是 $(\mathcal{X}\otimes\mathcal{A})/\mathcal{B}(\mathbb{R})$-可测这一事实可知，对所有 $l,i \in \mathbb{N}_0$，$\theta \in \Theta$，以下结论均成立：

$$\begin{aligned}&\mathbb{X}\times\Omega \ni (x,\omega) \mapsto f(x, V_l^{(\theta,l,i)}(x,\omega)) - \mathbb{I}^{\mathbb{N}}(l) f(x, V^{(\theta,-l,i)}_{\max\{l-1,0\}}(x,\omega)) \in \mathbb{R} \quad \text{and} \\ &\mathbb{X}\times\Omega \ni (x,\omega) \mapsto f(x, V_l^{0}(x,\omega)) - \mathbb{I}^{\mathbb{N}}(l) f(x, V^{1}_{\max\{l-1,0\}}(x,\omega)) \in \mathbb{R}\end{aligned} \tag{52}$$

是同分布随机场。[5, 引理 2.5] 证明，对所有 $l,i \in \mathbb{N}_0, \theta \in \Theta$，以下结论均成立：

$$\begin{aligned}&\mathbb{X}\times A\times\Omega \ni (x,a,\omega) \mapsto f\big(X^{(\theta,l,i),x,a}(\omega), V_l^{(\theta,l,i)}(X^{(\theta,l,i),x,a}(\omega),\omega)\big) \\ &\qquad - \mathbb{I}^{\mathbb{N}}(l) f\big(X^{(\theta,l,i),x,a}(\omega), V^{(\theta,-l,i)}_{\max\{l-1,0\}}(X^{(\theta,l,i),x,a}(\omega),\omega)\big) \in \mathbb{R} \quad \text{and} \\ &\mathbb{X}\times A\times\Omega \ni (x,a,\omega) \mapsto f\big(X^{0,x,a}(\omega), V_l^{0}(X^{0,x,a}(\omega),\omega)\big) \\ &\qquad - \mathbb{I}^{\mathbb{N}}(l) f\big(X^{0,x,a}(\omega), V^{1}_{\max\{l-1,0\}}(X^{0,x,a}(\omega),\omega)\big) \in \mathbb{R}\end{aligned} \tag{53}$$

是同分布随机场。这意味着对所有 $l,i \in \mathbb{N}_0$，$(x,a) \in \mathbb{X} \times A, \theta \in \Theta$均满足：

$$\begin{aligned}&\mathrm{Var}\Big[f\big(X^{(\theta,l,i),x,a}, V_l^{(\theta,l,i)}(X^{(\theta,l,i),x,a})\big) - \mathbb{I}^{\mathbb{N}}(l) f\big(X^{(\theta,l,i),x,a}, V^{(\theta,-l,i)}_{\max\{l-1,0\}}(X^{(\theta,l,i),x,a})\big)\Big] \\ &= \mathrm{Var}\Big[f\big(X^{0,x,a}, V_l^{0}(X^{0,x,a})\big) - \mathbb{I}^{\mathbb{N}}(l) f\big(X^{0,x,a}, V^{1}_{\max\{l-1,0\}}(X^{0,x,a})\big)\Big].\end{aligned} \tag{54}$$

由此以及（50）可知，对所有 $n \in \mathbb{N}$，$(x,a) \in \mathbb{X} \times A, \theta \in \Theta$ 均有

$$\begin{aligned}\mathrm{Var}[V_n^\theta(x)(a)] &= \sum_{l=0}^{n-1} \frac{1}{M^{n-l}} \mathrm{Var}\Big[f\big(X^{0,x,a}, V_l^{0}(X^{0,x,a})\big) - \mathbb{I}^{\mathbb{N}}(l) f\big(X^{0,x,a}, V^{1}_{\max\{l-1,0\}}(X^{0,x,a})\big)\Big] \\ &\le \sum_{l=0}^{n-1} \frac{1}{M^{n-l}} \mathbb{E}\Big[\big|f\big(X^{0,x,a}, V_l^{0}(X^{0,x,a})\big) - \mathbb{I}^{\mathbb{N}}(l) f\big(X^{0,x,a}, V^{1}_{\max\{l-1,0\}}(X^{0,x,a})\big)\big|^2\Big].\end{aligned} \tag{55}$$

结合这一点、对所有 $(x,a) \in \mathbb{X}\times A$ 均有 $\big(V_0^0(x)\big)(a) = 0$ 这一事实，以及对所有 $n \in \mathbb{N}$, $r_1, r_2, \ldots,$ $r_n \in [0,\infty)$ 均有 $\sqrt{r_1 + r_2 + \cdots + r_n} \leq \sqrt{r_1} + \sqrt{r_2} + \cdots + \sqrt{r_n}$ 这一事实，可证明对所有 $n \in \mathbb{N}$，(x,a) $\in \mathbb{X} \times A$, $\theta \in \Theta$ 均有

$$
\begin{aligned}
&\big(\operatorname{Var}\big[\big(V_n^\theta(x)\big)(a)\big]\big)^{\frac{1}{2}} \\
&\quad\leq \sum_{l=0}^{n-1} \frac{1}{\sqrt{M^{n-l}}}\Big(\mathbb{E}\Big[\big|f\big(X^{0,x,a}, V_l^0(X^{0,x,a})\big) - \mathbb{I}^{\mathbb{N}}(l) f\big(X^{0,x,a}, V_{\max\{l-1,0\}}^1(X^{0,x,a})\big)\big|^2\Big]\Big)^{\frac{1}{2}} \\
&\quad= \sum_{l=1}^{n-1} \frac{1}{\sqrt{M^{n-l}}}\Big(\mathbb{E}\Big[\big|f\big(X^{0,x,a}, V_l^0(X^{0,x,a})\big) - \mathbb{I}^{\mathbb{N}}(l) f\big(X^{0,x,a}, V_{l-1}^1(X^{0,x,a})\big)\big|^2\Big]\Big)^{\frac{1}{2}} \\
&\qquad+ \frac{1}{\sqrt{M^n}}\Big(\mathbb{E}\big[|f(X^{0,x,a},0)|^2\big]\Big)^{\frac{1}{2}}. \qquad (56)
\end{aligned}
$$

对于所有 $x \in \mathbb{X}$, $r, s \in \mathbb{R}^A$，假设 $|f(x,r) - f(x,s)| \leq L \max_{a\in A} |r(a) - s(a)|$ 成立，并且结合三角不等式可知，对于所有 $n \in \mathbb{N}$, $(x,a) \in \mathbb{X} \times A$, $\theta \in \Theta$，有

$$
\begin{aligned}
&\big(\operatorname{Var}\big[\big(V_n^\theta(x)\big)(a)\big]\big)^{\frac{1}{2}} \\
&\quad\leq \frac{1}{\sqrt{M^n}}\Big(\mathbb{E}\big[|f(X^{0,x,a},0)|^2\big]\Big)^{\frac{1}{2}} + \sum_{l=1}^{n-1} \frac{L}{\sqrt{M^{n-l}}}\Big(\mathbb{E}\Big[\big\|V_l^0(X^{0,x,a}) - V_{l-1}^1(X^{0,x,a})\big\|_\infty^2\Big]\Big)^{\frac{1}{2}} \\
&\quad\leq \sum_{l=1}^{n-1} \frac{L}{\sqrt{M^{n-l}}}\Big[\big(\mathbb{E}\big[\|V_l^0(X^{0,x,a}) - v(X^{0,x,a})\|_\infty^2\big]\big)^{\frac{1}{2}} + \big(\mathbb{E}\big[\|v(X^{0,x,a}) - V_{l-1}^1(X^{0,x,a})\|_\infty^2\big]\big)^{\frac{1}{2}}\Big] \\
&\qquad+ \frac{1}{\sqrt{M^n}}\Big(\mathbb{E}\big[|f(X^{0,x,a},0)|^2\big]\Big)^{\frac{1}{2}}. \qquad (57)
\end{aligned}
$$

对于所有 $l \in \mathbb{N}_0$, $(x,a) \in \mathbb{X} \times A$，$\Omega \ni \omega \mapsto V_l^0(X^{0,x,a}(\omega)\ \omega) \in \mathbb{R}^A$ 和 $\Omega \ni \omega \mapsto V_l^1(X^{0,x,a}(\omega)$ $\omega) \in \mathbb{R}^A$ 同分布这一事实，确保了对于所有 $l \in \mathbb{N}_0$, $(x,a) \in \mathbb{X} \times A$，$\Omega \ni \omega \mapsto \|v(X^{0,x,a}(\omega))$ $- V_l^0(X^{0,x,a}(\omega)\ \omega)\|_\infty^2 \in [0,\infty)$ 和 $\Omega \ni \omega \mapsto \|v(X^{0,x,a}(\omega)) - V_l^1(X^{0,x,a}(\omega)\ \omega)\|_\infty^2 \in [0,\infty)$ 同分布。由此以及 (57) 可得，对于所有$n \in \mathbb{N}$, $(x,a) \in \mathbb{X} \times A$, $\theta \in \Theta$，有

$$
\begin{aligned}
&\big(\operatorname{Var}\big[\big(V_n^\theta(x)\big)(a)\big]\big)^{\frac{1}{2}} \\
&\leq \sum_{l=1}^{n-1} \frac{L}{\sqrt{M^{n-l}}}\Big[\big(\mathbb{E}\big[\|V_l^0(X^{0,x,a}) - v(X^{0,x,a})\|_\infty^2\big]\big)^{\frac{1}{2}} + \big(\mathbb{E}\big[\|v(X^{0,x,a}) - V_{l-1}^0(X^{0,x,a})\|_\infty^2\big]\big)^{\frac{1}{2}}\Big] \\
&\quad+ \frac{1}{\sqrt{M^n}}\Big(\mathbb{E}\big[|f(X^{0,x,a},0)|^2\big]\Big)^{\frac{1}{2}} \\
&= \frac{1}{\sqrt{M^n}}\Big(\mathbb{E}\big[|f(X^{0,x,a},0)|^2\big]\Big)^{\frac{1}{2}} + \sum_{l=1}^{n-1} \frac{L}{\sqrt{M^{n-l}}}\big(\mathbb{E}\big[\|V_l^0(X^{0,x,a}) - v(X^{0,x,a})\|_\infty^2\big]\big)^{\frac{1}{2}} \\
&\quad+ \sum_{l=0}^{n-2} \frac{L\sqrt{M}}{\sqrt{M^{n-l}}}\big(\mathbb{E}\big[\|V_l^0(X^{0,x,a}) - v(X^{0,x,a})\|_\infty^2\big]\big)^{\frac{1}{2}} \\
&= \frac{1}{\sqrt{M^n}}\Big(\mathbb{E}\big[|f(X^{0,x,a},0)|^2\big]\Big)^{\frac{1}{2}} + \mathbb{I}^{[2,\infty)}(n)\frac{L}{\sqrt{M}}\big(\mathbb{E}\big[\|V_{n-1}^0(X^{0,x,a}) - v(X^{0,x,a})\|_\infty^2\big]\big)^{\frac{1}{2}} \\
&\quad+ \mathbb{I}^{[2,\infty)}(n)\frac{L\sqrt{M}}{\sqrt{M^n}}\big(\mathbb{E}\big[\|v(X^{0,x,a})\|_\infty^2\big]\big)^{\frac{1}{2}} + \sum_{l=1}^{n-2} \frac{L(1+\sqrt{M})}{\sqrt{M^{n-l}}}\big(\mathbb{E}\big[\|V_l^0(X^{0,x,a}) - v(X^{0,x,a})\|_\infty^2\big)^{\frac{1}{2}}. \qquad (58)
\end{aligned}
$$

至此，引理 3.7 的证明完成。 □

3.5 MLFP 近似的非递归误差界

命题 *3.8*。假设设定 *3.1*，假设 A 是有限的，令 $c_f, c_v, c_{\mathfrak{w}}\, L \in [0,\infty)$，令 $\mathfrak{w}\colon \mathbb{X} \to (0,\infty)$ 为 $\mathcal{X}/\mathcal{B}((0,\infty))$-可测，令 $v\colon \mathbb{X} \to \mathbb{R}^A$ 为 $\mathcal{X}/\widetilde{\mathcal{A}}$-可测，令 $c = \frac{3}{2} max\left\{\frac{c_v}{c_{\mathfrak{w}}}, c_v L + c_f, \frac{|A| c_f}{c_{\mathfrak{w}}|A|L+1}\right\}$，假设对所有 $(x,a) \in \mathbb{X}\times A,\ r,s \in \mathbb{R}^A$，均有 $|f(x,r) - f(x,s)| \le L\, max_{b\in A}|r(b) - s(b)|$，

$\left(\mathbb{E}\left[|f(X^{0,x,a}, 0)|^2\right]\right)^{\frac{1}{2}} \le c_f \mathfrak{w}(x)$，$\left(\mathbb{E}\left[\|v(X^{0,x,a})\|_\infty^2\right]\right)^{\frac{1}{2}} \le c_v \mathfrak{w}(x)$，$\left(\mathbb{E}\left[\left|\mathfrak{w}(X^{0,x,a})\right|^2\right]\right)^{\frac{1}{2}} \le c_{\mathfrak{w}} \mathfrak{w}(x)$，且 $\big(v(x)\big)(a) = \mathbb{E}\left[f(X^{0,x,a}, v(X^{0,x,a}))\right]$。那么，对所有$n \in \mathbb{N}$, $x \in \mathbb{X}$，有

$$\left(\frac{\mathbb{E}\left[\|v(x) - V_n^\theta(x)\|_\infty^2\right]}{|\mathfrak{w}(x)|^2}\right)^{1/2} \le c\left(\frac{c_{\mathfrak{w}} L(1+|A|M^{-\frac{1}{2}}) + M^{-\frac{1}{2}} + \sqrt{\left(c_{\mathfrak{w}} L(1+|A|M^{-\frac{1}{2}}) + M^{-\frac{1}{2}}\right)^2 + 4c_{\mathfrak{w}} L M^{-\frac{1}{2}}(|A|-1)}}{2}\right)^n. \quad (59)$$

命题 *3.8* 的证明。 根据三角不等式、引理 3.4 以及 A 有限这一假设，对于所有 $n \in \mathbb{N}$、 $x \in \mathbb{X}$、$\theta \in \Theta$，均有

$$\begin{aligned}
\left(\mathbb{E}\left[\|v(x) - V_n^\theta(x)\|_\infty^2\right]\right)^{\frac{1}{2}} &= \left(\mathbb{E}\left[\left\|v(x) - \mathbb{E}\left[V_n^\theta(x)\right] + \mathbb{E}\left[V_n^\theta(x)\right] - V_n^\theta(x)\right\|_\infty^2\right]\right)^{\frac{1}{2}} \\
&\le \left(\mathbb{E}\left[\left\|v(x) - \mathbb{E}\left[V_n^\theta(x)\right]\right\|_\infty^2\right]\right)^{\frac{1}{2}} + \left(\mathbb{E}\left[\left\|V_n^\theta(x) - \mathbb{E}\left[V_n^\theta(x)\right]\right\|_\infty^2\right]\right)^{\frac{1}{2}} \\
&\le \left\|v(x) - \mathbb{E}\left[V_n^\theta(x)\right]\right\|_\infty + \sum_{a\in A}\left(\mathbb{E}\left[\left|V_n^\theta(x)(a) - \mathbb{E}\left[V_n^\theta(x)(a)\right]\right|^2\right]\right)^{\frac{1}{2}} \\
&= \left\|v(x) - \mathbb{E}\left[V_n^\theta(x)\right]\right\|_\infty + \sum_{a\in A}\left(\operatorname{Var}[V_n^\theta(x)(a)]\right)^{\frac{1}{2}}. \qquad (60)
\end{aligned}$$

关于 $X^0\colon\ \mathbb{X}\times\Omega \to \mathbb{X}^A$是（$\mathcal{X}\otimes\mathcal{F}^0$）$/\widetilde{\mathcal{X}}$-可测的假设，关于 σ-代数 $\mathcal{F}^\theta$、 $\theta \in \Theta$相互独立的假设，关于 A 为有限的假设，以及引理 3.2和解体定理 [31, 引理 2.2] 证明了，对所有 $l \in \mathbb{N}$，均有（x， a）$\in \mathbb{X}\times A$

$$\begin{aligned}
\mathbb{E}\left[\|V_l^0(X^{0,x,a}) - v(X^{0,x,a})\|_\infty^2\right] &= \int_{\mathbb{X}} \mathbb{E}\left[\|V_l^0(y) - v(y)\|_\infty^2\right]\left(X^{0,x,a}(\mathbb{P})\right)(dy) \\
&\le \left(\sup_{y\in\mathbb{X}} \frac{\mathbb{E}\left[\|V_l^0(y) - v(y)\|_\infty^2\right]}{|\mathfrak{w}(y)|^2}\right) c_{\mathfrak{w}}^2 |\mathfrak{w}(x)|^2. \qquad (61)
\end{aligned}$$

由此、引理 3.6、引理 3.7和（60）可知，对所有 $n\in\mathbb{N}$、 $x\in\mathbb{X}$，均有

$$\begin{aligned}
&\big(\mathbb{E}\big[\|v(x)-V_n^0(x)\|_\infty^2\big]\big)^{\frac12}\\
&\quad\le c_{\mathfrak w}\mathfrak w(x)L\bigg(\sup_{y\in\mathbb{X}}\frac{\mathbb{E}\big[\|v(y)-V_{n-1}^0(y)\|_\infty^2\big]}{|\mathfrak w(y)|^2}\bigg)^{\frac12}+\sum_{a\in A}\bigg[\frac{1}{\sqrt{M^n}}\big(\mathbb{E}\big[|f(X^{0,x,a},0)|^2\big]\big)^{\frac12}\\
&\qquad+{}_{\mathbb{I}^{[2,\infty)(n)}}\frac{L\sqrt{M}}{\sqrt{M^n}}\big(\mathbb{E}\big[\|v(X^{0,x,a})\|_\infty^2\big]\big)^{\frac12}+{}_{\mathbb{I}^{[2,\infty)}(n)}\frac{c_{\mathfrak w}\mathfrak w(x)L}{\sqrt{M}}\bigg(\sup_{y\in\mathbb{X}}\frac{\mathbb{E}\big[\|v(y)-V_{n-1}^0(y)\|_\infty^2\big]}{|\mathfrak w(y)|^2}\bigg)^{\frac12}\\
&\qquad+\sum_{l=1}^{n-2}\frac{c_{\mathfrak w}\mathfrak w(x)L(1+\sqrt{M})}{\sqrt{M^{n-l}}}\bigg(\sup_{y\in X}\frac{\mathbb{E}\big[\|v(y)-V_l^0(y)\|_\infty^2\big]}{|\mathfrak w(y)|^2}\bigg)^{\frac12}\bigg]\\
&\quad\le c_{\mathfrak w}\mathfrak w(x)L\bigg(\sup_{y\in\mathbb{X}}\frac{\mathbb{E}\big[\|v(y)-V_{n-1}^0(y)\|_\infty^2\big]}{|\mathfrak w(y)|^2}\bigg)^{\frac12}+\frac{|A|\mathfrak w(x)}{\sqrt{M^n}}\big(c_f+{}_{\mathbb{I}^{[2,\infty)}}(n)c_vL\sqrt{M}\big)\\
&\qquad+{}_{\mathbb{I}^{[2,\infty)}}(n)\frac{c_{\mathfrak w}|A|L\mathfrak w(x)}{\sqrt{M}}\bigg(\sup_{y\in\mathbb{X}}\frac{\mathbb{E}\big[\|v(y)-V_{n-1}^0(y)\|_\infty^2\big]}{|\mathfrak w(y)|^2}\bigg)^{\frac12}\\
&\qquad+c_{\mathfrak w}\mathfrak w(x)|A|L(1+\sqrt{M})\sum_{l=1}^{n-2}\frac{1}{\sqrt{M^{n-l}}}\bigg(\sup_{y\in\mathbb{X}}\frac{\mathbb{E}\big[\|v(y)-V_l^0(y)\|_\infty^2\big]}{|\mathfrak w(y)|^2}\bigg)^{\frac12}.
\end{aligned}\tag{62}$$

对于所有 $n\in\mathbb{N}_0$，令

$$F_n=\bigg(\sup_{x\in\mathbb{X}}\frac{\mathbb{E}\big[\|v(x)-V_n^0(x)\|_\infty^2\big]}{|\mathfrak w(x)|^2}\bigg)^{\frac12}\quad\text{and}\quad a_n=M^{\frac n2}F_n.\tag{63}$$

引理 3.4 保证，对于所有 $n\in\mathbb{N}_0$，都有 $a_n,F_n\in[0,\infty)$。注意，式 (62) 表明，对于所有 $n\in\mathbb{N}\cap[2,\infty)$，

$$\begin{aligned}
F_n&\le|A|M^{-\frac n2}\big(c_f+c_vLM^{\frac12}\big)+c_{\mathfrak w}L(1+|A|M^{-\frac12})\bigg(\sup_{y\in\mathbb{X}}\frac{\mathbb{E}\big[\|v(y)-V_{n-1}^0(y)\|_\infty^2\big]}{|\mathfrak w(y)|^2}\bigg)^{\frac12}\\
&\quad+c_{\mathfrak w}|A|L(1+M^{\frac12})\sum_{l=1}^{n-2}M^{-\frac{n-l}{2}}\bigg(\sup_{y\in\mathbb{X}}\frac{\mathbb{E}\big[\|v(y)-V_l^0(y)\|_\infty^2\big]}{|\mathfrak w(y)|^2}\bigg)^{\frac12}\\
&=|A|M^{-\frac n2}\big(c_f+c_vLM^{\frac12}\big)+c_{\mathfrak w}L(1+|A|M^{-\frac12})F_{n-1}+c_{\mathfrak w}|A|L(1+M^{\frac12})\sum_{l=1}^{n-2}M^{-\frac{n-l}{2}}F_l.
\end{aligned}\tag{64}$$

这意味着，对于所有 $n\in\mathbb{N}\cap[2,\infty)$，

$$\begin{aligned}
a_n&\le|A|(c_f+c_vLM^{\frac12})+c_{\mathfrak w}M^{\frac12}L(1+|A|M^{-\frac12})M^{\frac{n-1}{2}}F_{n-1}+c_{\mathfrak w}|A|L(1+M^{\frac12})\sum_{l=1}^{n-2}M^{\frac l2}F_l\\
&=|A|(c_f+c_vLM^{\frac12})+c_{\mathfrak w}L(M^{\frac12}+|A|)a_{n-1}+c_{\mathfrak w}|A|L(1+M^{\frac12})\sum_{l=1}^{n-2}a_l.
\end{aligned}\tag{65}$$

此外，成立的是

$$a_0=F_0=\bigg(\sup_{y\in\mathbb{X}}\frac{\mathbb{E}\big[\|v(y)\|_\infty^2\big]}{|\mathfrak w(y)|^2}\bigg)^{\frac12}=\sup_{(y,a)\in\mathbb{X}\times A}\frac{\big|\big(v(y)\big)(a)\big|}{|\mathfrak w(y)|}.\tag{66}$$

假设对于所有 $(x,a)\in\mathbb{X}\times A$，都有 $\big(v(x)\big)(a)=\mathbb{E}\big[f(X^{0,x,a},v(X^{0,x,a}))\big]$，再结合詹森不等式、三角不等式、假设对于所有 $x\in\mathbb{X}$, $r,s\in\mathbb{R}^A$，都有 $|f(x,r)-f(x,s)|\le L\max_{a\in A}|r(a)-s(a)|$，以及假设对于所有 $(x,a)\in\mathbb{X}\times A$，都有 $\big(\mathbb{E}\big[|f(X^{0,x,a},0|^2\big]\big)^{\frac12}\le c_f\mathfrak{w}(x)$ and $\big(\mathbb{E}\big[\|v(X^{0,x,a})\|_\infty^2\big)^{\frac12}\le c_v\mathfrak{w}(x)$，可证明对于所有 $(x,a)\in\mathbb{X}\times A$，都有

$$\begin{aligned}\frac{\big|\big(v(x)\big)(a)\big|}{|\mathfrak{w}(x)|}&=\frac{1}{\mathfrak{w}(x)}\big(\big|\mathbb{E}\big[f(X^{0,x,a},v(X^{0,x,a}))\big]\big|^2\big)^{\frac12}\le\frac{1}{\mathfrak{w}(x)}\big(\mathbb{E}\big[\big|f(X^{0,x,a},v(X^{0,x,a}))\big|^2\big]\big)^{\frac12}\\&\le\frac{1}{\mathfrak{w}(x)}\Big[L\big(\mathbb{E}\big[\sup_{b\in A}|v(X^{0,x,a})(b)|^2\big]\big)^{\frac12}+\big(\mathbb{E}\big[|f(X^{0,x,a},0)|^2\big]\big)^{\frac12}\Big]\le c_vL+c_f\end{aligned}\tag{67}$$

这一点与 (66) 确保了 $a_0\le c_vL+c_f$。注意，(62) 蕴含 $a_1\le M^{\frac12}\big(c_{\mathfrak{w}}La_0+|A|M^{-\frac12}c_f\big)\le c_{\mathfrak{w}}LM^{\frac12}(c_vL+c_f)+|A|c_f$。此外，(65) 确保 $a_2\le|A|(c_f+c_vLM^{\frac12})+c_{\mathfrak{w}}L(M^{\frac12}+|A|)a_1$。令 $(\xi_n)_{n\in\mathbb{N}_0}$、$(b_n)_{n\in\mathbb{N}_0}\subseteq\mathbb{R}$ 满足对所有 $n\in\mathbb{N}_0$，

$$\begin{aligned}b_0=\xi_0&=\max\Big\{\frac{c_v}{c_{\mathfrak{w}}},c_vL+c_f,\frac{|A|c_f}{c_{\mathfrak{w}}|A|L+1}\Big\},\quad b_1=|A|c_f-(c_{\mathfrak{w}}|A|L+1)b_0,\\b_2&=|A|LM^{\frac12}(c_v-c_{\mathfrak{w}}b_0),\quad b_{n+3}=0,\quad \xi_1=c_{\mathfrak{w}}LM^{\frac12}b_0+|A|c_f,\\\xi_{n+2}&=|A|(c_f+c_vLM^{\frac12})+c_{\mathfrak{w}}L(M^{\frac12}+|A|)\xi_{n+1}+c_{\mathfrak{w}}|A|L(1+M^{\frac12})\sum_{l=1}^n\xi_l.\end{aligned}\tag{68}$$

结合 (62)、(65)、(68) 和归纳法，可得对所有 $n\in\mathbb{N}_0$，$a_n\le\xi_n$。注意，(68) 确保对所有 $n\in\mathbb{N}_0$，均有

$$\begin{aligned}\xi_0=b_0,\quad \xi_1&=b_1+(c_{\mathfrak{w}}L(M^{\frac12}+|A|)+1)\xi_0,\\\xi_{n+2}&=b_{n+2}+(c_{\mathfrak{w}}L(M^{\frac12}+|A|)+1)\xi_{n+1}+c_{\mathfrak{w}}LM^{\frac12}(|A|-1)\xi_n.\end{aligned}\tag{69}$$

令 x_1、$x_2\in\mathbb{R}$ 满足

$$\begin{aligned}x_1&=\frac{c_{\mathfrak{w}}L(M^{\frac12}+|A|)+1-\sqrt{\big(c_{\mathfrak{w}}L(M^{\frac12}+|A|)+1\big)^2+4c_{\mathfrak{w}}LM^{\frac12}(|A|-1)}}{2},\\x_2&=\frac{c_{\mathfrak{w}}L(M^{\frac12}+|A|)+1+\sqrt{\big(c_{\mathfrak{w}}L(M^{\frac12}+|A|)+1\big)^2+4c_{\mathfrak{w}}LM^{\frac12}(|A|-1)}}{2}.\end{aligned}\tag{70}$$

注意，对所有 $i\in\{1,2\}$，均有 $x_i^2=\big(c_{\mathfrak{w}}L(M^{\frac12}+|A|)+1\big)x_i+c_{\mathfrak{w}}LM^{\frac12}(|A|-1)$。此外，注意到 A 非空这一假设可推出 $c_{\mathfrak{w}}L(M^{\frac12}+|A|)+1\ge1$ 和 $c_{\mathfrak{w}}LM^{\frac12}(|A|-1)\ge0$。这确保了 $x_2>0\ge x_1$。这一点与 (70) 蕴含 $|x_2|\ge|x_1|$。因此，有

$$\frac{|x_2|}{|x_2-x_1|}\le1\quad\text{and}\quad\frac{|x_1|}{|x_2-x_1|}\le\frac12.\tag{71}$$

由此、$|x_1|\le|x_2|$这一事实以及三角不等式可得，对所有 $n\in\mathbb{N}_0$，均有

$$\frac{x_2^{n+1}-x_1^{n+1}}{x_2-x_1}\le\frac{|x_2|^{n+1}}{x_2-x_1}+\frac{|x_1|^{n+1}}{x_2-x_1}\le\frac{x_2}{x_2-x_1}x_2^n+\frac{|x_1|}{x_2-x_1}x_2^n\le\frac32x_2^n.\tag{72}$$

[34, 引理 2.1] 中的离散 Gronwall 型两步递推（在 [34, 引理 2.1]的记号下，取 $\kappa \leftarrow c_{\mathfrak{w}} L(M^{\frac{1}{2}}+|A|)+1$, $\lambda \leftarrow c_{\mathfrak{w}} L M^{\frac{1}{2}}(|A|-1)$, $(a_k)_{k\in\mathbb{N}_0} \leftarrow (\xi_k)_{k\in\mathbb{N}_0}$, $(b_k)_{k\in\mathbb{N}_0} \leftarrow (b_k)_{k\in\mathbb{N}_0}$, $x_{1/2} \leftarrow x_{1/2}$ ）表明，对所有 $n \in \mathbb{N}_0$，均有

$$\xi_n = \frac{1}{x_2 - x_1}\Big(b_0(x_2^{n+1} - x_1^{n+1}) + b_1(x_2^n - x_1^n) + b_2(x_2^{\max\{n-1,0\}} - x_1^{\max\{n-1,0\}})\Big). \tag{73}$$

注意到 (68) 保证了 $b_1 \le 0$ 和 $b_2 \le 0$。这一点、 $x_2 > x_1$这一事实，以及对所有 $n \in \mathbb{N}_0$ 均有 $a_n \le \xi_n$ 这一事实，证明了对所有 $n \in \mathbb{N}_0$ 均有

$$a_n \le b_0 \frac{x_2^{n+1} - x_1^{n+1}}{x_2 - x_1}. \tag{74}$$

结合这一点与 (72)，可知对所有 $n \in \mathbb{N}_0$ 均有

$$F_n \le \frac{3}{2} b_0 \big(M^{-\frac{1}{2}} x_2\big)^n. \tag{75}$$

此外，注意到

$$\begin{aligned} M^{-\frac{1}{2}} x_2 &= M^{-\frac{1}{2}} \left(\frac{c_{\mathfrak{w}} L(M^{\frac{1}{2}} + |A|) + 1 + \sqrt{\big(c_{\mathfrak{w}} L(M^{\frac{1}{2}} + |A|) + 1\big)^2 + 4 c_{\mathfrak{w}} L M^{\frac{1}{2}}(|A| - 1)}}{2} \right) \\ &= \frac{c_{\mathfrak{w}} L(1 + |A| M^{-\frac{1}{2}}) + M^{-\frac{1}{2}} + \sqrt{\big(c_{\mathfrak{w}} L(1 + |A| M^{-\frac{1}{2}}) + M^{-\frac{1}{2}}\big)^2 + 4 c_{\mathfrak{w}} L M^{-\frac{1}{2}}(|A| - 1)}}{2} \end{aligned} \tag{76}$$

由此、(63) 和 (75) 可知，对所有 $x \in \mathbb{X}$、 $n \in \mathbb{N}$ 均有

$$\begin{aligned} &\left(\frac{\mathbb{E}\big[\|v(x) - V_n^\theta(x)\|_\infty^2\big]}{|\mathfrak{w}(x)|^2} \right)^{\frac{1}{2}} \\ &\le c \left(\frac{c_{\mathfrak{w}} L(1 + |A| M^{-\frac{1}{2}}) + M^{-\frac{1}{2}} + \sqrt{\big(c_{\mathfrak{w}} L(1 + |A| M^{-\frac{1}{2}}) + M^{-\frac{1}{2}}\big)^2 + 4 c_{\mathfrak{w}} L M^{-\frac{1}{2}}(|A| - 1)}}{2} \right)^n. \end{aligned} \tag{77}$$

至此，命题 3.8 的证明完成。 □

推论 *3.9*。假设设定 *3.1*，假设 A 是有限的，令 c_f、 c_v、 $c_{\mathfrak{w}}$、 $L \in [0,\infty)$ 满足 $c_{\mathfrak{w}} L < 1$，令 $M > (1 + c_{\mathfrak{w}} L(2|A| - 1))^2 (1 - c_{\mathfrak{w}} L)^{-2}$，令 $\mathfrak{w}\colon \mathbb{X} \to (0,\infty)$ 为 $\mathcal{X}/\mathcal{B}((0,\infty))$-可测，令$v\colon \mathbb{X} \to \mathbb{R}^A$ 为 $\mathcal{X}/\widetilde{\mathcal{A}}$-可测，假设对所有 $(x,a) \in \mathbb{X} \times A$、 r、 $s \in \mathbb{R}^A$ 满足 $|f(x,r) - f(x,s)| \le L \max_{b\in A} |r(b) - s(b)|$、 $\big(\mathbb{E}\big[|f(X^{0,x,a},0)|^2\big]\big)^{\frac{1}{2}} \le c_f \mathfrak{w}(x)$， $\big(\mathbb{E}\big[\|v(X^{0,x,a})\|_\infty^2\big]\big)^{\frac{1}{2}} \le c_v \mathfrak{w}(x)$， $\big(\mathbb{E}\big[|\mathfrak{w}(X^{0,x,a})|^2\big]\big)^{\frac{1}{2}} \le c_{\mathfrak{w}} \mathfrak{w}(x)$，以及 $(v(x))(a) = \mathbb{E}\big[f\big(X^{0,x,a}, v(X^{0,x,a})\big)\big]$。则有

$$\limsup_{n\to\infty} \left[\sup_{x\in\mathbb{X}} \left(\frac{\mathbb{E}\big[\|v(x) - V_n^0(x)\|_\infty^2\big]}{|\mathfrak{w}(x)|^2} \right)^{\frac{1}{2}} \right] = 0. \tag{78}$$

推论 *3.9* 的证明。令 $c \in [0,\infty)$ 满足 $c = \frac{3}{2} \max\big\{\frac{c_v}{c_{\mathfrak{w}}}, c_v L + c_f, \frac{|A| c_f}{c_{\mathfrak{w}}|A| L + 1}\big\}$。命题 3.8 表明，对于所有 $n \in \mathbb{N}$, $x \in \mathbb{X}$，均有

$$\begin{aligned} &\left(\frac{\mathbb{E}\big[\|v(x) - V_n^0(x)\|_\infty^2\big]}{|\mathfrak{w}(x)|^2} \right)^{\frac{1}{2}} \\ &\le c \left(\frac{c_{\mathfrak{w}} L(1 + |A| M^{-\frac{1}{2}}) + M^{-\frac{1}{2}} + \sqrt{\big(c_{\mathfrak{w}} L(1 + |A| M^{-\frac{1}{2}}) + M^{-\frac{1}{2}}\big)^2 + 4 c_{\mathfrak{w}} L M^{-\frac{1}{2}}(|A| - 1)}}{2} \right)^n. \end{aligned} \tag{79}$$

注意，假设 $M > (1 + c_{\mathfrak{w}} L(2|A| - 1))^2 (1 - c_{\mathfrak{w}} L)^{-2}$ 蕴含 $M^{-\frac{1}{2}} < (1 - c_{\mathfrak{w}} L)(1 + c_{\mathfrak{w}} L(2|A| - 1))^{-1}$。这就证明了

$$\begin{aligned} c_{\mathfrak{w}} L(|A| - 1) M^{-\frac{1}{2}} &< \frac{c_{\mathfrak{w}} L(|A| - 1)(1 - c_{\mathfrak{w}} L)}{1 + c_{\mathfrak{w}} L(2|A| - 1)} \\ &= 1 - \Big(c_{\mathfrak{w}} L + (1 - c_{\mathfrak{w}} L) - \frac{c_{\mathfrak{w}} L(|A| - 1)(1 - c_{\mathfrak{w}} L)}{1 + c_{\mathfrak{w}} L(2|A| - 1)} \Big) \\ &= 1 - \Big(c_{\mathfrak{w}} L + \big(1 + c_{\mathfrak{w}} L(2|A| - 1) - c_{\mathfrak{w}} L(|A| - 1) \big) \frac{1 - c_{\mathfrak{w}} L}{1 + c_{\mathfrak{w}} L(2|A| - 1)} \Big) \\ &= 1 - \Big(c_{\mathfrak{w}} L + \big(1 + c_{\mathfrak{w}} |A| L \big) \frac{1 - c_{\mathfrak{w}} L}{1 + c_{\mathfrak{w}} L(2|A| - 1)} \Big) \\ &< 1 - \big(c_{\mathfrak{w}} L + (1 + c_{\mathfrak{w}} |A| L) M^{-\frac{1}{2}} \big) = 1 - \big(c_{\mathfrak{w}} L (1 + |A| M^{-\frac{1}{2}}) + M^{-\frac{1}{2}} \big). \end{aligned} \tag{80}$$

假设 A 非空，便可确保 $|A| \leq 2|A| - 1$。这与假设 $M > (1 + c_{\mathfrak{w}} L(2|A| - 1))^2 (1 - c_{\mathfrak{w}} L)^{-2}$ 共同推出 $M > (1 + c_{\mathfrak{w}} L|A|)^2 (1 - c_{\mathfrak{w}} L)^{-2}$。因此，有 $c_{\mathfrak{w}} L(1 + |A| M^{-\frac{1}{2}}) + M^{-\frac{1}{2}} = c_{\mathfrak{w}} L + (1 + c_{\mathfrak{w}} |A| L) M^{-\frac{1}{2}} < 1 < 2$。由此以及 (80) 可得

$$\begin{aligned} &\frac{c_{\mathfrak{w}} L(1 + |A| M^{-\frac{1}{2}}) + M^{-\frac{1}{2}} + \sqrt{\Big(c_{\mathfrak{w}} L(1 + |A| M^{-\frac{1}{2}}) + M^{-\frac{1}{2}} \Big)^2 + 4 c_{\mathfrak{w}} L M^{-\frac{1}{2}} (|A| - 1)}}{2} \\ &\quad < \frac{c_{\mathfrak{w}} L(1 + |A| M^{-\frac{1}{2}}) + M^{-\frac{1}{2}} + \sqrt{\Big(c_{\mathfrak{w}} L(1 + |A| M^{-\frac{1}{2}}) + M^{-\frac{1}{2}} \Big)^2 + 4 - 4 \Big(c_{\mathfrak{w}} L(1 + |A| M^{-\frac{1}{2}}) + M^{-\frac{1}{2}} \Big)}}{2} \\ &\quad = \frac{c_{\mathfrak{w}} L(1 + |A| M^{-\frac{1}{2}}) + M^{-\frac{1}{2}} + \sqrt{\Big(2 - c_{\mathfrak{w}} L(1 + |A| M^{-\frac{1}{2}}) + M^{-\frac{1}{2}} \Big)^2}}{2} = 1. \end{aligned} \tag{81}$$

结合 (79) 和 (81) 可得

$$\limsup_{n \to \infty} \left[\sup_{x \in \mathbb{X}} \left(\frac{\mathbb{E}\big[\|v(x) - V_n^0(x)\|_\infty^2 \big]}{|\mathfrak{w}(x)|^2} \right)^{\frac{1}{2}} \right] = 0. \tag{82}$$

至此，推论 3.9 的证明完成。 □

4 MLFP 近似的计算复杂度分析

4.1 函数不动点方程的 MLFP 近似

定理 4.1。 设 $M \in \mathbb{N}$，设 $\Theta = \bigcup_{n \in \mathbb{N}} \mathbb{Z}^n$，设 A 为有限非空集，设 $\kappa \in [0, \infty)$，设 $\mathfrak{D}$ 为非空集合，设 λ_d、L_d、$\mathfrak{R}_d \in [0, \infty)$、$d \in \mathfrak{D}$，设 $(\Omega, \mathcal{F}, \mathbb{P})$ 为概率空间，设 $(\mathbb{X}_d, \mathcal{X}_d)$、$d \in \mathfrak{D}$ 为非空可测空间，对于每个 $d \in \mathfrak{D}$，设 $\mathfrak{w}_d \colon \mathbb{X}_d \to (0, \infty)$ 为 $\mathcal{X}_d / \mathcal{B}((0, \infty))$-可测的；对于每个 $d \in \mathfrak{D}$，设 $f_d \colon \mathbb{X}_d \times \mathbb{R}^A \to \mathbb{R}$，为 $(\mathcal{X}_d \otimes (\bigotimes_{a \in A} \mathcal{B}(\mathbb{R}))) / \mathcal{B}(\mathbb{R})$-可测的；对于每个 $d \in \mathfrak{D}$，设 $(\mathcal{F}_d^\theta)_{\theta \in \Theta}$ 为 $\mathcal{F}$ 中相互独立的子-σ-代数；对于每个 $d \in \mathfrak{D}$，设 $X_d^\theta \colon \mathbb{X}_d \times \Omega \to \mathbb{X}_d^A$、$\theta \in \Theta$ 为独立同分布随机场，且满足对所有 $d \in \mathfrak{D}$、$\theta \in \Theta$，X_d^θ 都是 $(\mathcal{X}_d \otimes \mathcal{F}_d^\theta) / (\bigotimes_{a \in A} \mathcal{X}_d)$-可测的；假设对所有 $d \in \mathfrak{D}$、$(x, a) \in \mathbb{X}_d \times A$、$r, s \in \mathbb{R}^A$，有 $|f_d(x, r) - f_d(x, s)| \leq L_d \max_{b \in A} |r(b) - s(b)|$，$\big(\mathbb{E}\big[|f_d(X^0, {}^{x\,a}d, 0)|^2 \big] \big)^{\frac{1}{2}} \leq \kappa \mathfrak{w}_d(x)$，$\big(\mathbb{E}\big[|\mathfrak{w}_d(X^0, {}^{x\,a}d, \cdot)|^2 \big] \big)^{\frac{1}{2}} \leq \lambda_d \mathfrak{w}_d(x)$，$\sup_{u \in \mathfrak{D}} \lambda_u L_u < 1$；假设 $M > (1 + (\sup_{d \in \mathfrak{D}} \lambda_d L_d)(2|A| - 1))^2 (1 - (\sup_{d \in \mathfrak{D}} \lambda_d L_d))^2$；对于每个 $d \in \mathfrak{D}$，设 $V_{n,d}^\theta \colon \mathbb{X}_d \times \Omega \to \mathbb{R}^A$、$n \in \mathbb{N}_0$、$\theta \in \Theta$，满足对所有 $n \in \mathbb{N}_0$、$(x, a) \in \mathbb{X}_d \times A$、$\theta \in \Theta$

that

$$\big(V_{n,d}^{\theta}(x)\big)(a)=\sum_{l=0}^{n-1}\frac{1}{M^{n-l}}\sum_{i=1}^{M^{n-l}}f_d\big(X_d^{(\theta,l,i),x,a},V_{l,d}^{(\theta,l,i)}(X_d^{(\theta,l,i),x,a})\big)$$
$$-\mathbb{I}^{\mathbb{N}}(l)f_d\big(X_d^{(\theta,l,i),x,a},V_{\max\{l-1,0\},d}^{(\theta,-l,i)}(X_d^{(\theta,l,i),x,a})\big),\tag{83}$$

并设 $\mathfrak{C}_{n,\,d}\in[0,\infty)$、 $n\in\mathbb{N}_0$、 $d\in\mathfrak{D}$，满足对所有 $n\in\mathbb{N}_0$、 $d\in\mathfrak{D}$，有

$$\mathfrak{C}_{n,d}\le\sum_{l=0}^{n-1}M^{n-l}\big(\mathfrak{R}_d+\mathfrak{C}_{l,d}+\mathbb{I}^{\mathbb{N}}(l)\mathfrak{C}_{\max\{l-1,0\},d}\big).\tag{84}$$

则有以下结论:

(*i*) 对于每个 $d\in\mathfrak{D}$，都存在唯一函数 $v_d\colon\mathbb{X}_d\to\mathbb{R}^A$，它是 $\mathcal{X}_d/(\bigotimes_{a\in A}\mathcal{B}(\mathbb{R}))$-可测的，并且满足：对于所有 $(x,a)\in\mathbb{X}_d\times A$，有 $\sup_{y\in\mathbb{X}_d}\frac{\|v_d(x)-v_d(y)\|_\infty}{\mathfrak{w}_d(x,y)}<\infty$，$\mathbb{E}\big[|f_d(X^{0,x,a}_d,v_d(X^{0,x,a}_d))|\big]<\infty$，并且

$$\big(v_d(x)\big)(a)=\mathbb{E}\big[f_d(X_d^{0,x,a},v_d(X_d^{0,x,a}))\big].\tag{85}$$

(*ii*) 存在 $N\colon(0,1]\to\mathbb{N}$ 以及 $c\in\mathbb{R}$，使得对于所有 $d\in\mathfrak{D}$，$\varepsilon\in(0,1]$，均有 $\mathfrak{C}_{N_\varepsilon,d}\le c\mathfrak{R}_d\varepsilon^{-c}$，并且

$$\sup_{x\in\mathbb{X}_d}\left(\frac{\mathbb{E}\big[\|v_d(x)-V^0_{N_\varepsilon,d}(x)\|_\infty^2\big]}{|\mathfrak{w}_d(x)|^2}\right)^{\frac12}\le\varepsilon.\tag{86}$$

定理 *4.1* 的证明。令 $\alpha,\beta,\gamma\in\mathbb{R}\cup\{\infty\}$ 满足

$$\alpha=\frac12\sup_{d\in\mathfrak{D}}\Big[\lambda_dL_d(1+|A|M^{-\frac12})+M^{-\frac12}+\sqrt{\big(\lambda_dL_d(1+|A|M^{-\frac12})+M^{-\frac12}\big)^2+4M^{-\frac12}\lambda_dL_d(|A|-1)}\Big],$$
$$\beta=\frac{\ln(3M)}{\ln(\alpha^{-1})},\quad\text{and}\quad\gamma=\frac32\sup_{d\in\mathfrak{D}}\left(\max\left\{\frac{\kappa}{1-\lambda_dL_d},\frac{\kappa\lambda_dL_d}{1-\lambda_dL_d}+\kappa,\frac{|A|\kappa}{\lambda_d|A|L_d+1}\right\}\right).\tag{87}$$

令 $N\colon(0,1]\to\mathbb{N}\cup\{\infty\}$ 满足对所有 $\varepsilon\in(0,1]$，均有 $N_\varepsilon=\min(\{n\in\mathbb{N}\colon\gamma\alpha^n\le\varepsilon\}\cup\{\infty\})$。假设 $\sup_{d\in\mathfrak{D}}\lambda_dL_d<1$ 可确保 $\gamma\in[0,\infty)$。由三角不等式可得

$$\alpha=\frac12\sup_{d\in\mathfrak{D}}\Big[\lambda_dL_d(1+|A|M^{-\frac12})+M^{-\frac12}+\sqrt{\big(\lambda_dL_d(1+|A|M^{-\frac12})+M^{-\frac12}\big)^2+4M^{-\frac12}\lambda_dL_d(|A|-1)}\Big],$$
$$\le\frac12\Bigg[(\sup_{d\in\mathfrak{D}}\lambda_dL_d)(1+|A|M^{-\frac12})+M^{-\frac12}$$
$$+\sqrt{\big((\sup_{d\in\mathfrak{D}}\lambda_dL_d)(1+|A|M^{-\frac12})+M^{-\frac12}\big)^2+4M^{-\frac12}(\sup_{d\in\mathfrak{D}}\lambda_dL_d)(|A|-1)}\Bigg].\tag{88}$$

假设 $M^{-\frac{1}{2}} < \frac{1-(\sup_{d\in\mathfrak{D}}\lambda_d L_d)}{1+(\sup_{d\in\mathfrak{D}}\lambda_d L_d)(2|A|-1)}$ 可确保

$$\begin{aligned}
1-\big((\sup_{d\in\mathfrak{D}}\lambda_d L_d)(1+|A|M^{-\frac{1}{2}})+M^{-\frac{1}{2}}\big) &= 1-(\sup_{d\in\mathfrak{D}}\lambda_d L_d)-(1+|A|(\sup_{d\in\mathfrak{D}}\lambda_d L_d))M^{-\frac{1}{2}} \\
&> (1-(\sup_{d\in\mathfrak{D}}\lambda_d L_d))-(1+|A|(\sup_{d\in\mathfrak{D}}\lambda_d L_d))\frac{1-(\sup_{d\in\mathfrak{D}}\lambda_d L_d)}{1+(\sup_{d\in\mathfrak{D}}\lambda_d L_d)(2|A|-1)} \\
&= \Big(1+(\sup_{d\in\mathfrak{D}}\lambda_d L_d)(2|A|-1)-\big(|A|(\sup_{d\in\mathfrak{D}}\lambda_d L_d)+1\big)\Big)\frac{1-(\sup_{d\in\mathfrak{D}}\lambda_d L_d)}{1+(\sup_{d\in\mathfrak{D}}\lambda_d L_d)(2|A|-1)} \\
&= \Big((\sup_{d\in\mathfrak{D}}\lambda_d L_d)(|A|-1)\Big)\frac{1-(\sup_{d\in\mathfrak{D}}\lambda_d L_d)}{1+(\sup_{d\in\mathfrak{D}}\lambda_d L_d)(2|A|-1)} \\
&> (\sup_{d\in\mathfrak{D}}\lambda_d L_d)(|A|-1)M^{-\frac{1}{2}}. \qquad (89)
\end{aligned}$$

A 非空这一假设意味着 $|A|\le 2|A|-1$。因此，有 $M > \frac{(1+(\sup_{d\in\mathfrak{D}}\lambda_d L_d)(2|A|-1))^2}{(1+|A|(\sup_{t\in T}\lambda_d L_d)^2} \ge \frac{(1-(\sup_{d\in\mathfrak{D}}\lambda_d L_d))^2}{(1+|A|(\sup_{t\in T}\lambda_d L_d)^2}$。这就证明了

$$(\sup_{d\in\mathfrak{D}}\lambda_d L_d)+(|A|(\sup_{d\in\mathfrak{D}}\lambda_d L_d)+1)M^{-\frac{1}{2}}<1. \qquad (90)$$

结合这一点、(88) 和 (89)，可得

$$\begin{aligned}
\alpha &\le \frac{1}{2}\Bigg[(\sup_{d\in\mathfrak{D}}\lambda_d L_d)(1+|A|M^{-\frac{1}{2}})+M^{-\frac{1}{2}} \\
&\quad +\sqrt{\big((\sup_{d\in\mathfrak{D}}\lambda_d L_d)(1+|A|M^{-\frac{1}{2}})+M^{-\frac{1}{2}}\big)^2+4M^{-\frac{1}{2}}(\sup_{d\in\mathfrak{D}}\lambda_d L_d)(|A|-1)}\Bigg] \\
&< \frac{1}{2}\Bigg[(\sup_{d\in\mathfrak{D}}\lambda_d L_d)(1+|A|M^{-\frac{1}{2}})+M^{-\frac{1}{2}} \\
&\quad +\sqrt{\big((\sup_{d\in\mathfrak{D}}\lambda_d L_d)(1+|A|M^{-\frac{1}{2}})+M^{-\frac{1}{2}}\big)^2+4\big(1-(\sup_{d\in\mathfrak{D}}\big(\lambda_d L_d)(1+|A|M^{-\frac{1}{2}})+M^{-\frac{1}{2}}\big)\big)}\Bigg] \\
&= \frac{1}{2}\Bigg[(\sup_{d\in\mathfrak{D}}\lambda_d L_d)(1+|A|M^{-\frac{1}{2}})+M^{-\frac{1}{2}}+\sqrt{\big(2-\big((\sup_{d\in\mathfrak{D}}\lambda_d L_d)(1+|A|M^{-\frac{1}{2}})+M^{-\frac{1}{2}}\big)\big)^2}\Bigg] \\
&= 1. \qquad (91)
\end{aligned}$$

这与 (87) 确保了 $\alpha\in[M^{-\frac{1}{2}},1)$。假设 $M > \frac{(1+(\sup_{d\in\mathfrak{D}}\lambda_d L_d)(2|A|-1))^2}{(1-(\sup_{d\in\mathfrak{D}}\lambda_d L_d))^2}$ 意味着 $M>1$。因此，有 $[M^{-\frac{1}{2}},1)\neq\emptyset$。事实 $\alpha\in(0,1)$ 表明，对所有 $\varepsilon\in(0,1]$，都有 $N_\varepsilon\in\mathbb{N}$。事实 $\alpha\in[M^{-\frac{1}{2}},1)$ 以及事实 $M>1$ 确保了 $\beta\in(2,\infty)$。推论 2.3 得出，对所有 $d\in\mathfrak{D}$，存在唯一函数 $v_d\colon \mathbb{X}_d\to\mathbb{R}^A$，它是 $\mathcal{X}_d/(\bigotimes_{a\in A}\mathcal{B}(\mathbb{R}))$-可测的，并且对所有 $d\in\mathfrak{D}$、$(x,a)\in\mathbb{X}_d\times A$，满足 $\sup_{y\in\mathbb{X}_d}\frac{\|v_d(y)\|_\infty}{\mathfrak{w}_d(y)}<\infty$，$\mathbb{E}\big[|f_d(X^{0,x,a}_d,v_d(X^{0,x,a}_d))|\big]<\infty$，且 $v_d(x,a)=\mathbb{E}\big[f_d(X^{0,x,a}_d,v_d(X^{0,x,a}_d))\big]$。这证明了第 (i) 项。引理 2.4 意味着，对所有 $d\in\mathfrak{D}$、$(x,a)\in\mathbb{X}_d\times A$，有 $\big(\mathbb{E}\big[\|v_d(X^{0,x,a}_d)\|^2_\infty\big]\big)^{\frac{1}{2}}\le\frac{\kappa\lambda_d}{1-\lambda_d L_d}\mathfrak{w}_d(x)$。命题 3.8 证明了，对所有$n\in\mathbb{N}$、$d\in\mathfrak{D}$、$x\in\mathbb{X}_d$，都有

$$\begin{aligned}
&\left(\frac{\mathbb{E}\big[\|v_d(X^{0,x,a}_d)-V^0_{n,d}(X^{0,x,a}_d)\|^2_\infty\big]}{|\mathfrak{w}_d(x)|^2}\right)^{\frac{1}{2}} \qquad (92)\\
&\qquad\le\gamma\left(\frac{\lambda_d L_d(1+|A|M^{-\frac{1}{2}})+M^{-\frac{1}{2}}+\sqrt{(\lambda_d L_d(1+|A|M^{-\frac{1}{2}})+M^{-\frac{1}{2}})^2+4\lambda_d L_d M^{-\frac{1}{2}}(|A|-1)}}{2}\right)^n.
\end{aligned}$$

这表明，对所有 $n \in \mathbb{N}$、 $d \in \mathfrak{D}$ ，都有

$$\sup_{x \in \mathbb{X}_d} \left(\frac{\mathbb{E}\big[\|v_d(x) - V_{n,d}^0(x)\|_\infty^2\big]}{|\mathfrak{w}_d(x)|^2} \right)^{\frac{1}{2}} \leq \gamma \alpha^n. \tag{93}$$

注意， [3, 引理 3.14] （在 $M \leftarrow M$, $\alpha \leftarrow \mathfrak{R}_d + \mathfrak{C}_{0,d}$, $\beta \leftarrow \mathfrak{R}_d$, $(C_n) \leftarrow (\mathfrak{C}_{n,d})$ 取$d \in \mathfrak{D}$ 并采用 [3, 引理 3.14]中的记号时应用）证明，对所有 $n \in \mathbb{N}$, $d \in \mathfrak{D}$ ，均有

$$\mathfrak{C}_{n,d} \leq \left(\frac{\mathfrak{R}_d + \mathfrak{C}_{0,d} + \mathfrak{R}_d + \mathfrak{C}_{0,d}}{2} \right) (3M)^n = \mathfrak{R}_d (3M)^n. \tag{94}$$

由此、(93) 以及 [3, 引理 3.15] （在 $m \leftarrow 1$, $\alpha \leftarrow \alpha$, $\beta \leftarrow 3M$, $\kappa_1 \leftarrow \gamma$, $\kappa_2 \leftarrow \mathfrak{R}_d$,$N \leftarrow N$ 取 $d \in \mathfrak{D}$ 并采用 [3, 引理 3.15]中的记号时应用）可知，对所有 $d \in \mathfrak{D}$, $\varepsilon \in (0, 1]$ ，均有

$$\sup_{x \in \mathbb{X}_d} \left(\frac{\mathbb{E}\big[\|v_d(x) - V_{n,d}^0(x)\|_\infty^2\big]}{|\mathfrak{w}_d(x)|^2} \right)^{\frac{1}{2}} \leq \varepsilon \quad \text{and} \quad \mathfrak{C}_{N_\varepsilon,d} \leq 3M\mathfrak{R}_d \max\{1, \gamma\}^\beta \frac{1}{\varepsilon^\beta}. \tag{95}$$

令 $c = \mathsf{max}\{\beta, 3M \mathsf{max}\{1, \gamma\}^\beta\}$。因此，对所有 $d \in \mathfrak{D}$, $\varepsilon \in (0, 1]$ ，均有 $\mathfrak{C}_{N_\varepsilon,d} \leq c\mathfrak{R}_d \varepsilon^{-c}$。定理 4.1 的证明至此完成。 □

4.2 最优控制问题贝尔曼方程的 MLFP 近似

推论 *4.2*。设 $M \in \mathbb{N}$, $\kappa \in [0, \infty)$, $\Theta = \bigcup_{n \in \mathbb{N}} \mathbb{Z}^n$，设 A 为有限非空集，设$\mathfrak{D}$ 为非空集合，设 $(\Omega\ \mathcal{F}, \mathbb{P})$ 为概率空间，设 $\lambda_d, \delta_d, \mathfrak{R}_d \in [0, \infty)$, $d \in \mathfrak{D}$, $\mathbb{X}_d, \mathcal{X}_d)$, $d \in \mathfrak{D}$ 为非空可测空间；对于每个 $d \in \mathfrak{D}$，令 $\mathfrak{w}_d$: $\mathbb{X}_d \to (0, \infty)$ 为$\mathcal{X}_d/\mathcal{B}((0, \infty))$-可测；对于每个 $d \in \mathfrak{D}$，令 g_d: $\mathbb{X}_d \times A \to \mathbb{R}$ 为 $(\mathcal{X}_d \otimes 2^A)/\mathcal{B}(\mathbb{R})$-可测；对于每个 $d \in \mathfrak{D}$ $\mathcal{F}_d^\theta{}_{\theta \in \Theta}$ σ $\mathcal{F}$ $d \in \mathfrak{D}$ $X_d^\theta =$，令 $(\ )$ 为 的 $\big(X_d^{\theta\ x\ a}\ \omega\ \big)_{x \in \mathbb{X}_d\ a \in A\ \omega \in}$ $\mathbb{X}_d \times A \times \quad \to \mathbb{X}_d$ $\theta \in \Theta$ 独立子 -代数；对于每个，令，，$(\ )$，，$\Omega : \Omega$，，为独立同分布随机场，并满足对所有$d \in \mathfrak{D}$, $\theta \in \Theta$， X_d^θ均为$(\mathcal{X}_d \otimes 2^A \otimes \mathcal{F}_d^\theta)/\mathcal{X}_d$-可测；假设 $M > (1 + (\sup_{d \in \mathfrak{D}} \lambda_d \delta_d)(2|A| - 1))^2 (1 - (\sup_{d \in \mathfrak{D}} \lambda_d \delta_d))^2$ ；假设对所有 $d \in \mathfrak{D}$ $x\ a\ \in \mathbb{X}_d \times A$ $_{b \in A} |g_d\ x\ b\ | \leq \kappa \mathfrak{w}_d\ x$ $\big(\mathbb{E}\big[\ |\mathfrak{w}_d\ X_d^{0\ x\ a}|^2\ \big]\big)^{\frac{1}{2}} \leq$，$(\ )$，有 $\lambda_d \mathfrak{w}_d\ x$ $_{u \in \mathfrak{D}} \lambda_u \delta_u < 1$ $d \in \mathfrak{D}$ $Q_{n\ d}^\theta$ $\mathbb{X}_d \times A \times \quad \to \mathbb{R}$ $\theta \in \Theta$ $n \in \mathbb{N}_0$ $max\,(\ )\,(\ )$，$(\ ,\ )\,(\ )$，sup；对于每个，令，$: \Omega$，，，满足对所有 $n \in \mathbb{N}_0$, $(x, a) \in \mathbb{X}_d \times A$, $\theta \in \Theta$，有

$$Q_{n,d}^\theta(x, a) = g_d(x, a) + \sum_{l=0}^{n-1} \frac{\delta_d}{M^{n-l}} \sum_{i=1}^{M^{n-l}} \max_{b \in A} \big\{ Q_{l,d}^{(\theta,l,i)}(X_d^{(\theta,l,i),x,a}, b) \big\} \tag{96}$$
$$- {}_{\mathbb{I}^{\mathbb{N}}}(l) \max_{b \in A} \big\{ Q_{\max\{l-1,0\},d}^{(\theta,-l,i)}(X_d^{(\theta,l,i),x,a}, b) \big\},$$

并令 $\mathfrak{C}_{n,d} \in [0, \infty)$、 $n \in \mathbb{N}_0$、 $d \in \mathfrak{D}$ 满足对所有 $n \in \mathbb{N}_0$、 $d \in \mathfrak{D}$ ，均有

$$\mathfrak{C}_{n,d} \leq \sum_{l=0}^{n-1} M^{n-l} \big(\mathfrak{R}_d + \mathfrak{C}_{l,d} + {}_{\mathbb{I}^{\mathbb{N}}}(l) \mathfrak{C}_{\max\{l-1,0\},d} \big). \tag{97}$$

则以下结论成立：

(i) 对于每个 $d \in \mathfrak{D}$，存在唯一函数 Q_d: $\mathbb{X}_d \times A \to \mathbb{R}$，它是 $(\mathcal{X}_d \otimes 2^A)/\mathcal{B}(\mathbb{R})$-可测的，并且对所有 $(x, a) \in \mathbb{X}_d \times A$，满足 $sup_{y \in \mathbb{X}_d} max_{b \in A} |Q_d($ $_{\mathfrak{w}_d\ y}{}^{y\ b}|\ ,\ )\ (\ ) < \infty, \mathbb{E}\big[\big| max_{b \in A} Q_d(X^0, {}^{x\ a}_d\ ,\ , b)\big|\big] < \infty$，且

$$Q_d(x, a) = g_d(x, a) + \delta_d \mathbb{E}\big[\max_{b \in A} Q_d(X_d^{0,x,a}, b) \big]. \tag{98}$$

(ii) 存在 $N\colon(0,1]\to\mathbb{N}$ 和 $c\in\mathbb{R}$，使得对所有 $d\in\mathfrak{D}$, $\varepsilon\in(0,1]$，均有 $\mathfrak{C}_{N_\varepsilon,d}\le c\mathfrak{R}_d\varepsilon^{-c}$，且

$$\sup_{x\in\mathbb{X}_d}\left(\frac{\mathbb{E}\big[\max_{a\in A}|Q_d(x,a)-Q^0_{N_\varepsilon,d}(x,a)|^2\big]}{|\mathfrak{w}_d(x)|^2}\right)^{\frac12}\le\varepsilon. \tag{99}$$

推论 *4.2* 的证明。注意，对于每个 $d\in\mathfrak{D}$，函数 $\mathbb{X}_d\times\mathbb{R}^A\ni(x,r)\mapsto\delta_d\max_{a\in A}\{g_d(x,a)+r(a)\}\in\mathbb{R}$ 都是 $(\mathcal{X}_d\otimes(\bigotimes_{a\in A}\mathcal{B}(\mathbb{R})))/\mathcal{B}(\mathbb{R})$-可测的。此外，对于所有 $d\in\mathfrak{D}$, $x\in\mathbb{X}_d$, $r,s\in\mathbb{R}^A$，都有

$$\left|\delta_d\max_{a\in A}\{g_d(x,a)+r(a)\}-\delta_d\max_{a\in A}\{g_d(x,a)+s(a)\}\right|\le\delta_d\max_{a\in A}|r(a)-s(a)|. \tag{100}$$

假设对于所有 $d\in\mathfrak{D}$, $(x,a)\in\mathbb{X}_d\times A$ 均有 $\max_{b\in A}|g_d(x,b)|\le\kappa\mathfrak{w}_d(x)$，$\big(\mathbb{E}\big[\big|\mathfrak{w}_d(X^{0,x,a}_d)\big|^2\big]\big)^{\frac12}\le\lambda_d\mathfrak{w}_d(x)$，且 $\sup_{u\in\mathfrak{D}}\lambda_u\delta_u<1$，则对于所有 $d\in\mathfrak{D}$, $(x,a)\in\mathbb{X}_d\times A$ 均有

$$\left(\mathbb{E}\left[\left|\delta_d\max_{b\in A}\big\{g_d(X^{0,x,a}_d,b)\big\}\right|^2\right]\right)^{\frac12}\le\delta_d\kappa\left(\mathbb{E}\left[\big|\mathfrak{w}_d(X^{0,x,a}_d)\big|^2\right]\right)^{\frac12}\le\delta_d\kappa\lambda_d\mathfrak{w}_d(x)\le\kappa\mathfrak{w}_d(x). \tag{101}$$

对每个 $d\in\mathfrak{D}$，令 $R^\theta_{n,d}\colon\ \mathbb{X}_d\times A\to\mathbb{R}$、$\theta\in\Theta$、$n\in\mathbb{N}_0$，对所有 $n\in\mathbb{N}_0$、$(x,a)\in\mathbb{X}_d\times A$、$\theta\in\Theta$ 满足

$$\begin{aligned}R^\theta_{n,d}(x,a)=\sum_{l=0}^{n-1}\frac{\delta_d}{M^{n-l}}\sum_{i=1}^{M^{n-l}}&\max_{b\in A}\Big\{g_d(X^{(\theta,l,i),x,a}_d,b)+R^{(\theta,l,i)}_{l,d}\big(X^{(\theta,l,i),x,a}_d,b\big)\Big\}\\&-\mathbb{I}^{\mathbb{N}}(l)\max_{b\in A}\Big\{g_d(X^{(\theta,l,i),x,a}_d,b)+R^{(\theta,-l,i)}_{\max\{l-1,0\},d}\big(X^{(\theta,l,i),x,a}_d,b\big)\Big\}.\end{aligned} \tag{102}$$

由此以及定理 4.1（应用时取 $M\leftarrow M$、$\Theta\leftarrow\Theta$、$A\leftarrow A$、$\kappa\leftarrow\kappa$、$(\Omega,\mathcal{F},\mathbb{P})\leftarrow(\Omega,\mathcal{F},\mathbb{P})$、$\mathfrak{D}\leftarrow\mathfrak{D}$、$(\lambda_d)_{d\in\mathfrak{D}}\leftarrow(\lambda_d)_{d\in\mathfrak{D}}$、$(L_d)_{d\in\mathfrak{D}}\leftarrow(\delta_d)_{d\in\mathfrak{D}}$、$(\mathfrak{R}_d)_{d\in\mathfrak{D}}\leftarrow(\mathfrak{R}_d)_{d\in\mathfrak{D}}$、$(\mathbb{X}_d,\mathcal{X}_d)_{d\in\mathfrak{D}}\leftarrow(\mathbb{X}_d,\mathcal{X}_d)_{d\in\mathfrak{D}}$、$(\mathfrak{w}_d)_{d\in\mathfrak{D}}\leftarrow(\mathfrak{w}_d)_{d\in\mathfrak{D}}$、$(f_d)_{d\in\mathfrak{D}}\leftarrow(\mathbb{X}_d\times\mathbb{R}^A\ni(x,r)\mapsto\delta_d\max_{a\in A}\{g_d(x,a)+r(a)\}\in\mathbb{R})_{d\in\mathfrak{D}}$、$(\mathcal{F}^\theta_d)_{d\in\mathfrak{D}}$、$\theta\in\Theta\leftarrow(\mathcal{F}^\theta_d)_{d\in\mathfrak{D}}$、$\theta\in\Theta$、$(X^\theta_d)_{d\in\mathfrak{D}}$、$\theta\in\Theta\leftarrow(X^\theta_d)_{d\in\mathfrak{D}}$、$\theta\in\Theta$、$(\mathfrak{C}_{n,d})_{n\in\mathbb{N}_0}$、$d\in\mathfrak{D}\leftarrow(\mathfrak{C}_{n,d})_{n\in\mathbb{N}_0}$、$d\in\mathfrak{D}$、$(V^\theta_{n,d})_{n\in\mathbb{N}_0}$、$d\in\mathfrak{D}$、$\theta\in\Theta\leftarrow(R^\theta_{n,d})_{n\in\mathbb{N}_0}$、$d\in\mathfrak{D}$、$\theta\in\Theta$，采用定理 4.1 的记号）可得

(1) 对每个 $d\in\mathfrak{D}$，都存在唯一函数 $R_d\colon\ \mathbb{X}_d\times A\to\mathbb{R}$，它是 $(\mathcal{X}_d\otimes 2^A)/\mathcal{B}(\mathbb{R})$-可测的，并且对所有 $(x,a)\in\mathbb{X}_d\times A$，均满足 $\sup_{y\in\mathbb{X}_d}\max_{b\in A}|R_d(y,b)|/\mathfrak{w}_d(y)<\infty$，$\mathbb{E}\big[\big|\max_{b\in A}\{g_d(X^{0,x,a}_d,b)+R_d\big(X^{0,x,a}_d,b\big)\}\big|\big]<\infty$，

$$R_d(x,a)=\delta_d\mathbb{E}\big[\max_{b\in A}\big\{g_d(X^{0,x,a}_d,b)+R_d\big(X^{0,x,a}_d,b\big)\big\}\big], \tag{103}$$

and

(2) 存在 $N\colon(0,1]\to\mathbb{N}$ 和 $c\in\mathbb{R}$，使得对于所有 $d\in\mathfrak{D}$, $\varepsilon\in(0,1]$，均有 $\mathfrak{C}_{N_\varepsilon,d}\le c\mathfrak{R}_d\varepsilon^{-c}$ 并且

$$\sup_{x\in\mathbb{X}_d}\left(\frac{\mathbb{E}\big[\max_{a\in A}|R_d(x,a)-R^0_{N_\varepsilon,d}(x,a)|^2\big]}{|\mathfrak{w}_d(x)|^2}\right)^{\frac12}\le\varepsilon. \tag{104}$$

对于每个 $d \in \mathfrak{D}$，令 $Q_d\colon \mathbb{X}_d \times A \to \mathbb{R}$，使得对于所有 $(x,a) \in \mathbb{X}_d \times A$，均有 $Q_d(x,a) = g_d(x,a) + R_d(x,a)$。这一点与条目 (1) 一起确保对于所有 $d \in \mathfrak{D}$，$(x,a) \in \mathbb{X}_d \times A$，均有

$$\begin{aligned} Q_d(x,a) &= g_d(x,a) + R_d(x,a) = g_d(x,a) + \delta_d \mathbb{E}\left[\max_{b\in A}\left\{g_d(X_d^{0,x,a},b) + R_d(X_d^{0,x,a},b)\right\}\right] \\ &= g_d(x,a) + \delta_d \mathbb{E}\left[\max_{b\in A}\left\{Q_d(X_d^{0,x,a},b)\right\}\right]. \end{aligned} \tag{105}$$

此外，注意到以下假设：对于所有 $d \in \mathfrak{D}$，$x \in \mathbb{X}_d$，均有 $\max_{a\in A}|g_d(x,a)| \le \kappa \mathfrak{w}_d(x)$；结合条件 (1) 可知，对于所有 $d \in \mathfrak{D}, (x,a) \in \mathbb{X}_d \times A$，均有 $\sup_{y\in\mathbb{X}_d} \max_{b\in A} |Q_d(x,a) - Q_d(y,b)| / (\mathfrak{w}_d(x) + \mathfrak{w}_d(y)) < \infty$ 且 $\mathbb{E}\left[\left|\max_{b\in A}\left\{Q_d(X_d^{0,x,a},b)\right\}\right|\right] < \infty$。~~此外，对于每个~~ $d \in \mathfrak{D}$，令 $S_d\colon \mathbb{X}_d \times A \to \mathbb{R}$ 为 $(\mathcal{X}_d \otimes 2^A)/\mathcal{B}(\mathbb{R})$-可测的，并满足对于所有 $d \in \mathfrak{D}, (x,a) \in \mathbb{X}_d \times A$，有 $\sup_{y\in\mathbb{X}_d} \max_{b\in A} |S_d(x,a) - S_d(y,b)| / (\mathfrak{w}_d(x) + \mathfrak{w}_d(y)) < \infty$，$\mathbb{E}\left[\left|\max_{b\in A}\left\{S_d(X_d^{0,x,a},b)\right\}\right|\right] < \infty$，且 $S_d(x,a) = g_d(x,a) + \delta_d \mathbb{E}\left[\max_{b\in A}\left\{S_d(X_d^{0,x,a},b)\right\}\right]$。对于所有 $d \in \mathfrak{D}, (x,a) \in \mathbb{X}_d \times A$，均有

$$\begin{aligned} S_d(x,a) - g_d(x,a) &= \delta_d \mathbb{E}\left[\max_{b\in A}\left\{S_d(X_d^{0,x,a},b)\right\}\right] \\ &= \delta_d \mathbb{E}\left[\max_{b\in A}\left\{g_d(X_d^{0,x,a},b) + S_d(X_d^{0,x,a},b) - g_d(X_d^{0,x,a},b)\right\}\right]. \end{aligned} \tag{106}$$

条目 (1) 表明，对于所有 $d \in \mathfrak{D}$，都有 $S_d - g_d = R_d$。由此可得，对于所有 $d \in \mathfrak{D}$，都有 $S_d = Q_d$。这证明了条目 (i)。注意，归纳法和 (102) 表明，对于所有 $d \in \mathfrak{D}, n \in \mathbb{N}_0$，$(x,a) \in \mathbb{X}_d \times A$, $\theta \in \Theta$ 都有 $Q_{n,d}^{\theta}(x,a) = g_d(x,a) + R_{n,d}^{\theta}(x,a)$。结合这一点和条目 (2)，可得对于所有 $d \in \mathfrak{D}, \varepsilon \in (0,1]$，都有

$$\sup_{x\in\mathbb{X}_d}\left(\frac{\mathbb{E}\left[\max_{a\in A}|Q_d(x,a) - Q_{N_\varepsilon,d}^0(x,a)|^2\right]}{|\mathfrak{w}_d(x)|^2}\right)^{\frac{1}{2}} = \sup_{x\in\mathbb{X}_d}\left(\frac{\mathbb{E}\left[\max_{a\in A}|R_d(x,a) - R_{N_\varepsilon,d}^0(x,a)|^2\right]}{|\mathfrak{w}_d(x)|^2}\right)^{\frac{1}{2}} \le \varepsilon. \tag{107}$$

这证明了条目 (ii)。至此，推论 4.2 的证明完成。 □

4.3 最优停止问题的 Bellman 方程的 MLFP 近似

推论 *4.3*。设 $M \in \mathbb{N}$，设 $\Theta = \bigcup_{n\in\mathbb{N}} \mathbb{Z}^n$，设 $(\Omega, \mathcal{F}, \mathbb{P})$ 为概率空间，设 $\mathfrak{D}$ 为非空集合，设 $\delta_d, \mathfrak{R}_d \in [0,\infty)$, $d \in \mathfrak{D}$，设 $(\mathbb{X}_d, \mathcal{X}_d)$, $d \in \mathfrak{D}$ 为非空 *Borel*空间，对于每个$d \in \mathfrak{D}$，设 $g_d\colon \mathbb{X}_d \to \mathbb{R}$ 和 $G_d\colon \mathbb{X}_d \to \mathbb{R}$ 为 $\mathcal{X}_d/\mathcal{B}(\mathbb{R})$-可测，对于每个 $d \in \mathfrak{D}$，设 $(\mathcal{F}_d^{\theta})_{\theta\in\Theta}$ 为 $\mathcal{F}$ 的独立子-σ-代数，对于每个 $d \in \mathfrak{D}$，设 $X_d^{\theta} = (X_d^{\theta,x}(\omega))_{x\in\mathbb{X}_d,\,\omega\in\Omega}\colon \mathbb{X}_d \times \Omega \to \mathbb{X}_d$, $\theta \in \Theta$ 为独立同分布随机场，且满足对于所有 $d \in \mathfrak{D}, \theta \in \Theta$，$X_d^{\theta}$ 都是 $(\mathcal{X}_d \otimes \mathcal{F}_d^{\theta})/\mathcal{X}_d$-可测的。假设 $\sup_{d\in\mathfrak{D}} \delta_d < 1$，假设 $M > (1+3(\sup_{d\in\mathfrak{D}} \delta_d))^2 (1-(\sup_{d\in\mathfrak{D}} \delta_d))^2$，假设 $\sup_{d\in\mathfrak{D}} (\sup_{x\in\mathbb{X}_d} |g_d(x)| + |G_d(x)|) < \infty$，对于每个 $d \in \mathfrak{D}$，设 $Q_{n,d}^{\theta}\colon \mathbb{X}_d \times \Omega \to \mathbb{R}$, $n \in \mathbb{N}_0$, $\theta \in \Theta$，满足对于所有 $n \in \mathbb{N}_0$, $x \in \mathbb{X}_d$, $\theta \in \Theta$，都有

$$\begin{aligned} Q_{n,d}^{\theta}(x) = g_d(x) &+ \sum_{l=0}^{n-1} \frac{\delta_d}{M^{n-l}} \sum_{i=1}^{M^{n-l}} \max\left\{G_d\big(X_d^{(\theta,l,i),x}\big), Q_{l,d}^{(\theta,l,i)}\big(X_d^{(\theta,l,i),x}\big)\right\} \\ &- \mathbb{I}^{\mathbb{N}}(l) \max\left\{G_d\big(X_d^{(\theta,l,i),x}\big), Q_{\max\{l-1,0\},d}^{(\theta,-l,i)}\big(X_d^{(\theta,l,i),x}\big)\right\} \end{aligned} \tag{108}$$

并令 $\mathfrak{C}_{n,d} \in [0,\infty)$、$n \in \mathbb{N}_0$、$d \in \mathfrak{D}$ 对所有 $n \in \mathbb{N}_0$、$d \in \mathfrak{D}$ 均满足

$$\mathfrak{C}_{n,d} \le \sum_{l=0}^{n-1} M^{n-l}\left(\mathfrak{R}_d + \mathfrak{C}_{l,d} + \mathbb{I}^{\mathbb{N}}(l)\mathfrak{C}_{\max\{l-1,0\},d}\right). \tag{109}$$

则有以下结论:

(i) 对于每个 $d\in\mathfrak{D}$，都存在唯一函数 $Q_d\colon \mathbb{X}_d\to\mathbb{R}$，它是 $\mathcal{X}_d/\mathcal{B}(\mathbb{R})$-可测的，并且对所有 $x\in\mathbb{X}_d$ 都满足 $\sup_{y\in\mathbb{X}_d}|Q_d(y)|<\infty$，且

$$Q_d(x)=g_d(x)+\delta_d\mathbb{E}\left[\max\left\{G_d(X_d^{0,x}),Q_d(X_d^{0,x})\right\}\right]. \tag{110}$$

(ii) 存在 $N\colon(0,1]\to\mathbb{N}$ 和 $c\in\mathbb{R}$，使得对所有 $d\in\mathfrak{D}$，$\varepsilon\in(0,1]$ 都有$\mathfrak{C}_{N_\varepsilon,d}\le c\mathfrak{R}_d\varepsilon^{-c}$，且

$$\sup_{x\in\mathbb{X}_d}\left(\mathbb{E}\left[\left|Q_d(x)-Q^0_{N_\varepsilon,d}(x)\right|^2\right]\right)^{\frac{1}{2}}\le\varepsilon. \tag{111}$$

推论 *4.3* 的证明。设 $\mathbb{Y}_d$、$d\in\mathfrak{D}$ 为非空集合，满足存在 $\Upsilon\in\bigcap_{d\in\mathfrak{D}}\mathbb{Y}_d$，且对所有 $d\in\mathfrak{D}$ 都有 $\mathbb{Y}_d\setminus\{\Upsilon\}=\mathbb{X}_d$。对于每个 $d\in\mathfrak{D}$，令 $\mathcal{Y}_d=\sigma_{\mathbb{Y}_d}(\mathcal{X}_d)$；对于每个 $d\in\mathfrak{D}$，$\theta\in\Theta$，令 $Y_d^\theta=(Y^{\theta,y\ a}_{\ \ ,d})_{y\in\mathbb{Y}_d,\,a\in\{0,1\},\,\omega\in\Omega}\colon\mathbb{Y}_d\times\{0,1\}\times\Omega\to\mathbb{Y}_d$ 满足对所有$d\in\mathfrak{D}$，$\theta\in\Theta$，$(y,\ a)\in\mathbb{Y}_d\times\{0,1\}$，$\omega\in\Omega$，有

$$Y_d^{\theta,y,a}(\omega)=\begin{cases}\Upsilon & :(y=\Upsilon)\vee(a=0),\\ X_d^{\theta,y}(\omega) & :(y\neq\Upsilon)\wedge(a=1).\end{cases} \tag{112}$$

注意，对于所有 $d\in\mathfrak{D}$，都有 Y_d^θ、$\theta\in\Theta$ 是独立同分布随机场，并且对于每个 $d\in\mathfrak{D}$，$\theta\in\Theta$ 都是 $(\mathcal{Y}_d\otimes 2^{\{0,1\}}\otimes\mathcal{F}_d^\theta)/\mathcal{Y}_d$-可测的。对于每个 $d\in\mathfrak{D}$，令 $h_d\colon\mathbb{Y}_d\times\{0,1\}\to\mathbb{R}$ 满足对所有 $d\in\mathfrak{D}$，$(y,a)\in\mathbb{Y}_d\times\{0,1\}$ 均有

$$h_d(y,a)=\begin{cases}0 & :y=\Upsilon,\\ g_d(y) & :y\neq\Upsilon,\ a=1,\\ G_d(y) & :y\neq\Upsilon,\ a=0.\end{cases} \tag{113}$$

注意，对于所有 $d\in\mathfrak{D}$，h_d 都是 $(\mathcal{Y}_d\otimes 2^{\{0,1\}})/\mathcal{B}(\mathbb{R})$-可测的。此外，对于所有 $d\in\mathfrak{D}$，$(y,a)\in\mathbb{X}_d\times\{0,1\}$，都有 $|h_d(y,a)|\le\max\{|g_d(y)|,|G_d(y)|\}\le\sup_{u\in\mathfrak{D}}\left(\sup_{x\in\mathbb{X}_d}|g_u(x)|+|G_u(x)|\right)$。对于每个 $d\in\mathfrak{D}$，令 $q^\theta_{n,\ d}\colon\mathbb{Y}_d\times\{0,1\}\times\Omega\to\mathbb{Y}_d$，$\theta\in\Theta$，$n\in\mathbb{N}_0$ 满足对所有 $n\in\mathbb{N}_0$，$\theta\in\Theta$，$(y,a)\in\mathbb{Y}_d\times\{0,1\}$ 均有

$$\begin{aligned}q^\theta_{n,d}(y,a)=h_d(y,a)+\sum_{l=0}^{n-1}\frac{\delta_d}{M^{n-l}}\sum_{i=1}^{M^{n-l}}&\max_{b\in\{0,1\}}\left\{q^{(\theta,l,i)}_{l,d}(Y_d^{(\theta,l,i),y,a},b)\right\}\\ &-\mathbb{I}^{\mathbb{N}}(l)\max_{b\in\{0,1\}}\left\{q^{(\theta,-l,i)}_{\max\{l-1,0\},d}(Y_d^{(\theta,l,i),y,a},b)\right\}.\end{aligned} \tag{114}$$

推论 4.2（应用于 $A\leftarrow\{0,1\}$，$\kappa\leftarrow\sup_{d\in\mathfrak{D}}\left(\sup_{x\in\mathbb{X}_d}|g_d(x)|+|G_d(x)|\right)$，$g_d\leftarrow h_d$，$(\mathbb{X}_d,\mathcal{X}_d)\leftarrow(\mathbb{Y}_d,\mathcal{Y}_d)$，$Q^\theta_{n,d}\leftarrow q^\theta_{n,d}$ 在推论 4.2 的记号下的 $d\in\mathfrak{D}$）得出如下结果:

(1) 对于每个 $d\in\mathfrak{D}$，都存在一个唯一函数 $q_d\colon\mathbb{Y}_d\times\{0,1\}\to\mathbb{R}$，它是 $(\mathcal{Y}_d\otimes 2^{\{0,1\}})/\mathcal{B}(\mathbb{R})$-可测的，并且对所有 $(y,a)\in\mathbb{Y}_d\times\{0,1\}$，都有 $\sup_{(z,b)\in\mathbb{Y}_d\times\{0,1\}}|q_d(z,b)|<\infty$，且

$$q_d(y,a)=h_d(y,a)+\delta_d\mathbb{E}\big[\max_{b\in\{0,1\}}q_d(Y_d^{0,y,a},b)\big]. \tag{115}$$

(2) 存在 $N\colon (0,1] \to \mathbb{N}$ 和 $c \in \mathbb{R}$，使得对所有 $d \in \mathfrak{D}$, $\varepsilon \in (0,1]$，都有$\mathfrak{C}_{N_\varepsilon,\, d} \le c\mathfrak{R}_d\varepsilon^c$，且

$$\sup_{y\in\mathbb{Y}_d}\left(\mathbb{E}\Big[\max_{a\in\{0,1\}}|q_d(y,a)-q^0_{N_\varepsilon,d}(y,a)|^2\Big]\right)^{\frac{1}{2}} \le \varepsilon. \tag{116}$$

注意，对于所有 $d\in\mathfrak{D}$，$(y,a)\in\mathbb{Y}_d\setminus\{\Upsilon\}\times\{0,1\}$，都有 $q_d(\Upsilon,a)=0$，且 $q_d(y,0)=G_d(y)$。此外，对于所有 $d\in\mathfrak{D}$, $y\in\mathbb{Y}_d\setminus\{\Upsilon\}$，都有

$$\begin{aligned} q_d(y,1) &= h_d(y,1)+\delta_d\mathbb{E}\big[\max\big\{q_d(Y_d^{0,y,1},0),q_d(Y_d^{0,y,1},1)\big\}\big] \\ &= g_d(y)+\delta_d\mathbb{E}\left[\max\left\{G_d(X_d^{0,y}),q_d(X_d^{0,y},1)\right\}\right]. \end{aligned} \tag{117}$$

由此及条目 (1) 可知，对于所有 $d\in\mathfrak{D}$，都存在唯一函数 $Q_d\colon \mathbb{X}_d\to\mathbb{R}$，它是 $\mathcal{X}_d/\mathcal{B}(\mathbb{R})$-可测的，并且对所有 $x\in\mathbb{X}_d$，都有 $\sup_{y\in\mathbb{X}_d}|Q_d(y)|<\infty$，且

$$Q_d(x)=g_d(x)+\delta_d\mathbb{E}\left[\max\left\{G_d(X_d^{0,x}),Q_d(X_d^{0,x})\right\}\right]. \tag{118}$$

这证明了条目 (i)。此外，注意，对于所有 $d\in\mathfrak{D}$, $\theta\in\Theta$，$(y,a)\in\mathbb{Y}_d\setminus\{\Upsilon\}\times\{0,1\}$，都有 $\mathbb{P}$-几乎处处成立：$q^\theta_{0,d}(\Upsilon,a)=0$, $q^\theta_{0,d}(y,0)=G_d(y)$，且 $q^\theta_{0,d}(y,1)=g_d(y)$。由此及归纳法可得，对于所有 $d\in\mathfrak{D}$, $\theta\in\Theta$, $n\in\mathbb{N}_0$, $(y,a)\in\mathbb{Y}_d\setminus\{\Upsilon\}\times\{0,1\}$，都有 $\mathbb{P}$-几乎处处成立：$q^\theta_{n,d}(\Upsilon,a)=0$，且 $q^\theta_{n,d}(y,0)=G_d(y)$。结合这些结果、(112)、(113) 和 (114) 可知，对于所有 $d\in\mathfrak{D}$,$\theta\in\Theta$, $n\in\mathbb{N}_0$, $y\in\mathbb{Y}_d\setminus\{\Upsilon\}=\mathbb{X}_d$，都有 $\mathbb{P}$-几乎处处成立：

$$\begin{aligned} q^\theta_{n,d}(y,1) &= h_d(y,1)+\sum_{l=0}^{n-1}\frac{\delta_d}{M^{n-l}}\sum_{i=1}^{M^{n-l}}\max_{b\in\{0,1\}}\left\{q^{(\theta,l,i)}_{l,d}(Y_d^{(\theta,l,i),y,1},b)\right\} \\ &\qquad - \mathbb{I}^{\mathbb{N}}(l)\max_{b\in\{0,1\}}\left\{q^{(\theta,-l,i)}_{\max\{l-1,0\},d}(Y_d^{(\theta,l,i),y,1},b)\right\} \\ &= g_d(y)+\sum_{l=0}^{n-1}\frac{\delta_d}{M^{n-l}}\sum_{i=1}^{M^{n-l}}\max\left\{G_d(X_d^{(\theta,l,i),y}),q^{(\theta,l,i)}_{l,d}(X_d^{(\theta,l,i),y},1)\right\} \\ &\qquad - \mathbb{I}^{\mathbb{N}}(l)\max\left\{G_d(X_d^{(\theta,l,i),y}),q^{(\theta,-l,i)}_{\max\{l-1,0\},d}(X_d^{(\theta,l,i),y},1)\right\}. \end{aligned} \tag{119}$$

由此及（108）可得，对所有 $d\in\mathfrak{D}$, $\theta\in\Theta$, $n\in\mathbb{N}_0$, $y\in\mathbb{Y}_d\setminus\{\Upsilon\}=\mathbb{X}_d$，$\mathbb{P}$-几乎处处有

$$q^\theta_{n,d}(y,1)=Q^\theta_{n,d}(y). \tag{120}$$

结合此结果与（116）可知，对所有 $d\in\mathfrak{D}$, $\varepsilon\in(0,1]$，有

$$\sup_{x\in\mathbb{X}_d}\left(\mathbb{E}\Big[\big|Q_d(x)-Q^0_{N_\varepsilon,d}(x)\big|^2\Big]\right)^{\frac{1}{2}} \le \sup_{y\in\mathbb{Y}_d}\left(\mathbb{E}\Big[\max_{a\in\{0,1\}}\big|q_d(y,a)-q^0_{N_\varepsilon,d}(y,a)\big|^2\Big]\right)^{\frac{1}{2}} \le \varepsilon. \tag{121}$$

这就证明了条目（ii）。至此，推论 4.3 的证明完成。 □

致谢

本研究部分由德国研究基金会（DFG，德国研究基金会）通过研究资助项目 KR 5294/2-1 资助。我们也衷心感谢卓越集群 EXC 2044-390685587「明斯特数学：动力学—几何—结构」，该项目由德国研究基金会（DFG，德国研究基金会）资助。

参考文献

[1] Andersen, L., and Broadie, M. 用于多维美式期权定价的原始-对偶模拟算法。*Management Science 50*, 9 (2004), 1222–1234.[2] Arulkumaran, K., Deisenroth, M. P., Brundage, M., and Bharath, A. A. 深度强化学习：简要综述。*IEEE Signal Processing Magazine 34*, 6 (2017), 26–38.[3] Beck, C., Gonon, L., and Jentzen, A. 克服高维半线性椭圆偏微分方程数值逼近中的维数灾难。*arXiv:2003.00596* (2020).[4] Beck, C., Hornung, F., Hutzenthaler, M., Jentzen, A., and Kruse, T. 通过截断全历史递归多层Picard逼近，克服Allen-Cahn偏微分方程数值逼近中的维数灾难。*Journal of Numerical Mathematics 28*, 4 (2020), 197–222.[5] Beck, C., Jentzen, A., and Kruse, T. 用于高维迭代嵌套期望的多项式运行时间非线性蒙特卡罗方法。*arXiv:2009.13989* (2020).[6]Becker, S., Braunwarth, R., Hutzenthaler, M., Jentzen, A., and von Wurstemberger, P. 高维偏微分方程组全历史递归多层Picard逼近的数值模拟。*Communica-tions in Computational Physics 28*, 5 (2020), 2109–2138.[7] Becker, S., Cheridito, P., and Jentzen, A. 深度最优停止。*Journal ofMachine Learning Research 20* (2019), 74.[8] Becker, S., Cheridito, P., Jentzen, A., and Welti, T. 使用深度学习求解高维最优停止问题。*European Journal of Applied Mathematics32*, 3 (2021), 470–514.[9] Bellman, R. 动态规划。Princeton Landmarks in Mathematics。普林斯顿大学出版社，新泽西州普林斯顿，2010年。重印自1957年版，并附有Stuart Dreyfus撰写的新序言。[10] Belomestny, D., Bender, C., and Schoenmakers, J. 通过非嵌套蒙特卡罗获得Bermudan乘积的真实上界。*Mathematical Finance: An International Journalof Mathematics, Statistics and Financial Economics 19*, 1 (2009), 53–71.[11] Bertsekas, D. 强化学习与最优控制。Athena Scientific，2019年。[12] Bertsekas, D., and Shreve, S. E. 随机最优控制：离散时间情形，第5卷。Athena Scientific，1996年。[13] Bertsekas, D., and Tsitsiklis, J. N. 神经动态规划。Athena Scientific，1996年。[14] Bray, R. L. 关于“利用随机化打破维数灾难”的评论。*Econometrica 90*, 4 (2022), 1915–1929.[15] Broadie, M., and Glasserman, P. 使用模拟为美式证券定价。*Journal of economic dynamics and control 21*, 8-9 (1997), 1323–1352.

[16]Broadie, M. 和 Glasserman, P. 一种用于高维美式期权定价的随机网格方法。计算金融学杂志 *7* (2004年)，35–72。[17] Chow, C.-S. 和 Tsitsiklis, J. N. 动态规划的复杂度。复杂性杂志 *5*, 4 (1989)，466–488。[18] Chow, C.-S. 和 Tsitsiklis, J. N. 一种用于离散时间随机控制的最优单向多重网格算法。*IEEE* 自动控制汇刊 *36*, 8 (1991)，898–914。[19] Desai, V. V.、Farias, V. F. 和 Moallemi, C. C. 最优停止问题的逐路径优化。*Management Science 58*, 12 (2012)，2292–2308。[20] E, W.、Hutzenthaler, M.、Jentzen, A. 和 Kruse, T. 关于高维非线性抛物型偏微分方程和高维非线性向后随机微分方程的多层 Picard 数值逼近。*Journal of Sci- entific Computing 79*, 3 (2019年)，1534–1571。[21] E, W.、Hutzenthaler, M.、Jentzen, A. 和 Kruse, T. 用于求解光滑半线性抛物型热方程的多层 Picard 迭代。偏微分方程与应用 *2*, 6 (2021年)，80。[22] François-Lavet, V.、Henderson, P.、Islam, R.、Bellemare, M. G. 和 Pineau, J. 深度强化学习简介。机器学习领域基础与趋势® *11*, 3-4 (2018)，219–354。[23] Giles, M. 使用 Milstein 格式改进多层蒙特卡洛收敛性。载于*2006*年蒙特卡洛与准蒙特卡洛方法。施普林格，2008年，第343–358页。[24] Giles, M. B.、Jentzen, A. 和 Welti, T. 广义多层 Picard 逼近。*arXiv:1911.03188* (2019年)。[25]Gonon, L. 最优停止问题的深度神经网络表达能力。*arXiv:2210.10443* (2022年)。[26] Haugh, M. B. 和 Kogan, L. 美式期权定价：一种对偶方法。运筹学 *52*, 2 (2004年)，258–270。[27] Heinrich, S. 积分方程全局求解的蒙特卡洛复杂度。复杂性杂志 *14*, 2 (1998)，151–175。[28] Heinrich, S. 多层蒙特卡洛方法。载于国际大规模科学计算会议 (2001年)，施普林格，第58–67页。[29] Hutzenthaler, M.、Jentzen, A. 和 Kruse, T. 克服数值逼近中抛物型偏微分方程梯度相关非线性项所导致的维数灾难。*Found. Comput. Math.* (2021年)，1–62。[30] Hutzenthaler, M.、Jentzen, A.、Kruse, T. 和 Nguyen, T. A. 用于具有 Lipschitz 非线性项的高维半线性二阶偏微分方程的多层 Picard 逼近。*arXiv:2009.02484* (2020年)。

[31] Hutzenthaler, M.、Jentzen, A.、Kruse, T.、Nguyen, T. A. 和 von Wurstemberger, P. 克服半线性抛物型偏微分方程数值逼近中的维数灾难。英国皇家学会会刊 *A* 辑 *476*, 2244 (2020年)，20190630。[32] Hutzenthaler, M.、Jentzen, A. 和 von Wurstemberger, P. 克服具有违约风险的金融衍生品近似定价中的维数灾难。*Electronic Journal of Probability 25* (2020年)，论文编号101，73。[33]Hutzenthaler, M. 和 Kruse, T. 对具有梯度相关非线性项的高维半线性抛物型偏微分方程的多层 Picard 逼近。*SIAM* 数值分析期刊 *58*, 2 (2020年)，929–961。[34] Hutzenthaler, M.、Kruse, T. 和 Nguyen, T. A. McKean–Vlasov 随机微分方程的多层 Picard 逼近。*Journal of Mathematical Analysis andApplications 507*, 1 (2022年)，125761。[35] Klenke, A. 概率论，第2版。施普林格伦敦，2014年。[36] Kristensen, D.、Mogensen, P. K.、Moon, J. M. 和 Schjerning, B. 使用平滑和筛法求解动态离散选择模型。计量经济学杂志 *223*, 2 (2021年)，328–360。[37] Li, Y. 深度强化学习：综述。*arXiv:1701.07274* (2017年)。[38] Longstaff, F. A. 和 Schwartz, E. S. 通过模拟为美式期权估值：一种简单的最小二乘法。金融研究评论 *14*, 1 (2001)，113–147。[39] Novak, E. 和 Ritter, K. 维数灾难与一种通用数值积分方法。载于多元逼近与样条函数（曼海姆，*1996*年），第125卷，国际数值数学丛书 Birkhäuser，巴塞尔，1997年，第177–187页。[40] Novak, E. 和 Woźniakowski, H. 多元问题的可处理性。第*1*卷：线性信息，第6卷，欧洲数学学会数学专著丛书。欧洲数学学会（EMS），苏黎世，2008年。[41] Powell, W. B. 近似动态规划：解决维数灾难，第703卷。约翰·威立父子公司，2007年。[42] Puterman, M. L. 马尔可夫决策过程：离散随机动态规划。约翰·威立父子公司，2014年。[43] Rogers, L. C. 美式期权的蒙特卡罗估值。*Mathematical Finance 12*, 3 (2002)，271–286。[44] Rummery, G. A. 和 Niranjan, M. 使用联结主义系统进行在线 *Q*学习，第37卷。剑桥大学工程系，英国剑桥，1994年。[45] Rust, J. 利用随机化打破维数灾难。*Econometrica: Journalof the Econometric Society* (1997年)，487–516。[46] Rust, J. P. 比较用于求解连续状态、无限时域马尔可夫决策问题的策略迭代方法：采用随机、准随机和确定性离散化。可见于 *SSRN*: *http://dx.doi.org/10.2139/ssrn.37768* (1997年)。

[47] Sutton, R. S. 通过时序差分方法进行预测学习。机器学习 *3* (1988), 9–44.[48] Sutton, R. S.，和 Barto, A. G. 强化学习：导论。麻省理工学院出版社，2018.[49] Tsitsiklis, J. N.，和 Van Roy, B. 马尔可夫过程的最优停止：希尔伯特空间理论、近似算法，以及在高维金融衍生品定价中的应用。*IEEE* 自动控制汇刊 *44*, 10 (1999), 1840–1851.[50] Tsitsiklis, J. N.，和 Van Roy, B. 复杂美式期权定价的回归方法。*IEEE* 神经网络汇刊 *12*, 4 (2001), 694–703.[51] WATKINS, C. 从延迟奖励中学习。博士论文。剑桥大学(1989).[52] Watkins, C.，和 Dayan, P. Q学习。机器学习 *8*, 3-4 (1992), 279–292.