Christian Beck, Arnulf Jentzen, Konrad Kleinberg, and Thomas Kruse

Nonlinear Monte Carlo methods with polynomial runtime for Bellman equations of discrete time high-dimensional stochastic optimal control problems

March 6, 2023

Nonlinear Monte Carlo methods with polynomial runtime for Bellman equations of discrete time high-dimensional stochastic optimal control problems

Christian Beck[1], Arnulf Jentzen[2,3],
Konrad Kleinberg[4], and Thomas Kruse[5]

[1] Department of Mathematics, ETH Zurich,
Switzerland; e-mail: christian.beck@yahoo.de
[2] Applied Mathematics: Institute for Analysis and Numerics,
University of Münster, Germany; e-mail: ajentzen@uni-muenster.de
[3] School of Data Science and Shenzhen Research Institute of Big Data,
The Chinese University of Hong Kong, Shenzhen,
China; e-mail: ajentzen@cuhk.edu.ch
[4] Department of Mathematics & Informatics, University of Wuppertal,
Germany; e-mail: kleinberg@uni-wuppertal.de
[5] Department of Mathematics & Informatics, University of Wuppertal,
Germany; e-mail: tkruse@uni-wuppertal.de

March 6, 2023

Abstract

Discrete time *stochastic optimal control* problems and *Markov decision processes* (MDPs), respectively, serve as fundamental models for problems that involve sequential decision making under uncertainty and as such constitute the theoretical foundation of *reinforcement learning*. In this article we study the numerical approximation of MDPs with infinite time horizon, finite control set, and general state spaces. Our set-up in particular covers infinite-horizon optimal stopping problems of discrete time Markov processes. A key tool to solve MDPs are *Bellman equations* which characterize the value functions of the MDPs and determine the optimal control strategies. By combining ideas from the *full-history recursive multilevel Picard approximation method*, which was recently introduced to solve certain nonlinear partial differential equations, and ideas from *Q-learning* we introduce a class of suitable *nonlinear Monte Carlo methods* and prove that the proposed methods do overcome the *curse of dimensionality* in the numerical approximation of the solutions of Bellman equations and the associated discrete time stochastic optimal control problems.

Contents

1 Introduction

Reinforcement learning is an active research field in machine learning and has important applications in many areas which involve sequential decision making such as economics, engineering, finance, healthcare, logistics, and robotics (see, e.g., Sutton & Barto [48], Bertsekas [11], and Bertsekas & Tistsiklis [13] for overviews of the field and its application areas). In particular the combination of deep neural networks and reinforcement learning, i.e., deep reinforcement learning, has achieved remarkable success in complex decision-making problems during recent years (see, e.g., Li [37], François-Lavet et al. [22], and Arulkumaran et al. [2] for survey articles). The mathematical foundations of reinforcement learning are provided by the theory of *stochastic optimal control* and, in particular, *Markov decision processes* (MDPs; see, e.g., Bertsekas & Shreve [12], Powell [41], and Puterman [42]).

The basis of many reinforcement learning algorithms such as temporal difference learning (see, e.g., Sutton [47]), Q-learning (see, e.g., Watkins [51] and Watkins & Dayan [52]), and the SARSA algorithm (see, e.g., Rummery & Niranjan [44]) is formed by the stochastic dynamic programming principle. It reduces the problem of making the optimal decision at a given state to solving a particular functional equation – the so-called *Bellman equation*. The approximative solution of Bellman equations in high dimensional state spaces is a notoriously difficult challenge due to the *curse of dimensionality* (cf., e.g., Bellman [9], Novak & Wozniakowski [40, Chapter 1], and Novak & Ritter [39]).

In this work we introduce *nonlinear Monte Carlo methods* for MDPs with infinite time horizon and finite control set that are polynomially tractable in the sense that the computational effort of the algorithm to approximatively compute the solution of the Bellman equation grows at most polynomially in the reciprocal $1/\epsilon$ of the prescribed approximation accuracy $\epsilon \in (0, 1]$ and the dimension $d \in \mathbb{N} = \{1, 2, 3, \ldots\}$ of the underlying state space. In particular, the proposed methods overcome the curse of dimensionality in the numerical approximation of solutions of Bellman equations.

More formally, we consider MDPs that are specified by a measurable space $(\mathbb{X}, \mathcal{X})$ (typically $(\mathbb{R}^d, \mathcal{B}(\mathbb{R}^d))$ with $d \in \mathbb{N}$ but our framework also covers discrete state spaces), a finite control set A, a discount factor $\delta \in (0, 1)$, transition kernels $\kappa_a \colon \mathbb{X} \times \mathcal{X} \to [0, 1]$, $a \in A$, and a measurable and bounded reward function $g \colon \mathbb{X} \times A \to \mathbb{R}$. The set of strategies $\mathcal{A}$ consists of all measurable

functions $\alpha\colon \mathbb{X} \to A$ (under appropriate assumptions the restriction to such non-randomized, stationary strategies is without loss of generality, see, e.g., [12, Chapter 9]). Every $\alpha \in \mathcal{A}$ defines a Markov process $X^\alpha = (X_k^\alpha)_{k\in\mathbb{N}_0}$ on a filtered probability space $(\Omega, \mathcal{F}, (\mathcal{F}_k)_{k\in\mathbb{N}_0}, \mathbb{P})$ with state space $(\mathbb{X}, \mathcal{X})$ such that $\mathbb{P}[X_{k+1}^\alpha \in B|\mathcal{F}_k] = \kappa_{\alpha(X_k^\alpha)}(B, X_k^\alpha)$ for all $k \in \mathbb{N}_0$, $B \in \mathcal{X}$. The expected gain of a strategy $\alpha \in \mathcal{A}$ when starting in $x \in \mathbb{X}$ is

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

The value function $v\colon \mathbb{X} \to \mathbb{R}$ is given by $v(x) = \sup_{\alpha\in\mathcal{A}} J(x, \alpha)$, $x \in \mathbb{X}$. The dynamic programming principle ensures that under appropriate conditions v satisfies for all $x \in \mathbb{X}$ that

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

where it holds for all $x \in \mathbb{X}$, $a \in A$ that the random variable $X^{x,a}\colon \Omega \to \mathbb{X}$ has distribution $\kappa_a(x, \cdot)$ (see, e.g., [12, Chapter 9] for a proof). An important subclass of these MDPs is given by optimal stopping problems in discrete time with infinite time horizon. For such problems the control set A consists of the two elements "stop" and "continue" and once the control "stop" is taken the Markov process jumps to a hold state from which it cannot leave and generates no further gains.

There is a large number of numerical approximation methods for Bellman equations of the form (2) that have been proposed and analyzed in the scientific literature. Deterministic numerical approximation methods for Bellman equations suffer in general from the curse of dimensionality. Indeed, Chow & Tsitsiklis [17, 18] show for all $d \in \mathbb{N}$ that in the case where the state space $\mathbb{X}$ is given by the d-dimensional unit cube $[0,1]^d$ deterministic numerical approximation methods need at least $O(\epsilon^{-2d})$ computational operations to approximate the solution of the Bellman equation with precision $\epsilon \in (0,1]$. Rust [45] shows that under certain assumptions it is possible to overcome this curse of dimensionality by allowing for randomized algorithms. Rust's method consists of randomly sampling grid points in the state space and performing value iteration on this stochastic grid. The method is model-based as it requires explicit knowledge of the transition density. In the companion paper [46] Rust points out that convergence of his method might break down in situations where the transition density has spikes. Also Kristensen et al. [36] report a dramatic rise in the method's variance as the state dimension increases. This is theoretically confirmed by Bray [14] who proves that Rust's method only overcomes the curse of dimensionality in the special case where the MDP is equivalent to an MDP where all but a vanishingly small fraction of state variables behave like history-independent uniform random variables. Variants of Rust's method for optimal stopping problems are designed by Broadie & Glasserman in [15] and [16]. The literature on optimal stopping comprises a variety of further randomized algorithms. We refer, for example, to [38], [49], and [50] for regression-based algorithms for optimal stopping problems, we refer, for example, to [43], [1], [26], [19], and [10] for duality-based algorithms for optimal stopping problems, and we refer, for example to [7], [8], and [25] for deep learning-based algorithms for optimal stopping problems.

In this paper we introduce nonlinear Monte Carlo algorithms that overcome the curse of dimensionality in the approximation of Bellman equations, that require only very weak regularity assumptions and that are model-free in the sense that they do not use explicit knowledge of the transition kernel κ but only need access to independent realizations of the one step transitions $X^{x,a}$ for arbitrary actions $a \in A$ and states $x \in \mathbb{X}$. Our approach combines ideas from *Q-learning* and the recently introduced *full-history recursive multilevel Picard* (MLP) approximations which have been proven to overcome the curse of dimensionality in the numerical approximation of certain semilinear partial differential equations (PDEs) (see, e.g., [3–6, 20, 21, 24, 29–33]). *Q*-learning

is based on the idea to switch the order of expectation and maximization in (2). Formally, the Q-function satisfies for all $x \in \mathbb{X}$, $a \in A$ that $Q(x, a) = g(x, a) + \delta \, \mathbb{E}[v(X^{x,a})]$. This together with (2) implies for all $x \in \mathbb{X}$ that $v(x) = \max_{a \in A} Q(x, a)$ and, hence, that for all $x \in \mathbb{X}$, $a \in A$ it holds that

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

Under appropriate conditions one can show that Q is the unique solution of this fixed-point equation and that the sequence of fixed-point iterates $Q_n \colon \mathbb{X} \times A \to \mathbb{R}$, $n \in \mathbb{N}_0$, which satisfies for all $n \in \mathbb{N}$, $x \in \mathbb{X}$, $a \in A$ that $Q_0(x, a) = g(x, a)$ and

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

converges to Q. We next employ a central idea of MLP approximations and decompose the iterates into multilevels to obtain for all $n \in \mathbb{N}_0$, $x \in \mathbb{X}$, $a \in A$ that

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

In this telescope expansion, we apply a fundamental idea of Heinrich [27, 28] and Giles [23] and approximate the expected values by Monte Carlo averages with different degrees of accuracy at different levels $l \in \{1, \ldots, n\}$. The convergence of $(Q_l)_{l \in \mathbb{N}}$ ensures that for large $l \in \{1, \ldots, n\}$ the difference between Q_l and Q_{l-1} is small and hence we use for large $l \in \{1, \ldots, n\}$ less Monte Carlo samples to approximate the expected value $\mathbb{E}\left[(\max_{b \in A} Q_l(X^{x,a}, b)) - (\max_{b \in A} Q_{l-1}(X^{x,a}, b))\right]$ than for small $l \in \{1, \ldots, n\}$. More specifically, we fix $M \in \mathbb{N}$ and use M^{n-l} independent Monte Carlo samples to approximate the expected value at level $l \in \{1, \ldots, n\}$ which leads to the full-history recursive multilevel fixed-point (MLFP) approximation scheme in (6) below.

To briefly sketch the contribution of this article within this introductory section, we now present in the following result, Theorem 1.1 below, a special case of Theorem 4.1, the main result of this article. Below Theorem 1.1 we explain in words the statement of Theorem 1.1 as well as the mathematical objects appearing in Theorem 1.1.

Theorem 1.1. *Let A be a nonempty set, for every $d \in \mathbb{N}$ let $g_d \colon \mathbb{R}^d \times A \to \mathbb{R}$ be $(\mathcal{B}(\mathbb{R}^d) \otimes 2^A)/\mathcal{B}(\mathbb{R})$-measurable, assume $\sup_{d \in \mathbb{N}} \sup_{x \in \mathbb{R}^d} \sup_{a \in A} |g_d(x, a)| < \infty$, let $(\Omega, \mathcal{F}, \mathbb{P})$ be a probability space, let $\Theta = \cup_{n \in \mathbb{N}} \mathbb{Z}^n$, for every $d \in \mathbb{N}$ let $\mathcal{F}_d^\theta \subseteq \mathcal{F}$, $\theta \in \Theta$, be independent sub-sigma-algebras of $\mathcal{F}$, for every $d \in \mathbb{N}$ let $X_d^\theta = (X_d^{\theta,x,a}(\omega))_{(x,a,\omega) \in \mathbb{R}^d \times A \times \Omega} \colon \mathbb{R}^d \times A \times \Omega \to \mathbb{R}^d$, $\theta \in \Theta$, be i.i.d. random fields which satisfy for all $d \in \mathbb{N}$, $\theta \in \Theta$ that X_d^θ is $(\mathcal{B}(\mathbb{R}^d) \otimes 2^A \otimes \mathcal{F}_d^\theta)/\mathcal{B}(\mathbb{R}^d)$-measurable, for every $d \in \mathbb{N}$ let $\delta_d \in [0, 1)$, $\mathcal{R}_d \in [0, \infty)$, let $M \in \mathbb{N} \cap [\sup_{d \in \mathbb{N}}(4|A|^2(1 - \delta_d)^{-2}), \infty)$, for every $d \in \mathbb{N}$ let $\mathcal{Q}_{d,n}^\theta \colon \mathbb{R}^d \times A \times \Omega \to \mathbb{R}$, $n \in \mathbb{N}_0$, $\theta \in \Theta$, satisfy for all $n \in \mathbb{N}_0$, $\theta \in \Theta$, $x \in \mathbb{R}^d$, $a \in A$ that*

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

and for every $d \in \mathbb{N}$ let $\mathcal{C}_{d,n} \in [0, \infty)$, $n \in \mathbb{N}_0$, satisfy for all $n \in \mathbb{N}_0$ that

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

Then

4

(i) *it holds for all $d \in \mathbb{N}$ that there exists a unique bounded $(\mathcal{B}(\mathbb{R}^d) \otimes 2^A)/\mathcal{B}(\mathbb{R})$-measurable $Q_d \colon \mathbb{R}^d \times A \to \mathbb{R}$ which satisfies for all $x \in \mathbb{R}^d$, $a \in A$ that*

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

and

(ii) *there exist $N \colon (0,1] \to \mathbb{N}$ and $c \in \mathbb{R}$ such that for all $d \in \mathbb{N}$, $\varepsilon \in (0,1]$ it holds that $\mathcal{C}_{d,N_\varepsilon} \leq c \mathcal{R}_d \varepsilon^{-c}$ and*

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

Theorem 1.1 is an immediate consequence from Corollary 4.2 in Section 4 below. Corollary 4.2, in turn, follows from Theorem 4.1, which is the main result of this article. In the following we add some comments on the mathematical objects appearing in Theorem 1.1 above.

In Theorem 1.1 we introduce in (6) a Monte Carlo-type approximation algorithm for a sequence of MDPs indexed by the dimension $d \in \mathbb{N}$ of the state space. To formulate the proposed Monte Carlo-type approximation algorithm in (6) we need, roughly speaking, sufficiently many independent random variables which are indexed over a sufficiently large index set. This sufficiently large index set is provided through the set $\Theta = \cup_{n \in \mathbb{N}} \mathbb{Z}^n$ introduced in Theorem 1.1. The triple $(\Omega, \mathcal{F}, \mathbb{P})$ in Theorem 1.1 is the probability space on which the random variables are defined. In Theorem 1.1 we consider for every $d \in \mathbb{N}$ an MDP with state space $(\mathbb{R}^d, \mathcal{B}(\mathbb{R}^d))$. We assume that all elements of the sequence of MDPs have a common control set A. For every $d \in \mathbb{N}$, $x \in \mathbb{R}^d$, $a \in A$ the one-step transition of the controlled Markov chain of the MDP with state space $(\mathbb{R}^d, \mathcal{B}(\mathbb{R}^d))$ is given by the random variable $X_d^{0,x,a} \colon \Omega \to \mathbb{R}^d$. In the language of MDPs for every $d \in \mathbb{N}$, $a \in A$ the transition kernel $\kappa_{d,a} \colon \mathbb{R}^d \times \mathcal{B}(\mathbb{R}^d) \to [0,1]$ is thus determined by the distribution of $X_d^{0,\cdot,a}$. For every $d \in \mathbb{N}$ the function $g_d \colon \mathbb{R}^d \times A \to \mathbb{R}$ introduced in the first line of Theorem 1.1 is the reward function of the MDP with state space $(\mathbb{R}^d, \mathcal{B}(\mathbb{R}^d))$. In Theorem 1.1 we assume that the functions g_d, $d \in \mathbb{N}$, are uniformly bounded in $d \in \mathbb{N}$. For every $d \in \mathbb{N}$ the real number $\delta_d \in [0,1)$ introduced in Theorem 1.1 is the discount factor of the MDP with state space $(\mathbb{R}^d, \mathcal{B}(\mathbb{R}^d))$.

Item (i) in Theorem 1.1 establishes the essentially well-known result that under the above assumptions for every $d \in \mathbb{N}$ the Bellman equation (2) for the Q-function associated to the MDP with state space $(\mathbb{R}^d, \mathcal{B}(\mathbb{R}^d))$ has a unique solution $Q_d \colon \mathbb{R}^d \times A \to \mathbb{R}$.

In (6) in Theorem 1.1 we specify the MLFP approximation scheme which we propose to approximate the solution of the Bellman equation (2). The MLFP approximations $\mathcal{Q}_{d,n}^\theta$, $d \in \mathbb{N}$, $n \in \mathbb{N}_0$, $\theta \in \Theta$, are indexed by the dimension $d \in \mathbb{N}$, by the number $n \in \mathbb{N}_0$ of fixed-point iterates and by a parameter $\theta \in \Theta$ which is different for different appearences of MLFP approximations in (6). As random input sources the MLFP approximation scheme proposed in (6) employs the random variables $X_d^{\theta,x,a} \colon \Omega \to \mathbb{R}^d$, $\theta \in \Theta \setminus \{0\}$, $d \in \mathbb{N}$, $x \in \mathbb{R}^d$, $a \in A$. Note that for every $\theta \in \Theta \setminus \{0\}$, $d \in \mathbb{N}$, $x \in \mathbb{R}^d$, $a \in A$ the random variable $X_d^{\theta,x,a} \colon \Omega \to \mathbb{R}^d$ which is used as random input source of the MLFP approximation scheme proposed in (6) and the random variable $X_d^{0,x,a} \colon \Omega \to \mathbb{R}^d$ which is used to formulate the Bellman equation (2) are identically distributed. The parameter $\theta \in \Theta$ ensures that different appearances of MLFP approximations in (6) are independent and this ensures that (6) can be implemented with recursive function calls. The natural number $M \in \mathbb{N} \cap [\sup_{d \in \mathbb{N}}(4|A|^2(1-\delta_d)^{-2}), \infty]$ in Theorem 1.1 determines the number of Monte Carlo samples used in the definition of the MLFP approximation $\mathcal{Q}_{d,n}^\theta$, $d \in \mathbb{N}$, $n \in \mathbb{N}_0$, $\theta \in \Theta$ in (6). The assumption that the natural number M is an element of the set

$\mathbb{N} \cap [\sup_{d \in \mathbb{N}}(4|A|^2(1 - \delta_d)^{-2}), \infty]$ ensures that $\sup_{d \in \mathbb{N}}(4|A|^2(1 - \delta_d)^{-2}) < \infty$. This implies that the control set A is finite and that the sequence $(\delta_d)_{d \in \mathbb{N}} \subseteq [0, 1)$ is bounded away from 1 in the sense that $\sup_{d \in \mathbb{N}} \delta_d < 1$. Note that by a suitable identification this framework also covers the case where for each $d \in \mathbb{N}$ there is an individual control set A_d such that $\sup_{d \in \mathbb{N}} |A_d| < \infty$.

For every $d \in \mathbb{N}$ the nonnegative real number $\mathcal{R}_d \in [0, \infty)$ in Theorem 1.1 is understood as an upper bound of the computational cost to compute one realization of any of the random variables $X_d^{\theta,x} \colon \Omega \to (\mathbb{R}^d)^A$, $\theta \in \Theta$, $x \in \mathbb{R}^d$. The real numbers $\mathcal{C}_{d,n} \in [0, \infty)$, $d \in \mathbb{N}$, $n \in \mathbb{N}_0$, in (7) in Theorem 1.1 model the computational costs of the MLFP approximation scheme in (6). More specifically, for every $d \in \mathbb{N}$, $n \in \mathbb{N}_0$ the real number $\mathcal{C}_{d,n} \in [0, \infty)$ represents an upper bound of the computational costs to compute the realizations of all random variables $X_d^{\theta,x} \colon \Omega \to (\mathbb{R}^d)^A$, $\theta \in \Theta$, $x \in \mathbb{R}^d$, required to compute one realization of $\mathcal{Q}_{d,n}^0(0, \cdot) \colon \Omega \to \mathbb{R}^A$.

Item (ii) in Theorem 1.1 proves that the solutions of the Bellman equations in (2) can be approximated by means of the MLFP approximation scheme in (6) with a computational cost which grows at most polynomially in the reciprocal $1/\epsilon$ of the prescribed approximation accuracy $\epsilon \in (0, 1]$ and linearly in the computational cost $\mathcal{R}_d$ to compute one realization of any of the random variables $X_d^{\theta,x} \colon \Omega \to (\mathbb{R}^d)^A$, $\theta \in \Theta$, $x \in \mathbb{R}^d$, where $d \in \mathbb{N}$ is the dimension of the state space of the associated MDP. In particular, if the computational cost to compute one realization of any of the random variables $X_d^{\theta,x} \colon \Omega \to (\mathbb{R}^d)^A$, $\theta \in \Theta$, $x \in \mathbb{R}^d$, grows at most polynomially in the dimension $d \in \mathbb{N}$ of the state space of the associated MDP (as it is often the case in practical applications), then the MLFP approximation scheme in (6) overcomes the curse of dimensionality for the approximation of the solutions of the Bellman equations in (2). However, we would like to point out that the constant $c \in \mathbb{R}$ appearing in item (ii) in Theorem 1.1 may become arbitrary large if $\sup_{d \in \mathbb{N}} \delta_d$ is close to 1. Thus the computational cost of the nonlinear Monte Carlo methods in (6) may become impractical even so the methods in (6) provably overcome the curse of dimensionality.

In the following we also add some comments on generalizations and variants of Theorem 1.1 presented in this article. While Theorem 1.1 considers a sequence of MDPs indexed by the dimension $d \in \mathbb{N}$ of the Euclidean state spaces $(\mathbb{R}^d, \mathcal{B}(\mathbb{R}^d))$, Corollary 4.2 considers a family of MDPs with general index set $\mathfrak{D}$ and general measurable state spaces $(\mathbb{X}_d, \mathcal{X}_d)$, $d \in \mathfrak{D}$. While Theorem 1.1 considers MDPs with bounded reward functions, Corollary 4.2 allows for unbounded reward functions. While Theorem 1.1 and Corollary 4.2 consider Bellman equations of MDPs, Theorem 4.1 considers more general functional fixed-point equations (we refer to (85) in Theorem 4.1 for details). Corollary 4.3 proves that a variant of the MLFP approximation scheme (see (108) for details) overcomes the curse of dimensionality for the approximation of the solutions of Bellman equations for optimal stopping problems (see (110) for details).

The remainder of this article is organized as follows. In Section 2 below we establish existence, uniqueness, and integrability properties for solutions of functional fixed-point equations. In Section 3 below we introduce MLFP approximations for solutions of functional fixed-point equations, we study measurability, distributional, and integrability properties for the introduced MLFP approximations and we establish recursive and subsequently non-recursive upper bounds for the L^2-distances between the exact solutions of the considered functional fixed-point equations and the proposed MLFP approximations. In Section 4 we combine the existence, uniqueness, and regularity properties for solutions of functional fixed-point equations, which we have established in Section 2, with the error analysis for MLFP approximations for functional fixed-point equations, which we have established in Section 3, to obtain a computational complexity analysis for MLFP approximations for functional fixed-point equations and for Bellman equations of MDPs and optimal stopping problems.

2 Existence and uniqueness of solutions of functional fixed-point equations

Definition 2.1. Let $(\mathbb{X}, \mathcal{X})$, $(\mathbb{Y}, \mathcal{Y})$ be a nonempty measurable spaces, let A be a nonempty set, and let $\kappa_a \colon \mathbb{X} \times \mathcal{Y} \to [0,1]$, $a \in A$, satisfy for all $a \in A$, $M \in \mathcal{Y}$ that $\mathbb{X} \ni x \mapsto \kappa_a(x,M) \in [0,1]$ is $\mathcal{X}/\mathcal{B}([0,1])$-measurable and for all $x \in \mathbb{X}$, $a \in A$ that $\mathcal{Y} \ni M \mapsto \kappa_a(x,M) \in [0,1]$ is a probability measure on $(\mathbb{Y}, \mathcal{Y})$. Then we say that $(\kappa_a)_{a \in A}$ is a family of stochastic kernels from $(\mathbb{X}, \mathcal{X})$ to $(\mathbb{Y}, \mathcal{Y})$.

Lemma 2.2. *Let* $c, L \in [0, \infty)$ *with* $cL < 1$, *let* $(\mathbb{X}, \mathcal{X})$ *be a nonempty measurable space, let* A *be a nonempty countable set, let* $\kappa = (\kappa_a)_{a \in A}$ *be a family of stochastic kernels from* $(\mathbb{X}, \mathcal{X})$ *to* $(\mathbb{X}, \mathcal{X})$, *let* $\mathbb{R}^A = \{r \colon A \to \mathbb{R}\}$, *let* $\widetilde{\mathcal{A}} = \bigotimes_{a \in A} \mathcal{B}(\mathbb{R})$, *let* $f \colon \mathbb{X} \times \mathbb{R}^A \to \mathbb{R}$ *be* $(\mathcal{X} \otimes \widetilde{\mathcal{A}})/\mathcal{B}(\mathbb{R})$-*measurable, let* $\mathfrak{w} \colon \mathbb{X} \to (0, \infty)^A$ *be* $\mathcal{X}/\bigotimes_{a \in A} \mathcal{B}((0, \infty))$-*measurable, let*

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

assume for all $x \in \mathbb{X}$, $a \in A$, $r, s \in \mathbb{R}^A$ *that* $|f(x,r) - f(x,s)| \leq L \sup_{b \in A} |r(b) - s(b)|$, $\int_{\mathbb{X}} \sup_{b \in A} (\mathfrak{w}(y))(b) \kappa_a(x, dy) \leq c(\mathfrak{w}(x))(a)$, *and* $\sup_{(t,b) \in \mathbb{X} \times A} [|(\mathfrak{w}(t))(b)|^{-1} \int_{\mathbb{X}} |f(y,0)| \kappa_b(t, dy)] < \infty$. *Then there exists a unique* $v \in \mathfrak{W}$ *which satisfies for all* $x \in \mathbb{X}$, $a \in A$ *that*

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

Proof of Lemma 2.2. Let $\|\cdot\| \colon \mathfrak{W} \to [0, \infty)$ satisfy for all $u \in \mathfrak{W}$ that $\|u\| = \sup_{(x,a) \in \mathbb{X} \times A} \frac{|(u(x))(a)|}{|(\mathfrak{w}(x))(a)|}$. Note that $\|\cdot\|$ is a norm on $\mathfrak{W}$ and $(\mathfrak{W}, \|\cdot\|)$ is a Banach space. The assumption that for all $x \in X$, $r, s \in \mathbb{R}^A$ it holds that $|f(x,r) - f(x,s)| \leq L \sup_{a \in A} |r(a) - s(a)|$ and the assumption that for all $x \in \mathbb{X}$, $a \in A$ it holds that $\int_{\mathbb{X}} \sup_{b \in A} (\mathfrak{w}(y))(b) \kappa_a(x, dy) \leq c(\mathfrak{w}(x))(a)$, yield that for all $u \in \mathfrak{W}$, $x \in \mathbb{X}$, $a \in A$ it holds that

$$\frac{1}{(\mathfrak{w}(x))(a)} \left| \int_{\mathbb{X}} f(y, u(y)) \kappa_a(x, dy) \right|$$

$$\leq \frac{1}{(\mathfrak{w}(x))(a)} \int_{\mathbb{X}} |f(y, u(y)) - f(y, 0)| \kappa_a(x, dy) + \frac{1}{(\mathfrak{w}(x))(a)} \int_{\mathbb{X}} |f(y, 0)| \kappa_a(x, dy)$$

$$\leq \frac{L}{(\mathfrak{w}(x))(a)} \int_{\mathbb{X}} \sup_{b \in A} |(u(y))(b)| \kappa_a(x, dy) + \frac{1}{(\mathfrak{w}(x))(a)} \int_{\mathbb{X}} |f(y, 0)| \kappa_a(x, dy)$$

$$= \frac{L}{(\mathfrak{w}(x))(a)} \int_{\mathbb{X}} \sup_{b \in A} \left\{ \frac{|(u(y))(b)|}{(\mathfrak{w}(y))(b)} (\mathfrak{w}(y))(b) \right\} \kappa_a(x, dy) + \frac{1}{(\mathfrak{w}(x))(a)} \int_{\mathbb{X}} |f(y, 0)| \kappa_a(x, dy)$$

$$\leq \frac{L \|u\|}{(\mathfrak{w}(x))(a)} \int_{\mathbb{X}} \sup_{b \in A} (\mathfrak{w}(y))(b) \kappa_a(x, dy) + \frac{1}{(\mathfrak{w}(x))(a)} \int_{\mathbb{X}} |f(y, 0)| \kappa_a(x, dy)$$

$$\leq cL \|u\| + \frac{1}{(\mathfrak{w}(x))(a)} \int_{\mathbb{X}} |f(y, 0)| \kappa_a(x, dy). \quad (12)$$

The assumption that $\sup_{(x,a) \in \mathbb{X} \times A} \left[|(\mathfrak{w}(x))(a)|^{-1} \int_{\mathbb{X}} |f(x,0)| \kappa_a(x, dy) \right] < \infty$ demonstrates that for all $u \in \mathfrak{W}$ it holds that

$$\sup_{(x,a) \in \mathbb{X} \times A} \frac{1}{(\mathfrak{w}(x))(a)} \left| \int_{\mathbb{X}} f(y, u(y)) \kappa_a(x, dy) \right|$$

$$\leq cL \|u\| + \sup_{(x,a) \in \mathbb{X} \times A} \frac{1}{(\mathfrak{w}(x))(a)} \int_{\mathbb{X}} |f(y, 0)| \kappa_a(x, dy) < \infty. \quad (13)$$

7

[35, Lemma 14.20] ensures that for all $u \in \mathfrak{W}$ it holds that the map $\mathbb{X} \ni x \mapsto \big[A \ni a \mapsto \int_{\mathbb{X}} f(y, u(y)) \kappa_a(x, dy) \in \mathbb{R} \big] \in \mathbb{R}^A$ is $\mathcal{X}/\widetilde{\mathcal{A}}$-measurable. Let $\Phi \colon \mathfrak{W} \to \mathfrak{W}$ be the function which satisfies for all $u \in \mathfrak{W}$, $(x, a) \in \mathbb{X} \times A$ that

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

The assumption that for all $x \in \mathbb{X}$, $r, s \in \mathbb{R}^A$ it holds that $|f(x, r) - f(x, s)| \leq L \sup_{a \in A} |r(a) - s(a)|$, the assumption that $\sup_{(x,a) \in \mathbb{X} \times A} \big[|(\mathfrak{w}(x))(a)|^{-1} \int_{\mathbb{X}} |f(x, 0)| \kappa_a(x, dy) \big] < \infty$, and the assumption that for all $(x, a) \in \mathbb{X} \times A$ it holds that $\int_{\mathbb{X}} \sup_{b \in A} (\mathfrak{w}(y))(b) \kappa_a(x, dy) \leq c(\mathfrak{w}(x))(a)$, ensure that for all $u, v \in \mathfrak{W}, (x, a) \in \mathbb{X} \times A$ it holds that

$$\frac{1}{(\mathfrak{w}(x))(a)} \big| [\Phi(u)](x)(a) - [\Phi(v)](x)(a) \big|$$

$$\leq \frac{1}{(\mathfrak{w}(x))(a)} \int_{\mathbb{X}} \big| f(y, u(y)) - f(y, v(y)) \big| \kappa_a(x, dy)$$

$$\leq \frac{L}{(\mathfrak{w}(x))(a)} \int_{\mathbb{X}} \sup_{b \in A} |(u(y))(b) - (v(y))(b)| \kappa_a(x, dy)$$

$$= \frac{L}{(\mathfrak{w}(x))(a)} \int_{\mathbb{X}} \sup_{b \in A} \left\{ \frac{|(u(y))(b) - (v(y))(b)|}{(\mathfrak{w}(y))(b)} (\mathfrak{w}(y))(b) \right\} \kappa_a(x, dy)$$

$$\leq \frac{L \|u - v\|}{(\mathfrak{w}(x))(a)} \int_{\mathbb{X}} \sup_{b \in A} (\mathfrak{w}(y))(b) \kappa_a(x, dy)$$

$$\leq cL \|u - v\|. \tag{15}$$

This in turn proves that for all $u, v \in \mathfrak{W}$ it holds that

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

The assumption that $cL < 1$ shows that Φ is a contraction. Hence Banach's fixed-point theorem proves that there exists a unique function $v \in \mathfrak{W}$ such that $v = \Phi(v)$. The fact that $v \in \mathfrak{W}$ ensures that for all $(x, a) \in \mathbb{X} \times A$ it holds that $\int_{\mathbb{X}} |f(y, v(y))| \kappa_a(x, dy) < \infty$. The proof of Lemma 2.2 is thus completed. $\qquad\square$

Corollary 2.3. *Let $c, L \in [0, \infty)$ with $cL < 1$, let $(\mathbb{X}, \mathcal{X})$ be a nonempty measurable space, let $(\Omega, \mathcal{F}, \mathbb{P})$ be a probability space, let A be a nonempty countable set, let $\mathbb{R}^A = \{r \colon A \to \mathbb{R}\}$, let $\widetilde{\mathcal{A}} = \bigotimes_{a \in A} \mathcal{B}(\mathbb{R})$, let $\mathbb{X}^A = \{q \colon A \to \mathbb{X}\}$, let $\widetilde{\mathcal{X}} = \bigotimes_{a \in A} \mathcal{X}$, let $X = \big(X^{x,a}(\omega) \big)_{x \in \mathbb{X}, \, a \in A, \, \omega \in \Omega} \colon \mathbb{X} \times \Omega \to \mathbb{X}^A$ be $(\mathcal{X} \otimes \mathcal{F})/\widetilde{\mathcal{X}}$-measurable, let $f \colon \mathbb{X} \times \mathbb{R}^A \to \mathbb{R}$ be $(\mathcal{X} \otimes \widetilde{\mathcal{A}})/\mathcal{B}(\mathbb{R})$-measurable, let $\mathfrak{w} \colon \mathbb{X} \to (0, \infty)^A$ be $\mathcal{X}/\bigotimes_{a \in A} \mathcal{B}((0, \infty))$-measurable, let*

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

assume that for all $(x, a) \in \mathbb{X} \times A$, $r, s \in \mathbb{R}^A$ it holds that $|f(x, r) - f(x, s)| \leq L \sup_{b \in A} |r(b) - s(b)|$, $\mathbb{E}[\sup_{b \in A} (\mathfrak{w}(X^{x,a}))(b)] \leq c(\mathfrak{w}(x))(a)$, and $\sup_{(y,b) \in \mathbb{X} \times A} \big[|(\mathfrak{w}(y))(b)|^{-1} \mathbb{E}[|f(X^{y,b}, 0)|] \big] < \infty$. Then there exists a unique function $v \in \mathfrak{W}$ which satisfies for all $(x, a) \in \mathbb{X} \times A$ that

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

Proof of Corollary 2.3. Let $\kappa_a\colon \mathbb{X} \times \mathcal{X} \to [0,1]$, $a \in A$, satisfy for all $(x,a) \in \mathbb{X} \times A$, $M \in \mathcal{X}$, that $\kappa_a(x,M) = \mathbb{P}[X^{x,a} \in M]$. The fact that X is $(\mathcal{X} \otimes \mathcal{F})/\widetilde{\mathcal{X}}$-measurable ensures that for all $(x,a) \in \mathbb{X} \times A$ it holds that $X^{x,a}\colon \Omega \to \mathbb{X}$ is $\mathcal{F}/\mathcal{X}$-measurable. This implies that for all $(x,a) \in \mathbb{X} \times A$ it holds that $\mathcal{X} \ni M \mapsto \kappa_a(x,M) = \mathbb{P}[X^{x,a} \in M] \in [0,1]$ is a probability measure on $(\mathbb{X}, \mathcal{X})$. The fact that X is $(\mathcal{X} \otimes \mathcal{F})/\widetilde{\mathcal{X}}$-measurable and [35, Theorem 14.16] imply that for all $a \in A$, $M \in \mathcal{X}$ it holds that $\mathbb{X} \ni x \mapsto \kappa_a(x,M) = \mathbb{P}[X^{x,a} \in M]$ is $\mathcal{X}/\mathcal{B}([0,1])$-measurable. Hence it holds that $(\kappa_a)_{a \in A}$ is a familiy of stochastic kernels form $(\mathbb{X}, \mathcal{X})$ to $(\mathbb{X}, \mathcal{X})$. The assumptions that for all $(x,a) \in \mathbb{X} \times A$ it holds that $\mathbb{E}\big[\sup_{b \in A}\big(\mathfrak{w}(X^{x,a})\big)(b)\big] \le c\big(\mathfrak{w}(x)\big)(a)$, and $\sup_{(y,b) \in \mathbb{X} \times A}\big[|\big(\mathfrak{w}(y)\big)(b)|^{-1}\mathbb{E}[|f(X^{y,b},0)|]\big] < \infty$, prove that for all $(x,a) \in \mathbb{X} \times A$ it holds that

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

and

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

Lemma 2.2 implies that there exists a unique function $v \in \mathfrak{W}$ such that for all $(x,a) \in \mathbb{X} \times A$ it holds that

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

and

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

The proof of Corollary 2.3 is thus completed. $\qquad\square$

Lemma 2.4. *Let $c_f, c_\mathfrak{w}, L \in [0,\infty)$ with $c_\mathfrak{w}L < 1$, let $(\mathbb{X}, \mathcal{X})$ be a nonempty measurable space, let $(\Omega, \mathcal{F}, \mathbb{P})$ be a probability space, let A be a nonempty countable set, let $\mathbb{R}^A = \{r\colon A \to \mathbb{R}\}$, let $\widetilde{\mathcal{A}} = \bigotimes_{a \in A}\mathcal{B}(\mathbb{R})$, let $\mathbb{X}^A = \{q\colon A \to \mathbb{X}\}$, let $\widetilde{\mathcal{X}} = \bigotimes_{a \in A}\mathcal{X}$, let $X = \big(X^{x,a}(\omega)\big)_{x \in \mathbb{X},\, a \in A,\, \omega \in \Omega}\colon \mathbb{X} \times \Omega \to \mathbb{X}^A$ be $(\mathcal{X} \otimes \mathcal{F})/\widetilde{\mathcal{X}}$-measurable, let $f\colon \mathbb{X} \times \mathbb{R}^A \to \mathbb{R}$ be $(\mathcal{X} \otimes \widetilde{\mathcal{A}})/\mathcal{B}(\mathbb{R})$-measurable, let $\mathfrak{w}\colon \mathbb{X} \to (0,\infty)^A$ be $\mathcal{X}/\bigotimes_{a \in A}\mathcal{B}((0,\infty))$-measurable, let $v\colon \mathbb{X} \to \mathbb{R}^A$ be $\mathcal{X}/\widetilde{\mathcal{A}}$-measurable, assume that for all $(x,a) \in \mathbb{X} \times A$, $r,s \in \mathbb{R}^A$ it holds that $|f(x,r) - f(x,s)| \le L\sup_{b \in A}|r(b) - s(b)|$, $\sup_{(y,b) \in \mathbb{X} \times A}\big[|\big(\mathfrak{w}(y)\big)(b)|^{-1}|\big(v(y)\big)(b)|\big] < \infty$, $\big(\mathbb{E}\big[\sup_{b \in A}|\big(\mathfrak{w}(X^{x,a})\big)(b)|^2\big]\big)^{\frac{1}{2}} \le c_\mathfrak{w}\big(\mathfrak{w}(x)\big)(a)$, $\big(\mathbb{E}\big[|f(X^{x,a},0)|^2\big]\big)^{\frac{1}{2}} \le c_f\big(\mathfrak{w}(x)\big)(a)$, and $\big(v(x)\big)(a) = \mathbb{E}[f(X^{x,a},v(X^{x,a}))]$. Then it holds that*

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

Proof of Lemma 2.4. Jensen's inequality, the triangle inequality, and the assumption that for all $x \in \mathbb{X}$, $r,s \in \mathbb{R}^A$ it holds that $|f(x,r) - f(x,s)| \le L\sup_{a \in A}|r(a) - s(a)|$ prove that for all

9

$(x, a) \in \mathbb{X} \times A$ it holds that

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

Combining this, the assumption that $\sup_{(x,a) \in \mathbb{X} \times A}[|(\mathfrak{w}(x))(a)|^{-1}|(v(x))(a)|] < \infty$, and the assumption that $c_{\mathfrak{w}}L < 1$ shows that

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

This and the assumption that for all $(x, a) \in \mathbb{X} \times A$ it holds that $\mathbb{E}\left[|\sup_{b \in A}(\mathfrak{w}(X^{x,a}))(b)|^2\right]^{\frac{1}{2}} \leq c_{\mathfrak{w}}(\mathfrak{w}(x))(a)$ imply that for all $(x, a) \in \mathbb{X} \times A$ it holds that

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

Taking the supremum over $\mathbb{X} \times A$ yields

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

This completes the proof of Lemma 2.4. $\qquad\square$

3 Full-history recursive multilevel fixed-point (MLFP) approximations

3.1 Mathematical description of MLFP approximations

Setting 3.1. *Let $M \in \mathbb{N}$, let $\Theta = \bigcup_{n \in \mathbb{N}} \mathbb{Z}^n$, let $(\mathbb{X}, \mathcal{X})$ be a nonempty measurable space, let $(\Omega, \mathcal{F}, \mathbb{P})$ be a probability space, let A be a nonempty set, let $\mathbb{R}^A = \{r : A \to \mathbb{R}\}$, let $\widetilde{\mathcal{A}} = \bigotimes_{a \in A} \mathcal{B}(\mathbb{R})$, let $\|\cdot\|_\infty : \mathbb{R}^A \to [0, \infty]$ satisfy for all $r \in \mathbb{R}^A$ that $\|r\|_\infty = \sup_{a \in A}|r(a)|$, let $\mathbb{X}^A = \{q : A \to \mathbb{X}\}$, let $\widetilde{\mathcal{X}} = \bigotimes_{a \in A} \mathcal{X}$, let $f : \mathbb{X} \times \mathbb{R}^A \to \mathbb{R}$ be $(\mathcal{X} \otimes \widetilde{\mathcal{A}})/\mathcal{B}(\mathbb{R})$-measurable, let $(\mathcal{F}^\theta)_{\theta \in \Theta}$, with $\mathcal{F}^\theta \subseteq \mathcal{F}$, $\theta \in \Theta$, be independent σ-algebras, let $X^\theta = \left(X^{\theta,x,a}(\omega)\right)_{x \in \mathbb{X}, a \in A, \omega \in \Omega} : \mathbb{X} \times \Omega \to \mathbb{X}^A$,*

10

$\theta \in \Theta$, be i.i.d. random fields, such that for all $\theta \in \Theta$ it holds that X^θ is $(\mathcal{X} \otimes \mathcal{F}^\theta)/\widetilde{\mathcal{X}}$-measurable, let $V_n^\theta \colon \mathbb{X} \times \Omega \to \mathbb{R}^A$, $n \in \mathbb{N}_0$, $\theta \in \Theta$, satisfy for all $n \in \mathbb{N}_0$, $\theta \in \Theta$, $x \in \mathbb{X}$, $a \in A$ that

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

3.2 Measurability and distributional properties for MLFP approximations

Lemma 3.2. *Assume Setting 3.1. It holds for all $n \in \mathbb{N}_0$, $\theta \in \Theta$ that V_n^θ is $(\mathcal{X} \otimes \sigma(\bigcup_{\eta \in \Theta} \mathcal{F}^{(\theta,\eta)}))/\widetilde{\mathcal{A}}$-measurable and it holds for all $n \in \mathbb{N}_0$, $\theta, \vartheta \in \Theta$ that*

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

is $(\mathcal{X} \otimes \sigma(\bigcup_{\eta \in \Theta} \mathcal{F}^{(\vartheta,\eta)} \cup \mathcal{F}^\theta))/\widetilde{\mathcal{A}}$-measurable.

Proof of Lemma 3.2. Note that (27) implies that for all $\theta \in \Theta$, $x \in \mathbb{X}$, $a \in A$ it holds that $\big(V_0^\theta(x)\big)(a) = 0$. Hence for all $\theta \in \Theta$ it holds that V_0^θ is $(\mathcal{X} \otimes \sigma(\bigcup_{\eta \in \Theta} \mathcal{F}^{(\theta,\eta)}))/\widetilde{\mathcal{A}}$-measurable. Fix $n \in \mathbb{N}$ and assume for all $l \in \{0, 1, \ldots, n-1\}$, $\theta \in \Theta$ that V_l^θ is $(\mathcal{X} \otimes \sigma(\bigcup_{\eta \in \Theta} \mathcal{F}^{(\theta,\eta)}))/\widetilde{\mathcal{A}}$-measurable. Observe that for all $l \in \{0, 1, \ldots, n-1\}$, $\theta \in \Theta$ it holds that

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

is $(\mathcal{X} \otimes \sigma(\bigcup_{\eta \in \Theta} \mathcal{F}^{(\theta,\eta)}))/(\mathcal{X} \otimes \widetilde{\mathcal{A}})$-measurable. Since for all $\theta \in \Theta$ it holds that X^θ is $(\mathcal{X} \otimes \mathcal{F}^\theta)/\widetilde{\mathcal{X}}$-measurable it follows that for all $\theta \in \Theta$, $a \in A$ it holds that $\mathbb{X} \times \Omega \ni (x, \omega) \mapsto X^{\theta,x,a}(\omega) \in \mathbb{X}$ is $(\mathcal{X} \otimes \mathcal{F}^\theta)/\mathcal{X}$-measurable. This implies that for all $\theta, \vartheta \in \Theta$, $a \in A$ it holds that $\mathbb{X} \times \Omega \ni (x, \omega) \mapsto (X^{\theta,x,a}(\omega), \omega) \in \mathbb{X} \times \Omega$ is $(\mathcal{X} \otimes \sigma(\bigcup_{\eta \in \Theta} \mathcal{F}^{(\vartheta,\eta)} \cup \mathcal{F}^\theta))/(\mathcal{X} \otimes \sigma(\bigcup_{\eta \in \Theta} \mathcal{F}^{(\vartheta,\eta)}))$-measurable. This and (29) ensure that for all $l \in \{0, 1, \ldots, n-1\}$, $\theta, \vartheta \in \Theta$, $a \in A$ it holds that $\mathbb{X} \times \Omega \ni (x, \omega) \mapsto \big(X^{\theta,x,a}(\omega), V_l^\vartheta(X^{\theta,x,a}(\omega), \omega)\big) \in \mathbb{X} \times \mathbb{R}^A$ is $(\mathcal{X} \otimes \sigma(\bigcup_{\eta \in \Theta} \mathcal{F}^{(\vartheta,\eta)} \cup \mathcal{F}^\theta))/(\mathcal{X} \otimes \widetilde{\mathcal{A}})$-measurable. Therefore for all $l \in \{0, 1, \ldots, n-1\}$, $\theta, \vartheta \in \Theta$, $a \in A$ it holds that $\mathbb{X} \times \Omega \ni (x, \omega) \mapsto f\big(X^{\theta,x,a}(\omega), V_l^\vartheta(X^{\theta,x,a}(\omega), \omega)\big) \in \mathbb{R}$ is $(\mathcal{X} \otimes \sigma(\bigcup_{\eta \in \Theta} \mathcal{F}^{(\vartheta,\eta)} \cup \mathcal{F}^\theta))/\mathcal{B}(\mathbb{R})$-measurable. This implies that for all $l \in \{0, 1, \ldots, n-1\}$, $\theta, \vartheta \in \Theta$ it holds that

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

is $(\mathcal{X} \otimes \sigma(\bigcup_{\eta \in \Theta} \mathcal{F}^{(\vartheta,\eta)} \cup \mathcal{F}^\theta))/\widetilde{\mathcal{A}}$-measurable. Note that for all $l \in \{0, 1, \ldots, n-1\}$, $i \in \mathbb{N}$, $\theta \in \Theta$ it holds that $\sigma(\bigcup_{\eta \in \Theta} \mathcal{F}^{(\theta,l,i,\eta)} \cup \mathcal{F}^{(\theta,l,i)}) \subseteq \sigma(\bigcup_{\eta \in \Theta} \mathcal{F}^{(\theta,\eta)})$ and $\sigma(\bigcup_{\eta \in \Theta} \mathcal{F}^{(\theta,-l,i,\eta)} \cup \mathcal{F}^{(\theta,-l,i)}) \subseteq \sigma(\bigcup_{\eta \in \Theta} \mathcal{F}^{(\theta,\eta)})$. This and (30) ensure that for all $\theta \in \Theta$ it holds that $V_n^\theta \colon \mathbb{X} \times \Omega \to \mathbb{R}^A$ is $(\mathcal{X} \otimes \sigma(\bigcup_{\eta \in \Theta} \mathcal{F}^{(\theta,\eta)}))/\widetilde{\mathcal{A}}$-measurable. Induction hence proves that for all $n \in \mathbb{N}_0$, $\theta \in \Theta$ it holds that V_n^θ is $(\mathcal{X} \otimes \sigma(\bigcup_{\eta \in \Theta} \mathcal{F}^{(\theta,\eta)}))/\widetilde{\mathcal{A}}$-measurable. This and (30) imply that for all $n \in \mathbb{N}_0$, $\theta, \vartheta \in \Theta$ it holds that $\mathbb{X} \times \Omega \ni (x, \omega) \mapsto \big[A \ni a \mapsto f\big(X^{\theta,x,a}(\omega), V_n^\vartheta(X^{\theta,x,a}(\omega), \omega)\big) \in \mathbb{R} \big] \in \mathbb{R}^A$ is $(\mathcal{X} \otimes \sigma(\bigcup_{\eta \in \Theta} \mathcal{F}^{(\vartheta,\eta)} \cup \mathcal{F}^\theta))/\widetilde{\mathcal{A}}$-measurable. The proof of Lemma 3.2 is thus completed. $\square$

Lemma 3.3. *Assume Setting 3.1. Then for all $n \in \mathbb{N}_0$ it holds that V_n^θ, $\theta \in \Theta$, are identically distributed random fields.*

Proof of Lemma 3.3. For all $\theta \in \Theta$, $(x,a) \in \mathbb{X} \times A$ it holds that $\big(V_0^\theta(x)\big)(a) = 0$. Therefore $V_0^\theta, \theta \in \Theta$, are identically distributed random fields. Fix $n \in \mathbb{N}$ and assume that for all $l \in \{0, 1, \ldots, n-1\}$ it holds that V_l^θ, $\theta \in \Theta$, are identically distributed random fields. Lemma 3.2 and [5, Lemma 2.6] ensure for all $l \in \{0, 1, \ldots, n-1\}$, $i \in \mathbb{N}$, $\theta \in \Theta$ that $\mathbb{X} \times \Omega \ni (x,\omega) \mapsto (V_l^{(\theta,l,i)}(x,\omega), V_{\max\{l-1,0\}}^{(\theta,-l,i)}(x,\omega)) \in \mathbb{R}^A \times \mathbb{R}^A$ and $\mathbb{X} \times \Omega \ni (x,\omega) \mapsto (V_l^{(0,l,i)}(x,\omega), V_{\max\{l-1,0\}}^{(0,-l,i)}(x,\omega)) \in \mathbb{R}^A \times \mathbb{R}^A$ are identically distributed random fields. This ensures that for all $l \in \{0, 1, \ldots, n-1\}$, $i \in \mathbb{N}$, $\theta \in \Theta$ it holds that $\mathbb{X} \times \Omega \ni (x,\omega) \mapsto f(x, V_l^{(\theta,l,i)}(x,\omega)) - \mathbb{1}_{\mathbb{N}}(l) f(x, V_{\max\{l-1,0\}}^{(\theta,-l,i)}(x,\omega)) \in \mathbb{R}$ and $\mathbb{X} \times \Omega \ni (x,\omega) \mapsto f(x, V_l^{(0,l,i)}(x,\omega)) - \mathbb{1}_{\mathbb{N}}(l) f(x, V_{\max\{l-1,0\}}^{(0,-l,i)}(x,\omega)) \in \mathbb{R}$ are identically distributed random fields. Combining this, the assumption that for all $\theta \in \Theta$ it holds that X^θ is $(\mathcal{X} \otimes \mathcal{F}^\theta)/\widetilde{\mathcal{X}}$-measurable, and [5, Lemma 2.5] establishes that for all $l \in \{0, 1, \ldots, n-1\}$, $i \in \mathbb{N}$, $\theta \in \Theta$ it holds that

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

and

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

are identically distributed random fields. This implies that for all $l \in \{0, 1, \ldots, n-1\}$, $i \in \mathbb{N}$, $\theta \in \Theta$ it holds that

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

and

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

are identically distributed random fields. Let $g_k \colon (\mathbb{R}^A)^k \to \mathbb{R}^A$, $k \in \mathbb{N}$, satisfy for all $k \in \mathbb{N}$, $r_1, \ldots, r_k \in \mathbb{R}^A$ that $g_k(r_1, \ldots, r_k) = \sum_{j=1}^{k} r_j$. Note that for all $k \in \mathbb{N}$ it holds that g_k is $\widetilde{\mathcal{A}}^{\otimes k}/\widetilde{\mathcal{A}}$-measurable. Let $U_{l,i}^\theta \colon \mathbb{X} \times \Omega \to \mathbb{R}^A$, $l \in \{0, 1, \ldots, n-1\}$, $i \in \mathbb{N}$, $\theta \in \Theta$, satisfy for all $l \in \{0, 1, \ldots, n-1\}$, $i \in \mathbb{N}$, $\theta \in \Theta$, $(x,\omega) \in \mathbb{X} \times \Omega$ that

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

Lemma 3.2 demonstrates that for all $l \in \{0, 1, \ldots, n-1\}$, $i \in \mathbb{N}$, $\theta \in \Theta$ it holds that $U_{l,i}^\theta$ is $(\mathcal{X} \otimes \sigma(\bigcup_{\eta \in \Theta} \mathcal{F}^{(\theta,l,i,\eta)} \cup \mathcal{F}^{(\theta,-l,i,\eta)} \cup \mathcal{F}^{(\theta,l,i)}))/\widetilde{\mathcal{A}}$-measurable. Note that (27) ensures for all $(x,\omega) \in \mathbb{X} \times \Omega$, $\theta \in \Theta$ that $V_n^\theta(x,\omega) = \sum_{l=0}^{n-1} \frac{1}{M^{n-l}} \sum_{i=1}^{M^{n-l}} U_{l,i}^\theta(x,\omega)$. Let $\widetilde{M} = \sum_{j=1}^{n} M^j \in \mathbb{N}$, let $Y_n^\theta \colon \mathbb{X} \times \Omega \to (\mathbb{R}^A)^{\widetilde{M}}$, $\theta \in \Theta$, satisfy for all $\theta \in \Theta$ that

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

Observe that for all $(x, \omega) \in \mathbb{X} \times \Omega$, $\theta \in \Theta$ it holds that

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

The fact that for all $l \in \{0, 1, \ldots, n-1\}$, $i \in \mathbb{N}$, $\theta \in \Theta$ it holds that $U_{l,i}^\theta$ is $(\mathcal{X} \otimes \sigma(\bigcup_{\eta \in \Theta} \mathcal{F}^{(\theta,l,i,\eta)} \cup \mathcal{F}^{(\theta,-l,i,\eta)} \cup \mathcal{F}^{(\theta,l,i)}))/\widetilde{\mathcal{A}}$-measurable, the assumption that $(\mathcal{F}^\theta)_{\theta \in \Theta}$ are independent σ-algebras, and [5, Lemma 2.6] prove for all $\theta \in \Theta$ that Y_n^θ and Y_n^0 are identically distributed random fields. This implies for all $\theta \in \Theta$ that V_n^θ and V_n^0 are identically distributed random fields. Induction and the fact that V_0^θ, $\theta \in \Theta$, are identically distributed random fields prove that for all $n \in \mathbb{N}_0$ it holds that V_n^θ, $\theta \in \Theta$, are identically distributed random fields. This completes the proof of Lemma 3.3. $\qquad\square$

3.3 Integrability properties for MLFP approximations

Lemma 3.4. *Assume Setting 3.1, assume A is finite, let $L \in [0, \infty)$, let $\mathfrak{w} \colon \mathbb{X} \to (0, \infty)$ be $\mathcal{X}/\mathcal{B}((0, \infty))$-measurable, assume that for all $(x, a) \in \mathbb{X} \times A$, $r, s \in \mathbb{R}^A$ it holds that $|f(x, r) - f(x, s)| \leq L \max_{b \in A} |r(b) - s(b)|$, and $\big(\mathbb{E}\big[|f(X^{0,x,a}, 0)|^2 + |\mathfrak{w}(X^{0,x,a})|^2\big]\big)^{\frac{1}{2}} \leq L\mathfrak{w}(x)$. Then it holds for all $n \in \mathbb{N}_0$ that $\sup_{x \in \mathbb{X}} \big(\mathbb{E}\big[\max_{a \in A} |(V_n^0(x))(a)|^2\big] |\mathfrak{w}(x)|^{-2}\big) < \infty$.*

Proof of Lemma 3.4. For all $(x, a) \in \mathbb{X} \times A$, $\theta \in \Theta$ it holds that $(V_0^\theta(x))(a) = 0$. Hence it holds that $\sup_{x \in \mathbb{X}}(\mathbb{E}\big[\|V_0^0(x)\|_\infty^2\big] |\mathfrak{w}(x)|^{-2}) < \infty$. Fix $n \in \mathbb{N}$ and assume for all $l \in \{0, 1, \ldots, n-1\}$ that $\sup_{x \in \mathbb{X}}(\mathbb{E}\big[\|V_l^0(x)\|_\infty^2\big] |\mathfrak{w}(x)|^{-2}) < \infty$. Note that for all $m \in \mathbb{N}$ it holds that

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

Lemma 3.3 and Jensen's inequality prove that for all $m \in \mathbb{N}$, $(x, a) \in \mathbb{X} \times A$, $\theta \in \Theta$ it holds that

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

The triangle inequality, Jensen's inequality, the assumption that for all $x \in \mathbb{X}$, $r, s \in \mathbb{R}^A$ it holds that $|f(x, r) - f(x, s)| \leq L \max_{a \in A} |r(a) - s(a)|$, and the assumption that for all $(x, a) \in \mathbb{X} \times A$ it holds that $\big(\mathbb{E}\big[|f(X^{0,x,a}, 0)|^2\big]\big)^{\frac{1}{2}} \leq L\mathfrak{w}(x)$ ensure that for all $m \in \mathbb{N}_0$, $(x, a) \in \mathbb{X} \times A$, $\theta, \vartheta \in \Theta$ it holds that

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

13

Lemma 3.3 ensures for all $m \in \mathbb{N}$, $\vartheta \in \Theta$ that V_m^ϑ and V_m^0 are identically distributed. Combining this and the assumption that A is finite demonstrates that for all $m \in \mathbb{N}$, $\vartheta \in \Theta$ it holds that $\mathbb{X} \times \Omega \ni (x, \omega) \mapsto \|V_m^\vartheta(x, w)\|_\infty^2 \in [0, \infty)$ and $\mathbb{X} \times \Omega \ni (x, \omega) \mapsto \|V_m^0(x, \omega)\|_\infty^2 \in [0, \infty)$ are identically distributed. This, Lemma 3.2, the assumption that for all $\theta \in \Theta$ it holds that X^θ is $(\mathcal{X} \otimes \mathcal{F}^\theta)/\widetilde{\mathcal{X}}$-measurable, the assumption that the σ-algebras $\mathcal{F}^\theta$, $\theta \in \Theta$, are independent, and [31, Lemma 2.2] imply that for all $m \in \mathbb{N}_0$, $(x, a) \in \mathbb{X} \times A$, $\theta, \vartheta \in \Theta$ with $\theta \notin \bigcup_{\eta \in \Theta} \{(\vartheta, \eta)\}$ it holds that

$$\mathbb{E}\big[\|V_m^\vartheta(X^{\theta, x, a})\|_\infty^2\big] = \int_\mathbb{X} \mathbb{E}\big[\|V_m^\vartheta(y)\|_\infty^2\big] \big(X^{\theta, x, a}(\mathbb{P})\big)(dy)$$

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

This, (40), the assumption that for all $(x, a) \in \mathbb{X} \times A$ it holds that $(\mathbb{E}[|\mathfrak{w}(X^{0, x, a})|^2])^{\frac{1}{2}} \leq L\mathfrak{w}(x)$, and the assumption that for all $l \in \{0, 1, \ldots, n-1\}$ it holds that $\sup_{x \in \mathbb{X}}(\mathbb{E}[\|V_l^0(x)\|_\infty^2]|\mathfrak{w}(x)|^{-2}) < \infty$ prove for all $l \in \{0, 1, \ldots, n-1\}$, $(x, a) \in \mathbb{X} \times A$, $\theta, \vartheta \in \Theta$ with $\theta \notin \bigcup_{\eta \in \Theta} \{(\vartheta, \eta)\}$ that

$$\frac{\mathbb{E}\big[|f(X^{\theta, x, a}, V_l^\vartheta(X^{\theta, x, a}))|^2\big]}{|\mathfrak{w}(x)|^2} \leq \frac{2L^2}{|\mathfrak{w}(x)|^2} \int_\mathbb{X} \mathbb{E}\big[\|V_l^0(y)\|_\infty^2\big] \big(X^{0, x, a}(\mathbb{P})\big)(dy) + 2L^2$$

$$\leq 2L^2 \bigg(\sup_{y \in \mathbb{X}} \frac{\mathbb{E}\big[\|V_l^0(y)\|_\infty^2\big]}{|\mathfrak{w}(y)|^2}\bigg) \bigg(\frac{\mathbb{E}\big[|\mathfrak{w}(X^{0, x, a})|^2\big]}{|\mathfrak{w}(x)|^2}\bigg) + 2L^2$$

$$\leq 2L^4 \bigg(\sup_{y \in \mathbb{X}} \frac{\mathbb{E}\big[\|V_l^0(y)\|_\infty^2\big]}{|\mathfrak{w}(y)|^2}\bigg) + 2L^2. \tag{42}$$

This ensures that for all $l \in \{0, 1, \ldots, n-1\}$, $\theta, \vartheta \in \Theta$ with $\theta \notin \bigcup_{\eta \in \Theta} \{(\vartheta, \eta)\}$ that

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

This and (39) demonstrate for all $a \in A$ that $\sup_{x \in \mathbb{X}} \big(\mathbb{E}[|(V_n^0(x))(a)|^2]|\mathfrak{w}(x)|^{-2}\big) < \infty$. This, the assumption that A is finite, and (38) show that $\sup_{x \in \mathbb{X}} \big(\mathbb{E}[\|V_n^0(x)\|_\infty^2]|\mathfrak{w}(x)|^{-2}\big) < \infty$. Induction hence proves that for all $n \in \mathbb{N}_0$ it holds that $\sup_{x \in \mathbb{X}} \big(\mathbb{E}[\max_{a \in A} |(V_n^0(x))(a)|^2]|\mathfrak{w}(x)|^{-2}\big) < \infty$. The proof of Lemma 3.4 is thus completed. $\square$

Lemma 3.5. *Assume Setting 3.1, assume A is finite, let $L \in [0, \infty)$, let $\mathfrak{w} \colon \mathbb{X} \to (0, \infty)$ be $\mathcal{X}/\mathcal{B}((0, \infty))$-measurable, assume that for all $(x, a) \in \mathbb{X} \times A$, $r, s \in \mathbb{R}^A$ it holds that $|f(x, r) - f(x, s)| \leq L \max_{b \in A} |r(b) - s(b)|$, $(\mathbb{E}[|f(X^{0, x, a}, 0)|^2 + |\mathfrak{w}(X^{0, x, a})|^2])^{\frac{1}{2}} \leq L\mathfrak{w}(x)$. Then it holds for all $n \in \mathbb{N}$, $(x, a) \in \mathbb{X} \times A$, $\theta \in \Theta$ that $\mathbb{E}[|f(X^{0, x, a}, V_{n-1}^0(X^{0, x, a}))| + |(V_n^\theta(x))(a)|] < \infty$ and $\mathbb{E}[(V_n^\theta(x))(a)] = \mathbb{E}[f(X^{0, x, a}, V_{n-1}^0(X^{0, x, a}))]$.*

Proof of Lemma 3.5. The assumption that for all $x \in \mathbb{X}$, $r, s \in \mathbb{R}^A$ it holds that $|f(x, r) - f(x, s)| \leq L \max_{a \in A} |r(a) - s(a)|$, Lemma 3.3, Lemma 3.4, the fact that $0 \notin \bigcup_{\eta \in \Theta} \{(0, \eta)\}$, the assumption that $(\mathbb{E}[|f(X^{0, x, a}, 0)|^2])^{\frac{1}{2}} \leq L\mathfrak{w}(x)$, and [31, Lemma 2.2] prove that for all $n \in \mathbb{N}$, $(x, a) \in \mathbb{X} \times A$, $\theta \in \Theta$ it holds that $\mathbb{E}[|f(X^{0, x, a}, V_{n-1}^0(X^{0, x, a}))|^2 + |(V_n^\theta(x))(a)|^2] < \infty$. This ensures for all $n \in \mathbb{N}$, $(x, a) \in \mathbb{X} \times A$, $\theta \in \Theta$ that $\mathbb{E}[|f(X^{0, x, a}, V_{n-1}^0(X^{0, x, a}))| + |(V_n^\theta(x))(a)|] < \infty$. This, Lemma 3.2, the assumption that X^θ, $\theta \in \Theta$, are i.i.d. random fields, Lemma 3.3, and [31, Lemma 2.2] demonstrate that for all $n \in \mathbb{N}_0$, $(x, a) \in \mathbb{X} \times A$, $\theta, \vartheta \in \Theta$ with $\theta \notin \bigcup_{\eta \in \Theta} \{(\vartheta, \eta)\}$ it holds that

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

14

This establishes for all $n \in \mathbb{N}$, $(x, a) \in \mathbb{X} \times A$, $\theta \in \Theta$ that

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

The proof of Lemma 3.5 is thus completed. $\qquad\square$

3.4 Recursive error bounds for MLFP approximations

Lemma 3.6. *Assume Setting 3.1, assume A is finite, let $c, L \in [0, \infty)$, let $\mathfrak{w} \colon \mathbb{X} \to (0, \infty)$ be $\mathcal{X}/\mathcal{B}((0, \infty))$-meaurable, let $v \colon \mathbb{X} \to \mathbb{R}^A$ be $\mathcal{X}/\widetilde{\mathcal{A}}$-measurable, assume that for all $(x, a) \in \mathbb{X} \times A$, $r, s \in \mathbb{R}^A$ it holds that $|f(x, r) - f(x, s)| \leq L\max_{b \in A} |r(b) - s(b)|$, $\big(\mathbb{E}\big[|f(X^{0,x,a}, 0)|^2 + \|v(X^{0,x,a})\|_\infty^2 + |\mathfrak{w}(X^{0,x,a})|^2\big]\big)^{\frac{1}{2}} \leq c\mathfrak{w}(x)$, and $\big(v(x)\big)(a) = \mathbb{E}[f(X^{0,x,a}, v(X^{0,x,a}))]$. Then it holds for all $n \in \mathbb{N}$, $(x, a) \in \mathbb{X} \times A$, $\theta \in \Theta$ that*

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

Proof of Lemma 3.6. The assumption that for all $(x, a) \in \mathbb{X} \times A$ it holds that $\big(v(x)\big)(a) = \mathbb{E}\big[f(X^{0,x,a}, v(X^{0,x,a}))\big]$, Lemma 3.5, the assumption that for all $x \in \mathbb{X}$, $r, s \in \mathbb{R}^A$ it holds that $|f(x, r) - f(x, s)| \leq L\max_{a \in A} |r(a) - s(a)|$, Lemma 3.4, and the Cauchy-Schwarz inequality imply that for all $n \in \mathbb{N}$, $(x, a) \in \mathbb{X} \times A$, $\theta \in \Theta$ it holds that

$$\big|v(x)(a) - \mathbb{E}\big[V_n^\theta(x)(a)\big]\big| \leq \mathbb{E}\big[\big|f\big(X^{0,x,a}, v(X^{0,x,a})\big) - f\big(X^{0,x,a}, V_{n-1}^0(X^{0,x,a})\big)\big|\big]$$
$$\leq L\mathbb{E}\Big[\big\|v(X^{0,x,a}) - V_{n-1}^0(X^{0,x,a})\big\|_\infty\Big]$$
$$\leq L\left(\mathbb{E}\Big[\big\|v(X^{0,x,a}) - V_{n-1}^0(X^{0,x,a})\big\|_\infty^2\Big]\right)^{\frac{1}{2}}. \tag{47}$$

The proof of Lemma 3.6 is thus completed. $\qquad\square$

Lemma 3.7. *Assume Setting 3.1, assume A is finite, let $c, L \in [0, \infty)$, let $\mathfrak{w} \colon \mathbb{X} \to (0, \infty)$ be $\mathcal{X}/\mathcal{B}((0, \infty))$-measurable, let $v \colon \mathbb{X} \to \mathbb{R}^A$ be $\mathcal{X}/\widetilde{\mathcal{A}}$-measurable, assume that for all $(x, a) \in \mathbb{X} \times A$, $r, s \in \mathbb{R}^A$ it holds that $|f(x, r) - f(x, s)| \leq L\max_{b \in A} |r(b) - s(b)|$, $\big(\mathbb{E}\big[|f(X^{0,x,a}, 0)|^2 + \|v(X^{0,x,a})\|_\infty^2 + |\mathfrak{w}(X^{0,x,a})|^2\big]\big)^{\frac{1}{2}} \leq c\mathfrak{w}(x)$, and $\big(v(x)\big)(a) = \mathbb{E}\big[f(X^{0,x,a}, v(X^{0,x,a}))\big]$. Then it holds for all $n \in \mathbb{N}$, $(x, a) \in \mathbb{X} \times A$, $\theta \in \Theta$ that*

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

15

Proof of Lemma 3.7. Lemma 3.4 ensures for all $n \in \mathbb{N}$, $(x,a) \in \mathbb{X} \times A$, $\theta \in \Theta$ that $\mathrm{Var}\left[(V_n^\theta(x))(a)\right] < \infty$. Lemma 3.2 yields that for all $l, i \in \mathbb{N}_0$, $(x,a) \in \mathbb{X} \times A$, $\theta \in \Theta$ it holds that

$$
\Omega \ni \omega \mapsto f\big(X^{(\theta,l,i),x,a}(\omega), V_l^{(\theta,l,i)}(X^{(\theta,l,i),x,a}(\omega),\omega)\big)
$$
$$
- \mathbb{1}_{\mathbb{N}}(l) f\big(X^{(\theta,l,i),x,a}(\omega), V_{\max\{l-1,0\}}^{(\theta,-l,i)}(X^{(\theta,l,i),x,a}(\omega),\omega)\big) \in \mathbb{R} \qquad (49)
$$

is $\sigma\big(\bigcup_{\eta \in \Theta} \mathcal{F}^{(\theta,l,i,\eta)} \cup \mathcal{F}^{(\theta,-l,i,\eta)} \cup \mathcal{F}^{(\theta,l,i)}\big)/\mathcal{B}(\mathbb{R})$-measurable. This demonstrates that for all $n \in \mathbb{N}_0$, $(x,a) \in \mathbb{X} \times A$ it holds that

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

Lemma 3.2, Lemma 3.3, and [5, Lemma 2.6] ensure that for all $l, i \in \mathbb{N}_0$, $\theta \in \Theta$ it holds that

$$
\mathbb{X} \times \Omega \ni (x,\omega) \mapsto \big(V_l^{(\theta,l,i)}(x,\omega), V_{\max\{l-1,0\}}^{(\theta,-l,i)}(x,\omega)\big) \in \mathbb{R}^A \times \mathbb{R}^A \quad \text{and}
$$
$$
\mathbb{X} \times \Omega \ni (x,\omega) \mapsto \big(V_l^0(x,\omega), V_{\max\{l-1,0\}}^1(x,\omega)\big) \in \mathbb{R}^A \times \mathbb{R}^A. \qquad (51)
$$

are identically distributed random fields. This and the fact that f is $(\mathcal{X} \otimes \widetilde{\mathcal{A}})/\mathcal{B}(\mathbb{R})$-measurable show that for all $l, i \in \mathbb{N}_0$, $\theta \in \Theta$ it holds that

$$
\mathbb{X} \times \Omega \ni (x,\omega) \mapsto f(x, V_l^{(\theta,l,i)}(x,\omega)) - \mathbb{1}_{\mathbb{N}}(l) f(x, V_{\max\{l-1,0\}}^{(\theta,-l,i)}(x,\omega)) \in \mathbb{R} \quad \text{and}
$$
$$
\mathbb{X} \times \Omega \ni (x,\omega) \mapsto f(x, V_l^0(x,\omega)) - \mathbb{1}_{\mathbb{N}}(l) f(x, V_{\max\{l-1,0\}}^1(x,\omega)) \in \mathbb{R} \qquad (52)
$$

are identically distributed random fields. [5, Lemma 2.5] proves that for all $l, i \in \mathbb{N}_0$, $\theta \in \Theta$ it holds that

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

are identically distributed random fields. This implies for all $l, i \in \mathbb{N}_0$, $(x,a) \in \mathbb{X} \times A$, $\theta \in \Theta$ that

$$
\mathrm{Var}\left[f\big(X^{(\theta,l,i),x,a}, V_l^{(\theta,l,i)}(X^{(\theta,l,i),x,a})\big) - \mathbb{1}_{\mathbb{N}}(l) f\big(X^{(\theta,l,i),x,a}, V_{\max\{l-1,0\}}^{(\theta,-l,i)}(X^{(\theta,l,i),x,a})\big)\right]
$$
$$
= \mathrm{Var}\left[f\big(X^{0,x,a}, V_l^0(X^{0,x,a})\big) - \mathbb{1}_{\mathbb{N}}(l) f\big(X^{0,x,a}, V_{\max\{l-1,0\}}^1(X^{0,x,a})\big)\right]. \qquad (54)
$$

This and (50) ensure that for all $n \in \mathbb{N}$, $(x,a) \in \mathbb{X} \times A$, $\theta \in \Theta$ it holds that

$$
\mathrm{Var}[V_n^\theta(x)(a)] = \sum_{l=0}^{n-1} \frac{1}{M^{n-l}} \mathrm{Var}\left[f\big(X^{0,x,a}, V_l^0(X^{0,x,a})\big) - \mathbb{1}_{\mathbb{N}}(l) f\big(X^{0,x,a}, V_{\max\{l-1,0\}}^1(X^{0,x,a})\big)\right]
$$
$$
\leq \sum_{l=0}^{n-1} \frac{1}{M^{n-l}} \mathbb{E}\left[\left|f\big(X^{0,x,a}, V_l^0(X^{0,x,a})\big) - \mathbb{1}_{\mathbb{N}}(l) f\big(X^{0,x,a}, V_{\max\{l-1,0\}}^1(X^{0,x,a})\big)\right|^2\right]. \qquad (55)
$$

Combining this, the fact that for all $(x, a) \in \mathbb{X} \times A$ it holds that $\left(V_0^0(x)\right)(a) = 0$, and the fact that for all $n \in \mathbb{N}$, $r_1, r_2, \ldots, r_n \in [0, \infty)$ it holds that $\sqrt{r_1 + r_2 + \cdots + r_n} \leq \sqrt{r_1} + \sqrt{r_2} + \cdots + \sqrt{r_n}$ prove that for all $n \in \mathbb{N}$, $(x, a) \in \mathbb{X} \times A$, $\theta \in \Theta$ it holds that

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

The assumption that for all $x \in \mathbb{X}$, $r, s \in \mathbb{R}^A$ it holds that $|f(x, r) - f(x, s)| \leq L \max_{a \in A} |r(a) - s(a)|$ and the triangle inequality imply that for all $n \in \mathbb{N}$, $(x, a) \in \mathbb{X} \times A$, $\theta \in \Theta$ it holds that

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

The fact that for all $l \in \mathbb{N}_0$, $(x, a) \in \mathbb{X} \times A$ it holds that $\Omega \ni \omega \mapsto V_l^0(X^{0,x,a}(\omega), \omega) \in \mathbb{R}^A$ and $\Omega \ni \omega \mapsto V_l^1(X^{0,x,a}(\omega), \omega) \in \mathbb{R}^A$ are identically distributed ensures for all $l \in \mathbb{N}_0$, $(x, a) \in \mathbb{X} \times A$ that $\Omega \ni \omega \mapsto \|v(X^{0,x,a}(\omega)) - V_l^0(X^{0,x,a}(\omega), \omega)\|_\infty^2 \in [0, \infty)$ and $\Omega \ni \omega \mapsto \|v(X^{0,x,a}(\omega)) - V_l^1(X^{0,x,a}(\omega), \omega)\|_\infty^2 \in [0, \infty)$ are identically distributed. This and (57) establish that for all $n \in \mathbb{N}$, $(x, a) \in \mathbb{X} \times A$, $\theta \in \Theta$ it holds that

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

The proof of Lemma 3.7 is thus completed. $\qquad\square$

3.5 Non-recursive error bounds for MLFP approximations

Proposition 3.8. *Assume Setting 3.1, assume A is finite, let $c_f, c_v, c_{\mathfrak{w}}, L \in [0, \infty)$, let $\mathfrak{w} \colon \mathbb{X} \to (0, \infty)$ be $\mathcal{X}/\mathcal{B}((0, \infty))$-measurable, let $v \colon \mathbb{X} \to \mathbb{R}^A$ be $\mathcal{X}/\widetilde{\mathcal{A}}$-measurable, let $c = \frac{3}{2} \max\left\{ \frac{c_v}{c_{\mathfrak{w}}}, c_v L + c_f, \frac{|A| c_f}{c_{\mathfrak{w}} |A| L + 1} \right\}$, assume that for all $(x, a) \in \mathbb{X} \times A$, $r, s \in \mathbb{R}^A$ it holds that $|f(x, r) - f(x, s)| \le L \max_{b \in A} |r(b) - s(b)|$, $\left(\mathbb{E}[|f(X^{0,x,a}, 0)|^2] \right)^{\frac{1}{2}} \le c_f \mathfrak{w}(x)$, $\left(\mathbb{E}[\|v(X^{0,x,a})\|_\infty^2] \right)^{\frac{1}{2}} \le c_v \mathfrak{w}(x)$, $\left(\mathbb{E}[|\mathfrak{w}(X^{0,x,a})|^2] \right)^{\frac{1}{2}} \le c_{\mathfrak{w}} \mathfrak{w}(x)$, and $(v(x))(a) = \mathbb{E}[f(X^{0,x,a}, v(X^{0,x,a}))]$. Then it holds for all $n \in \mathbb{N}$, $x \in \mathbb{X}$ that*

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

Proof of Proposition 3.8. The triangle inequality, Lemma 3.4, and the assumption that A is finite ensure that for all $n \in \mathbb{N}$, $x \in \mathbb{X}$, $\theta \in \Theta$ it holds that

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

The assumption that $X^0 \colon \mathbb{X} \times \Omega \to \mathbb{X}^A$ is $(\mathcal{X} \otimes \mathcal{F}^0)/\widetilde{\mathcal{X}}$-measurable, the assumption that the σ-algebras $\mathcal{F}^\theta$, $\theta \in \Theta$, are independent, the assumption that A is finite, Lemma 3.2, and the disintegration theorem [31, Lemma 2.2] prove that for all $l \in \mathbb{N}$, $(x, a) \in \mathbb{X} \times A$ it holds that

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

This, Lemma 3.6, Lemma 3.7 and (60) imply that for all $n \in \mathbb{N}$, $x \in \mathbb{X}$ it holds that

$$
\left(\mathbb{E}\big[\|v(x) - V_n^0(x)\|_\infty^2\big]\right)^{\frac{1}{2}}
$$

$$
\leq c_{\mathfrak{w}} \mathfrak{w}(x) L \left(\sup_{y \in \mathbb{X}} \frac{\mathbb{E}\big[\|v(y) - V_{n-1}^0(y)\|_\infty^2\big]}{|\mathfrak{w}(y)|^2} \right)^{\frac{1}{2}} + \sum_{a \in A} \left[\frac{1}{\sqrt{M^n}} \left(\mathbb{E}\big[|f(X^{0,x,a}, 0)|^2\big]\right)^{\frac{1}{2}} \right.
$$

$$
+ \mathbb{1}_{[2,\infty)(n)} \frac{L\sqrt{M}}{\sqrt{M^n}} \left(\mathbb{E}\big[\|v(X^{0,x,a})\|_\infty^2\big]\right)^{\frac{1}{2}} + \mathbb{1}_{[2,\infty)}(n) \frac{c_{\mathfrak{w}} \mathfrak{w}(x) L}{\sqrt{M}} \left(\sup_{y \in \mathbb{X}} \frac{\mathbb{E}\big[\|v(y) - V_{n-1}^0(y)\|_\infty^2\big]}{|\mathfrak{w}(y)|^2} \right)^{\frac{1}{2}}
$$

$$
\left. + \sum_{l=1}^{n-2} \frac{c_{\mathfrak{w}} \mathfrak{w}(x) L (1 + \sqrt{M})}{\sqrt{M^{n-l}}} \left(\sup_{y \in X} \frac{\mathbb{E}\big[\|v(y) - V_l^0(y)\|_\infty^2\big]}{|\mathfrak{w}(y)|^2} \right)^{\frac{1}{2}} \right]
$$

$$
\leq c_{\mathfrak{w}} \mathfrak{w}(x) L \left(\sup_{y \in \mathbb{X}} \frac{\mathbb{E}\big[\|v(y) - V_{n-1}^0(y)\|_\infty^2\big]}{|\mathfrak{w}(y)|^2} \right)^{\frac{1}{2}} + \frac{|A| \mathfrak{w}(x)}{\sqrt{M^n}} \left(c_f + \mathbb{1}_{[2,\infty)}(n) c_v L \sqrt{M}\right)
$$

$$
+ \mathbb{1}_{[2,\infty)}(n) \frac{c_{\mathfrak{w}} |A| L \mathfrak{w}(x)}{\sqrt{M}} \left(\sup_{y \in \mathbb{X}} \frac{\mathbb{E}\big[\|v(y) - V_{n-1}^0(y)\|_\infty^2\big]}{|\mathfrak{w}(y)|^2} \right)^{\frac{1}{2}}
$$

$$
+ c_{\mathfrak{w}} \mathfrak{w}(x) |A| L (1 + \sqrt{M}) \sum_{l=1}^{n-2} \frac{1}{\sqrt{M^{n-l}}} \left(\sup_{y \in \mathbb{X}} \frac{\mathbb{E}\big[\|v(y) - V_l^0(y)\|_\infty^2\big]}{|\mathfrak{w}(y)|^2} \right)^{\frac{1}{2}}. \tag{62}
$$

For all $n \in \mathbb{N}_0$ let

$$
F_n = \left(\sup_{x \in \mathbb{X}} \frac{\mathbb{E}\big[\|v(x) - V_n^0(x)\|_\infty^2\big]}{|\mathfrak{w}(x)|^2} \right)^{\frac{1}{2}} \quad \text{and} \quad a_n = M^{\frac{n}{2}} F_n. \tag{63}
$$

Lemma 3.4 ensures that for all $n \in \mathbb{N}_0$ it holds that $a_n, F_n \in [0, \infty)$. Note that (62) shows for all $n \in \mathbb{N} \cap [2, \infty)$ that

$$
F_n \leq |A| M^{-\frac{n}{2}} \left(c_f + c_v L M^{\frac{1}{2}}\right) + c_{\mathfrak{w}} L (1 + |A| M^{-\frac{1}{2}}) \left(\sup_{y \in \mathbb{X}} \frac{\mathbb{E}\big[\|v(y) - V_{n-1}^0(y)\|_\infty^2\big]}{|\mathfrak{w}(y)|^2} \right)^{\frac{1}{2}}
$$

$$
+ c_{\mathfrak{w}} |A| L (1 + M^{\frac{1}{2}}) \sum_{l=1}^{n-2} M^{-\frac{n-l}{2}} \left(\sup_{y \in \mathbb{X}} \frac{\mathbb{E}\big[\|v(y) - V_l^0(y)\|_\infty^2\big]}{|\mathfrak{w}(y)|^2} \right)^{\frac{1}{2}}
$$

$$
= |A| M^{-\frac{n}{2}} \left(c_f + c_v L M^{\frac{1}{2}}\right) + c_{\mathfrak{w}} L (1 + |A| M^{-\frac{1}{2}}) F_{n-1} + c_{\mathfrak{w}} |A| L (1 + M^{\frac{1}{2}}) \sum_{l=1}^{n-2} M^{-\frac{n-l}{2}} F_l. \tag{64}
$$

This implies for all $n \in \mathbb{N} \cap [2, \infty)$ that

$$
a_n \leq |A|(c_f + c_v L M^{\frac{1}{2}}) + c_{\mathfrak{w}} M^{\frac{1}{2}} L (1 + |A| M^{-\frac{1}{2}}) M^{\frac{n-1}{2}} F_{n-1} + c_{\mathfrak{w}} |A| L (1 + M^{\frac{1}{2}}) \sum_{l=1}^{n-2} M^{\frac{l}{2}} F_l
$$

$$
= |A|(c_f + c_v L M^{\frac{1}{2}}) + c_{\mathfrak{w}} L (M^{\frac{1}{2}} + |A|) a_{n-1} + c_{\mathfrak{w}} |A| L (1 + M^{\frac{1}{2}}) \sum_{l=1}^{n-2} a_l. \tag{65}
$$

Moreover it holds that

$$
a_0 = F_0 = \left(\sup_{y \in \mathbb{X}} \frac{\mathbb{E}\big[\|v(y)\|_\infty^2\big]}{|\mathfrak{w}(y)|^2} \right)^{\frac{1}{2}} = \sup_{(y,a) \in \mathbb{X} \times A} \frac{|(v(y))(a)|}{|\mathfrak{w}(y)|}. \tag{66}
$$

19

The assumption that for all $(x, a) \in \mathbb{X} \times A$ it holds that $(v(x))(a) = \mathbb{E}\big[f(X^{0,x,a}, v(X^{0,x,a}))\big]$, Jensen's inequality, the triangle inequality, the assumption that for all $x \in \mathbb{X}$, $r, s \in \mathbb{R}^A$ it holds that $|f(x, r) - f(x, s)| \leq L \max_{a \in A} |r(a) - s(a)|$, and the assumption that for all $(x, a) \in \mathbb{X} \times A$ it holds that $\big(\mathbb{E}\big[|f(X^{0,x,a}, 0)|^2\big]\big)^{\frac{1}{2}} \leq c_f \mathfrak{w}(x)$ and $\big(\mathbb{E}\big[\|v(X^{0,x,a})\|_\infty^2\big]\big)^{\frac{1}{2}} \leq c_v \mathfrak{w}(x)$ prove for all $(x, a) \in \mathbb{X} \times A$ that

$$
\frac{|(v(x))(a)|}{|\mathfrak{w}(x)|} = \frac{1}{\mathfrak{w}(x)}\big(\big|\mathbb{E}[f(X^{0,x,a}, v(X^{0,x,a}))]\big|^2\big)^{\frac{1}{2}} \leq \frac{1}{\mathfrak{w}(x)}\big(\mathbb{E}\big[\big|f(X^{0,x,a}, v(X^{0,x,a}))\big|^2\big]\big)^{\frac{1}{2}}
$$
$$
\leq \frac{1}{\mathfrak{w}(x)}\Big[L\big(\mathbb{E}\big[\sup_{b \in A}|v(X^{0,x,a})(b)|^2\big]\big)^{\frac{1}{2}} + \big(\mathbb{E}\big[|f(X^{0,x,a}, 0)|^2\big]\big)^{\frac{1}{2}}\Big] \leq c_v L + c_f \qquad (67)
$$

This and (66) ensure that $a_0 \leq c_v L + c_f$. Note that (62) implies that $a_1 \leq M^{\frac{1}{2}}\big(c_\mathfrak{w} L a_0 + |A| M^{-\frac{1}{2}} c_f\big) \leq c_\mathfrak{w} L M^{\frac{1}{2}}(c_v L + c_f) + |A| c_f$. Moreover (65) ensures $a_2 \leq |A|(c_f + c_v L M^{\frac{1}{2}}) + c_\mathfrak{w} L(M^{\frac{1}{2}} + |A|)a_1$. Let $(\xi_n)_{n \in \mathbb{N}_0}$, $(b_n)_{n \in \mathbb{N}_0} \subseteq \mathbb{R}$ satisfy for all $n \in \mathbb{N}_0$ that

$$
b_0 = \xi_0 = \max\left\{\frac{c_v}{c_\mathfrak{w}}, c_v L + c_f, \frac{|A| c_f}{c_\mathfrak{w}|A|L + 1}\right\}, \quad b_1 = |A|c_f - (c_\mathfrak{w}|A|L + 1)b_0,
$$
$$
b_2 = |A|LM^{\frac{1}{2}}(c_v - c_\mathfrak{w} b_0), \quad b_{n+3} = 0, \quad \xi_1 = c_\mathfrak{w} LM^{\frac{1}{2}} b_0 + |A|c_f,
$$
$$
\xi_{n+2} = |A|(c_f + c_v LM^{\frac{1}{2}}) + c_\mathfrak{w} L(M^{\frac{1}{2}} + |A|)\xi_{n+1} + c_\mathfrak{w}|A|L(1 + M^{\frac{1}{2}})\sum_{l=1}^{n} \xi_l. \qquad (68)
$$

Combining (62), (65), (68), and induction establishes for all $n \in \mathbb{N}_0$ that $a_n \leq \xi_n$. Observe that (68) ensures that for all $n \in \mathbb{N}_0$ it holds that

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

Let $x_1, x_2 \in \mathbb{R}$ satisfy

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

Note that for all $i \in \{1, 2\}$ it holds that $x_i^2 = \big(c_\mathfrak{w} L(M^{\frac{1}{2}} + |A|) + 1\big)x_i + c_\mathfrak{w} LM^{\frac{1}{2}}(|A| - 1)$. Moreover note that the assumption that A is nonempty yields that $c_\mathfrak{w} L(M^{\frac{1}{2}} + |A|) + 1 \geq 1$ and $c_\mathfrak{w} LM^{\frac{1}{2}}(|A| - 1) \geq 0$. This ensures that $x_2 > 0 \geq x_1$. This and (70) imply that $|x_2| \geq |x_1|$. Hence it holds that

$$
\frac{|x_2|}{|x_2 - x_1|} \leq 1 \quad \text{and} \quad \frac{|x_1|}{|x_2 - x_1|} \leq \frac{1}{2}. \qquad (71)
$$

This, the fact that $|x_1| \leq |x_2|$, and the triangle inequality prove that for all $n \in \mathbb{N}_0$ it holds that

$$
\frac{x_2^{n+1} - x_1^{n+1}}{x_2 - x_1} \leq \frac{|x_2|^{n+1}}{x_2 - x_1} + \frac{|x_1|^{n+1}}{x_2 - x_1} \leq \frac{x_2}{x_2 - x_1}x_2^n + \frac{|x_1|}{x_2 - x_1}x_2^n \leq \frac{3}{2}x_2^n. \qquad (72)
$$

The discrete Gronwall-type two-step recursion in [34, Lemma 2.1] (applied with $\kappa \leftarrow c_{\mathfrak{w}} L(M^{\frac{1}{2}} + |A|) + 1$, $\lambda \leftarrow c_{\mathfrak{w}} L M^{\frac{1}{2}}(|A| - 1)$, $(a_k)_{k\in\mathbb{N}_0} \leftarrow (\xi_k)_{k\in\mathbb{N}_0}$, $(b_k)_{k\in\mathbb{N}_0} \leftarrow (b_k)_{k\in\mathbb{N}_0}$, $x_{1/2} \leftarrow x_{1/2}$ in the notation of [34, Lemma 2.1]) demonstrates that for all $n \in \mathbb{N}_0$ it holds that

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

Note that (68) ensures that $b_1 \leq 0$ and $b_2 \leq 0$. This, the fact that $x_2 > x_1$, and the fact that for all $n \in \mathbb{N}_0$ it holds that $a_n \leq \xi_n$ proves that for all $n \in \mathbb{N}_0$ it holds that

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

Combining this and (72) implies that for all $n \in \mathbb{N}_0$ it holds that

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

Furthermore, observe that

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

$$= \frac{c_{\mathfrak{w}} L(1 + |A| M^{-\frac{1}{2}}) + M^{-\frac{1}{2}} + \sqrt{\big(c_{\mathfrak{w}} L(1 + |A| M^{-\frac{1}{2}}) + M^{-\frac{1}{2}}\big)^2 + 4c_{\mathfrak{w}} L M^{-\frac{1}{2}}(|A| - 1)}}{2} \tag{76}$$

This, (63), and (75) prove that for all $x \in \mathbb{X}$, $n \in \mathbb{N}$ it holds that

$$\left(\frac{\mathbb{E}\big[\|v(x) - V_n^\theta(x)\|_\infty^2\big]}{|\mathfrak{w}(x)|^2}\right)^{\frac{1}{2}} \tag{77}$$

$$\leq c\left(\frac{c_{\mathfrak{w}} L(1 + |A| M^{-\frac{1}{2}}) + M^{-\frac{1}{2}} + \sqrt{\big(c_{\mathfrak{w}} L(1 + |A| M^{-\frac{1}{2}}) + M^{-\frac{1}{2}}\big)^2 + 4c_{\mathfrak{w}} L M^{-\frac{1}{2}}(|A| - 1)}}{2}\right)^n.$$

The proof of Proposition 3.8 is thus completed. $\qquad\square$

Corollary 3.9. *Assume Setting 3.1, assume A is finite, let $c_f, c_v, c_{\mathfrak{w}}, L \in [0, \infty)$ with $c_{\mathfrak{w}} L < 1$, let $M > (1 + c_{\mathfrak{w}} L(2|A| - 1))^2(1 - c_{\mathfrak{w}} L)^{-2}$, let $\mathfrak{w}\colon \mathbb{X} \to (0, \infty)$ be $\mathcal{X}/\mathcal{B}((0, \infty))$-measurable, let $v\colon \mathbb{X} \to \mathbb{R}^A$ be $\mathcal{X}/\widetilde{\mathcal{A}}$-measurable, assume for all $(x, a) \in \mathbb{X} \times A$, $r, s \in \mathbb{R}^A$ that $|f(x, r) - f(x, s)| \leq L \max_{b\in A} |r(b) - s(b)|$, $\big(\mathbb{E}\big[|f(X^{0,x,a}, 0)|^2\big]\big)^{\frac{1}{2}} \leq c_f \mathfrak{w}(x)$, $\big(\mathbb{E}\big[\|v(X^{0,x,a})\|_\infty^2\big]\big)^{\frac{1}{2}} \leq c_v \mathfrak{w}(x)$, $\big(\mathbb{E}\big[|\mathfrak{w}(X^{0,x,a})|^2\big]\big)^{\frac{1}{2}} \leq c_{\mathfrak{w}} \mathfrak{w}(x)$, and $(v(x))(a) = \mathbb{E}\big[f\big(X^{0,x,a}, v(X^{0,x,a})\big)\big]$. Then it holds that*

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

Proof of Corollary 3.9. Let $c \in [0, \infty)$ satisfy $c = \frac{3}{2}\max\big\{\frac{c_v}{c_{\mathfrak{w}}}, c_v L + c_f, \frac{|A| c_f}{c_{\mathfrak{w}} |A| L + 1}\big\}$. Proposition 3.8 establishes that for all $n \in \mathbb{N}$, $x \in \mathbb{X}$ it holds that

$$\left(\frac{\mathbb{E}\big[\|v(x) - V_n^0(x)\|_\infty^2\big]}{|\mathfrak{w}(x)|^2}\right)^{\frac{1}{2}} \tag{79}$$

$$\leq c\left(\frac{c_{\mathfrak{w}} L(1 + |A| M^{-\frac{1}{2}}) + M^{-\frac{1}{2}} + \sqrt{\big(c_{\mathfrak{w}} L(1 + |A| M^{-\frac{1}{2}}) + M^{-\frac{1}{2}}\big)^2 + 4c_{\mathfrak{w}} L M^{-\frac{1}{2}}(|A| - 1)}}{2}\right)^n.$$

21

Observe that the assumption that $M > (1 + c_{\mathfrak{w}}L(2|A| - 1))^2(1 - c_{\mathfrak{w}}L)^{-2}$ implies $M^{-\frac{1}{2}} < (1 - c_{\mathfrak{w}}L)(1 + c_{\mathfrak{w}}L(2|A| - 1))^{-1}$. This proves that

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

The assumption that A is nonempty ensures that $|A| \leq 2|A| - 1$. This and the assumption that $M > (1 + c_{\mathfrak{w}}L(2|A| - 1))^2(1 - c_{\mathfrak{w}}L)^{-2}$ yield $M > (1 + c_{\mathfrak{w}}L|A|)^2(1 - c_{\mathfrak{w}}L)^{-2}$. Therefore it holds that $c_{\mathfrak{w}}L(1 + |A|M^{-\frac{1}{2}}) + M^{-\frac{1}{2}} = c_{\mathfrak{w}}L + (1 + c_{\mathfrak{w}}|A|L)M^{-\frac{1}{2}} < 1 < 2$. This and (80) establish that

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

Combining (79) and (81) implies that

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

The proof of Corollary 3.9 is thus completed. $\qquad\square$

4 Computational complexity analysis for MLFP approximations

4.1 MLFP approximations for functional fixed-point equations

Theorem 4.1. *Let $M \in \mathbb{N}$, let $\Theta = \bigcup_{n\in\mathbb{N}} \mathbb{Z}^n$, let A be a finite nonempty set, let $\kappa \in [0, \infty)$, let $\mathfrak{D}$ be a nonempty set, let $\lambda_d, L_d, \mathfrak{R}_d \in [0, \infty)$, $d \in \mathfrak{D}$, let $(\Omega, \mathcal{F}, \mathbb{P})$ be a probability space, let $(\mathbb{X}_d, \mathcal{X}_d)$, $d \in \mathfrak{D}$, be nonempty measurable spaces, for every $d \in \mathfrak{D}$ let $\mathfrak{w}_d \colon \mathbb{X}_d \to (0, \infty)$ be $\mathcal{X}_d/\mathcal{B}((0, \infty))$-measurable, for every $d \in \mathfrak{D}$ let $f_d \colon \mathbb{X}_d \times \mathbb{R}^A \to \mathbb{R}$, be $(\mathcal{X}_d \otimes (\bigotimes_{a\in A}\mathcal{B}(\mathbb{R})))/\mathcal{B}(\mathbb{R})$-measurable, for every $d \in \mathfrak{D}$ let $(\mathcal{F}_d^\theta)_{\theta\in\Theta}$ be independent sub-σ-algebras of $\mathcal{F}$, for every $d \in \mathfrak{D}$ let $X_d^\theta \colon \mathbb{X}_d \times \Omega \to \mathbb{X}_d^A$, $\theta \in \Theta$, be i.i.d. random fields which satisfy for all $d \in \mathfrak{D}$, $\theta \in \Theta$ that X_d^θ is $(\mathcal{X}_d \otimes \mathcal{F}_d^\theta)/(\bigotimes_{a\in A}\mathcal{X}_d)$-measurable, assume for all $d \in \mathfrak{D}$, $(x, a) \in \mathbb{X}_d \times A$, $r, s \in \mathbb{R}^A$ that $|f_d(x, r) - f_d(x, s)| \leq L_d \max_{b\in A}|r(b) - s(b)|$, $(\mathbb{E}[|f_d(X_d^{0,x,a}, 0)|^2])^{\frac{1}{2}} \leq \kappa\mathfrak{w}_d(x)$, $(\mathbb{E}[|\mathfrak{w}_d(X_d^{0,x,a})|^2])^{\frac{1}{2}} \leq \lambda_d\mathfrak{w}_d(x)$, $\sup_{u\in\mathfrak{D}}\lambda_u L_u < 1$, assume $M > \frac{(1+(\sup_{d\in\mathfrak{D}}\lambda_d L_d)(2|A|-1))^2}{(1-(\sup_{d\in\mathfrak{D}}\lambda_d L_d))^2}$, for every $d \in \mathfrak{D}$ let $V_{n,d}^\theta \colon \mathbb{X}_d \times \Omega \to \mathbb{R}^A$, $n \in \mathbb{N}_0$, $\theta \in \Theta$, satisfy for all $n \in \mathbb{N}_0$, $(x, a) \in \mathbb{X}_d \times A$, $\theta \in \Theta$*

that

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

and let $\mathfrak{C}_{n,d} \in [0,\infty)$, $n \in \mathbb{N}_0$, $d \in \mathfrak{D}$, satisfy for all $n \in \mathbb{N}_0$, $d \in \mathfrak{D}$ that

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

Then the following holds:

(i) *For every $d \in \mathfrak{D}$ there exists a unique function $v_d \colon \mathbb{X}_d \to \mathbb{R}^A$ which is $\mathcal{X}_d/(\bigotimes_{a \in A} \mathcal{B}(\mathbb{R}))$-measurable and satisfies for all $(x,a) \in \mathbb{X}_d \times A$ that $\sup_{y \in \mathbb{X}_d} \frac{\|v_d(y)\|_{\infty}}{\mathfrak{w}_d(y)} < \infty$, $\mathbb{E}\big[|f_d(X_d^{0,x,a}, v_d(X_d^{0,x,a}))|\big] < \infty$, and*

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

(ii) *There exist $N \colon (0,1] \to \mathbb{N}$ and $c \in \mathbb{R}$ such that for all $d \in \mathfrak{D}$, $\varepsilon \in (0,1]$ it holds that $\mathfrak{C}_{N_{\varepsilon},d} \le c\mathfrak{R}_d \varepsilon^{-c}$ and*

$$\sup_{x \in \mathbb{X}_d} \left(\frac{\mathbb{E}\big[\|v_d(x) - V_{N_{\varepsilon},d}^0(x)\|_{\infty}^2\big]}{|\mathfrak{w}_d(x)|^2}\right)^{\frac{1}{2}} \le \varepsilon. \tag{86}$$

Proof of Theorem 4.1. Let $\alpha, \beta, \gamma \in \mathbb{R} \cup \{\infty\}$ satisfy

$$\alpha = \frac{1}{2}\sup_{d \in \mathfrak{D}}\left[\lambda_d L_d(1 + |A|M^{-\frac{1}{2}}) + M^{-\frac{1}{2}} + \sqrt{\big(\lambda_d L_d(1 + |A|M^{-\frac{1}{2}}) + M^{-\frac{1}{2}}\big)^2 + 4M^{-\frac{1}{2}}\lambda_d L_d(|A| - 1)}\right],$$

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

Let $N \colon (0,1] \to \mathbb{N} \cup \{\infty\}$ satisfy for all $\varepsilon \in (0,1]$ that $N_{\varepsilon} = \min(\{n \in \mathbb{N} : \gamma\alpha^n \le \varepsilon\} \cup \{\infty\})$. The assumption that $\sup_{d \in \mathfrak{D}} \lambda_d L_d < 1$ ensures that $\gamma \in [0,\infty)$. The triangle inequality yields

$$\alpha = \frac{1}{2}\sup_{d \in \mathfrak{D}}\left[\lambda_d L_d(1 + |A|M^{-\frac{1}{2}}) + M^{-\frac{1}{2}} + \sqrt{\big(\lambda_d L_d(1 + |A|M^{-\frac{1}{2}}) + M^{-\frac{1}{2}}\big)^2 + 4M^{-\frac{1}{2}}\lambda_d L_d(|A| - 1)}\right],$$

$$\le \frac{1}{2}\left[(\sup_{d \in \mathfrak{D}}\lambda_d L_d)(1 + |A|M^{-\frac{1}{2}}) + M^{-\frac{1}{2}}\right.$$

$$\left. + \sqrt{\big((\sup_{d \in \mathfrak{D}}\lambda_d L_d)(1 + |A|M^{-\frac{1}{2}}) + M^{-\frac{1}{2}}\big)^2 + 4M^{-\frac{1}{2}}(\sup_{d \in \mathfrak{D}}\lambda_d L_d)(|A| - 1)}\right]. \tag{88}$$

23

The assumption that $M^{-\frac{1}{2}} < \frac{1-(\sup_{d\in\mathfrak{D}}\lambda_d L_d)}{1+(\sup_{d\in\mathfrak{D}}\lambda_d L_d)(2|A|-1)}$ ensures that

$$1 - \left((\sup_{d\in\mathfrak{D}}\lambda_d L_d)(1+|A|M^{-\frac{1}{2}})+M^{-\frac{1}{2}}\right) = 1 - (\sup_{d\in\mathfrak{D}}\lambda_d L_d) - (1+|A|(\sup_{d\in\mathfrak{D}}\lambda_d L_d))M^{-\frac{1}{2}}$$

$$> (1-(\sup_{d\in\mathfrak{D}}\lambda_d L_d)) - (1+|A|(\sup_{d\in\mathfrak{D}}\lambda_d L_d))\frac{1-(\sup_{d\in\mathfrak{D}}\lambda_d L_d)}{1+(\sup_{d\in\mathfrak{D}}\lambda_d L_d)(2|A|-1)}$$

$$= \left(1+(\sup_{d\in\mathfrak{D}}\lambda_d L_d)(2|A|-1)-\left(|A|(\sup_{d\in\mathfrak{D}}\lambda_d L_d)+1\right)\right)\frac{1-(\sup_{d\in\mathfrak{D}}\lambda_d L_d)}{1+(\sup_{d\in\mathfrak{D}}\lambda_d L_d)(2|A|-1)}$$

$$= \left((\sup_{d\in\mathfrak{D}}\lambda_d L_d)(|A|-1)\right)\frac{1-(\sup_{d\in\mathfrak{D}}\lambda_d L_d)}{1+(\sup_{d\in\mathfrak{D}}\lambda_d L_d)(2|A|-1)}$$

$$> (\sup_{d\in\mathfrak{D}}\lambda_d L_d)(|A|-1)M^{-\frac{1}{2}}. \tag{89}$$

The assumption that A is nonempty implies that $|A| \leq 2|A|-1$. Hence in holds that $M > \frac{(1+(\sup_{d\in\mathfrak{D}}\lambda_d L_d)(2|A|-1))^2}{(1-(\sup_{d\in\mathfrak{D}}\lambda_d L_d))^2} \geq \frac{(1+|A|(\sup_{t\in T}\lambda_d L_d)^2}{(1-(\sup_{d\in\mathfrak{D}}\lambda_d L_d))^2}$. This establishes that

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

Combining this, (88), and (89) demonstrates that

$$\alpha \leq \frac{1}{2}\Bigg[(\sup_{d\in\mathfrak{D}}\lambda_d L_d)(1+|A|M^{-\frac{1}{2}})+M^{-\frac{1}{2}}$$

$$+ \sqrt{\left((\sup_{d\in\mathfrak{D}}\lambda_d L_d)(1+|A|M^{-\frac{1}{2}})+M^{-\frac{1}{2}}\right)^2+4M^{-\frac{1}{2}}(\sup_{d\in\mathfrak{D}}\lambda_d L_d)(|A|-1)}\Bigg]$$

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

$$+ \sqrt{\left((\sup_{d\in\mathfrak{D}}\lambda_d L_d)(1+|A|M^{-\frac{1}{2}})+M^{-\frac{1}{2}}\right)^2+4\left(1-(\sup_{d\in\mathfrak{D}}\left(\lambda_d L_d)(1+|A|M^{-\frac{1}{2}})+M^{-\frac{1}{2}}\right)\right)}\Bigg]$$

$$= \frac{1}{2}\Bigg[(\sup_{d\in\mathfrak{D}}\lambda_d L_d)(1+|A|M^{-\frac{1}{2}})+M^{-\frac{1}{2}}+\sqrt{\left(2-\left((\sup_{d\in\mathfrak{D}}\lambda_d L_d)(1+|A|M^{-\frac{1}{2}})+M^{-\frac{1}{2}}\right)\right)^2}\Bigg]$$

$$= 1. \tag{91}$$

This and (87) ensure that $\alpha \in [M^{-\frac{1}{2}},1)$. The assumption that $M > \frac{(1+(\sup_{d\in\mathfrak{D}}\lambda_d L_d)(2|A|-1))^2}{(1-(\sup_{d\in\mathfrak{D}}\lambda_d L_d))^2}$ implies that $M > 1$. Hence it holds that $[M^{-\frac{1}{2}},1) \neq \emptyset$. The fact that $\alpha \in (0,1)$ demonstrates that for all $\varepsilon \in (0,1]$ it holds that $N_\varepsilon \in \mathbb{N}$. The fact that $\alpha \in [M^{-\frac{1}{2}},1)$ and the fact that $M > 1$ ensure that $\beta \in (2,\infty)$. Corollary 2.3 yields that for all $d \in \mathfrak{D}$ there exists a unique function $v_d\colon \mathbb{X}_d \to \mathbb{R}^A$ which is $\mathcal{X}_d/(\bigotimes_{a\in A}\mathcal{B}(\mathbb{R}))$-measurable and satisfies for all $d \in \mathfrak{D}$, $(x,a) \in \mathbb{X}_d \times A$ that $\sup_{y\in\mathbb{X}_d}\frac{\|v_d(y)\|_\infty}{\mathfrak{w}_d(y)} < \infty$, $\mathbb{E}\big[|f_d(X_d^{0,x,a}v_d(X_d^{0,x,a}))|\big] < \infty$, and $v_d(x,a) = \mathbb{E}\big[f_d(X_d^{0,x,a},v_d(X_d^{0,x,a}))\big]$. This proves item (i). Lemma 2.4 implies for all $d \in \mathfrak{D}$, $(x,a) \in \mathbb{X}_d \times A$ that $\big(\mathbb{E}\big[\|v_d(X_d^{0,x,a})\|_\infty^2\big]\big)^{\frac{1}{2}} \leq \frac{\kappa\lambda_d}{1-\lambda_d L_d}\mathfrak{w}_d(x)$. Proposition 3.8 establishes for all $n \in \mathbb{N}$, $d \in \mathfrak{D}$, $x \in \mathbb{X}_d$ that

$$\left(\frac{\mathbb{E}\big[\|v_d(X_d^{0,x,a})-V_{n,d}^0(X_d^{0,x,a})\|_\infty^2\big]}{|\mathfrak{w}_d(x)|^2}\right)^{\frac{1}{2}} \tag{92}$$

$$\leq \gamma\left(\frac{\lambda_d L_d(1+|A|M^{-\frac{1}{2}})+M^{-\frac{1}{2}}+\sqrt{(\lambda_d L_d(1+|A|M^{-\frac{1}{2}})+M^{-\frac{1}{2}})^2+4\lambda_d L_d M^{-\frac{1}{2}}(|A|-1)}}{2}\right)^n.$$

This demonstrates for all $n \in \mathbb{N}$, $d \in \mathfrak{D}$ that

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

Note that [3, Lemma 3.14] (applied with $M \leftarrow M$, $\alpha \leftarrow \mathfrak{R}_d + \mathfrak{C}_{0,d}$, $\beta \leftarrow \mathfrak{R}_d$, $(C_n) \leftarrow (\mathfrak{C}_{n,d})$ for $d \in \mathfrak{D}$ in the notation of [3, Lemma 3.14]) proves that for all $n \in \mathbb{N}$, $d \in \mathfrak{D}$ it holds that

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

This, (93), and [3, Lemma 3.15] (applied with $m \leftarrow 1$, $\alpha \leftarrow \alpha$, $\beta \leftarrow 3M$, $\kappa_1 \leftarrow \gamma$, $\kappa_2 \leftarrow \mathfrak{R}_d$, $N \leftarrow N$ for $d \in \mathfrak{D}$ in the notation of [3, Lemma 3.15]) implies for all $d \in \mathfrak{D}$, $\varepsilon \in (0,1]$ that

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

Let $c = \max\{\beta, 3M \max\{1,\gamma\}^\beta\}$. Hence it holds for all $d \in \mathfrak{D}$, $\varepsilon \in (0,1]$ that $\mathfrak{C}_{N_\varepsilon,d} \leq c\mathfrak{R}_d \varepsilon^{-c}$. This completes the proof of Theorem 4.1. $\qquad\square$

4.2 MLFP approximations for Bellman equations of optimal control problems

Corollary 4.2. *Let $M \in \mathbb{N}$, $\kappa \in [0,\infty)$, $\Theta = \bigcup_{n \in \mathbb{N}} \mathbb{Z}^n$, let A be a finite nonempty set, let $\mathfrak{D}$ be a nonempty set, let $(\Omega, \mathcal{F}, \mathbb{P})$ be a probability space, let $\lambda_d, \delta_d, \mathfrak{R}_d \in [0,\infty)$, $d \in \mathfrak{D}$, let $(\mathbb{X}_d, \mathcal{X}_d)$, $d \in \mathfrak{D}$, be nonempty measurable spaces, for every $d \in \mathfrak{D}$ let $\mathfrak{w}_d \colon \mathbb{X}_d \to (0,\infty)$ be $\mathcal{X}_d/\mathcal{B}((0,\infty))$-measurable, for every $d \in \mathfrak{D}$ let $g_d \colon \mathbb{X}_d \times A \to \mathbb{R}$ be $(\mathcal{X}_d \otimes 2^A)/\mathcal{B}(\mathbb{R})$-measurable, for every $d \in \mathfrak{D}$ let $(\mathcal{F}_d^\theta)_{\theta \in \Theta}$ be independent sub-σ-algebras of $\mathcal{F}$, for every $d \in \mathfrak{D}$ let $X_d^\theta = \big(X_d^{\theta,x,a}(\omega)\big)_{x \in \mathbb{X}_d, a \in A, \omega \in \Omega} \colon \mathbb{X}_d \times A \times \Omega \to \mathbb{X}_d$, $\theta \in \Theta$, be i.i.d. random fields which satisfy for all $d \in \mathfrak{D}$, $\theta \in \Theta$ that X_d^θ is $(\mathcal{X}_d \otimes 2^A \otimes \mathcal{F}_d^\theta)/\mathcal{X}_d$-measurable, assume $M > \frac{(1+(\sup_{d \in \mathfrak{D}} \lambda_d \delta_d)(2|A|-1))^2}{(1-(\sup_{d \in \mathfrak{D}} \lambda_d \delta_d))^2}$, assume for all $d \in \mathfrak{D}$, $(x,a) \in \mathbb{X}_d \times A$ that $\max_{b \in A} |g_d(x,b)| \leq \kappa \mathfrak{w}_d(x)$, $\big(\mathbb{E}\big[\big|\mathfrak{w}_d(X_d^{0,x,a})\big|^2 \big]\big)^{\frac{1}{2}} \leq \lambda_d \mathfrak{w}_d(x)$, $\sup_{u \in \mathfrak{D}} \lambda_u \delta_u < 1$, for every $d \in \mathfrak{D}$ let $Q_{n,d}^\theta \colon \mathbb{X}_d \times A \times \Omega \to \mathbb{R}$, $\theta \in \Theta$, $n \in \mathbb{N}_0$, satisfy for all $n \in \mathbb{N}_0$, $(x,a) \in \mathbb{X}_d \times A$, $\theta \in \Theta$ that*

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

and let $\mathfrak{C}_{n,d} \in [0,\infty)$, $n \in \mathbb{N}_0$, $d \in \mathfrak{D}$, satisfy for all $n \in \mathbb{N}_0$, $d \in \mathfrak{D}$ that

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

Then the following holds:

(i) For every $d \in \mathfrak{D}$ there exists a unique function $Q_d \colon \mathbb{X}_d \times A \to \mathbb{R}$ which is $(\mathcal{X}_d \otimes 2^A)/\mathcal{B}(\mathbb{R})$-measurable and satisfies for all $(x,a) \in \mathbb{X}_d \times A$ that $\sup_{y \in \mathbb{X}_d} \frac{\max_{b \in A} |Q_d(y,b)|}{\mathfrak{w}_d(y)} < \infty$, $\mathbb{E}\big[\big| \max_{b \in A} Q_d(X_d^{0,x,a}, b)\big|\big] < \infty$, and

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

(ii) *There exist* $N\colon (0,1] \to \mathbb{N}$ *and* $c \in \mathbb{R}$ *such that for all* $d \in \mathfrak{D}$, $\varepsilon \in (0,1]$ *it holds that* $\mathfrak{C}_{N_\varepsilon,d} \leq c\mathfrak{R}_d \varepsilon^{-c}$ *and*

$$\sup_{x \in \mathbb{X}_d} \left(\frac{\mathbb{E}\left[\max_{a \in A} |Q_d(x,a) - Q^0_{N_\varepsilon,d}(x,a)|^2 \right]}{|\mathfrak{w}_d(x)|^2} \right)^{\frac{1}{2}} \leq \varepsilon. \tag{99}$$

Proof of Corollary 4.2. Note that for every $d \in \mathfrak{D}$ it holds that the function $\mathbb{X}_d \times \mathbb{R}^A \ni (x,r) \mapsto \delta_d \max_{a \in A} \{g_d(x,a) + r(a)\} \in \mathbb{R}$ is $(\mathcal{X}_d \otimes (\bigotimes_{a \in A} \mathcal{B}(\mathbb{R})))/\mathcal{B}(\mathbb{R})$-measurable. Moreover, for all $d \in \mathfrak{D}$, $x \in \mathbb{X}_d$, $r, s \in \mathbb{R}^A$ it holds that

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

The assumption that for all $d \in \mathfrak{D}$, $(x,a) \in \mathbb{X}_d \times A$ it holds that $\max_{b \in A} |g_d(x,b)| \leq \kappa \mathfrak{w}_d(x)$, $(\mathbb{E}[|\mathfrak{w}_d(X_d^{0,x,a})|^2])^{\frac{1}{2}} \leq \lambda_d \mathfrak{w}_d(x)$, and $\sup_{u \in \mathfrak{D}} \lambda_u \delta_u < 1$ yields for all $d \in \mathfrak{D}$, $(x,a) \in \mathbb{X}_d \times A$ that

$$\left(\mathbb{E}\left[\left| \delta_d \max_{b \in A} \{ g_d(X_d^{0,x,a}, b) \} \right|^2 \right] \right)^{\frac{1}{2}} \leq \delta_d \kappa \left(\mathbb{E}\left[|\mathfrak{w}_d(X_d^{0,x,a})|^2 \right] \right)^{\frac{1}{2}} \leq \delta_d \kappa \lambda_d \mathfrak{w}_d(x) \leq \kappa \mathfrak{w}_d(x). \tag{101}$$

For every $d \in \mathfrak{D}$ let $R^\theta_{n,d}\colon \mathbb{X}_d \times A \to \mathbb{R}$, $\theta \in \Theta$, $n \in \mathbb{N}_0$, satisfy for all $n \in \mathbb{N}_0$, $(x,a) \in \mathbb{X}_d \times A$, $\theta \in \Theta$ that

$$R^\theta_{n,d}(x,a) = \sum_{l=0}^{n-1} \frac{\delta_d}{M^{n-l}} \sum_{i=1}^{M^{n-l}} \max_{b \in A} \left\{ g_d(X_d^{(\theta,l,i),x,a}, b) + R^{(\theta,l,i)}_{l,d}(X_d^{(\theta,l,i),x,a}, b) \right\}$$
$$- \mathbb{1}_{\mathbb{N}}(l) \max_{b \in A} \left\{ g_d(X_d^{(\theta,l,i),x,a}, b) + R^{(\theta,-l,i)}_{\max\{l-1,0\},d}(X_d^{(\theta,l,i),x,a}, b) \right\}. \tag{102}$$

This and Theorem 4.1 (applied with $M \leftarrow M$, $\Theta \leftarrow \Theta$, $A \leftarrow A$, $\kappa \leftarrow \kappa$, $(\Omega,\mathcal{F},\mathbb{P}) \leftarrow (\Omega,\mathcal{F},\mathbb{P})$, $\mathfrak{D} \leftarrow \mathfrak{D}$, $(\lambda_d)_{d \in \mathfrak{D}} \leftarrow (\lambda_d)_{d \in \mathfrak{D}}$, $(L_d)_{d \in \mathfrak{D}} \leftarrow (\delta_d)_{d \in \mathfrak{D}}$, $(\mathfrak{R}_d)_{d \in \mathfrak{D}} \leftarrow (\mathfrak{R}_d)_{d \in \mathfrak{D}}$, $(\mathbb{X}_d, \mathcal{X}_d)_{d \in \mathfrak{D}} \leftarrow (\mathbb{X}_d, \mathcal{X}_d)_{d \in \mathfrak{D}}$, $(\mathfrak{w}_d)_{d \in \mathfrak{D}} \leftarrow (\mathfrak{w}_d)_{d \in \mathfrak{D}}$, $(f_d)_{d \in \mathfrak{D}} \leftarrow (\mathbb{X}_d \times \mathbb{R}^A \ni (x,r) \mapsto \delta_d \max_{a \in A} \{g_d(x,a) + r(a)\} \in \mathbb{R})_{d \in \mathfrak{D}}$, $(\mathcal{F}^\theta_d)_{d \in \mathfrak{D}, \theta \in \Theta} \leftarrow (\mathcal{F}^\theta_d)_{d \in \mathfrak{D}, \theta \in \Theta}$, $(X^\theta_d)_{d \in \mathfrak{D}, \theta \in \Theta} \leftarrow (X^\theta_d)_{d \in \mathfrak{D}, \theta \in \Theta}$, $(\mathfrak{C}_{n,d})_{n \in \mathbb{N}_0, d \in \mathfrak{D}} \leftarrow (\mathfrak{C}_{n,d})_{n \in \mathbb{N}_0, d \in \mathfrak{D}}$, $(V^\theta_{n,d})_{n \in \mathbb{N}_0, d \in \mathfrak{D}, \theta \in \Theta} \leftarrow (R^\theta_{n,d})_{n \in \mathbb{N}_0, d \in \mathfrak{D}, \theta \in \Theta}$ in the notation of Theorem 4.1) yield the that

(1) for every $d \in \mathfrak{D}$ there exists a unique function $R_d\colon \mathbb{X}_d \times A \to \mathbb{R}$ which is $(\mathcal{X}_d \otimes 2^A)/\mathcal{B}(\mathbb{R})$-measurable and satisfies for all $(x,a) \in \mathbb{X}_d \times A$ that $\sup_{y \in \mathbb{X}_d} \frac{\max_{b \in A} |R_d(y,b)|}{\mathfrak{w}_d(y)} < \infty$, $\mathbb{E}[|\max_{b \in A} \{ g_d(X_d^{0,x,a}, b) + R_d(X_d^{0,x,a}, b) \}|] < \infty$,

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

and

(2) there exist $N\colon (0,1] \to \mathbb{N}$ and $c \in \mathbb{R}$ such that for all $d \in \mathfrak{D}$, $\varepsilon \in (0,1]$ it holds that $\mathfrak{C}_{N_\varepsilon,d} \leq c\mathfrak{R}_d \varepsilon^{-c}$ and

$$\sup_{x \in \mathbb{X}_d} \left(\frac{\mathbb{E}\left[\max_{a \in A} |R_d(x,a) - R^0_{N_\varepsilon,d}(x,a)|^2 \right]}{|\mathfrak{w}_d(x)|^2} \right)^{\frac{1}{2}} \leq \varepsilon. \tag{104}$$

For every $d \in \mathfrak{D}$ let $Q_d \colon \mathbb{X}_d \times A \to \mathbb{R}$ satisfy for all $(x, a) \in \mathbb{X}_d \times A$ that $Q_d(x, a) = g_d(x, a) + R_d(x, a)$. This and item (1) ensure for all $d \in \mathfrak{D}$, $(x, a) \in \mathbb{X}_d \times A$ that

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

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

Moreover, note that the assumption that for all $d \in \mathfrak{D}$, $x \in \mathbb{X}_d$ it holds that $\max_{a \in A}|g_d(x, a)| \le \kappa \mathfrak{w}_d(x)$ and item (1) demonstrate that for all $d \in \mathfrak{D}$, $(x, a) \in \mathbb{X}_d \times A$ it holds that $\sup_{y \in \mathbb{X}_d} \frac{\max_{b \in A}|Q_d(y,b)|}{\mathfrak{w}_d(y)} < \infty$ and $\mathbb{E}\left[|\max_{b \in A}\{Q_d(X_d^{0,x,a}, b)\}|\right] < \infty$. Furthermore, for every $d \in \mathfrak{D}$ let $S_d \colon \mathbb{X}_d \times A \to \mathbb{R}$ be $(\mathcal{X}_d \otimes 2^A)/\mathcal{B}(\mathbb{R})$-measurable and satisfy for all $d \in \mathfrak{D}$, $(x, a) \in \mathbb{X}_d \times A$ that $\sup_{y \in \mathbb{X}_d} \frac{\max_{b \in A}|S_d(y,b)|}{\mathfrak{w}_d(y)} < \infty$, $\mathbb{E}\left[|\max_{b \in A}\{S_d(X_d^{0,x,a}, b)\}|\right] < \infty$, and $S_d(x, a) = g_d(x, a) + \delta_d \mathbb{E}\left[\max_{b \in A}\{S_d(X_d^{0,x,a}, b)\}\right]$. It holds for all $d \in \mathfrak{D}$, $(x, a) \in \mathbb{X}_d \times A$ that

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

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

Item (1) implies that for all $d \in \mathfrak{D}$ it holds that $S_d - g_d = R_d$. This establishes for all $d \in \mathfrak{D}$ that $S_d = Q_d$. This proves item (i). Note that induction and (102) demonstrate that for all $d \in \mathfrak{D}$, $n \in \mathbb{N}_0$, $(x, a) \in \mathbb{X}_d \times A$, $\theta \in \Theta$ it holds that $Q_{n,d}^\theta(x, a) = g_d(x, a) + R_{n,d}^\theta(x, a)$. Combining this and item (2) implies for all $d \in \mathfrak{D}$, $\varepsilon \in (0, 1]$ that

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

This proves item (ii). The proof of Corollary 4.2 is thus completed. $\qquad\square$

4.3 MLFP approximations for Bellman equations of optimal stopping problems

Corollary 4.3. *Let $M \in \mathbb{N}$, let $\Theta = \bigcup_{n \in \mathbb{N}} \mathbb{Z}^n$, let $(\Omega, \mathcal{F}, \mathbb{P})$ be a probability space, let $\mathfrak{D}$ be a nonempty set, let $\delta_d, \mathfrak{R}_d \in [0, \infty)$, $d \in \mathfrak{D}$, let $(\mathbb{X}_d, \mathcal{X}_d)$, $d \in \mathfrak{D}$, be nonempty Borel spaces, for every $d \in \mathfrak{D}$ let $g_d \colon \mathbb{X}_d \to \mathbb{R}$ and $G_d \colon \mathbb{X}_d \to \mathbb{R}$ be $\mathcal{X}_d/\mathcal{B}(\mathbb{R})$-measurable, for every $d \in \mathfrak{D}$ let $(\mathcal{F}_d^\theta)_{\theta \in \Theta}$ be independent sub-σ-algebras of $\mathcal{F}$, for every $d \in \mathfrak{D}$ let $X_d^\theta = (X_d^{\theta,x}(\omega))_{x \in \mathbb{X}_d, \omega \in \Omega} \colon \mathbb{X}_d \times \Omega \to \mathbb{X}_d$, $\theta \in \Theta$, be i.i.d. random fields which satisfy for all $d \in \mathfrak{D}$, $\theta \in \Theta$ that X_d^θ is $(\mathcal{X}_d \otimes \mathcal{F}_d^\theta)/\mathcal{X}_d$-measurable, assume $\sup_{d \in \mathfrak{D}} \delta_d < 1$, assume $M > \frac{(1+3(\sup_{d \in \mathfrak{D}} \delta_d))^2}{(1 - (\sup_{d \in \mathfrak{D}} \delta_d))^2}$, assume $\sup_{d \in \mathfrak{D}}\left(\sup_{x \in \mathbb{X}_d}|g_d(x)| + |G_d(x)|\right) < \infty$, for every $d \in \mathfrak{D}$ let $Q_{n,d}^\theta \colon \mathbb{X}_d \times \Omega \to \mathbb{R}$, $n \in \mathbb{N}_0$, $\theta \in \Theta$, satisfy for all $n \in \mathbb{N}_0$, $x \in \mathbb{X}_d$, $\theta \in \Theta$ that*

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

$$- \mathbb{1}_{\mathbb{N}}(l) \max\left\{G_d\big(X_d^{(\theta,l,i),x}\big), Q_{\max\{l-1,0\},d}^{(\theta,-l,i)}\big(X_d^{(\theta,l,i),x}\big)\right\}$$

and let $\mathfrak{C}_{n,d} \in [0, \infty)$, $n \in \mathbb{N}_0$, $d \in \mathfrak{D}$, satisfy for all $n \in \mathbb{N}_0$, $d \in \mathfrak{D}$ that

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

Then the following holds:

(i) *For every $d \in \mathfrak{D}$ there exists a unique function $Q_d \colon \mathbb{X}_d \to \mathbb{R}$ which is $\mathcal{X}_d/\mathcal{B}(\mathbb{R})$-measurable and satisfies for all $x \in \mathbb{X}_d$ that $\sup_{y \in \mathbb{X}_d} |Q_d(y)| < \infty$, and*

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

(ii) *There exist $N \colon (0,1] \to \mathbb{N}$ and $c \in \mathbb{R}$ such that for all $d \in \mathfrak{D}$, $\varepsilon \in (0,1]$ it holds that $\mathfrak{C}_{N_\varepsilon,d} \leq c\mathfrak{R}_d\varepsilon^{-c}$ and*

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

Proof of Corollary 4.3. Let $\mathbb{Y}_d$, $d \in \mathfrak{D}$, be nonempty sets which satisfy that there exists $\Upsilon \in \bigcap_{d \in \mathfrak{D}} \mathbb{Y}_d$, and that for all $d \in \mathfrak{D}$ it holds that $\mathbb{Y}_d \setminus \{\Upsilon\} = \mathbb{X}_d$, for every $d \in \mathfrak{D}$ let $\mathcal{Y}_d = \sigma_{\mathbb{Y}_d}(\mathcal{X}_d)$, for every $d \in \mathfrak{D}$, $\theta \in \Theta$ let $Y_d^\theta = (Y_d^{\theta,y,a}(\omega))_{y \in \mathbb{Y}_d,\, a \in \{0,1\},\, \omega \in \Omega} \colon \mathbb{Y}_d \times \{0,1\} \times \Omega \to \mathbb{Y}_d$ satisfy for all $d \in \mathfrak{D}$, $\theta \in \Theta$, $(y,a) \in \mathbb{Y}_d \times \{0,1\}$, $\omega \in \Omega$ that

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

Note that for all $d \in \mathfrak{D}$ it holds that Y_d^θ, $\theta \in \Theta$, are i.i.d. random fields and that for every $d \in \mathfrak{D}$, $\theta \in \Theta$ it holds that Y_d^θ is $(\mathcal{Y}_d \otimes 2^{\{0,1\}} \otimes \mathcal{F}_d^\theta)/\mathcal{Y}_d$-measurable. For every $d \in \mathfrak{D}$ let $h_d \colon \mathbb{Y}_d \times \{0,1\} \to \mathbb{R}$ satisfy for all $d \in \mathfrak{D}$, $(y,a) \in \mathbb{Y}_d \times \{0,1\}$ that

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

Note that for all $d \in \mathfrak{D}$ it holds that h_d is $(\mathcal{Y}_d \otimes 2^{\{0,1\}})/\mathcal{B}(\mathbb{R})$-measurable. Moreover, for all $d \in \mathfrak{D}$, $(y,a) \in \mathbb{X}_d \times \{0,1\}$ it holds that $|h_d(y,a)| \leq \max\{|g_d(y)|, |G_d(y)|\} \leq \sup_{u \in \mathfrak{D}}\left(\sup_{x \in \mathbb{X}_d} |g_u(x)| + |G_u(x)|\right)$. For every $d \in \mathfrak{D}$ let $q_{n,d}^\theta \colon \mathbb{Y}_d \times \{0,1\} \times \Omega \to \mathbb{Y}_d$, $\theta \in \Theta$, $n \in \mathbb{N}_0$, satisfy for all $n \in \mathbb{N}_0$, $\theta \in \Theta$, $(y,a) \in \mathbb{Y}_d \times \{0,1\}$ that

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

Corollary 4.2 (applied with $A \leftarrow \{0,1\}$, $\kappa \leftarrow \sup_{d \in \mathfrak{D}}\left(\sup_{x \in \mathbb{X}_d} |g_d(x)| + |G_d(x)|\right)$, $g_d \leftarrow h_d$, $(\mathbb{X}_d, \mathcal{X}_d) \leftarrow (\mathbb{Y}_d, \mathcal{Y}_d)$, $Q_{n,d}^\theta \leftarrow q_{n,d}^\theta$ for $d \in \mathfrak{D}$ in the notation of Corollary 4.2) establishes the following:

(1) For every $d \in \mathfrak{D}$ there exists a unique function $q_d \colon \mathbb{Y}_d \times \{0,1\} \to \mathbb{R}$ which is $(\mathcal{Y}_d \otimes 2^{\{0,1\}})/\mathcal{B}(\mathbb{R})$-measurable and satisfies for all $(y,a) \in \mathbb{Y}_d \times \{0,1\}$ that $\sup_{(z,b) \in \mathbb{Y}_d \times \{0,1\}} |q_d(z,b)| < \infty$ and

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

28

(2) There exist $N\colon (0,1] \to \mathbb{N}$ and $c \in \mathbb{R}$ such that for all $d \in \mathfrak{D}$, $\varepsilon \in (0,1]$ it holds that $\mathfrak{C}_{N_\varepsilon,d} \leq c\mathfrak{R}_d\varepsilon^c$ and

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

Note that for all $d \in \mathfrak{D}$, $(y,a) \in \mathbb{Y}_d \backslash \{\Upsilon\} \times \{0,1\}$ it holds that $q_d(\Upsilon, a) = 0$ and $q_d(y, 0) = G_d(y)$. Moreover, for all $d \in \mathfrak{D}$, $y \in \mathbb{Y}_d \setminus \{\Upsilon\}$ it holds that

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

This and item (1) demonstrate for all $d \in \mathfrak{D}$ that there exists a unique function $Q_d\colon \mathbb{X}_d \to \mathbb{R}$ which is $\mathcal{X}_d/\mathcal{B}(\mathbb{R})$-measurable and satisfies for all $x \in \mathbb{X}_d$ that $\sup_{y \in \mathbb{X}_d} |Q_d(y)| < \infty$ and

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

This proves item (i). Furthermore, note that for all $d \in \mathfrak{D}$, $\theta \in \Theta$, $(y,a) \in \mathbb{Y}_d \setminus \{\Upsilon\} \times \{0,1\}$ it holds $\mathbb{P}$-a.s. that $q^\theta_{0,d}(\Upsilon, a) = 0$, $q^\theta_{0,d}(y, 0) = G_d(y)$, and $q^\theta_{0,d}(y, 1) = g_d(y)$. This and induction yield that for all $d \in \mathfrak{D}$, $\theta \in \Theta$, $n \in \mathbb{N}_0$, $(y,a) \in \mathbb{Y}_d\backslash\{\Upsilon\}\times\{0,1\}$ it holds $\mathbb{P}$-a.s. that $q^\theta_{n,d}(\Upsilon, a) = 0$ and $q^\theta_{n,d}(y, 0) = G_d(y)$. Combining this, (112), (113), and (114) demonstrates that for all $d \in \mathfrak{D}$, $\theta \in \Theta$, $n \in \mathbb{N}_0$, $y \in \mathbb{Y}_d \setminus \{\Upsilon\} = \mathbb{X}_d$ it holds $\mathbb{P}$-a.s. that

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

This and (108) yield for all $d \in \mathfrak{D}$, $\theta \in \Theta$, $n \in \mathbb{N}_0$, $y \in \mathbb{Y}_d \setminus \{\Upsilon\} = \mathbb{X}_d$ it holds $\mathbb{P}$-a.s. that

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

Combining this and (116) implies that for all $d \in \mathfrak{D}$, $\varepsilon \in (0,1]$ it holds that

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

This establishes item (ii). The proof of Corollary 4.3 is thus completed. $\qquad\square$

Acknowledgements

References

[1] ANDERSEN, L., AND BROADIE, M. Primal-dual simulation algorithm for pricing multidimensional American options. *Management Science 50*, 9 (2004), 1222–1234.

[2] ARULKUMARAN, K., DEISENROTH, M. P., BRUNDAGE, M., AND BHARATH, A. A. Deep reinforcement learning: A brief survey. *IEEE Signal Processing Magazine 34*, 6 (2017), 26–38.

[3] BECK, C., GONON, L., AND JENTZEN, A. Overcoming the curse of dimensionality in the numerical approximation of high-dimensional semilinear elliptic partial differential equations. *arXiv:2003.00596* (2020).

[4] BECK, C., HORNUNG, F., HUTZENTHALER, M., JENTZEN, A., AND KRUSE, T. Overcoming the curse of dimensionality in the numerical approximation of Allen-Cahn partial differential equations via truncated full-history recursive multilevel Picard approximations. *Journal of Numerical Mathematics 28*, 4 (2020), 197–222.

[5] BECK, C., JENTZEN, A., AND KRUSE, T. Nonlinear Monte Carlo methods with polynomial runtime for high-dimensional iterated nested expectations. *arXiv:2009.13989* (2020).

[6] BECKER, S., BRAUNWARTH, R., HUTZENTHALER, M., JENTZEN, A., AND VON WURSTEMBERGER, P. Numerical simulations for full history recursive multilevel Picard approximations for systems of high-dimensional partial differential equations. *Communications in Computational Physics 28*, 5 (2020), 2109–2138.

[7] BECKER, S., CHERIDITO, P., AND JENTZEN, A. Deep optimal stopping. *Journal of Machine Learning Research 20* (2019), 74.

[8] BECKER, S., CHERIDITO, P., JENTZEN, A., AND WELTI, T. Solving high-dimensional optimal stopping problems using deep learning. *European Journal of Applied Mathematics 32*, 3 (2021), 470–514.

[9] BELLMAN, R. *Dynamic programming*. Princeton Landmarks in Mathematics. Princeton University Press, Princeton, NJ, 2010. Reprint of the 1957 edition, With a new introduction by Stuart Dreyfus.

[10] BELOMESTNY, D., BENDER, C., AND SCHOENMAKERS, J. True upper bounds for Bermudan products via non-nested Monte Carlo. *Mathematical Finance: An International Journal of Mathematics, Statistics and Financial Economics 19*, 1 (2009), 53–71.

[11] BERTSEKAS, D. *Reinforcement learning and optimal control*. Athena Scientific, 2019.

[12] BERTSEKAS, D., AND SHREVE, S. E. *Stochastic optimal control: the discrete-time case*, vol. 5. Athena Scientific, 1996.

[13] BERTSEKAS, D., AND TSITSIKLIS, J. N. *Neuro-dynamic programming*. Athena Scientific, 1996.

[14] BRAY, R. L. A comment on "Using randomization to break the curse of dimensionality". *Econometrica 90*, 4 (2022), 1915–1929.

[15] BROADIE, M., AND GLASSERMAN, P. Pricing American-style securities using simulation. *Journal of economic dynamics and control 21*, 8-9 (1997), 1323–1352.

[16] Broadie, M., and Glasserman, P. A stochastic mesh method for pricing high-dimensional American options. *Journal of Computational Finance 7* (2004), 35–72.

[17] Chow, C.-S., and Tsitsiklis, J. N. The complexity of dynamic programming. *Journal of complexity 5*, 4 (1989), 466–488.

[18] Chow, C.-S., and Tsitsiklis, J. N. An optimal one-way multigrid algorithm for discrete-time stochastic control. *IEEE transactions on automatic control 36*, 8 (1991), 898–914.

[19] Desai, V. V., Farias, V. F., and Moallemi, C. C. Pathwise optimization for optimal stopping problems. *Management Science 58*, 12 (2012), 2292–2308.

[20] E, W., Hutzenthaler, M., Jentzen, A., and Kruse, T. On multilevel Picard numerical approximations for high-dimensional nonlinear parabolic partial differential equations and high-dimensional nonlinear backward stochastic differential equations. *Journal of Scientific Computing 79*, 3 (2019), 1534–1571.

[21] E, W., Hutzenthaler, M., Jentzen, A., and Kruse, T. Multilevel picard iterations for solving smooth semilinear parabolic heat equations. *Partial Differential Equations and Applications 2*, 6 (2021), 80.

[22] François-Lavet, V., Henderson, P., Islam, R., Bellemare, M. G., and Pineau, J. An introduction to deep reinforcement learning. *Foundations and Trends® in Machine Learning 11*, 3-4 (2018), 219–354.

[23] Giles, M. Improved multilevel Monte Carlo convergence using the Milstein scheme. In *Monte Carlo and Quasi-Monte Carlo Methods 2006*. Springer, 2008, pp. 343–358.

[24] Giles, M. B., Jentzen, A., and Welti, T. Generalised multilevel Picard approximations. *arXiv:1911.03188* (2019).

[25] Gonon, L. Deep neural network expressivity for optimal stopping problems. *arXiv:2210.10443* (2022).

[26] Haugh, M. B., and Kogan, L. Pricing American options: a duality approach. *Operations Research 52*, 2 (2004), 258–270.

[27] Heinrich, S. Monte Carlo complexity of global solution of integral equations. *Journal of Complexity 14*, 2 (1998), 151–175.

[28] Heinrich, S. Multilevel Monte Carlo methods. In *International Conference on Large-Scale Scientific Computing* (2001), Springer, pp. 58–67.

[29] Hutzenthaler, M., Jentzen, A., and Kruse, T. Overcoming the curse of dimensionality in the numerical approximation of parabolic partial differential equations with gradient-dependent nonlinearities. *Found. Comput. Math.* (2021), 1–62.

[30] Hutzenthaler, M., Jentzen, A., Kruse, T., and Nguyen, T. A. Multilevel Picard approximations for high-dimensional semilinear second-order PDEs with Lipschitz nonlinearities. *arXiv:2009.02484* (2020).

[31] Hutzenthaler, M., Jentzen, A., Kruse, T., Nguyen, T. A., and von Wurstemberger, P. Overcoming the curse of dimensionality in the numerical approximation of semilinear parabolic partial differential equations. *Proceedings of the Royal Society A 476*, 2244 (2020), 20190630.

[32] Hutzenthaler, M., Jentzen, A., and von Wurstemberger, P. Overcoming the curse of dimensionality in the approximative pricing of financial derivatives with default risks. *Electron. J. Probab. 25* (2020), Paper No. 101, 73.

[33] Hutzenthaler, M., and Kruse, T. Multilevel Picard approximations of high-dimensional semilinear parabolic differential equations with gradient-dependent nonlinearities. *SIAM Journal on Numerical Analysis 58*, 2 (2020), 929–961.

[34] Hutzenthaler, M., Kruse, T., and Nguyen, T. A. Multilevel Picard approximations for McKean-Vlasov stochastic differential equations. *Journal of Mathematical Analysis and Applications 507*, 1 (2022), 125761.

[35] Klenke, A. *Probability Theory*, 2 ed. Springer London, 2014.

[36] Kristensen, D., Mogensen, P. K., Moon, J. M., and Schjerning, B. Solving dynamic discrete choice models using smoothing and sieve methods. *Journal of Econometrics 223*, 2 (2021), 328–360.

[37] Li, Y. Deep reinforcement learning: An overview. *arXiv:1701.07274* (2017).

[38] Longstaff, F. A., and Schwartz, E. S. Valuing American options by simulation: a simple least-squares approach. *The review of financial studies 14*, 1 (2001), 113–147.

[39] Novak, E., and Ritter, K. The curse of dimension and a universal method for numerical integration. In *Multivariate approximation and splines (Mannheim, 1996)*, vol. 125 of *Internat. Ser. Numer. Math.* Birkhäuser, Basel, 1997, pp. 177–187.

[40] Novak, E., and Woźniakowski, H. *Tractability of multivariate problems. Vol. 1: Linear information*, vol. 6 of *EMS Tracts in Mathematics*. European Mathematical Society (EMS), Zürich, 2008.

[41] Powell, W. B. *Approximate Dynamic Programming: Solving the curses of dimensionality*, vol. 703. John Wiley & Sons, 2007.

[42] Puterman, M. L. *Markov decision processes: discrete stochastic dynamic programming*. John Wiley & Sons, 2014.

[43] Rogers, L. C. Monte Carlo valuation of American options. *Mathematical Finance 12*, 3 (2002), 271–286.

[44] Rummery, G. A., and Niranjan, M. *On-line Q-learning using connectionist systems*, vol. 37. University of Cambridge, Department of Engineering Cambridge, UK, 1994.

[45] Rust, J. Using randomization to break the curse of dimensionality. *Econometrica: Journal of the Econometric Society* (1997), 487–516.

[46] Rust, J. P. A comparison of policy iteration methods for solving continuous-state, infinite-horizon Markovian decision problems using random, quasi-random, and deterministic discretizations. *Available at SSRN: http://dx.doi.org/10.2139/ssrn.37768* (1997).

[47] SUTTON, R. S. Learning to predict by the methods of temporal differences. *Machine learning 3* (1988), 9–44.

[48] SUTTON, R. S., AND BARTO, A. G. *Reinforcement learning: An introduction.* MIT press, 2018.

[49] TSITSIKLIS, J. N., AND VAN ROY, B. Optimal stopping of Markov processes: Hilbert space theory, approximation algorithms, and an application to pricing high-dimensional financial derivatives. *IEEE Transactions on Automatic Control 44*, 10 (1999), 1840–1851.

[50] TSITSIKLIS, J. N., AND VAN ROY, B. Regression methods for pricing complex American-style options. *IEEE Transactions on Neural Networks 12*, 4 (2001), 694–703.

[51] WATKINS, C. Learning from delayed rewards. *Ph. D. thesis. University of Cambridge* (1989).

[52] WATKINS, C., AND DAYAN, P. Q-learning. *Machine learning 8*, 3-4 (1992), 279–292.