Empirical portfolio selection strategies with proportional transaction costs

László Györfi, *Fellow, IEEE* Harro Walk

***Abstract*— Discrete time growth optimal investment in stock markets with proportional transactions costs is considered. The market process is modeled by a first order Markov process. Not assuming that the distribution of the market process is known, we show empirical investment strategies such that, in the long run, the growth rate on trajectories achieves the maximum with probability 1.**

***Index Terms*— portfolio selection, log-optimal investment, proportional transaction cost, dynamic optimization.**

I. Introduction

The purpose of this paper is to investigate sequential investment strategies for financial markets such that the strategies are allowed to use information collected from the past of the market and determine, at the beginning of a trading period, a portfolio, that is, a way to distribute their current capital among the available assets. The goal of the investor is to maximize his wealth on the long run. If there is no transaction cost then under the only assumption that the daily price relatives form a stationary and ergodic process the best strategy (called log-optimum strategy) can be constructed in full knowledge of the distribution of the entire process, see Algoet and Cover [1]. Moreover, Györfi and Schäfer [11], Györfi, Lugosi and Udina [10] and Györfi, Udina and Walk [13] constructed empirical (data driven) growth optimum strategies in case of unknown distributions. The empirical results show that the performance of these empirical investment strategies measured on past NYSE data is solid, and sometimes even spectacular.

Papers dealing with growth optimal investment with transaction costs in discrete time setting are seldom. Iyengar and Cover [22] formulated the problem of horse race markets, where in every market period one of the assets has positive pay off and all the others pay nothing. Their model included proportional transaction costs and they used a long run expected average reward criterion. There are results for more general markets as well. Sass and Schäl [27] investigated the numeraire portfolio in context of bond and stock as assets. Iyengar [20], [21] investigated growth optimal investment with several assets assuming independent and identically distributed (i.i.d.) sequence of asset returns. Bobryk and Stettner [4] considered the case of portfolio selection with consumption, when there are two assets, a bond and a stock. Furthermore, long run expected discounted reward and i.i.d asset returns were assumed.

Knowing the distribution of the market process, in the case of discrete time and finite order stationary Markov market process Schäfer [28] considered the maximization of the long run expected average growth rate with several assets and proportional transaction costs. Györfi and Vajda [14], and Györfi and Walk [15] extended the expected growth optimality mentioned above to the almost sure (a.s.) growth optimality.

This paper considers long term optimal trading strategies on Markovian markets when proportional transactions costs are to be paid after each buy or sell operation. The main result of the paper is two constructions of purely empirical strategies that achieve the best possible rate of growth of net capital of the investor when the market behaves as a stationary Markov process satisfying some mild regularity conditions. For the first trading strategy, the asymptotic optimality is proved if the state space of the relative prices is finite (Theorem 1). For a modification of this strategy, it is possible to extend the optimality to general state space (Theorem 2).

II. Mathematical setup: investment with transaction cost

Consider a market consisting of d assets. The evolution of the market in time is represented by a sequence of market vectors $\mathbf{s}_1, \mathbf{s}_2, \ldots \in \mathbb{R}_+^d$, where

$$\mathbf{s}_i = (s_i^{(1)}, \ldots, s_i^{(d)})$$

such that the j-th component $s_i^{(j)}$ of $\mathbf{s}_i$ denotes the price of the j-th asset at the end of the i-th trading period. ($s_0^{(j)} = 1$.)

In order to apply the usual prediction techniques for time series analysis one has to transform the sequence $\{\mathbf{s}_i\}$ into a sequence of return vectors $\{\mathbf{x}_i\}$ as follows:

$$\mathbf{x}_i = (x_i^{(1)}, \ldots, x_i^{(d)})$$

such that

$$x_i^{(j)} = \frac{s_i^{(j)}}{s_{i-1}^{(j)}}.$$

Thus, the j-th component $x_i^{(j)}$ of the return vector $\mathbf{x}_i$ denotes the amount obtained at the end of the i-th trading period after investing a unit capital in the j-th asset.

The investor is allowed to diversify his capital at the beginning of each trading period according to a portfolio vector $\mathbf{b} = (b^{(1)}, \ldots b^{(d)})^T$. The j-th component $b^{(j)}$ of $\mathbf{b}$ denotes the proportion of the investor's capital invested in asset j. Throughout the paper we assume that the portfolio vector $\mathbf{b}$ has nonnegative components with $\sum_{j=1}^d b^{(j)} = 1$. The fact that $\sum_{j=1}^d b^{(j)} = 1$ means that the investment strategy is self financing and consumption of capital is excluded. The non-negativity of the components of $\mathbf{b}$ means that short selling and buying stocks on margin are not permitted. To make the analysis feasible, some simplifying assumptions are used that need to be taken into account. We assume that assets are arbitrarily divisible and all assets are available in unbounded quantities at the current price at any given trading period. We also assume that the behavior of the market is not affected by the actions of the investor using the strategies under investigation.

For $j \le i$ we abbreviate by $\mathbf{x}_j^i$ the array of return vectors $(\mathbf{x}_j, \ldots, \mathbf{x}_i)$. Denote by Δ_d the simplex of all vectors $\mathbf{b} \in \mathbb{R}_+^d$ with nonnegative components summing up to one. An *investment strategy* is a sequence $\mathbf{B}$ of functions

$$\mathbf{b}_i : \left(\mathbb{R}_+^d\right)^{i-1} \to \Delta_d \ , \qquad i = 1, 2, \ldots$$

The work was supported in part by the Computer and Automation Research Institute of the Hungarian Academy of Sciences and by the PASCAL2 Network of Excellence under EC grant no. 216886.

L. Györfi is with Department of Computer Science and Information Theory, Budapest University of Technology and Economics, Magyar tudósok körútja 2., Budapest, Hungary, H-1117. (e-mail: gyorfi@szit.bme.hu).

H. Walk is with Department of Mathematics, Universität Stuttgart, Pfaffenwaldring 57, D-70569 Stuttgart, Germany. (e-mail: harro.walk@t-online.de)

so that $\mathbf{b}_i(\mathbf{x}_1^{i-1})$ denotes the portfolio vector chosen by the investor on the i-th trading period, upon observing the past behavior of the market. We write $\mathbf{b}(\mathbf{x}_1^{i-1}) = \mathbf{b}_i(\mathbf{x}_1^{i-1})$ to ease the notation.

The derivations in this paper can be extended to any compact set Δ_d. For example, one may allow short selling or leverage. Under the Condition (iii) below we can create no-ruin conditions, while for no transaction cost, the empirical results on NYSE data show that for short selling there is no gain and for leverage the increase of the growth rate is spectacular (cf. Horváth and Urbán [19]).

In this section our presentation of the transaction cost problem utilizes the formulation in Kalai and Blum [23] and Schäfer [28] and Györfi and Vajda [14]. Let S_n denote the gross wealth at the end of trading period n, $n = 0, 1, 2, \cdots$, i.e., it is the wealth before paying the transaction cost, while N_n stands for the net wealth at the end of trading period n, i.e., it is the wealth after paying the transaction cost. Without loss of generality we let the investor's initial capital S_0 be 1 dollar. Using the above notations, for the trading period n, the net wealth N_{n-1} can be invested according to the portfolio $\mathbf{b}_n$, therefore the gross wealth S_n at the end of trading period n is

$$S_n = N_{n-1} \sum_{j=1}^{d} b_n^{(j)} x_n^{(j)} = N_{n-1} \langle \mathbf{b}_n \, , \, \mathbf{x}_n \rangle \, ,$$

where $\langle \cdot \, , \, \cdot \rangle$ denotes inner product.

At the beginning of a new market period (day) $n + 1$, the investor sets up his new portfolio, i.e. buys/sells stocks according to the actual portfolio vector $\mathbf{b}_{n+1}$. During this rearrangement, he has to pay transaction cost, therefore at the beginning of a new market day $n+1$ the net wealth N_n in the portfolio $\mathbf{b}_{n+1}$ is less than S_n.

The rate of proportional transaction costs (commission factors) levied on one asset are denoted by $0 < c_s < 1$ and $0 < c_p < 1$, i.e., the sale of 1 dollar worth of asset i nets only $1 - c_s$ dollars, and similarly we take into account the purchase of an asset such that the purchase of 1 dollar's worth of asset i costs an extra c_p dollars. We consider the special case when the rate of costs is constant over the assets.

We describe the transaction cost to be paid when select the portfolio $\mathbf{b}_{n+1}$. Before rearranging the capitals, at the j-th asset there are $b_n^{(j)} x_n^{(j)} N_{n-1}$ dollars, while after rearranging the investor's wealth should be $b_{n+1}^{(j)} N_n$ dollars. If $b_n^{(j)} x_n^{(j)} N_{n-1} \geq b_{n+1}^{(j)} N_n$ then one has to sell and the transaction cost at the j-th asset is

$$c_s \left(b_n^{(j)} x_n^{(j)} N_{n-1} - b_{n+1}^{(j)} N_n \right),$$

otherwise one has to buy and the transaction cost at the j-th asset is

$$c_p \left(b_{n+1}^{(j)} N_n - b_n^{(j)} x_n^{(j)} N_{n-1} \right).$$

Let x^+ denote the positive part of x. Thus, the gross wealth S_n decomposes to the sum of the net wealth and cost in the following - self-financing - way

$$\begin{aligned} N_n = S_n \quad &- \quad \sum_{j=1}^{d} c_s \left(b_n^{(j)} x_n^{(j)} N_{n-1} - b_{n+1}^{(j)} N_n \right)^+ \\ &- \quad \sum_{j=1}^{d} c_p \left(b_{n+1}^{(j)} N_n - b_n^{(j)} x_n^{(j)} N_{n-1} \right)^+ , \end{aligned}$$

or equivalently

$$\begin{aligned} S_n = N_n \quad &+ \quad c_s \sum_{j=1}^{d} \left(b_n^{(j)} x_n^{(j)} N_{n-1} - b_{n+1}^{(j)} N_n \right)^+ \\ &+ \quad c_p \sum_{j=1}^{d} \left(b_{n+1}^{(j)} N_n - b_n^{(j)} x_n^{(j)} N_{n-1} \right)^+ . \end{aligned}$$

Dividing both sides by S_n and introducing ratio

$$w_n = \frac{N_n}{S_n},$$

$0 < w_n < 1$, we get

$$\begin{aligned} 1 = w_n \quad &+ \quad c_s \sum_{j=1}^{d} \left(\frac{b_n^{(j)} x_n^{(j)}}{\langle \mathbf{b}_n \, , \, \mathbf{x}_n \rangle} - b_{n+1}^{(j)} w_n \right)^+ \\ &+ \quad c_p \sum_{j=1}^{d} \left(b_{n+1}^{(j)} w_n - \frac{b_n^{(j)} x_n^{(j)}}{\langle \mathbf{b}_n \, , \, \mathbf{x}_n \rangle} \right)^+ . \end{aligned} \tag{1}$$

For given previous return vector $\mathbf{x}_n$ and portfolio vector $\mathbf{b}_n$, there is a portfolio vector $\tilde{\mathbf{b}}_{n+1} = \tilde{\mathbf{b}}_{n+1}(\mathbf{b}_n, \mathbf{x}_n)$ for which there is no trading:

$$\tilde{b}_{n+1}^{j} = \frac{b_n^{(j)} x_n^{(j)}}{\langle \mathbf{b}_n \, , \, \mathbf{x}_n \rangle} \tag{2}$$

such that there is no transaction cost, i.e., $w_n = 1$.

For arbitrary fixed portfolio vectors $\mathbf{b}_n$, $\mathbf{b}_{n+1}$, and return vector $\mathbf{x}_n$ there exists a unique cost factor $w_n \in [0, 1)$, i.e. the portfolio is self financing. The value of cost factor w_n at day n is determined by portfolio vectors $\mathbf{b}_n$ and $\mathbf{b}_{n+1}$ as well as by return vector $\mathbf{x}_n$, i.e.,

$$w_n = w(\mathbf{b}_n, \mathbf{b}_{n+1}, \mathbf{x}_n),$$

for some function w. If we want to rearrange our portfolio substantially, then our net wealth decreases more considerably, however, it remains positive. Note also, that the cost does not restrict the set of new portfolio vectors, i.e., the optimization algorithm searches for optimal vector $\mathbf{b}_{n+1}$ within the whole simplex Δ_d. The value of the cost factor ranges between

$$\frac{1 - c_s}{1 + c_p} \leq w_n \leq 1.$$

For the sake of simplicity we consider the special case of $c_s = c_p =: c$, while the general case can be treated in a similar manner. Then

$$\begin{aligned} c_s \left(b_n^{(j)} x_n^{(j)} N_{n-1} - b_{n+1}^{(j)} N_n \right)^+ + c_p \left(b_{n+1}^{(j)} N_n - b_n^{(j)} x_n^{(j)} N_{n-1} \right)^+ \\ = c \left| b_n^{(j)} x_n^{(j)} N_{n-1} - b_{n+1}^{(j)} N_n \right| . \end{aligned}$$

Starting with an initial wealth $S_0 = 1$ and $w_0 = 1$, wealth S_n at the closing time of the n-th market day becomes

$$\begin{aligned} S_n &= N_{n-1} \langle \mathbf{b}_n \, , \, \mathbf{x}_n \rangle \\ &= w_{n-1} S_{n-1} \langle \mathbf{b}_n \, , \, \mathbf{x}_n \rangle \\ &= \prod_{i=1}^{n} [w(\mathbf{b}_{i-1}, \mathbf{b}_i, \mathbf{x}_{i-1}) \langle \mathbf{b}_i \, , \, \mathbf{x}_i \rangle]. \end{aligned}$$

Introduce the notation

$$g(\mathbf{b}_{i-1}, \mathbf{b}_i, \mathbf{x}_{i-1}, \mathbf{x}_i) = \log(w(\mathbf{b}_{i-1}, \mathbf{b}_i, \mathbf{x}_{i-1}) \langle \mathbf{b}_i \, , \, \mathbf{x}_i \rangle),$$

then the average growth rate becomes

$$\begin{aligned}\frac{1}{n}\log S_n &= \frac{1}{n}\sum_{i=1}^{n}\log(w(\mathbf{b}_{i-1},\mathbf{b}_i,\mathbf{x}_{i-1})\langle\mathbf{b}_i\,,\,\mathbf{x}_i\rangle)\\ &= \frac{1}{n}\sum_{i=1}^{n} g(\mathbf{b}_{i-1},\mathbf{b}_i,\mathbf{x}_{i-1},\mathbf{x}_i). \qquad (3)\end{aligned}$$

Our aim is to maximize this average growth rate.

Farias et al. [5] considered a special averaged cost, where there is no memory in the portfolios:

$$\frac{1}{n}\sum_{i=1}^{n} g(\mathbf{b}_{i-1},\mathbf{x}_{i-1},\mathbf{x}_i).$$

Moreover, both the return vectors $\mathbf{x}_i$ and the portfolio vectors $\mathbf{b}_i$ may take finitely many values. However, in their scheme more generally the trading can influence the prices.

In the sequel $\mathbf{x}_i$ will be a realization of a random variable $\mathbf{X}_i$, and we assume the following
Conditions:

(i) $\{\mathbf{X}_i\}$ is a homogeneous and first order Markov process,

(ii) the Markov kernel is continuous, which means that for $\mu(H|\mathbf{x})$ being the Markov kernel defined by

$$\mu(H|\mathbf{x}) := \mathbb{P}\{\mathbf{X}_2 \in H \mid \mathbf{X}_1 = \mathbf{x}\}$$

we assume that the Markov kernel is continuous in total variation, i.e.,

$$V(\mathbf{x},\mathbf{x}') := \sup_{H\in\mathcal{H}} |\mu(H|\mathbf{x}) - \mu(H|\mathbf{x}')| \to 0$$

if $\mathbf{x}' \to \mathbf{x}$ such that $\mathcal{H}$ denotes the Borel σ-algebra, further

$$V(\mathbf{x},\mathbf{x}') < 1 \text{ for all } \mathbf{x},\mathbf{x}' \in [a_1,a_2]^d,$$

(iii) there exist $0 < a_1 < 1 < a_2 < \infty$ such that $a_1 \le X^{(j)} \le a_2$ for all $j = 1,\dots,d$.

Schäfer [28] considered the scheme, where $\{\mathbf{X}_i\}$ is a k-th order stationary Markov process with known k, while the situation of unknown k can be treated via machine learning combination of experts of degrees. However, the experiments on 19 NYSE assets of Györfi, Ottucsák and Urbán [12] showed that because of curse of dimensionality there is no gain for considering k-th order Markov modeling with $k > 1$.

We note that Conditions (ii) and (iii) imply uniform continuity of V and thus

$$\sup_{\mathbf{x},\mathbf{x}'\in[a_1,a_2]^d} V(\mathbf{x},\mathbf{x}') = \max_{\mathbf{x},\mathbf{x}'\in[a_1,a_2]^d} V(\mathbf{x},\mathbf{x}') < 1. \qquad (4)$$

Condition (iii) implies that the bankrupt is not possible. For the NYSE daily data, Condition (iii) is satisfied with $a_1 = 0.7$ and with $a_2 = 1.2$ (cf. Fernholz [7], Horváth and Urbán [19]).

From this point on assume that $\mathbf{b}_i$ is a function of the past return vectors: $\mathbf{b}_i = \mathbf{b}_i(\mathbf{X}_1^{i-1})$. Let's use the decomposition

$$\frac{1}{n}\log S_n = I_n + J_n, \qquad (5)$$

where I_n is

$$\frac{1}{n}\sum_{i=1}^{n}(g(\mathbf{b}_{i-1},\mathbf{b}_i,\mathbf{X}_{i-1},\mathbf{X}_i) - \mathbb{E}\{g(\mathbf{b}_{i-1},\mathbf{b}_i,\mathbf{X}_{i-1},\mathbf{X}_i)|\mathbf{X}_1^{i-1}\})$$

and

$$J_n = \frac{1}{n}\sum_{i=1}^{n}\mathbb{E}\{g(\mathbf{b}_{i-1},\mathbf{b}_i,\mathbf{X}_{i-1},\mathbf{X}_i)|\mathbf{X}_1^{i-1}\}.$$

I_n is an average of martingale differences. Under the Condition (iii), the random variable $g(\mathbf{b}_{i-1},\mathbf{b}_i,\mathbf{X}_{i-1},\mathbf{X}_i)$ is bounded ($|g(\mathbf{b}_{i-1},\mathbf{b}_i,\mathbf{X}_{i-1},\mathbf{X}_i)| \le c < \infty$), therefore I_n is an average of bounded martingale differences, which converges to 0 almost surely, since according to Chow's theorem (cf. Theorem 3.3.1 in Stout [29])

$$\sum_{i=1}^{\infty}\frac{\mathbb{E}\{g(\mathbf{b}_{i-1},\mathbf{b}_i,\mathbf{X}_{i-1},\mathbf{X}_i)^2\}}{i^2} \le \sum_{i=1}^{\infty}\frac{c^2}{i^2} < \infty$$

implies that

$$I_n \to 0 \qquad (6)$$

almost surely. Thus, the asymptotic maximization of the average growth rate $\frac{1}{n}\log S_n$ is equivalent to the maximization of J_n.

Under the condition (i), we have that

$$\begin{aligned}&\mathbb{E}\{g(\mathbf{b}_{i-1},\mathbf{b}_i,\mathbf{X}_{i-1},\mathbf{X}_i)|\mathbf{X}_1^{i-1}\}\\ =\;& \mathbb{E}\{\log(w(\mathbf{b}_{i-1},\mathbf{b}_i,\mathbf{X}_{i-1})\langle\mathbf{b}_i\,,\,\mathbf{X}_i\rangle)|\mathbf{X}_1^{i-1}\}\\ =\;& \log w(\mathbf{b}_{i-1},\mathbf{b}_i,\mathbf{X}_{i-1}) + \mathbb{E}\{\log\langle\mathbf{b}_i\,,\,\mathbf{X}_i\rangle\,|\mathbf{X}_1^{i-1}\}\\ =\;& \log w(\mathbf{b}_{i-1},\mathbf{b}_i,\mathbf{X}_{i-1}) + \mathbb{E}\{\log\langle\mathbf{b}_i\,,\,\mathbf{X}_i\rangle\,|\mathbf{b}_i,\mathbf{X}_{i-1}\}\\ \stackrel{\text{def}}{=}\;& v(\mathbf{b}_{i-1},\mathbf{b}_i,\mathbf{X}_{i-1}),\end{aligned}$$

therefore the maximization of the average growth rate $\frac{1}{n}\log S_n$ is asymptotically equivalent to the maximization of

$$J_n = \frac{1}{n}\sum_{i=1}^{n} v(\mathbf{b}_{i-1},\mathbf{b}_i,\mathbf{X}_{i-1}). \qquad (7)$$

The terms in the average J_n have a memory, which transforms the problem into a dynamic programming setup (cf. Merhav et al. [25]).

III. Growth optimal portfolio selection algorithms

An essential tool in the definition and investigation of portfolio selection algorithms under transaction costs are optimality equations of Bellman type. First we present an informal and heuristic way to them in our context of portfolio selection. Later on a rigorous treatment will be given.

Let us start with a finite-horizon problem concerning J_N defined by (7): For fixed integer $N > 0$, maximize

$$\begin{aligned}&\mathbb{E}\{N\cdot J_N \mid \mathbf{b}_0 = \mathbf{b},\mathbf{X}_0 = \mathbf{x}\}\\ =\;& \mathbb{E}\left\{\sum_{i=1}^{N} v(\mathbf{b}_{i-1},\mathbf{b}_i,\mathbf{X}_{i-1}) \mid \mathbf{b}_0 = \mathbf{b},\mathbf{X}_0 = \mathbf{x}\right\}\end{aligned}$$

by suitable choice of $\mathbf{b}_1,\dots,\mathbf{b}_N$. For general problems of dynamic programming (dynamic optimization), Bellman [3], p. 89, formulates his famous principle of optimality as follows: "An optimality policy has the property that whatever the initial state and initial decisions are, the remaining decisions must constitute an optimal policy with regard to the state resulting from the first decision."

By this principle, which for stochastic models is not so obvious as it seems (cf. pp. 14, 15 in Hinderer [18]), one can show: If the functions $G_0, G_1, \dots, G_N$ on $\Delta_d \times [a_1,a_2]^d$

are defined by the so-called dynamic programming equations (optimality equations, Bellman equations)

$$\begin{aligned} G_N(\mathbf{b},\mathbf{x}) &:= 0, \\ G_n(\mathbf{b},\mathbf{x}) &:= \max_{\mathbf{b}'}\left[v(\mathbf{b},\mathbf{b}',\mathbf{x})+\mathbb{E}\{G_{n+1}(\mathbf{b}',\mathbf{X}_2)\mid \mathbf{X}_1=\mathbf{x}\}\right] \end{aligned}$$

$(n = N-1, N-2, \dots, 0)$ with maximizer $\mathbf{b}'_n = G_n(\mathbf{b},\mathbf{x})$. Setting

$$F^n := G_{N-n}$$

$(n = 0, 1, \dots, N)$, one can write these backward equations in the forward form

$$\begin{aligned} F^0(\mathbf{b},\mathbf{x}) &:= 0, \\ F^n(\mathbf{b},\mathbf{x}) &:= \max_{\mathbf{b}'}\left[v(\mathbf{b},\mathbf{b}',\mathbf{x})+\mathbb{E}\{F^{n-1}(\mathbf{b}',\mathbf{X}_2)\mid \mathbf{X}_1=\mathbf{x}\}\right] \end{aligned} \tag{8}$$

$(n = 1, 2, \dots, N)$ with maximizer $F^n(\mathbf{b},\mathbf{x}) = G_{N-n}(\mathbf{b},\mathbf{x})$, where the choices $\mathbf{b}_n = F^n(\mathbf{b}_{n-1},\mathbf{X}_{n-1})$ are optimal.

For the situations, which are favorite for the investor, one has $F^n(\mathbf{b},\mathbf{x}) \to \infty$ as $n\to\infty$, which does not allow distinguishing between the qualities of competing choice sequences in the infinite-horizon case. If one considers (8) as a Value Iteration formula, then the underlying Bellman type equation

$$F^\infty(\mathbf{b},\mathbf{x}) = \max_{\mathbf{b}'}\left\{v(\mathbf{b},\mathbf{b}',\mathbf{x})+\mathbb{E}\{F^\infty(\mathbf{b}',\mathbf{X}_2)\mid \mathbf{X}_1=\mathbf{x}\}\right\}$$

has, roughly speaking, the degenerate solution $F^\infty = \infty$. Therefore one uses a discount factor $0<\delta<1$ and arrives at the discounted Bellman equation

$$F_\delta(\mathbf{b},\mathbf{x}) = \max_{\mathbf{b}'}\left\{v(\mathbf{b},\mathbf{b}',\mathbf{x})+(1-\delta)\mathbb{E}\{F_\delta(\mathbf{b}',\mathbf{X}_2)\mid \mathbf{X}_1=\mathbf{x}\}\right\}. \tag{9}$$

Its solution allows to solve the discounted problem maximizing

$$\begin{aligned} &\mathbb{E}\left\{\sum_{i=0}^{\infty}(1-\delta)^i v(\mathbf{b}_{i-1},\mathbf{b}_i,\mathbf{X}_{i-1})\mid \mathbf{b}_0=\mathbf{b},\mathbf{X}_0=\mathbf{x}\right\} \\ = &\sum_{i=0}^{\infty}(1-\delta)^i\mathbb{E}\left\{v(\mathbf{b}_{i-1},\mathbf{b}_i,\mathbf{X}_{i-1})\mid \mathbf{b}_0=\mathbf{b},\mathbf{X}_0=\mathbf{x}\right\}. \end{aligned}$$

The classic Hardy-Littlewood theorem (see, e.g., Theorem 97, together with Theorem 55 in [16]) states that for a real valued bounded sequence a_n, $n = 1, 2, \dots$,

$$\lim_{\delta\downarrow 0}\delta\sum_{i=0}^{\infty}(1-\delta)^i a_i$$

exists if and only if

$$\lim_{n\to\infty}\frac{1}{n}\sum_{i=0}^{n-1}a_i$$

exists and that then the limits are equal. Therefore, for maximizing

$$\lim_{n\to\infty}\frac{1}{n}\sum_{i=0}^{n-1}\mathbb{E}\left\{v(\mathbf{b}_{i-1},\mathbf{b}_i,\mathbf{X}_{i-1})\mid \mathbf{b}_0=\mathbf{b},\mathbf{X}_0=\mathbf{x}\right\},$$

(if it exists), it is important to solve the equation (9) for small δ. Letting $\delta\downarrow 0$, (9) with solution F^*_δ leads to the non-discounted Bellman equation

$$\lambda + F(\mathbf{b},\mathbf{x}) = \max_{\mathbf{b}'}\left\{v(\mathbf{b},\mathbf{b}',\mathbf{x})+\mathbb{E}\{F(\mathbf{b}',\mathbf{X}_2)\mid \mathbf{X}_1=\mathbf{x}\}\right\} \tag{10}$$

with a real constant λ. The interpretation of (8) as Value Iteration motivates solving (9) and (10) also by Value Iterations $F_{\delta,n}$ (see below) with discount factors $\delta > 0$. As to the corresponding problems in Markov Control theory we refer to Hernández-Lerma and Lasserre [17].

Let $B = B(\Delta_d\times[a_1,a_2]^d)$ and $C = C(\Delta_d\times[a_1,a_2]^d)$ be the Banach spaces of bounded measurable and of continuous functions F, respectively, defined on the compact set $\Delta_d\times[a_1,a_2]^d$ with the sup norm $\|\cdot\|_\infty$. Convergence with respect to $\|\cdot\|_\infty$ means uniform convergence. Let $0<\delta<1$ denote a discount factor. For such a δ, let

$$M^*_\delta : C\to C$$

be the operator which transforms each function $F\in C$ into a function $M^*_\delta F\in C$ defined by

$$\begin{aligned} &(M^*_\delta F)(\mathbf{b},\mathbf{x}) \\ = &\max_{\mathbf{b}'}\left\{v(\mathbf{b},\mathbf{b}',\mathbf{x})+(1-\delta)\mathbb{E}\{F(\mathbf{b}',\mathbf{X}_2)\mid \mathbf{X}_1=\mathbf{x}\}\right\} \end{aligned}$$

$((\mathbf{b},\mathbf{x})\in\Delta_d\times[a_1,a_2]^d)$. By Conditions (ii) and (iii), in fact $M^*_\delta F\in C$. The discounted Bellman equation (9) can be written in the form

$$F_\delta = M^*_\delta F_\delta.$$

Because of $0<\delta<1$, Banach's fixed point theorem yields that this equation has a unique solution (cf. Schäfer [28]). The so-called Value Iteration may result in the solution: for fixed $0<\delta<1$, put

$$F_{\delta,0} = 0$$

and

$$\begin{aligned} &F_{\delta,k+1}(\mathbf{b},\mathbf{x}) \\ = &\max_{\mathbf{b}'}\left\{v(\mathbf{b},\mathbf{b}',\mathbf{x})+(1-\delta)\mathbb{E}\{F_{\delta,k}(\mathbf{b}',\mathbf{X}_2)\mid \mathbf{X}_1=\mathbf{x}\}\right\}, \end{aligned}$$

$k = 0, 1, \dots$. Then Banach's fixed point theorem implies that the value iteration converges uniformly to the unique solution.

Knowing the distributions of the return vectors Schäfer [28], and Györfi and Vajda [14] introduced portfolio $\{\bar{\mathbf{b}}_i\}$ with capital $\bar{S}_n$ such that it is optimal in the sense that for any portfolio strategy $\{\mathbf{b}_i\}$ with capital S_n,

$$\liminf_{n\to\infty}\left(\frac{1}{n}\mathbb{E}\{\log\bar{S}_n\}-\frac{1}{n}\mathbb{E}\{\log S_n\}\right)\geq 0.$$

and

$$\liminf_{n\to\infty}\left(\frac{1}{n}\log\bar{S}_n-\frac{1}{n}\log S_n\right)\geq 0$$

a.s. Györfi and Walk [15] proved that a solution $(\lambda = W^*_c, F)$ of the (non-discounted) Bellman equation (10) exists, where $W^*_c\in\mathbb{R}$ is unique. W^*_c is the maximum growth rate (see below). If (W^*_c, F) is a solution then $(W^*_c, F+const)$ is a solution, too, therefore we introduce a standardized solution:

$$F - \max_{\mathbf{b},\mathbf{x}} F(\mathbf{b},\mathbf{x}),$$

which is again in C and has maximum value 0.

Again, knowing the distributions of the return vectors Györfi and Walk [15] introduced portfolio selection rules such that if S^*_n denotes the wealth at period n using these portfolios then

$$\lim_{n\to\infty}\frac{1}{n}\log S^*_n = W^*_c$$

a.s., while if S_n denotes the wealth at period n using any other portfolio then

$$\limsup_{n\to\infty} \frac{1}{n} \log S_n \le W_c^*$$

a.s.

Next we introduce an empirical (data driven) partitioning-based portfolio selection rule. Without transaction cost it was studied in Györfi and Schäfer [11]. Let $\mathcal{P}_n = \{A_{n,j}, j = 1, 2, \ldots\}$ be a sequence of cubic partitions of $\mathbb{R}^d$ with the side length of the cubic cells $h_n \downarrow 0$. For $\mathbf{x} \in \mathbb{R}^d$, set

$$A_n(\mathbf{x}) := A_{n,j} \text{ if } \mathbf{x} \in A_{n,j}.$$

Choose a sequence $0 < \delta_n < 1$ such that

$$\delta_n \downarrow 0, \quad \liminf_n n^\tau \delta_n > 0 \text{ for some } 0 < \tau < 1/2, \quad \frac{\delta_{n+1}}{\delta_n} \to 1,$$

e.g.,

$$\delta_n = \frac{1}{n^\tau}.$$

Set

$$F_1 := 0$$

and, with

$$\begin{aligned}
&(M_n F_n)(\mathbf{b}, \mathbf{x}) \\
:= \quad & \max_{\tilde{\mathbf{b}}} \Bigg\{ \log w(\mathbf{b}, \tilde{\mathbf{b}}, \mathbf{x}) + \frac{\sum_{i=2}^n \log \left\langle \tilde{\mathbf{b}}\,,\, \mathbf{X}_i \right\rangle I_{\mathbf{X}_{i-1} \in A_n(\mathbf{x})}}{\sum_{i=2}^n I_{\mathbf{X}_{i-1} \in A_n(\mathbf{x})}} \\
& + (1 - \delta_n) \frac{\sum_{i=2}^n F_n(\tilde{\mathbf{b}}, \mathbf{X}_i) I_{\mathbf{X}_{i-1} \in A_n(\mathbf{x})}}{\sum_{i=2}^n I_{\mathbf{X}_{i-1} \in A_n(\mathbf{x})}} \Bigg\} \qquad (11)
\end{aligned}$$

(with a void sum being 0 and $0/0 := 0$), iterate

$$F_{n+1} := M_n F_n - \sup_{\mathbf{b},\mathbf{x}} (M_n F_n)(\mathbf{b}, \mathbf{x}) \qquad (12)$$

$(n = 1, 2, \ldots)$. Put

$$\mathbf{b}_1 := \{1/d, \ldots, 1/d\}$$

and

$$\begin{aligned}
\mathbf{b}_{n+1} \quad := \quad & \arg\max_{\tilde{\mathbf{b}}} \Bigg\{ \log w(\mathbf{b}_n, \tilde{\mathbf{b}}, \mathbf{X}_n) \\
& + \frac{\sum_{i=2}^n \log \left\langle \tilde{\mathbf{b}}\,,\, \mathbf{X}_i \right\rangle I_{\mathbf{X}_{i-1} \in A_n(\mathbf{X}_n)}}{\sum_{i=2}^n I_{\mathbf{X}_{i-1} \in A_n(\mathbf{X}_n)}} \\
& + (1 - \delta_n) \frac{\sum_{i=2}^n F_n(\tilde{\mathbf{b}}, \mathbf{X}_i) I_{\mathbf{X}_{i-1} \in A_n(\mathbf{X}_n)}}{\sum_{i=2}^n I_{\mathbf{X}_{i-1} \in A_n(\mathbf{X}_n)}} \Bigg\}.
\end{aligned}$$

In the realistic case that the state space of the Markov process $(\mathbf{X}_n)$ is a finite set D of rational vectors (components being quotients of integer-valued \$-amounts) containing $\mathbf{e} = (1, \ldots, 1)$, the second part of Condition (ii) is fulfilled under the plausible assumption $\mu(\{\mathbf{e}\}|\mathbf{x}) > 0$ for all $\mathbf{x} \in D$. Another example for finite state Markov process is when one rounds down the components of $\mathbf{x}$ to a grid applying, for example, a grid size 0.00001. Under mild condition the Markov process is irreducible and aperiodic, e.g., assume that asset prices (in \$) are given by natural numbers and the d-tuple $\mathbf{s}$ of asset prices at the end of a trading period changes to a d-tuple $\mathbf{s}^*$ of asset prices at the end of the next trading period with positive probability for all $\mathbf{s}, \mathbf{s}^*$, where Condition (iii) is fulfilled. Then the Markov process $\mathbf{X}_n$ is really irreducible and aperiodic, since the state $\mathbf{e}$ is aperiodic because of $\mu(\{\mathbf{e}\}|\mathbf{e}) > 0$ and thus by irreducibility each state is aperiodic.

Theorem 1: Assume that the Markov process $\mathbf{X}_n$ takes values in a finite state space D and it is irreducible and aperiodic. Under the Conditions (i), (ii) and (iii), if S_n denotes the wealth at period n using the portfolio $\{\mathbf{b}_n\}$ then

$$\lim_{n\to\infty} \frac{1}{n} \log S_n = W_c^*$$

a.s.

One can comprehend a more general situation. Let the homogeneous first order Markov process $\{\mathbf{X}_n\}_{n\geq 1}$ on a state space $[a_1, a_2]^d$ be (Harris-)recurrent and strongly aperiodic. According to Athreya and Ney ([2], with references) this means the following: there exists a (measurable) set $A \subset [a_1, a_2]^d$, a probability measure ϕ on A, a number $0 < \lambda < \infty$ such that

$$\mathbb{P}\{\mathbf{X}_n \in A \text{ for some } n \ge 2 \mid \mathbf{X}_1 = \mathbf{x}\} = 1$$

for each $\mathbf{x} \in [a_1, a_2]^d$, and

$$\mu(U \mid \mathbf{x}) \ge \lambda \phi(U)$$

(μ is the Markov kernel) for each $\mathbf{x} \in A$ and each (measurable) set $U \subset A$.

We modify the partitioning-based portfolio selection rule to a k_n-nearest neighbor (k_n-NN) based rule. It is assumed that ties occur with probability zero. Because of the possibility of including a randomizer component into the return vector, this tie condition is not crucial (see, e.g., Györfi et al [9], pp. 86, 87). Choose $k_n = \lfloor n^K \rfloor$, $\delta_n = n^{-\tau}$ with $0 < \tau < K < 1$. We shall quantize the random variables: Choose a sequence $\{T_n\}$ of finite subsets of $[a_1, a_2]^d$ such that $T_n \uparrow$, $\cup_n T_n$ is dense in $[a_1, a_2]^d$, $card(T_n) = \lfloor n^\rho \rfloor$ with $0 < \rho < K$. Let

$$\mathbf{X}_{n,i} := \arg\min_{\mathbf{x} \in T_n} \|\mathbf{x} - \mathbf{X}_i\|.$$

Now set

$$F_1' := 0$$

and, with

$$I_{n,i}(\mathbf{x}) := I_{\{\mathbf{X}_{i-1} \text{ is among the } k_{n-1} \text{ NNs of } \mathbf{x} \text{ in } \{\mathbf{X}_1, \ldots, \mathbf{X}_{n-1}\}\}},$$

put

$$\begin{aligned}
&(Q_n F)(\mathbf{b}, \mathbf{x}) \\
:= \quad & \sup_{\tilde{\mathbf{b}}} \Bigg\{ \log w(\mathbf{b}, \tilde{\mathbf{b}}, \mathbf{x}) + \frac{1}{k_{n-1}} \sum_{i=2}^n \log \left\langle \tilde{\mathbf{b}}\,,\, \mathbf{X}_{n,i} \right\rangle I_{n,i}(\mathbf{x}) \\
& + \frac{1 - \delta_n}{k_{n-1}} \sum_{i=2}^n F(\tilde{\mathbf{b}}, \mathbf{X}_{n,i}) I_{n,i}(\mathbf{x}) \Bigg\},
\end{aligned}$$

$F \in B$ (with a void sum being 0 and $0/0 := 0$), iterate

$$F_{n+1}' := Q_n F_n' - W_n', \qquad (13)$$

where

$$W_n' = \sup_{\mathbf{b},\mathbf{x}} (Q_n F_n')(\mathbf{b}, \mathbf{x}),$$

$(n = 1, 2, \ldots)$. Put

$$\mathbf{b}_1' := \{1/d, \ldots, 1/d\}$$

and

$$\begin{aligned}
\mathbf{b}'_{n+1} \quad := \quad & \arg\max_{\tilde{\mathbf{b}}} \Big\{ \log w(\mathbf{b}'_n, \tilde{\mathbf{b}}, \mathbf{X}_n) \\
& + \frac{1}{k_{n-1}} \sum_{i=2}^{n} \log \left\langle \tilde{\mathbf{b}}\,,\, \mathbf{X}_{n,i} \right\rangle I_{n,i}(\mathbf{x}) \\
& + \frac{1-\delta_n}{k_{n-1}} \sum_{i=2}^{n} F'_n(\tilde{\mathbf{b}}, \mathbf{X}_{n,i}) I_{n,i}(\mathbf{x}) \Big\}.
\end{aligned}$$

Theorem 2: Assume that the Markov process $\mathbf{X}_n$ is recurrent and strongly aperiodic. Under the Conditions (i), (ii) and (iii), if S'_n denotes the wealth at period n using the portfolio $\{\mathbf{b}'_n\}$ then

$$\lim_{n\to\infty} \frac{1}{n} \log S'_n = W^*_c$$

a.s.

IV. Proofs

Proof of Theorem 1.

Step 1. In general, an irreducible denumerable homogeneous Markov chain is either transient or null-recurrent or positive-recurrent. But here, because of finite state space, only the third case is possible (cf. XV.6, Theorem 4 in Feller [6]). (Feller uses the terminology "persistent" instead of "recurrent".) Then by the ergodic theorem of Markov chains, for all fixed $m = 0, 1, \ldots$ and $\mathbf{x}, \mathbf{x}' \in D$,

$$\begin{aligned}
& \mathbb{P}\{\mathbf{X}_n = \mathbf{x}' \mid \mathbf{X}_m = \mathbf{x}\} \to \pi(\mathbf{x}') \\
:= \quad & \lim_{n\to\infty} \mathbb{P}\{\mathbf{X}_n = \mathbf{x}'\} = (\text{ mean recurrent time of } \mathbf{x}')^{-1} > 0
\end{aligned}$$

for $n \to \infty$ (cf. XV.7, Theorem in Feller [6]). According to Facts 4 and 3 in Rosenthal [26], all these convergences have an exponential rate. This means that $\mathbf{X}_n$ is ϕ-mixing with mixing coefficients $\phi_k \le c' e^{-c'' k}$ for some $c' > 0, c'' > 0$ (cf. Definition 2.2.1 in Györfi et al [8]). For a bounded function $F : \Delta_d \times D \to \mathbb{R}$ we show that

$$\begin{aligned}
& \sup_{\mathbf{b}\in\Delta_d, \mathbf{x}\in D} \Bigg| \frac{\sum_{i=2}^n F(\mathbf{b}, \mathbf{X}_i) I_{\mathbf{X}_{i-1}\in A_n(\mathbf{x})}}{\sum_{i=2}^n I_{\mathbf{X}_{i-1}\in A_n(\mathbf{x})}} \\
& \quad - \mathbb{E}\{F(\mathbf{b}, \mathbf{X}_2) \mid \mathbf{X}_1 = \mathbf{x}\} \Bigg| \\
\le \quad & E'_n \sup_{\mathbf{b}\in\Delta_d, \mathbf{x}\in D} \mathbb{E}\{|F(\mathbf{b}, \mathbf{X}_2)| \mid \mathbf{X}_1 = \mathbf{x}\}
\end{aligned}$$

a.s. with random variables $E'_n = o(n^{-\tau})$ independent of F. We note

$$\mathbb{E}\{F(\mathbf{b}, \mathbf{X}_2) \mid \mathbf{X}_1 = \mathbf{x}\} = \sum_{\mathbf{x}'\in D} F(\mathbf{b}, \mathbf{x}') \mu(\{\mathbf{x}'\} \mid \mathbf{x}),$$

$\mathbf{b} \in \Delta_d, \mathbf{x} \in D$. Further for $\mathbf{b} \in \Delta_d, \mathbf{x} \in D$ and n sufficiently large (independent of $F, \mathbf{b}, \mathbf{x}$) we have

$$\begin{aligned}
& \frac{\sum_{i=2}^n F(\mathbf{b}, \mathbf{X}_i) I_{\mathbf{X}_{i-1}\in A_n(\mathbf{x})}}{\sum_{i=2}^n I_{\mathbf{X}_{i-1}\in A_n(\mathbf{x})}} \\
= \quad & \frac{\sum_{i=2}^n F(\mathbf{b}, \mathbf{X}_i) I_{\mathbf{X}_{i-1}=\mathbf{x}}}{\sum_{i=2}^n I_{\mathbf{X}_{i-1}=\mathbf{x}}} \\
= \quad & \frac{\sum_{i=2}^n \sum_{\mathbf{x}'\in D, \mu(\{\mathbf{x}'\}|\mathbf{x})>0} F(\mathbf{b}, \mathbf{x}') I_{\mathbf{X}_i=\mathbf{x}', \mathbf{X}_{i-1}=\mathbf{x}}}{\sum_{i=2}^n I_{\mathbf{X}_{i-1}=\mathbf{x}}} \\
= \quad & \sum_{\mathbf{x}'\in D, \mu(\{\mathbf{x}'\}|\mathbf{x})>0} F(\mathbf{b}, \mathbf{x}') \\
& \cdot \Bigg(\frac{\frac{1}{n}\sum_{i=2}^n (I_{\mathbf{X}_i=\mathbf{x}', \mathbf{X}_{i-1}=\mathbf{x}} - \mathbb{P}\{\mathbf{X}_i = \mathbf{x}', \mathbf{X}_{i-1} = \mathbf{x}\})}{\frac{1}{n}\sum_{i=2}^n [(I_{\mathbf{X}_{i-1}=\mathbf{x}} - \mathbb{P}\{\mathbf{X}_{i-1} = \mathbf{x}\}) + \mathbb{P}\{\mathbf{X}_{i-1} = \mathbf{x}\}]} \\
& + \frac{\frac{1}{n}\sum_{i=2}^n \mathbb{P}\{\mathbf{X}_i = \mathbf{x}', \mathbf{X}_{i-1} = \mathbf{x}\}}{\frac{1}{n}\sum_{i=2}^n [(I_{\mathbf{X}_{i-1}=\mathbf{x}} - \mathbb{P}\{\mathbf{X}_{i-1} = \mathbf{x}\}) + \mathbb{P}\{\mathbf{X}_{i-1} = \mathbf{x}\}]} \Bigg)
\end{aligned}$$

a.s., since $I_{\mathbf{X}_i=\mathbf{x}', \mathbf{X}_{i-1}=\mathbf{x}} = 0$ a.s. in case $\mu(\{\mathbf{x}'\} \mid \mathbf{x}) = 0$. The sequence $(\mathbf{X}_{n-1}, \mathbf{X}_n)$ is ϕ-mixing with exponential convergence rate of mixing coefficients ϕ'_k, thus $\Phi := \sum_{k=1}^\infty \phi'_k < \infty$. We use Collomb's exponential inequality (see Theorem 2.2.1 in Györfi et al. [8]) noticing

$$\frac{1}{n^{1-\tau}} \left| I_{\mathbf{X}_i=\mathbf{x}', \mathbf{X}_{i-1}=\mathbf{x}} - \mathbb{P}\{\mathbf{X}_i = \mathbf{x}', \mathbf{X}_{i-1} = \mathbf{x}\} \right| \le \frac{1}{n^{1-\tau}}$$

and

$$\begin{aligned}
& \mathbb{E}\left\{ \left(\frac{1}{n^{1-\tau}} \left| I_{\mathbf{X}_i=\mathbf{x}', \mathbf{X}_{i-1}=\mathbf{x}} - \mathbb{P}\{\mathbf{X}_i = \mathbf{x}', \mathbf{X}_{i-1} = \mathbf{x}\} \right| \right)^2 \right\} \\
\le \quad & \frac{1}{n^{2(1-\tau)}}
\end{aligned}$$

and obtain for $\epsilon > 0$

$$\mathbb{P}\left\{ \left| \frac{1}{n^{1-\tau}} \sum_{i=2}^n (I_{\mathbf{X}_i=\mathbf{x}', \mathbf{X}_{i-1}=\mathbf{x}} - \mathbb{P}\{\mathbf{X}_i = \mathbf{x}', \mathbf{X}_{i-1} = \mathbf{x}\}) \right| > \epsilon \right\}$$

$$\le e^{3\sqrt{e} n \phi'_m / m - \alpha\epsilon + 6\alpha^2 n (1+4\Phi)/n^{2(1-\tau)}}$$

with $\alpha > 0$, $1 \le m \le n-1$, $\alpha m / n^{1-\tau} \le 1/4$. Choosing

$$m = \lfloor n^\sigma \rfloor$$

with $\tau < \sigma < 1 - \tau$ and

$$\alpha = \frac{n^{1-\tau}}{4m},$$

the right-hand side for $n = 2, 3, \ldots$ is bounded from above by

$$e^{3\sqrt{e}(n-1)\phi'_{\lfloor (n-1)^\sigma \rfloor} / \lfloor (n-1)^\sigma \rfloor - (n-1)^{1-\tau-\sigma}\epsilon/4 + 3(1+4\Phi) n^{1-2\sigma}/8}$$

(where $n\phi'_{\lfloor n^\sigma \rfloor} / \lfloor n^\sigma \rfloor \to 0$), which converges to 0 exponentially fast. Thus

$$\frac{1}{n} \sum_{i=2}^n (I_{\mathbf{X}_i=\mathbf{x}', \mathbf{X}_{i-1}=\mathbf{x}} - \mathbb{P}\{\mathbf{X}_i = \mathbf{x}', \mathbf{X}_{i-1} = \mathbf{x}\}) = o(n^{-\tau})$$

a.s. Further, by homogeneity of the Markov chain $\mathbf{X}_n$ and the exponential convergence rate of $\mathbb{P}\{\mathbf{X}_n = \mathbf{x}'\}$ mentioned above

$$\begin{aligned} & \left|\frac{1}{n}\sum_{i=2}^{n}\mathbb{P}\{\mathbf{X}_i = \mathbf{x}', \mathbf{X}_{i-1} = \mathbf{x}\} - \mu(\{\mathbf{x}'\} \mid \mathbf{x})\pi(\mathbf{x})\right| \\ = \quad & \mu(\{\mathbf{x}'\} \mid \mathbf{x})\left|\frac{1}{n}\sum_{i=2}^{n}\mathbb{P}\{\mathbf{X}_{i-1} = \mathbf{x}\} - \pi(\mathbf{x})\right| \\ \leq \quad & \mu(\{\mathbf{x}'\} \mid \mathbf{x})\left(\frac{1}{n}\sum_{i=2}^{\infty}|\mathbb{P}\{\mathbf{X}_{i-1} = \mathbf{x}\} - \pi(\mathbf{x})| + \pi(\mathbf{x})/n\right) \\ = \quad & O(1/n). \end{aligned}$$

Because the state space D is finite, a.s. the rates of convergence are uniform with respect to $\mathbf{x}, \mathbf{x}' \in D$. The argument concerning $\frac{1}{n}\sum_{i=2}^{n} I_{\mathbf{X}_{i-1}=\mathbf{x}}$ is analogous, but even simpler.

$$\begin{aligned} & \frac{\sum_{i=2}^{n} F(\mathbf{b}, \mathbf{X}_i) I_{\mathbf{X}_{i-1}\in A_n(\mathbf{x})}}{\sum_{i=2}^{n} I_{\mathbf{X}_{i-1}\in A_n(\mathbf{x})}} \\ = \quad & \sum_{\mathbf{x}'\in D, \mu(\{\mathbf{x}'\}|\mathbf{x})>0} F(\mathbf{b}, \mathbf{x}')\frac{\mu(\{\mathbf{x}'\} \mid \mathbf{x})\pi(\mathbf{x}) + o(n^{-\tau})}{\pi(\mathbf{x}) + o(n^{-\tau})} \\ = \quad & \sum_{\mathbf{x}'\in D} F(\mathbf{b}, \mathbf{x}')\mu(\{\mathbf{x}'\} \mid \mathbf{x})(1 + o(n^{-\tau})) \\ = \quad & \mathbb{E}\{F(\mathbf{b}, \mathbf{X}_2) \mid \mathbf{X}_1 = \mathbf{x}\}(1 + o(n^{-\tau})) \end{aligned}$$

uniformly with respect to $\mathbf{x} \in D$ and $\mathbf{b} \in \Delta_d$ a.s., since the o-terms depend only on $\mathbf{x}$, not on $\mathbf{b}$ or F. This yields the assertion.

Step 2. With B and C as in Section III and with M_n defined by (11), we show that F_n converges in B to a set of solutions (in C) of the Bellman equation (10) a.s., further

$$W_n := \max_{\mathbf{b},\mathbf{x}}(M_n F_n)(\mathbf{b}, \mathbf{x}) \to W_c^* \tag{14}$$

a.s. For $0 \leq \delta < 1$ and for $F \in B$, define the operator

$$\begin{aligned} & (M_\delta^* F)(\mathbf{b}, \mathbf{x}) \\ := \quad & \sup_{\mathbf{b}'}\left\{v(\mathbf{b}, \mathbf{b}', \mathbf{x}) + (1-\delta)\mathbb{E}\{F(\mathbf{b}', \mathbf{X}_2) \mid \mathbf{X}_1 = \mathbf{x}\}\right\}. \end{aligned} \tag{15}$$

By continuity assumption (ii), with restriction on C, this leads to an operator

$$M_\delta^* : C \to C.$$

(See Schäfer [28] p.114.) The operator $M_\delta^* : B \to B$ is continuous, even Lipschitz continuous with Lipschitz constant $1 - \delta$. Indeed, for $F, F' \in B$ from the representation

$$\begin{aligned} & (M_\delta^* F)(\mathbf{b}, \mathbf{x}) \\ = \quad & v(\mathbf{b}, \mathbf{b}_F^*(\mathbf{b}, \mathbf{x}), \mathbf{x}) + (1-\delta)\mathbb{E}\{F(\mathbf{b}_F^*(\mathbf{b}, \mathbf{x}), \mathbf{X}_2) \mid \mathbf{X}_1 = \mathbf{x}\}, \end{aligned}$$

without loss of generality assuming that sup is attained, and from the corresponding representation of $(M_\delta^* F')(\mathbf{b}, \mathbf{x})$ one obtains

$$\begin{aligned} & (M_\delta^* F')(\mathbf{b}, \mathbf{x}) \\ \geq \quad & v(\mathbf{b}, \mathbf{b}_F^*(\mathbf{b}, \mathbf{x}), \mathbf{x}) + (1-\delta)\mathbb{E}\{F'(\mathbf{b}_F^*(\mathbf{b}, \mathbf{x}), \mathbf{X}_2) \mid \mathbf{X}_1 = \mathbf{x}\} \\ \geq \quad & v(\mathbf{b}, \mathbf{b}_F^*(\mathbf{b}, \mathbf{x}), \mathbf{x}) + (1-\delta)\mathbb{E}\{F(\mathbf{b}_F^*(\mathbf{b}, \mathbf{x}), \mathbf{X}_2) \mid \mathbf{X}_1 = \mathbf{x}\} \\ & -(1-\delta)||F - F'||_\infty \\ = \quad & (M_\delta^* F)(\mathbf{b}, \mathbf{x}) - (1-\delta)||F - F'||_\infty \end{aligned}$$

for all $(\mathbf{b}, \mathbf{x}) \in \Delta_d \times [a_1, a_2]^d$, therefore

$$||M_\delta^* F - M_\delta^* F'||_\infty \leq (1-\delta)||F - F'||_\infty.$$

It can be easily checked that

$$||M_{\delta_{n+1}}^* F_{n+1}' - M_{\delta_n}^* F_{n+1}'||_\infty \leq (\delta_n - \delta_{n+1})||F_{n+1}'||_\infty. \tag{16}$$

From Step 1, noticing

$$L := \sup_{\mathbf{b}\in\Delta_d, \mathbf{x}\in D}|\log\langle\mathbf{b}\,,\,\mathbf{x}\rangle| < \infty,$$

we obtain

$$||M_n F_n - M_{\delta_n}^* F_n||_\infty \leq E_n(1 + ||F_n||_\infty) \tag{17}$$

a.s. with random variables

$$E_n := (2 + L)E_n' = o(n^{-\tau}).$$

Because of (17) it holds

$$\begin{aligned} & |F_{n+1}(\bar{\mathbf{b}}, \bar{\mathbf{x}}) - F_{n+1}(\mathbf{b}, \mathbf{x})| \\ = \quad & |(M_n F_n)(\bar{\mathbf{b}}, \bar{\mathbf{x}}) - (M_n F_n)(\mathbf{b}, \mathbf{x})| \\ \leq \quad & |(M_{\delta_n}^* F_n)(\bar{\mathbf{b}}, \bar{\mathbf{x}}) - (M_{\delta_n}^* F_n)(\mathbf{b}, \mathbf{x})| + 2E_n(1 + ||F_n||_\infty) \\ \leq \quad & \max_{\mathbf{b}'}|v(\bar{\mathbf{b}}, \mathbf{b}', \bar{\mathbf{x}}) - v(\mathbf{b}, \mathbf{b}', \bar{\mathbf{x}})| \\ & + \max_{\mathbf{b}'}|v(\mathbf{b}, \mathbf{b}', \bar{\mathbf{x}}) - v(\mathbf{b}, \mathbf{b}', \mathbf{x})| \\ & + V(\mathbf{x}, \bar{\mathbf{x}})||F_n||_\infty + 2E_n(1 + ||F_n||_\infty) \end{aligned} \tag{18}$$

a.s. Then, because of boundedness of v,

$$||F_{n+1}||_\infty \leq const + \max_{\mathbf{x},\bar{\mathbf{x}}} V(\mathbf{x}, \bar{\mathbf{x}})||F_n||_\infty + 2E_n(1 + ||F_n||_\infty)$$

a.s. Noticing $E_n \to 0$ a.s. and $\max_{\mathbf{x},\bar{\mathbf{x}}} V(\mathbf{x}, \bar{\mathbf{x}}) < 1$, one obtains

$$||F_n||_\infty \leq E < \infty \tag{19}$$

a.s. with some random variable E. With

$$\begin{aligned} & E_n^* := E_{n+1}(1 + ||F_{n+1}||_\infty) + E_n(1 + ||F_n||_\infty) \\ \leq \quad & (E_{n+1} + E_n)(1 + E) = o(n^{-\tau}) \end{aligned}$$

a.s. (by (19)), the Lipschitz continuity of $M_{\delta_n}^*$ with Lipschitz constant $1 - \delta_n$, (16) for F_{n+1}, (19) and the conditions on δ_n we obtain that

$$\begin{aligned} & ||F_{n+2} - F_{n+1}||_\infty \\ = \quad & ||M_{n+1}F_{n+1} - M_n F_n||_\infty \\ \leq \quad & ||M_{\delta_{n+1}}^* F_{n+1} - M_{\delta_n}^* F_n||_\infty + E_n^* \\ \leq \quad & ||M_{\delta_n}^* F_{n+1} - M_{\delta_n}^* F_n||_\infty + ||M_{\delta_{n+1}}^* F_{n+1} - M_{\delta_n}^* F_{n+1}||_\infty \\ & + E_n^* \\ \leq \quad & (1-\delta_n)||F_{n+1} - F_n||_\infty + (\delta_n - \delta_{n+1})||F_{n+1}||_\infty + E_n^* \\ \leq \quad & (1-\delta_n)||F_{n+1} - F_n||_\infty + \left(\left(1 - \frac{\delta_{n+1}}{\delta_n}\right)E + \frac{E_n^*}{\delta_n}\right)\delta_n \\ \leq \quad & (1-\delta_n)||F_{n+1} - F_n||_\infty + o(1)\delta_n \end{aligned}$$

a.s., leading to

$$||F_{n+2} - F_{n+1}||_\infty \to 0 \tag{20}$$

a.s. (cf. Lemma 1(c) in Walk and Zsidó [30]). Now let $\{\delta_{n_k}\}$ be an arbitrary subsequence of $\{\delta_n\}$. From Condition (ii), (18) and (19) we obtain

$$\sup_{i\geq j}|F_i(\bar{\mathbf{b}}, \bar{\mathbf{x}}) - F_i(\mathbf{b}, \mathbf{x})| \to 0$$

a.s. when $(\bar{\mathbf{b}}, \bar{\mathbf{x}}) \to (\mathbf{b}, \mathbf{x})$ and $j \to \infty$, even uniformly with respect to $(\mathbf{b}, \mathbf{x})$. This together with (19) yields existence of a random subsequence $\{\delta_{n_{k_\ell}}\}$ and of a random function F^* with realizations in C (bounded, where $\max_{\mathbf{b},\mathbf{x}} F^*(\mathbf{b},\mathbf{x}) = 0$) such that

$$||F_{n_{k_\ell}} - F^*||_\infty \to 0 \tag{21}$$

a.s. as $\ell \to \infty$ (cf. Ascoli-Arzelá theorem and its proof, [31]). Thus, by continuity of M_0^*,

$$||M_0^* F_{n_{k_\ell}} - M_0^* F^*||_\infty \to 0 \tag{22}$$

a.s. as $\ell \to \infty$. By (12),

$$F_{n_{k_\ell}} + (F_{n_{k_\ell}+1} - F_{n_{k_\ell}}) = M_0^* F_{n_{k_\ell}} + (M_{n_{k_\ell}} F_{n_{k_\ell}} - M_0^* F_{n_{k_\ell}}) - W_{n_{k_\ell}}.$$

(20) implies that

$$||F_{n_{k_\ell}+1} - F_{n_{k_\ell}}||_\infty \to 0$$

a.s. We notice

$$\begin{aligned} & ||M_{n_{k_\ell}} F_{n_{k_\ell}} - M_0^* F_{n_{k_\ell}}||_\infty \\ \leq \; & ||M_{n_{k_\ell}} F_{n_{k_\ell}} - M^*_{\delta_{n_{k_\ell}}} F_{n_{k_\ell}}||_\infty + ||M^*_{\delta_{n_{k_\ell}}} F_{n_{k_\ell}} - M_0^* F_{n_{k_\ell}}||_\infty \\ \leq \; & E_{n_{k_\ell}}(1 + ||F_{n_{k_\ell}}||_\infty) + \delta_{n_{k_\ell}} ||F_{n_{k_\ell}}||_\infty \\ \to \; & 0 \end{aligned}$$

a.s. (by (17), (16) and (19)). This together with (21) and (22) yields a.s. convergence of $W_{n_{k_\ell}}$ and

$$\lim_\ell W_{n_{k_\ell}} + F^* = M_0^* F^*$$

a.s. This equation means that a.s. the realizations of F^* solve the Bellman equation (10) such that

$$\lim_\ell W_{n_{k_\ell}} = W_c^*$$

a.s. This yields the assertion.

Step 3. We show the assertion of Theorem 1. Noticing that F_n depends on $\mathbf{X}_1, \ldots, \mathbf{X}_{n-1}$ and that $\mathbf{b}_{n+1}$ depends on $\mathbf{X}_1, \ldots, \mathbf{X}_n$, Step 1 together with a.s. uniform boundedness of F_n (by (19)) and the assumption that $\mathbf{X}_n$ is a homogeneous first order Markov chain yields

$$\frac{\sum_{i=2}^n F_n(\mathbf{b}_{n+1}, \mathbf{X}_i) I_{\mathbf{X}_{i-1} \in A_n(\mathbf{X}_n)}}{\sum_{i=2}^n I_{\mathbf{X}_{i-1} \in A_n(\mathbf{X}_n)}} - \mathbb{E}\{F_n(\mathbf{b}_{n+1}, \mathbf{X}_{n+1}) \mid \mathbf{X}_1^n\} \to 0 \tag{23}$$

a.s., further

$$\begin{aligned} & \mathbb{E}\{\log \langle \mathbf{b}_{n+1}, \mathbf{X}_{n+1} \rangle \mid \mathbf{b}_{n+1}, \mathbf{X}_n\} \\ = \; & \mathbb{E}\{\log \langle \mathbf{b}_{n+1}, \mathbf{X}_{n+1} \rangle \mid \mathbf{X}_1^n\} \\ = \; & \frac{\sum_{i=2}^n \log \langle \mathbf{b}_{n+1}, \mathbf{X}_i \rangle I_{\mathbf{X}_{i-1} \in A_n(\mathbf{X}_n)}}{\sum_{i=2}^n I_{\mathbf{X}_{i-1} \in A_n(\mathbf{X}_n)}} + o(1) \end{aligned} \tag{24}$$

a.s. Because of (5), (6), (7) and (24) it is enough to prove

$$\begin{aligned} T_N \; := \; & \frac{1}{N} \sum_{n=1}^N \Big(\log w(\mathbf{b}_n, \mathbf{b}_{n+1}, \mathbf{X}_n) \\ & + \frac{\sum_{i=2}^n \log \langle \mathbf{b}_{n+1}, \mathbf{X}_i \rangle I_{\mathbf{X}_{i-1} \in A_n(\mathbf{X}_n)}}{\sum_{i=2}^n I_{\mathbf{X}_{i-1} \in A_n(\mathbf{X}_n)}} \Big) \to W_c^* \end{aligned} \tag{25}$$

a.s. Thus,

$$\begin{aligned} & W_n + F_{n+1}(\mathbf{b}_n, \mathbf{X}_n) \\ = \; & \Big(\log w(\mathbf{b}_n, \mathbf{b}_{n+1}, \mathbf{X}_n) \\ & + \frac{\sum_{i=2}^n \log \langle \mathbf{b}_{n+1}, \mathbf{X}_i \rangle I_{\mathbf{X}_{i-1} \in A_n(\mathbf{X}_n)}}{\sum_{i=2}^n I_{\mathbf{X}_{i-1} \in A_n(\mathbf{X}_n)}} \Big) \\ & + (1 - \delta_n) \frac{\sum_{i=2}^n F_n(\mathbf{b}_{n+1}, \mathbf{X}_i) I_{\mathbf{X}_{i-1} \in A_n(\mathbf{X}_n)}}{\sum_{i=2}^n I_{\mathbf{X}_{i-1} \in A_n(\mathbf{X}_n)}}. \end{aligned}$$

Then

$$\begin{aligned} T_N \; = \; & \frac{1}{N} \sum_{n=1}^N W_n + \frac{1}{N} \sum_{n=1}^N F_{n+1}(\mathbf{b}_n, \mathbf{X}_n) \\ & - \frac{1}{N} \sum_{n=1}^N (1 - \delta_n) \frac{\sum_{i=2}^n F_n(\mathbf{b}_{n+1}, \mathbf{X}_i) I_{\mathbf{X}_{i-1} \in A_n(\mathbf{X}_n)}}{\sum_{i=2}^n I_{\mathbf{X}_{i-1} \in A_n(\mathbf{X}_n)}}. \end{aligned}$$

Without loss of generality we may assume that E in (19) is a constant. Otherwise we suitably truncate F_n having an exceptional set of arbitrarily small probability measure. By (19) and (23) together with $\delta_n \to 0$ we obtain

$$\begin{aligned} T_N \; = \; & \frac{1}{N} \sum_{n=1}^N W_n \\ & + \frac{1}{N} \sum_{n=1}^N (F_{n+1}(\mathbf{b}_n, \mathbf{X}_n) - \mathbb{E}\{F_n(\mathbf{b}_{n+1}, \mathbf{X}_{n+1}) \mid \mathbf{X}_1^n\}) \\ & + o(1) \end{aligned}$$

a.s. This together with (14), (20) and (19) implies that

$$\begin{aligned} T_N \; = \; & W_c^* \\ & + \frac{1}{N} \sum_{n=1}^N (F_n(\mathbf{b}_{n+1}, \mathbf{X}_{n+1}) - \mathbb{E}\{F_n(\mathbf{b}_{n+1}, \mathbf{X}_{n+1}) \mid \mathbf{X}_1^n\}) \\ & + o(1) \end{aligned}$$

a.s. By (19), Chow's theorem yields that the middle term of the right hand side a.s. converges to 0. Thus (25) is obtained. ∎

Sketch of the proof of Theorem 2.

Step 1. Athreya and Ney state ([2], Theorem (4.1), (i)): if the homogeneous first order Markov process $\{\mathbf{X}_n\}_{n \geq 1}$ is recurrent and strongly aperiodic, with invariant probability measure π (i.e., $\int \mu(\cdot \mid \mathbf{x}) \pi(d\mathbf{x}) = \pi$), then

$$\sup_{D \subset [a_1, a_2]^d} |\mathbb{P}\{\mathbf{X}_n \in D \mid \mathbf{X}_1 = \mathbf{x}\} - \pi(D)| \to 0$$

for each $\mathbf{x} \in [a_1, a_2]^d$.

In our situation

$$\begin{aligned} & \sup_{D \subset [a_1, a_2]^d} |\mathbb{P}\{\mathbf{X}_n \in D \mid \mathbf{X}_1 = \mathbf{x}\} - \mathbb{P}\{\mathbf{X}_n \in D \mid \mathbf{X}_1 = \mathbf{x}'\}| \\ \leq \; & \sup_{D \subset [a_1, a_2]^d} |\mu(D \mid \mathbf{x}) - \mu(D \mid \mathbf{x}')| \to 0 \; (\mathbf{x}' \to \mathbf{x}) \end{aligned}$$

by Condition (ii). Therefore even

$$\sup_{\mathbf{x}, D} |\mathbb{P}\{\mathbf{X}_n \in D \mid \mathbf{X}_1 = \mathbf{x}\} - \pi(D)| \to 0 \tag{26}$$

as $n \to \infty$. Thus $\{\mathbf{X}_n\}$ is ϕ-mixing. Also $\{(\mathbf{X}_n, \mathbf{X}_{n-1}); n \geq 2\}$ is ϕ-mixing. Let $\mathcal{A}$ be the system of closed spheres $S \subset (0, \infty)^d$ with centers in $[a_1, a_2]^d$. For each $F \in B$, with

$$V_{F,n,\mathbf{b},S} := \frac{1}{k_{n-1}} \sum_{i=2}^n F(\mathbf{b}, \mathbf{X}_{n,i}) I_{\{\mathbf{X}_{i-1} \in S\}}$$

we have

$$\sup_{S\in\mathcal{A}}\sup_{\mathbf{b}\in\Delta_d}|V_{F,n,\mathbf{b},S}-\mathbb{E}\{V_{F,n,\mathbf{b},S}\}|\leq E_n'\|F\|_\infty \tag{27}$$

a.s. with random variables $E_n'\to 0$ independent of F. This is obtained by an application of Collomb's exponential inequality (cf. Györfi et al. [8], pp. 19, 20) for the ϕ-mixing sequence $\{(\mathbf{X}_n,\mathbf{X}_{n-1})\}$, which yields

$$\begin{aligned}
&\mathbb{P}\left\{\sup_{\|F\|_\infty>0}\sup_{S,\mathbf{b}}\frac{1}{\|F\|_\infty}|V_{F,n,\mathbf{b},S}-\mathbb{E}\{V_{F,n,\mathbf{b},S}\}|>\epsilon\right\}\\
\leq\;&\sum_{\mathbf{x}\in T_n}\mathbb{P}\Big\{\sup_{S\in\mathcal{A}}\Big|\frac{1}{n}\sum_{i=2}^n(I_{\{\mathbf{X}_{n,i}=\mathbf{x},\mathbf{X}_{i-1}\in S\}}\\
&-\mathbb{E}\{I_{\{\mathbf{X}_{n,i}=\mathbf{x},\mathbf{X}_{i-1}\in S\}}\})\Big|>\frac{\epsilon k_n}{n\cdot card(T_n)}\Big\}\\
\leq\;&c_1 card(T_n)n^{d+2}e^{-c_2\epsilon k_n/card(T_n)}
\end{aligned}$$

with $c_1,c_2\in(0,\infty)$ depending on ϵ, where the factor n^{d+2} in the right-hand side follows from the Vapnik-Chervonenkis theory (cf. Kohler et al. [24], p. 689). Now the Borel-Cantelli lemma yields (27). For $\mathbf{x}\in[a_1,a_2]^d$, set

$$\hat{h}_n(\mathbf{x}):=\min\left\{h>0;\sum_{i=2}^n I_{\{\mathbf{X}_{i-1}\in S_{\mathbf{x},h}\}}\geq k_{n-1}\right\}.$$

Introduce the notation

$$\nu_i(H):=\mathbb{P}\{\mathbf{X}_i\in H\}.$$

For $F=1$, (27) yields

$$\frac{1}{k_{n-1}}\sum_{i=2}^n\nu_{i-1}(S_{\mathbf{x},\hat{h}_n(\mathbf{x})})\to 1 \tag{28}$$

uniformly with respect to $\mathbf{x}$, a.s., further by

$$\lim_n\mathbb{E}\left\{\frac{1}{n-1}\sum_{i=2}^n I_{\{\mathbf{X}_{i-1}\in S\}}\right\}=\lim_n\mathbb{P}\{\mathbf{X}_n\in S\}=\pi(S),\ S\in\mathcal{A},$$

we get that

$$\sup_{\mathbf{x}\in supp(\pi)}\hat{h}_n(\mathbf{x})\to 0 \tag{29}$$

a.s.

Step 2. We show that F_n' a.s. converges in B to the set of solutions ($\in C$) of the Bellman equation (10), further

$$W_n'=\sup_{\mathbf{b},\mathbf{x}}(Q_nF_n')(\mathbf{b},\mathbf{x})\to W_c^*$$

a.s. For $F\in B$, set

$$(R_n'F)(\mathbf{b},\mathbf{x},h):=\frac{1}{k_{n-1}}\sum_{i=2}^n\mathbb{E}\{F(\mathbf{b},\mathbf{X}_{n,i})I_{\{\mathbf{X}_{i-1}\in S_{\mathbf{x},h}\}}\}$$

and

$$(R_n^*F)(\mathbf{b},\mathbf{x}):=(R_n'F)(\mathbf{b},\mathbf{x},\hat{h}_n(\mathbf{x}))$$

and

$$(R_0F)(\mathbf{b},\mathbf{x}):=\mathbb{E}\{F(\mathbf{b},\mathbf{X}_2)\mid\mathbf{X}_1=\mathbf{x}\}$$

and

$$\begin{aligned}
(Q_n^*F)(\mathbf{b},\mathbf{x})\;:=\;&\sup_{\mathbf{b}'}\{\log w(\mathbf{b},\mathbf{b}',\mathbf{x})+(R_n^*\log\langle\cdot\,,\,\cdot\rangle)(\mathbf{b}',\mathbf{x})\\
&+(1-\delta_n)(R_n^*F)(\mathbf{b}',\mathbf{x})\}
\end{aligned}$$

and

$$\begin{aligned}
(Q_0F)(\mathbf{b},\mathbf{x})\;:=\;&(M_0^*F)(\mathbf{b},\mathbf{x})\\
=\;&\sup_{\mathbf{b}'}\{\log w(\mathbf{b},\mathbf{b}',\mathbf{x})+(R_0\log\langle\cdot\,,\,\cdot\rangle)(\mathbf{b}',\mathbf{x})\\
&+(R_0F)(\mathbf{b}',\mathbf{x})\}.
\end{aligned}$$

For each $F\in B$, with $L':=\max_{\mathbf{b},\mathbf{x}}\log\langle\mathbf{b}\,,\,\mathbf{x}\rangle<\infty$, we have

$$\begin{aligned}
&|(Q_n^*F)(\mathbf{b},\mathbf{x})-(Q_n^*F)(\bar{\mathbf{b}},\bar{\mathbf{x}})|\\
\leq\;&\sup_{\mathbf{b}'}|\log w(\bar{\mathbf{b}},\mathbf{b}',\bar{\mathbf{x}})-\log w(\mathbf{b},\mathbf{b}',\bar{\mathbf{x}})|\\
&+\sup_{\mathbf{b}'}|\log w(\mathbf{b},\mathbf{b}',\bar{\mathbf{x}})-\log w(\mathbf{b},\mathbf{b}',\mathbf{x})|\\
&+E_n''(1+\|F\|_\infty)+(L'+\|F\|_\infty)V(\mathbf{x},\mathbf{x}')
\end{aligned} \tag{30}$$

a.s. with random variables $E_n''\to 0$ independently of F, $\mathbf{b}$, $\mathbf{x}$. To obtain this, we notice that with random variables $E_n'''\to 0$

$$\begin{aligned}
&\Big|(R_n^*F)(\mathbf{b}',\mathbf{x})-\frac{1}{k_{n-1}}\sum_{i=2}^n\mathbb{E}\{F(\mathbf{b}',\mathbf{X}_{n,i})\mid\mathbf{X}_{i-1}=\mathbf{x}\}\\
&\cdot\nu_{i-1}(S_{\mathbf{x},\hat{h}_n(\mathbf{x})})\Big|\\
\leq\;&\|F\|_\infty\frac{1}{k_{n-1}}\sum_{i=2}^n\int_{S_{\mathbf{x},\hat{h}_n(\mathbf{x})}}V(\mathbf{z},\mathbf{x})P_{\mathbf{X}_{i-1}}(d\mathbf{z})\\
\leq\;&E_n'''\|F\|_\infty
\end{aligned}$$

a.s. because of Condition (i), (29) and (28), further

$$\begin{aligned}
&\frac{1}{k_{n-1}}\Big|\sum_{i=2}^n\big[\mathbb{E}\{F(\mathbf{b}',\mathbf{X}_{n,i})\mid\mathbf{X}_{i-1}=\bar{\mathbf{x}}\}\\
&\cdot\nu_{i-1}(S_{\bar{\mathbf{x}},\hat{h}_n(\bar{\mathbf{x}})})\Big|\\
&-\mathbb{E}\{F(\mathbf{b}',\mathbf{X}_{n,i})\mid\mathbf{X}_{i-1}=\mathbf{x}\}\nu_{i-1}(S_{\mathbf{x},\hat{h}_n(\mathbf{x})})\big]\Big|\\
\leq\;&\frac{1}{k_{n-1}}\|F\|_\infty V(\mathbf{x},\bar{\mathbf{x}})\sum_{i=2}^n\nu_{i-1}(S_{\mathbf{x},\hat{h}_n(\mathbf{x})})\\
&+\frac{1}{k_{n-1}}\|F\|_\infty\frac{1}{k_{n-1}}\Big|\sum_{i=2}^n\big(\nu_{i-1}(S_{\mathbf{x},\hat{h}_n(\mathbf{x})})\\
&-\nu_{i-1}(S_{\bar{\mathbf{x}},\hat{h}_n(\bar{\mathbf{x}})})\big)\Big|\\
\leq\;&\|F\|_\infty V(\mathbf{x},\bar{\mathbf{x}})(1+E_n''')
\end{aligned}$$

a.s. because of (28), especially also for $\langle\cdot\,,\,\cdot\rangle$. (27) and (30) yield

$$\begin{aligned}
&\big|F_{n+1}'(\bar{\mathbf{b}},\bar{\mathbf{x}})-F_{n+1}'(\mathbf{b},\mathbf{x})\big|\\
\leq\;&\sup_{\mathbf{b}'}|\log w(\bar{\mathbf{b}},\mathbf{b}',\bar{\mathbf{x}})-\log w(\mathbf{b},\mathbf{b}',\bar{\mathbf{x}})|\\
&+\sup_{\mathbf{b}'}|\log w(\mathbf{b},\mathbf{b}',\bar{\mathbf{x}})-\log w(\mathbf{b},\mathbf{b}',\mathbf{x})|\\
&+(L'+\|F_n'\|_\infty)V(\mathbf{x},\bar{\mathbf{x}})+E_n^{(4)}(1+\|F_n'\|_\infty)
\end{aligned} \tag{31}$$

a.s. with random variables $E_n^{(4)}\to 0$. This corresponds to (18) in the proof of Theorem 1. By (31) we obtain

$$\|F_n'\|_\infty\leq E'<\infty \tag{32}$$

with some random variable E', which corresponds to (19). Via

$$\|Q_nF-Q_nF'\|_\infty\leq+(1-\delta_n)\|F-F'\|_\infty$$

and

$$\|Q_{n+1}F-Q_nF\|_\infty\leq\frac{2L'+2\|F\|_\infty}{k_{n-1}}+(\delta_n-\delta_{n+1})\|F\|_\infty+\frac{\|F\|_\infty}{k_nk_{n-1}},$$

($F,F'\in B$), by (32) we obtain

$$\|F_{n+2}'-F_{n+1}'\|_\infty\to 0 \tag{33}$$

a.s., which corresponds to (20). We notice

$$||Q_n^* F - Q_n^* F'||_\infty \leq +(1-\delta_n)||F - F'||_\infty \tag{34}$$

$(F, F' \in B)$. Further for each $F \in C$

$$(R_n^* F)(\mathbf{b}, \mathbf{x}) \to (R_0 F)(\mathbf{b}, \mathbf{x}) \tag{35}$$

uniformly with respect to $\mathbf{b}, \mathbf{x}$, a.s., because the left hand side equals

$$\int F(\mathbf{b}, \mathbf{y})\mu(\mathbf{x} \mid d\mathbf{y}) \frac{1}{k_{n-1}} \sum_{i=2}^{n} \nu_{i-1}(S_{\mathbf{x}, \hat{h}_n(\mathbf{x})}) + o(1)$$

uniformly with respect to $\mathbf{b}, \mathbf{x}$, a.s., by (28) and by Condition (ii), (29) and once more (28). Now let $\{\delta_{n_k}\}$ be an arbitrary subsequence of $\{\delta_n\}$. We argue as in Step 2 of the proof of Theorem 1 using (31) and (32) instead of (18) and (19), respectively, and obtain as there

$$||F'_{n_{k_\ell}} - F^*||_\infty \to 0 \tag{36}$$

a.s. for a random subsequence n_{k_ℓ} of indices and a bounded random function F^* with realizations in C. We notice

$$F'_{n_{k_\ell}} + (F'_{n_{k_\ell}+1} - F'_{n_{k_\ell}}) = Q_0 F^* + (Q_{n_{k_\ell}} F'_{n_{k_\ell}} - Q_0 F^*) - W'_{n_{k_\ell}},$$

further (36), (33) and

$$\begin{aligned} & ||Q_{n_{k_\ell}} F'_{n_{k_\ell}} - Q_0 F^*||_\infty \\ \leq\ & ||Q_{n_{k_\ell}} F'_{n_{k_\ell}} - Q^*_{n_{k_\ell}} F'_{n_{k_\ell}}||_\infty + ||Q^*_{n_{k_\ell}} F'_{n_{k_\ell}} - Q^*_{n_{k_\ell}} F^*||_\infty \\ & + ||Q^*_{n_{k_\ell}} F^* - Q_0 F^*||_\infty \\ \leq\ & E_n^{(4)}(1 + ||F'_{n_{k_\ell}}||_\infty) + ||F'_{n_{k_\ell}} - F^*||_\infty + o(1) \\ \to\ & 0 \end{aligned}$$

a.s. (by (27) – compare (31) –, (32), (34), (35) and (36)). This yields

$$\lim_\ell W'_{n_{k_\ell}} + F^* = W_c^* + F^* = M_0^* F^*,$$

and thus the assertion.

Step 3. We show the assertion of Theorem 2. By (27) and (32) we obtain

$$(R_n F'_n)(\mathbf{b}_{n+1}, \mathbf{X}_n) = (R_n^* F'_n)(\mathbf{b}_{n+1}, \mathbf{X}_n) + o(1) \tag{37}$$

a.s. Further we notice that for $F'_n \in B$, $F^* \in C$ and $||F'_n - F^*||_\infty \to 0$ we have

$$\begin{aligned} & ||R_n^* F'_n - R_0 F'_n||_\infty \\ \leq\ & ||R_n^* F^* - R_0 F^*||_\infty + ||R_n^* F'_n - R_n^* F^*||_\infty \\ & + ||R_0 F'_n - R_0 F^*||_\infty \\ \leq\ & o(1) + ||F'_n - F^*||_\infty + ||F'_n - F^*||_\infty \\ \to\ & 0 \end{aligned}$$

a.s. because of (35) and (34). Thus by (36) and use of subsequences of subsequences we obtain

$$(R_n^* F'_n)(\mathbf{b}_{n+1}, \mathbf{X}_n) = (R_0 F'_n)(\mathbf{b}_{n+1}, \mathbf{X}_n) + o(1) \tag{38}$$

a.s. (37), (38) and

$$(R_0 F'_n)(\mathbf{b}_{n+1}, \mathbf{X}_n) = \mathbb{E}\{F'_n(\mathbf{b}_{n+1}, \mathbf{X}_{n+1}) \mid \mathbf{X}_1^n\}$$

yield

$$(R_n F'_n)(\mathbf{b}_{n+1}, \mathbf{X}_n) - \mathbb{E}\{F'_n(\mathbf{b}_{n+1}, \mathbf{X}_{n+1}) \mid \mathbf{X}_1^n\} \to 0$$

a.s. (compare (22)). An analogous result holds for $\langle \cdot\,, \cdot \rangle$ instead of F'_n (compare (23)). Now the assertion follows as in the final part of the proof of Theorem 1.

V. Acknowledgement

The authors wish to thank the reviewers for helpful suggestions which led to an improvement of the paper. We also thank Michael Kohler and Adam Krzyzak for the valuable discussion.

References

[1] P. Algoet, and T. Cover, "Asymptotic optimality asymptotic equipartition properties of log-optimum investments," *Annals of Probability*, 16, pp. 876-898, 1988.

[2] K. B. Athreya, and P. Ney, "A new approach to the limit theory of recurrent Markov chains," *Transactions of the American Mathematical Society*, 245, pp. 493-501, 1978.

[3] R. Bellman, *Dynamic Programming*, Princeton: Princeton University Press, 1957.

[4] R. V. Bobryk, and L. Stettner, "Discrete time portfolio selection with proportional transaction costs," *Probability and Mathematical Statistics*, 19, pp. 235-248, 1999.

[5] V. F. Farias, C. C. Moallemi, B. Van Roy, T. Weissman, "Universal reinforcement learning," *IEEE Transactions on Information Theory*, 56, pp. 2441-2454, May 2010.

[6] W. Feller, *An Introduction to Probability Theory and its Applications, Vol.1* , New York: John Wiley, 1968.

[7] E. R. Fernholz, *Stochastic Portfolio Theory,* New York: Springer, 2000.

[8] L. Györfi, W. Härdle, P. Sarda, and Ph. Vieu, *Nonparametric Curve Estimation from Time Series*, Lecture Notes in Statistics, Berlin: Springer-Verlag, 1989.

[9] Györfi, L., Kohler, M., Krzyżak, A. and Walk, H. *A Distribution-free Theory of Nonparametric Regression*, New York: Springer-Verlag, 2002.

[10] L. Györfi, G. Lugosi, and F. Udina, "Nonparametric kernel-based sequential investment strategies," *Mathematical Finance*, 16, pp. 337-357, 2006.

[11] L. Györfi, and D. Schäfer, "Nonparametric prediction," in *Advances in Learning Theory: Methods, Models and Applications*, eds. J. A. K. Suykens, G. Horváth, S. Basu, C. Micchelli, and J. Vandevalle. IOS Press, pp. 339-354, 2003.

[12] L. Györfi, G. Ottucsák and A. Urbán, "Empirical log-optimal portfolio selections: a survey" in *Machine Learning for Financial Engineering*, eds. L. Györfi, G. Ottucsák, H. Walk, pp. 79-116, Imperial College Press, 2011.

[13] L. Györfi, F. Udina, and H. Walk, "Nonparametric nearest-neighbor-based sequential investment strategies", *Statistics and Decisions*, 26, pp. 145-157, 2008.

[14] L. Györfi, and I. Vajda, "Growth optimal portfolio selection strategies with transaction costs", in *Algorithmic Learning Theory*, Freund, Y., Györfi, L., Turán, G., Zeugmann, Th. (Eds.), Lecture Notes in Artificial Intelligence, Vol. LNAI 5254, Springer, pp. 108-122, 2008.

[15] L. Györfi, and H. Walk, "Log-optimal portfolio selection strategies with proportional transaction costs," in *Machine Learning for Financial Engineering*, eds. L. Györfi, G. Ottucsák, H. Walk, Imperial College Press, pp. 117-150, 2011.

[16] G. H. Hardy, *Divergent Series*, London: Oxford University Press, 1949.

[17] O. Hernández-Lerma and J. B. Lasserre, *Discrete-Time Markov Control Processes: Basic Optimality Criteria*, New York: Springer, 1996.

[18] K. Hinderer, *Foundations of Non-Stationary Dynamic Programming with Discrete Time Parameter*, Berlin: Springer-Verlag, Berlin, 1970.

[19] M. Horváth and A. Urbán, "Growth optimal portfolio selection with short selling and leverage", in *Machine Learning for Financial Engineering*, eds. L. Györfi, G. Ottucsák, H. Walk, Imperial College Press, pp. 151-176, 2011.

[20] G. Iyengar, "Discrete time growth optimal investment with costs," *Working Paper*, 2002 http://www.columbia.edu/∼gi10/Papers/stochastic.pdf

[21] G. Iyengar, "Universal investment in markets with transaction costs," *Mathematical Finance*, 15, pp. 359-371, 2005.

[22] G. Iyengar, and T. Cover, "Growth optimal investment in horse race markets with costs," *IEEE Transactions on Information Theory*, 46, pp. 2675-2683, 2000.

[23] A. Kalai, and A.Blum, "Universal portfolios with and without transaction costs," *Proceedings of the 10th Annual Conference on Learning Theory.*, pp. 309-313, 1997.

[24] M. Kohler, A. Krzyzak, and H. Walk, "Estimation of the essential supremum of a regression function," *Statistics and Probability Letters*, 81, pp. 685-693, 2011.

[25] N. Merhav, E. Ordentlich, G. Seroussi, and M. J. Weinberger, "On sequential strategies for loss functions with memory," *IEEE Transactions on Information Theory*, 48, pp. 1947-1958, 2002.

[26] J. S. Rosenthal, "Convergence rates of Markov chains," *SIAM Review*, 37, pp. 387-405, 1995.

[27] J. Sass, and M. Schäl, "The numeraire portfolio under proportional transaction cost," Working paper, 2010.

[28] D. Schäfer, *Nonparametric Estimation for Finantial Investment under Log-Utility,* PhD Dissertation, Mathematical Institute, Universität Stuttgart, Aachen: Shaker Verlag, 2002. http://www.szit.bme.hu/∼oti/portfolio/articles/dominik.pdf

[29] W. F. Stout, *Almost Sure Convergence*, New York: Academic Press, 1974.

[30] H. Walk, and L. Zsidó, "Convergence of the Robbins-Monro method for linear problems in a Banach space," *Journal of Mathematical Analysis and Applications*, 139:152–177, 1989.

[31] K. Yosida, *Functional Analysis*, 2nd ed. Berlin: Springer-Verlag, 1968.