Statistics & Decisions 99, 901–9999 (2004)

An Asymptotic Analysis of the Mean-Variance Portfolio Selection

György Ottucsák, István Vajda

Received: Month-1 99, 2003; Accepted: Month-2 99, 2004

Summary: This paper gives an asymptotic analysis of the mean-variance (Markowitz-type) portfolio selection under mild assumptions on the market behavior. Theoretical results show the rate of underperformance of the risk aware Markowitz-type portfolio strategy in growth rate compared to the log-optimal portfolio strategy, which does not have explicit risk control. Statements are given with and without full knowledge of the statistical properties of the underlying process generating the market, under the only assumption that the market is stationary and ergodic. The experiments show how the achieved wealth depends on the coefficient of absolute risk aversion measured on past NYSE data.

1 Introduction

The goal of the Markowitz's portfolio strategy is the optimization of the asset allocation in financial markets in order to achieve optimal trade-off between the return and the risk (variance). For the *static model* (one-period model), the classical solution of the mean-variance optimization problem was given by Markowitz [16] and Merton [17]. Compared to *expected utility models*, it offers an intuitive explanation for diversification. However, most of the mean-variance analysis handles only static models contrary to the expected utility models, whose literature is rich in *multiperiod models*.

In the multiperiod models (investment strategies) the investor is allowed to rebalance his portfolio at the beginning of each trading period. More precisely, investment strategies use information collected from the past of the market and determine a portfolio, that is, a way to distribute the current capital among the available assets. The goal of the investor is to maximize his utility in the long run. Under the assumption that the daily price relatives form a stationary and ergodic process the *asymptotic growth rate of the wealth* has a well-defined maximum which can be achieved in full knowledge of the distribution of the entire process, see Algoet and Cover [3]. This maximum can be achieved by the *log-optimal strategy*, which maximizes the conditional expectation of *log-utility*.

AMS 1991 subject classification: 90A09, 90A10
Key words and phrases: sequential investment, kernel-based estimation, mean-variance investment, log-optimal investment

For an investor plausibly arises the following question: how much the loss in the average growth rate compared to the asymptotically best rate if a mean-variance portfolio optimization is followed in each trading period.

The first theoretical result connecting the expected utility and mean-variance analysis was shown by Tobin [19], for quadratic utility function. Grauer [9] compared the log-optimal strategies and the mean-variance analysis in the one-period model for various specifications of the state return distributions and his experiments revealed that these two portfolios have almost the same performance when the returns come from the normal distribution. Kroll, Levy and Markowitz [15] conducted a similar study. Merton [17] developed a continuous time mean-variance analysis and showed that log-optimal portfolio is instantaneously mean-variance efficient when asset prices are log-normal. Hakansson and Ziemba [13] followed another method in defining a dynamic mean-variance setting, where the log-optimal portfolio can be chosen as a risky mutual fund. They considered a finite time horizon model and the asset price behavior was determined by a Wiener-process.

In this paper, we follow a similar approach to [17], however in discrete time setting for general stationary and ergodic processes.

To determine a Markowitz-type portfolio, knowledge of the infinite dimensional statistical characterization of the process is required. In contrast, for log-optimal portfolio setting *universally consistent* strategies are known, which can achieve growth rate achievable in the full knowledge of distributions, however without knowing these distributions. These strategies are universal with respect to the class of all stationary and ergodic processes as it was proved by Algoet [1]. Györfi and Schäfer [11], Györfi, Lugosi, Udina [10] constructed a practical kernel based algorithm for the same problem.

We present a strategy achieving the same growth rate as the Markowitz-type strategy for the general class of stationary and ergodic processes, however without assuming the knowledge of the statistical characterization of the process. This strategy is risk averse in the sense that in each time period it carries out the mean-variance optimization.

We also present an experimental performance analysis for the data sets of New York Stock Exchange (NYSE) spanning a twenty-two-year period with thirty six stocks included, which set was presented in [10].

The rest of the paper is organized as follows. In Section 2 the mathematical model is described. The investigated portfolio strategies are defined in Section 3. In the next section a lower bound on the performance of Markowitz-type strategy is shown. Section 5 presents the kernel-based Markowitz-type sequential investment strategy and its main consistency properties. Numerical results based on NYSE data set are shown in Section 6. The proofs are given in Section 7.

2 Setup, the market model

The model of stock market investigated in this paper is the one considered, among others, by Breiman [7], Algoet and Cover [3]. Consider a market of d assets. A *market vector* $\mathbf{x} = (x^{(1)}, \dots, x^{(d)}) \in \mathbb{R}_+^d$ is a vector of d nonnegative numbers representing price relatives for a given trading period. That is, the j-th component $x^{(j)} \geq 0$ of $\mathbf{x}$ expresses the ratio of the opening prices of asset j. In other words, $x^{(j)}$ is the factor by which capital invested in the j-th asset grows during the trading period.

The investor is allowed to diversify his capital at the beginning of each trading period according to a portfolio vector $\mathbf{b} = (b^{(1)}, \dots b^{(d)})$. 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. Let S_0 denote the investor's initial capital. Then at the end of the trading period the investor's wealth becomes

$$S_1 = S_0 \sum_{j=1}^{d} b^{(j)} x^{(j)} = S_0 \langle \mathbf{b}\, , \, \mathbf{x} \rangle \, ,$$

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

The evolution of the market in time is represented by a sequence of market vectors $\mathbf{x}_1, \mathbf{x}_2, \dots \in \mathbb{R}_+^d$, where the j-th component $x_i^{(j)}$ of $\mathbf{x}_i$ denotes the amount obtained after investing a unit capital in the j-th asset on the i-th trading period. For $j \leq i$ we abbreviate by $\mathbf{x}_j^i$ the array of market vectors $(\mathbf{x}_j, \dots, \mathbf{x}_i)$ and 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

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

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.

Starting with an initial wealth S_0, after n trading periods, the investment strategy $\mathbf{B}$ achieves the wealth

$$S_n = S_0 \prod_{i=1}^{n} \left\langle \mathbf{b}(\mathbf{x}_1^{i-1}) \, , \, \mathbf{x}_i \right\rangle = S_0 e^{\sum_{i=1}^{n} \log \left\langle \mathbf{b}(\mathbf{x}_1^{i-1}) \, , \, \mathbf{x}_i \right\rangle} = S_0 e^{n W_n(\mathbf{B})}.$$

where $W_n(\mathbf{B})$ denotes the *average growth rate*

$$W_n(\mathbf{B}) \stackrel{\text{def}}{=} \frac{1}{n} \sum_{i=1}^{n} \log \left\langle \mathbf{b}(\mathbf{x}_1^{i-1}) \, , \, \mathbf{x}_i \right\rangle \, .$$

Obviously, maximization of $S_n = S_n(\mathbf{B})$ and maximization of $W_n(\mathbf{B})$ are equivalent.

In this paper we assume that the market vectors are realizations of a random process, and describe a statistical model. Our view is completely nonparametric in that the only assumption we use is that the market is stationary and ergodic, allowing arbitrarily complex distributions. More precisely, assume that $\mathbf{x}_1, \mathbf{x}_2, \dots$ are realizations of the random vectors $\mathbf{X}_1, \mathbf{X}_2, \dots$ drawn from the vector-valued stationary and ergodic process $\{\mathbf{X}_n\}_{-\infty}^{\infty}$. The sequential investment problem, under these conditions, has been considered by, e.g., Breiman [7], Algoet and Cover [3], Algoet [1, 2], Györfi and Schäfer [11], Györfi, Lugosi, Udina [10].

3 The Markowitz-type and the log-optimal portfolio selection

3.1 The Markowitz-type portfolio selection

In his seminal paper Markowitz [16] used expected utility and the investigation was restricted to single period investment with i.i.d returns. In contrary we work with a multi-period model (investment strategy) and we assume general stationary and ergodic model for the market returns. Consequently it is more adequate for us to use conditional expected utility. In the special case of i.i.d. returns it is simplified to the standard expected utility. For the sake of the distinction from the standard Markowitz approach we will call our utility function *Markowitz-type utility function*.

Let the following formula define the conditional expected value of the Markowitz-type utility function:

$$\begin{aligned}\mathbb{E}\left\{U_M(\left\langle \mathbf{b}(\mathbf{X}_1^{n-1})\,,\,\mathbf{X}_n\right\rangle,\lambda)|\mathbf{X}_1^{n-1}\right\} &\stackrel{\text{def}}{=} \mathbb{E}\{\left\langle \mathbf{b}(\mathbf{X}_1^{n-1})\,,\,\mathbf{X}_n\right\rangle|\mathbf{X}_1^{n-1}\}\\ &\quad-\lambda\mathbf{Var}\left\{\left\langle \mathbf{b}(\mathbf{X}_1^{n-1})\,,\,\mathbf{X}_n\right\rangle|\mathbf{X}_1^{n-1}\right\},\end{aligned}$$

where $\mathbf{X}_n$ is the market vector for n-th day, $\mathbf{b}(\mathbf{X}_1^{n-1}) \in \Delta_d$ and $\lambda \in [0,\infty)$ is the constant coefficient of absolute risk aversion of the investor. The conditional expected value of the Markowitz-type utility function can be expressed after some algebra in following form:

$$\begin{aligned}&\mathbb{E}\{U_M(\left\langle \mathbf{b}(\mathbf{X}_1^{n-1})\,,\,\mathbf{X}_n\right\rangle,\lambda)|\mathbf{X}_1^{n-1}\}\\ &= (1-2\lambda)\mathbb{E}\left\{\left\langle \mathbf{b}(\mathbf{X}_1^{n-1})\,,\,\mathbf{X}_n\right\rangle-1|\mathbf{X}_1^{n-1}\right\}-\lambda\mathbb{E}\{(\left\langle \mathbf{b}(\mathbf{X}_1^{n-1})\,,\,\mathbf{X}_n\right\rangle-1)^2|\mathbf{X}_1^{n-1}\}\\ &\quad+1-\lambda+\lambda\mathbb{E}^2\left\{\left\langle \mathbf{b}(\mathbf{X}_1^{n-1})\,,\,\mathbf{X}_n\right\rangle|\mathbf{X}_1^{n-1}\right\}.\end{aligned}$$

Accordingly let the Markowitz-type utility function be defined as follows

$$\begin{aligned}&U_M(\left\langle \mathbf{b}(\mathbf{X}_1^{n-1})\,,\,\mathbf{X}_n\right\rangle,\lambda)\\ &\stackrel{\text{def}}{=} (1-2\lambda)(\left\langle \mathbf{b}(\mathbf{X}_1^{n-1})\,,\,\mathbf{X}_n\right\rangle-1)-\lambda(\left\langle \mathbf{b}(\mathbf{X}_1^{n-1})\,,\,\mathbf{X}_n\right\rangle-1)^2+1-\lambda\\ &\quad+\lambda\mathbb{E}^2\{\left\langle \mathbf{b}(\mathbf{X}_1^{n-1})\,,\,\mathbf{X}_n\right\rangle|\mathbf{X}_1^{n-1}\}.\end{aligned}$$

Hence the Markowitz-type portfolio strategy be defined by $\bar{\mathbf{B}}_\lambda^* = \{\bar{\mathbf{b}}_\lambda^*(\cdot)\}$, where

$$\bar{\mathbf{b}}_\lambda^*(\mathbf{X}_1^{n-1}) = \arg\max_{\mathbf{b}\in\Delta_d}\mathbb{E}\left\{U_M(\left\langle \mathbf{b}(\mathbf{X}_1^{n-1})\,,\,\mathbf{X}_n\right\rangle,\lambda)|\mathbf{X}_1^{n-1}\right\}.$$

Let $\bar{S}_{n,\lambda}^* = S_n(\bar{\mathbf{B}}_\lambda^*)$ denote the capital achieved by Markowitz-type portfolio strategy $\bar{\mathbf{B}}_\lambda^*$, after n trading periods.

3.2 The log-optimal portfolio selection

Now we briefly introduce the log-optimal portfolio selection. The fundamental limits, first published in [3], [1, 2] reveal that the so-called *log-optimal portfolio* $\mathbf{B}^* = \{\mathbf{b}^*(\cdot)\}$

is the best possible choice for the maximization of S_n. More precisely, on trading period n let $\mathbf{b}^*(\cdot)$ be such that

$$\mathbf{b}^*(\mathbf{X}_1^{n-1}) = \arg\max_{\mathbf{b}\in\Delta_d} \mathbb{E}\left\{ \log\left\langle \mathbf{b}(\mathbf{X}_1^{n-1}), \mathbf{X}_n \right\rangle \middle| \mathbf{X}_1^{n-1} \right\}.$$

If $S_n^* = S_n(\mathbf{B}^*)$ denotes the capital achieved by a log-optimal portfolio strategy $\mathbf{B}^*$, after n trading periods, then for any other investment strategy $\mathbf{B}$ with capital $S_n = S_n(\mathbf{B})$ and for any stationary and ergodic process $\{\mathbf{X}_n\}_{-\infty}^{\infty}$,

$$\limsup_{n\to\infty} \frac{1}{n}\log\frac{S_n}{S_n^*} \le 0 \quad \text{a.s.}$$

and

$$\lim_{n\to\infty} \frac{1}{n}\log S_n^* = W^* \quad \text{a.s.},$$

where

$$W^* = \mathbb{E}\left\{ \log\left\langle \mathbf{b}^*(\mathbf{X}_{-\infty}^{-1}), \mathbf{X}_0 \right\rangle \right\}$$

is the maximal possible growth rate of any investment strategy. Thus, (almost surely) no investment strategy can have a higher growth rate than a log-optimal portfolio.

3.3 Connection of the Markowitz-type and the log-optimal portfolio: an intuitive argument

As the main tool we apply the *semi-log* function introduced by Györfi, Urbán and Vajda [12]. The semi-log function is the second order Taylor expansion of $\log z$ at $z = 1$

$$h(z) \stackrel{\text{def}}{=} z - 1 - \frac{1}{2}(z-1)^2.$$

The semi log-optimal portfolio strategy is $\tilde{\mathbf{B}}^* = \{\tilde{\mathbf{b}}^*\}$, where

$$\tilde{\mathbf{b}}^*(\mathbf{X}_1^{n-1}) \stackrel{\text{def}}{=} \arg\max_{\mathbf{b}\in\Delta_d} \mathbb{E}\left\{ h\left(\left\langle \mathbf{b}(\mathbf{X}_1^{n-1}), \mathbf{X}_n \right\rangle\right) | \mathbf{X}_1^{n-1} \right\}.$$

Applying the semi-log approximation we get:

$$\begin{aligned}
&\mathbb{E}\left\{ \log\left\langle \mathbf{b}(\mathbf{X}_1^{n-1}), \mathbf{X}_n \right\rangle \middle| \mathbf{X}_1^{n-1} \right\} \\
&\approx \mathbb{E}\left\{ h\left(\left\langle \mathbf{b}(\mathbf{X}_1^{n-1}), \mathbf{X}_n \right\rangle\right) | \mathbf{X}_1^{n-1} \right\} \\
&= \mathbb{E}\left\{ \left(\left\langle \mathbf{b}(\mathbf{X}_1^{n-1}), \mathbf{X}_n \right\rangle - 1\right) - \frac{1}{2}\left(\left\langle \mathbf{b}(\mathbf{X}_1^{n-1}), \mathbf{X}_n \right\rangle - 1\right)^2 \middle| \mathbf{X}_1^{n-1} \right\}. \qquad (3.1)
\end{aligned}$$

According to formula (3.1) we introduce a few simplifying notations for the conditional expected value

$$E_n(\mathbf{b}) \stackrel{\text{def}}{=} \mathbb{E}\left\{ \left\langle \mathbf{b}(\mathbf{X}_1^{n-1}), \mathbf{X}_n \right\rangle \middle| \mathbf{X}_1^{n-1} \right\},$$

for the conditional second order moment

$$E_n(\mathbf{b})^2 \stackrel{\text{def}}{=} \mathbb{E}\left\{ \left(\left\langle \mathbf{b}(\mathbf{X}_1^{n-1}), \mathbf{X}_n \right\rangle\right)^2 \middle| \mathbf{X}_1^{n-1} \right\},$$

and finally for the variance

$$V_n(\mathbf{b}) \stackrel{\mathrm{def}}{=} E_n(\mathbf{b})^2 - E_n^2(\mathbf{b}).$$

Hence we get

$$\begin{aligned}
\tilde{\mathbf{b}}^*(\mathbf{X}_1^{n-1}) &= \arg\max_{\mathbf{b}\in\Delta_d} \left(2E_n(\mathbf{b}) - \frac{1}{2}E_n(\mathbf{b})^2 - \frac{3}{2}\right)\\
&= \arg\max_{\mathbf{b}\in\Delta_d} \left(2E_n(\mathbf{b}) - \frac{1}{2}\left(E_n(\mathbf{b})^2 - E_n^2(\mathbf{b})\right) - \frac{1}{2}E_n^2(\mathbf{b})\right)\\
&= \arg\max_{\mathbf{b}\in\Delta_d} \left(E_n(\mathbf{b})(4 - E_n(\mathbf{b})) - V_n(\mathbf{b})\right)\\
&= \arg\max_{\mathbf{b}\in\Delta_d} \left(\frac{E_n(\mathbf{b})}{\lambda_n} - V_n(\mathbf{b})\right) \stackrel{\mathrm{def}}{=} \bar{\mathbf{b}}^*_{\lambda_n}(\mathbf{X}_1^{n-1}),
\end{aligned}$$

where

$$\lambda_n \stackrel{\mathrm{def}}{=} \frac{1}{4 - E_n(\mathbf{b})}$$

is the coefficient of risk aversion of the investor, which is here a function of conditional expected value. Note that parameter λ_n dynamically changes in time, which means as if the Markowitz's investor would dynamically adjust his coefficient of absolute risk aversion to the past performance of the portfolio.

So far we sketched a basic relationship between the Markowitz-type and semi-log portfolio selection. The relationship between the log-optimal and semi-log optimal approach was examined in [12]. We present formal, rigorous analysis for the comparison of the investment strategies in the subsequent sections.

4 Comparison of Markowitz-type and log-optimal portfolio selection in case of known distribution

Naturally arises the question, how much we lose in the long run if we are risk averse investors compared to the log-optimal investment. The theorem in this section shows an upper bound on the loss in case of known distribution, if our hypothetical investor is risk averse with parameter λ. More precisely for an arbitrary λ we give the growth rate of the Markowitz-type strategy compared to the maximal possible growth rate (which is achieved by the log-optimal investment) in asymptotic sense.

The only assumptions, which we use in our analysis is that the market vectors $\{\mathbf{X}_n\}_{-\infty}^{\infty}$ come from a stationary and ergodic process, for which

$$a \le X_n^{(j)} \le \frac{1}{a} \tag{4.1}$$

for all $j = 1, \dots, d$, where $0 < a < 1$.

Theorem 4.1 *For any stationary and ergodic process $\{\mathbf{X}_n\}_{-\infty}^{\infty}$, for which (4.1) holds, for all $\lambda \in \left[0, \frac{1}{2}\right)$*

$$\begin{aligned}
W^* &\geq \liminf_{n\to\infty} \frac{1}{n} \log \bar{S}^*_{n,\lambda} \\
&\geq W^* + \frac{2\lambda a^{-1}}{1-2\lambda} \mathbb{E}\left\{ \min_m \mathbb{E}\left\{ 1 + \log(X_0^{(m)}) - X_0^{(m)} \middle| \mathbf{X}_{-\infty}^{-1} \right\} \right\} \\
&\quad - \left| \frac{2\lambda - \frac{1}{2}}{1-2\lambda} \right| \mathbb{E}\left\{ \max_m \mathbb{E}\left\{ (X_0^{(m)} - 1)^2 \middle| \mathbf{X}_{-\infty}^{-1} \right\} - \left(\min_m \mathbb{E}\left\{ |X_0^{(m)} - 1| \middle| \mathbf{X}_{-\infty}^{-1} \right\} \right)^2 \right\} \\
&\quad - \frac{a^{-3}+1}{3(1-2\lambda)} \mathbb{E}\left\{ \max_m \mathbb{E}\left\{ \left| X_0^{(m)} - 1 \right|^3 \middle| \mathbf{X}_{-\infty}^{-1} \right\} \right\} \qquad \text{a.s.}
\end{aligned}$$

Remark 4.2 Under the assumption of Theorem 4.1 with slight modification of the proof one can show same result for all $\lambda \in \left(\frac{1}{2}, \infty\right)$.

Remark 4.3 We have made some experiments on pairs of stocks of NYSE. For IBM and Coca-Cola we got the following estimates on the magnitude of the terms on the right-hand side of the statement of Theorem 4.1 respectively 10^{-5}, 10^{-4}, 10^{-6}. Furthermore the order of W^* is $9 \cdot 10^{-4}$. With optimized λ one could push the difference between W^* and $\liminf_{n\to\infty} \frac{1}{n} \log \bar{S}^*_{n,\lambda}$ below 1% of W^*.

5 Nonparametric kernel-based Markowitz-type strategy

To determine a Markowitz-type portfolio, knowledge of the infinite dimensional statistical characterization of the process is required. In contrast, for log-optimal portfolio setting *universally consistent* strategies are known, which can achieve growth rate achievable in the full knowledge of distributions, however without knowing these distributions. Roughly speaking, an investment strategy $\mathbf{B}$ is called universally consistent with respect to a class of stationary and ergodic processes $\{\mathbf{X}_n\}_{-\infty}^{\infty}$, if for each process in the class,

$$\lim_{n\to\infty} \frac{1}{n} \log S_n(\mathbf{B}) = W^* \quad \text{a.s.}$$

The surprising fact that there exists a strategy, universally consistent with respect to the class of all stationary and ergodic processes with $\mathbb{E}\left\{ |\log X^{(j)}| \right\} < \infty$ for all $j = 1, \ldots, d$, was first proved by Algoet [1] and by Györfi and Schäfer [11]. Györfi, Lugosi, Udina [10] introduced kernel-based strategies, here we describe a "moving-window" version, corresponding to an uniform kernel function in order to keep the notations simple.

Define an infinite array of experts $\mathbf{H}^{(k,\ell)} = \{\mathbf{h}^{(k,\ell)}(\cdot)\}$, where k, ℓ are positive integers. For fixed positive integers k, ℓ, choose the radius $r_{k,\ell} > 0$ such that for any fixed k,

$$\lim_{\ell\to\infty} r_{k,\ell} = 0 \,.$$

Then, for $n > k+1$, define the expert $\mathbf{h}^{(k,\ell)}$ as follows. Let J_n be the locations of matches:

$$J_n = \left\{ k < i < n : \|\mathbf{x}_{i-k}^{i-1} - \mathbf{x}_{n-k}^{n-1}\| \leq r_{k,\ell} \right\},$$

where $\|\cdot\|$ denotes the Euclidean norm. Put

$$\mathbf{h}^{(k,\ell)}(\mathbf{x}_1^{n-1}) = \arg\max_{\mathbf{b}\in\Delta_d} \prod_{\{i\in J_n\}} \langle \mathbf{b}\,,\, \mathbf{x}_i\rangle \ , \tag{5.1}$$

if the product is non-void, and $\mathbf{b}_0 = (1/d, \ldots, 1/d)$ otherwise.

These experts are mixed as follows: let $\{q_{k,\ell}\}$ be a probability distribution over the set of all pairs (k, ℓ) of positive integers such that for all k, ℓ, $q_{k,\ell} > 0$. If $S_n(\mathbf{H}^{(k,\ell)})$ is the capital accumulated by the elementary strategy $\mathbf{H}^{(k,\ell)}$ after n periods when starting with an initial capital $S_0 = 1$, then, after period n, the investor's capital becomes

$$S_n(\mathbf{B}) = \sum_{k,\ell} q_{k,\ell} S_n(\mathbf{H}^{(k,\ell)}) \ . \tag{5.2}$$

Györfi, Lugosi and Udina [10] proved that the kernel-based portfolio scheme $\mathbf{B}$ is universally consistent with respect to the class of all ergodic processes such that $\mathbb{E}\{|\log X^{(j)}|\} < \infty$, for $j = 1, 2, \ldots d$.

Equation (5.1) can be formulated in an equivalent form:

$$\mathbf{h}^{(k,\ell)}(\mathbf{x}_1^{n-1}) = \arg\max_{\mathbf{b}\in\Delta_d} \sum_{\{i\in J_n\}} \log \langle \mathbf{b}\,,\, \mathbf{x}_i\rangle \ .$$

Next we introduce the kernel-based Markowitz-type experts $\bar{\mathbf{H}}_\lambda^{(k,\ell)} = \{\bar{\mathbf{h}}_\lambda^{(k,\ell)}(\cdot)\}$ as follows:

$$\bar{\mathbf{h}}_\lambda^{(k,\ell)}(\mathbf{x}_1^{n-1}) = \arg\max_{\mathbf{b}\in\Delta_d} \Bigg((1-2\lambda) \sum_{\{i\in J_n\}} (\langle \mathbf{b}\,,\, \mathbf{x}_i\rangle - 1) - \lambda \sum_{\{i\in J_n\}} (\langle \mathbf{b}\,,\, \mathbf{x}_i\rangle - 1)^2 + \frac{\lambda}{|J_n|} \left(\sum_{\{i\in J_n\}} \langle \mathbf{b}\,,\, \mathbf{x}_i\rangle \right)^2 \Bigg) \ . \tag{5.3}$$

The Markowitz-type kernel-based strategy $\bar{\mathbf{B}}_\lambda$ is the mixture of the experts $\{\bar{\mathbf{H}}_\lambda^{(k,\ell)}\}$ according to (5.2).

Theorem 5.1 *For any stationary and ergodic process $\{\mathbf{X_n}\}_{-\infty}^{\infty}$, for which (4.1) holds, for $\bar{S}_{n,\lambda} = S_n(\bar{\mathbf{B}}_\lambda)$ and for all $\lambda \in \left[0, \frac{1}{2}\right)$ we get*

$$\begin{aligned}
&\liminf_{n\to\infty} \frac{1}{n} \bar{S}_{n,\lambda} \\
&\geq W^* + \frac{2\lambda a^{-1}}{1-2\lambda} \mathbb{E}\left\{ \min_m \mathbb{E}\left\{ 1 + \log(X_0^{(m)}) - X_0^{(m)} \middle| \mathbf{X}_{-\infty}^{-1} \right\} \right\} \\
&\quad - \left| \frac{2\lambda - \frac{1}{2}}{1-2\lambda} \right| \mathbb{E}\left\{ \max_m \mathbb{E}\left\{ (X_0^{(m)} - 1)^2 \middle| \mathbf{X}_{-\infty}^{-1} \right\} - \left(\min_m \mathbb{E}\left\{ |X_0^{(m)} - 1| \middle| \mathbf{X}_{-\infty}^{-1} \right\} \right)^2 \right\} \\
&\quad - \frac{a^{-3}+1}{3(1-2\lambda)} \mathbb{E}\left\{ \max_m \mathbb{E}\left\{ \left| X_0^{(m)} - 1 \right|^3 \middle| \mathbf{X}_{-\infty}^{-1} \right\} \right\} \qquad \text{a.s.}
\end{aligned}$$

Remark 5.2 Under the assumption of Theorem 5.1 with slight modification of the proof one can show same result for all $\lambda \in \left(\frac{1}{2}, \infty\right)$.

6 Simulation

The theoretical results assume stationarity and ergodicity of the market. No test is known, which could decide whether a market satisfies these properties or not. Therefore the practical usefulness of these assumptions should be judged based on the numerical results the investment strategies lead to.

We tested the investment strategies on a standard set of New York Stock Exchange data used by Cover [8], Singer [18], Hembold, Schapire, Singer, and Warmuth [14], Blum and Kalai [4], Borodin, El-Yaniv, and Gogan [5], and others. The NYSE data set includes daily prices of 36 assets along a 22-year period (5651 trading days) ending in 1985.

We show the wealth achieved by the strategies by investing in the pairs of NYSE stocks used in Cover [8]: Iroqouis-Kin Ark, Com. Met.-Mei. Corp., Com. Met-Kin Ark and IBM-Coca-Cola. The results of the simulation are shown in the Table 6.1.

All the proposed strategies use an infinite array of experts. In practice we take a finite array of size $K \times L$. In all cases select $K = 5$ and $L = 10$. We choose the uniform distribution $\{q_{k,\ell}\} = 1/(KL)$ over the experts in use, and the radius

$$r_{k,\ell}^2 = 0.0001 \cdot d \cdot k \cdot \ell,$$

($k = 1, \ldots, K$ and $\ell = 1, \ldots, L$).

Table 6.1 shows the performance of the nonparametric kernel based Markowitz-type strategy for different values of parameter λ. Note that the Table 6.1 contains a row with parameter $\lambda = 0.5$ which is not covered in the Theorem 5.1, however kernel-based Markowitz-type strategy can be used in this case too. The k and ℓ parameters of the *best performing experts* given in columns $2-5$ are different: (2,10), (3,10), (2,8) and (1,1) respectively. The best performance of pairs of stocks was attained at different values of parameter λ (underlined in Table 6.1). The average in Table 6.1 means the performance of the nonparametric kernel based Markowitz-type strategies averaging through λ. Log-optimal and semi-log-optimal denote the performance of the best performing k and ℓ expert of the log-optimal investment and of the semi-log investment for the given pairs of stocks.

Note that the wealth is achieved by the Markowitz-type strategy with parameter value $\lambda = 0$ is not the best overall, although it does not consider the market risk of stocks which it invests in.

Note that the presence of the stock Kin Ark makes the wealth of these strategies explode as it was noted in [10]. This is interesting, since the overall growth of Kin Ark in the reported period is quite modest. The reason is that somehow the variations of the price relatives of this asset turn out to be well predictable by at least one expert and that suffices to produce this explosive growth.

Markowitz	$S_n(\bar{\mathbf{H}}_\lambda^{(2,10)})$	$S_n(\bar{\mathbf{H}}_\lambda^{(3,10)})$	$S_n(\bar{\mathbf{H}}_\lambda^{(2,8)})$	$S_n(\bar{\mathbf{H}}_\lambda^{(1,1)})$
Pairs of stocks.	Iro-Kin	Com-Mei	Com-Kin	IBM-Cok
λ=0.00	2.61e+11	7.13e+03	1.75e+12	1.52e+02
0.05	2.75e+11	6.73e+03	1.73e+12	1.57e+02
0.10	2.51e+11	5.74e+03	1.76e+12	1.62e+02
0.15	2.45e+11	5.44e+03	1.61e+12	1.67e+02
0.20	2.60e+11	5.87e+03	1.56e+12	1.69e+02
0.25	2.97e+11	6.12e+03	1.54e+12	1.72e+02
0.30	3.09e+11	5.66e+03	1.57e+12	1.73e+02
0.33	3.26e+11	5.47e+03	1.75e+12	1.73e+02
0.35	3.32e+11	5.46e+03	1.67e+12	1.73e+02
0.40	3.76e+11	5.45e+03	1.70e+12	1.73e+02
0.45	3.62e+11	5.12e+03	1.85e+12	1.79e+02
0.50	3.44e+11	4.82e+03	1.92e+12	1.83e+02
0.55	3.23e+11	4.07e+03	1.74e+12	1.88e+02
0.60	2.64e+11	3.28e+03	1.49e+12	1.94e+02
0.65	2.05e+11	2.62e+03	1.12e+12	2.04e+02
0.70	1.49e+11	1.97e+03	7.40e+11	2.13e+02
0.75	7.30e+10	1.53e+03	3.65e+11	2.20e+02
0.80	1.80e+10	1.19e+03	9.15e+10	2.23e+02
0.85	1.18e+09	7.02e+02	4.17e+09	2.05e+02
0.90	8.68e+06	3.30e+02	2.01e+07	1.70e+02
0.95	1.73e+04	1.38e+02	5.83e+04	7.59e+01
Average	2.22e+11	4040	1.24e+12	177
Log-optimal	$S_n(\mathbf{H}^{(2,10)})$	$S_n(\mathbf{H}^{(3,10)})$	$S_n(\mathbf{H}^{(2,8)})$	$S_n(\mathbf{H}^{(1,1)})$
	3.6e+11	4765	1.9e+12	182.4
Semi-log-opt.	$S_n(\tilde{\mathbf{H}}^{(2,10)})$	$S_n(\tilde{\mathbf{H}}^{(3,10)})$	$S_n(\tilde{\mathbf{H}}^{(2,8)})$	$S_n(\tilde{\mathbf{H}}^{(1,1)})$
	3.6e+11	4685	1.9e+12	182.6

Table 6.1 Wealth achieved by the strategies by investing in the pairs of NYSE stocks used in Cover [8].

7 Proofs

The proof of Theorem 4.1 uses the following two auxiliary results and four other lemmas. The first is known as Breiman's generalized ergodic theorem.

Lemma 7.1 (BREIMAN [6]). *Let $Z = \{Z_i\}_{-\infty}^{\infty}$ be a stationary and ergodic process. For each positive integer i, let T^i denote the operator that shifts any sequence $\{\dots, z_{-1}, z_0, z_1, \dots\}$ by i digits to the left. Let $f_1, f_2, \dots$ be a sequence of real-valued functions such that* $\lim_{n\to\infty} f_n(Z) = f(Z)$ *almost surely for some function f. Assume* *that* $\mathbb{E}\{\sup_n |f_n(Z)|\} < \infty$. *Then*

$$\lim_{n\to\infty} \frac{1}{n}\sum_{i=1}^{n} f_i(T^i Z) = \mathbb{E}\{f(Z)\} \qquad \text{a.s.}$$

The next lemma is a slight modifications of the results due to Algoet and Cover [3, Theorems 3 and 4]. The modified statements are in Györfi, Urbán and Vajda [12].

Lemma 7.2 *Let* $\mathbf{Q}_{n\in\mathcal{N}\cup\{\infty\}}$ *be a family of regular probability distributions over the set* $\mathbb{R}_+^d$ *of all market vectors such that* $a \le U_n^{(j)} \le \frac{1}{a}$ *for any coordinate of a random market vector where* $0 < a < 1$ *and* $\mathbf{U}_n = (U_n^{(1)}, \ldots, U_n^{(d)})$ *distributed according to* $\mathbf{Q}_n$. *Let* $h, g \in C_0[a, \frac{1}{a}]$, *where* C_0 *denotes the set of continuous functions. In addition, let* $\bar{\mathbf{B}}^*(\mathbf{Q}_n)$ *be the set of all Markowitz-type portfolios with respect to* $\mathbf{Q}_n$, *that is, the set of all portfolios* $\mathbf{b}$ *that attain* $\max_{\mathbf{b}\in\Delta_d}\{\mathbb{E}_{\mathbf{Q}_n}\{h\langle \mathbf{b}\,,\,\mathbf{U}_n\rangle\} + \mathbb{E}^2_{\mathbf{Q}_n}\{g\langle \mathbf{b}\,,\,\mathbf{U}_n\rangle\}\}$. *Consider an arbitrary sequence* $\mathbf{b}_n \in \bar{\mathbf{B}}^*(\mathbf{Q}_n)$. *If*

$$\mathbf{Q}_n \to \mathbf{Q}_\infty \quad \textit{weakly as } n\to\infty$$

then, for $\mathbf{Q}_\infty$*-almost all* $\mathbf{u}$,

$$\lim_{n\to\infty} \langle \mathbf{b}_n\,,\,\mathbf{u}\rangle \to \left\langle \bar{\mathbf{b}}^*\,,\,\mathbf{u}\right\rangle$$

where the right-hand side is constant as $\bar{\mathbf{b}}^*$ *ranges over* $\bar{\mathbf{B}}^*(\mathbf{Q}_\infty)$.

For the proof of Theorem 4.1 we need the following lemmas:

Lemma 7.3 *For any stationary and ergodic process* $\{\mathbf{X_n}\}_{-\infty}^{\infty}$, *for which (4.1) holds and* $p \in C_0[a, \frac{1}{a}]$, *we get*

$$\lim_{n\to\infty} \frac{1}{n}\sum_{i=1}^{n} \max_j \mathbb{E}\left\{ p\left(X_i^{(j)}\right) \middle| \mathbf{X}_1^{i-1}\right\} = \mathbb{E}\left\{ \max_j \mathbb{E}\left\{ p\left(X_0^{(j)}\right) \middle| \mathbf{X}_{-\infty}^{-1}\right\}\right\} \qquad \text{a.s.}$$

Proof. Let us introduce notations

$$\bar{w}_n \stackrel{\text{def}}{=} \max_j \mathbb{E}\left\{ p(X_0^{(j)}) \mid \mathbf{X}_{-n+1}^{-1}\right\}$$

where $n = 1, 2, \ldots$ and

$$g(\mathbf{X}_n) \stackrel{\text{def}}{=} \left(p(X_n^{(1)}), \ldots, p(X_n^{(d)})\right).$$

Note that

$$\max_{\mathbf{b}\in\Delta_d} \mathbb{E}\left\{ \left\langle \mathbf{b}(\mathbf{X}_{-n+1}^{-1})\,,\,g(\mathbf{X}_0)\right\rangle \mid \mathbf{X}_{-n+1}^{-1}\right\} = \bar{w}_n \tag{7.1}$$

and $\bar{w}_n$ is measurable with respect to $\mathbf{X}_{-n+1}^{-1}$.

First we show that $\{\bar{w}_n\}$ is a sub-martingale, that is, $\mathbb{E}\left\{\bar{w}_{n+1} \mid \mathbf{X}_{-n+1}^{-1}\right\} \geq \bar{w}_n$. If a portfolio is $\mathbf{X}_{-n+1}^{-1}$-measurable, then it is also $\mathbf{X}_{-n}^{-1}$-measurable, therefore we obtain

$$\begin{aligned}
\bar{w}_n &= \max_{\mathbf{b}\in\Delta_d} \mathbb{E}\left\{\langle \mathbf{b}\,,\, g(\mathbf{X}_0)\rangle \mid \mathbf{X}_{-n+1}^{-1}\right\} \\
&= \max_{\mathbf{b}\in\Delta_d} \mathbb{E}\left\{\mathbb{E}\left\{\langle \mathbf{b}\,,\, g(\mathbf{X}_0)\rangle \mid \mathbf{X}_{-n}^{-1}\right\} \mid \mathbf{X}_{-n+1}^{-1}\right\} \\
&\leq \mathbb{E}\left\{\max_{\mathbf{b}\in\Delta_d} \mathbb{E}\left\{\langle \mathbf{b}\,,\, g(\mathbf{X}_0)\rangle \mid \mathbf{X}_{-n}^{-1}\right\} \mid \mathbf{X}_{-n+1}^{-1}\right\} \\
&= \mathbb{E}\left\{\bar{w}_{n+1} \mid \mathbf{X}_{-n+1}^{-1}\right\},
\end{aligned}$$

where in the last equation we applied formula (7.1). Thus $\bar{w}_n$ is a submartingale and $\mathbb{E}|\bar{w}_n|_+ \leq \infty$, because of $p \in C_0\left[a, \frac{1}{a}\right]$. Then we can apply convergence theorem of submartingales and we conclude that there exists a random variable $\bar{w}_\infty$ such that

$$\lim_{n\to\infty} \bar{w}_n = \bar{w}_\infty \qquad \text{a.s.}$$

We apply Lemma 7.1 with $f_i(\mathbf{X}) \stackrel{\text{def}}{=} \bar{w}_i(\mathbf{X})$ parameter, then we get

$$\lim_{n\to\infty} \frac{1}{n}\sum_{i=1}^{n} \max_j \mathbb{E}\left\{p(X_i^{(j)})|\mathbf{X}_1^{i-1}\right\} = \mathbb{E}\left\{\max_j \mathbb{E}\left\{p(X_0^{(j)})|\mathbf{X}_{-\infty}^{-1}\right\}\right\} \qquad \text{a.s.}$$

because of

$$f_i(T^i\mathbf{X}) = \bar{w}_i(T^i\mathbf{X}) = \max_j E\left(p(X_0^{(j)}) \mid \mathbf{X}_1^{i-1}\right).$$

and $\mathbb{E}\left\{\sup_i |f_i(\mathbf{X})|\right\} < \infty$. The latter one follows from that $p(\cdot)$ is bounded. ∎

Lemma 7.4 *Let $Z_1, \ldots, Z_n$ a sequence of random variables then we get the following upper and lower bound for the logarithmic function of Z_n if $0 \leq \lambda < \frac{1}{2}$*

$$\begin{aligned}
&\frac{U_M(Z_n,\lambda) + g(Z_n,\lambda) - \lambda\mathbb{E}^2\{Z_n|Z_1^{n-1}\} + \frac{1}{3Z_n^3}(Z_n-1)^3}{1-2\lambda} \leq \log Z_n \\
&\qquad\qquad \leq \frac{U_M(Z_n,\lambda) + g(Z_n,\lambda) - \lambda\mathbb{E}^2\{Z_n|Z_1^{n-1}\} + \frac{1}{3}(Z_n-1)^3}{1-2\lambda}
\end{aligned}$$

where

$$g(Z_n,\lambda) = \left(2\lambda - \frac{1}{2}\right)(Z_n-1)^2 - 1 + \lambda.$$

Proof. To show the relationship between the log- and the Markowitz-utility function we use the Taylor expansion of the logarithmic function

$$\begin{aligned}
U_M(Z_n,\lambda) &- (1-2\lambda)\log Z_n \\
&= \left(\frac{1}{2} - 2\lambda\right)(Z_n-1)^2 + 1 - \lambda + \lambda\mathbb{E}^2\{Z_n|Z_1^{n-1}\} + R_2, \qquad (7.2)
\end{aligned}$$

where $R_2 = \frac{(Z_n-1)^3}{3(Z^*)^3}$ $(Z_n^* \in [\min\{Z_n, 1\}, \max\{Z_n, 1\}])$ is the Lagrange remainder. Then using

$$\frac{1}{3Z_n^3}(Z_n-1)^3 \le R_2 \le \frac{1}{3}(Z_n-1)^3$$

we obtain the statement of the lemma. ∎

Lemma 7.5 *Let* $\mathbf{X}$ *a random market vector satisfying (4.1). Then for any* $\mathbf{b}'$ *and* $\mathbf{b}''$ *portfolios*

$$\left| \frac{(\langle \mathbf{b}', \mathbf{X}\rangle - 1)^3}{\langle \mathbf{b}', \mathbf{X}\rangle^3} - (\langle \mathbf{b}'', \mathbf{X}\rangle - 1)^3 \right| \le (a^{-3}+1) \max_m |X^{(m)} - 1|^3.$$

Proof. First we show

$$\frac{(\langle \mathbf{b}', \mathbf{X}\rangle - 1)^3}{\langle \mathbf{b}', \mathbf{X}\rangle^3} - (\langle \mathbf{b}'', \mathbf{X}\rangle - 1)^3 \ge -(a^{-3}+1) \max_m |X^{(m)} - 1|^3. \tag{7.3}$$

If $\langle \mathbf{b}', \mathbf{X}\rangle < \langle \mathbf{b}'', \mathbf{X}\rangle$ then

$$\begin{aligned}
&\frac{(\langle \mathbf{b}', \mathbf{X}\rangle - 1)^3}{\langle \mathbf{b}', \mathbf{X}\rangle^3} - (\langle \mathbf{b}'', \mathbf{X}\rangle - 1)^3 \\
&= -\left| \frac{(\langle \mathbf{b}', \mathbf{X}\rangle - 1)^3}{\langle \mathbf{b}', \mathbf{X}\rangle^3} - (\langle \mathbf{b}'', \mathbf{X}\rangle - 1)^3 \right| \\
&\ge -\frac{|\langle \mathbf{b}', \mathbf{X}\rangle - 1|^3}{\langle \mathbf{b}', \mathbf{X}\rangle^3} - |\langle \mathbf{b}'', \mathbf{X}\rangle - 1|^3 \\
&\ge -\max\left\{ |\langle \mathbf{b}', \mathbf{X}\rangle - 1|^3, |\langle \mathbf{b}'', \mathbf{X}\rangle - 1|^3 \right\} (\langle \mathbf{b}', \mathbf{X}\rangle^{-3} + 1).
\end{aligned} \tag{7.4}$$

Let bound the terms in the maximum by Jensen's inequality,

$$|\langle \mathbf{b}, \mathbf{X}\rangle - 1|^3 = \left| \sum_{m=1}^d b^{(m)}(X^{(m)}-1) \right|^3 \le \sum_{m=1}^d b^{(m)} \left| X^{(m)} - 1 \right|^3 \le \max_m \left| X^{(m)} - 1 \right|^3, \tag{7.5}$$

use $\langle \mathbf{b}', \mathbf{X}\rangle^{-3} \le a^{-3}$ and plug these bounds into (7.4) we obtain (7.3).
If $\left\langle \mathbf{b}', \mathbf{X} \right\rangle \ge \left\langle \mathbf{b}'', \mathbf{X} \right\rangle$ then

$$\begin{aligned}
\frac{(\langle \mathbf{b}', \mathbf{X}\rangle - 1)^3}{\langle \mathbf{b}', \mathbf{X}\rangle^3} - (\langle \mathbf{b}'', \mathbf{X}\rangle - 1)^3 &\ge \left(\langle \mathbf{b}', \mathbf{X}\rangle^{-3} - 1 \right) (\langle \mathbf{b}', \mathbf{X}\rangle - 1)^3 \\
&= -\left| \langle \mathbf{b}', \mathbf{X}\rangle^{-3} - 1 \right| |\langle \mathbf{b}', \mathbf{X}\rangle - 1|^3 \\
&\ge -\left| a^{-3} - 1 \right| \max_m \left| X^{(m)} - 1 \right|^3 \\
&\ge -\left(a^{-3} + 1 \right) \max_m \left| X^{(m)} - 1 \right|^3 .
\end{aligned}$$

With similarly argument we obtain

$$\frac{(\langle \mathbf{b}', \mathbf{X}\rangle - 1)^3}{\langle \mathbf{b}', \mathbf{X}\rangle^3} - (\langle \mathbf{b}'', \mathbf{X}\rangle - 1)^3 \le (a^{-3}+1)\max_m \left|X^{(m)} - 1\right|^3 .$$

∎

Corollary 7.6 (of Lemma 7.5) *Let $\{\mathbf{X}_n\}_{-\infty}^{\infty}$ be a stationary and ergodic process, satisfying (4.1). Then for any $\mathbf{b}'$ and $\mathbf{b}''$ portfolios*

$$\mathbb{E}\left\{\left|\frac{(\langle \mathbf{b}', \mathbf{X}_n\rangle - 1)^3}{\langle \mathbf{b}', \mathbf{X}_n\rangle^3} - (\langle \mathbf{b}'', \mathbf{X}_n\rangle - 1)^3\right| \middle| \mathbf{X}_1^{n-1}\right\} \le (a^{-3}+1)\max_m \mathbb{E}\left\{|X_n^{(m)} - 1|^3 \middle| \mathbf{X}_1^{n-1}\right\}.$$

Proof. In the proof of Lemma 7.5 instead of equation (7.5) use the following

$$\begin{aligned}\mathbb{E}\left\{|\langle \mathbf{b}, \mathbf{X}_n\rangle - 1|^3 \,|\mathbf{X}_1^{n-1}\right\} &= \mathbb{E}\left\{\left|\sum_{m=1}^{d} b^{(m)}(X_n^{(m)} - 1)\right|^3 \middle| \mathbf{X}_1^{n-1}\right\} \\ &\le \sum_{m=1}^{d} b^{(m)}\mathbb{E}\left\{\left|X_n^{(m)} - 1\right|^3 \middle| \mathbf{X}_1^{n-1}\right\} \\ &\le \max_m \mathbb{E}\left\{\left|X_n^{(m)} - 1\right|^3 \middle| \mathbf{X}_1^{n-1}\right\}.\end{aligned}$$

∎

Corollary 7.7 (of Lemma 7.5) *Let $\mathbf{X}$ be a random market vector satisfying (4.1). Then for any $\mathbf{b}'$ and $\mathbf{b}''$ portfolios*

$$\left|\frac{(\langle \mathbf{b}', \mathbf{X}\rangle - 1)^3}{\langle \mathbf{b}', \mathbf{X}\rangle^3} - \mathbb{E}\left\{(\langle \mathbf{b}'', \mathbf{X}\rangle - 1)^3\right\}\right| \le (a^{-3}+1)\max_m \left|X^{(m)} - 1\right|^3 .$$

Proof. In the proof of Lemma 7.5 instead of considering cases $\langle \mathbf{b}', \mathbf{X}\rangle < \langle \mathbf{b}'', \mathbf{X}\rangle$ and $\langle \mathbf{b}', \mathbf{X}\rangle \ge \langle \mathbf{b}'', \mathbf{X}\rangle$, we split according to $\langle \mathbf{b}', \mathbf{X}\rangle < \mathbb{E}\{\langle \mathbf{b}'', \mathbf{X}\rangle\}$ and $\langle \mathbf{b}', \mathbf{X}\rangle \ge \mathbb{E}\{\langle \mathbf{b}'', \mathbf{X}\rangle\}$. The proof of the two cases goes on the same way as in Lemma 7.5. ∎

Lemma 7.8 *Let $\{\mathbf{X}_n\}_{-\infty}^{\infty}$ be a stationary and ergodic process then*

$$\mathbb{E}\left\{\left\langle \mathbf{b}^*(\mathbf{X}_1^{n-1}), \mathbf{X}_n\right\rangle - \langle \mathbf{b}, \mathbf{X}_n\rangle \mid \mathbf{X}_1^{n-1}\right\} \ge \min_m \mathbb{E}\left\{1 + \log(X_n^{(m)}) - X_n^{(m)} \mid \mathbf{X}_1^{n-1}\right\},$$

where $\mathbf{b}^*(\mathbf{X}_1^{n-1})$ is the log-optimal portfolio and $\mathbf{b} \in \Delta_d$ is an arbitrary portfolio.

Proof. Let us bound from below the first term of the statement

$$\mathbb{E}\left\{\left\langle \mathbf{b}^*(\mathbf{X}_1^{n-1}), \mathbf{X}_n\right\rangle \mid \mathbf{X}_1^{n-1}\right\} \geq e^{\mathbb{E}\left\{\log\left\langle \mathbf{b}^*(\mathbf{X}_1^{n-1}), \mathbf{X}_n\right\rangle \mid \mathbf{X}_1^{n-1}\right\}} \tag{7.6}$$

$$\geq e^{\mathbb{E}\left\{\log\left\langle \mathbf{b}(\mathbf{X}_1^{n-1}), \mathbf{X}_n\right\rangle \mid \mathbf{X}_1^{n-1}\right\}} \tag{7.7}$$

$$\geq 1 + \sum_{m=1}^{d} b^{(m)}\mathbb{E}\left\{\log X_n^{(m)} \mid \mathbf{X}_1^{n-1}\right\}, \tag{7.8}$$

where (7.6) follows from the Jensen inequality, (7.7) comes from the definition of $\mathbf{b}^*$ and (7.8) because of $e^x \geq 1 + x$. Plug this into the left side of the statement we get

$$\begin{aligned}
&\mathbb{E}\left\{\left\langle \mathbf{b}^*(\mathbf{X}_1^{n-1}), \mathbf{X}_n\right\rangle - \left\langle \mathbf{b}(\mathbf{X}_1^{n-1}), \mathbf{X}_n\right\rangle \mid \mathbf{X}_1^{n-1}\right\} \\
&\geq \sum_{m=1}^{d} b^{(m)}\mathbb{E}\left\{1 + \log X_n^{(m)} - X_n^{(m)} \mid \mathbf{X}_1^{n-1}\right\} \\
&\geq \min_m \mathbb{E}\left\{1 + \log X_n^{(m)} - X_n^{(m)} \mid \mathbf{X}_1^{n-1}\right\}
\end{aligned}$$

because of $\mathbb{E}\left\{1 + \log X_n^{(m)} - X_n^{(m)} \middle| \mathbf{X}_1^{n-1}\right\} \leq 0$. ∎

We are now ready to prove Theorem 4.1. For convenience in the proof of both theorems we use the notations $\bar{\mathbf{b}}$ instead of $\bar{\mathbf{b}}_\lambda$, $\bar{\mathbf{h}}^{(k,\ell)}$ instead of $\bar{\mathbf{h}}_\lambda^{(k,\ell)}$, $\bar{\mathbf{B}}$ instead of $\bar{\mathbf{B}}_\lambda$ and $\bar{S}_n$ instead of $\bar{S}_{n,\lambda}$.

Proof of Theorem 4.1. The λ parameter is fixed. Use Lemma 7.4 with $Z_n \stackrel{\text{def}}{=} \left\langle \bar{\mathbf{b}}^*(\mathbf{X}_1^{n-1}), \mathbf{X}_n\right\rangle$, where $\bar{\mathbf{b}}^*(\mathbf{X}_1^{n-1})$ is the Markowitz-type portfolio, then we get

$$\begin{aligned}
&(1-2\lambda)\mathbb{E}\left\{\log\left\langle \bar{\mathbf{b}}^*(\mathbf{X}_1^{n-1}), \mathbf{X}_n\right\rangle \mid \mathbf{X}_1^{n-1}\right\} - \mathbb{E}\left\{g(\left\langle \bar{\mathbf{b}}^*(\mathbf{X}_1^{n-1}), \mathbf{X}_n\right\rangle, \lambda) \mid \mathbf{X}_1^{n-1}\right\} \\
&\quad +\lambda\mathbb{E}^2\{\left\langle \bar{\mathbf{b}}^*(\mathbf{X}_1^{n-1}), \mathbf{X}_n\right\rangle \mid \mathbf{X}_1^{n-1}\} - \frac{1}{3}\left(\mathbb{E}\left\{\frac{\left(\left\langle \bar{\mathbf{b}}^*(\mathbf{X}_1^{n-1}), \mathbf{X}_n\right\rangle - 1\right)^3}{\left\langle \bar{\mathbf{b}}^*(\mathbf{X}_1^{n-1}), \mathbf{X}_n\right\rangle^3}\middle| \mathbf{X}_1^{n-1}\right\}\right) \\
&\geq \mathbb{E}\left\{U_M(\left\langle \bar{\mathbf{b}}^*(\mathbf{X}_1^{n-1}), \mathbf{X}_n\right\rangle, \lambda) \mid \mathbf{X}_1^{n-1}\right\} \\
&\geq \mathbb{E}\left\{U_M(\left\langle \mathbf{b}^*(\mathbf{X}_1^{n-1}), \mathbf{X}_n\right\rangle, \lambda) \mid \mathbf{X}_1^{n-1}\right\} \\
&\geq (1-2\lambda)\mathbb{E}\left\{\log\left\langle \mathbf{b}^*(\mathbf{X}_1^{n-1}), \mathbf{X}_n\right\rangle \mid \mathbf{X}_1^{n-1}\right\} - \mathbb{E}\left\{g(\left\langle \mathbf{b}^*(\mathbf{X}_1^{n-1}), \mathbf{X}_n\right\rangle, \lambda) \mid \mathbf{X}_1^{n-1}\right\} \\
&\quad +\lambda\mathbb{E}^2\{\left\langle \mathbf{b}^*(\mathbf{X}_1^{n-1}), \mathbf{X}_n\right\rangle \mid \mathbf{X}_1^{n-1}\} - \frac{1}{3}\mathbb{E}\left\{(\left\langle \mathbf{b}^*(\mathbf{X}_1^{n-1}), \mathbf{X}_n\right\rangle - 1)^3 \mid \mathbf{X}_1^{n-1}\right\}.
\end{aligned}$$

After rearranging the above inequalities, we get

$$\begin{aligned}
&(1-2\lambda)\mathbb{E}\left\{\log\left\langle \bar{\mathbf{b}}^*(\mathbf{X}_1^{n-1}), \mathbf{X}_n\right\rangle \mid \mathbf{X}_1^{n-1}\right\} \\
&\geq (1-2\lambda)\mathbb{E}\left\{\log\left\langle \mathbf{b}^*(\mathbf{X}_1^{n-1}), \mathbf{X}_n\right\rangle \mid \mathbf{X}_1^{n-1}\right\} \\
&\quad +\lambda\left(\mathbb{E}^2\{\left\langle \mathbf{b}^*(\mathbf{X}_1^{n-1}), \mathbf{X}_n\right\rangle \mid \mathbf{X}_1^{n-1}\} - \mathbb{E}^2\{\left\langle \bar{\mathbf{b}}^*(\mathbf{X}_1^{n-1}), \mathbf{X}_n\right\rangle \mid \mathbf{X}_1^{n-1}\}\right) \\
&\quad +\mathbb{E}\left\{g(\left\langle \bar{\mathbf{b}}^*(\mathbf{X}_1^{n-1}), \mathbf{X}_n\right\rangle, \lambda) - g(\left\langle \mathbf{b}^*(\mathbf{X}_1^{n-1}), \mathbf{X}_n\right\rangle, \lambda) \mid \mathbf{X}_1^{n-1}\right\} \\
&\quad +\frac{1}{3}\mathbb{E}\left\{\frac{(\left\langle \bar{\mathbf{b}}^*(\mathbf{X}_1^{n-1}), \mathbf{X}_n\right\rangle - 1)^3}{\left\langle \bar{\mathbf{b}}^*(\mathbf{X}_1^{n-1}), \mathbf{X}_n\right\rangle^3} - (\left\langle \mathbf{b}^*(\mathbf{X}_1^{n-1}), \mathbf{X}_n\right\rangle - 1)^3 \middle| \mathbf{X}_1^{n-1}\right\}.
\end{aligned} \tag{7.9}$$

Taking the arithmetic average on both sides of the inequality over trading periods $1, \dots, n$, then

$$\begin{aligned}
&\frac{1}{n}\sum_{i=1}^{n} \mathbb{E}\left\{\log\left\langle \bar{\mathbf{b}}^*(\mathbf{X}_1^{i-1}), \mathbf{X}_i\right\rangle \mid \mathbf{X}_1^{i-1}\right\} \\
&\geq \frac{1}{n}\sum_{i=1}^{n} \mathbb{E}\left\{\log\left\langle \mathbf{b}^*(\mathbf{X}_1^{i-1}), \mathbf{X}_i\right\rangle \mid \mathbf{X}_1^{i-1}\right\} \\
&\quad + \frac{1}{n}\sum_{i=1}^{n} \frac{\lambda\left(\mathbb{E}^2\{\left\langle \mathbf{b}^*(\mathbf{X}_1^{i-1}), \mathbf{X}_i\right\rangle \mid \mathbf{X}_1^{i-1}\} - \mathbb{E}^2\{\left\langle \bar{\mathbf{b}}^*(\mathbf{X}_1^{i-1}), \mathbf{X}_i\right\rangle \mid \mathbf{X}_1^{i-1}\}\right)}{1-2\lambda} \\
&\quad + \frac{1}{n}\sum_{i=1}^{n} \frac{\mathbb{E}\left\{g(\left\langle \bar{\mathbf{b}}^*(\mathbf{X}_1^{i-1}), \mathbf{X}_i\right\rangle, \lambda) - g(\left\langle \mathbf{b}^*(\mathbf{X}_1^{i-1}), \mathbf{X}_i\right\rangle, \lambda) \mid \mathbf{X}_1^{i-1}\right\}}{1-2\lambda} \\
&\quad + \frac{1}{3n}\sum_{i=1}^{n} \frac{\mathbb{E}\left\{\frac{\left(\left\langle \bar{\mathbf{b}}^*(\mathbf{X}_1^{i-1}), \mathbf{X}_i\right\rangle - 1\right)^3}{\left\langle \bar{\mathbf{b}}^*(\mathbf{X}_1^{i-1}), \mathbf{X}_i\right\rangle^3} - (\left\langle \mathbf{b}^*(\mathbf{X}_1^{i-1}), \mathbf{X}_i\right\rangle - 1)^3 \middle| \mathbf{X}_1^{i-1}\right\}}{1-2\lambda}. \qquad (7.10)
\end{aligned}$$

We derive simple bounds for the last three additive parts of the above inequality. First, because of $\lambda < \frac{1}{2}$

$$\begin{aligned}
&\mathbb{E}^2\{\left\langle \mathbf{b}^*(\mathbf{X}_1^{i-1}), \mathbf{X}_i\right\rangle \mid \mathbf{X}_1^{i-1}\} - \mathbb{E}^2\{\left\langle \bar{\mathbf{b}}^*(\mathbf{X}_1^{i-1}), \mathbf{X}_i\right\rangle \mid \mathbf{X}_1^{i-1}\} \\
&= \mathbb{E}\{\left\langle \mathbf{b}^*(\mathbf{X}_1^{i-1}), \mathbf{X}_i\right\rangle + \left\langle \bar{\mathbf{b}}^*(\mathbf{X}_1^{i-1}), \mathbf{X}_i\right\rangle \mid \mathbf{X}_1^{i-1}\} \\
&\quad \cdot \mathbb{E}\{\left\langle \mathbf{b}^*(\mathbf{X}_1^{i-1}), \mathbf{X}_i\right\rangle - \left\langle \bar{\mathbf{b}}^*(\mathbf{X}_1^{i-1}), \mathbf{X}_i\right\rangle \mid \mathbf{X}_1^{i-1}\} \\
&\geq \mathbb{E}\{\left\langle \mathbf{b}^*(\mathbf{X}_1^{i-1}), \mathbf{X}_i\right\rangle + \left\langle \bar{\mathbf{b}}^*(\mathbf{X}_1^{i-1}), \mathbf{X}_i\right\rangle \mid \mathbf{X}_1^{i-1}\} \\
&\quad \cdot \min_m \mathbb{E}\left\{1 + \log(X_i^{(m)}) - X_i^{(m)} \mid \mathbf{X}_1^{i-1}\right\} \\
&\geq 2a^{-1} \min_m \mathbb{E}\left\{1 + \log(X_i^{(m)}) - X_i^{(m)} \mid \mathbf{X}_1^{i-1}\right\}. \qquad (7.11)
\end{aligned}$$

Second,

$$\begin{aligned}
&\mathbb{E}\left\{\frac{g(\left\langle \bar{\mathbf{b}}^*(\mathbf{X}_1^{i-1}), \mathbf{X}_i\right\rangle, \lambda) - g(\left\langle \mathbf{b}^*(\mathbf{X}_1^{i-1}), \mathbf{X}_i\right\rangle, \lambda)}{1-2\lambda} \middle| \mathbf{X}_1^{i-1}\right\} \\
&\qquad \geq -\left|\frac{2\lambda - \frac{1}{2}}{1-2\lambda}\right| \left(\max_m \mathbb{E}\left\{(X_i^{(m)} - 1)^2 \mid \mathbf{X}_1^{i-1}\right\}\right. \\
&\qquad\qquad \left. - \left(\min_m \mathbb{E}\left\{|X_i^{(m)} - 1| \middle| \mathbf{X}_{-\infty}^{-1}\right\}\right)^2\right) \qquad (7.12)
\end{aligned}$$

for all value of λ. And finally, we use Corollary 7.6,

$$\begin{aligned}
&\frac{1}{3(1-2\lambda)}\mathbb{E}\left\{\frac{(\left\langle \bar{\mathbf{b}}^*(\mathbf{X}_1^{i-1}), \mathbf{X}_i\right\rangle - 1)^3}{\left\langle \bar{\mathbf{b}}^*(\mathbf{X}_1^{i-1}), \mathbf{X}_i\right\rangle^3} - (\left\langle \mathbf{b}^*(\mathbf{X}_1^{i-1}), \mathbf{X}_i\right\rangle - 1)^3 \middle| \mathbf{X}_1^{i-1}\right\} \\
&\qquad \geq -\frac{a^{-3}+1}{3(1-2\lambda)} \max_m \mathbb{E}\left\{|X_n^{(m)} - 1|^3 |\mathbf{X}_1^{i-1}\right\}. \qquad (7.13)
\end{aligned}$$

Consider the following decomposition

$$\frac{1}{n}\log \bar{S}_n^* = \bar{Y}_n + \bar{V}_n,$$

where

$$\bar{Y}_n = \frac{1}{n}\sum_{i=1}^{n}\left(\log\left\langle \bar{\mathbf{b}}^*(\mathbf{X}_1^{i-1}), \mathbf{X}_i\right\rangle - \mathbb{E}\left\{\log\left\langle \bar{\mathbf{b}}^*(\mathbf{X}_1^{i-1}), \mathbf{X}_i\right\rangle |\mathbf{X}_1^{i-1}\right\}\right)$$

and

$$\bar{V}_n = \frac{1}{n}\sum_{i=1}^{n}\mathbb{E}\left\{\log\left\langle \bar{\mathbf{b}}^*(\mathbf{X}_1^{i-1}), \mathbf{X}_i\right\rangle |\mathbf{X}_1^{i-1}\right\}.$$

It can be shown that $\bar{Y}_n \to 0$ a.s., since it is an average of bounded martingale differences. So

$$\liminf_{n\to\infty} \bar{V}_n = \liminf_{n\to\infty}\frac{1}{n}\log \bar{S}_n^*. \tag{7.14}$$

Similarly, consider the following decomposition

$$\frac{1}{n}\log S_n^* = Y_n + V_n,$$

where

$$Y_n = \frac{1}{n}\sum_{i=1}^{n}\left(\log\left\langle \mathbf{b}^*(\mathbf{X}_1^{i-1}), \mathbf{X}_i\right\rangle - \mathbb{E}\left\{\log\left\langle \mathbf{b}^*(\mathbf{X}_1^{i-1}), \mathbf{X}_i\right\rangle |\mathbf{X}_1^{i-1}\right\}\right)$$

and

$$V_n = \frac{1}{n}\sum_{i=1}^{n}\mathbb{E}\left\{\log\left\langle \mathbf{b}^*(\mathbf{X}_1^{i-1}), \mathbf{X}_i\right\rangle |\mathbf{X}_1^{i-1}\right\}.$$

Again, it can be shown that $Y_n \to 0$ a.s. Therefore

$$\lim_{n\to\infty} V_n = \lim_{n\to\infty}\frac{1}{n}\log S_n^*. \tag{7.15}$$

Taking the limes inferior of both sides of (7.10) as n goes to infinity and applying equalities (7.11), (7.12), (7.13), (7.14), (7.15) and Lemma 7.3, we obtain

$$\begin{aligned}
W^* &\geq \liminf_{n\to\infty}\frac{1}{n}\log \bar{S}_n^* \\
&\geq W^* + \frac{\lambda}{1-2\lambda}2a^{-1}\mathbb{E}\left\{\min_m \mathbb{E}\left\{1+\log(X_0^{(m)}) - X_0^{(m)}\Big|\mathbf{X}_{-\infty}^{-1}\right\}\right\} \\
&\quad - \left|\frac{2\lambda-\frac{1}{2}}{1-2\lambda}\right|\mathbb{E}\left\{\max_m \mathbb{E}\left\{(X_0^{(m)}-1)^2\Big|\mathbf{X}_{-\infty}^{-1}\right\}\right\} \\
&\quad - \frac{a^{-3}+1}{3(1-2\lambda)}\mathbb{E}\left\{\max_m \mathbb{E}\left\{\left|X_0^{(m)}-1\right|^3\Big|\mathbf{X}_{-\infty}^{-1}\right\}\right\}
\end{aligned}$$

as desired.

■

For the proof of Theorem 5.1 we need following two lemmas. The first is a simple modification of Theorem 1 in [10].

Lemma 7.9 *Under the conditions of Theorem 5.1, for any k, ℓ integers and for any fixed λ*

$$\lim_{n\to\infty} \frac{1}{n}\sum_{i=1}^{n} U_M\left(\left\langle \bar{\mathbf{h}}^{(k,\ell)}(\mathbf{X}_1^{i-1})\,,\,\mathbf{X}_i\right\rangle, \lambda\right) = \mathbb{E}\left\{U_M\left(\left\langle \bar{\mathbf{b}}^*_{k,\ell}(\mathbf{X}_{-k}^{-1})\,,\,\mathbf{X}_0\right\rangle, \lambda\right)\right\}$$

where $\bar{\mathbf{b}}^*_{k,\ell}(\mathbf{X}_{-k}^{-1})$ is the Markowitz-type portfolio with respect to the limit distribution $\mathbb{P}^{*(k,\ell)}_{\mathbf{X}_{-k}^{-1}}$.

Proof. Let the integers k, ℓ and the vector $\mathbf{s} = \mathbf{s}_{-k}^{-1} \in \mathbb{R}_+^{dk}$ be fixed. Let $\mathbb{P}^{(k,\ell)}_{j,\mathbf{s}}$ denote the (random) measure concentrated on $\{\mathbf{X}_i : 1-j+k \le i \le 0, \|\mathbf{X}_{i-k}^{i-1} - \mathbf{s}\| \le r_{k,\ell}\}$ defined by

$$\mathbb{P}^{(k,\ell)}_{j,\mathbf{s}}(A) = \frac{\sum_{\{i:1-j+k\le i\le 0, \|\mathbf{X}_{i-k}^{i-1}-\mathbf{s}\|\le r_{k,\ell}\}} \mathbb{I}_A(\mathbf{X}_i)}{|\{i : 1-j+k \le i \le 0, \|\mathbf{X}_{i-k}^{i-1} - \mathbf{s}\| \le r_{k,\ell}\}|}\,, \qquad A \subset \mathbb{R}_+^d$$

where $\mathbb{I}_A$ denotes the indicator function of the set A. If the above set of $\mathbf{X}_i$'s is empty, then let $\mathbb{P}^{(k,\ell)}_{j,\mathbf{s}} = \delta_{(1,\dots,1)}$ be the probability measure concentrated on the vector $(1,\dots,1)$. Györfi, Lugosi, Udina [10] proved that for all $\mathbf{s}$, with probability one,

$$\mathbb{P}^{(k,\ell)}_{j,\mathbf{s}} \to \mathbb{P}^{*(k,\ell)}_{\mathbf{s}} = \begin{cases} \mathbb{P}_{\mathbf{X}_0 | \|\mathbf{X}_{-k}^{-1}-\mathbf{s}\| \le r_{k,\ell}} & \text{if } \mathbb{P}(\|\mathbf{X}_{-k}^{-1} - \mathbf{s}\| \le r_{k,\ell}) > 0, \\ \delta_{(1,\dots,1)} & \text{if } \mathbb{P}(\|\mathbf{X}_{-k}^{-1} - \mathbf{s}\| \le r_{k,\ell}) = 0 \end{cases} \tag{7.16}$$

weakly as $j \to \infty$ where $\mathbb{P}_{\mathbf{X}_0 | \|\mathbf{X}_{-k}^{-1}-\mathbf{s}\| \le r_{k,\ell}}$ denotes the distribution of the vector $\mathbf{X}_0$ conditioned on the event $\|\mathbf{X}_{-k}^{-1} - \mathbf{s}\| \le r_{k,\ell}$.

By definition, $\bar{\mathbf{b}}^{(k,\ell)}(\mathbf{X}_{1-j}^{-1}, \mathbf{s})$ is the Markowitz-type portfolio with respect to the probability measure $\mathbb{P}^{(k,\ell)}_{j,\mathbf{s}}$. Let $\bar{\mathbf{b}}^*_{k,\ell}(\mathbf{s})$ denote the Markowitz-type portfolio with respect to the limit distribution $\mathbb{P}^{*(k,\ell)}_{\mathbf{s}}$. Then, using Lemma 7.2, we infer from (7.16) that, as j tends to infinity, we have the almost sure convergence

$$\lim_{j\to\infty} \left\langle \bar{\mathbf{b}}^{(k,\ell)}(\mathbf{X}_{1-j}^{-1}, \mathbf{s})\,,\,\mathbf{x}_0\right\rangle = \left\langle \bar{\mathbf{b}}^*_{k,\ell}(s)\,,\,\mathbf{x}_0\right\rangle$$

for $\mathbb{P}^{*(k,\ell)}_{\mathbf{s}}$-almost all $\mathbf{x}_0$ and hence for $\mathbb{P}_{\mathbf{X}_0}$-almost all $\mathbf{x}_0$. Since $\mathbf{s}$ was arbitrary, we obtain

$$\lim_{j\to\infty} \left\langle \bar{\mathbf{b}}^{(k,\ell)}(\mathbf{X}_{1-j}^{-1}, \mathbf{X}_{-k}^{-1})\,,\,\mathbf{x}_0\right\rangle = \left\langle \bar{\mathbf{b}}^*_{k,\ell}(\mathbf{X}_{-k}^{-1})\,,\,\mathbf{x}_0\right\rangle \quad \text{a.s.} \tag{7.17}$$

Next we apply Lemma 7.1 for the utility function

$$f_i(\mathbf{x}_{-\infty}^{\infty}) \stackrel{\text{def}}{=} U_M\left(\left\langle \bar{\mathbf{h}}^{(k,\ell)}(\mathbf{x}_{1-i}^{-1})\,,\,\mathbf{x}_0\right\rangle, \lambda\right) \quad = U_M\left(\left\langle \bar{\mathbf{b}}^{(k,\ell)}(\mathbf{x}_{1-i}^{-1}, \mathbf{x}_{-k}^{-1})\,,\,\mathbf{x}_0\right\rangle, \lambda\right)$$

defined on $\mathbf{x}_{-\infty}^{\infty} = (\ldots, \mathbf{x}_{-1}, \mathbf{x}_0, \mathbf{x}_1, \ldots)$. Because of $\left|f_i(\mathbf{X}_{-\infty}^{\infty})\right| < \infty$ and

$$\lim_{i\to\infty} f_i(\mathbf{X}_{-\infty}^{\infty}) = U_M(\left\langle \bar{\mathbf{b}}_{k,\ell}^*(\mathbf{X}_{-k}^{-1})\, ,\, \mathbf{X}_0 \right\rangle, \lambda) \quad \text{a.s.}$$

by (7.17). As $n \to \infty$, Lemma 7.1 yields

$$\begin{aligned} \frac{1}{n}\sum_{i=1}^{n} f_i(T^i\mathbf{X}_{-\infty}^{\infty}) &= \frac{1}{n}\sum_{i=1}^{n} U_M\left(\left\langle \bar{\mathbf{h}}^{(k,\ell)}(\mathbf{X}_1^{i-1})\, ,\, \mathbf{X}_i \right\rangle, \lambda\right) \\ &\to \mathbb{E}\left\{U_M\left(\left\langle \bar{\mathbf{b}}_{k,\ell}^*(\mathbf{X}_{-k}^{-1})\, ,\, \mathbf{X}_0 \right\rangle, \lambda\right)\right\} \end{aligned}$$

as desired. ∎

Lemma 7.10 *Under the conditions of Theorem 5.1 there exists a portfolio* $\mathbf{b}_c$ *for all* $c \in \mathcal{C}$ *for which*

$$\mathbb{E}^2\left\{\langle \mathbf{b}_c\, ,\, \mathbf{X}_0 \rangle\right\} = c,$$

where $\mathcal{C} \stackrel{\text{def}}{=} \left[\left(\min_m \mathbb{E}(\mathbf{X}_0^{(m)})\right)^2, \left(\max_m \mathbb{E}(\mathbf{X}_0^{(m)})\right)^2\right]$ *and furthermore*

$$\liminf_{n\to\infty} \frac{1}{n}\sum_{i=1}^{n} \mathbb{E}^2\left\{\left\langle \mathbf{h}^{(k,\ell)}(\mathbf{X}_1^{i-1})\, ,\, \mathbf{X}_i \right\rangle |\mathbf{X}_1^{i-1}\right\} \subseteq \mathcal{C}.$$

Proof. The proof has two steps. First we show, that there exists a portfolio for all $c \in \mathcal{C}$ whose expected value is c. Let $c \in \mathcal{C}$ and $\mathbb{E}^2\left\{\langle \mathbf{b}_c\, ,\, \mathbf{X}_0 \rangle\right\} = c$ then

$$\mathbb{E}^2\left\{\left\langle \mathbf{b}_c(\mathbf{X}_{-k}^{-1})\, ,\, \mathbf{X}_0 \right\rangle\right\} = \left(\sum_{m=1}^{d} b_c^{(m)} \mathbb{E}\left\{\mathbf{X}_0^{(m)}\right\}\right)^2 \stackrel{\text{def}}{=} \left(\sum_{m=1}^{d} b_c^{(m)} e_m\right)^2.$$

Denote $m' = \arg\min_m e_m$ and $m'' = \arg\max_m e_m$ then let $\mathbf{b}_c$ portfolio is the following $b_c^{(m')} + b_c^{(m'')} = 1$ for all other m $b_c^{(m)} = 0$. So for all $c \in \mathcal{C}$ there exists $b_c^{(m')}$ and $b_c^{(m'')}$ that $\left(b_c^{(m')} e_{m'} + b_c^{(m'')} e_{m''}\right)^2 = c$ because of the continuity. As the second step we have to prove that

$$\liminf_{n\to\infty} \frac{1}{n}\sum_{i=1}^{n} \mathbb{E}^2\left\{\left\langle \mathbf{h}^{(k,\ell)}(\mathbf{X}_1^{i-1})\, ,\, \mathbf{X}_i \right\rangle |\mathbf{X}_1^{i-1}\right\} \subseteq \mathcal{C}.$$

It is easy to show from Lemma 7.9 and Lemma 7.1 that as $n \to \infty$

$$\frac{1}{n}\sum_{i=1}^{n} \mathbb{E}\left\{\left\langle \bar{\mathbf{h}}^{(k,\ell)}(\mathbf{X}_1^{i-1})\, ,\, \mathbf{X}_i \right\rangle |X_1^{i-1}\right\} \to \mathbb{E}\left\{\left\langle \bar{\mathbf{b}}_{k,\ell}^*(\mathbf{X}_{-k}^{-1})\, ,\, \mathbf{X}_0 \right\rangle\right\},$$

from which the statement of the lemma follows with argument of contradiction. ∎

Proof of Theorem 5.1. Without loss of generality we may assume $S_0 = 1$, so

$$\begin{aligned}
\liminf_{n\to\infty} W_n(\bar{\mathbf{B}}) &= \liminf_{n\to\infty} \frac{1}{n} \log S_n(\bar{\mathbf{B}}) \\
&= \liminf_{n\to\infty} \frac{1}{n} \log \left(\sum_{k,\ell} q_{k,\ell} S_n(\bar{\mathbf{H}}^{(k,\ell)}) \right) \\
&\geq \liminf_{n\to\infty} \frac{1}{n} \log \left(\sup_{k,\ell} q_{k,\ell} S_n(\bar{\mathbf{H}}^{(k,\ell)}) \right) \\
&= \liminf_{n\to\infty} \frac{1}{n} \sup_{k,\ell} \left(\log q_{k,\ell} + \log S_n(\bar{\mathbf{H}}^{(k,\ell)}) \right) \\
&= \liminf_{n\to\infty} \sup_{k,\ell} \left(W_n(\bar{\mathbf{H}}^{(k,\ell)}) + \frac{\log q_{k,\ell}}{n} \right) \\
&\geq \sup_{k,\ell} \liminf_{n\to\infty} \left(W_n(\bar{\mathbf{H}}^{(k,\ell)}) + \frac{\log q_{k,\ell}}{n} \right) \\
&= \sup_{k,\ell} \liminf_{n\to\infty} W_n(\bar{\mathbf{H}}^{(k,\ell)}),.
\end{aligned} \tag{7.18}$$

Because of Lemma 7.4, we can write

$$\begin{aligned}
W_n(\bar{\mathbf{H}}^{(k,\ell)}) &= \frac{1}{n} \sum_{i=1}^{n} \log \left\langle \bar{\mathbf{h}}^{(k,\ell)}(\mathbf{X}_1^{i-1}) \,,\, \mathbf{X}_i \right\rangle \\
&\geq \frac{1}{1-2\lambda} \left(\frac{1}{n} \sum_{i=1}^{n} U_M \left(\left\langle \bar{\mathbf{h}}^{(k,\ell)}(\mathbf{X}_1^{i-1}) \,,\, \mathbf{X}_i \right\rangle, \lambda \right) \right. \\
&\quad -\lambda \frac{1}{n} \sum_{i=1}^{n} \mathbb{E}^2 \left\{ \left\langle \bar{\mathbf{h}}^{(k,\ell)}(\mathbf{X}_1^{i-1}) \,,\, \mathbf{X}_i \right\rangle \mid \mathbf{X}_1^{i-1} \right\} \\
&\quad + \frac{1}{n} \sum_{i=1}^{n} g \left(\left\langle \bar{\mathbf{h}}^{(k,\ell)}(\mathbf{X}_1^{i-1}) \,,\, \mathbf{X}_i \right\rangle, \lambda \right) \\
&\quad \left. + \frac{1}{3n} \sum_{i=1}^{n} \frac{\left(\left\langle \bar{\mathbf{h}}^{(k,\ell)}(\mathbf{X}_1^{i-1}) \,,\, \mathbf{X}_i \right\rangle - 1 \right)^3}{\left\langle \bar{\mathbf{h}}^{(k,\ell)}(\mathbf{X}_1^{i-1}) \,,\, \mathbf{X}_i \right\rangle^3} \right)
\end{aligned} \tag{7.19}$$

First, we calculate the limes of the first term, because of Lemma 7.9, we get

$$\begin{aligned}
\lim_{n\to\infty} \frac{1}{n} \sum_{i=1}^{n} U_M \left(\left\langle \bar{\mathbf{h}}^{(k,\ell)}(\mathbf{X}_1^{i-1}) \,,\, \mathbf{X}_i \right\rangle, \lambda \right) &= \mathbb{E} \left\{ U_M \left(\left\langle \bar{\mathbf{b}}^*_{k,\ell}(\mathbf{X}_{-k}^{-1}) \,,\, \mathbf{X}_0 \right\rangle, \lambda \right) \right\} \\
&\stackrel{\text{def}}{=} \bar{\epsilon}_{k,\ell}
\end{aligned} \tag{7.20}$$

where $\bar{\mathbf{b}}^*_{k,\ell}(\mathbf{X}_{-k}^{-1})$ is a Markowitz-type portfolio with respect to the limit distribution $\mathbb{P}^{*(k,\ell)}_{\mathbf{X}_{-k}^{-1}}$. Let $\mathbf{b}^*_{k,\ell}(\mathbf{X}_{-k}^{-1})$ denote a log-optimal portfolio with respect to the limit distribution $\mathbb{P}^{*(k,\ell)}_{\mathbf{X}_{-k}^{-1}}$. Then

$$\begin{aligned}
\bar{\epsilon}_{k,\ell} &= \mathbb{E}\left\{U_M\left(\left\langle \bar{\mathbf{b}}^*_{k,\ell}(\mathbf{X}_{-k}^{-1}), \mathbf{X}_0\right\rangle, \lambda\right)\right\} \\
&\geq \mathbb{E}\left\{U_M\left(\left\langle \mathbf{b}^*_{k,\ell}(\mathbf{X}_{-k}^{-1}), \mathbf{X}_0\right\rangle, \lambda\right)\right\} \\
&\geq (1-2\lambda)\mathbb{E}\left\{\log\left\langle \mathbf{b}^*_{k,\ell}(\mathbf{X}_{-k}^{-1}), \mathbf{X}_0\right\rangle\right\} + \lambda\mathbb{E}^2\left\{\left\langle \mathbf{b}^*_{k,\ell}(\mathbf{X}_{-k}^{-1}), \mathbf{X}_0\right\rangle\right\} \\
&\quad -\mathbb{E}\left\{g\left(\left\langle \mathbf{b}^*_{k,\ell}(\mathbf{X}_{-k}^{-1}), \mathbf{X}_0\right\rangle, \lambda\right)\right\} - \frac{1}{3}\mathbb{E}\left\{\left(\left\langle \mathbf{b}^*_{k,\ell}(\mathbf{X}_{-k}^{-1}), \mathbf{X}_0\right\rangle - 1\right)^3\right\} \quad (7.21)
\end{aligned}$$

where last inequality follows from Lemma 7.4. Let us $\epsilon_{k,\ell} \stackrel{\text{def}}{=} \mathbb{E}\left\{\log\left\langle \mathbf{b}^*_{k,\ell}(\mathbf{X}_{-k}^{-1}), \mathbf{X}_0\right\rangle\right\}$. Combine (7.19), (7.20) and (7.21), then we obtain

$$\begin{aligned}
&\liminf_{n\to\infty} W_n(\bar{\mathbf{H}}^{(k,\ell)}) \\
&\geq \epsilon_{k,\ell} + \liminf_{n\to\infty} \frac{\lambda\sum_{i=1}^n\left(\mathbb{E}^2\left\{\left\langle \mathbf{b}^*_{k,\ell}(\mathbf{X}_{-k}^{-1}), \mathbf{X}_0\right\rangle - \mathbb{E}^2\left\{\left\langle \bar{\mathbf{h}}^{(k,\ell)}(\mathbf{X}_1^{i-1}), \mathbf{X}_i\right\rangle \mid \mathbf{X}_1^{i-1}\right\}\right\}\right)}{n(1-2\lambda)} \\
&\quad + \liminf_{n\to\infty} \frac{\sum_{i=1}^n\left(g\left(\left\langle \bar{\mathbf{h}}^{(k,\ell)}(\mathbf{X}_1^{i-1}), \mathbf{X}_i\right\rangle, \lambda\right) - \mathbb{E}\left\{g\left(\left\langle \mathbf{b}^*_{k,\ell}(\mathbf{X}_{-k}^{-1}), \mathbf{X}_0\right\rangle, \lambda\right)\right\}\right)}{n(1-2\lambda)} \\
&\quad + \liminf_{n\to\infty} \frac{\sum_{i=1}^n\left(\frac{\left(\left\langle \bar{\mathbf{h}}^{(k,\ell)}(\mathbf{X}_1^{i-1}), \mathbf{X}_i\right\rangle - 1\right)^3}{\left\langle \bar{\mathbf{h}}^{(k,\ell)}(\mathbf{X}_1^{i-1}), \mathbf{X}_i\right\rangle^3} - \mathbb{E}\left\{\left(\left\langle \mathbf{b}^*_{k,\ell}(\mathbf{X}_{-k}^{-1}), \mathbf{X}_0\right\rangle - 1\right)^3\right\}\right)}{3n(1-2\lambda)}.
\end{aligned}$$

Now bound the three additive terms separately. First,

$$\begin{aligned}
&\mathbb{E}^2\left\{\left\langle \mathbf{b}^*_{k,\ell}(\mathbf{X}_{-k}^{-1}), \mathbf{X}_0\right\rangle\right\} - \liminf_{n\to\infty}\frac{1}{n}\sum_{i=1}^n \mathbb{E}^2\left\{\left\langle \bar{\mathbf{h}}^{(k,\ell)}(\mathbf{X}_1^{i-1}), \mathbf{X}_i\right\rangle \mid \mathbf{X}_1^{i-1}\right\} \\
&= \mathbb{E}^2\left\{\left\langle \mathbf{b}^*_{k,\ell}(\mathbf{X}_{-k}^{-1}), \mathbf{X}_0\right\rangle\right\} - \mathbb{E}^2\left\{\langle \mathbf{b}', \mathbf{X}_0\rangle\right\} \quad (7.22) \\
&= \mathbb{E}\left\{\left\langle \mathbf{b}^*_{k,\ell}(\mathbf{X}_{-k}^{-1}), \mathbf{X}_0\right\rangle + \langle \mathbf{b}', \mathbf{X}_0\rangle\right\}\mathbb{E}\left\{\left\langle \mathbf{b}^*_{k,\ell}(\mathbf{X}_{-k}^{-1}), \mathbf{X}_0\right\rangle - \langle \mathbf{b}', \mathbf{X}_0\rangle\right\} \\
&\geq \mathbb{E}\left\{\left\langle \mathbf{b}^*_{k,\ell}(\mathbf{X}_{-k}^{-1}), \mathbf{X}_0\right\rangle + \langle \mathbf{b}', \mathbf{X}_0\rangle\right\} \\
&\quad \cdot\mathbb{E}\left\{\min_m \mathbb{E}\left\{1 + \log(X_0^{(m)}) - X_0^{(m)} \mid X_{-\infty}^{-1}\right\}\right\} \quad (7.23) \\
&\geq 2a^{-2}\mathbb{E}\left\{\min_m \mathbb{E}\left\{1 + \log(X_0^{(m)}) - X_0^{(m)}\middle| X_{-\infty}^{-1}\right\}\right\},
\end{aligned}$$

where (7.22) is true because of Lemma 7.10 ($\mathbf{b}'$ is a fix portfolio vector). (7.23) follows from Lemma 7.8.

For the second and third term we use (7.12) and Corollary 2 of Lemma 7.5 and also stationarity then we get,

$$\begin{aligned}
\liminf_{n\to\infty} W_n(\bar{\mathbf{H}}^{(k,\ell)}) &\geq \epsilon_{k,\ell} + \frac{2\lambda a^{-1}}{1-2\lambda}\mathbb{E}\left\{\min_m \mathbb{E}\left\{1 + \log(X_0^{(m)}) - X_0^{(m)}\middle|\mathbf{X}_{-\infty}^{-1}\right\}\right\} \\
&\quad - \left|\frac{2\lambda - \frac{1}{2}}{1-2\lambda}\right|\mathbb{E}\left\{\max_m \mathbb{E}\left\{(X_0^{(m)} - 1)^2\middle|\mathbf{X}_{-\infty}^{-1}\right\}\right\} \\
&\quad - \frac{a^{-3}+1}{3(1-2\lambda)}\mathbb{E}\left\{\max_m \mathbb{E}\left\{\left|X_0^{(m)} - 1\right|^3\middle|\mathbf{X}_{-\infty}^{-1}\right\}\right\}. \quad (7.24)
\end{aligned}$$

Györfi, Lugosi and Udina [10] proved that

$$\sup_{k,\ell} \epsilon_{k,\ell} = W^*,$$

therefore, by (7.18) and (7.24) the proof of the theorem is finished.

■

References

[1] P. Algoet. Universal schemes for prediction, gambling, and portfolio selection. *Annals of Probability*, 20:901–941, 1992.

[2] P. Algoet. The strong law of large numbers for sequential decisions under uncertainity. *IEEE Transactions on Information Theory*, 40:609–634, 1994.

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

[4] A. Blum and A. Kalai. Universal portfolios with and without transaction costs. *Machine Learning*, 35:193–205, 1999.

[5] A. Borodin, R. El-Yaniv, and V. Gogan. On the competitive theory and practice of portfolio selection (extended abstract). In *Proc. of the 4th Latin American Symposium on Theoretical Informatics (LATIN'00)*, pages 173–196, Punta del Este, Uruguay, 2000.

[6] L. Breiman. The individual ergodic theorem of information theory. *Annals of Mathematical Statistics*, 28:809–811, 1957. Correction. *Annals of Mathematical Statistics*, 31:809–810, 1960.

[7] L. Breiman. Optimal gambling systems for favorable games. In *Proc. of the Fourth Berkeley Symposium on Mathematical Statistics and Probability*, pages 65–78, Berkeley, 1961. University of California Press.

[8] T.M. Cover. Universal portfolios. *Mathematical Finance*, 1:1–29, 1991.

[9] R. Grauer. A comparison of growth optimal investment and mean variance investment policies. *Journal of Financial and Quantiative Analysis*, 16:1–21, 1981.

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

[11] L. Györfi and D. Schäfer. Nonparametric prediction. In J. A. K. Suykens, G. Horváth, S. Basu, C. Micchelli, and J. Vandevalle, editors, *Advances in Learning Theory: Methods, Models and Applications*, pages 339–354. IOS Press, NATO Science Series, 2003.

[12] L. Györfi, A. Urbán, and I. Vajda. Kernel-based semi-log-optimal empirical portfolio selection strategies, 2006. Accepted to International Journal of Theoretical and Applied Finance.

[13] N. H. Hakansson and W. T. Ziemba. Capital growth theory. In R. Jarrow, V. Maksimovic, and W. Ziemba, editors, *Handbooks in Operations Research and Management Science: Finance*, volume 9, pages 61–86. Elsevier Science, North-Holland, Amsterdam, 1995.

[14] D. P. Helmbold, R. E. Schapire, Y. Singer, and M. K. Warmuth. On-line portfolio selection using multiplicative updates. *Mathematical Finance*, 8:325–344, 1998.

[15] Y. Kroll, H. Levy, and H. Markowitz. Mean variance versus direct utility maximization. *Journal of Finance*, 39:47–75, 1984.

[16] H. Markowitz. Portfolio selection. *Journal of Finance*, 7(1):77–91, 1952.

[17] R. C. Merton. An intertemporal capital asset pricing model. *Econometrica*, 41(5):867–887, 1973.

[18] Y. Singer. Switching portfolios. *International Journal of Neural Systems*, 8:445–455, May 1997.

[19] James Tobin. Liquidity preference as behavior towards risk. *The Review of Economic studies*, (67):65–86, Feb. 1958.

György Ottucsák
Department of Computer Science and Information Theory
Budapest University of Technology and Economics
Magyar Tudósok körútja 2
H–1117 Budapest, Hungary
oti@szit.bme.hu

István Vajda
Department of Computer Science and Information Theory
Budapest University of Technology and Economics
Magyar Tudósok körútja 2
H–1117 Budapest, Hungary
vajda@szit.bme.hu