Received: 1 November 2016 | Accepted: 1 August 2018

DOI: 10.1111/mafi.12201

WILEY MATHEMATICAL FINANCE An International Journal of Mathematics, Statistics and Financial Economics

ORIGINAL ARTICLE

# Cover's universal portfolio, stochastic portfolio theory, and the numéraire portfolio

Christa Cuchiero[1] | Walter Schachermayer[1] | Ting-Kam Leonard Wong[2]

[1]University of Vienna, Vienna, Austria

[2]Department of Statistical Sciences, University of Toronto, Toronto, Ontario, Canada

**Correspondence**
Christa Cuchiero, Fakultät für Mathematik, Oskar-Morgenstern-Platz 1, 1090 Wien, Austria.
Email: christa.cuchiero@univie.ac.at

**Funding information**
Austrian Science Fund (FWF), Grants P25815 and P28861; Vienna Science and Technology Fund (WWTF), Grants MA14-008 and MA16-021

**Abstract**

Cover's celebrated theorem states that the long-run yield of a properly chosen "universal" portfolio is almost as good as that of the best *retrospectively chosen* constant rebalanced portfolio. The "universality" refers to the fact that this result is *model-free*, that is, not dependent on an underlying stochastic process. We extend Cover's theorem to the setting of stochastic portfolio theory: the market portfolio is taken as the numéraire, and the rebalancing rule need not be constant anymore but may depend on the current state of the stock market. By fixing a stochastic model of the stock market this model-free result is complemented by a comparison with the numéraire portfolio. Roughly speaking, under appropriate assumptions the asymptotic growth rate coincides for the three approaches mentioned in the title of this paper. We present results in both discrete and continuous time.

**KEYWORDS**

Diffusions on the unit simplex, ergodic Markov process, functionally generated portfolios, long-only portfolios, log-optimal portfolio, stochastic portfolio theory, universal portfolio

## 1 | INTRODUCTION

In Fernholz and Karatzas (2009), the question was raised whether there is a relation between Cover's theory of universal portfolio (which appeared as the very first paper of the present journal, see

Cover, 1991) and stochastic portfolio theory (SPT henceforth) as initiated by Fernholz (see Fernholz, 2002 and the references therein). After all, both theories ask for general recipes for choosing in a *preference-free* way good (at least in the long run) portfolios among $d$ assets, whose prices over time are given by

$$S = \left(S_t^1, \dots, S_t^d\right).$$

Here, the time $t$ varies in $\mathbb{T}$, where $\mathbb{T}$ stands either for $\mathbb{N} = \{0, 1, \dots\}$ (discrete time) or $\mathbb{R}_+ = [0, \infty)$ (continuous time). In many cases, $S$ is modeled by a stochastic process defined on some probability space. We note, however, that one may also consider a *model-free* approach where $S = (s_t^1, \dots s_t^d)_{t\in\mathbb{T}}$ is just a deterministic trajectory with values in $(0, \infty)^d$. Indeed, Cover and Ordentlich's discrete time results in Cover (1991) and Cover and Ordentlich (1996) are formulated in this model-free sense. The situation is more subtle in continuous time due to stochastic integration. Jamshidian (1992), extended Cover's universal portfolio to continuous time under a setting of Itô processes satisfying some asymptotic stability conditions.

In SPT, one also seeks robust investment strategies. More precisely, the strategies should be constructed using only observable quantities (such as market weights and their quadratic variations) and should not depend on quantities that are nonobservable or difficult to estimate. In particular, no drift estimation is involved which is usually required in expected utility maximization. These are exactly the principles behind the concept of *functionally generated portfolios* (see Fernholz, 2002, chapter 3). Although in most of the literature an Itô process setting is assumed, much of SPT can be developed in a model-free setting as done by Pal and Wong (2016) in discrete time and by Schied, Speiser, and Voloshchenko (2016) in continuous time. The reason why it works in continuous time is that the value processes of functionally generated portfolios can be defined without stochastic integration.

In this paper, we connect the two theories and provide additionally a comparison with the numéraire portfolio, which corresponds to the classical log-optimal portfolio.[1] Relationships between the two theories were studied in the recent papers by Ichiba and Brod (2014), and Brod (2014) as well as Wong (2015). In particular, Wong (2015) extends Cover's approach to the family of functionally generated portfolios in discrete time and shows that the distribution of wealth in this family satisfies a pathwise large deviation principle.

## 1.1 | Summary and discussion of the main results

In this paper, we work under the setting of SPT. Namely, the market portfolio is taken as the benchmark, or "numéraire," so that the primary assets are the market weights which take values in the open $d$-simplex defined by $\Delta^d = \{x \in (0,1)^d \mid \sum_{i=1}^d x^i = 1\}$. Its closure is denoted by $\bar{\Delta}^d = \{x \in [0,1]^d \mid \sum_{i=1}^d x^i = 1\}$. This enables us to analyze strategies which depend on the market weights, and the performance of *relative* wealth with respect to the market portfolio.

### 1.1.1 | Discrete time

We start by summarizing our results in discrete time. We extend Cover's universal portfolio to a class of $M$-Lipschitz portfolio maps denoted by $\mathscr{L}^M$. Each element of $\mathscr{L}^M$ maps the market weights to *long-only* portfolio weights in $\bar{\Delta}^d$ (see Definition 3.1).

Denoting by $(V_t^\pi)_{t=0}^\infty$ the *relative wealth process* corresponding to a portfolio strategy[2] $(\pi_t)_{t=1}^\infty$, we are interested in comparing the asymptotic growth rates

$$\lim_{T\to\infty} \frac{1}{T} \log\left(V_T^\pi\right)$$

for certain "optimal" portfolio choices $\pi$. More precisely, under suitable conditions we establish asymptotic equality of the growth rates of the following portfolios:

- the *best retrospectively chosen portfolio* at time $T$ in the class $\mathscr{L} := \bigcup_{M=1}^{\infty} \mathscr{L}^M$ (in this context $V_T^{*,M}$ will denote the relative wealth at time $T$ achieved by investing according to the best strategy in $\mathscr{L}^M$ over the time interval $[0, T]$);
- the *analog of Cover's universal portfolio* whose relative wealth process $(V_t(\nu))_{t=0}^{\infty}$ is defined in (17) (here $\nu$ is a probability measure on $\mathscr{L}$ with *full support* on each $\mathscr{L}^M$);
- the *log-optimal portfolio* among the class of *long-only* strategies, whose relative wealth process is denoted by $(\widehat{V}_t)_{t=0}^{\infty}$.

The first two portfolios can be compared in a model-free way (see Theorem 3.9). To compare them with the log-optimal portfolio, we have to introduce a probabilistic setting. Our main result can then be roughly stated as follows:

**Theorem 1.1.** *Let $(\mu_t)_{t=0}^{\infty}$ be a time-homogenous ergodic Markov process in discrete time describing the dynamics of the market weights. Then*

$$\lim_{M\to\infty} \lim_{T\to\infty} \frac{1}{T} \log\left(V_T^{*,M}\right) = \lim_{T\to\infty} \frac{1}{T} \log(V_T(\nu)) = \lim_{T\to\infty} \frac{1}{T} \log(\widehat{V}_T) \tag{1}$$

*holds almost surely.*

Intuitively, this theorem says that a suitable *full support* mixture of strategies (given by the universal portfolio) is asymptotically as good as the best one chosen with hindsight, and the log-optimal portfolio constructed with full knowledge of the underlying process.

### 1.1.2 | Continuous time

Theorem 1.1, which involves Lipschitz portfolio maps, cannot be extended directly to continuous time because of stochastic integrals. Instead, we consider *functionally generated portfolios* (see Section 4) whose relative wealth processes can be defined in a pathwise manner (see, e.g., Schied et al., 2016). This choice not only allows model-free considerations but also perfectly connects Cover's theory with SPT in continuous time. By replacing the set $\mathscr{L}^M$ by certain spaces of functionally generated portfolios and assuming that the log-optimal portfolio is functionally generated, we get essentially the same theorem as above.

Apart from the work by Jamshidian (1992), universal portfolio theory has only been studied sparingly in continuous time; see, for example, the paper Ichiba, Papathanakos, Banner, Karatzas, and Fernholz (2011) which studied the performance of the universal portfolio under the "Hybrid Atlas" model. To the best of our knowledge, generalizations to nonparametric families of portfolio maps (in continuous time) have not been considered so far. In this sense, our results significantly extend the continuous-time literature.

Although our approach focuses on the mathematical aspects, universal portfolio strategies have also been studied extensively in an algorithmic framework. See Li and Hoi (2014) for a recent survey and in particular Hazan and Kale (2015).

### 1.1.3 | Discussion of the results

Our model-free approach has clear advantages over classical ones which heavily rely on a particular model choice. Even in the case when the model class (e.g., the Heston model or Lévy models) is

correctly specified, model parameters cannot be estimated precisely and always come with a confidence interval. So, in practice, the *estimated optimal* portfolio is always different from the *true optimal* one. Our results support the idea that a Bayesian average in the spirit of Cover's universal portfolio is, in the long run, better than a suboptimal estimate.

As for the original theorems of Cover and Jamshidian, a valid criticism is of course that we only establish asymptotic equality on a first-order log-return basis. As such, a lot of important information is lost in the limit. However, one cannot expect to obtain any information on higher order terms unless further quantitative assumptions are made on the considered models. Cover's aim and also the goal of the present paper is to be as model-free as possible.[3] Nevertheless, it is of great theoretical and practical interest to strengthen the asymptotic results to quantitative ones under suitable additional conditions. We hope to address this important question in future research.

The remainder of the paper is organized as follows. In Section 2, we provide a brief overview (in discrete time for convenience) of the main topics of this paper, that is, Cover's theorem, the setting of SPT, and the log-optimal portfolio. In Section 3, we establish Theorem 1.1 in discrete time (see Theorem 3.10 and Corollary 3.11), whereas Section 4 is dedicated to proving the corresponding statements in continuous time in the setting of functionally generated portfolios and—for the comparison with the log-optimal portfolio—under the assumption that the market weights follow an ergodic Itô diffusion (see Theorem 4.11 and Corollary 4.13). Some auxiliary and technical proofs are gathered in the Appendix.

## 2 | OVERVIEW OF THE THREE PORTFOLIOS

For expositional simplicity, time is discrete in this section.

### 2.1 | Cover's universal portfolio

Cover's insight reveals that the "wisdom of hindsight" does not give significant advantages over a properly chosen "universal" portfolio constructed using only historical and current prices of the assets. The relevant optimality criterion here is the asymptotic growth rate of the portfolio.

Let us sketch this—at first glance surprising—result in a particularly easy setting (compare Cover, 1991; Cover & Ordentlich, 1996): Fix $T \in \mathbb{N}$ and think of an investor who at time $T$ looks back which stock she should have bought at time $t = 0$ (by investing her initial endowment and subsequently holding the stock). There is an obvious solution: pick $i \in \{1, \dots, d\}$ which maximizes the normalized logarithmic return

$$\frac{1}{T}\left(\log\left(S_T^i\right) - \log\left(S_0^i\right)\right). \tag{2}$$

The problem with this trading strategy is, of course, that we have to make our choice at time $t = 0$ instead of $t = T$. Here is the remedy (compare, e.g., Blum & Kalai, 1999): at time $t = 0$ simply divide the initial endowment, say $1e$, into $d$ portions of $\frac{1}{d}e$, invest each portion in each of the stocks and then hold the resulting portfolio. At time $T$, the normalized logarithmic return satisfies[4]

$$\frac{1}{T}\log(V_T) \geq \frac{1}{T}\log\left(\frac{1}{d}\sum_{j=1}^{d}\frac{S_T^j}{S_0^j}\right) \geq \frac{1}{T}\left(\log\left(S_T^i\right) - \log\left(S_0^i\right) - \log d\right), \tag{3}$$

where again $i$ denotes the stock which performed best during the time interval $[0, T]$. Hence, the difference between (2) and (3) can be bounded by $\frac{\log(d)}{T}$ which tends to zero as $T \to \infty$. Hence this buy-and-hold portfolio, which corresponds to a universal portfolio in the sense of Cover, has asymptotically the same normalized logarithmic return as the—only retrospectively known—best performing stock.

Instead of these "pure" investments, Cover considered a more ambitious setting, namely, all *constant rebalanced portfolio strategies*: let $b = (b^1, \dots, b^d) \in \bar{\Delta}^d$, that is, $b^j \geq 0$ and $\sum_{j=1}^d b^j = 1$. The value of the corresponding constant rebalanced portfolio $(V_t(b))_{t=0}^\infty$ starting at $V_0(b) = 1$ is defined by holding throughout the proportion $b^j$ of the current wealth in stock $j$, so that $V_0(b) = 1$ and

$$\frac{V_{t+1}(b)}{V_t(b)}(s) = \sum_{j=1}^d b^j \frac{s_{t+1}^j}{s_t^j} \tag{4}$$

for each trajectory $s = ((s_t^j)_{j=1}^d)_{t=0}^\infty \subset (0, \infty)^d$ of the stocks.

Fix again $T$ and define the quantity $V_T^*$ by

$$V_T^*(s) = \max_{b \in \bar{\Delta}^d} V_T(b)(s), \tag{5}$$

which is a function of the trajectory $s = (s_t^1, \dots, s_t^d)_{t=0}^T$. Again, the idea is that, with hindsight, that is, knowing $(s_t^1, \dots, s_t^d)_{t=0}^T$, one considers the best weight $b \in \bar{\Delta}^d$ which attains the maximum (5). Cover's goal is to construct a portfolio which generates wealth that performs asymptotically as well as the process $(V_T^*)_{T=0}^\infty$ as $T \to \infty$, uniformly over all price paths. For this reason, the portfolio is said to be *universal*. In order to do so, let $\nu$ be a probability measure on $\bar{\Delta}^d$ which replaces the previous uniform distribution over the $d$ stocks. The universal portfolio is built by investing at time 0 the portion $d\nu(b)$ of initial capital in the constant rebalanced portfolio $V(b)$ and by subsequently following the constant rebalanced portfolio process $(V_t(b))_{t=0}^T$. The explicit formula for the wealth is

$$V_t(\nu)(s) = \int_{\bar{\Delta}^d} V_t(b)(s) d\nu(b), \tag{6}$$

where $V_t(b)$ is defined by (4). The portfolio weight of the corresponding *universal portfolio* is given by the wealth-weighted average

$$b_t^\nu(s) = \frac{\int_{\bar{\Delta}^d} b V_t(b)(s) d\nu(b)}{\int_{\bar{\Delta}^d} V_t(b)(s) d\nu(b)}. \tag{7}$$

Let us now recall Cover's celebrated result:

**Theorem 2.1.** *(Cover, 1991): Let $\nu$ be a probability measure on $\bar{\Delta}^d$ with full support. Then*

$$\lim_{T \to \infty} \frac{1}{T} \left( \log(V_T(\nu)(s)) - \log\left(V_T^*(s)\right) \right) = 0 \tag{8}$$

*for* all *trajectories $s = (s_t^1, \dots, s_t^d)_{t=0}^\infty$ for which there are constants $0 < c \leq C < \infty$ such that*

$$c \leq \frac{s_{t+1}^j}{s_t^j} \leq C, \qquad \textit{for all} \quad j = 1, \dots, d \quad \textit{and all} \quad t \in \mathbb{N}. \tag{9}$$

The proof is given in the Appendix.

*Remark* 2.2. As shown by Cover and Ordentlich (1996), the condition (9) can be dropped at least when $\nu$ is the uniform or Dirichlet $(\frac{1}{2}, \cdots, \frac{1}{2})$ distribution on $\Delta^d$ (see also Blum & Kalai, 1999, for an elegant proof in case of the uniform distribution).

*Remark* 2.3. Let $\mathcal{M}_1(\bar{\Delta}^d)$ be the set of probability measures on $\bar{\Delta}^d$. For each $\mu \in \mathcal{M}_1(\bar{\Delta}^d)$, consider the value $\int_{\bar{\Delta}^d} V_T(b)(s)d\mu(b)$ of the mixture portfolio with initial measure $\mu$. Note that the constant rebalanced portfolio $V_T(b)$ corresponds to the case where $\mu$ is the point mass at $b$. It is easy to see that

$$\sup_{\mu \in \mathcal{M}_1(\bar{\Delta})} \int_{\bar{\Delta}^d} V_T(b)(s)d\mu(b) = V_T^*(s),$$

where $V_T^*(s)$ is defined by (5). It follows that the universal portfolio (6) (with initial measure $\nu$) is still asymptotically optimal in the larger class

$$\left\{ \left( \frac{\int_{\bar{\Delta}^d} bV_t(b)d\mu(b)}{\int_{\bar{\Delta}^d} V_t(b)(s)d\mu(b)} \right)_{t \geq 0} \mid \mu \in \mathcal{M}_1(\bar{\Delta}^d) \right\}. \tag{10}$$

## 2.2 | SPT, portfolio maps, and the corresponding universal portfolio

In SPT, we let $(s^1, \ldots, s^d)$ denote the market capitalizations of the stocks rather than their prices. Then we define the vector of *market weights* $(\mu^1, \ldots, \mu^d) \in \Delta^d$ by

$$(\mu^1, \ldots, \mu^d) = \left( \frac{s^1}{s^1 + \cdots + s^d}, \ldots, \frac{s^d}{s^1 + \cdots + s^d} \right).$$

This amounts to taking the *market portfolio* (whose value at time $t$ is $\sum_{j=1}^d s_t^j$) as the *numéraire* (compare Delbaen & Schachermayer, 1995 and Fernholz & Karatzas, 2010a).

The *relative wealth process* $(V_t^\pi)_{t=0}^\infty$, expressed in units of the market portfolio and starting at $V_0 = 1$, is obtained by the following recursive relation:[5]

$$\frac{V_{t+1}^\pi}{V_t^\pi} = \sum_{j=1}^d \pi_{t+1}^j \frac{\mu_{t+1}^j}{\mu_t^j}. \tag{11}$$

In general, we allow all predictable, admissible trading strategies $(\pi_t)_{t=1}^\infty$, where the portfolio weight $\pi_t$ is used over the time interval $[t-1, t]$. In this paper, all trading strategies are fully invested in the equity market, that is, the portfolio weights sum to 1 for all $t$. In particular, the strategies do not lend or borrow money. Henceforth, all wealth processes are measured in units of the market portfolio.

We will focus on trading strategies defined by (deterministic) *portfolio maps*. These are (Borel) measurable functions

$$\pi : \Delta^d \to \bar{\Delta}^d, \tag{12}$$

which associate to the current market capitalization $\mu_t = (\mu_t^1, \ldots, \mu_t^d)$ the weights $(\pi(\mu_t) = (\pi^1(\mu_t), \ldots, \pi^d(\mu_t))$ according to which an agent distributes current wealth among the $d$ stocks at time $t$. The *constant rebalanced portfolio strategies* considered by Cover correspond to the constant functions $\pi : \Delta^d \to \bar{\Delta}^d$.

In this paper, we extend Cover's theory of *constant rebalanced portfolios* to certain families of *portfolio maps*. First, we note that Cover's and Jamshidian's definition of a universal portfolio as in (7) and (6) can be easily extended to a general setting. Let $\mathcal{G}$ denote some appropriate space of portfolio maps, $\mathcal{B}(\mathcal{G})$ its Borel $\sigma$-algebra and $\nu$ some probability measure on $\mathcal{G}$.

**Definition 2.4.** Let $\nu$ be a probability measure on $(\mathcal{G}, \mathcal{B}(\mathcal{G}))$. Then, the corresponding universal portfolio at time $t$ is given by the wealth-weighted average

$$\pi_t^\nu = \frac{\int_{\mathcal{G}} \pi V_t^\pi d\nu(\pi)}{\int_{\mathcal{G}} V_t^\pi d\nu(\pi)}. \tag{13}$$

From (11), it is easily seen that the wealth generated by $\pi^\nu$ is given by

$$V_T(\nu) = \int_{\mathcal{G}} V_T^\pi d\nu(\pi). \tag{14}$$

## 2.3 | The log-optimal portfolio

To define the log-optimal portfolio, we consider a probabilistic setting. The stock price process $S = (S_t^1, \dots, S_t^d)_{t=0}^\infty$ and the corresponding relative market capitalizations $\mu = (\mu_t^1, \dots, \mu_t^d)_{t=0}^\infty$ are now assumed to be stochastic processes defined on a filtered probability space $(\Omega, \mathcal{F}, (\mathcal{F}_t)_{t=0}^\infty, \mathbb{P})$.

There is a large literature on the log-optimal portfolio (see, e.g., Becherer, 2001; Karatzas & Kardaras, 2007 and the references given there). For a fixed horizon $T$, this portfolio is by definition the maximizer of the expected logarithmic growth rate

$$\mathbb{E}[\log(V_T^\pi)] = \mathbb{E}\left[\sum_{t=0}^{T-1} \log\left(\sum_{j=1}^d \pi_{t+1}^j \frac{\mu_{t+1}^j}{\mu_t^j}\right)\right] \tag{15}$$

over all predictable, admissible trading strategies $(\pi_t)_{t=1}^T$. Under mild assumptions on the process a unique optimizer exists; see, for example, Becherer (2001) and Kramkov and Schachermayer (1999).

To connect the log-optimal portfolio with universal portfolios in the sense of Definition 2.4, we need appropriate assumptions. We will assume that $\mu$ is a time-homogenous Markov process, and we will restrict to *long-only* portfolios in the optimization of (15). These imply that the optimal portfolio in (15) (over the set of predictable processes taking values in $\bar{\Delta}^d$) has the form $\pi_t = \pi(\mu_{t-1})$, where $\pi : \Delta^d \mapsto \bar{\Delta}^d$ as in (12). We denote the corresponding optimizer by $\widehat{\pi}$.

The Markovian assumption can be motivated by the stability of capital distributions of equity markets (see Fernholz, 2002, chapter 5). In SPT, this led to systems of interacting Brownian particles whose dynamics depend on their relative rankings. Under suitable conditions, these systems show behaviors observed in large equity markets. See, for example, Banner, Fernholz, and Karatzas (2005) and Ichiba et al. (2011) for "Atlas"-type models and the references therein.[6] We also refer to Kardaras and Robertson (2012) which studies the growth optimal portfolio in a Markovian setting with uncertainties.

# 3 | A COMPARISON OF THE THREE APPROACHES—THE DISCRETE TIME CASE

Throughout this section, we work in discrete time and assume that the market weights are described by a $d$-dimensional path $\mu = (\mu_t)_{t=0}^{\infty}$ with values in $\Delta^d$. We consider as far as possible a model-free approach, but will introduce a probabilistic setting when the log-optimal portfolio is involved.

## 3.1 | Definitions of the portfolios

We start by defining rigorously, in the present setting, the three portfolios introduced in Sections 1 and 2.

### 3.1.1 | The best retrospectively chosen portfolio

Consider Cover's theme of choosing retrospectively at time $T$ a strategy which is optimal within a certain class of strategies, in our case portfolio maps $\pi : \Delta^d \to \bar{\Delta}^d$. A moment's reflection reveals that it does not make sense to allow to choose among *all measurable* functions $\pi : \Delta^d \to \bar{\Delta}^d$. Indeed, there is no restriction to choose $\pi$ such that $\pi(\mu_t) = e_{j(t)}$, where $j(t) \in \{1, \dots, d\}$ maximizes $\mu_{t+1}^j / \mu_t^j$. This is asking for too much clairvoyance and does not allow for meaningful results (compare Cover & Ordentlich, 1996; and Blum & Kalai, 1999, section 5).

However, it does make sense (economically as well as mathematically) to restrict to more regular trading strategies. In particular, we work with the following set of $M$-Lipschitz portfolio maps. For $\epsilon > 0$, we let $\bar{\Delta}_\epsilon^d$ denote the set of $x \in \Delta^d$ satisfying $x^j \geq \frac{\epsilon}{d}$, for $j = 1, \dots, d$. Also we let $\| \cdot \|_1$ be the usual 1-norm.

**Definition 3.1.** For $M > 0$, we denote by $\mathscr{L}^M$ the set of all $M$-Lipschitz functions $\Delta^d \to \bar{\Delta}^d_{M^{-1}}$, that is, $\|\pi(x) - \pi(y)\|_1 \leq L\|x - y\|_1$, $x, y \in \Delta^d$.

*Remark* 3.2. The set $\mathscr{L}^M$ of $M$-Lipschitz functions $\pi : \Delta^d \to \bar{\Delta}^d_{M^{-1}}$ is a compact metric space with respect to the topology of uniform convergence induced by the norm $\|\pi\|_\infty = \sup\{\|\pi(x)\|_1 : x \in \Delta^d\}$.

*Remark* 3.3. Instead of Lipschitz functions we could just as well consider other compact function spaces, for example, Hölder spaces equipped with a proper norm. This is done in the context of functionally generated portfolios in Section 4.

The retrospectively chosen best performing portfolio among the above Lipschitz maps is defined as follows:

**Definition 3.4.** For a given trajectory $(\mu_t)_{t=0}^T \in (\Delta^d)^{T+1}$, we define

$$V_T^{*,M} = \sup_{\pi \in \mathscr{L}^M} V_T^\pi = \sup_{\pi \in \mathscr{L}^M} \prod_{t=0}^{T-1} \left( \sum_{j=1}^d \pi^j(\mu_t) \frac{\mu_{t+1}^j}{\mu_t^j} \right). \tag{16}$$

By compactness (see Remark 3.2) and continuity of the map $\pi \mapsto V_T^\pi$, there exists an optimizer $\pi^{*,M} \in \mathscr{L}^M$ (not necessarily unique) such that $V_T^{*,M} = V_T^{\pi^{*,M}}$, thus the sup above can be replaced by max.

### 3.1.2 | The universal portfolio

Our aim is to find a *predictable* process $\pi^M = (\pi_t^M)_{t=1}^\infty$, that is, one which depends only on the history of the market weights, such that the performance of $(V_t^{\pi^M})_{t=0}^\infty$ is asymptotically as good as that of $(V_t^{*,M})_{t=0}^\infty$. This can be achieved by the universal portfolio introduced in Definition 2.4, where the $\mathcal{G}$ is now $\mathscr{L}^M$ as in Definition 3.1. As $\mathscr{L}^M$ is a compact metric space, we may find a (Borel) probability measure $\nu$ on $(\mathscr{L}_M, \|\cdot\|_\infty)$ with full support; this will be essential for establishing an analog to Theorem 2.1. The (relative) wealth of the universal portfolio is given, as in (14), by

$$V_T^M(\nu) = \int_{\mathscr{L}^M} V_T^\pi \, d\nu(\pi). \tag{17}$$

### 3.1.3 | The log-optimal portfolio

In order to relate the universal portfolio to the (long-only) log-optimal portfolio, we assume that $\mu = (\mu_t)_{t=0}^\infty$ is a time-homogeneous Markov process (see Section 2.3). Here is a precise statement.

**Assumption 3.5.** The process $\mu$ is a time homogeneous, ergodic Markov process with a unique invariant measure $\varrho$ on the open simplex $\Delta^d$.

We denote the transition kernel of the chain by $(\varrho(x,\cdot))_{x\in\Delta^d}$, that is, for all Borel sets $A \subseteq \bar{\Delta}^d$ we have $\mathbb{P}[\mu_{t+1} \in A|\mathcal{F}_t] = \varrho(\mu_t, A)$. For further notions concerning ergodic Markov processes, we refer to Eberle (2016).

The *long-only log-optimal* trading strategy $\widehat{\pi}$, as noted above, is given in terms of a portfolio map. Given that $\mu_t = x \in \Delta^d$, we know the conditional law $\varrho(x,\cdot)$ of $\mu_{t+1}$. We therefore choose $\widehat{\pi}(x) \in \bar{\Delta}^d$ as the maximizer

$$\widehat{\pi}(x) = \operatorname*{arg\,max}_{p\in\bar{\Delta}^d} \left( \int_{\Delta^d} \log\left( \left\langle p, \frac{y}{x} \right\rangle \right) \varrho(x, dy) \right) \tag{18}$$

and assume that $\widehat{\pi}(\cdot)$ can be chosen to be measurable (here $\langle , \rangle$ denotes the Euclidean dot product). For $x \in \Delta^d$, define the number $L(x)$ as the value of the optimization problem (18), that is,

$$L(x) = \max_{p\in\bar{\Delta}^d} \left( \int_{\Delta_d} \log\left( \left\langle p, \frac{y}{x} \right\rangle \right) \varrho(x, dy) \right) = \int_{\Delta_d} \log\left( \left\langle \widehat{\pi}(x), \frac{y}{x} \right\rangle \right) \varrho(x, dy). \tag{19}$$

Considering $\pi(x) = x$ (which corresponds to the market portfolio) we clearly have $L(x) \geq 0$ for each $x \in \Delta^d$. We obtain the a.s. relation

$$L(x) = \mathbb{E}\left[ \log\left( \frac{\widehat{V}_{t+1}}{\widehat{V}_t} \right) \middle| \mu_t = x \right],$$

where $\widehat{V} = (\widehat{V}_t)_{t=0}^\infty$ denotes the long-only log-optimal wealth process $V^{\widehat{\pi}}$ defined by the portfolio map $\widehat{\pi}$ via (11).

**Assumption 3.6.** Using the above notation we assume that

$$L := \int_{\Delta^d} L(x) d\varrho(x) < \infty. \tag{20}$$

Applying Birkhoff's ergodic theorem for discrete time Markov processes (see Eberle, 2016, theorem 2.2, section 2.1.4), we have the following result.

**Theorem 3.7.** *Under Assumptions 3.5 and 3.6, we have that, for $\varrho$-a.e. starting value $\mu_0 \in \Delta^d$,*

$$\lim_{T\to\infty} \frac{1}{T} \log(\widehat{V}_T) = L, \tag{21}$$

*the limit holding true a.s. as well as in $L^1$.*

*More generally, let $\pi : \Delta^d \to \bar{\Delta}^d$ be any measurable portfolio map such that*

$$L^\pi := \int_{\Delta^d} \left( \int_{\Delta^d} \log \left( \left\langle \pi(x), \frac{y}{x} \right\rangle \right) \varrho(x, dy) \right) d\varrho(x) > -\infty. \tag{22}$$

*We then have, for $\varrho$-a.s. starting value $\mu_0$, that*

$$\lim_{T\to\infty} \frac{1}{T} \log \left( V_T^\pi \right) = L^\pi \tag{23}$$

*a.s. as well as in $L^1$.*

In general, there is little reason why the function $\widehat{\pi}$ should have better regularity properties than being just measurable. On the other hand, we may approximate $\widehat{\pi}$ by more regular functions, in particular by functions in $\mathscr{L}^M$. This will be crucial for comparing the asymptotic growth rates. The following result is intuitively obvious, but the proof turns out to be quite technical and will be given in the Appendix.

**Lemma 3.8.** *Under Assumptions 3.5 and 3.6, for any $\epsilon > 0$ there exist $M > 0$ and an $M$-*Lipschitz function *$\pi_{Lip} \in \mathscr{L}^M$ such that*

$$L^{\pi_{Lip}} > L - \epsilon,$$

*where $L$ and $L^\pi$ are given in (20) and (22), respectively. In particular, we have $L = \sup_M \sup_{\pi\in\mathscr{L}^M} L^\pi$.*

## 3.2 | Asymptotically equivalent growth rates

We are now ready to compare the asymptotic performance of the three approaches. We first establish an analog of Theorem 2.1.

**Theorem 3.9.** *Fix $M > 0$ and a Borel probability measure $\nu$ with full support on $\mathscr{L}^M$. For every trajectory $(\mu_t)_{t=0}^\infty$ in $\Delta^d$, we have*

$$\lim_{T\to\infty} \frac{1}{T} \left( \log \left( V_T^{*,M} \right) - \log \left( V_T^M(\nu) \right) \right) = 0. \tag{24}$$

*Proof.* The inequality "$\geq$" is obvious. For the reverse inequality, we follow the argument of Blum and Kalai (1999). As $\mathscr{L}^M$ is compact and $\nu$ has full support, it is not difficult to see that for any $\eta > 0$, there exists $\delta > 0$ such that every $\eta$-neighborhood of a point $\pi \in \mathscr{L}^M$ has $\nu$-measure bigger than $\delta$.

Let a trajectory $(\mu_t)_{t=0}^\infty$ in $\Delta^d$ be given. For a fixed time $T$, let $\pi^{*,M} \in \mathscr{L}^M$ be an optimizer of (16). Consider a portfolio map $\pi^M \in \mathscr{L}^M$ with $\|\pi^M - \pi^{*,M}\|_\infty < \eta$, that is, such that, for every $x \in \Delta^d$ we have $\|\pi^M(x) - \pi^{*,M}(x)\|_1 = \sum_{j=1}^d |\pi^M(x)^j - \pi^{*,M}(x)^j| < \eta$.

Choose $\eta > 0$ small enough so that $\alpha = \eta M d < 1$ and define, for $x \in \Delta^d$,

$$\widetilde{\pi}(x) = \frac{1}{\alpha}\pi^M(x) - \frac{1-\alpha}{\alpha}\pi^{*,M}(x). \tag{25}$$

Rearranging, we have

$$\pi^M(x) = (1-\alpha)\pi^{*,M}(x) + \alpha\widetilde{\pi}(x). \tag{26}$$

It is easy to see that $\widetilde{\pi}$ maps $\Delta^d$ into $\bar{\Delta}^d$.

Using (26), we have the estimate

$$\begin{aligned}\frac{1}{T}\log V_T^{\pi^M} &= \frac{1}{T}\sum_{t=0}^{T-1}\log\left(\left\langle \pi^M(\mu_t), \frac{\mu_{t+1}}{\mu_t}\right\rangle\right)\\ &\geq \frac{1}{T}\sum_{t=0}^{T-1}\log\left(\left\langle (1-\alpha)\pi^{*,M}(\mu_t), \frac{\mu_{t+1}}{\mu_t}\right\rangle\right)\\ &= \frac{1}{T}\log\left(V_T^{*,M}\right) + \log(1-\alpha).\end{aligned} \tag{27}$$

Fix $\epsilon > 0$. Choosing $\eta > 0$ sufficiently small we can make $\alpha = \eta M d$ small enough such that the final term is bigger than $-\epsilon$. Summing up, we have

$$\frac{1}{T}\left[\log\left(V_T^{*,M}\right) - \log\left(V_T^{\pi^M}\right)\right] < \epsilon \tag{28}$$

whenever $\|\pi^M - \pi^{*,M}\|_\infty < \eta$.

Denote by $B = B_\eta(\pi^{*,M})$ the $\|\cdot\|_\infty$-ball with radius $\eta$ in $\mathscr{L}^M$ which has $\nu$-measure at least $\delta > 0$, where $\delta$ only depends on $\eta$. As each element $\pi^M$ of $B$ satisfies (28) we have

$$\frac{1}{T}\log\left(V_T^M(\nu)\right) \geq \frac{\log(\delta)}{T} + \frac{1}{T}\log\left(V_T^{*,M}\right) - \epsilon. \tag{29}$$

Now (24) is proved by sending in (29) $T$ to infinity and letting $\epsilon$ to zero. □

Note that in Theorem 3.9 we do not need the uniform boundedness condition (9) (compare this result with Wong, 2015, lemma 3.3). We now combine Lemma 3.8 (which is probabilistic) with Theorem 3.9 (which is pathwise) to obtain—under suitable assumptions–equality of the asymptotic performance among the three portfolios. We first consider the space $\mathscr{L}^M$ for a fixed $M$. In Corollary 3.11, we then formulate a result for $\mathscr{L} = \bigcup_M \mathscr{L}^M$.

**Theorem 3.10.** *Let $\Omega = (\Delta^d)^{\mathbb{N}}$ be the canonical path space equipped with its natural filtration and a probability measure $\mathbb{P}$. Define $\mu = (\mu_t)_{t=0}^\infty$ to be the canonical process, that is, $\mu_t(\omega) = \omega_t$, which takes values in $\Delta^d$ and satisfies Assumptions 3.5 and 3.6. Moreover, let $M > 0$ be a fixed Lipschitz constant for the space $\mathscr{L}^M$. Consider the following objects that are defined for each trajectory $(\mu_t)_{t=0}^\infty$:*[7]

- **(i)** *Define for each $T \in \mathbb{N}$ the portfolio map $\pi^{*,M} \in \mathscr{L}^M$ as well as the corresponding wealth $V_T^{*,M} := V_T^{\pi^{*,M}}$ as in (16).*
- **(ii)** *Fix a probability measure $\nu$ on $\mathscr{L}^M$ with full support and consider the wealth process of the universal portfolio $(V_t^M(\nu))_{t=0}^\infty$ as of (17).*

**(iii)** *Define the log-optimal portfolio among the portfolio maps $\pi \in \mathscr{L}^M$ by*

$$\widehat{\pi}^M = \operatorname*{arg\,max}_{\pi \in \mathscr{L}^M} \int_{\Delta^d} \left[ \int_{\Delta^d} \log\left( \left\langle \pi(x), \frac{y}{x} \right\rangle \right) \varrho(x, dy) \right] d\varrho(x) \tag{30}$$

*and the corresponding wealth process $(\widehat{V}_t^M)_{t=0}^\infty = (V_t^{\widehat{\pi}^M})_{t=0}^\infty$ via (11).*

*Then, we have $\mathbb{P}$-a.s.*

$$\liminf_{T\to\infty} \frac{1}{T} \log\left(V_T^{*,M}\right) = \liminf_{T\to\infty} \frac{1}{T} \log\left(V_T^M(\nu)\right) = \lim_{T\to\infty} \frac{1}{T} \log\left(\widehat{V}_T^M\right) = \sup_{\pi\in\mathscr{L}^M} L^\pi, \tag{31}$$

*where $L^\pi$ is given in (22). In addition, the first equality holds for* all *trajectories $(\mu_t)_{t=0}^\infty$ in $\Delta^d$.*

*Proof.* We first note that $\widehat{\pi}^M$ is well-defined; simply use the compactness of $\mathscr{L}^M$ with respect to $\|\cdot\|_\infty$ (compare the proof of Lemma 3.8). Note also that by the ergodic theorem (Theorem 3.7), we have for each $\pi \in \mathscr{L}^M$

$$\lim_{T\to\infty} \frac{1}{T} \log V_T^\pi = L^\pi \quad \mathbb{P}\text{-a.s.},$$

where $L^\pi$ is defined by (22). In particular, as $\widehat{\pi}^M \in \mathscr{L}^M$ by definition, we have

$$\lim_{T\to\infty} \frac{1}{T} \log \widehat{V}_T^M = \sup_{\pi\in\mathscr{L}^M} L^\pi \quad \mathbb{P}\text{-a.s.} \tag{32}$$

That the first equality in (31) holds for *all* trajectories $(\mu_t)_{t=0}^\infty$ in $\Delta^d$ was shown in Theorem 3.9. For each fixed $T \in \mathbb{N}$, we obviously have

$$\frac{1}{T} \log\left(\widehat{V}_T^M\right) \le \frac{1}{T} \log\left(V_T^{*,M}\right) \quad \mathbb{P}\text{-a.s.} \tag{33}$$

Using (32), (33), and Theorem 3.9 we thus have $\mathbb{P}$-a.s.

$$\sup_{\pi\in\mathscr{L}^M} L^\pi = \lim_{T\to\infty} \frac{1}{T} \log\left(\widehat{V}_T^M\right) \le \liminf_{T\to\infty} \frac{1}{T} \log\left(V_T^{*,M}\right) = \liminf_{T\to\infty} \frac{1}{T} \log\left(V_T^M(\nu)\right). \tag{34}$$

On the other hand, by the definition of $(\widehat{V}_t^M)_{t=0}^\infty$ as the log-optimizer within the class $\mathscr{L}^M$, we have

$$\mathbb{E}\left[\log\left(V_T^M(\nu)\right)\right] \le \sup_{\pi\in\mathscr{L}^M} \mathbb{E}\left[\log\left(V_T^\pi\right)\right] = \mathbb{E}\left[\log\left(\widehat{V}_T^M\right)\right]. \tag{35}$$

To see this, note that the universal portfolio is given by (13). By the time-homogenous Markovianity it is thus sufficient to dominate the left-hand side of (35) by taking the supremum over elements in $\mathscr{L}^M$.

Combining now (35), Theorem 3.7 and (34) yields that

$$\begin{aligned}
\mathbb{E}\left[\liminf_{T\to\infty} \frac{1}{T} \log\left(V_T^M(\nu)\right)\right] &\le \liminf_{T\to\infty} \frac{1}{T} \mathbb{E}\left[\log V_T^M(\nu)\right] \\
&\le \lim_{T\to\infty} \frac{1}{T} \mathbb{E}\left[\log\left(\widehat{V}_T^M\right)\right] \\
&= \lim_{T\to\infty} \frac{1}{T} \log\left(\widehat{V}_T^M\right)
\end{aligned}$$

$$\leq \liminf_{T\to\infty} \frac{1}{T} \log V_T^{*,M}$$

$$= \liminf_{T\to\infty} \frac{1}{T} \log \left(V_T^M(\nu)\right), \quad \mathbb{P}\text{-a.s.}$$

Here, the first inequality follows from Fatou's lemma (note here that $\frac{1}{T}\log(V_T^M(\nu))$ is bounded from below, see, e.g., (29)). From this we see that the quantity $\liminf_{T\to\infty} \frac{1}{T}\log(V_T^M(\nu))$ is $\mathbb{P}$-a.s. constant and equal to $\lim_{T\to\infty} \frac{1}{T}\log(\widehat{V}_T^M)$. This completes the proof of the theorem. □

Next we will send $M$ to infinity in the following way. For $M = 1, 2, 3, \ldots$ choose a measure $\nu^M$ on $\mathscr{L}^M$ with full support. Define $\nu = \sum_{M=1}^{\infty} 2^{-M}\nu^M$ and the wealth of the universal portfolio $V(\nu)$ as in (17) by

$$V_t(\nu) = \int_{\mathscr{L}} V_t^{\pi} d\nu(\pi), \qquad t \in \mathbb{N}. \tag{36}$$

where $\mathscr{L} = \bigcup_{M=1}^{\infty} \mathscr{L}^M$. Recall that $(\widehat{V}_t)_{t=0}^{\infty}$ is the wealth process of the (long-only) log-optimal portfolio (18).

**Corollary 3.11.** *Under the assumptions of Theorem 3.10 we have $\mathbb{P}$-a.s.*

$$\lim_{M\to\infty} \lim_{T\to\infty} \frac{1}{T} \log V_T^{*,M} = \lim_{T\to\infty} \frac{1}{T} \log V_T(\nu) = \lim_{T\to\infty} \frac{1}{T} \log \widehat{V}_T = L, \tag{37}$$

*where $L$ is defined in (20).*

*Proof.* Letting $M \to \infty$ in (31), we have

$$\lim_{M\to\infty} \liminf_{T\to\infty} \frac{1}{T} \log V_T^{*,M} = \lim_{M\to\infty} \sup_{\pi\in\mathscr{L}^M} L^{\pi} = L = \lim_{T\to\infty} \frac{1}{T} \log \hat{V}_T,$$

where the last equality follows from Theorem 3.7 and the second last follows from Lemma 3.8. By construction $V_T(\nu) \geq 2^{-M} V_T^M(\nu^M)$ for every $M$, so we have by Theorem 3.9 for every $M$

$$\liminf_{T\to\infty} \frac{1}{T} \log V_T(\nu) \geq \liminf_{T\to\infty} \frac{1}{T} \left(-M \log 2 + \log V_T^M(\nu^M)\right) = \liminf_{T\to\infty} \frac{1}{T} \log V_T^{*,M},$$

and hence also

$$\liminf_{T\to\infty} \frac{1}{T} \log V_T(\nu) \geq \lim_{M\to\infty} \liminf_{T\to\infty} \frac{1}{T} \log V_T^{*,M}.$$

Using the same argument as in the last part of the proof of Theorem 3.10, we get

$$\lim_{M\to\infty} \liminf_{T\to\infty} \frac{1}{T} \log V_T^{*,M} = \liminf_{T\to\infty} \frac{1}{T} \log V_T(\nu) = \lim_{T\to\infty} \frac{1}{T} \log \widehat{V}_T = L. \tag{38}$$

Now the corollary is proved if

$$\limsup_{T\to\infty} \frac{1}{T}(\log V_T(\nu) - \log \widehat{V}_T) = \limsup_{T\to\infty} \frac{1}{T} \log \left(\frac{V_T(\nu)}{\widehat{V}_T}\right) = 0 \tag{39}$$

holds $\mathbb{P}$-a.s. As by Lemma 3.12, $(\frac{V_t(\nu)}{\widehat{V}_t})_{t=0}^{\infty}$ is a nonnegative supermartingale, it converges $\mathbb{P}$-a.s. to a finite limit as $t \to \infty$. This in turn implies (39) and proves the assertion. □

**Lemma 3.12.** *The process $(\frac{V_t(\nu)}{\widehat{V}_t})_{t=0}^{\infty}$ is a nonnegative supermartingale.*

*Proof.* First note that for any $\pi : \Delta^d \to \bar{\Delta}^d$, $(\frac{V_t^\pi}{\widehat{V}_t})_{t=0}^{\infty}$ is a nonnegative supermartingale. Indeed, by Lemma 3.13 we have

$$E\left[\frac{V_{t+1}^\pi}{\widehat{V}_{t+1}}\Big|\mathcal{F}_t\right] = \frac{V_t^\pi}{\widehat{V}_t}\int_{\Delta^d}\frac{\langle\pi(\mu_t),\frac{y}{\mu_t}\rangle}{\langle\widehat{\pi}(\mu_t),\frac{y}{\mu_t}\rangle}\varrho(\mu_t,dy) \le \frac{V_t^\pi}{\widehat{V}_t}.$$

By Fubini's theorem, we get the supermartingale property of $(\frac{V_t(\nu)}{\widehat{V}_t})_{t=0}^{\infty}$,

$$\begin{aligned} E\left[\frac{V_{t+1}(\nu)}{\widehat{V}_{t+1}}\Big|\mathcal{F}_t\right] &= E\left[\int_{\mathscr{L}}\frac{V_{t+1}^\pi}{\widehat{V}_{t+1}}d\nu(\pi)\Big|\mathcal{F}_t\right] \\ &= \int_{\mathscr{L}} E\left[\frac{V_{t+1}^\pi}{\widehat{V}_{t+1}}\Big|\mathcal{F}_t\right]d\nu(\pi) \\ &\le \int_{\mathscr{L}}\frac{V_t^\pi}{\widehat{V}_t}d\nu(\pi) = \frac{V_t(\nu)}{\widehat{V}_t}. \end{aligned}$$

□

Here we establish the supermartingale property used in the previous proof.

**Lemma 3.13.** *Let $\widehat{\pi}$ be given by (18). Then for every $\pi : \Delta^d \to \bar{\Delta}^d$ and every $x \in \Delta^d$,*

$$\int_{\Delta^d}\frac{\langle\pi(x),\frac{y}{x}\rangle}{\langle\widehat{\pi}(x),\frac{y}{x}\rangle}\varrho(x,dy) \le 1.$$

*Proof.* We proceed as in the proof of Becherer (2001, proposition 4.3). Fix $\pi$ and $\alpha \in (0,1)$ and define $\pi^\alpha = \alpha\pi + (1-\alpha)\widehat{\pi}$. Then by the (long only) log-optimality of $\widehat{\pi}$ we have for every $x \in \Delta^d$

$$\begin{aligned} 0 &\le \int_{\Delta^d}\left(\log\left\langle\widehat{\pi}(x),\frac{y}{x}\right\rangle - \log\left\langle\pi^\alpha(x),\frac{y}{x}\right\rangle\right)\varrho(x,dy) = \int_{\Delta^d}\left(\int_{\left\langle\pi^\alpha(x),\frac{y}{x}\right\rangle}^{\left\langle\widehat{\pi}(x),\frac{y}{x}\right\rangle}\frac{1}{z}dz\right)\varrho(x,dy) \\ &\le \int_{\Delta^d}\frac{\langle\widehat{\pi}(x),\frac{y}{x}\rangle - \langle\pi^\alpha(x),\frac{y}{x}\rangle}{\langle\pi^\alpha(x),\frac{y}{x}\rangle}\varrho(x,dy) = \int_{\Delta^d}\frac{\langle\alpha(\widehat{\pi}(x)-\pi(x)),\frac{y}{x}\rangle}{\langle\pi^\alpha(x),\frac{y}{x}\rangle}\varrho(x,dy). \end{aligned}$$

Hence,

$$\int_{\Delta^d}\frac{\langle\pi(x),\frac{y}{x}\rangle}{\langle\pi^\alpha(x),\frac{y}{x}\rangle}\varrho(x,dy) \le \int_{\Delta^d}\frac{\langle\widehat{\pi}(x),\frac{y}{x}\rangle}{\langle\pi^\alpha(x),\frac{y}{x}\rangle}\varrho(x,dy) \le \int_{\Delta^d}\frac{\langle\widehat{\pi}(x),\frac{y}{x}\rangle}{\langle(1-\alpha)\widehat{\pi}(x),\frac{y}{x}\rangle}\varrho(x,dy),$$

where the last equality follows from $\pi^\alpha \geq (1-\alpha)\widehat{\pi}$. By Fatou's lemma, we therefore have

$$\int_{\Delta^d} \frac{\langle \pi(x), \frac{y}{x}\rangle}{\langle \widehat{\pi}(x), \frac{y}{x}\rangle} \varrho(x, dy) = \int_{\Delta^d} \lim_{\alpha\to 0} \frac{\langle \pi(x), \frac{y}{x}\rangle}{\langle \pi^\alpha(x), \frac{y}{x}\rangle} \varrho(x, dy) \leq \lim_{\alpha\to 0} \int_{\Delta^d} \frac{\langle \pi(x), \frac{y}{x}\rangle}{\langle \pi^\alpha(x), \frac{y}{x}\rangle} \varrho(x, dy)$$

$$\leq \lim_{\alpha\to 0} \frac{1}{1-\alpha} \int_{\Delta^d} \frac{\langle \widehat{\pi}(x), \frac{y}{x}\rangle}{\langle \widehat{\pi}(x), \frac{y}{x}\rangle} \varrho(x, dy) = 1.$$

□

# 4 | THE CONTINUOUS TIME CASE WITH FUNCTIONALLY GENERATED PORTFOLIOS

This section is dedicated to a similar analysis in continuous time and with *functionally generated* portfolio maps (Fernholz, 2002, chapter 3). Using the pathwise Itô calculus developed by Föllmer (1981), we can define the corresponding wealth processes in a pathwise manner for any continuous market path admitting a quadratic variation process. This allows us to define the best retrospectively chosen portfolio which is not well-defined in general (and in particular for the Lipschitz portfolio maps).

## 4.1 | Functionally generated portfolios

We consider the following set of concave functions. For some fixed $M > 0$ and $0 \leq \alpha \leq 1$, we define

$$\mathcal{G}^{M,\alpha} = \left\{ G \in C^{2,\alpha}(\bar{\Delta}^d), \text{concave such that } \|G\|_{C^{2,\alpha}} \leq M \text{ and } G \geq \frac{1}{M} \right\},$$

where $C^{2,\alpha}(\bar{\Delta}^d)$ denotes the Hölder space of 2-times continuously differentiable functions from $\bar{\Delta}^d \to \mathbb{R}$ whose derivatives are $\alpha$-Hölder continuous. That is,

$$C^{2,\alpha}(\bar{\Delta}^d) = \{G \in C^2(\bar{\Delta}^d) \mid \|G\|_{C^{2,\alpha}} < \infty\},$$

where

$$\|G\|_{C^{2,\alpha}} = \max_{|\mathbf{k}|\leq 2} \|D^{\mathbf{k}} G\|_\infty + \max_{|\mathbf{k}|=2} \sup_{x\neq y} \frac{|D^{\mathbf{k}} G(x) - D^{\mathbf{k}} G(y)|}{\|x-y\|^\alpha}$$

with $\mathbf{k}$ denoting a multi-index in $\mathbb{N}^2$. For $\alpha = 0$, the second term in this norm is left away. Note that $G$ is only defined on the simplex $\Delta^d$. In order that the partial derivatives are well-defined, we assume that each $G$ is extended to an open neighborhood of $\Delta^d$ such that $G(x) = G(x')$, where $x'$ is the orthogonal projection of $x$ onto $\Delta^d$. The choice of the extension is irrelevant.

Here is an analytical lemma whose proof is given in the appendix.

**Lemma 4.1.** *For any $M, \alpha > 0$, the set $\mathcal{G}^{M,\alpha}$ is compact with respect to $\|\cdot\|_{C^{2,0}}$.*

To the set of generating functions $\mathcal{G}^{M,\alpha}$, we associate now the set of functionally generated portfolios $\mathcal{FG}^{M,\alpha}$ in the spirit of Fernholz (2002) defined by

$$\mathcal{FG}^{M,\alpha} = \left\{ \pi^G : \Delta^d \to \bar{\Delta}^d,\right.$$

$$x \mapsto (\pi^G(x))^i = x^i \left( \frac{D^i G(x)}{G(x)} + 1 - \sum_{j=1}^{d} \frac{D^j G(x)}{G(x)} x^j \right),\ i = 1, \dots d,\ | \ G \in \mathcal{G}^{M,\alpha} \Bigg\}. \quad (40)$$

By the concavity of $G$, $\pi^G$ takes values in $\bar{\Delta}^d$, that is, it is long-only (see, e.g., Fernholz & Karatzas, 2009, remark 11.1). The corresponding wealth processes are denoted by $V^{\pi^G}$ or $V^G$.

For these portfolios, it is possible to obtain a pathwise expression for $V^{\pi^G}$. We refer the reader to Schied et al. (2016) for extensions of this pathwise approach to time-dependent and path-dependent generating functions. There, this is achieved by applying the functional Itô calculus developed by Dupire (2009) and Cont and Fournié (2010, 2013), which generalizes Föllmer's Itô calculus to path-dependent functionals. In this paper, we only consider functionally generated portfolio maps as defined in (40).

We adopt the notation of Schied et al. (2016) and fix a refining sequence of partitions $(\mathbb{T}_n)_{n=1}^{\infty}$ of $[0, \infty)$, that is, $\mathbb{T}_n = \{t_0, t_1, \dots\}$ is such that $0 = t_0^n < t_1^n < \cdots$ and $t_k^n \to \infty$ as $k \to \infty$, and $\mathbb{T}_1 \subset \mathbb{T}_2 \subset \cdots$. Moreover, the mesh of $\mathbb{T}_n$ tends to zero on each compact interval as $n \to \infty$. Furthermore, we denote the successor of $t \in \mathbb{T}_n$ by $t'$. That is, $t' = \min\{u \in \mathbb{T}_n \,|,\ u > t\}$. Throughout this section, the market weights are described by a $d$-dimensional continuous path $\mu = (\mu_t)_{t \geq 0}$ with values in $\Delta^d$. Here and henceforth we let $S_d^+$ be the set of $d \times d$ positive definite matrices.

**Assumption 4.2.** The path $(\mu_t)_{t \geq 0}$ admits a continuous $S_d^+$-valued quadratic variation $[\mu]$ along $(\mathbb{T}_n)$ in the sense of Föllmer (1981), that is, for any $1 \leq i, j \leq d$ and all $t \geq 0$ the sequence

$$\sum_{s \in \mathbb{T}_n, s \leq t} \left(\mu_{s'}^i - \mu_s^i\right) \left(\mu_{s'}^j - \mu_s^j\right)$$

converges to a finite limit, as $n \to \infty$, denoted $[\mu^i, \mu^j]_t$, such that $t \mapsto [\mu^i, \mu^j]_t$ is continuous.

The dynamics of the relative wealth process $V^{\pi^G}$ built by investing according to $\pi^G \in \mathcal{FG}^{M,\alpha}$ are given in this continuous-time case by

$$\frac{dV_t^{\pi^G}}{V_t^{\pi^G}} = \sum_{i=1}^{d} \left(\pi^G(\mu_t)\right)^i \frac{d\mu_t^i}{\mu_t^i} = \sum_{i=1}^{d} \frac{D^i G(\mu_t)}{G(\mu_t)} d\mu_t^i, \quad V_0^{\pi} = 1, \quad (41)$$

(compare (11) in the discrete time case), where the right-hand side has to be understood as Föllmer's pathwise integral (cf. equation (6) in Schied et al., 2016). Note that the second equality holds by the definition of $\pi^G$ and the fact that $\sum_{i=1}^{d} d\mu_t^i = 0$.

Using (41) and Föllmer's Itô calculus, we have the following pathwise version of Fernholz's (2002) master equation (also see Schied et al., 2016, theorem 2.9).

**Corollary 4.3.** *Let $G \in C^2(\bar{\Delta}^d)$ and $\pi^G$ be defined as in (40). Let $(\mu_t)_{t \geq 0}$ be a continuous path satisfying Assumption 4.2. Then $V^{\pi^G}$ satisfies*

$$V_T^{\pi^G} \equiv V_T^G = \frac{G(\mu_T)}{G(\mu_0)} e^{\mathfrak{g}([0,T])}, \quad 0 \leq T < \infty, \quad (42)$$

*where* $\mathfrak{g}(dt) = -\frac{1}{2G(\mu_t)} \sum_{i,j} D^{ij} G(\mu_t) d[\mu^i, \mu^j]_t$.

## 4.2 | Definitions of the portfolios

We again consider (i) the best retrospectively chosen portfolio, (ii) the universal portfolio, and (iii) the log-optimal portfolio. To define the log-optimal portfolio, we will restrict to a specific stochastic model introduced in Section 4.2.3. In Section 4.2.4, we derive the asymptotic growth rate for this model class under an additional ergodicity assumption.

### 4.2.1 | The best retrospectively chosen portfolio

We consider the set of functionally generated portfolios $\mathcal{FG}^{M,\alpha}$ and a given continuous path $(\mu_t)_{t\geq 0}$ satisfying Assumption 4.2. For $M, \alpha > 0$ fixed, we define

$$V_T^{*,M,\alpha} = \sup_{\pi^G \in \mathcal{FG}^{M,\alpha}} V_T^{\pi^G} = \sup_{G \in \mathcal{G}^{M,\alpha}} V_T^G. \tag{43}$$

We first prove that an optimizer exists by establishing the following continuity property whose proof can be found in the appendix.

**Lemma 4.4.** *Let $T, M, \alpha > 0$ be fixed and $(\mu_t)_{t\geq 0}$ be a continuous path satisfying Assumption 4.2. Consider the function $G \mapsto V_T^G$ where $V_T^G$ is given by (42). Then $G \mapsto V_T^G$ is continuous from $(\mathcal{G}^{M,\alpha}, \|\cdot\|_{C^{2,0}})$ to $\mathbb{R}$.*

**Proposition 4.5.** *Let $T$ be fixed and $(\mu_t)_{t\geq 0}$ be a continuous path satisfying Assumption 4.2. Let $V_T^{*,M,\alpha}$ be defined by (43). Then there exists an optimizer $G_T^* \in \mathcal{G}^{M,\alpha}$ and in turn a portfolio $\pi_T^*$ generated by $G_T^*$ such that*

$$V_T^{*,M,\alpha} = V_T^{\pi_T^*} = V_T^{G_T^*}.$$

*Proof.* This is simply a consequence of continuity as proved in Lemma 4.4 and compactness of $(\mathcal{G}^{M,\alpha}, \|\cdot\|_{C^{2,0}})$ as shown in Lemma 4.1. □

### 4.2.2 | Universal portfolio

To define the analog of Cover's/Jamshidian's portfolio in the present setting, let $m$ be a Borel probability measure on $(\mathcal{G}^{M,\alpha}, \|\cdot\|_{C^{2,0}})$. Consider the map

$$F : \mathcal{G}^{M,\alpha} \to \mathcal{FG}^{M,\alpha}; G \mapsto F(G) = \pi^G, \tag{44}$$

where $\pi^G$ is given by (40). Define now on $(\mathcal{FG}^{M,\alpha}, \|\cdot\|_\infty)$ a Borel probability measure $\nu$ via the pushforward $\nu = F_* m$. As in Definition 2.4, we then define the corresponding universal portfolio via

$$\pi_T^\nu = \frac{\int_{\mathcal{FG}^{M,\alpha}} \pi^G(\mu_T) V_T^{\pi^G} d\nu(\pi^G)}{\int_{\mathcal{FG}^{M,\alpha}} V_T^{\pi^G} d\nu(\pi^G)}. \tag{45}$$

Analogous to (14), the value of the universal portfolio is given by

$$V_T^{M,\alpha}(\nu) := V_T^{\pi^\nu} = \int_{\mathcal{FG}^{M,\alpha}} V_T^{\pi^G} d\nu(\pi^G) = \int_{\mathcal{G}^{M,\alpha}} V_T^G dm(G). \tag{46}$$

*Remark* 4.6. More precisely, we need to verify that the universal portfolio still allows for pathwise integration and that the value of the portfolio (as a pathwise integral) is given by the right-hand side

of (46). These claims can be easily checked using the definitions and results in Schied et al. (2016), so we omit the details.

### 4.2.3 | Functionally generated log-optimal portfolios

By definition, the log-optimal portfolios requires a stochastic model for the market weights. We suppose that $\mu = (\mu_t^1, \dots, \mu_t^d)_{t\geq 0}$ follows a *time-homogeneous Markovian Itô diffusion*, defined on $(\Omega, \mathcal{F}, (\mathcal{F}_t)_{t\geq 0}, \mathbb{P})$ with values in $\Delta^d$, given by

$$\mu_t = \mu_0 + \int_0^t c(\mu_s)\lambda(\mu_s)ds + \int_0^t \sqrt{c(\mu_s)}dW_s, \quad \mu_0 \in \Delta^d, \tag{47}$$

where $\sqrt{\cdot}$ denotes the matrix square root, $W$ is a $d$-dimensional Brownian motion, $\lambda$ is a Borel measurable function from $\Delta^d \to \mathbb{R}^d$, and $c$ is a Borel measurable function from $\Delta^d \to S_+^d$, satisfying

$$\int_0^T \lambda^\top(\mu_t)c(\mu_t)\lambda(\mu_t)dt < \infty, \quad \forall T \in [0, \infty), \tag{48}$$

$$c(x)\mathbf{1} = 0, \qquad \sum_{i,j} c^{ij}(x)\lambda(x)^j = 0, \quad \forall x \in \Delta^d. \tag{49}$$

The requirements in (49) are necessary to guarantee that the process $\mu$ lies in $\Delta^d$. Note that $(\mu_t)_{t\geq 0}$ given by (47) satisfies the so-called *structure condition* (see Schweizer, 1995) (because of (48) and the fact that the drift part is of form $\int_0^t c(\mu_s)\lambda(\mu_s)ds$). This structural condition characterizes the condition of "no unbounded profit with bounded risk" (NUPBR) in the case of continuous semimartingales (see, e.g., Hulley & Schweizer, 2010).

In this setting, the proportions of current (relative) wealth invested in each of the assets are described by processes $\pi$ in the following set:

$$\Pi = \{\pi \mid H^d\text{-valued, predictable, } R\text{-integrable}\}, \tag{50}$$

where the process $R$ is defined componentwise by $R_t^i = \int_0^t \frac{d\mu_s^i}{\mu_s^i}$. Here, $H^d$ denotes the hyperplane corresponding to portfolio weights that are *not* necessarily long-only, that is, $H^d = \{x \in \mathbb{R}^d \mid \sum_{j=1}^d x^j = 1\}$. Note that the set $\mathcal{FG}^{M,\alpha}$ is clearly a subset of long-only strategies in $\Pi$. The relative wealth process $V^\pi$ satisfies

$$\frac{dV_t^\pi}{V_t^\pi} = \sum_{i=1}^d \pi_t^i \frac{d\mu_t^i}{\mu_t^i}, \quad V_0^\pi = 1. \tag{51}$$

In contrast to Section 4.1, this is a usual stochastic integral because we are dealing with general integrands $\pi$. Note that we can also write

$$V_T^\pi = \mathcal{E}((\pi \bullet R))_T = \exp\left(\int_0^T \left(\frac{\pi}{\mu_t}\right)^\top d\mu_t - \frac{1}{2}\int_0^T \left(\frac{\pi}{\mu_t}\right)^\top c(\mu_t)\frac{\pi}{\mu_t}dt\right) \tag{52}$$

$$= \exp\left(\int_0^T \sum_{i=1}^d \pi^i \frac{d\mu_t^i}{\mu_t^i} - \frac{1}{2}\int_0^T \sum_{i,j} \frac{\pi^i}{\mu_t^i}\frac{\pi^j}{\mu_t^j}c^{ij}(\mu_t)dt\right),$$

where, for two vectors $p, q \in \mathbb{R}^d$, $p/q$ always denotes the componentwise quotient $(\frac{p^1}{q^1}, \dots, \frac{p^d}{q^d})$.

Next we consider the log-optimal portfolio defined by (15) (but in continuous time now). As in Fernholz and Karatzas (2010b, section 3.1), we derive the ratio of two wealth processes $V^\pi$ and $V^\theta$ for $\pi, \theta \in \Pi$. Using (51) (for the processes $\pi$ and $\theta$) and Itô's lemma, this ratio is given by

$$\begin{aligned} d\left(\frac{V_t^\pi}{V_t^\theta}\right) &= \frac{V_t^\pi}{V_t^\theta}\left(\frac{\pi_t}{\mu_t} - \frac{\theta_t}{\mu_t}\right)^\top \left(d\mu_t - c(\mu_t)\frac{\theta_t}{\mu_t}dt\right) \\ &= \frac{V_t^\pi}{V_t^\theta}\left(\frac{\pi_t}{\mu_t} - \frac{\theta_t}{\mu_t}\right)^\top \left(\sqrt{c(\mu_t)}dW_t + c(\mu_t)\left(\lambda(\mu_t) - \frac{\theta_t}{\mu_t}\right)dt\right). \end{aligned}$$

The finite variation part of the expression vanishes *for every* $\pi \in \Pi$ if we choose $\theta \in \Pi$ such that

$$c(\mu_t)\left(\frac{\theta_t}{\mu_t} - \lambda(\mu_t)\right) = 0, \quad \mathbb{P}\text{-a.s. for all } t \geq 0. \tag{53}$$

By passing from the scaled relative weights $\theta/\mu$ to ordinary portfolio weights via Fernholz and Karatzas (2010b, equation (5)), the generic solution of (53), which we denote by $\widehat{\pi}$,[8] is given by

$$\widehat{\pi}_t^i = \mu_t^i\left(\lambda^i(\mu_t) + 1 - \sum_{j=1}^d \mu_t^j \lambda^j(\mu_t)\right). \tag{54}$$

Let $\widehat{V}$ be the associated wealth process. From (53), the ratio $V_t^\pi/\widehat{V}_t$ is, for any $\pi \in \Pi$, a nonnegative local martingale and therefore a supermartingale. Hence $\widehat{V}_t$ yields the relative wealth process corresponding to the log-optimal portfolio (see, e.g., Fernholz & Karatzas, 2010b; Karatzas & Kardaras, 2007). Indeed, by the supermartingale property and Jensen's inequality

$$\mathbb{E}\left[\log\left(V_T^\pi\right) - \log\left(\widehat{V}_T\right)\right] = \mathbb{E}\left[\log\left(\frac{V_T^\pi}{\widehat{V}_T}\right)\right] \leq \log\left(\mathbb{E}\left[\frac{V_T^\pi}{\widehat{V}_T}\right]\right) \leq 0.$$

Thus $\mathbb{E}[\log(V_T^\pi)] \leq \mathbb{E}[\log(\widehat{V}_T)]$ for all $\pi \in \Pi$.

By (52), the expected value of the log-optimal portfolio is given by

$$\sup_{\pi\in\Pi} \mathbb{E}\left[\log V_T^\pi\right] = \frac{1}{2}\mathbb{E}\left[\int_0^T \lambda^\top(\mu_t)c(\mu_t)\lambda(\mu_t)dt\right].$$

So far we have optimized over *all* strategies in $\Pi$. In the sequel, we shall mainly consider suprema taken over smaller sets, in particular over $\mathcal{FG}^{M,\alpha}$. Note that in this case the optimizer will still be a function of the market weights due to the Markov property of $(\mu_t)_{t\geq 0}$.

In this context, let us also answer the question of when the log-optimal portfolio is functionally generated. This is needed to relate its asymptotic growth rate to the one of the best retrospectively chosen portfolio and the universal portfolio.

**Proposition 4.7.** *Let $(\mu_t)_{t\geq 0}$ be of the form (47). Then the log-optimal portfolio is generated by a differentiable function $G$, that is,*

$$\widehat{\pi}_t^i = \mu_t^i\left(\frac{D^iG(\mu_t)}{G(\mu_t)} + 1 - \sum_{j=1}^d \mu_t^j \frac{D^jG(\mu_t)}{G(\mu_t)}\right), \quad i = 1, \ldots, d,$$

*if the drift characteristic $\lambda$ satisfies*

$$\lambda(x) = \nabla \log G(x) = \frac{\nabla G(x)}{G(x)}, \quad x \in \Delta^d.$$

*Proof.* The assertion follows from expression (54). □

### 4.2.4 | Asymptotic growth rates for an ergodic market weights process

**Assumption 4.8.** The process $\mu$ as given in (47) is an ergodic process with stationary measure $\varrho$ on $\Delta^d$.

With this assumption we derive an expression of the asymptotic growth rate $\lim_{T\to\infty} \frac{1}{T} \log V_T^\pi$. For the precise notion of ergodicity in continuous time, we refer to Eberle (2016, section 2.2., theorem 2.4, and section 2.2.3). Assumption 4.8 is essentially satisfied under a mean reversion condition. Examples include polynomial models for the market weights staying in the interior of the simplex (see Cuchiero, 2019, theorem 5.1) with the subclass of volatility stabilized models (Fernholz & Karatzas, 2005).

In the following theorem, we consider portfolio maps which are not necessarily long-only, but can take values in the hyperplane $H^d$.

**Theorem 4.9.** *Under Assumption 4.8 the following statements hold true:*

**(i)** *Let $\pi : \Delta^d \to H^d$ be any ($\varrho$-measurable) portfolio map such that*

$$\int_{\Delta^d} \left| \left( \frac{\pi(x)}{x} \right)^\top c(x)\lambda(x) \right| \varrho(dx) < \infty,$$

$$Q^\pi := \int_{\Delta^d} \left( \frac{\pi(x)}{x} \right)^\top c(x) \left( \frac{\pi(x)}{x} \right) \varrho(dx) < \infty. \tag{55}$$

*We then have, for $\varrho$-a.e. starting value $\mu_0$, that*

$$\lim_{T\to\infty} \frac{1}{T} \log(V_T^\pi) = L^\pi := \int_{\Delta^d} \left( \frac{\pi(x)}{x} \right)^\top c(x)\lambda(x)\varrho(dx) - \frac{1}{2} \int_{\Delta^d} \left( \frac{\pi(x)}{x} \right)^\top c(x) \left( \frac{\pi(x)}{x} \right) \varrho(dx), \quad \mathbb{P}\text{-a.s.}$$

**(ii)** *Assume that $L := \frac{1}{2} \int_{\Delta^d} \lambda^\top(x)c(x)\lambda(x)\varrho(dx) < \infty$. Then, for $\varrho$-a.e. starting value $\mu_0$, it holds that*

$$\lim_{T\to\infty} \frac{1}{T} \log \widehat{V}_T = L, \quad \mathbb{P}\text{-a.s.}$$

The proof of Theorem 4.9 relies on the following lemma which is stated and proved in Fernholz (2002, lemma 1.3.2).

**Lemma 4.10.** *Let $M$ be a continuous local martingale such that*

$$\lim_{T\to\infty} \frac{1}{T^2} \langle M, M \rangle_T \log\log T = 0, \quad \mathbb{P}\text{-a.s.} \tag{56}$$

*Then $\lim_{T\to\infty} \frac{1}{T} M_T = 0$, $\mathbb{P}$-a.s.*

*Proof of Theorem* 4.9. Let us start by proving statement (i). By (51), $\log V_T^\pi$ reads as

$$\log V_T^\pi = \int_0^T \left(\frac{\pi(\mu_t)}{\mu_t}\right)^\top c(\mu_t)\lambda(\mu_t)dt - \frac{1}{2}\int_0^T \left(\frac{\pi(\mu_t)}{\mu_t}\right)^\top c(\mu_t)\frac{\pi(\mu_t)}{\mu_t}dt + \int_0^T \left(\frac{\pi(\mu_t)}{\mu_t}\right)^\top \sqrt{c(\mu_t)}dW_t. \tag{57}$$

The local martingale part

$$M_T^\pi := \int_0^T \left(\frac{\pi(\mu_t)}{\mu_t}\right)^\top \sqrt{c(\mu_t)}dW_t$$

satisfies Condition (56) of Lemma 4.10. Indeed, by the ergodic theorem in continuous time (see, e.g., Eberle, 2016, theorem 2.4 and section 2.2.3) and (55) we have

$$\frac{1}{T}\langle M^\pi, M^\pi\rangle_T = \frac{1}{T}\int_0^T \left(\frac{\pi(\mu_t)}{\mu_t}\right)^\top c(\mu_t)\frac{\pi(\mu_t)}{\mu_t}dt \overset{T\to\infty}{\to} Q^\pi < \infty, \quad \mathbb{P}\text{-a.s.}$$

Multiplying the left-hand side with $(\log\log T)/T$, therefore yields Condition (56) and

$$\frac{1}{T}M_T^\pi = \frac{1}{T}\int_0^T \left(\frac{\pi(\mu_t)}{\mu_t}\right)^\top \sqrt{c(\mu_t)}dW_t \to 0, \quad \mathbb{P}\text{-a.s.}$$

Hence, evoking again the ergodic theorem yields

$$\begin{aligned}\lim_{T\to\infty}\frac{1}{T}\log V_T^\pi &= \lim_{T\to\infty}\frac{1}{T}\left(\int_0^T \left(\frac{\pi(\mu_t)}{\mu_t}\right)^\top c(\mu_t)\lambda(\mu_t)dt - \frac{1}{2}\int_0^T \left(\frac{\pi(\mu_t)}{\mu_t}\right)^\top c(\mu_t)\frac{\pi(\mu_t)}{\mu_t}dt\right)\\ &= \int_{\Delta^d}\left(\frac{\pi(x)}{x}\right)^\top c(x)\lambda(x)\varrho(dx) - \frac{1}{2}\int_{\Delta^d}\left(\frac{\pi(x)}{x}\right)^\top c(x)\left(\frac{\pi(x)}{x}\right)\varrho(dx),\end{aligned}$$

$\mathbb{P}$-a.s. (and also in $L^1(\Omega, \mathcal{F}, P)$) and thus assertion (i).

Concerning statement (ii), note from (53) that the scaled relative weights corresponding to the log-optimal portfolio satisfy

$$c(x)\left(\frac{\widehat{\pi}(x)}{x} - \lambda(x)\right) = 0.$$

Thus, by (57) and (51), $\log \widehat{V}_T$ simplifies to

$$\log \widehat{V}_T = \frac{1}{2}\int_0^T \lambda^\top(\mu_t)c(\mu_t)\lambda(\mu_t)dt + \int_0^T \lambda^\top(\mu_t)\sqrt{c(\mu_t)}dW_t.$$

In this case, we have

$$\frac{1}{T}\langle M^{\widehat{\pi}}, M^{\widehat{\pi}}\rangle_T = \frac{1}{T}\int_0^T \lambda^\top(\mu_t)c(\mu_t)\lambda(\mu_t)dt \overset{T\to\infty}{\to} 2L, \quad \mathbb{P}\text{-a.s.},$$

which yields by the same argument as above

$$\frac{1}{T} M_T^{\hat{\pi}} = \frac{1}{T} \int_0^T \lambda^\top(\mu_t)\sqrt{c(\mu_t)} dW_t \to 0, \quad \mathbb{P}\text{-a.s.}$$

and in turn

$$\lim_{T\to\infty} \frac{1}{T} \log \widehat{V}_T = \lim_{T\to\infty} \frac{1}{2T} \int_0^T \lambda^\top(\mu_t) c(\mu_t) \lambda(\mu_t) dt = L, \quad \mathbb{P}\text{-a.s.}$$

□

## 4.3 | Asymptotically equivalent growth rates

As in discrete time, we will establish asymptotic equality of the growth rates of all three portfolio types introduced in Section 4.2. First, we compare the best retrospectively chosen portfolio with the universal one. For an analogous result in the context of optimal arbitrage, see theorem 4.5 of Kardaras and Robertson (2012).

**Theorem 4.11.** *Let $M, \alpha > 0$ be fixed and let $(\mu_t)_{t\geq 0}$ be a continuous path satisfying Assumption 4.2 such that for all $i \in \{1, \dots, d\}$*

$$\lim_{T\to\infty} \frac{1}{T} [\mu^i, \mu^i]_T < \infty. \tag{58}$$

*Consider a probability measure $m$ on $\mathcal{G}^{M,\alpha}$ with full support and set $\nu = F_* m$ with $F$ defined in (44). Then*

$$\lim_{T\to\infty} \frac{1}{T} \left( \log V_T^{*,M,\alpha} - \log V_T^{M,\alpha}(\nu) \right) = 0,$$

*where $V^{*,M,\alpha}$ and $V^{M,\alpha}(\nu)$ are defined in (43) and (46), respectively.*

*Proof.* The inequality "$\geq$" is obvious. For the converse inequality, we proceed similarly as in the previous section (using only generating functions). As $m$ has full support and $\mathcal{G}^{M,\alpha}$ is compact, we have that, for $\eta > 0$ there exists some $\delta > 0$, such that every $\eta$-neighborhood of a point $G \in \mathcal{G}^{M,\alpha}$ has $m$-measure bigger than $\delta$.

Let $T \geq 1$ and denote by $G_T^*$ the optimizer as of Proposition 4.5. Consider now a generating function $G$ such $\|G - G_T^*\|_{C^{2,0}} \leq \eta$. Then it follows from (A9) that

$$\begin{aligned} \frac{1}{T} \left( \log \left( V_T^G \right) - \log \left( V_T^{G_T^*} \right) \right) &\geq \frac{1}{T} \left( -2M\eta - \left( \frac{M}{2} d^2 \eta + \frac{M^3}{2} d^2 \eta \right) \max_i [\mu^i, \mu^i]_T \right) \\ &=: -K_T. \end{aligned} \tag{59}$$

Fix $\epsilon > 0$ and note that by assumption (58) and continuity of $T \mapsto \frac{1}{T}[u^i, u^i]_T$ on $[1, \infty)$, $\sup_{T\in[1,\infty)} \frac{1}{T}[\mu^i, \mu^i]_T$ can be bounded by some constant. Therefore, we can choose $\eta$ sufficiently small such that $K_T \leq \epsilon$ for all $T \geq 1$.

Denote by $B = B_\eta(G_T^*)$ the $\|\cdot\|_{C^{2,0}}$-ball with radius $\eta$ in $\mathcal{G}^{M,\alpha}$ which has $m$-measure at least $\delta > 0$, where $\delta$ only depends on $\eta$. We then may estimate using Jensen's inequality and (59)

$$\left(\frac{V_T^{M,\alpha}(\nu)}{V_T^{G_T^*}}\right)^{\frac{1}{T}} = \left(\frac{\int_{\mathcal{G}^{M,\alpha}} V_T^G m(dG)}{V_T^{G_T^*}}\right)^{\frac{1}{T}} \geq \left(\frac{\int_{B_\eta(G_T^*)} V_T^G m(dG)}{V_T^{G_T^*}}\right)^{\frac{1}{T}}$$

$$\geq \delta^{\frac{1}{T}-1} \frac{\int_{B_\eta(G_T^*)} (V_T^G)^{\frac{1}{T}} m(dG)}{(V_T^{G_T^*})^{\frac{1}{T}}} \geq \delta^{\frac{1}{T}} e^{-K_T} \geq \delta^{\frac{1}{T}} e^{-\epsilon}.$$

Letting $T \to \infty$ for any given $\epsilon$ (which determines $\eta$ and in turn $\delta$) yields the assertion. □

To compare the asymptotic performance with that of the log-optimal portfolio, we optimize over portfolio maps in $\mathcal{FG}^{M,\alpha}$ and suppose henceforth that $(\mu_t)_{t\geq 0}$ is of the form (47). Under Assumption 4.8 and from Theorem 4.9 define

$$\widehat{\pi}^{M,\alpha} := \underset{\pi^G \in \mathcal{FG}^{M,\alpha}}{\arg\max} \left( \int_{\Delta^d} \left(\frac{\pi^G(x)}{x}\right)^\top c(x)\lambda(x)\varrho(dx) \right. \tag{60}$$

$$\left. -\frac{1}{2}\int_{\Delta^d} \left(\frac{\pi^G(x)}{x}\right)^\top c(x) \left(\frac{\pi^G(x)}{x}\right) \varrho(dx) \right)$$

and the corresponding wealth process $\widehat{V}^{M,\alpha}$ by $\widehat{V}^{M,\alpha} = V^{\widehat{\pi}^{M,\alpha}}$, whenever $\widehat{\pi}^{M,\alpha}$ is well-defined. As

$$\sup_{\pi^G \in \mathcal{FG}^{M,\alpha}} \mathbb{E}\left[\log\left(V_T^{\pi^G}\right)\right]$$

yields $\widehat{\pi}^{M,\alpha}$ as optimizer for all $T > 0$, $\widehat{V}^{M,\alpha}$ corresponds to the log-optimal portfolio among functionally generated portfolios with generating function in $\mathcal{G}^{M,\alpha}$.

**Theorem 4.12.** *Let $M, \alpha > 0$ be fixed and let $(\mu_t)_{t\geq 0}$ be a stochastic process of the form (47) satisfying Assumption 4.8. Moreover, suppose that*

$$\int_{\Delta^d} c^{ii}(x)\varrho(dx) < \infty, \quad \textit{for all } i \in \{1, \dots, d\}, \tag{61}$$

$$\int_{\Delta^d} \max_{i\in\{1,\dots,d\}} |(c(x)\lambda(x))^i| \varrho(dx) < \infty. \tag{62}$$

*Consider a probability measure $m$ on $\mathcal{G}^{M,\alpha}$ with full support and set $\nu = F_* m$ with $F$ defined in (44). Then*

$$\liminf_{T\to\infty} \frac{1}{T} \log V_T^{*,M,\alpha} = \liminf_{T\to\infty} \frac{1}{T} \log V_T^{M,\alpha}(\nu) = \lim_{T\to\infty} \frac{1}{T} \log \widehat{V}_T^{M,\alpha}, \quad \mathbb{P}\textit{-a.s.} \tag{63}$$

*where $\widehat{V}_T^{M,\alpha}$ denotes the log-optimal portfolio among $\mathcal{FG}^{M,\alpha}$-maps defined via (60), $V^{*,M,\alpha}$ and $V^{M,\alpha}(\nu)$ are defined pathwise in (43) and (46), respectively.*

*Proof.* We first note that $\widehat{\pi}^{M,\alpha}$ is well-defined. Indeed, the map

$$G \mapsto \int_{\Delta^d} \left(\frac{\pi^G(x)}{x}\right)^\top c(x)\lambda(x)\varrho(dx) - \frac{1}{2}\int_{\Delta^d} \left(\frac{\pi^G(x)}{x}\right)^\top c(x) \left(\frac{\pi^G(x)}{x}\right) \varrho(dx)$$
$$= \int_{\Delta^d} \left(\frac{\nabla G(x)}{G(x)}\right)^\top c(x)\lambda(x)\varrho(dx) - \frac{1}{2}\int_{\Delta^d} \left(\frac{\nabla G(x)}{G(x)}\right)^\top c(x) \left(\frac{\nabla G(x)}{G(x)}\right) \varrho(dx)$$

is continuous from $(\mathcal{G}^{M,\alpha}, \|\cdot\|_{2,0})$ to $\mathbb{R}$. This together with compactness of $\mathcal{G}^{M,\alpha}$ with respect to $\|\cdot\|_{2,0}$ imply the well-definedness of $\widehat{\pi}^{M,\alpha}$.

Note also that (61) and (62) as well as the conditions on $G$ imply the assumptions of the ergodic theorem (Theorem 4.9). Hence, we have for each $\pi^G \in \mathcal{FG}^{M,\alpha}$ the $\mathbb{P}$-a.s. limit

$$\lim_{T\to\infty} \frac{1}{T} \log V_T^{\pi^G} = L^{\pi^G}.$$

In particular,

$$\lim_{T\to\infty} \frac{1}{T} \log \widehat{V}_T^{M,\alpha} = \sup_{\pi^G \in \mathcal{FG}^{M,\alpha}} L^{\pi^G} =: L^{M,\alpha} \tag{64}$$

holds $\mathbb{P}$-a.s.

Due to (61), we can now apply Theorem 4.11 which implies the first equality in (63). Moreover, we have by the definition of $V_T^{*,M,\alpha}$ for each fixed $T$ the inequality

$$\frac{1}{T} \log\left(\widehat{V}_T^{M,\alpha}\right) \le \frac{1}{T} \log\left(V_T^{*,M,\alpha}\right), \quad \mathbb{P}\text{-a.s.} \tag{65}$$

Using (64), (65), and Theorem 4.11, we thus have $\mathbb{P}$-a.s.,

$$L^{M,\alpha} = \lim_{T\to\infty} \frac{1}{T} \log\left(\widehat{V}_T^{M,\alpha}\right) \le \liminf_{T\to\infty} \frac{1}{T} \log\left(V_T^{*,M,\alpha}\right) = \liminf_{T\to\infty} \frac{1}{T} \log\left(V_T^{M,\alpha}(\nu)\right). \tag{66}$$

On the other hand, by the definition of $(\widehat{V}_t^{M,\alpha})_{t\ge 0}$ as log-optimizer within the class $\mathcal{FG}^{M,\alpha}$

$$\mathbb{E}\left[\log\left(V_T^{M,\alpha}(\nu)\right)\right] \le \sup_{\pi^G \in \mathcal{FG}^{M,\alpha}} \mathbb{E}\left[\log\left(V_T^{\pi^G}\right)\right] = \mathbb{E}\left[\log\left(\widehat{V}_T^{M,\alpha}\right)\right] \tag{67}$$

holds. Concerning the first inequality, note that the universal portfolio to build the wealth $V_T^{M,\alpha}(\nu)$ is given by (45). By the time-homogenous Markovianity it is thus sufficient to dominate the left-hand side of (67) by taking the supremum over elements in $\mathcal{FG}^{M,\alpha}$.

Combining now (67), Theorem 4.9 and (66) yields,

$$\begin{aligned}
\mathbb{E}\left[\liminf_{T\to\infty} \frac{1}{T} \log\left(V_T^{M,\alpha}(\nu)\right)\right] &\le \liminf_{T\to\infty} \frac{1}{T} \mathbb{E}\left[\log\left(V_T^{M,\alpha}(\nu)\right)\right] \\
&\le \lim_{T\to\infty} \frac{1}{T} \mathbb{E}\left[\log\left(\widehat{V}_T^{M,\alpha}\right)\right] \\
&= \lim_{T\to\infty} \frac{1}{T} \log\left(\widehat{V}_T^{M,\alpha}\right) \\
&\le \liminf_{T\to\infty} \frac{1}{T} \log\left(V_T^{M,\alpha}(\nu)\right), \quad \mathbb{P}\text{-a.s.},
\end{aligned}$$

where the first inequality follows from Fatou's lemma. From this, we see that

$$\liminf_{T\to\infty} \frac{1}{T} \log\left(V_T^{M,\alpha}(\nu)\right)$$

is $\mathbb{P}$-a.s. constant and equal to $\lim_{T\to\infty} \frac{1}{T}\log(\widehat{V}_T^{M,\alpha})$. Hence the assertion is proved. □

As in the previous section, we can formulate a result not depending explicitly on the constant $M$ on $\alpha$. Setting $\alpha = \frac{1}{M}$ we choose for $M = 1, 2, 3, \dots$ a measure $m^M$ on $\mathcal{G}^{M,\frac{1}{M}}$ with full support. Define $m = \sum_{M=1}^{\infty} 2^{-M} m^M$ and the process $V(\nu)$ by

$$V_T(\nu) = \int_{\bigcup_{M=1}^{\infty} \mathcal{G}^{M,\frac{1}{M}}} V_T^G m(dG).$$

In order to compare the performance with the one of the global log-optimal portfolio, whenever it is functionally generated, we combine the above results with Proposition 4.7.

**Corollary 4.13.** *Let $(\mu_t)_{t\geq 0}$ be a stochastic process of form (47) satisfying Assumption 4.8. Moreover, suppose that $\lambda$ and $c$ satisfy (61) and*

$$\lambda(x) = \frac{\nabla \widehat{G}(x)}{\widehat{G}(x)}, \tag{68}$$

$$L = \frac{1}{2}\int_{\Delta^d} \frac{\nabla \widehat{G}(x)}{\widehat{G}(x)} c(x) \frac{\nabla \widehat{G}(x)}{\widehat{G}(x)} \rho(dx) < \infty \tag{69}$$

*for some concave function $\widehat{G} \in C^2(\bar{\Delta}^d)$. Then we have $\mathbb{P}$-a.s.*

$$\lim_{M\to\infty}\lim_{T\to\infty} \frac{1}{T}\log\left(V_T^{*,M,\frac{1}{M}}\right) = \lim_{T\to\infty}\frac{1}{T}\log\left(V_T(\nu)\right) = \lim_{T\to\infty}\frac{1}{T}\log(\widehat{V}_T) = L. \tag{70}$$

*Proof.* Note first that $L$ is well-defined due to (69). Furthermore, note that for every $\varepsilon > 0$, there exists some $M > 0$ and some function $G \in \mathcal{G}^{M,\frac{1}{M}}$ such that

$$\lim_{T\to\infty}\frac{1}{T}\log\left(V_T^G\right) \geq \lim_{T\to\infty}\frac{1}{T}\log(\widehat{V}_T) + \varepsilon.$$

Indeed this simply follows from continuity of $G \mapsto V^G$ as asserted in Lemma 4.4 and by choosing $G \in \mathcal{G}^{M,\frac{1}{M}}$ close enough with respect to the $\|\cdot\|_{C^{2,0}}$ to the optimizing function $\widehat{G} \in C^2(\bar{\Delta}^d)$ whose generated portfolio yields $\widehat{V}$ due to (68) and Proposition 4.7. By Theorem 4.12, we can therefore conclude (following the proof of Corollary 3.11) that

$$\lim_{M\to\infty}\liminf_{T\to\infty}\frac{1}{T}\log\left(V_T^{*,M,\frac{1}{M}}\right) = \liminf_{T\to\infty}\frac{1}{T}\log(V_T(\nu)) = \lim_{T\to\infty}\frac{1}{T}\log(\widehat{V}_T) = L \tag{71}$$

holds true. As Theorem 4.11 implies that

$$\limsup_{T\to\infty}\frac{1}{T}\log V_T(\nu) = \lim_{M\to\infty}\limsup_{T\to\infty}\frac{1}{T}\log V_T^{*,M,\frac{1}{M}},$$

the assertion is proved if

$$\limsup_{T\to\infty} \frac{1}{T}\left(\log V_T(\nu) - \log \widehat{V}_T\right) = \limsup_{T\to\infty} \frac{1}{T}\log\left(\frac{V_T(\nu)}{\widehat{V}_T}\right) = 0, \quad \mathbb{P}\text{-a.s.} \tag{72}$$

By the considerations of Section 4.2.3 (see also Becherer, 2001, proposition 4.3), it follows that $(\frac{V_t(\nu)}{\widehat{V}_t})_{t\geq 0}$ is a nonnegative supermartingale. It converges $\mathbb{P}$-a.s. to a finite limit as $t \to \infty$. This in turn implies (72) and proves the statement. □

Finally, a similar result can be obtained by restricting the log-optimal portfolio to the class of $C^2$-functionally generated portfolios without imposing the drift condition in Proposition 4.7. We denote by $V_T^{fun}$ the wealth process of the log-optimal portfolio among concave $C^2$-functionally generated portfolios, that is, $\pi^{fun}$ is defined as in (60), however by taking the arg max over all concave $C^2$-functionally generated portfolios.

**Corollary 4.14.** *Let $(\mu_t)_{t\geq 0}$ be a stochastic process of form (47) satisfying Assumption 4.8. Moreover, suppose that (61) and (62) hold true. Then*

$$\lim_{M\to\infty} \liminf_{T\to\infty} \frac{1}{T}\log V_T^{*,M,\frac{1}{M}} = \liminf_{T\to\infty} \frac{1}{T}\log V_T(\nu) = \lim_{T\to\infty} \frac{1}{T}\log V_T^{fun}, \quad \mathbb{P}\text{-a.s.} \tag{73}$$

*Proof.* The proof is the same as the first part of Corollary 4.13 up to (71). Note that we cannot get rid of the lim inf because the supermartingale argument from the proof of Corollary 4.13 does not hold. □

## ACKNOWLEDGMENTS

The authors wish to thank the anonymous referees and the associated editor for their useful comments. Parts of this paper were written while Cuchiero was visiting ETH Zürich; she is grateful to the Forschungsinstitut für Mathematik.

## ENDNOTES

[1] Henceforth, we only use the terminology log-optimal portfolio.

[2] Here, the portfolio weight $\pi_t$ is chosen at time $t-1$ and is used over the time interval $[t-1, t]$.

[3] We are grateful to one of anonymous referees for pointing this out.

[4] Later we will use $V$ to denote instead the *relative* wealth of the portfolio.

[5] Here, it is assumed implicitly that the stocks do not pay dividends. This assumption is common in universal and SPT and allows us to focus on the main ideas.

[6] A comparison between the log-optimal portfolio and Cover's universal portfolio is studied in Ichiba et al. (2011).

[7] To simplify the notations, we will suppress $\omega$.

[8] By a slight abuse of notation, we here write $\widehat{\pi}$ although we do not restrict to long-only portfolios as in Section 2.3.

## REFERENCES

Banner, A. D., Fernholz, R., & Karatzas, I. (2005). Atlas models of equity markets. *Annals of Applied Probability*, *15*(4), 2296–2330.

Becherer, D. (2001). The numeraire portfolio for unbounded semimartingales. *Finance and Stochastics*, *5*(3), 327–341.

Blum, A., & Kalai, A. (1999). Universal portfolios with and without transaction costs. *Machine Learning*, *35*(3), 193–205.

Brod, M. R. (2014). *Generating the universal portfolio* (Master's thesis). Eidgenössische Technische Hochschule, Zürich.

Cont, R., & Fournie, D.-A. (2010). A functional extension of the Itô formula. *Comptes Rendus Mathematique*, *348*(1), 57–61.

Cont, R., & Fournié, D.-A. (2013). Functional Itô calculus and stochastic integral representation of martingales. *Annals of Probability*, *41*(1), 109–133.

Cover, T. M. (1991). Universal portfolios. *Mathematical Finance*, *1*(1), 1–29.

Cover, T. M., & Ordentlich, E. (1996). Universal portfolios with side information. *IEEE Transactions on Information Theory*, *42*(2), 348–363.

Cuchiero, C. (2019). Polynomial processes in stochastic portfolio theory. *Stochastic Processes and Their Applications*, *129*(5), 1829–1872. https://doi.org/10.1016/j.spa.2018.06.007

Delbaen, F., & Schachermayer, W. (1995). The no-arbitrage property under a change of numéraire. *Stochastics Stochastics Reports*, *53*(3–4), 213–226.

Dobrowolski, M. (2010). *Angewandte Funktionalanalysis: Funktionalanalysis, Sobolev-Räume und elliptische Differentialgleichungen*. Berlin, Heidelberg: Springer-Verlag.

Dupire, B. (2009). *Functional Itô calculus* (Bloomberg Portfolio Research Paper 2009-04).

Eberle, A. (2016). *Markov processes*. Lecture Notes, University Bonn.

Fernholz, D., & Karatzas, I. (2010a). On optimal arbitrage. *Annals of Applied Probability*, *20*(4), 1179–1204.

Fernholz, D., & Karatzas, I. (2010b). Probabilistic aspects of arbitrage. In C. Chiarella & A. Novikov (Eds.), *Contemporary Quantitative Finance* (pp. 1–17). Berlin: Springer.

Fernholz, E. R. (2002). *Stochastic Portfolio Theory*. New York: Springer.

Fernholz, R., & Karatzas, I. (2005). Relative arbitrage in volatility-stabilized markets. *Annals of Finance*, *1*(2), 149–177.

Fernholz, R., & Karatzas, I. (2009). Stochastic portfolio theory: An overview. In *Handbook of Numerical Analysis* (Vol. 15, pp. 89–167). Amsterdam: Elsevier.

Föllmer, H. (1981). Calcul d'Itô sans probabilités. *Séminaire de Probabilités de Strasbourg*, *15*, 143–150.

Hazan, E., & Kale, S. (2015). An online portfolio selection algorithm with regret logarithmic in price variation. *Mathematical Finance*, *25*(2), 288–310.

Hulley, H., & Schweizer, M. (2010). On minimal market models and minimal martingale measures. In C. Chiarella & A. Novikov (Eds.), *Contemporary Quantitative Finance* (pp. 35–51). Berlin, Heidelberg: Springer.

Ichiba, T., & Brod, M. (2014). *Some aspects of universal portfolios*. Presentation, Banff 2014.

Ichiba, T., Papathanakos, V., Banner, A., Karatzas, I., & Fernholz, R. (2011). Hybrid atlas models. *Annals of Applied Probability*, *21*(2), 609–644.

Jamshidian, F. (1992). Asymptotically optimal portfolios. *Mathematical Finance*, *2*(2), 131–150.

Karatzas, I., & Kardaras, C. (2007). The numéraire portfolio in semimartingale financial models. *Finance and Stochastics*, *11*(4), 447–493.

Kardaras, C., & Robertson, S. (2012). Robust maximization of asymptotic growth. *Annals of Applied Probability*, *22*(4), 1576–1610.

Kramkov, D., & Schachermayer, W. (1999). The asymptotic elasticity of utility functions and optimal investment in incomplete markets. *Annals of Applied Probability*, 904–950.

Li, B., & Hoi, S. C. (2014). Online portfolio selection: A survey. *ACM Computing Surveys*, *46*(3), 35:1–35:36.

Pal, S., & Wong, T.-K. L. (2016). The geometry of relative arbitrage. *Mathematics and Financial Economics*, *10*(3), 263–293.

Schied, A., Speiser, L., & Voloshchenko, I. (2016). Model-free portfolio theory and its functional master formula. Preprint arXiv:1606.03325.

Schweizer, M. (1995). On the minimal martingale measure and the Föllmer-Schweizer decomposition. *Stochastic Analysis and Applications*, *13*(5), 573–599.

Wong, T.-K. L. (2015). Universal portfolios in stochastic portfolio theory. Preprint arXiv:1510.02808.

**How to cite this article:** Cuchiero C, Schachermayer W, Wong T-K L. Cover's universal portfolio, stochastic portfolio theory, and the numéraire portfolio. *Mathematical Finance*. 2019;29:773–803. https://doi.org/10.1111/mafi.12201

## APPENDIX: PROOFS OF CERTAIN RESULTS AND LEMMAS

*Proof of Theorem* 2.1. Fix $T > 0$ and the trajectory $s = (s_t^1, \dots, s_t^d)_{t=0}^T \in (\mathbb{R}^d)^{T+1}$. For fixed $s$ the function $b \mapsto V_T(b)(s)$ is continuous on $\bar{\Delta}^d$. Hence there is $\bar{b} = \bar{b}(s) \in \bar{\Delta}^d$ such that

$$V_T^*(s) = V_T(\bar{b})(s). \tag{A1}$$

In fact, condition (9) implies that the sequence of functions $(b \mapsto \frac{1}{T} \log V_T(b))_{T=1}^\infty$ is Lipschitz on $\bar{\Delta}^d$, uniformly in $T \in \mathbb{N}$ and $s$ satisfying (9) for some fixed constants $C > c > 0$.

Indeed, consider the distance on $\bar{\Delta}^d$ defined by $\|b - \tilde{b}\|_1 = \sum_{j=1}^d |b^j - \tilde{b}^j|$. Then we may estimate

$$\left| \frac{1}{T} \log V_T(b) - \frac{1}{T} \log V_T(\tilde{b}) \right| \leq (\log(C) - \log(c)) \|b - \tilde{b}\|_1.$$

For $\epsilon > 0$, we may therefore define $\delta := \frac{c\epsilon}{C} > 0$ such that, for every $\delta$-neighborhood $U(\bar{b})$ around any $\bar{b} \in \bar{\Delta}^d$ we have

$$\frac{1}{T} \log V_T(b) \geq \frac{1}{T} \log V_T(\bar{b}) - \epsilon$$

for every $b \in U(\bar{b})$. If the probability measure $\nu$ has full support, we also may find $\eta = \eta(\epsilon, c, C) > 0$ such that each such $\delta$-neighborhood $U(\bar{b})$, where $\bar{b}$ runs through $\bar{\Delta}^d$, satisfies $\nu(U(\bar{b})) > \eta$. Using (A1) we therefore may conclude, similarly as in (A8), that (8) holds true, uniformly in $s = (s_t^1, \dots, s_t^d)_{t=0}^\infty$ satisfying (9) for some fixed constants $C > c > 0$. □

*Proof of Lemma* 3.8. Let $\hat{\pi} : \Delta^d \to \bar{\Delta}^d$ be the optimizer of (18) and define, for $0 < \epsilon < 1$,

$$\pi_\epsilon = (1 - \epsilon)\hat{\pi} + \epsilon \left( \frac{1}{d}, \dots, \frac{1}{d} \right).$$

Note that $\pi_\epsilon$ takes values in $\bar{\Delta}_\epsilon^d$ (see Definition 3.1), which is crucial for the subsequent arguments and the reason why we do not directly work with $\hat{\pi}$. Also note that, for $p \in \bar{\Delta}_\epsilon^d$, we have

$$\left\langle p, \frac{y}{x} \right\rangle = \sum_{j=1}^d p^j \frac{y^j}{x^j} \geq \frac{\epsilon}{d} \tag{A2}$$

for $x, y \in \Delta^d$, as at least one of the terms $\frac{y^j}{x^j}$ is greater than or equal to one.

The average performance $L^{\pi_\epsilon}$ defined via (22) for the portfolio map $\pi_\epsilon$ is still almost as good as the optimal average performance $L \equiv L^{\widehat{\pi}}$:

$$\begin{aligned} L^{\pi_\epsilon} &= \int_{\Delta^d} \left[ \int_{\Delta^d} \log\left( \left\langle \pi_\epsilon(x), \frac{y}{x} \right\rangle \right) \varrho(x, dy) \right] d\varrho(x) \\ &\geq \int_{\Delta^d} \left[ \int_{\Delta^d} \log\left( (1-\epsilon) \left\langle \widehat{\pi}(x), \frac{y}{x} \right\rangle \right) \varrho(x, dy) \right] d\varrho(x) \\ &\geq L + \log(1-\epsilon). \end{aligned} \tag{A3}$$

To approximate $\pi_\epsilon$ by a Lipschitz function $\pi_{Lip}$ taking its values in $\bar{\Delta}^d_\epsilon$, we need some preparation. By Assumption 3.6, we can find $\delta > 0$ such that, for $A \subseteq \Delta^d$,

$$\int_A \left[ \int_{\Delta^d} \left( \log\left(\frac{\epsilon}{d}\right) - \log\left( \left\langle \pi_\epsilon(x), \frac{y}{x} \right\rangle \right) \right) \varrho(x, dy) \right] d\varrho(x) > -\epsilon \tag{A4}$$

provided that $\varrho[A] < \delta$. In particular, we may find $\eta > 0$ such that

$$\int_{\Delta^d \setminus \bar{\Delta}^d_\eta} \left[ \int_{\Delta^d} \left( \log\left(\frac{\epsilon}{d}\right) - \log\left( \left\langle \pi_\epsilon(x), \frac{y}{x} \right\rangle \right) \right) d\varrho(x, y) \right] d\varrho(x) > -\epsilon. \tag{A5}$$

Now we find a Lipschitz function $\pi_{Lip} : \Delta^d \to \bar{\Delta}^d_\epsilon$ such that

$$\|\pi_{Lip}(x) - \pi_\epsilon(x)\|_1 = \sum_{j=1}^{d} |\pi_{Lip}(x)^j - \pi_\epsilon(x)^j| < \eta\epsilon^2 \tag{A6}$$

for all $x \in \Delta^d \setminus A$, where the exceptional set $A$ satisfies $\varrho[A] < \delta$. Indeed, the functions from $\mathbb{R}^d \to \bar{\Delta}^d_\epsilon$ which are continuously differentiable in a neighborhood of $\Delta^d$ are dense with respect to the $L^1(\mathbb{R}^d, \varrho; \mathbb{R}^d)$-norm. Let $M$ be a Lipschitz constant for $\pi_{Lip}$ such that $M^{-1} \leq \epsilon$.

To estimate $L^{\pi_{Lip}} - L^{\pi_\epsilon}$ we argue separately on the sets $\Delta^d \setminus \bar{\Delta}^d_\eta$, $A \cap \bar{\Delta}^d_\eta$ and $\bar{\Delta}^d_\eta \setminus A$. To start with the latter set note that, for $x \in \bar{\Delta}^d_\eta$ and $y \in \Delta^d$ we have that the function

$$p \mapsto \left\langle p, \frac{y}{x} \right\rangle = \sum_{i=1}^{n} p^j \frac{y^j}{x^j}, \qquad p \in \bar{\Delta}^d$$

is Lipschitz on $\bar{\Delta}^d$ with Lipschitz constant bounded by $(\frac{\eta}{d})^{-1}$. From (A6), we get

$$\begin{aligned} \int_{\bar{\Delta}^d_\eta \setminus A} \left[ \int_{\Delta^d} \left( \log\left( \left\langle \pi_{Lip}(x), \frac{y}{x} \right\rangle \right) - \log\left( \left\langle \pi_\epsilon(x), \frac{y}{x} \right\rangle \right) d\varrho(x, y) \right] d\varrho(x) \\ \geq -(\eta \cdot \epsilon^2) \left(\frac{\eta}{d}\right)^{-1} \left(\frac{\epsilon}{d}\right)^{-1} \geq -d^2\epsilon. \end{aligned} \tag{A7}$$

The term $(\frac{\epsilon}{d})^{-1}$ above comes from the fact that $\langle \pi_{Lip}(x), \frac{y}{x} \rangle$ as well as $\langle \pi_\epsilon(x), \frac{y}{x} \rangle$ takes values in $[\frac{\epsilon}{d}, \infty[$ and the function $z \mapsto \log(z)$ is Lipschitz on this set with constant $(\frac{\epsilon}{d})^{-1}$.

As regards the set $A \cap \bar{\Delta}_\eta^d$ we obtain from (A2) and (A4) the estimate

$$\int_{A\cap\bar{\Delta}_\eta^d}\left[\int_{\Delta^d}\left(\log\left(\left\langle\pi_{Lip}(x),\frac{y}{x}\right\rangle\right)-\log\left(\left\langle\pi_\epsilon(x),\frac{y}{x}\right\rangle\right)d\varrho(x,y)\right]d\varrho(x)\geq-\epsilon \tag{A8}$$

and a similar estimate holds true for the set $\Delta^d\backslash\bar{\Delta}_\eta^d$ by (A5). Hence, we obtain from (A3), (A7), and (A8)

$$L^{\pi_{Lip}}\geq L+\log(1-\epsilon)-d^2\epsilon-2\epsilon.$$

As $\epsilon>0$ is arbitrary, we have proved Lemma 3.8. □

*Proof of Lemma* 4.1. This follows from the fact that the embedding from $C^{2,\alpha}(\bar{\Delta}^d)\to C^{2,\alpha'}(\bar{\Delta}^d)$ is compact for $\alpha'<\alpha$ (see, e.g., Dobrowolski, 2010, satz 2.42). This means in particular that any bounded set in $C^{2,\alpha}(\bar{\Delta}^d)$ is totally bounded in $C^{2,0}(\bar{\Delta}^d)$, thus relatively compact. To prove compactness it thus suffices to prove that $\mathcal{G}^{M,\alpha}$ is closed. Take a sequence $G^n\in\mathcal{G}^{M,\alpha}$ converging to $G$ with respect to the $\|\cdot\|_{C^{2,0}}$ norm. Then, we can estimate $\|G\|_{C^{2,\alpha}}$ by

$$\begin{aligned}\|G\|_{C^{2,\alpha}}&=\|G\|_{C^{2,0}}+\max_{|\mathbf{k}|=2}\sup_{x\neq y}\frac{|D^{\mathbf{k}}G(x)-D^{\mathbf{k}}G(y)|}{\|x-y\|^\alpha}\\&\leq\|G-G^n\|_{C^{2,0}}+\|G^n\|_{C^{2,0}}\\&\quad+\max_{|\mathbf{k}|=2}\sup_{x\neq y}\frac{|D^{\mathbf{k}}G(x)-D^{\mathbf{k}}G^n(x)|+|D^{\mathbf{k}}G^n(x)-D^{\mathbf{k}}G^n(y)|+|D^{\mathbf{k}}G^n(y)-D^{\mathbf{k}}G(y)|}{\|x-y\|^\alpha}\end{aligned}$$

for any $n\in\mathbb{N}$. Letting $n\to\infty$ and using the fact that $\|G^n-G\|_{C^{2,0}}\to0$ yields $\|G\|_{C^{2,\alpha}}\leq M$. Similarly, we obtain $G\geq\frac{1}{M}$. This together with the fact that $G$ is concave as a limit of concave functions proves $G\in\mathcal{G}^{M,\alpha}$ and thus in turn compactness of $\mathcal{G}^{M,\alpha}$ with respect to $\|\cdot\|_{C^{2,0}}$. □

*Proof of Lemma* 4.4. For $G,\widetilde{G}\in\mathcal{G}^{M,\alpha}$, we have

$$\begin{aligned}\log\left(V_T^G\right)-\log\left(V_T^{\widetilde{G}}\right)&=\log(G(\mu_T))-\log(\widetilde{G}(\mu_T))-(\log(G(\mu_0))-\log(\widetilde{G}(\mu_0)))\\&\quad-\int_0^T\left(\sum_{i,j}\frac{D^{ij}G(\mu_t)}{2G(\mu_t)}-\frac{D^{ij}\widetilde{G}(\mu_t)}{2\widetilde{G}(\mu_t)}\right)d[\mu^i,\mu^j]_t\\&=\log(G(\mu_T))-\log(\widetilde{G}(\mu_T))-(\log(G(\mu_0))-\log(\widetilde{G}(\mu_0)))\\&\quad-\sum_{i,j}\int_0^T\left(\frac{D^{ij}G(\mu_t)-D^{ij}\widetilde{G}(\mu_t)}{2G(\mu_t)}+\frac{D^{ij}\widetilde{G}(\mu_t)\left(\widetilde{G}(\mu_t)-G(\mu_t)\right)}{2\widetilde{G}(\mu_t)G(\mu_t)}\right)d[\mu^i,\mu^j]_t.\end{aligned}$$

Hence, using the fact that $\|\widetilde{G}\|_{C^{2,0}}\leq M$ as well as $G\geq\frac{1}{M}$ and $\widetilde{G}\geq\frac{1}{M}$ and that $z\mapsto\log(z)$ is Lipschitz continuous on $[\frac{1}{M},\infty)$ with constant $M$, we obtain the estimate

$$\left|\log\left(V_T^G\right) - \log\left(V_T^{\widetilde{G}}\right)\right| \leq 2M\|G - \widetilde{G}\|_{C^{2,0}}$$

$$+ \left(\frac{M}{2}d^2\|G - \widetilde{G}\|_{C^{2,0}} + \frac{M^3}{2}d^2\|G - \widetilde{G}\|_{C^{2,0}}\right)\max_i [\mu^i, \mu^i]_T. \tag{A9}$$

This proves the asserted continuity. □