Online Portfolio Selection: A Survey

BIN LI, Wuhan University
STEVEN C. H. HOI, Nanyang Technological University

Online portfolio selection is a fundamental problem in computational finance, which has been extensively studied across several research communities, including finance, statistics, artificial intelligence, machine learning, and data mining. This article aims to provide a comprehensive survey and a structural understanding of online portfolio selection techniques published in the literature. From an online machine learning perspective, we first formulate online portfolio selection as a sequential decision problem, and then we survey a variety of state-of-the-art approaches, which are grouped into several major categories, including benchmarks, Follow-the-Winner approaches, Follow-the-Loser approaches, Pattern-Matching–based approaches, and Meta-Learning Algorithms. In addition to the problem formulation and related algorithms, we also discuss the relationship of these algorithms with the capital growth theory so as to better understand the similarities and differences of their underlying trading ideas. This article aims to provide a timely and comprehensive survey for both machine learning and data mining researchers in academia and quantitative portfolio managers in the financial industry to help them understand the state of the art and facilitate their research and practical applications. We also discuss some open issues and evaluate some emerging new trends for future research.

1. INTRODUCTION

Portfolio selection, aiming to optimize the allocation of wealth across a set of assets, is a fundamental research problem in computational finance and a practical engineering task in financial engineering. There are two major schools for investigating this problem—that is, the mean-variance theory [Markowitz 1952, 1959; Markowitz et al. 2000], mainly from the finance community, and the Capital Growth Theory (CGT) [Kelly 1956; Hakansson and Ziemba 1995], primarily originated from information theory. The mean-variance theory, widely known in the asset management industry, focuses on a single-period (batch) portfolio selection to trade off a portfolio's expected return (mean)

and risk (variance), which typically determines the optimal portfolios subject to the investor's risk-return profile. On the other hand, CGT focuses on multiple-period or sequential portfolio selection, aiming to maximize the portfolio's expected growth rate, or expected log return. Although both theories solve the task of portfolio selection, the latter is fitted to the "online" scenario, which naturally consists of multiple periods and is the focus of this article.

Online portfolio selection, which sequentially selects a portfolio over a set of assets in order to achieve certain targets, is a natural and important task for asset portfolio management. Aiming to maximize the cumulative wealth, several categories of algorithms have been proposed to solve this task. One category of algorithms—Follow the Winner—tries to asymptotically achieve the same growth rate (expected log return) as that of an optimal strategy, which is often based on the CGT. The second category—Follow the Loser—transfers the wealth from winning assets to losers, which seems contradictory to the common sense but empirically often achieves significantly better performance. Finally, the third category—Pattern Matching–based approaches—tries to predict the next market distribution based on a sample of historical data and explicitly optimizes the portfolio based on the sampled distribution. Although these three categories are focused on a single strategy (class), there are also some other strategies that focus on combining multiple strategies (classes)—Meta-Learning Algorithms (MLAs). As a brief summary, Table I outlines the list of main algorithms and corresponding references.

This article provides a comprehensive survey of online portfolio selection algorithms belonging to the mentioned categories. To the best of our knowledge, this is the first survey that includes these three categories and the MLAs as well. Moreover, we are the first to explicitly discuss the connection between the online portfolio selection algorithms and CGT, and illustrate their underlying trading ideas. In the following sections, we also clarify the scope of this article and discuss some related existing surveys in the literature.

1.1. Scope

In this survey, we focus on discussing the empirical motivating ideas of the online portfolio selection algorithms, while only skimming theoretical aspects (e.g., competitive analysis by El-Yaniv [1998] and Borodin et al. [2000] and asymptotical convergence analysis by Györfi et al. [2012]). Moreover, various other related issues and topics are excluded from this survey, as discussed next.

First of all, it is important to mention that the Portfolio Selection task in our survey differs from a great body of financial engineering studies [Kimoto et al. 1993; Merhav and Feder 1998; Cao and Tay 2003; Lu et al. 2009; Dhar 2011; Huang et al. 2011], which attempted to forecast financial time series by applying machine learning techniques and conduct single stock trading [Katz and McCormick 2000; Koolen and Vovk 2012], such as reinforcement learning [Moody et al. 1998; Moody and Saffell 2001; O et al. 2002], neural networks [Kimoto et al. 1993; Dempster et al. 2001], genetic algorithms [Mahfoud and Mani 1996; Allen and Karjalainen 1999; Mandziuk and Jaruszewicz 2011], decision trees [Tsang et al. 2004], and support vector machines [Tay and Cao 2002; Cao and Tay 2003; Lu et al. 2009], and boosting and expert weighting [Creamer 2007; Creamer and Freund 2007, 2010; Creamer 2012]. The key difference between these existing works and the subject area of this survey is that their learning goal is to make explicit predictions of future prices/trends and to trade on a single asset [Borodin et al. 2000, Section 6], whereas our goal is to directly optimize the allocation among a set of assets.

Second, this survey emphasizes the importance of "online" decision for portfolio selection, meaning that related market information arrives sequentially and the allocation decision must be made immediately. Due to the sequential (online) nature of this task,

Table I. General Classification for the State-of-the-Art Online Portfolio Selection Algorithms

Classifications	Algorithms	Representative References
Benchmarks	Buy And Hold	
	Best Stock	
	Constant Rebalanced Portfolios	Kelly [1956]; Cover [1991]
Follow the Winner	Universal Portfolios	Cover [1991]; Cover and Ordentlich [1996]
	Exponential Gradient	Helmbold et al. [1996, 1998]
	Follow the Leader	Gaivoronski and Stella [2000]
	Follow the Regularized Leader	Agarwal et al. [2006]
	Aggregating-Type Algorithms	Vovk and Watkins [1998]
Follow the Loser	Anti-Correlation	Borodin et al. [2003, 2004]
	Passive Aggressive Mean Reversion	Li et al. [2012]
	Confidence Weighted Mean Reversion	Li et al. [2011b, 2013]
	Online Moving Average Reversion	Li and Hoi [2012]
	Robust Median Reversion	Huang et al. [2013]
Pattern-Matching–Based Approaches	Nonparametric Histogram Log-Optimal Strategy	Györfi et al. [2006]
	Nonparametric Kernel-Based Log-Optimal Strategy	
	Nonparametric Nearest Neighbor Log-Optimal Strategy	Györfi et al. [2008]
	Correlation-Driven Nonparametric Learning Strategy	Li et al. [2011a]
	Nonparametric Kernel-Based Semi-Log-Optimal Strategy	Györfi et al. [2007]
	Nonparametric Kernel-Based Markowitz-Type Strategy	Ottucsák and Vajda [2007]
	Nonparametric Kernel-Based GV-Type Strategy	Györfi and Vajda [2008]
Meta-Learning Algorithms	Aggregating Algorithm	Vovk [1990], [1998]
	Fast Universalization Algorithm	Akcoglu et al. [2002, 2004]
	Online Gradient Updates	Das and Banerjee [2011]
	Online Newton Updates	
	Follow the Leading History	Hazan and Seshadhri [2009]

we mainly focus on the survey of multiperiod/sequential portfolio selection work, in which the portfolio is rebalanced to a specified allocation at the end of each trading period [Cover 1991], and the goal typically is to *maximize the expected* log *return* over a sequence of trading periods. We note that these works can be connected to the CGT [Kelly 1956], stemmed from the seminal paper of Kelly [1956] and further developed by Breiman [1960, 1961], Hakansson [1970, 1971], Thorp [1969, 1971], Bell and Cover [1980], Finkelstein and Whitley [1981], Algoet and Cover [1988], Barron and Cover [1988], MacLean et al. [1992], MacLean and Ziemba [1999], Ziemba and Ziemba [2007], MacLean et al. [2010], and others. It has been successfully applied to gambling [Thorp 1962, 1969, 1997], sports betting [Hausch et al. 1981; Ziemba and Hausch 1984, 2008; Thorp 1997], and portfolio investment [Thorp and Kassouf 1967; Rotando and Thorp 1992; Ziemba 2005]. We thus exclude the studies related to the mean-variance port-folio theory [Markowitz 1952, 1959], which were typically developed for single-period (batch) portfolio selection (with the exception of some extensions [Li and Ng 2000; Dai et al. 2010]).

Finally, this article focuses on surveying the algorithmic aspects and providing a structural understanding of the existing online portfolio selection strategies. To prevent loss of focus, we will not dig into theoretical details. In the literature, there is a large body of related work for the theory [MacLean et al. 2011]. Interested researchers can

explore the details of the theory from two exhaustive surveys [Thorp 1997; MacLean and Ziemba 2008], and its history from Poundstone [2005] and Györfi et al. [2012, Chapter 1].

1.2. Related Surveys

There exist several related surveys in this area, but none of them is comprehensive and timely enough for understanding the state-of-the-art of online portfolio selection research. For example, El-Yaniv [1998, Section 5] and Borodin et al. [2000] surveyed the online portfolio selection problem in the framework of competitive analysis. Using our classification in Table I, Borodin et al. mainly surveyed the benchmarks and two Follow-the-Winner algorithms—that is, Universal Portfolios and Exponential Gradient (refer to the details in Section 3.2). Although the competitive framework is important for the Follow-the-Winner category, both surveys are outdated in the sense that they do not include a number of state-of-the-art algorithms afterward. A recent survey by Györfi et al. [2012, Chapter 2] mainly surveyed Pattern-Matching–based approaches (i.e., the third category shown in Table I, which does not include the other categories in this area and is thus far from complete).

1.3. Organization

The remainder of this article is organized as follows. Section 2 formulates the problem of online portfolio selection formally and addresses several practical issues. Section 3 introduces the state-of-the-art algorithms, including Benchmarks in Section 3.1, the Follow-the-Winner approaches in Section 3.2, Follow-the-Loser approaches in Section 3.3, Pattern-Matching–based approaches in Section 3.4, and Meta-Learning algorithms in Section 3.5. Section 4 connects the existing algorithms with the CGT and also illustrates the essentials of their underlying trading ideas. Section 5 discusses several related open issues, and finally Section 6 concludes this survey and outlines some future directions.

2. PROBLEM SETTING

Consider a financial market with m assets, in which we invest our wealth over all assets in the market for a sequence of n trading periods. The market price change is represented by a m-dimensional *price relative vector* $\mathbf{x}_t \in \mathbb{R}_+^m, t = 1, \ldots, n$, where the ith element of tth price relative vector, $x_{t,i}$, denotes the ratio of tth closing price to last closing price for the ith assets. Thus, an investment in asset i on period t increases by a factor of $x_{t,i}$. We also denote the market price changes from period t_1 to t_2 ($t_2 > t_1$) by a *market window*, which consists of a sequence of price relative vectors $\mathbf{x}_{t_1}^{t_2} = \{\mathbf{x}_{t_1}, \ldots, \mathbf{x}_{t_2}\}$, where t_1 denotes the beginning period and t_2 denotes the ending period. One special market window starts from period 1 to n—that is, $\mathbf{x}_1^n = \{\mathbf{x}_1, \ldots, \mathbf{x}_n\}$.

At the beginning of the tth period, an investment is specified by a *portfolio vector* $\mathbf{b}_t, t = 1, \ldots, n$. The ith element of tth portfolio, $b_{t,i}$, represents the proportion of capital invested in the ith asset. Typically, we assume a portfolio is self-financed, and no margin/short is allowed. Thus, a portfolio satisfies the constraint that each entry is nonnegative and all entries sum up to one—that is, $\mathbf{b}_t \in \Delta_m$, where $\Delta_m = \{\mathbf{b} : \mathbf{b} \succeq 0, \mathbf{b}^\top \mathbf{1} = 1\}$. Here, $\mathbf{1}$ is the m-dimensional vector of all 1s, and $\mathbf{b}^\top \mathbf{1}$ denotes the inner product of $\mathbf{b}$ and $\mathbf{1}$. The investment procedure from period 1 to n is represented by a *portfolio strategy*, which is a sequence of mappings as follows:

$$\mathbf{b}_1 = \frac{1}{m}\mathbf{1}, \quad \mathbf{b}_t : \mathbb{R}_+^{m(t-1)} \to \Delta_m, \ t = 2, 3, \ldots, n,$$

where $\mathbf{b}_t = \mathbf{b}_t(\mathbf{x}_1^{t-1})$ denotes the portfolio computed from the past market window $\mathbf{x}_1^{t-1}$. Let us denote the portfolio strategy for n periods as $\mathbf{b}_1^n = \{\mathbf{b}_1, \ldots, \mathbf{b}_n\}$.

For the tth period, a portfolio manager apportions its capital according to portfolio $\mathbf{b}_t$ at the opening time and holds the portfolio until the closing time. Thus, the portfolio wealth will increase by a factor of $\mathbf{b}_t^\top \mathbf{x}_t = \sum_{i=1}^m b_{t,i} x_{t,i}$. Since this model uses price relatives and reinvests the capital, the portfolio wealth will increase multiplicatively. From period 1 to n, a portfolio strategy $\mathbf{b}_1^n$ increases the initial wealth S_0 by a factor of $\prod_{t=1}^n \mathbf{b}_t^\top \mathbf{x}_t$—that is, the final *cumulative wealth* after a sequence of n periods is

$$S_n\left(\mathbf{b}_1^n\right) = S_0 \prod_{t=1}^n \mathbf{b}_t^\top \mathbf{x}_t = S_0 \prod_{t=1}^n \sum_{i=1}^m b_{t,i} x_{t,i}.$$

Since the model assumes multiperiod investment, we define the *exponential growth rate* for a strategy $\mathbf{b}_1^n$ as

$$W_n\left(\mathbf{b}_1^n\right) = \frac{1}{n} \log S_n\left(\mathbf{b}_1^n\right) = \frac{1}{n} \sum_{t=1}^n \log \mathbf{b}_t \cdot \mathbf{x}_t.$$

Finally, let us combine all elements and formulate the online portfolio selection model. In a portfolio selection task, the decision maker is a portfolio manager, whose goal is to produce a portfolio strategy $\mathbf{b}_1^n$ in order to achieve certain targets. Following the principles conveyed by the algorithms in Table I, our target is to maximize the portfolio cumulative wealth S_n. The portfolio manager computes the portfolio strategy in a sequential fashion. On the beginning of period t, based on the previous market window $\mathbf{x}_1^{t-1}$, the portfolio manager learns a new portfolio vector $\mathbf{b}_t$ for the coming price relative vector $\mathbf{x}_t$, where the decision criterion varies among different managers/strategies. The portfolio $\mathbf{b}_t$ is scored using the portfolio period return $\mathbf{b}_t \cdot \mathbf{x}_t$. This procedure is repeated until period n, and the strategy is finally scored according to the portfolio cumulative wealth S_n. Algorithm 1 shows the framework of online portfolio selection, which serves as a general procedure to backtest any online portfolio selection algorithm.

ALGORITHM 1: Online portfolio selection framework.

Input: $\mathbf{x}_1^n$: Historical market sequence
Output: S_n: Final cumulative wealth

Initialize $S_0 = 1, \mathbf{b}_1 = \left(\frac{1}{m}, \ldots, \frac{1}{m}\right)$
for $t = 1, 2, \ldots, n$ **do**
    Portfolio manager computes a portfolio $\mathbf{b}_t$;
    Market reveals the market price relative $\mathbf{x}_t$;
    Portfolio incurs period return $\mathbf{b}_t^\top \mathbf{x}_t$ and updates cumulative return $S_t = S_{t-1} \times \left(\mathbf{b}_t^\top \mathbf{x}_t\right)$;
    Portfolio manager updates his/her online portfolio selection rules ;
end

In general, some assumptions are made in the previous widely adopted model:

(1) *Transaction cost*. We assume no transaction costs/taxes in the model.
(2) *Market liquidity*. We assume that one can buy and sell any quantity of any asset in its closing prices.
(3) *Impact cost*. W assume market behavior is not affected by any portfolio selection strategy.

To better understand these notions and the model presented let us illustrate with a classical example.

Example 2.1 (*Synthetic market by Cover and Gluss* [*1986*]).   Assume a two-asset market with cash and one volatile asset with the price relative sequence $\mathbf{x}_1^n = \{(1, 2), (1, \frac{1}{2}), (1, 2), \dots\}$. The 1st price relative vector $\mathbf{x}_1 = (1, 2)$ means that if we invest \$1 in the first asset, you will get \$1 at the end of period; if we invest \$1 in the second asset, we will get \$2 after the period.

Let a fixed proportion portfolio strategy be $\mathbf{b}_1^n = \{(\frac{1}{2}, \frac{1}{2}), (\frac{1}{2}, \frac{1}{2}), \dots\}$, which means that every day, the manager redistributes the capital equally among the two assets. For the 1st period, the portfolio wealth increases by a factor of $1 \times \frac{1}{2} + 2 \times \frac{1}{2} = \frac{3}{2}$. Initializing the capital with $S_0 = 1$, then the capital at the end of the 1st period equals $S_1 = S_0 \times \frac{3}{2} = \frac{3}{2}$. Similarly, $S_2 = S_1 \times (1 \times \frac{1}{2} + \frac{1}{2} \times \frac{1}{2}) = \frac{3}{2} \times \frac{3}{4} = \frac{9}{8}$. Thus, at the end of period n, the final cumulative wealth equals

$$
S_n\left(\mathbf{b}_1^n\right) =
\begin{cases}
\frac{9}{8}^{\frac{n}{2}} & n \text{ is even} \\
\frac{3}{2} \times \frac{9}{8}^{\frac{n-1}{2}} & n \text{ is odd}
\end{cases},
$$

and the exponential growth rate is

$$
W_n\left(\mathbf{b}_1^n\right) =
\begin{cases}
\frac{1}{2} \log \frac{9}{8} & n \text{ is even} \\
\frac{n-1}{2n} \log \frac{9}{8} + \frac{1}{n} \log \frac{3}{2} & n \text{ is odd}
\end{cases},
$$

which approaches $\frac{1}{2} \log \frac{9}{8} > 0$ if n is sufficiently large.

2.1. Transaction Cost

In reality, the most important and unavoidable issue is *transaction costs*. In this section, we model the transaction costs into our formulation, which enables us to evaluate online portfolio selection algorithms. However, we will not introduce strategies [Davis and Norman 1990; Iyengar and Cover 2000; Akian et al. 2001; Schäfer 2002; Györfi and Vajda 2008; Ormos and Urbán 2011] that directly solve the transaction costs issues.

The widely adopted transaction costs model is the *proportional transaction costs* model [Blum and Kalai 1999; Györfi and Vajda 2008], in which the incurred transaction cost is proportional to the wealth transferred during rebalancing. Let the brokers charge transaction costs on both buying and selling. At the beginning of the tth period, the portfolio manager intends to rebalance the portfolio from closing price adjusted portfolio $\hat{\mathbf{b}}_{t-1}$ to a new portfolio $\mathbf{b}_t$. Here, $\hat{\mathbf{b}}_{t-1}$ is calculated as $\hat{b}_{t-1,i} = \frac{b_{t-1,i} x_{t-1,i}}{\mathbf{b}_{t-1} \cdot \mathbf{x}_{t-1}}, i = 1, \dots, m$. Assuming two transaction cost rates $\gamma_b \in (0, 1)$ and $\gamma_s \in (0, 1)$, where γ_b denotes the transaction costs rate incurred during buying and γ_s denotes the transaction costs rate incurred during selling. After rebalancing, S_{t-1} will be decomposed into two parts—that is, the net wealth N_{t-1} in the new portfolio $\mathbf{b}_t$ and the transaction costs incurred during the buying and selling. If the wealth on asset i before rebalancing is higher than that after reblancing—that is, $\frac{b_{t-1,i} x_{t-1,i}}{\mathbf{b}_{t-1} \cdot \mathbf{x}_{t-1}} S_{t-1} \geq b_{t,i} N_{t-1}$— then there will be a selling rebalancing. Otherwise, a buying rebalancing is required. Formally,

$$
S_{t-1} = N_{t-1} + \gamma_s \sum_{i=1}^m \left(\frac{b_{t-1,i} x_{t-1,i}}{\mathbf{b}_{t-1} \cdot \mathbf{x}_{t-1}} S_{t-1} - b_{t,i} N_{t-1} \right)^+ + \gamma_b \sum_{i=1}^m \left(b_{t,i} N_{t-1} - \frac{b_{t-1,i} x_{t-1,i}}{\mathbf{b}_{t-1} \cdot \mathbf{x}_{t-1}} S_{t-1} \right)^+.
$$

Let us denote *transaction costs factor* [Györfi and Vajda 2008] as the ratio of net wealth after rebalancing to wealth before rebalancing—that is, $c_{t-1} = \frac{N_{t-1}}{S_{t-1}} \in (0, 1)$. Dividing

the previous equation by S_{t-1}, we get

$$1 = c_{t-1} + \gamma_s \sum_{i=1}^{m} \left(\frac{b_{t-1,i} x_{t-1,i}}{\mathbf{b}_{t-1} \cdot \mathbf{x}_{t-1}} - b_{t,i} c_{t-1} \right)^+ + \gamma_b \sum_{i=1}^{m} \left(b_{t,i} c_{t-1} - \frac{b_{t-1,i} x_{t-1,i}}{\mathbf{b}_{t-1} \cdot \mathbf{x}_{t-1}} \right)^+. \quad (1)$$

Clearly, given $\mathbf{b}_{t-1}$, $\mathbf{x}_{t-1}$, and $\mathbf{b}_t$, there exists a unique transaction costs factor for each rebalancing. Thus, we can denote c_{t-1} as a function, $c_{t-1} = c(\mathbf{b}_t, \mathbf{b}_{t-1}, \mathbf{x}_{t-1})$. Moreover, considering that the portfolio is in the simplex domain, then the factor ranges between $\frac{1-\gamma_s}{1+\gamma_b} \leq c_{t-1} \leq 1$.

Finally, for each period t, the wealth grows by a factor as

$$S_t = S_{t-1} \times c_{t-1} \times (\mathbf{b}_t \cdot \mathbf{x}_t),$$

and the final cumulative wealth after n periods equals

$$S_n = S_0 \prod_{t=1}^{n} c_{t-1} \times (\mathbf{b}_t \cdot \mathbf{x}_t),$$

where c_{t-1} is calculated as Equation (1).

3. ONLINE PORTFOLIO SELECTION APPROACHES

In this section, we survey the area of online portfolio selection. Algorithms in this area formulate the online portfolio selection task as in Section 2 and derive explicit portfolio update schemes for each period. Basically, the routine is to implicitly assume various price relative predictions and learn optimal portfolios.

In the subsequent sections, we mainly list the algorithms following Table I. In particular, we first introduce several benchmark algorithms in Section 3.1. Then, we introduce the algorithms with explicit update schemes in the subsequent three sections. We classify them based on the direction of the weight transfer. The first approach—Follow the Winner—tries to increase the relative weights of more successful experts/stocks, often based on their historical performance. On the contrary, the second approach—Follow the Loser—tries to increase the relative weights of less successful experts/stocks, or transfer the weights from winners to losers. The third approach—Pattern Matching—tries to build a portfolio based on some sampled similar historical patterns with no explicit weights transfer directions. After that, we survey MLAs, which can be applied to higher-level experts equipped with any existing algorithm.

3.1. Benchmarks

3.1.1. Buy-and-Hold Strategy. The most common baseline is the Buy-and-Hold (BAH) strategy, in which one invests wealth among a pool of assets with an initial portfolio $\mathbf{b}_1$ and holds the portfolio until the end. The manager only buys the assets at the beginning of the 1st period and does not rebalance in the following periods, and the portfolio holdings are implicitly changed following the market fluctuations. For example, at the end of the 1st period, the portfolio holding becomes $\frac{\mathbf{b}_1 \odot \mathbf{x}_1}{\mathbf{b}_1^\top \mathbf{x}_1}$, where $\odot$ denotes an elementwise product. In summary, the final cumulative wealth achieved by a BAH strategy is the initial portfolio weighted average of individual stocks' final wealth,

$$S_n\left(BAH\left(\mathbf{b}_1\right)\right) = \mathbf{b}_1 \cdot \left(\bigodot_{t=1}^{n} \mathbf{x}_t \right).$$

The BAH strategy with initial *uniform* portfolio $\mathbf{b}_1 = (\frac{1}{m}, \ldots, \frac{1}{m})$ is referred to as uniform BAH strategy, which is often adopted as a *market* strategy to produce a market index.

3.1.2. Best Stock Strategy. Another widely adopted benchmark is the Best Stock (Best) strategy, which is a special BAH strategy that puts all capital on the stock with best performance in hindsight. Clearly, its initial portfolio $\mathbf{b}^\circ$ in hindsight can be calculated as

$$\mathbf{b}^\circ = \operatorname*{arg\,max}_{\mathbf{b} \in \Delta_m} \mathbf{b} \cdot \left(\bigodot_{t=1}^{n} \mathbf{x}_t \right).$$

As a result, the final cumulative wealth achieved by the Best strategy can be calculated as

$$S_n\left(Best\right) = \max_{\mathbf{b} \in \Delta_m} \mathbf{b} \cdot \left(\bigodot_{t=1}^{n} \mathbf{x}_t \right) = S_n\left(BAH\left(\mathbf{b}^\circ\right)\right).$$

3.1.3. Constant Rebalanced Portfolios. Another more sophisticated benchmark strategy is the Constant Rebalanced Portfolio (CRP) strategy, which rebalances the portfolio to a fixed portfolio $\mathbf{b}$ every period. In particular, the portfolio strategy can be represented as $\mathbf{b}_1^n = \{\mathbf{b}, \mathbf{b}, \dots\}$. Thus, the cumulative portfolio wealth achieved by a CRP strategy after n periods is defined as

$$S_n(CRP(\mathbf{b})) = \prod_{t=1}^{n} \mathbf{b}^\top \mathbf{x}_t.$$

One special CRP strategy that rebalances to uniform portfolio $\mathbf{b} = (\frac{1}{m}, \dots, \frac{1}{m})$ each period is the Uniform Constant Rebalanced Portfolio (UCRP). It is possible to calculate an optimal offline portfolio for the CRP strategy as

$$\mathbf{b}^\star = \operatorname*{arg\,max}_{\mathbf{b}^n \in \Delta_m} \log S_n(CRP(\mathbf{b})) = \operatorname*{arg\,max}_{\mathbf{b} \in \Delta_m} \sum_{t=1}^{n} \log(\mathbf{b}^\top \mathbf{x}_t),$$

which is convex and can be efficiently solved. The CRP strategy with $\mathbf{b}^\star$ is denoted by Best Constant Rebalanced Portfolio (BCRP). BCRP achieves a final cumulative portfolio wealth and corresponding exponential growth rate defined as follows:

$$S_n(BCRP) = \max_{\mathbf{b} \in \Delta_m} S_n(CRP(\mathbf{b})) = S_n(CRP(\mathbf{b}^\star)),$$

$$W_n(BCRP) = \frac{1}{n} \log S_n(BCRP) = \frac{1}{n} \log S_n(CRP(\mathbf{b}^\star)).$$

Note that BCRP strategy is a hindsight strategy, which can only be calculated with complete market sequences. Cover [1991] proved the benefits of BCRP as a target—that is, BCRP exceeds the best stock, Value Line Index (geometric mean of component returns), and the Dow Jones Index (arithmetic mean of component returns, or BAH). Moreover, BCRP is invariant under permutations of the price relative sequences—that is, it does not depend on the order in which $\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_n$ occur.

Until now, let us compare BAH and CRP strategies by continuing Example 2.1.

Example 3.1 (*Synthetic Market by Cover and Gluss* [*1986*]). Assume a two-asset market with cash and one volatile asset with the price relative sequence $\mathbf{x}_1^n = \{(1, 2), (1, \frac{1}{2}), (1, 2), \dots\}$. Let us consider BAH with uniform initial portfolio $\mathbf{b}_1 = (\frac{1}{2}, \frac{1}{2})$ and the CRP with uniform portfolio $\mathbf{b} = (\frac{1}{2}, \frac{1}{2})$. Clearly, since no asset grows in the long run, the final wealth of BAH equals the uniform weighted summation of two assets, which roughly equals to 1 in the long run. On the other hand, according to the analysis of Example 2.1, the final cumulative wealth of CRP is roughly $\frac{9}{8}^{\frac{n}{2}}$, which increases

exponentially. Note that the BAH only rebalances on the 1st period, whereas the CRP rebalances every period. While market provides no return on the synthetic market, CRP can produce an exponentially increasing return. The underlying idea of CRP is to take advantage of the underlying volatility, or so-called volatility pumping [Luenberger 1998, Chapter 15].

Since CRP rebalances a fixed portfolio each period, its frequent transactions will incur high transaction costs. Helmbold et al. [1996, 1998] proposed a Semi-Constant Rebalanced Portfolio (Semi-CRP), which rebalances the portfolio on selected periods rather than every period.

One desired theoretical result for online portfolio selection is *universality* [Cover 1991]. An online portfolio selection algorithm *Alg* is *universal* if its average *(external) regret* [Stoltz and Lugosi 2005; Blum and Mansour 2007] for n periods asymptotically approaches 0,

$$\frac{1}{n}\text{regret}_n(Alg) = W_n(BCRP) - W_n(Alg) \longrightarrow 0, \text{ as } n \to \infty. \tag{2}$$

In other words, a universal portfolio selection algorithm asymptotically approaches the same exponential growth rate as BCRP strategy for arbitrary sequences of price relatives.

3.2. Follow-the-Winner Approaches

Follow the Winner is characterized by increasing the relative weights of more successful experts/stocks. Rather than targeting market and best stock, algorithms in this category often aim to track the BCRP strategy, which can be shown to be the optimal strategy in an i.i.d. market [Cover and Thomas 1991, Theorem 15.3.1]. In other words, such optimality motivates that universal portfolio selection algorithms approach the performance of the hindsight BCRP for arbitrary sequence of price relative vectors, called *individual sequences*.

3.2.1. Universal Portfolios. The basic idea of Universal Portfolio–type algorithms is to assign the capital to a single class of base experts, let the experts run, and finally pool their wealth. Strategies in this type are analogous to the BAH strategy. Their difference is that base BAH expert is the strategy investing on a single stock, and thus the number of experts is the same as that of stocks. In other words, BAH strategy buys the individual stocks and lets the stocks go and finally pools their individual wealth. On the other hand, the base expert in the Follow-the-Winner category can be any strategy class that invests in any set of stocks in the market. Besides, algorithms in this category are also similar to the MLAs further described in Section 3.5, although MLA generally applies to experts of multiple classes.

Cover [1991] proposed the Universal Portfolio (UP) strategy, and Cover and Ordentlich [1996] further refined the algorithm as μ-Weighted Universal Portfolio, in which μ denotes a given distribution on the space of valid portfolio Δ_m. Intuitively, Cover's UP operates similar to a Fund of Funds (FOF), and its main idea is to BAH the parameterized CRP strategies over the whole simplex domain. In particular, it initially invests a proportion of wealth $d\mu(\mathbf{b})$ to each portfolio manager operating CRP strategy with $\mathbf{b} \in \Delta_m$, and lets the CRP managers run. Then, at the end, each manager will grow his wealth to $S_n(\mathbf{b})d\mu(\mathbf{b})$. Finally, Cover's UP pools the individual experts' wealth over the continuum of portfolio strategies. Note that $S_n(\mathbf{b}) = e^{nW_n(\mathbf{b})}$, which means that the portfolio grows at an exponential rate of $W_n(\mathbf{b})$.

Formally, its update scheme [Cover and Ordentlich 1996, Definition 1] can be interpreted as a historical performance weighed average of all valid CRPs,

$$\mathbf{b}_{t+1} = \frac{\int_{\Delta_m} \mathbf{b} S_t(\mathbf{b}) d\mu(\mathbf{b})}{\int_{\Delta_m} S_t(\mathbf{b}) d\mu(\mathbf{b})}.$$

Note that at the beginning of period $t + 1$, one CRP manager's wealth (historical performance) equals to $S_t(\mathbf{b}) d\mu(\mathbf{b})$. Incorporating the initial wealth of $S_0 = 1$, the final cumulative wealth is the weighted average of CRP managers' wealth [Cover and Ordentlich 1996, Eq. (24)],

$$S_n(UP) = \int_{\Delta_m} S_n(\mathbf{b}) d\mu(\mathbf{b}). \tag{3}$$

One special case is that μ follows a uniform distribution, then the portfolio update reduces to Cover's UP [Cover 1991, Eq. (1.3)]. Another special case is Dirichlet $(\frac{1}{2}, \dots, \frac{1}{2})$ weighted UP [Cover and Ordentlich 1996], which is proved to be a more optimal allocation. Alternatively, if a loss function is defined as the negative logarithmic function of portfolio return, Cover's UP is actually an exponentially weighted average forecaster [Cesa-Bianchi and Lugosi 2006].

Cover [1991] showed that with suitable smoothness conditions the average of exponentials grows at the same exponential rate as the maximum, thus one can asymptotically approach BCRP's exponential growth rate. The regret achieved by Cover's UP is $O(m \log n)$, and its time complexity is $O(n^m)$, where m denotes the number of stocks and n refers to the number of periods. Cover and Ordentlich [1996] proved that the $(\frac{1}{2}, \dots, \frac{1}{2})$ weighted Universal Portfolios has the same scale of regret bound but a better constant term [Cover and Ordentlich 1996, Theorem 2].

As Cover's UP is based on an ideal market model, one research topic with respect to Cover's UP is to extend the algorithm with various realistic assumptions. Cover and Ordentlich [1996] extended the model to include side information, which can be instantiated experts' opinions, fundamental data, and so forth. Blum and Kalai [1999] took account of transaction costs for online portfolio selection and proposed a universal portfolio algorithm to handle the costs.

Another research topic is to generalize Cover's UP with different underlying base expert classes, rather than the CRP strategy. Jamshidian [1992] generalized the algorithm for continuous time market and derived the long-term performance of Cover's UP in this setting. Vovk and Watkins [1998] applied Aggregating Algorithm (AA) [Vovk 1990] to a finite number of arbitrary investment strategies. Cover's UP becomes a specialized case of AA when applied to an infinite number of CRPs. We will further investigate AA in Section 3.2.5. Ordentlich and Cover [1998] derived the lower bound of the final wealth achieved by any nonanticipating investment strategy to that of BCRP strategy. Cross and Barron [2003] generalized Cover's UP from CRP strategy class to any parameterized target class and proposed a universal strategy that costs a polynomial time. Akcoglu et al. [2002, 2004] extended Cover's UP from the parameterized CRP class to a wide class of investment strategies, including trading strategies operating on a single stock and portfolio strategies operating on the whole stock market. Kozat and Singer [2011] proposed a similar universal algorithm based on the class of Semi-CRPs [Helmbold et al. 1996, 1998], which provides good performance with transaction costs.

Rather than the previous analysis, various work has also been proposed to discuss the connection between Cover's UP with universal prediction [Feder et al. 1992], data compression [Rissanen 1983] and Markowitz's mean-variance theory [Markowitz 1952, 1959]. Algoet [1992] discussed the universal schemes for prediction, gambling, and portfolio selection. Cover [1996] and Ordentlich [1996] discussed the connection

of universal portfolio selection and data compression. Belentepe [2005] presented a statistical view of Cover's UP strategy and connected it with traditional Markowitz's mean-variance portfolio theory [Markowitz 1952]. The authors showed that by allowing short selling and leverage, UP is approximately equivalent to sequential mean-variance optimization; otherwise, the strategy is approximately equivalent to constrained sequential optimization. Although its update scheme is distributional free, UP implicitly estimates the multivariate mean and covariance matrix.

Cover's UP has a good theoretical performance guarantee; however, its implementation costs exponential time in the number of assets, which restricts its practical capability. To overcome this computational bottleneck, Kalai and Vempala [2002] presented an efficient implementation based on nonuniform random walks that are rapidly mixing. Their implementation requires a poly running time of $O(m^7 n^8)$, which is a substantial improvement of the original bound of $O(n^m)$.

3.2.2. Exponential Gradient. The strategies in the Exponential Gradient type generally focus on the following optimization problem:

$$\mathbf{b}_{t+1} = \arg\max_{\mathbf{b} \in \Delta_m} \quad \eta \log \mathbf{b} \cdot \mathbf{x}_t - R(\mathbf{b}, \mathbf{b}_t), \tag{4}$$

where $R(\mathbf{b}, \mathbf{b}_t)$ denotes a regularization term and $\eta > 0$ is the learning rate. One straightforward interpretation of the optimization is to track the stock with the best performance in the last period but keep the new portfolio close to the previous portfolio. This is obtained using the regularization term $R(\mathbf{b}, \mathbf{b}_t)$.

Helmbold et al. [1996, 1998] proposed the Exponential Gradient (EG) strategy, which is based on the algorithm proposed for the mixture estimation problem [Helmbold et al. 1997]. The EG strategy employs the relative entropy as the regularization term in Equation (4),

$$R(\mathbf{b}, \mathbf{b}_t) = \sum_{i=1}^{m} b_i \log \frac{b_i}{b_{t,i}}.$$

EG's formulation is thus convex in $\mathbf{b}$, however, it is hard to solve since log is nonlinear. Thus, the authors adopted log's first-order Taylor expansion at $\mathbf{b}_t$,

$$\log \mathbf{b} \cdot \mathbf{x}_t \approx \log(\mathbf{b}_t \cdot \mathbf{x}_t) + \frac{\mathbf{x}_t}{\mathbf{b}_t \cdot \mathbf{x}_t}(\mathbf{b} - \mathbf{b}_t),$$

with which the first term in Equation (4) becomes linear and easy to solve. Solving the optimization, the update rule [Helmbold et al. 1998, Equation (3.3)] becomes

$$b_{t+1,i} = b_{t,i} \exp\left(\eta \frac{x_{t,i}}{\mathbf{b}_t \cdot \mathbf{x}_t}\right)/Z, \quad i = 1, \ldots, m,$$

where Z denotes the normalization term such that the portfolio sums to 1.

The optimization problem in Equation (4) can also be solved using the Gradient Projection (GP) and Expectation Maximization (EM) method [Helmbold et al. 1997]. GP and EM adopt different regularization terms. In particular, GP adopts L2-norm regularization, and EM adopts χ^2 regularization:

$$R(\mathbf{b}, \mathbf{b}_t) = \begin{cases} \frac{1}{2} \sum_{i=1}^{m} (b_i - b_{t,i})^2 & \text{GP} \\ \frac{1}{2} \sum_{i=1}^{m} \frac{(b_i - b_{t,i})^2}{b_{t,i}} & \text{EM} \end{cases}.$$

The final update rule of GP [Helmbold et al. 1997, Eq. (5)] is

$$b_{t+1,i} = b_{t,i} + \eta \left(\frac{x_{t,i}}{\mathbf{b}_t \cdot \mathbf{x}_t} - \frac{1}{m} \sum_{i=1}^{m} \frac{x_{t,i}}{\mathbf{b}_t \cdot \mathbf{x}_t} \right),$$

and the update rule of EM [Helmbold et al. 1997, Eq. (7)] is

$$b_{t+1,i} = b_{t,i} \left(\eta \left(\frac{x_{t,i}}{\mathbf{b}_t \cdot \mathbf{x}_t} - 1 \right) + 1 \right),$$

which can also be viewed as the first-order approximation of EG's update formula.

The regret of the EG strategy can be bounded by $O(\sqrt{n \log m})$ with $O(m)$ running time per period. The regret is not as tight as that of Cover's UP; however, its linear running time substantially surpasses that of Cover's UP. Besides, the authors also proposed a variant, which has a regret bound of $O(m^{0.5}(\log m)^{0.25}n^{0.75})$. Although not proposed for online portfolio selection task, according to Helmbold et al. [1997], GP can straightforwardly achieve a regret of $O(\sqrt{mn})$, which is significantly worse than that of EG.

One key parameter for EG is the learning rate $\eta > 0$. In order to achieve the desired regret bound presented previously, η has to be small. However, as $\eta \to 0$, its update approaches uniform portfolio, and EG reduces to UCRP.

Das and Banerjee [2011] extended the EG algorithm to the sense of the MLA named Online Gradient Updates (OGU), which will be introduced in Section 3.5.3. OGU combines underlying experts such that the overall system can achieve the performance that is no worse than any convex combination of base experts.

To handle the case of nonzero transaction costs, Das et al. [2013] extended the GP algorithm by appending a L1 regularization to formulation, and proposed *Online Lazy Update* (OLU) algorithm.

3.2.3. Follow the Leader. Strategies in the Follow-the-Leader (FTL) approach try to track the BCRP strategy until time t—that is,

$$\mathbf{b}_{t+1} = \mathbf{b}_t^* = \arg\max_{\mathbf{b} \in \Delta_m} \sum_{j=1}^{t} \log (\mathbf{b} \cdot \mathbf{x}_j). \tag{5}$$

Clearly, this category follows the BCRP leader, and the ultimate leader is the BCRP over all periods.

Ordentlich [1996, Chapter 4.4] briefly mentioned a strategy to obtain portfolios by mixing the BCRP up to time t and uniform portfolio

$$\mathbf{b}_{t+1} = \frac{t}{t+1}\mathbf{b}_t^* + \frac{1}{t+1}\frac{1}{m}\mathbf{1}.$$

He also showed its worst case bound, which is slightly worse than that of Cover's UP.

Gaivoronski and Stella [2000] proposed Successive Constant Rebalanced Portfolios (SCRP) and Weighted Successive Constant Rebalanced Portfolios (WSCRP) for stationary markets. For each period, SCRP directly adopts the BCRP portfolio until now, that is,

$$\mathbf{b}_{t+1} = \mathbf{b}_t^*.$$

The authors further solved the optimal portfolio $\mathbf{b}_t^*$ via stochastic optimization [Birge and Louveaux 1997], resulting in the detail updates of SCRP [Gaivoronski and Stella 2000, Algorithm 1]. On the other hand, WSCRP outputs a convex combination of SCRP portfolio and last portfolio,

$$\mathbf{b}_{t+1} = (1 - \gamma)\mathbf{b}_t^* + \gamma\mathbf{b}_t,$$

where $\gamma \in [0, 1]$ represents the trade-off parameter.

The regret bounds achieved by SCRP [Gaivoronski and Stella 2000, Theorem 1] and WSCRP [Gaivoronski and Stella 2000, Theorem 4] are both $O(K^2 \log n)$, where K is a

uniform upper bound of the gradient of $\log \mathbf{b}^{\top}\mathbf{x}$ with respect to $\mathbf{b}$. It is straightforward to see that given the same assumption of upper/lower bound of price relatives as Cover's UP [Cover 1991, Theorem 6.1], the regret bound is on the same scale of Cover's UP, although the constant term is slightly worse.

Rather than assuming that the historical market is stationary, some algorithms assume that the historical market is nonstationary. Gaivoronski and Stella [2000] propose Variable Rebalanced Portfolios (VRP), which calculates the BCRP portfolio based on a latest sliding window. To be more specific, VRP updates its portfolio as follows:

$$\mathbf{b}_{t+1} = \arg\max_{\mathbf{b}\in\Delta_m} \sum_{j=t-W+1}^{t} \log(\mathbf{b}\cdot\mathbf{x}_j),$$

where W denotes a specified window size. Following their algorithms for CRP, they further proposed Successive Variable Rebalanced Portfolios (SVRP) and Weighted Successive Variable Rebalanced Portfolios (WSVRP). No theoretical results were given on the two algorithms.

Gaivoronski and Stella [2003] further generalized Gaivoronski and Stella [2000] and proposed Adaptive Portfolio Selection (APS) for online portfolio selection task. By changing the objective part, APS can handle three types of portfolio selection task—that is, adaptive Markowitz portfolio, log-optimal CRP, and index tracking. To handle the transaction cost issue, they proposed Threshold Portfolio Selection (TPS), which only rebalances the portfolio if the expected return of new portfolio exceeds that of a previous portfolio for more than a threshold.

3.2.4. Follow the Regularized Leader. Another category of approaches follows a similar idea as FTL, while adding a regularization term, and thus actually becomes the Follow the Regularized Leader (FTRL) approach. In generally, FTRL approaches can be formulated as follows:

$$\mathbf{b}_{t+1} = \arg\max_{\mathbf{b}\in\Delta_m} \sum_{\tau=1}^{t} \log(\mathbf{b}\cdot\mathbf{x}_\tau) - \frac{\beta}{2}R(\mathbf{b}), \tag{6}$$

where β denotes the trade-off parameter and $R(\mathbf{b})$ is a regularization term on $\mathbf{b}$. Note that here all historical information is captured in the first term, thus the regularization term only concerns the next portfolio, which is different from the EG algorithm. One typical regularization is an L2-norm—that is, $R(\mathbf{b}) = \|\mathbf{b}\|^2$.

Agarwal et al. [2006] proposed the Online Newton Step (ONS) by solving the optimization problem in Equation (6) with L2-norm regularization via online convex optimization technique [Zinkevich 2003; Hazan et al. 2006, 2007; Hazan 2006]. Similar to the Newton method for offline optimization, the basic idea is to replace the log term via its second-order Taylor expansion at $\mathbf{b}_t$, then solve the problem for closed-form update scheme. Finally, the ONS update rule [Agarwal et al. 2006, Lemma 2] is

$$\mathbf{b}_1 = \left(\frac{1}{m},\ldots,\frac{1}{m}\right), \quad \mathbf{b}_{t+1} = \Pi_{\Delta_m}^{\mathbf{A}_t}(\delta\mathbf{A}_t^{-1}\mathbf{p}_t),$$

with

$$\mathbf{A}_t = \sum_{\tau=1}^{t}\left(\frac{\mathbf{x}_\tau\mathbf{x}_\tau^{\top}}{(\mathbf{b}_\tau\cdot\mathbf{x}_\tau)^2}\right) + \mathbf{I}_m, \quad \mathbf{p}_t = \left(1+\frac{1}{\beta}\right)\sum_{\tau=1}^{t}\frac{\mathbf{x}_\tau}{\mathbf{b}_\tau\cdot\mathbf{x}_\tau},$$

where β is the trade-off parameter, δ is a scale term, and $\Pi_{\Delta_m}^{\mathbf{A}_t}(\cdot)$ is an exact projection to the simplex domain.

ONS iteratively updates the first- and second-order information and the portfolio with a time cost of $O(m^3)$, which is irrelevant to the number of historical instances t. The authors also proved ONS's regret bound [Agarwal et al. 2006, Theorem 1] of $O(m^{1.5} \log(mn))$, which is worse than Cover's UP and Dirichlet (1/2) weighted UP.

While FTRL or even the Follow-the-Winner category mainly focuses on the worst-case investing, Hazan and Kale [2009, 2012] linked the worst-case model with practically widely used average-case investing—that is, the Geometric Brownian Motion (GBM) model [Bachelier 1900; Osborne 1959; Cootner 1964], which is a probabilistic model of stock returns. The authors also designed an investment strategy that is universal in the worst case and is capable of exploiting the GBM model. Their algorithm, or so-called Exp-Concave-FTL, follows a slightly different form of optimization problem (6) with L2-norm regularization,

$$\mathbf{b}_{t+1} = \operatorname*{arg\,max}_{\mathbf{b} \in \Delta_m} \sum_{\tau=1}^{t} \log(\mathbf{b} \cdot \mathbf{x}_\tau) - \frac{1}{2}\|\mathbf{b}\|^2.$$

Similar to ONS, the optimization problem can be efficiently solved via online convex optimization technique. The authors further analyzed its regret bound and linked it with the GBM model. Linking the GBM model, the regret round [Hazan and Kale 2012, Theorem 1.1 and Corollary 1.2] is $O(m\log(Q+m))$, where Q denotes the quadratic variability, calculated as $n-1$ times the sample variance of the sequence of price relative vectors. Since Q is typically much smaller than n, the regret bound significantly improves the $O(m\log n)$ bound.

Besides the improved regret bound, the authors also discussed the relationship of their algorithm's performance to trading frequency. The authors asserted that increasing the trading frequency would decrease the variance of the minimum variance CRP—that is, the more frequently they trade, the more likely the payoff will be close to the expected value. On the other hand, the regret stays the same even if they trade more. Consequently, it is expected to see improved performance of such algorithm as the trading frequency increases [Agarwal et al. 2006].

Das and Banerjee [2011] further extended the FTRL approach to a generalized MLA—that is, Online Newton Update (ONU), which guarantees that the overall performance is no worse than any convex combination of its underlying experts.

3.2.5. Aggregating-Type Algorithms. Although BCRP is the optimal strategy for an i.i.d. market, the i.i.d. assumption is controversial in real markets, so the optimal portfolio may not belong to CRP or fixed fraction portfolio. Some algorithms have been designed to track a different set of experts. The algorithms in this category share similar idea to the MLAs in Section 3.5. However, here the base experts are of a special class—that is, individual expert that invests fully on a single stock—although in general MLAs often apply to more complex experts from multiple classes.

Vovk and Watkins [1998] applied the AA [Vovk 1990, 1997, 1999, 2001] to the online portfolio selection task, of which Cover's UP is a special case. The general setting for AA is to define a countable or finite set of base experts and sequentially allocate the resource among multiple base experts in order to achieve a good performance that is no worse than any fixed combination of underlying experts. Although its general form is shown in Section 3.5.1, its portfolio update formula [Vovk and Watkins 1998, Algorithm 1] for online portfolio selection is

$$\mathbf{b}_{t+1} = \frac{\int_{\Delta_m} \mathbf{b} \prod_{i=1}^{t-1} (\mathbf{b} \cdot \mathbf{x}_t)^\eta P_0(d\mathbf{b})}{\int_{\Delta_m} \prod_{i=1}^{t-1} (\mathbf{b} \cdot \mathbf{x}_t)^\eta P_0(d\mathbf{b})},$$

where $P_0(d\mathbf{b})$ denotes the prior weights of the experts. As a special case, Cover's UP corresponds to AA with uniform prior distribution and $\eta = 1$.

Singer [1997] proposed Switching Portfolios (SP) to track a changing market, in which the stock's behaviors may change frequently. Rather than the CRP class, SP decides a set of basic strategies—for example, the pure strategy that invests all wealth in one asset—and chooses a prior distribution over the set of strategies. Based on the actual return of each strategy and the prior distribution, SP is able to select a portfolio for each period. Upon this procedure, the author proposed two algorithms, both of which assume that the duration of using a basic strategy follows geometric distribution with a parameter of γ, which can be fixed or varied in time. With fixed γ, the first version of SP has an explicit update formula [Singer 1997, Eq. (6)],

$$\mathbf{b}_{t+1} = \left(1 - \gamma - \frac{\gamma}{m-1}\right)\mathbf{b}_t + \frac{\gamma}{m-1}.$$

With a varying γ, SP has no explicit update. The author also adopted the algorithm for transaction costs. Theoretically, the author further gave the lower bound of SP's logarithmic wealth with respect to any underlying switching regime in hindsight [Singer 1997, Theorem 2]. Empirical evaluation on Cover's two-stock pairs shows that SP can outperform UP, EG, and BCRP, in most cases.

Levina and Shafer [2008] proposed the Gaussian Random Walk (GRW) strategy, which switches among the base experts according to Gaussian distribution. Kozat and Singer [2007] extended SP to *piecewise* fixed fraction strategies, which partitions the periods into different segments and transits among these segments. The authors proved the piecewise universality of their algorithm, which can achieve the performance of the *optimal* piecewise fixed fraction strategy. Kozat and Singer [2008] extended Kozat and Singer [2007] to the case of transaction costs. Kozat and Singer [2009, 2010] further generalized Kozat and Singer [2007] to sequential decision problem. Kozat et al. [2008] proposed another piecewise universal portfolio selection strategy via context trees, and Kozat et al. [2011] generalized to sequential decision problem via tree weighting.

The most interesting thing is that switching portfolios adopts the notion of regime switching [Hamilton 1994, 2008], which is different from the underlying assumption of universal portfolio selection methods and seems to be more plausible than an i.i.d. market. The regime switching is also applied to some state-of-the-art trading strategies [Hardy 2001; Mlnařĺk et al. 2009]. However, this approach suffers from its distribution assumption, because geometric and Gaussian distributions do not seem to fit the market well [Mandelbrot 1963; Cont 2001]. This leads to other potential distributions that can better model the markets.

3.3. Follow-the-Loser Approaches

The underlying assumption for the optimality of BCRP strategy is that market is i.i.d., which however does not always hold for the real-world data and thus often results in inferior empirical performance, as observed in various previous works. Instead of tracking the winners, the Follow-the-Loser approach is often characterized by transferring the wealth from winners to losers. The underlying of this approach is *mean reversion* [Bondt and Thaler 1985; Poterba and Summers 1988; Lo and MacKinlay 1990], which means that the good (poor)-performing assets will perform poor (good) in the following periods.

To better understand the mean reversion principle, let us further analyze the behaviors of CRP in Example 3.1 [Li et al. 2012].

Example 3.2 (*Synthetic Market by Cover and Gluss [1986]*). As illustrated in Example 3.1, uniform CRP grows exponentially on the synthetic market. Now we

Table II. Example to Illustrate the Mean Reversion Trading Idea

# Period	Price Relative (A,B)	CRP	CRP Return	Portfolio Holdings	Notes
1	$(1, 2)$	$\left(\frac{1}{2}, \frac{1}{2}\right)$	$\frac{3}{2}$	$\left(\frac{1}{3}, \frac{2}{3}\right)$	$B \longrightarrow A$
2	$\left(1, \frac{1}{2}\right)$	$\left(\frac{1}{2}, \frac{1}{2}\right)$	$\frac{3}{4}$	$\left(\frac{2}{3}, \frac{1}{3}\right)$	$A \longrightarrow B$
3	$(1, 2)$	$\left(\frac{1}{2}, \frac{1}{2}\right)$	$\frac{3}{2}$	$\left(\frac{1}{3}, \frac{2}{3}\right)$	$B \longrightarrow A$
$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$

analyze its portfolio update behaviors, which follows mean reversion, as shown in Table II.

Suppose that the initial CRP portfolio is $(\frac{1}{2}, \frac{1}{2})$ and that at the end of the 1st period, the closing price adjusted portfolio holding becomes $(\frac{1}{3}, \frac{2}{3})$ and corresponding cumulative wealth increases by a factor of $\frac{3}{2}$. At the beginning of the 2nd period, the CRP manager rebalances the portfolio to initial uniform portfolio by transferring the wealth from good-performing stock (B) to poor-performing stock (A), which actually follows the mean reversion principle. Then, its cumulative wealth changes by a factor of $\frac{3}{4}$ and the portfolio holding at the end of the 2nd period becomes $(\frac{2}{3}, \frac{1}{3})$. At the beginning of the 3rd period, the wealth transfer with the mean reversion idea continues.

In summary, CRP implicitly assumes that if one stock performs poor (good), it tends to perform good (poor) in the subsequent period and thus transfers the weights from good-performing stocks to poor-performing stocks.

3.3.1. Anti-Correlation. Borodin et al. [2003, 2004] proposed a Follow-the-Loser portfolio strategy named Anti-Correlation (Anticor) strategy. Rather than no distributional assumption like Cover's UP, Anticor strategy assumes that market follows the mean reversion principle. To exploit the mean reversion property, it statistically makes a bet on the consistency of positive lagged cross-correlation and negative autocorrelation.

To obtain a portfolio for the $t + 1^{st}$ period, the Anticor algorithm adopts logarithmic price relatives [Hull 2008] in two specific market windows—that is, $\mathbf{y}_1 = \log(\mathbf{x}_{t-2w+1}^{t-w})$ and $\mathbf{y}_2 = \log(\mathbf{x}_{t-w+1}^{t})$. It then calculates the cross-correlation matrix between $\mathbf{y}_1$ and $\mathbf{y}_2$:

$$M_{cov}(i, j) = \frac{1}{w - 1} \left(\mathbf{y}_{1,i} - \bar{\mathbf{y}}_1\right)^{\top} \left(\mathbf{y}_{2,j} - \bar{\mathbf{y}}_2\right)$$

$$M_{cor}(i, j) = \begin{cases} \frac{M_{cov}(i,j)}{\sigma_1(i) * \sigma_2(j)} & \sigma_1(i), \sigma_2(j) \neq 0 \\ 0 & \text{otherwise} \end{cases}$$

Then, according to the cross-correlation matrix, Anticor algorithm transfers the wealth according to the mean reversion trading idea, or moves the proportions from the stocks increased more to the stocks increased less, and the corresponding amounts are adjusted according to the cross-correlation matrix. In particular, if asset i increases more than asset j and their sequences in the window are positively correlated, Anticor claims a transfer from asset i to j with the amount equals the cross correlation value ($M_{cor}(i, j)$) minus their negative autocorrelation values (min $\{0, M_{cor}(i, i)\}$ and min $\{0, M_{cor}(j, j)\}$). These transfer claims are finally normalized to keep the portfolio in the simplex domain.

Because of its mean reversion nature, it is difficult to obtain a useful bound such as the universal regret bound. Although heuristic and has no theoretical guarantee, Anticor empirically outperforms all other strategies at the time. On the other hand, although Anticor algorithm obtains good performance outperforming all algorithms at the time, its heuristic nature cannot fully exploit the mean reversion property. Thus, exploiting the property using systematic learning algorithms is highly desired.

3.3.2. Passive Aggressive Mean Reversion. Li et al. [2012] proposed Passive Aggressive Mean Reversion (PAMR) strategy, which exploits the mean reversion property with the Passive Aggressive (PA) online learning [Shalev-Shwartz et al. 2003; Crammer et al. 2006].

The main idea of PAMR is to design a loss function in order to reflect the mean reversion property—that is, if the expected return based on last price relative is larger than a threshold, the loss will linearly increase; otherwise, the loss is zero. In particular, the authors defined the ϵ-insensitive loss function for the tth period as

$$\ell_\epsilon\left(\mathbf{b};\mathbf{x}_t\right) = \begin{cases} 0 & \mathbf{b}\cdot\mathbf{x}_t \leq \epsilon \\ \mathbf{b}\cdot\mathbf{x}_t - \epsilon & \text{otherwise} \end{cases},$$

where $0 \leq \epsilon \leq 1$ is a sensitivity parameter to control the mean reversion threshold. Based on the loss function, PAMR passively maintains the last portfolio if the loss is zero; otherwise, it aggressively approaches a new portfolio that can force the loss zero. In summary, PAMR obtains the next portfolio via the following optimization problem:

$$\mathbf{b}_{t+1} = \underset{\mathbf{b}\in\Delta_m}{\arg\min}\, \frac{1}{2}\left\|\mathbf{b}-\mathbf{b}_t\right\|^2 \quad \text{s.t.} \quad \ell_\epsilon(\mathbf{b};\mathbf{x}_t) = 0. \tag{7}$$

Solving the optimization problem (7), PAMR has a clean closed form update scheme [Li et al. 2012, Proposition 1]:

$$\mathbf{b}_{t+1} = \mathbf{b}_t - \tau_t\left(\mathbf{x}_t - \bar{x}_t\mathbf{1}\right), \quad \tau_t = \max\left\{0, \frac{\mathbf{b}_t\cdot\mathbf{x}_t - \epsilon}{\|\mathbf{x}_t - \bar{x}_t\mathbf{1}\|^2}\right\}.$$

Since the authors ignored the nonnegativity constraint of the portfolio in the derivation, they also added a simplex projection step [Duchi et al. 2008]. The closed form update scheme clearly reflects the mean reversion trading idea by transferring the wealth from the good-performing stocks to the poor-performing stocks. It also coincides with the general form [Lo and MacKinlay 1990, Eq. (1)] of return-based contrarian strategies [Conrad and Kaul 1998; Lo 2008], except an adaptive multiplier τ_t. Besides the optimization problem (7), the authors also proposed two variants to avoid noise price relatives by introducing some nonnegative slack variables into optimization, which is similar to soft margin support vector machines.

Similar to the Anticor algorithm, due to PAMR's mean reversion nature, it is hard to obtain a meaningful theoretical regret bound. Nevertheless, PAMR achieves significant performance, beating all algorithms at the time, and shows its robustness along with the parameters. It also enjoys linear update time and runs extremely fast in the back tests, which show its practicability to large-scale real-world application.

The underlying idea is to exploit the single-period mean reversion, which is empirically verified by its evaluations on several real market datasets. However, PAMR suffers from drawbacks in risk management since it suffers significant performance degradation if the underlying single-period mean reversion fails to exist. Such drawback is clearly indicated by its performance in DJIA dataset [Borodin et al. 2003, 2004; Li et al. 2012].

3.3.3. Confidence Weighted Mean Reversion. Li et al. [2011b] proposed the Confidence Weighted Mean Reversion (CWMR) algorithm to further exploit the second-order portfolio information, which refers to the variance of portfolio weights (not price or price relative), following the mean reversion trading idea via Confidence Weighted (CW) online learning [Dredze et al. 2008; Crammer et al. 2008, 2009; Dredze et al. 2010].

The basic idea of CWMR is to model the portfolio vector as a multivariate Gaussian distribution with mean $\boldsymbol{\mu} \in \mathbb{R}^m$ and the diagonal covariance matrix $\boldsymbol{\Sigma} \in \mathbb{R}^{m\times m}$, which

has nonzero diagonal elements σ^2 and zero for off-diagonal elements. Whereas the mean represents the knowledge for the portfolio, the diagonal covariance matrix term stands for the confidence we have in the corresponding portfolio mean. Then, CWMR sequentially updates the mean and covariance matrix of the Gaussian distribution and draws portfolios from the distribution at the beginning of a period. In particular, the authors define $\mathbf{b}_t \in \mathcal{N}(\boldsymbol{\mu}_t, \boldsymbol{\Sigma}_t)$ and update the distribution parameters according to the similar idea of PA learning—that is, CWMR keeps the next distribution close to the last distribution in terms of Kullback-Leibler divergence if the probability of a portfolio return lower than ϵ is higher than a specified threshold. In summary, the optimization problem to be solved is

$$(\boldsymbol{\mu}_{t+1}, \boldsymbol{\Sigma}_{t+1}) = \underset{\boldsymbol{\mu} \in \Delta_m, \boldsymbol{\Sigma}}{\arg\min} \quad D_{KL}(\mathcal{N}(\boldsymbol{\mu}, \boldsymbol{\Sigma}) \| \mathcal{N}(\boldsymbol{\mu}_t, \boldsymbol{\Sigma}_t))$$
$$\text{s.t.} \quad \Pr[\boldsymbol{\mu} \cdot \mathbf{x}_t \leq \epsilon] \geq \theta.$$

To solve the optimization, Li et al. [2013] transformed the optimization problem using two techniques. One transformed optimization problem [Li et al. 2013, Eq. (3)] is

$$(\boldsymbol{\mu}_{t+1}, \boldsymbol{\Sigma}_{t+1}) = \arg\min \quad \frac{1}{2}\left(\log\left(\frac{\det\boldsymbol{\Sigma}_t}{\det\boldsymbol{\Sigma}}\right) + \text{Tr}\left(\boldsymbol{\Sigma}_t^{-1}\boldsymbol{\Sigma}\right) + (\boldsymbol{\mu}_t - \boldsymbol{\mu})^\top \boldsymbol{\Sigma}_t^{-1}(\boldsymbol{\mu}_t - \boldsymbol{\mu})\right)$$
$$\text{s. t.} \quad \epsilon - \boldsymbol{\mu}^\top\mathbf{x}_t \geq \phi\mathbf{x}_t^\top\boldsymbol{\Sigma}\mathbf{x}_t$$
$$\boldsymbol{\mu}^\top\mathbf{1} = 1, \quad \boldsymbol{\mu} \succeq 0.$$

Solving the presented optimization, one can obtain the closed form update scheme [Li et al. 2013, Proposition 4.1] as

$$\boldsymbol{\mu}_{t+1} = \boldsymbol{\mu}_t - \lambda_{t+1}\boldsymbol{\Sigma}_t(\mathbf{x}_t - \bar{x}_t\mathbf{1}), \quad \boldsymbol{\Sigma}_{t+1}^{-1} = \boldsymbol{\Sigma}_t^{-1} + 2\lambda_{t+1}\phi\mathbf{x}_t\mathbf{x}_t^\top,$$

where λ_{t+1} corresponds to the Lagrangian multiplier calculated by Equation (11) in Li et al. [2013] and $\bar{x}_t = \frac{\mathbf{1}^\top\boldsymbol{\Sigma}_t\mathbf{x}_t}{\mathbf{1}^\top\boldsymbol{\Sigma}_t\mathbf{1}}$ denotes the confidence weighted price relative average. Clearly, the update scheme reflects the mean reversion trading idea and can exploit both the first- and second-order information of a portfolio vector.

Similar to Anticor and PAMR, CWMR's mean reversion nature makes it hard to obtain a meaningful theoretical regret bound for the algorithm. Empirical performance shows that the algorithm can outperform the state of the art, including PAMR, which only exploits the first-order information of a portfolio vector. However, CWMR also exploits the single-period mean reversion, which suffers the same drawback as PAMR.

3.3.4. Online Moving Average Reversion. Observing that PAMR and CWMR implicitly assume *single-period* mean reversion, which causes one failure case on real dataset [Li et al. 2012, DJIA dataset], Li and Hoi [2012] defined a *multiple-period* mean reversion named Moving Average Reversion and proposed OnLine Moving Average Reversion (OLMAR) to exploit the multiple-period mean reversion.

The basic intuition of OLMAR is the observation that PAMR and CWMR implicitly predict next prices as last price—that is, $\hat{\mathbf{p}}_{t+1} = \mathbf{p}_{t-1}$, where $\mathbf{p}$ denotes the price vector corresponding to the related $\mathbf{x}$. Such extreme single-period prediction may cause some drawbacks that caused the failureof certain cases in Li et al. [2012]. Instead, the authors proposed a multiple-period mean reversion, which explicitly predicts the next price vector as the moving average within a window. They adopted simple moving average, which is calculated as $\text{MA}_t = \frac{1}{w}\sum_{i=t-w+1}^{t} \mathbf{p}_i$. Then, the corresponding next

price relative [Li and Hoi 2012, Eq. (1)] equals

$$\hat{\mathbf{x}}_{t+1}(w) = \frac{MA_t(w)}{\mathbf{p}_t} = \frac{1}{w}\left(1 + \frac{1}{\mathbf{x}_t} + \cdots + \frac{1}{\bigodot_{i=0}^{w-2}\mathbf{x}_{t-i}}\right), \tag{8}$$

where w is the window size and $\bigodot$ denotes elementwise product.

Then, they adopted Passive Aggressive online learning [Crammer et al. 2006] to learn a portfolio, which is similar to PAMR:

$$\mathbf{b}_{t+1} = \underset{\mathbf{b}\in\Delta_m}{\arg\min} \quad \frac{1}{2}\|\mathbf{b} - \mathbf{b}_t\|^2 \text{ s.t. } \mathbf{b} \cdot \hat{\mathbf{x}}_{t+1} \geq \epsilon.$$

Different from PAMR, its formulation follows the basic intuitive of investment—that is, to achieve a good performance based on the prediction. Solving the algorithm is similar to PAMR, and we ignore its solution. At the time, OLMAR achieves the best results among all existing algorithms [Li and Hoi 2012], especially on certain datasets that failed PAMR and CWMR.

3.3.5. Robust Median Reversion. As existing mean reversion algorithms do not consider noises and outliers in the data, they often suffer from estimation errors, which lead to nonoptimal portfolios and subsequent poor performance in practice. To handle the noises and outliers, Huang et al. [2013] proposed to exploit mean reversion via a robust L_1-median estimator and designed a novel portfolio selection strategy called Robust Median Reversion (RMR).

The basic idea of RMR is to explicitly estimate next price vector via a robust L_1-median estimator at the end of tth period—that is, $\hat{\mathbf{p}}_{t+1} = L_1 med_{t+1}(w) = \boldsymbol{\mu}_{t+1}$, where w is a window size and $\boldsymbol{\mu}$ is calculated by solving a optimization [Weber 1929, Fermat-Weber problem],

$$\boldsymbol{\mu}_{t+1} = \underset{\boldsymbol{\mu}}{\arg\min} \sum_{i=0}^{w-1} \|\mathbf{p}_{t-i} - \boldsymbol{\mu}\|.$$

In other words, L_1-median is the point with minimal sum of Euclidean distance to the k given price vectors. The solution to the optimization problem is unique [Weiszfeld 1937] if the data points are not collinear. Therefore, the expected price relative with the L_1-median estimator becomes

$$\hat{x}_{t+1}(w) = \frac{L_1 med_{t+1}(w)}{\mathbf{p}_t} = \frac{\boldsymbol{\mu}_{t+1}}{\mathbf{p}_t}. \tag{9}$$

Then RMR follows the similar portfolio optimization method as OLMAR [Li and Hoi 2012] to learn an optimal portfolio. Empirically, RMR outperforms the state-of-the-art on most datasets.

3.4. Pattern-Matching Based Approaches

Besides the two categories of Follow the Winner/Follow the Loser, another type of strategies may utilize both winners and losers, which is based on pattern matching. This category mainly covers nonparametric sequential investment strategies, which guarantee universal consistency (i.e., the corresponding trading rules are of growth optimal for any stationary and ergodic market process). Note that different from the optimality of BCRP for the i.i.d. market, which motivates the Follow-the-Winner approaches, Pattern-Matching–based approaches consider the non-i.i.d. market and maximize the conditional expectation of *log*-return given past observations (cf. Algoet and Cover [1988]). For the non-i.i.d. market, there is a big difference between the optimal

growth rate and the growth rate of BCRP. For example, for NYSE datasets during 1962–2006, the Average Annual Yield (AAY) of BCRP is about 20%, whereas the strategies in this category have AAY of more than 30% (cf. Györfi et al. [2012, Chapter 2]). Grounded on nonparametric prediction [Györfi and Schäfer 2003], this category consists of several pattern-matching–based investment strategies [Györfi et al. 2006, 2007, 2008; Li et al. 2011a]. Moreover, some techniques are also applied to the sequential prediction problem [Biau et al. 2010].

Now let us describe the main idea of the Pattern-Matching–based approaches [Györfi et al. 2006], which consists of two steps: the Sample Selection step and the Portfolio Optimization step.[1] The first step—Sample Selection—selects an index set C of similar historical price relatives, whose corresponding price relatives will be used to predict the next price relative. After locating the similarity set, each sample price relative $\mathbf{x}_i, i \in C$ is assigned with a probability $P_i, i \in C$. Existing methods often set the probabilities to uniform probability $P_i = \frac{1}{|C|}$, where $|\cdot|$ denotes the cardinality of a set. Besides uniform probability, it is possible to design a different probability setting. The second step—Portfolio Optimization—is to learn an optimal portfolio based on the similarity set obtained in the first step—that is,

$$\mathbf{b}_{t+1} = \underset{\mathbf{b} \in \Delta_m}{\arg\max} \quad U(\mathbf{b}; C),$$

where $U(\mathbf{b}; C)$ is a specified utility function of $\mathbf{b}$ based on C. One particular utility function is the log utility—for instance, $U(\mathbf{b}; C) = \sum_{i \in C} \log \mathbf{b}^\top \mathbf{x}_i$, which is usually the default utility. In case of an empty similarity set, a uniform portfolio is adopted as the optimal portfolio.

In the following sections, we concretize the Sample Selection step in Section 3.4.1 and the Portfolio Optimization step in Section 3.4.2. We further combine the two steps in order to formulate specific online portfolio selection algorithms in Section 3.4.3.

3.4.1. Sample Selection Techniques. The general idea in this step is to select similar samples from historical price relatives by comparing the preceding market windows of two price relatives. Suppose that we are going to locate the price relatives that are similar to next price relative $\mathbf{x}_{t+1}$. The basic routine is to iterate all historic price relative vectors $\mathbf{x}_i, i = w + 1, \ldots, t$ and count $\mathbf{x}_i$ as similar one, if the preceding market window $\mathbf{x}_{i-w}^{i-1}$ is similar to the latest market window $\mathbf{x}_{t-w+1}^t$. The set C is maintained to contain the indexes of similar price relatives. Note that the market window is a $w \times m$-matrix, and the similarity between two market windows is often calculated on the concatenated $w \times m$-vector. The Sample Selection procedure ($C(\mathbf{x}_1^t, w)$) is further illustrated in Algorithm 2.

Nonparametric *histogram-based* sample selection [Györfi and Schäfer 2003] predefines a set of discretized partitions, partitions both the latest market window $\mathbf{x}_{t-w+1}^t$ and the historical market window $\mathbf{x}_{i-w}^{i-1}, i = w + 1, \ldots, t$, and finally chooses price relative vectors whose $\mathbf{x}_{i-w}^{i-1}$ is in the same partition as $\mathbf{x}_{t-w+1}^t$. In particular, given a partition $P_\ell = A_{j,\ell}, j = 1, 2, \ldots, d_\ell$ of $\mathbb{R}_+^m$ into d_ℓ disjoint sets and a corresponding discretization function $G_\ell(\mathbf{x}) = j$, if $\mathbf{x} \in A_{\ell,j}$, we can define the similarity set as

$$C_H(\mathbf{x}_1^t, w) = \left\{ w < i < t + 1 : G_\ell(\mathbf{x}_{t-w+1}^t) = G_\ell(\mathbf{x}_{i-w}^{i-1}) \right\}.$$

Note that ℓ is adopted to aggregate multiple experts.

---

[1]Here we only introduce the key idea. All algorithms in this category consist of an additional aggregation step, which is a special case of MLAs in Section 3.5.

ALGORITHM 2: Sample selection framework ($C(\mathbf{x}_1^t, w)$).

Input: $\mathbf{x}_1^t$: Historical market sequence; w: window size;
Output: C: Index set of similar price relatives.

Initialize $C = \emptyset$;
if $t \leq w + 1$ **then**
  | return;
end
for $i = w + 1, w + 2, \ldots, t$ **do**
  | **if** $\mathbf{x}_{i-w}^{i-1}$ *is similar to* $\mathbf{x}_{t-w+1}^t$ **then**
  | | $C = C \cup \{i\}$;
  | **end**
end

Nonparametric *kernel-based* sample selection [Györfi et al. 2006] identifies the similarity set by comparing two market windows via Euclidean distance:

$$C_K\big(\mathbf{x}_1^t, w\big) = \Big\{ w < i < t + 1 : \|\mathbf{x}_{t-w+1}^t - \mathbf{x}_{i-w}^{i-1}\| \leq \frac{c}{\ell} \Big\},$$

where c and ℓ are the thresholds used to control the number of similar samples. Note that the authors adopted two threshold parameters for theoretical analysis.

Nonparametric *nearest neighbor–based* sample selection [Györfi et al. 2008] searches the price relatives whose preceding market windows are within the ℓ nearest neighbor of latest market window in terms of Euclidean distance:

$$C_N\big(\mathbf{x}_1^t, w\big) = \big\{ w < i < t + 1 : \mathbf{x}_{i-w}^{i-1} \text{ is among the } \ell \text{ NNs of } \mathbf{x}_{t-w+1}^t \big\},$$

where ℓ is a threshold parameter.

Correlation-driven nonparametric sample selection [Li et al. 2011a] identifies the linear similarity among two market windows via correlation coefficient:

$$C_C\big(\mathbf{x}_1^t, w\big) = \left\{ w < i < t + 1 : \frac{cov\big(\mathbf{x}_{i-w}^{i-1}, \mathbf{x}_{t-w+1}^t\big)}{std\big(\mathbf{x}_{i-w}^{i-1}\big) std\big(\mathbf{x}_{t-w+1}^t\big)} \geq \rho \right\},$$

where ρ is a predefined correlation coefficient threshold.

3.4.2. Portfolio Optimization Techniques. The second step of the Pattern-Matching–based approaches is to construct an optimal portfolio based on the similar set C. Two main approaches are the Kelly's CGT and Markowitz's mean-variance theory. Next, we illustrate several techniques adopted in these approaches.

Györfi et al. [2006] proposed to figure out a *log-optimal* (Kelly) portfolio based on similar price relatives located in the first step, which is clearly following the CGT. Given a similarity set, the log-optimal utility function is defined as

$$U_L\big(\mathbf{b}; C\big(\mathbf{x}_1^t\big)\big) = \mathbb{E}\big\{ \log \mathbf{b} \cdot \mathbf{x} | \mathbf{x}_i, i \in C\big(\mathbf{x}_1^t\big) \big\} = \sum_{i \in C(\mathbf{x}_1^t)} P_i \log \mathbf{b} \cdot \mathbf{x}_i,$$

where P_i denotes the probability assigned to a similar price relative $\mathbf{x}_i, i \in C(\mathbf{x}_1^t)$. Györfi et al. [2006] assume a uniform probability among the similar samples, thus it is equivalent to the following utility function:

$$U_L\big(\mathbf{b}; C\big(\mathbf{x}_1^t\big)\big) = \sum_{i \in C(\mathbf{x}_1^t)} \log \mathbf{b} \cdot \mathbf{x}_i.$$

Györfi et al. [2007] introduced *semi-log-optimal* strategy, which approximates log in the log-optimal utility function aiming to release the computational issue, and Vajda [2006] presented theoretical analysis and proved its universal consistency. The semi-log-optimal utility function is defined as

$$U_S\big(\mathbf{b}; C(\mathbf{x}_1^t)\big) = \mathbb{E}\big\{ f(\mathbf{b} \cdot \mathbf{x}) \big| \mathbf{x}_i, i \in C(\mathbf{x}_1^t)\big\} = \sum_{i \in C(\mathbf{x}_1^t)} P_i \, f(\mathbf{b} \cdot \mathbf{x}_i),$$

where $f(\cdot)$ is defined as the second-order Taylor expansion of $\log z$ with respect to $z = 1$—that is, $f(z) = z - 1 - \frac{1}{2}(z-1)^2$. Györfi et al. [2007] assume a uniform probability among the similar samples, thus, equivalently,

$$U_S\big(\mathbf{b}; C(\mathbf{x}_1^t)\big) = \sum_{i \in C(\mathbf{x}_1^t)} f\left(\mathbf{b} \cdot \mathbf{x}_i\right).$$

Ottucsák and Vajda [2007] proposed the nonparametric *Markowitz-type* strategy, which is a further generalization of the semi-log-optimal strategy. The basic idea of the Markowitz-type strategy is to represent portfolio return using Markowitz's idea to trade off between portfolio mean and variance. To be specific, the Markowitz-type utility function is defined as

$$
\begin{aligned}
U_M\big(\mathbf{b}; C(\mathbf{x}_1^t)\big) &= \mathbb{E}\left\{\mathbf{b} \cdot \mathbf{x} \middle| \mathbf{x}_i, i \in C(\mathbf{x}_1^t)\right\} - \lambda \mathrm{Var}\left\{\mathbf{b} \cdot \mathbf{x} \middle| \mathbf{x}_i, i \in C(\mathbf{x}_1^t)\right\} \\
&= \mathbb{E}\left\{\mathbf{b} \cdot \mathbf{x} \middle| \mathbf{x}_i, i \in C(\mathbf{x}_1^t)\right\} - \lambda \mathbb{E}\left\{(\mathbf{b} \cdot \mathbf{x})^2 \middle| \mathbf{x}_i, i \in C(\mathbf{x}_1^t)\right\} \\
&\quad + \lambda\left(\mathbb{E}\left\{\mathbf{b} \cdot \mathbf{x} \middle| \mathbf{x}_i, i \in C(\mathbf{x}_1^t)\right\}\right)^2,
\end{aligned}
$$

where λ is a trade-off parameter. In particular, a simple numerical transformation shows that semi-log-optimal portfolio is an instance of the *log-optimal* utility function with a specified λ.

To solve the problem with transaction costs, Györfi and Vajda [2008] propose a *GV-type* utility function (Algorithm 2 in Györfi and Vajda [2008]; their Algorithm 1 follows the same procedure as log-optimal utility) by incorporating the transaction costs, as follows:

$$U_T\left(\mathbf{b}; C\left(\mathbf{x}_1^t\right)\right) = \mathbb{E}\left\{\log \mathbf{b} \cdot \mathbf{x} + \log c\left(\mathbf{b}_t, \mathbf{b}, \mathbf{x}_t\right)\right\},$$

where $c(\cdot) \in (0, 1)$ is the transaction cost factor in Equation (1), which represents the remaining proportion after transaction costs imposed by the market. The details of the calculation of the factor are illustrated in Section 2.1. According to a uniform probability assumption of the similarity set, it is equivalent to calculate

$$U_T\left(\mathbf{b}; C\left(\mathbf{x}_1^t\right)\right) = \sum_{i \in C(\mathbf{x}_1^t)} \left(\log \mathbf{b} \cdot \mathbf{x} + \log c\left(\mathbf{b}_t, \mathbf{b}, \mathbf{x}_t\right)\right).$$

In any of the procedures presented, if the similarity set is nonempty, we can gain an optimal portfolio based on the similar price relatives and their probability. In case of an empty set, we can choose either uniform portfolio or previous portfolio.

3.4.3. Combinations. In this section, let us combine the first and second steps and describe the detail algorithms in the Pattern-Matching–based approaches. Table III shows existing combinations, where "—" means that no algorithm is proposed to exploit the combination.

One default utility function is the log-optimal function. Györfi and Schäfer [2003] introduced the *nonparametric histogram-based log-optimal* investment strategy (B^{H}), which combines the histogram-based sample selection and log-optimal utility function

Table III. Pattern-Matching–Based Approaches: Sample Selection and Portfolio Optimization

Portfolio Optimization	Histogram	Sample Selection Techniques		
		Kernel	Nearest Neighbor	Correlation
Log-optimal	B^H: $C_H + U_L$	B^K: $C_K + U_L$	B^{NN}: $C_N + U_L$	CORN: $C_C + U_L$
Semi-log-optimal	—	B^S: $C_K + U_S$	—	—
Markowitz-type	—	B^M: $C_K + U_M$	—	—
GV-type	—	B^{GV}: $C_K + U_R$	—	—

and proved its universal consistency. Györfi et al. [2006] presented the *nonparametric kernel-based log-optimal* investment strategy (B^K), which combines the kernel-based sample selection and log-optimal utility function and proved its universal consistency. Györfi et al. [2008] proposed the *nonparametric nearest neighbor log-optimal* investment strategy (B^{NN}), which combines the nearest neighbor sample selection and log-optimal utility function and proved its universal consistency. Li et al. [2011a] created the *correlation-driven nonparametric* learning approach (CORN) by combining the correlation-driven sample selection and log-optimal utility function and showed its superior empirical performance over the previous three combinations. Besides the log-optimal utility function, several algorithms using different utility functions have been proposed. Györfi et al. [2007] proposed the *nonparametric kernel-based semi-log-optimal* investment strategy (B^S) by combining the kernel-based sample selection and semi-log-optimal utility function to ease the computation of (B^K). Ottucsák and Vajda [2007] proposed the *nonparametric kernel-based Markowitz-type* investment strategy (B^M) by combining the kernel-based sample selection and Markowitz-type utility function to make trade-offs between the return (mean) and risk (variance) of expected portfolio return. Györfi and Vajda [2008] proposed the *nonparametric kernel-based GV-type* investment strategy (B^{GV}) by combining the kernel-based sample selection and GV-type utility function to construct portfolios in case of transaction costs. If the sequence of relative price vectors is a first-order Markov chain with known distributions, then their strategies are growth optimal. For unknown distributions, Györfi and Walk [2012] introduced empirical growth optimal algorithms. Ormos and Urbán [2011] empirically analyzed the performance of log-optimal portfolio strategies with transaction costs.

Note that this section only introduces the key steps (or individual expert) in the Pattern-Matching–based approaches, whereas all algorithms presented previously also consist of an additional aggregation step. With different parameters (w, ℓ, or ρ), one can get a family of portfolios, which are then aggregated into a final portfolio using exponential weighting [Györfi et al. 2006]. Such rule is actually a meta-algorithm, which we will introduce in the following section.

3.5. Meta-Learning Algorithms

Another category of research in the area of online portfolio selection is the MLA [Das and Banerjee 2011], which is closely related to expert learning [Cesa-Bianchi and Lugosi 2006] in the machine learning community. This is directly applicable to an FOF, which delegates its portfolio assets to other funds. In general, MLA assumes several base experts, either from the same strategy class or different classes. Each expert outputs a portfolio vector for the coming period, and MLA combines these portfolios to form a final portfolio, which is used for the next rebalancing. MLAs are similar to algorithms in Follow-the-Winner approaches; however, they are proposed to handle a broader class of experts, which CRP can serve as one special case. On the one hand, the MLA system can be used to smooth the final performance with respect to all underlying experts, especially when base experts are sensitive to certain environments/parameters. On the

other hand, combining a universal strategy and a heuristic algorithm, where it is not easy to obtain a theoretical bound, such as Anticor, can provide the universal property to the whole MLA system. Finally, MLA is able to combine all existing algorithms, thus providing a much broader area of application.

3.5.1. Aggregating Algorithms. Besides the algorithms discussed in Section 3.2.5, the AA [Vovk 1990; Vovk and Watkins 1998] can also be generalized to include more sophisticated base experts. Given a learning rate $\eta > 0$, a measurable set of experts A, and a prior distribution P_0 that assigns the initial weights to the experts, AA defines a loss function as $\ell(\mathbf{x}, \gamma)$ and $\gamma_t(\theta)$ as the action chosen by expert θ at time t. At the beginning of each period $t = 1, 2, \ldots$, AA updates the experts' weights as

$$P_{t+1}(A) = \int_A \beta^{\ell(\mathbf{x}_t, \gamma_t(\theta))} P_t(d\theta),$$

where $\beta = e^{-\eta}$ and P_t denotes the weights to the experts at time t.

3.5.2. Fast Universalization. Akcoglu et al. [2002, 2004] proposed Fast Universalization (FU), which extends Cover's UP [Cover 1991] from a parameterized CRP class to a wide class of investment strategies, including trading strategies operating on a single stock and portfolio strategies allocating wealth among the whole stock market. FU's basic idea is to evenly split the wealth among a set of base experts, let these experts operate on their own, and finally pool their wealth. FU's update is similar to that of Cover's UP, and it also asymptotically achieves the wealth equal to an optimal fixed convex combination of base experts. In cases in which all experts are CRPs, FU is reduced to Cover's UP.

Besides the universalization in the continuous parameter space, various discrete BAH combinations have been adopted by various existing algorithms. Rewritten in discrete form, its update can be straightforwardly obtained. For example, Borodin et al. [2003, 2004] adopted BAH strategy to combine Anticor experts with a finite number of window sizes. Li et al. [2012] combined PAMR experts with a finite number of mean reversion thresholds. Moreover, all Pattern-Matching–based approaches in Section 3.4 used BAH to combine their underlying experts, also with a finite number of window sizes.

3.5.3. Online Gradient and Newton Updates. Das and Banerjee [2011] proposed two meta-optimization algorithms, OGU and ONU, which are natural extensions of EG and ONS, respectively. Since their updates and proofs are similar to their precedents, here we ignore their updates. Theoretically, OGU and ONU can achieve the growth rate as the optimal convex combination of the underlying experts. Particularly, if any base expert is universal, the final meta-system enjoys the universal property. Such a property is useful because an MLA can incorporate a heuristic algorithm and a universal algorithm, whereby the final system enjoys the performance while keeping the universal property.

3.5.4. Follow the Leading History. Hazan and Seshadhri [2009] proposed a Follow the Leading History (FLH) algorithm for changing environments. FLH can incorporate various universal base experts, such as the ONS algorithm. Its basic idea is to maintain a working set of finite experts, which are dynamically flowed in and dropped out according to their performance, and allocate the weights among the active working experts with an MLA—for example, the Herbster-Warmuth algorithm [Herbster and Warmuth 1998]. Different from other MLAs where experts operate from the same beginning, FLH adopts experts starting from different periods. Theoretically, the FLH algorithm with universal methods is universal. Empirically, FLH equipped with ONS can significantly outperform ONS.

4. CONNECTION WITH CAPITAL GROWTH THEORY

Most online portfolio selection algorithms introduced earlier can be interestingly connected to the CGT. In this section, we first introduce the CGT for portfolio selection and then connect the previous algorithms to the CGT in order to reveal their underlying trading principles.

4.1. Capital Growth Theory for Portfolio Selection

Originally introduced in the context of gambling, CGT [Hakansson and Ziemba 1995] (also referred to as Kelly investment [Kelly 1956] or Growth Optimal Portfolio (GOP) [Györfi et al. 2012]) can generally be adopted for online portfolio selection. Breiman [1961] generalized Kelly criterion to multiple investment opportunities. Thorp [1971] and Hakansson [1971] focused on the theory of Kelly criterion or logarithmic utility for the portfolio selection problem. Now let us briefly introduce the theory for portfolio selection [Thorp 1971].

The basic procedure of CGT for portfolio selection is to maximize the expected log return for each period. It involves two related steps: prediction and portfolio optimization. For the prediction step, CGT receives the predicted distribution of price relative combinations $\hat{\mathbf{x}}_{t+1} = (\hat{x}_{t+1,1}, \ldots, \hat{x}_{t+1,m})$, which can be obtained as follows. For each investment i, one can predict a finite number of distinct values and corresponding probabilities. Let the range of $\hat{x}_{t+1,i}$ be $\{r_{i,1}; \ldots; r_{i,N_i}\}, i = 1, \ldots, m$, and the corresponding probability for each possible value $r_{i,j}$ be $p_{i,j}$. Based on these predictions, one can estimate their joint vectors and corresponding joint probabilities. In this way, there are in total $\prod_{i=1}^{m} N_i$ possible prediction combinations, each of which is in the form of $\hat{\mathbf{x}}_{t+1}^{(k_1,k_2,\ldots,k_m)} = [\hat{x}_{t+1,1} = r_{1,k_1}$ and $\hat{x}_{t+1,2} = r_{2,k_2}$ and $\ldots$ and $\hat{x}_{t+1,m} = r_{m,k_m}]$ with a probability of $p^{(k_1,k_2,\ldots,k_m)} = \prod_{j=1}^{m} p_{j,k_j}$. Given these predictions, CGT tries to obtain an optimal portfolio that maximizes the expected log return,

$$\mathbb{E} \log S = \sum p^{(k_1,k_2,\ldots,k_m)} \log \left(\mathbf{b} \cdot \hat{\mathbf{x}}_{t+1}^{(k_1,k_2,\ldots,k_m)} \right)$$
$$= \sum \left[p^{(k_1,k_2,\ldots,k_m)} \log \left(b_1 r_{1,k_1} + \cdots + b_m r_{m,k_m} \right) \right],$$

where the summation is over all $\prod_{i=1}^{m} N_i$ price relative combinations. Obviously, maximizing the equation is concave in $\mathbf{b}$ and thus can be efficiently solved via convex optimization [Boyd and Vandenberghe 2004].

4.2. Online Portfolio Selection and Capital Growth Theory

Most existing online portfolio selection algorithms have close connection with the CGT (or Kelly criterion). Although the theory provides a theoretically guaranteed framework for asset allocation, online portfolio selection algorithms mainly connect to the theory from two different aspects.

The first connection is established between CGT and the universal portfolio selection scheme, which mainly includes several Follow-the-Winner approaches. Kelly criterion aims to maximize the exponential growth rate of an investment scheme, whereas the universal portfolio selection scheme tries to maximize the exponential growth rate relative to the rate of BCRP. Although they have different objectives, they are somehow connected as the target of universal portfolio selection scheme is one special case of Kelly criterion. Cover and Thomas [1991, Theorem 15.3.1] showed that if the market sequence (price relative vectors) is i.i.d., then the maximum performance is achieved by an optimal CRP in hindsight, or the BCRP. We further rewrite the BCRP strategy in the form of Kelly criterion, as shown in the first row of Table IV. There after, Cover [1991] set BCRP as a target and proposed the universal portfolio selection scheme. Note that as introduced in Section 3.1.3, the gap between their cumulative exponential

Table IV. Online Portfolio Selection and the Capital Growth Theory

Algorithms	$\hat{\mathbf{x}}_{t+1}$	Prob.	Capital Growth Forms
BCRP*	$\mathbf{x}_i, i = 1, \ldots, n$	$1/n$	$\mathbf{b}_{t+1} = \arg\max_{\mathbf{b} \in \Delta_m} \frac{1}{n} \sum_{i=1}^{n} \log \mathbf{b} \cdot \mathbf{x}_i$
EG	$\mathbf{x}_t$	100%	$\mathbf{b}_{t+1} = \arg\max_{\mathbf{b} \in \Delta_m} \log \mathbf{b} \cdot \mathbf{x}_t - \lambda R(\mathbf{b}, \mathbf{b}_t)$
PAMR	$\frac{1}{\mathbf{x}_t}$	100%	$\mathbf{b}_{t+1} = \arg\min_{\mathbf{b} \in \Delta_m} \mathbf{b} \cdot \mathbf{x}_t + \lambda R(\mathbf{b}, \mathbf{b}_t)$
CWMR	$\frac{1}{\mathbf{x}_t}$	100%	$\mathbf{b}_{t+1} = \arg\min_{\mathbf{b} \in \Delta_m} P(\mathbf{b} \cdot \mathbf{x}_t) + \lambda R(\mathbf{b}, \mathbf{b}_t)$
OLMAR/RMR	Eq. (8)/Eq. (9)	100%	$\mathbf{b}_{t+1} = \arg\max_{\mathbf{b} \in \Delta_m} \mathbf{b} \cdot \hat{\mathbf{x}}_{t+1} - \lambda R(\mathbf{b}, \mathbf{b}_t)$
$B^H/B^K/B^{NN}$/CORN	$\mathbf{x}_i, i \in C_t$	$1/|C_t|$	$\mathbf{b}_{t+1} = \arg\max_{\mathbf{b} \in \Delta_m} \frac{1}{|C_t|} \sum_{i \in C_t} \log \mathbf{b} \cdot \mathbf{x}_i$
B^{GV}	$\mathbf{x}_i, i \in C_t$	$1/|C_t|$	$\mathbf{b}_{t+1} = \arg\max_{\mathbf{b} \in \Delta_m} \frac{1}{|C_t|} \sum_{i \in C_t} (\log \mathbf{b} \cdot \mathbf{x}_i + \log c(\cdot))$
FTL	$\mathbf{x}_i, i = 1, \ldots, t$	$1/t$	$\mathbf{b}_{t+1} = \arg\max_{\mathbf{b} \in \Delta_m} \frac{1}{t} \sum_{i=1}^{t} \log \mathbf{b} \cdot \mathbf{x}_i$
FTRL	$\mathbf{x}_i, i = 1, \ldots, t$	$1/t$	$\mathbf{b}_{t+1} = \arg\max_{\mathbf{b} \in \Delta_m} \frac{1}{t} \sum_{i=1}^{t} \log \mathbf{b} \cdot \mathbf{x}_i - \lambda R(\mathbf{b})$

growth rates is termed *regret*. Such connection also coincides with competitive analysis in Borodin et al. [2000].

In particular, the first four algorithms in the Follow-the-Winner category (i.e., Universal Portfolios, Exponential Gradient, Follow the Leader, and Follow the Regularized Leader) all release regret bounds whose daily average asymptotically approaches zero as the trading period goes to infinity. In other words, these algorithms can achieve the same exponential growth rate as BCRP, which is CGT optimal in an i.i.d. market. Although the Aggregating-type Algorithms extend online portfolio selection from the CRP class to other strategy classes, they may not be optimal relative to BCRP.

The second connection explicitly adopts the idea of CGT for online portfolio selection, as shown in Section 4.1. For each period, one algorithm requires the predicted price relative combinations and their corresponding probabilities. Without loss of generality, let us make a portfolio decision for the $t + 1^{st}$ period. Table IV summarizes their rewritten formulations. Note that some algorithms in the first connection (EG, ONS, etc.) can also be rewritten to this form, although their objectives are different from CGT. We present their implicit market distributions, denoted by their values ($\hat{\mathbf{x}}_{t+1}$) and probabilities (Prob.), in the second and third columns, respectively. We then rewrite all algorithms following the CGT—that is, to maximize the expected log return for the $t+1^{st}$ period—in the fourth column. The regularization terms are denoted as $R(\mathbf{b}, \mathbf{b}_t)$, which preserves the information of last portfolio vector ($\mathbf{b}_t$), and $R(\mathbf{b})$, which constrains the variability of a portfolio vector. Based on the number of predictions, we can categorize most existing algorithms into three categories.

The first category, including EG/PAMR/CWMR/OLMAR/RMR, implicitly or explicitly predicts a single scenario with certainty and tries to select an optimal portfolio. Note that the capital growth forms of PAMR and CWMR are rewritten from their original forms while keeping their essential ideas. Moreover, PAMR, CWMR, and OLMAR all ignore the log utility function, as adding the log utility function follows the same idea but causes the convexity issue. Although such single prediction is risky, all of these algorithms adopt regularization terms, such as $R(\mathbf{b}, \mathbf{b}_t) = \|\mathbf{b} - \mathbf{b}_t\|^2$, to ensure that next portfolio is not far from current one, which indeed reduces the risk.

The second category, including Pattern-Matching–based approaches, predicts multiple scenarios that are deemed similar to the next price relative vector. In particular, it expects the next price relative to be $\mathbf{x}_i, i \in C$ with a uniform probability of $\frac{1}{|C|}$, where C denotes the similarity set. Then, algorithms in this category try to maximize its expected log return in terms of the similarity set, which is consistent with the CGT and results in an optimal fixed fraction portfolio. Note that several algorithms in the Pattern-Matching–based approaches, including B^S, B^M, and B^{GV}, adopt different portfolio optimization approaches, which we do not count in here.

Table V. Online Portfolio Selection and Their Underlying Trading Principles

Principles	Algorithms	$\hat{\mathbf{x}}_{t+1}$	Prob.	$\mathbb{E}\left\{\hat{\mathbf{x}}_{t+1}\right\}$
Momentum	EG	$\mathbf{x}_t$	100%	$\mathbf{x}_t$
	FTL/FTRL	$\mathbf{x}_i, i = 1, \ldots, t$	$1/t$	$\frac{1}{t}\sum_{i=1}^{t}\mathbf{x}_i$
Mean reversion	CRP/UP/AA	n/a	n/a	n/a
	Anticor	n/a	n/a	n/a
	PAMR/CWMR	$\frac{1}{\mathbf{x}_t}$	100%	$\frac{1}{\mathbf{x}_t}$
	OLMAR	Eq. (8)	100%	Eq. (8)
	RMR	Eq. (9)	100%	Eq. (9)
Mixed	$B^H/B^K/B^{NN}/B^{GV}/CORN$	$\mathbf{x}_i, i \in C_t$	$1/|C_t|$	$\frac{1}{|C_t|}\sum_{i \in C_t}\mathbf{x}_i$

The third category, including FTL and FTRL, implicitly predicts the next scenario as all historical price relatives. In particular, it predicts that the next price relative vector equals $\mathbf{x}_i, i = 1, \ldots, t$ with a uniform probability of $\frac{1}{t}$. Based on such prediction, strategies in this category aim to maximize the expected log return and subtract a regularization term for FTRL. Note that different from the regularization terms in the first category, the regularization term in this category, such as $R(\mathbf{b}) = \|\mathbf{b}\|^2$, only controls the deviation of next portfolio. This is due to the fact that the predictions already contain all available information.

Note that the two connections are not exclusive. For example, some universal portfolio selection algorithms (EG, ONS, etc.) show both connections. On the one hand, their formulations can be explicitly rewritten to Kelly's form. On the other hand, their motivations follow the first connection, which is validated by their theoretical results.

4.3. Underlying Trading Principles

Besides the aspect of the CGT, most existing algorithms also follow certain trading ideas to implicitly or explicitly predict their next price relatives. Table V summarizes their underlying trading ideas via three trading principles: momentum, mean reversion, and others (e.g., nonparametric prediction).

Momentum strategy [Chan et al. 1996; Rouwenhorst 1998; Moskowitz and Grinblatt 1999; Lee and Swaminathan 2000; George and Hwang 2004; Cooper et al. 2004] assumes winners (losers) will still be winners (losers) in the following period. By observing algorithms' underlying prediction schemes, we can classify EG/FTL/FTRL as this category. Although EG assumes that next price relative vector will be the same as last one, FTL and FTRL assume that next price relative is expected to be the average of all historical price relative vectors.

In contrast, a mean reversion strategy [Bondt and Thaler 1985, 1987; Poterba and Summers 1988; Jegadeesh 1991; Chaudhuri and Wu 2003] assumes that winners (losers) will be losers (winners) in the following period. Clearly, CRP and UP, Anticor, and PAMR/CWMR belong to this category. Here, note that UP is an expert combination of the CRP strategies, and we classify it by its implicit assumption on the underlying stocks. If we observe from the perspective of experts, UP transfers wealth from CRP losers to CRP winners, which is actually momentum. Moreover, PAMR and CWMR's expected price relative vector is implicitly the inverse of last vector, which is in the opposite of EG.

Other trading ideas, including the Pattern-Matching–based approaches, cannot be classified as the two categories mentioned previously. For example, for the Pattern-Matching approaches, their average of the price relatives in a similarity set may be regarded as either momentum or mean reversion. Besides, the classification of AA depends on the type of underlying experts. From an experts' perspective, AA always transfers the wealth from loser experts to winner experts, which is momentum strategy. From stocks' perspective, which is the assumption in Table V, the classification of

AA coincides with that of its underlying experts. That is, if the underlying experts are single-stock strategy, which is momentum, then we view AA's trading idea as momentum. On the other hand, if the underlying experts are CRP strategy, which follows the mean reversion principle, we regard AA's trading idea as mean reversion.

5. CHALLENGES AND FUTURE DIRECTIONS

The online portfolio selection task is a special and important case of asset management. Although existing algorithms perform well theoretically or empirically in back tests, researchers have encountered several challenges in designing the algorithms. In this section, we focus on the two consecutive steps in the online portfolio selection: prediction and portfolio optimization. In particular, we illustrate open challenges in the prediction step in Section 5.1 and list several other challenges in the portfolio optimization step in Section 5.2. There are a lot of opportunities in this area, and it warrants further exploration.

5.1. Accurate Prediction by Advanced Techniques

As we analyzed in Section 4.2, existing algorithms implicitly assume various prediction schemes. Although current assumptions can result in good performance, they are far from perfect. Thus, the challenges for the prediction step are often related to the design of more subtle prediction schemes in order to produce more accurate predictions of the price relative distribution.

—*Searching patterns*. In the Pattern-Matching–based approaches, despite the many sample selection techniques introduced, efficiently recognizing patterns in the financial markets is often challenging. Moreover, existing algorithms always assume uniform probability on the similar samples, whereas it is a challenge to assign appropriate probability, hoping to predict more accurately. Finally, existing algorithms only consider the similarity between two market windows with the same length and same interval; however, locating patterns with varying timing [Rabiner and Levinson 1981; Sakoe and Chiba 1990; Keogh 2002] is also potentially attractive.
—*Utilizing stylized facts of returns*. In econometrics, there exist a lot of *stylized facts*, which refer to consistent empirical findings that are accepted as truth. One stylized fact is related to autocorrelations in various assets' returns.[2] It is often observed that some stocks/indices show positive daily/weekly/monthly autocorrelations [Fama 1970; Lo and MacKinlay 1999; Llorente et al. 2002], whereas some others have negative daily autocorrelations [Lo and MacKinlay 1988, 1990; Jegadeesh 1990]. An open challenge is to predict future price relatives utilizing their autocorrelations.
—*Utilizing stylized facts in absolute/square returns*. Another stylized fact [Taylor 2005] is that the autocorrelations in absolute and squared returns are much higher than those exhibited by simple returns. This fact indicates that there may be a consistent nonlinear relationship within the time series, which may be exploited by various machine learning techniques to boost the prediction accuracy. However, in the current prediction step, such information is rarely exploited, thus constituting a challenge.
—*Utilizing calendar effects*. It is well known that there exist some *calendar effects*, such as January effect or turn-of-the-year effect [Rozeff and Kinney Jr. 1976; Haugen and Lakonishok 1987; Moller and Zilca 2008] and holiday effect [Fields 1934; Brockman and Michayluk 1998; Dzhabarov and Ziemba 2010]. No existing algorithm exploits such information, which can potentially provide better predictions. Thus, another open challenge is to take advantage of these calendar effects in the prediction step.

---

[2]The econometrics community often uses the net return, which equals price relative minus one.

—*Exploiting additional information*. Although most existing prediction schemes focus solely on the price relative (or price), there exists other useful side information, such as volume, fundamental, and experts' opinions. Cover and Ordentlich [1996] presented a preliminary model to incorporate the information, which is, however, far from applicability. Thus, it is an open challenge to incorporate other sources of information in order to facilitate the prediction of next price relatives.

5.2. Open Issues of Portfolio Optimization

Portfolio optimization is the subsequent step for online portfolio selection. Although the CGT is effective in maximizing final cumulative wealth, which is the aim of our task, it often incurs high risk [Thorp 1997], which is sometimes unacceptable for an investor. Incorporating the risk concern within online portfolio selection is another open issue, which is currently not taken into account.

—*Incorporating risk*. Mean-variance theory [Markowitz 1952] adopts variance as a proxy to risk. However, simply adding variance may not efficiently trade off between return and risk. Thus, one challenge is to exploit an effective proxy to risk and efficiently trade off between them in the scenario of online portfolio selection.

—*Utilizing* "optimal *f*". One recent advancement in money management is "optimal *f*" [Vince 1990, 1992, 1995, 2007, 2009], which is proposed to handle the drawbacks of Kelly's theory. Optimal *f* can reduce the risk of Kelly's approach; however, it requires an additional estimation of drawdown [Magdon-Ismail and Atiya 2004], which is also difficult. Thus, this poses one challenge to explore the power of optimal *f* and efficiently incorporate it to the area of online portfolio selection.

—*Loosening portfolio constraints*. Current portfolios are generally constrained in a simplex domain, which means that the portfolio is self-financed without margin/shorting. Besides current long-only portfolios, there also exist long/short portfolios [Jacobs and Levy 1993], which allow short selling and margin. Cover [1991] proposed a proxy to evaluate an algorithm when margin is allowed, by adding additional margin components for all assets. Moreover, the empirical results on NYSE data [Györfi et al. 2012, Chapter 4] show that there is no gain for online portfolio selection with short selling; however, with leverage, there is a spectacular increase in the growth rate. Yet, current methods are still in their infancy and far from application. Thus, the challenge is to develop effective algorithms when margin and shorting are allowed.

—*Extending transaction costs*. To make an algorithm practical, one has to consider some practical issues, such as transaction costs. Although several online portfolio selection models with transaction costs [Blum and Kalai 1999; Iyengar 2005; Györfi and Vajda 2008] have been proposed, they cannot be explicitly conveyed in an algorithmic perspective, which is hard to understand. One challenge is to extend current algorithms to the cases when transaction costs are taken into account.

—*Extending market liquidity*. Although all published algorithms claim that in the back tests they choose blue chip stocks, which have the highest liquidity, it cannot solve the concern of market liquidity. Completely solving this problem may involve paper trading or real trading, which is difficult for the community. Besides, no algorithm has ever considered this issue in its algorithm formulation. The challenge here is to accurately model the market liquidity and then design efficient algorithms.

6. CONCLUSIONS

This article conducted a survey on the online portfolio selection problem, an interdisciplinary topic of machine learning and finance. With the focus on algorithmic aspects, we began by formulating the task as a sequential decision learning problem and then further grouped the existing algorithms into five major categories: Benchmarks,

Follow the Winner, Follow the Loser, Pattern-Matching–based approaches, and MLAs. After presenting the surveys of individual algorithms, we further connected them to the CGT so as to better understand the essentials of their underlying trading ideas. Finally, we outlined some open challenges for future research investigations. We note that although quite a few algorithms have been proposed in the literature, many open research problems remain unsolved and deserve further exploration. It is our hope that this survey article facilitates researchers to understand the state of the art in this area and potentially inspires more fruitful future studies.

ACKNOWLEDGMENTS

REFERENCES

A. Agarwal, E. Hazan, S. Kale, and R. E. Schapire. 2006. Algorithms for Portfolio Management Based on the Newton Method. In *Proceedings of International Conference on Machine Learning*. 9–16.

K. Akcoglu, P. Drineas, and M. Kao. 2002. Fast universalization of investment strategies with provably good relative returns. In *Proceedings of International Colloquium on Automata, Languages and Programming*. 888–900.

K. Akcoglu, P. Drineas, and M. Kao. 2004. Fast Universalization of Investment Strategies. *SIAM J. Comput.* 34 (2004), 1–22.

M. Akian, A. Sulem, and M. I. Taksar. 2001. Dynamic Optimization of Long-Term Growth Rate for a Portfolio with Transaction Costs and Logarithmic Utility. *Mathematical Finance* 11, 2 (2001), 153–188.

P. Algoet. 1992. Universal Schemes for Prediction, Gambling and Portfolio Selection. *Annals of Probability* 20, 2 (1992), 901–941.

P. H. Algoet and T. M. Cover. 1988. Asymptotic Optimality and Asymptotic Equipartition Properties of Log-Optimum Investment. *Annals of Probability* 16, 2 (1988), 876–898.

F. Allen and R. Karjalainen. 1999. Using Genetic Algorithms to Find Technical Trading Rules. *Journal of Financial Economics* 51 (1999), 245–271.

L. Bachelier. 1900. Théorie de la spéculation. *Annales Scientifiques de l'École Normale Supérieure* 3, 17 (1900), 21–86.

A. R. Barron and T. M. Cover. 1988. A Bound on the Financial Value of Information. *IEEE Transactions on Information Theory* 34, 5 (1988), 1097–1100.

C. Y. Belentepe. 2005. *A Statistical View of Universal Portfolios*. Ph.D. Dissertation. University of Pennsylvania.

R. M. Bell and T. M. Cover. 1980. Competitive Optimality of Logarithmic Investment. *Mathematics of Operations Research* 5, 2 (1980), 161–162.

G. Biau, K. Bleakley, L. Györfi, and G. Ottucsák. 2010. Nonparametric Sequential Prediction of Time Series. *Journal of Nonparametric Statistics* 22, 3 (2010), 297–317.

J. R. Birge and F. Louveaux. 1997. *Introduction to Stochastic Programming*. Springer, New York.

A. Blum and A. Kalai. 1999. Universal Portfolios with and without Transaction Costs. *Machine Learning* 35, 3 (1999), 193–205.

A. Blum and Y. Mansour. 2007. From External to Internal Regret. *Journal of Machine Learning Research* 8 (2007), 1307–1324.

W. F. M. Bondt and R. Thaler. 1985. Does the Stock Market Overreact? *The Journal of Finance* 40, 3 (1985), 793–805.

W. F. M. Bondt and R. Thaler. 1987. Further Evidence on Investor Overreaction and Stock Market Seasonality. *Journal of Finance* 42, 3 (1987), 557–581.

A. Borodin, R. El-Yaniv, and V. Gogan. 2000. On the Competitive Theory and Practice of Portfolio Selection (Extended Abstract). In *Proceedings of the Latin American Symposium on Theoretical Informatics*. 173–196.

A. Borodin, R. El-Yaniv, and V. Gogan. 2003. Can We Learn to Beat the Best Stock. In *Proceedings of the Annual Conference on Neural Information Processing Systems*.

A. Borodin, R. El-Yaniv, and V. Gogan. 2004. Can We Learn to Beat the Best Stock. *Journal of Artificial Intelligence Research* 21 (2004), 579–594.

S. Boyd and L. Vandenberghe. 2004. *Convex Optimization*. Cambridge University Press, New York.

L. Breiman. 1960. Investment Policies for Expanding Businesses Optimal in a Long-Run Sense. *Naval Research Logistics Quarterly* 7, 4 (1960), 647–651.

L. Breiman. 1961. Optimal Gambling Systems for Favorable Games. In *Proceedings of the Berkeley Symposium on Mathematical Statistics and Probability* 1 (1961), 65–78.

P. Brockman and D. Michayluk. 1998. The Persistent Holiday Effect: Additional Evidence. *Applied Economics Letters* 5 (1998), 205–209.

L. J. Cao and F. E. H. Tay. 2003. Support Vector Machine with Adaptive Parameters in Financial Time Series Forecasting. *IEEE Transactions on Neural Networks* 14, 6 (2003), 1506–1518.

N. Cesa-Bianchi and G. Lugosi. 2006. *Prediction, Learning, and Games*. Cambridge University Press, New York.

L. K. C. Chan, N. Jegadeesh, and J. Lakonishok. 1996. Momentum Strategies. *Journal of Finance* 51, 5 (1996), 1681–1713.

K. Chaudhuri and Y. Wu. 2003. Mean Reversion in Stock Prices: Evidence from Emerging Markets. *Managerial Finance* 29 (2003), 22–37.

J. Conrad and G. Kaul. 1998. An Anatomy of Trading Strategies. *Review of Financial Studies* 11, 3 (1998), 489–519.

R. Cont. 2001. Empirical Properties of Asset Returns: Stylized Facts and Statistical Issues. *Quantitative Finance* 1, 2 (2001), 223–236.

M. J. Cooper, R. C. Gutierrez, and A. Hameed. 2004. Market States and Momentum. *The Journal of Finance* 59, 3 (2004), 1345–1365.

P. H. Cootner. 1964. *The Random Character of Stock Market Prices*. MIT Press.

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

T. M. Cover. 1996. Universal Data Compression and Portfolio Selection. In *Proceedings of the Annual IEEE Symposium on Foundations of Computer Science*. 534–538.

T. M. Cover and D. H. Gluss. 1986. Empirical Bayes Stock Market Portfolios. *Advances in Applied Mathematics* 7, 2 (1986), 170–181.

T. M. Cover and E. Ordentlich. 1996. Universal Portfolios with Side Information. *IEEE Transactions on Information Theory* 42, 2 (1996), 348–363.

T. M. Cover and J. A. Thomas. 1991. *Elements of Information Theory*. Wiley-Interscience, New York.

K. Crammer, O. Dekel, J. Keshet, S. Shalev-Shwartz, and Y. Singer. 2006. Online Passive-Aggressive Algorithms. *Journal of Machine Learning Research* 7 (2006), 551–585.

K. Crammer, M. Dredze, and A. Kulesza. 2009. Multi-Class Confidence Weighted Algorithms. In *Proceedings of the Conference on Empirical Methods in Natural Language Processing*. 496–504.

K. Crammer, M. Dredze, and F. Pereira. 2008. Exact Convex Confidence-Weighted Learning. In *Proceedings of the Annual Conference on Neural Information Processing Systems*. 345–352.

G. Creamer. 2007. *Using Boosting for Automated Planning And Trading Systems*. Ph.D. Dissertation. Columbia University.

G. Creamer. 2012. Model Calibration and Automated Trading Agent for Euro Futures. *Quantitative Finance* 12, 4 (2012), 531–545.

G. Creamer and Y. Freund. 2007. A Boosting Approach for Automated Trading. *Journal of Trading* 2, 3 (2007), 84–96.

G. Creamer and Y. Freund. 2010. Automated Trading with Boosting and Expert Weighting. *Quantitative Finance* 4, 10 (2010).

J. E. Cross and A. R. Barron. 2003. Efficient Universal Portfolios for Past-Dependent Target Classes. *Mathematical Finance* 13, 2 (2003), 245–276.

M. Dai, Z. Q. Xu, and X. Y. Zhou. 2010. Continuout-Time Mean-Variance Portfolio Selection with Proportional Transaction Costs. *SIAM Journal on Financial Mathematics* 1, 1 (2010), 96–125.

P. Das and A. Banerjee. 2011. Meta Optimization and Its Application to Portfolio Selection. In *Proceedings of International Conference on Knowledge Discovery and Data Mining*.

P. Das, N. Johnson, and A. Banerjee. 2013. Online Lazy Updates for Portfolio Selection with Transaction Costs. In *Proceedings of the 27th Conference on Artificial Intelligence*.

M. H. A. Davis and A. R. Norman. 1990. Portfolio Selection with Transaction Costs. *Mathematics of Operations Research* 15, 4 (1990), 676–713.

M. A. H. Dempster, T. W. Payne, Y. Romahi, and G. W. P. Thompson. 2001. Computational Learning Techniques for Intraday FX Trading Using Popular Technical Indicators. *IEEE Transactions on Neural Networks* 12, 4 (2001), 744–754.

V. Dhar. 2011. Prediction in Financial Markets: The Case for Small Disjuncts. *ACM Transactions on Intelligent Systems and Technology* 2, 3 (2011), 19:1–19:22.

M. Dredze, K. Crammer, and F. Pereira. 2008. Confidence-Weighted Linear Classification. In *Proceedings of the International Conference on Machine Learning*. 264–271.

M. Dredze, A. Kulesza, and K. Crammer. 2010. Multi-Domain Learning by Confidence-Weighted Parameter Combination. *Machine Learning* 79, 1–2 (2010), 123–149.

J. Duchi, S. Shalev-Shwartz, Y. Singer, and T. Chandra. 2008. Efficient Projections onto the l_1-ball for Learning in High Dimensions. In *Proceedings of the International Conference on Machine Learning*. 272–279.

C. Dzhabarov and W. T. Ziemba. 2010. Do Seasonal Anomalies Still Work? *Journal of Portfolio Management* 36, 3 (2010), 93–104.

R. El-Yaniv. 1998. Competitive Solutions for Online Financial Problems. *ACM Computing Survey* 30, 1 (1998), 28–69.

E. F. Fama. 1970. Efficient Capital Markets: A Review of Theory and Empirical Work. *Journal of Finance* 25, 2 (1970), 383–417.

M. Feder, N. Merhav, and M. Gutman. 1992. Universal Prediction of Individual Sequences. *IEEE Transactions on Information Theory* 38, 4 (1992), 1258–1270.

M. J. Fields. 1934. Security Prices and Stock Exchange Holidays in Relation to Short Selling. *Journal of Business of the University of Chicago* 7, 4 (1934), 328–338.

M. Finkelstein and R. Whitley. 1981. Optimal Strategies for Repeated Games. *Advances in Applied Probability* 13, 2 (1981), 415–428.

A. A. Gaivoronski and F. Stella. 2000. Stochastic Nonstationary Optimization for Finding Universal Portfolios. *Annals of Operationas Research* 100 (2000), 165–188.

A. A. Gaivoronski and F. Stella. 2003. On-line Portfolio Selection Using Stochastic Programming. *Journal of Economic Dynamics and Control* 27, 6 (2003), 1013–1043.

T. J. George and C.-Y. Hwang. 2004. The 52-Week High and Momentum Investing. *Journal of Finance* 59, 5 (2004), 2145–2176.

L. Györfi, G. Lugosi, and F. Udina. 2006. Nonparametric Kernel-Based Sequential Investment Strategies. *Mathematical Finance* 16, 2 (2006), 337–357.

L. Györfi, G. Ottucsák, and H. Walk. 2012. *Machine Learning for Financial Engineering*. Imperial College Press.

L. Györfi and D. Schäfer. 2003. Nonparametric Prediction. In *Advances in Learning Theory: Methods, Models and Applications*, J. A. K. Suykens, G. Horváth, S. Basu, C. Micchelli, and J. Vandevalle (Eds.). IOS Press, Amsterdam, Netherlands, 339–354.

L. Györfi, F. Udina, and H. Walk. 2008. Nonparametric Nearest Neighbor Based Empirical Portfolio Selection Strategies. *Statistics and Decisions* 26, 2 (2008), 145–157.

L. Györfi, A. Urbán, and I. Vajda. 2007. Kernel-Based Semi-Log-Optimal Empirical Portfolio Selection Strategies. *International Journal of Theoretical and Applied Finance* 10, 3 (2007), 505–516.

L. Györfi and I. Vajda. 2008. Growth Optimal Investment with Transaction Costs. In *Proceedings of the International Conference on Algorithmic Learning Theory*. 108–122.

L. Györfi and H. Walk. 2012. Empirical Portfolio Selection Strategies with Proportional Transaction Costs. *IEEE Transactions on Information Theory* 58, 10 (2012), 6320–6331.

N. H Hakansson. 1970. Optimal Investment and Consumption Strategies under Risk for a Class of Utility Functions. *Econometrica* 38, 5 (1970), 587–607.

N. H. Hakansson. 1971. Capital Growth and the Mean-Variance Approach to Portfolio Selection. *Journal of Financial and Quantitative Analysis* 6, 1 (1971), 517–557.

N. H. Hakansson and W. T. Ziemba. 1995. Capital Growth Theory. In *Handbooks in OR & MS*. Elsevier Science.

J. D. Hamilton. 1994. *Time Series Analysis*. Princeton University Press, Princeton, NJ.

J. D. Hamilton. 2008. *New Palgrave Dictionary of Economics*. Palgrave McMillan Ltd.

M. R. Hardy. 2001. A Regime-Switching Model of Long-Term Stock Returns. *North American Actuarial Journal Society of Acutaries* 5, 2 (2001), 41–53.

R. A. Haugen and J. Lakonishok. 1987. *The Incredible January Effect: The Stock Market's Unsolved Mystery*. Dow Jones-Irwin, Homewood, IL.

D. B. Hausch, W. T. Ziemba, and M. Rubinstein. 1981. Efficiency of the Market for Racetrack Betting. *Management Science* 27, 12 (1981), 1435–1452.

E. Hazan. 2006. *Efficient Algorithms for Online Convex Optimization and Their Applications*. Ph.D. Dissertation. Princeton University.

E. Hazan, A. Agarwal, and S. Kale. 2007. Logarithmic Regret Algorithms for Online Convex Optimization. *Machine Learning* 69, 2–3 (2007), 169–192.

E. Hazan, A. Kalai, S. Kale, and A. Agarwal. 2006. Logarithmic Regret Algorithms for Online Convex Optimization. In *Proceedings of the Annual Conference on Learning Theory*. 499–513.

E. Hazan and S. Kale. 2009. On Stochastic and Worst-case Models for Investing. In *Proceedings of the Annual Conference on Neural Information Processing Systems*. 709–717.

E. Hazan and S. Kale. 2012. An Online Portfolio Selection Algorithm With Regret Logarithmic In Price Variation. *Mathematical Finance* (2012).

E. Hazan and C. Seshadhri. 2009. Efficient Learning Algorithms for Changing Environments. In *Proceedings of the International Conference on Machine Learning*. 393–400.

D. P. Helmbold, R. E. Schapire, Y. Singer, and M. K. Warmuth. 1996. On-Line Portfolio Selection Using Multiplicative Updates. In *Proceedings of the International Conference on Machine Learning*. 243–251.

D. P. Helmbold, R. E. Schapire, Y. Singer, and M. K. Warmuth. 1997. A Comparison of New and Old Algorithms for a Mixture Estimation Problem. *Machine Learning* 27, 1 (1997), 97–119.

D. P. Helmbold, R. E. Schapire, Y. Singer, and M. K. Warmuth. 1998. On-Line Portfolio Selection Using Multiplicative Updates. *Mathematical Finance* 8, 4 (1998), 325–347.

M. Herbster and M. K. Warmuth. 1998. Tracking the Best Expert. *Machine Learning* 32, 2 (1998), 151–178.

D. Huang, J. Zhou, B. Li, S. C. H. Hoi, and S. Zhou. 2013. Robust Median Reversion Strategy for On-Line Portfolio Selection. In *Proceedings of the 23rd International Joint Conference on Artificial Intelligence*.

S.-H. Huang, S.-H. Lai, and S.-H. Tai. 2011. A Learning-based Contrarian Trading Strategy via a Dual-Classifier Model. *ACM Transactions on Intelligent Systems and Technology* 2, 3 (2011), 20:1–20:20.

J. C. Hull. 2008. *Options, Futures, and Other Derivatives* (7 ed.). Prentice Hall, Upper Saddle River, NJ.

G. Iyengar. 2005. Universal Investment in Markets with Transaction Costs. *Mathematical Finance* 15, 2 (2005), 359–371.

G. N. Iyengar and T. M. Cover. 2000. Growth Optimal Investment in Horse Race Markets with Costs. *IEEE Transactions on Information Theory* 46, 7 (2000), 2675–2683.

B. I. Jacobs and K. N. Levy. 1993. Long/Short Equity Investing. *Journal of Portfolio Management* 20, 1 (1993), 52–63.

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

N. Jegadeesh. 1990. Evidence of Predictable Behavior of Security Returns. *Journal of Finance* 45, 3 (1990), 881–898.

N. Jegadeesh. 1991. Seasonality in Stock Price Mean Reversion: Evidence from the U.S. and the U.K. *Journal of Finance* 46, 4 (1991), 1427–1444.

A. Kalai and S. Vempala. 2002. Efficient Algorithms for Universal Portfolios. *Journal of Machine Learning Research* 3 (2002), 423–440.

J. O. Katz and D. L. McCormick. 2000. *The Encyclopedia of Trading Strategies*. McGraw-Hill, New York.

J. Kelly Jr. 1956. A New Interpretation of Information Rate. *Bell Systems Technical Journal* 35 (1956), 917–926.

E. Keogh. 2002. Exact Indexing of Dynamic Time Warping. In *Proceedings of the 28th International Conference on Very Large Data Bases*. 406–417.

T. Kimoto, K. Asakawa, M. Yoda, and M. Takeoka. 1993. Stock Market Prediction System with Modular Neural Networks. *Neural Networks in Finance and Investing* (1993), 343–357.

W. M. Koolen and V. Vovk. 2012. Buy Low, Sell High. In *Proceedings of International Conference on Algorithmic Learning Theory*. 335–349.

S. S. Kozat and A. C. Singer. 2007. Universal Constant Rebalanced Portfolios with Switching. In *Proceedings of the International Conference on Acoustics, Speech, and Signal Processing*. 1129–1132.

S. S. Kozat and A. C. Singer. 2008. Universal Switching Portfolios under Transaction Costs. In *Proceedings of the International Conference on Acoustics, Speech, and Signal Processing*. 5404–5407.

S. S. Kozat and A. C. Singer. 2009. Switching Strategies for Sequential Decision Problems With Multiplicative Loss With Application to Portfolios. *IEEE Transactions on Signal Processing* 57, 6 (2009), 2192–2208.

S. S. Kozat and A. C. Singer. 2010. Universal Randomized Switching. *IEEE Transactions on Signal Processing* 58 (2010), 3.

S. S. Kozat and A. C. Singer. 2011. Universal Semiconstant Rebalanced Portfolios. *Mathematical Finance* 21, 2 (2011), 293–311.

S. S. Kozat, A. C. Singer, and A. J. Bean. 2008. Universal Portfolios via Context Trees. In *Proceedings of the International Conference on Acoustics, Speech, and Signal Processing*. 2093–2096.

S. S. Kozat, A. C. Singer, and A. J. Bean. 2011. A Tree-Weighting Approach to Sequential Decision Problems with Multiplicative Loss. *Signal Processing* 91, 4 (2011), 890–905.

C. M. C. Lee and B. Swaminathan. 2000. Price Momentum and Trading Volume. *Journal of Finance* 55 (2000), 2017–2069.

T. Levina and G. Shafer. 2008. Portfolio Selection and Online Learning. *International Journal of Uncertainty, Fuzziness and Knowledge-Based Systems* 16, 4 (2008), 437–473.

B. Li and S. C. H. Hoi. 2012. On-Line Portfolio Selection with Moving Average Reversion. In *Proceedings of the International Conference on Machine Learning*.

B. Li, S. C. H. Hoi, and V. Gopalkrishnan. 2011a. CORN: Correlation-driven Nonparametric Learning Approach for Portfolio Selection. *ACM Transactions on Intelligent Systems and Technology* 2, 3 (2011), 21:1–21:29.

B. Li, S. C. H. Hoi, P. Zhao, and V. Gopalkrishnan. 2011b. Confidence Weighted Mean Reversion Strategy for On-Line Portfolio Selection. In *Proceedings of the International Conference on Artificial Intelligence and Statistics*.

B. Li, S. C. H. Hoi, P. Zhao, and V. Gopalkrishnan. 2013. Confidence Weighted Mean Reversion Strategy for Online Portfolio Selection. *ACM Transactions on Knowledge Discovery from Data* 7, 1 (2013), 4:1–4:38.

B. Li, P. Zhao, S. C. H. Hoi, and V. Gopalkrishnan. 2012. PAMR: Passive aggressive mean reversion strategy for portfolio selection. *Machine Learning* 87, 2 (2012), 221–258.

D. Li and W.-L. Ng. 2000. Optimal Dynamic Portfolio Selection: Multiperiod Mean-Variance Formulation. *Mathematical Finance* 10, 3 (2000), 387–406.

G. Llorente, R. Michaely, G. Saar, and J. Wang. 2002. Dynamic Volume-Return Relation of Individual Stocks. *Review of Financial Studies* 15, 4 (2002), 1005–1047.

A. W. Lo. 2008. Where Do Alphas Come From? A Measure of the Value of Active Investment Management. *Journal of Investment Management* 6 (2008), 1–29.

A. W. Lo and A. C. MacKinlay. 1988. Stock Market Prices Do Not Follow Random Walks: Evidence from a Simple Specification Test. *Review of Financial Studies* 1, 1 (1988), 41–66.

A. W. Lo and A. C. MacKinlay. 1990. When Are Contrarian Profits Due to Stock Market Overreaction? *Review of Financial Studies* 3, 2 (1990), 175–205.

A. W. Lo and A. C. MacKinlay. 1999. *A Non-Random Walk Down Wall Street*. Princeton University Press, Princeton, NJ.

C.-J. Lu, T.-S. Lee, and C.-C. Chiu. 2009. Financial Time Series Forecasting Using Independent Component Analysis and Support Vector Regression. *Decision Support Systems* 47 (2009), 115–125. Issue 2.

D. G. Luenberger. 1998. *Investment Science*. Oxford University Press.

L. MacLean, E. Thorp, and W. Ziemba. 2010. Long-term Capital Growth: The Good and Bad Properties of the Kelly and Fractional Kelly Capital Growth Criteria. *Quantitative Finance* 10, 7 (2010), 681–687.

L. C. MacLean, E. O. Thorp, and W. T. Ziemba. 2011. *The Kelly Capital Growth Investment Criterion: Theory and Practice*. Vol. 3. World Scientific Publishing.

L. C. MacLean and W. T. Ziemba. 1999. Growth versus Security Tradeoffs in Dynamic Investment Snalysis. *Annals of Operations Research* 85 (1999), 193–225.

L. C. MacLean and W. T. Ziemba. 2008. Capital Growth: Theory and Practice. In *Handbook of Asset and Liability Management*, S. A. Zenios and W. T. Ziemba (Eds.). North-Holland, 429–473.

L. C. MacLean, W. T. Ziemba, and G. Blazenko. 1992. Growth versus Security in Dynamic Investment Analysis. *Management Science* 38, 11 (1992), 1562–1585.

M. Magdon-Ismail and A. Atiya. 2004. Maximum Drawdown. *Risk Magazine* 10 (2004), 99–102.

S. Mahfoud and G. Mani. 1996. Financial Forecasting Using Genetic Algorithms. *Applied Artificial Intelligence* 10 (1996), 543–565.

B. Mandelbrot. 1963. The Variation of Certain Speculative Prices. *Journal of Business* 36, 4 (1963), 394–419.

H. Markowitz. 1952. Portfolio Selection. *Journal of Finance* 7, 1 (1952), 77–91.

H. Markowitz. 1959. *Portfolio Selection: Efficient Diversification of Investments*. Wiley, New York.

H. Markowitz, G. P. Todd, and W. F. Sharpe. 2000. *Mean-Variance Analysis in Portfolio Choice and Capital Markets*. Wiley.

J. Mandziuk and M. Jaruszewicz. 2011. Neuro-genetic System for Stock Index Prediction. *Journal of Intelligent & Fuzzy Systems* 22 (2011), 93–123.

N. Merhav and M. Feder. 1998. Universal Prediction. *IEEE Transactions on Information Theory* 44, 6 (1998), 2124–2147.

H. Mlnařĺk, S. Ramamoorthy, and R. Savani. 2009. Multi-Strategy Trading Utilizing Market Regimes. In *Advances in Machine Learning for Computational Finance Workshop*.

N. Moller and S. Zilca. 2008. The Evolution of the January Effect. *Journal of Banking & Finance* 32, 3 (2008), 447–457.

J. Moody and M. Saffell. 2001. Learning to Trade via Direct Reinforcement. *IEEE Transactions on Neural Networks* 12, 4 (2001), 875–889.

J. Moody, L. Wu, Y. Liao, and M. Saffell. 1998. Performance Functions and Reinforcement Learning For Trading Systems and Portfolios. *Journal of Forecasting* 17 (1998), 441–471.

T. J. Moskowitz and M. Grinblatt. 1999. Do Industries Explain Momentum? *The Journal of Finance* 54, 4 (1999), 1249–1290.

J. O, J.-W. Lee, and B.-T. Zhang. 2002. Stock Trading System Using Reinforcement Learning with Cooperative Agents. In *Proceedings of the 19th International Conference on Machine Learning*. 451–458.

E. Ordentlich. 1996. *Universal Investment and Universal Data Compression*. Ph.D. Dissertation. Stanford University.

E. Ordentlich and T. M. Cover. 1998. The Cost of Achieving the Best Portfolio in Hindsight. *Mathematics of Operations Research* 23, 4 (1998), 960–982.

M. Ormos and A. Urbán. 2011. Performance Analysis of Log-Optimal Portfolio Strategies with Transaction Costs. *Quantitative Finance* (2011), 1–11.

M. F. M. Osborne. 1959. Brownian Motion in the Stock Market. *Operations Research* 7, 2 (1959), 145–173.

G. Ottucsák and I. Vajda. 2007. An Asymptotic Analysis of the Mean-Variance portfolio selection. *Statistics and Decisions* 25 (2007), 63–88.

J. M. Poterba and L. H. Summers. 1988. Mean Reversion in Stock Prices: Evidence and Implications. *Journal of Financial Economics* 22, 1 (1988), 27–59.

W. Poundstone. 2005. *Fortune's Formula: The Untold Story of the Scientific Betting System That Beat the Casinos and Wall Street*. Hill and Wang, New York.

Michael S. Rozeff and William R. Kinney Jr. 1976. Capital Market Seasonality: The Case of Stock Returns. *Journal of Financial Economics* 3, 4 (1976), 379–402.

L. Rabiner and S. E. Levinson. 1981. Isolated and Connected Word Recognition–Theory and Selected Applications. *IEEE Transactions on Communications* 29, 5 (1981), 621–659.

J. Rissanen. 1983. A Universal Data Compression System. *IEEE Transactions on Information Theory* 29, 5 (1983), 656–663.

L. M. Rotando and E. O. Thorp. 1992. The Kelly Criterion and the Stock Market. *Amer. Math. Monthly* (1992), 922–931.

K. G. Rouwenhorst. 1998. International Momentum Strategies. *Journal of Finance* 53, 1 (1998), 267–284.

H. Sakoe and S. Chiba. 1990. Dynamic Programming Algorithm Optimization for Spoken Word Recognition. In *Readings in Speech Recognition*. 159–165.

D. Schäfer. 2002. *Nonparametric Estimation for Fnancial Investment under Log-Utility*. Ph.D. Dissertation. Mathematical Institute, Universität Stuttgart.

S. Shalev-Shwartz, K. Crammer, O. Dekel, and Y. Singer. 2003. Online Passive-Aggressive Algorithms. In *Proceedings of the Annual Conference on Neural Information Processing Systems*.

Y. Singer. 1997. Switching Portfolios. *International Journal of Neural Systems* 8, 4 (1997), 488–495.

G. Stoltz and G. Lugosi. 2005. Internal Regret in On-Line Portfolio Selection. *Machine Learning* 59, 1–2 (2005), 125–159.

F. E. H. Tay and L. J. Cao. 2002. Modified Support Vector Machines in Financial Time Series Forecasting. *Neurocomputing* 48 (2002), 847–861.

S. Taylor. 2005. *Asset Price Dynamics, Volatility, and Prediction*. Princeton University Press, Princeton, NJ.

E. O. Thorp. 1962. *Beat the Dealer: A Winning Strategy for the Game of Twenty-One*. Blaisdell Publishing, New York.

E. O. Thorp. 1969. Optimal Gambling Systems for Favorable Games. *Review of the International Statistical Institute* 37, 3 (1969), 273–293.

E. O. Thorp. 1971. Portfolio Choice and the Kelly Criterion. In *Business and Economics Section of the American Statistical Association*. 215–224.

E. O. Thorp. 1997. The Kelly Criterion In Blackjack, Sports Betting, and The Stock Market. In *Proceedings of the International Conference on Gambling and Risk Taking*.

E. O. Thorp and S. T. Kassouf. 1967. *Beat the Market: A Scientific Stock Market System*. Random House, New York.

E. Tsang, P. Yung, and J. Li. 2004. EDDIE-Automation, A Decision Support Tool for Financial Forecasting. *Decision Support Systems* 37 (2004), 559–565.

I. Vajda. 2006. Analysis of Semi-Log-Optimal Investment Strategies. In *Proceedings of Prague Stochastic*, M. Huskova and M. Janzura (Eds.). Matfyzpress, 719–727.

R. Vince. 1990. *Portfolio Management Formulas: Mathematical Trading Methods for the Futures, Options, and Stock Markets*. Wiley, Hoboken, NJ.

R. Vince. 1992. *The Mathematics of Money Management: Risk Analysis Techniques for Traders*. Wiley, Hoboken, NJ.

R. Vince. 1995. *The New Money Management: A Framework for Asset Allocation*. Wiley, Hoboken, NJ.

R. Vince. 2007. *The Handbook of Portfolio Mathematics: Formulas for Optimal Allocation & Leverage*. Wiley, Hoboken, NJ.

R. Vince. 2009. *The Leverage Space Trading Model: Reconciling Portfolio Management Strategies and Economic Theory*. Wiley.

V. Vovk. 1990. Aggregating Strategies. In *Proceedings of the Annual Conference on Learning Theory*. 371–386.

V. Vovk. 1997. Derandomizing Stochastic Prediction Strategies. In *Proceedings of the 10th Annual Conference on Computational Learning Theory*. 32–44.

V. Vovk. 1999. Derandomizing Stochastic Prediction Strategies. *Machine Learning* 35 (1999), 247–282.

V. Vovk. 2001. Competitive On-Line Statistics. *International Statistical Review/Revue Internationale de Statistique* 69, 2 (2001), 213–248.

V. Vovk and C. Watkins. 1998. Universal Portfolio Selection. In *Proceedings of the Annual Conference on Learning Theory*. 12–23.

A. Weber. 1929. *Theory of the Location of Industries*. The University of Chicago Press, Chicago.

E. Weiszfeld. 1937. Sur le point pour lequel la somme des distances de n points donnes est minimum. *Tohoku Mathematical Journal* 43 (1937), 355–386.

R. E. S. Ziemba and W. T. Ziemba. 2007. *Scenarios for Risk Management and Global Investment Strategies*. John Wiley & Sons.

W. T. Ziemba. 2005. The Symmetric Downside-Risk Sharpe Ratio and the Evaluation of Great INvestors and Speculators. *Journal of Portfolio Management* 32, 1 (2005), 108–122.

W. T. Ziemba and D. B. Hausch. 1984. *Beat the racetrack*. Harcourt Brace Jovanovich.

W. T. Ziemba and D. B. Hausch. 2008. The Dr. Z Betting System in England. In *Efficiency of Racetrack Betting Markets*. World Scientific.

M. Zinkevich. 2003. Online Convex Programming and Generalized Infinitesimal Gradient Ascent. In *Proceedings of the International Conference on Machine Learning*.