3-2013

Confidence Weighted Mean Reversion Strategy for Online Portfolio Selection

Bin LI
Nanyang Technological University

Steven C. H. HOI
Singapore Management University, CHHOI@smu.edu.sg

Peilin ZHAO
Nanyang Technological University

Vivekanand Gopalkrishnan
Deloitte Analytics Institute (Asia)

DOI: https://doi.org/10.1145/2435209.2435213

Published in ACM Transactions on Knowledge Discovery from Data, Volume 7 Issue 1, March 2013, Article 4.
https://doi.org/10.1145/2435209.2435213

Confidence Weighted Mean Reversion Strategy for Online Portfolio Selection

BIN LI, STEVEN C. H. HOI, and PEILIN ZHAO, Nanyang Technological University
VIVEKANAND GOPALKRISHNAN, Deloitte Analytics Institute (Asia)

Online portfolio selection has been attracting increasing attention from the data mining and machine learning communities. All existing online portfolio selection strategies focus on the first order information of a portfolio vector, though the second order information may also be beneficial to a strategy. Moreover, empirical evidence shows that relative stock prices may follow the mean reversion property, which has not been fully exploited by existing strategies. This article proposes a novel online portfolio selection strategy named *Confidence Weighted Mean Reversion* (CWMR). Inspired by the mean reversion principle in finance and confidence weighted online learning technique in machine learning, CWMR models the portfolio vector as a Gaussian distribution, and sequentially updates the distribution by following the mean reversion trading principle. CWMR's closed-form updates clearly reflect the mean reversion trading idea. We also present several variants of CWMR algorithms, including a CWMR mixture algorithm that is theoretical universal. Empirically, CWMR strategy is able to effectively exploit the power of mean reversion for online portfolio selection. Extensive experiments on various real markets show that the proposed strategy is superior to the state-of-the-art techniques. The experimental testbed including source codes and data sets is available online.[1]

Categories and Subject Descriptors: J.1 [**Computer Applications**]: Administrative Data Processing—*Financial*; J.4 [**Computer Applications**]: Social and Behavioral Sciences—*Economics*; I.2.6 [**Artificial Intelligence**]: Learning

General Terms: Design, Algorithms, Economics, Experimentation

Additional Key Words and Phrases: Portfolio selection, mean reversion, confidence weighted learning, online learning

1. INTRODUCTION

Online portfolio selection (PS), also termed sequential portfolio selection, aims to determine a practical strategy for investing wealth among a set of assets to achieve some financial objectives in the long run. The finance community has mainly addressed

[1] http://www.cais.ntu.edu.sg/~chhoi/CWMR/

A short version of this work appeared in the 14th International Conference on Artificial Intelligence and Statistics (AISTATS 2011) [Li et al. 2011b].
This work was supported by Singapore MOE tier 1 project (RG33/11).
Authors' addresses: B. Li, S. C. H. Hoi, and P. Zhao, School of Computer Engineering, Nanyang Technological University, Singapore; email: {s080061, chhoi, zhao0106}@ntu.edu.sg; V. Gopalkrishnan, Deloitte Analytics Institute (Asia), Singapore; email: vivek@deloitte.com.

this problem by maximizing risk-adjusted returns [Elton et al. 1995; Markowitz 1952; Sharpe 1963, 1964]. On the other hand, this problem has also been actively investigated by exploring data mining and machine learning techniques that aim to maximize the logarithmic compound return or growth rate. These techniques include work in the information theory community [Breiman 1961; Cover 1991; Kelly 1956; Ordentlich and Cover 1996; Thorp 1971], the data mining, and machine learning communities [Agarwal et al. 2006; Borodin et al. 2004; Das and Banerjee 2011; Györfi et al. 2006, 2008; Helmbold et al. 1998; Li and Hoi 2012; Li et al. 2011a].

One popular trading assumption [Agarwal et al. 2006; Helmbold et al. 1998] is that current well-performing stocks would also perform well in the following trading period, which is often known as the trend following principle. However, empirical evidence [Jegadeesh 1990] indicates that such trends could often be violated, especially in the short term. This observation leads to a counter strategy of buying poor-performing stocks and selling well-performing stocks. Such a trading principle is known as mean reversion, which has been adopted by some existing methods [Borodin et al. 2004; Cover 1991].

One classical strategy that exploits the mean reversion trading idea is Constant Rebalanced Portfolios (CRP) [Cover 1991], which redistributes the wealth among all stocks based on a given portfolio at the end of each trading period. Although nicely grounded in theory, CRP's passive scheme is somewhat limited in achieving good performance. One recent study shows that the best CRP strategy in hindsight empirically performs significantly worse than an anticorrelation algorithm (Anticor) [Borodin et al. 2004], which redistributes the wealth by heuristically exploiting mean reversion via statistical correlations. This calls for a powerful learning method to actively exploit the mean reversion property. Besides, we notice that all existing strategies (refer to Section 3 for a review) only exploit the first order information of a portfolio vector, while the change in the portfolio distribution could be better reflected in both first order and second order information, that is, mean and variance.

To address these drawbacks, we present a new online portfolio selection strategy named Confidence Weighted Mean Reversion (CWMR). In short, CWMR models the portfolio vector as a Gaussian distribution and sequentially updates the distribution by applying online learning techniques to exploit the mean reversion trading principle. Unlike existing work, CWMR learns both first and second order information of a portfolio vector by exploiting the mean reversion property in the financial markets using the powerful Confidence Weighted (CW) online learning algorithm [Crammer et al. 2008; Dredze et al. 2008]. In order to provide a theoretical guarantee (refer to Section 3.3 for a review) for the proposed algorithm, we also create a mixture algorithm that mixes CWMR with other regret-bounded algorithms, such that the mixture algorithm is universal.

The key salient features of the proposed CWMR strategy are threefold.

— It is the first online portfolio selection approach that exploits the second order information of a portfolio (not the second order information of price).
— It can effectively exploit the mean reversion property of financial markets by applying confidence weighted learning technique.
— The proposed CWMR mixture algorithm has a safety guarantee (regret bound) and is a universal strategy.

Through an extensive set of numerical experiments on a variety of up-to-date real testbeds, we show that the proposed CWMR algorithms significantly surpass a number of state-of-the-art strategies in terms of both cumulative return and risk-adjusted return. The experiments on high frequency data, which is new to the online portfolio selection community, supports the assertion that the mean reversion principle is stronger in the short term markets. The experiments also show that CWMR is robust with respect to different parameter settings, and can withstand moderate transaction costs.

The rest of this article is organized as follows. Section 2 formally formulates the online portfolio selection problem. Section 3 reviews and analyzes related work. Section 4 presents the proposed CWMR algorithm and its mixture extension. Section 5 evaluates the empirical performance of the proposed algorithms on real historical stock markets. Finally, Section 6 concludes this work with future directions.

2. PROBLEM SETTING

This article tackles a focused problem in finance, to make it more accessible, we first introduce an abstract data mining problem and then formulate an online portfolio selection problem. Supposing a decision maker makes decisions in a sequential manner. At time t, given $t-1$ m-dimensional vectors $\mathbf{x}_1, \ldots, \mathbf{x}_{t-1}$, he/she wants to calculate an m-dimensional vector $\mathbf{b}_t$, denoting weights on the next vector $\mathbf{x}_t$ according to some criteria. As a result, from time 1 to time n, the decision process will produce n decision variables $\mathbf{b}_1, \ldots, \mathbf{b}_n$, which correspond to variables $\mathbf{x}_1, \ldots, \mathbf{x}_n$, respectively. The decision maker is finally scored based on certain problem-dependent criteria, depending on all decision variables.

Now let us consider the online portfolio selection problem. We want to invest over a financial market with m assets for n trading periods. On the tth period, the assets' price changes are represented by a positive *price relative vector*, that is, $\mathbf{x}_t \in \mathbb{R}_+^m$. The element x_{ti} denotes the ratio of closing price to last closing price of the ith asset at the end of the tth trading period; thus an investment in asset i on the tth period increases by a factor of x_{ti}. Let us use $\mathbf{x}^n = \{\mathbf{x}_1, \ldots, \mathbf{x}_n\}$ to denote the sequence of vectors for n periods.

An investment in the market at the beginning of the tth period is specified by a *portfolio vector* $\mathbf{b}_t = \left(b_{t1}, \ldots, b_{tm}\right)$, where b_{ti} represents the proportion of wealth invested in the ith asset. Typically, we assume the portfolio is self-financed and no margin/short is allowed[2], therefore each entry of the portfolio is non-negative and adds up to one, that is, $\mathbf{b}_t \in \Delta_m$, where $\Delta_m = \left\{\mathbf{b}_t : \mathbf{b}_t \in \mathbb{R}_+^m, \sum_{i=1}^m b_{ti} = 1\right\}$. The investment procedure is represented by a *portfolio strategy*, that is, a sequence of mappings $\mathbf{b}_t : \mathbb{R}_+^{m(t-1)} \to \Delta_m, t = 1, 2, \ldots$, where $\mathbf{b}_t = \mathbf{b}_t\left(\mathbf{x}_1, \ldots, \mathbf{x}_{t-1}\right)$ is the portfolio used on the tth period, given past market price relative sequence $\mathbf{x}^{t-1} = \left\{\mathbf{x}_1, \ldots, \mathbf{x}_{t-1}\right\}$. Let us denote by $\mathbf{b}^n = \left\{\mathbf{b}_1, \ldots, \mathbf{b}_n\right\}$, the portfolio strategy for a sequence of n trading periods.

On the tth period, an investment with portfolio vector $\mathbf{b}_t$ produces a *portfolio period return* $\mathbf{s}_t$, that is, the wealth increases by a factor of $\mathbf{s}_t = \mathbf{b}_t^\top \mathbf{x}_t = \sum_{i=1}^m b_{ti} x_{ti}$. Since we reinvest and adopt price relative, the portfolio wealth would increase multiplicatively. Thus, after n trading periods, the investment according to a portfolio strategy $\mathbf{b}^n$ produces a *portfolio cumulative wealth* $\mathbf{S}_n$, which increases the initial wealth by a factor of $\prod_{t=1}^n \mathbf{b}_t^\top \mathbf{x}_t$, that is,

$$\mathbf{S}_n\left(\mathbf{b}^n, \mathbf{x}^n\right) = \mathbf{S}_0 \prod_{t=1}^n \mathbf{b}_t^\top \mathbf{x}_t,$$

where $\mathbf{S}_0$ denotes the initial wealth, and is set to \$1 for convenience in this article.

[2]In other words, we assume long only portfolios, while one can extend the model into non-long only portfolios, as done by Cover and Ordentlich [1998] and Vovk and Watkins [1998].

Algorithm 1: Online Portfolio Selection framework.

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

Initialize $\mathbf{S}_0 = 1, \mathbf{b}_1 = \left(\frac{1}{m}, \ldots, \frac{1}{m}\right)$
for $t = 1, 2, \ldots, n$ **do**
  Portfolio manager learns the 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 $\mathbf{S}_t = \mathbf{S}_{t-1} \times \left(\mathbf{b}_t^\top \mathbf{x}_t\right)$;
  Portfolio manager updates his/her online portfolio selection rules ;
end

Finally, we formulate the online portfolio selection problem following the previously mentioned abstract problem. In a portfolio selection task the decision maker is the portfolio manager, whose goal is to produce a portfolio strategy $\mathbf{b}^n$ to satisfy certain targets. In this study, the target is to maximize the portfolio cumulative wealth $\mathbf{S}_n$. The portfolio manager computes the portfolios in a sequential fashion. On each period t, the manager has access to all previous sequences of price relative vectors $\mathbf{x}^{t-1}$. Then the portfolio manager computes a new portfolio vector $\mathbf{b}_t$ for the coming price relative vector $\mathbf{x}_t$, where the decision criterion varies among different managers. The vector $\mathbf{b}_t$ is scored using the portfolio period return s_t. This procedure is repeated until the ending period n, and the portfolio strategy is finally scored according to the portfolio cumulative wealth $\mathbf{S}_n$. Algorithm 1 shows the framework of online portfolio selection.

In general, some assumptions are made in the preceding widely adopted model.

(1) Transaction cost. We assume no transaction cost/tax exists in the model.
(2) Market liquidity. We assume that one can buy and sell the desired quantities at last closing prices.
(3) Impact cost. We assume the market behavior is not affected by the portfolio selection strategy.

The preceding assumptions are not trivial. The implications and effects of these assumptions will be further analyzed and discussed in Section 5.4.4 and Section 5.5.

3. RELATED WORK

3.1. Benchmark Approaches

The most common baseline is *Buy-And-Hold* (BAH) strategy, that is, one invests wealth among a pool of assets with a fixed initial portfolio $\mathbf{b}_1$ and holds the portfolio. The BAH strategy with a uniform initial portfolio $\mathbf{b}_1 = \left(\frac{1}{m}, \ldots, \frac{1}{m}\right)$ is referred to as *uniform BAH* strategy, which is adopted as market strategy to produce the market index in this study. Contrary to the static BAH strategy, active trading strategies usually change portfolios regularly during entire trading periods. A classical strategy is *Constant Rebalanced Portfolios* (CRP) [Cover and Gluss 1986], which keeps a fixed fraction of an investor's wealth in each of the assets every trading period. Assuming a CRP strategy with $\mathbf{b}$, that is, $\mathbf{b}^n = \{\mathbf{b}, \ldots, \mathbf{b}\}$, then the wealth achieved by the strategy is $\mathbf{S}_n\left(\mathbf{b}^n, \mathbf{x}^n\right) = \mathbf{S}_0 \prod_{t=1}^n \mathbf{b}^\top \mathbf{x}_t$. The best possible CRP strategy is often called *Best CRP*

(BCRP), which is only a hindsight strategy. The CRP strategy can take advantage of market fluctuations for active trading, and its underlying idea is the mean reversion principle, intuitively known as "*Buy Low, Sell High*." To handle transaction cost, Blum and Kalai [1999] proposed *semi-CRP*, which partially balances between potential return and potential transaction cost and rebalances to initial portfolio at the end of any subset of the trading periods rather than every trading period.

3.2. Online Learning

Online learning has been extensively studied in data mining and machine learning [Cesa-Bianchi et al. 2004; Crammer et al. 2006; Crammer and Singer 2003; Rosenblatt 1958; Wang et al. 2012; Zhao et al. 2011]. In this article, we apply online learning techniques to perform the online portfolio selection task since it perfectly matches the online nature of the task. In literature, a classical online learning algorithm is *Perceptron* [Freund and Schapire 1999; Rosenblatt 1958], which performs simple additive updates when an incoming example is misclassified. Recently a number of online learning algorithms have been proposed based on the criterion of maximum margin [Crammer and Singer 2003; Crammer et al. 2006; Gentile 2001; Kivinen et al. 2004; Li and Long 1999]. For example, *Relaxed Online Maximum Margin* algorithm [Li and Long 1999] repeatedly chooses the hyperplanes that correctly classify the existing training example with the maximum margin; *Passive Aggressive* (PA) [Crammer et al. 2006] algorithm updates the classifier using the maximum margin principle when the prediction loss of a new example is nonzero. Unlike the existing online learners that update only the weight vector of classifiers, the *Confidence Weighted* (CW) [Crammer et al. 2008, 2009; Dredze et al. 2008] algorithm updates both the classifier weight vector and the estimate of their parameters confidence. In particular, it models the uncertainty of a classification function with a Gaussian distribution over the weight vector and updates the mean and covariance of the distribution using every incoming example. Confidence weighted learning has demonstrated superior classification performance in comparison to the state-of-the-art online learning algorithms [Crammer et al. 2008, 2009; Dredze et al. 2008]. This motivates us to apply the idea of confidence weighted learning to tackle the online learning task for portfolio selection.

3.3. Online Portfolio Selection Strategy

The online portfolio selection problem has been extensively studied in several communities. In general, a strategy usually learns to compete with a target class of strategies. Following Cesa-Bianchi and Lugosi [2006], given the target class of strategies $\mathcal{Q} = \left\{Q^1, Q^2, \dots\right\}$, each element denotes one online strategy, we define the *worst-case logarithmic wealth ratio* achieved by strategy P as,

$$\mathbf{W}_n\left(P, \mathcal{Q}\right) = \sup_{\mathbf{x}^n} \sup_{Q\in\mathcal{Q}} \ln \frac{\mathbf{S}_n\left(Q, \mathbf{x}^n\right)}{\mathbf{S}_n\left(P, \mathbf{x}^n\right)}.$$

One can arbitrarily choose any class of strategy to target, for example, the widely adopted CRP strategy class (see Section 3.1 for the description of CRP strategy), or mixture of different classes of strategies.

Since the best CRP strategy is the optimal strategy in an independ identically distributed market (see Cover and Thomas [1991], Theorem 15.3.1), the online portfolio selection community always chooses CRP strategy $\mathcal{B}$ as a target class, which means that it compares a strategy with any possible CRP strategy in the simplex domain. In this case, the worst-case logarithmic wealth ratio becomes the *regret* [Cover 1991] of a strategy P, that is,

$$\text{regret}\,(P) = W_n\,(P, \mathcal{B}) = \sup_{\mathbf{x}^n} \sup_{B \in \Delta_m} \ln \frac{\mathbf{S}_n\,(B, \mathbf{x}^n)}{\mathbf{S}_n\,(P, \mathbf{x}^n)},$$

where $\mathbf{S}_n\,(B, \mathbf{x}^n)$ denotes the cumulative wealth achieved by a CRP strategy B and thus $\sup_{B \in \Delta_m} \mathbf{S}_n\,(B, \mathbf{x}^n)$ is obviously the wealth achieved by Best CRP strategy (BCRP). An online portfolio selection strategy P is *universal* if

$$\lim_{n \to \infty} \frac{1}{n} W_n\,(P, \mathcal{B}) \leq 0.$$

In other words, for arbitrary price relative sequences $\mathbf{x}^n$, a universal portfolio selection algorithm can asymptotically achieve no regret with respect to Best CRP strategy.

Cover [1991] proposed *Universal Portfolio* (UP) strategy, where the portfolio is the historical performance-weighted average of all CRP experts. The regret achieved by Cover's UP strategy is $\mathrm{O}(m \log n)$, and its running time complexity is $\mathrm{O}(n^m)$, which limits the practical applications. Kalai and Vempala [2002] presented a polynomial time-efficient implementation, which takes $\mathrm{O}(m^7 n^8)$. The UP strategy was further enhanced by Cover and Ordentlich [1996], who took into account various side information (fundamental data, experts' opinions, etc.). Cross and Barron [2003] proposed a new universal strategy, tracking the best-in-hindsight wealth achievable within target classes of linearly parameterized portfolio sequences, which are more general than the class of CRP strategy. Belentepe [2005] presented a statistical view of Cover's UP, showing that it is approximately equivalent to a constrained sequential portfolio optimization, thereby connecting Cover's UP strategy with traditional mean-variance portfolio theory [Markowitz 1952].

Helmbold et al. [1998] proposed *Exponential Gradient* (EG) strategy, which updates the portfolio using multiplicative updates. In essence, EG strategy attempts to maximize the expected logarithmic portfolio period return estimated by last price relative, and minimize the deviation from last portfolio. The regret achieved by EG strategy is $\mathrm{O}(\sqrt{n} \log m)$ with $\mathrm{O}(mn)$ running time. While its regret is not as tight as that of UP's, its linear time makes it more attractive in real large-scale applications.

Recently, online convex optimization has been applied on the portfolio selection problem [Agarwal and Hazan 2005; Agarwal et al. 2006]. *Online Newton Step* (ONS) strategy [Agarwal et al. 2006] aims to maximize the expected logarithmic cumulative wealth estimated by all historical price relatives [Gaivoronski and Stella 2000, 2003] and minimize the variation of next portfolio. ONS exploits the second order information of log function and applies it to an online learning scenario. It achieves a regret of $\mathrm{O}(m \log n)$, which is the same as Cover's UP, and has running time complexity of $\mathrm{O}(m^3 n)$. Following ONS, Hazan and Seshadhri [2009] recently proposed an adaptive regret approach. Moreover, Hazan and Kale [2009] linked ONS-type strategies (or "follow the leader" in online learning literature) for investing with probabilistic models of stock price returns, namely, Geometric Brownian Motion (GBM), and improved the regret to $\mathrm{O}(m \log Q)$, where Q is the quadratic variability of a price relative sequence and typically smaller than n.

Borodin et al. [2004] proposed a nonuniversal strategy named *Anticorrelation* (Anticor). Unlike previous approaches, Anticor takes advantage of the statistical properties of the financial markets. The underlying motivation is to bet on the consistency of positive lagged cross-correlation and negative autocorrelation. It exploits the statistical information from historical price relatives and adopts the mean reversion trading idea to transfer wealth among assets. Although without any theoretical guarantee, Anticor can outperform other existing strategies in most cases. Unlike the greedy algorithm by the Anticor strategy, Li et al. [2012] very recently proposed *Passive Aggressive Mean Reversion* (PAMR) strategy to actively exploit the mean reversion property and the first order information of a portfolio, which produces better performance than Anticor. To solve the drawbacks caused by the underlying single-period mean reversion of PAMR, Li and Hoi [2012] proposed *Moving Average Reversion* (MAR), which is a multiperiod mean reversion, and *Online Moving Average Reversion* (OLMAR) to exploit MAR. Empirically, OLMAR performs better than PAMR, especially on certain datasets that failed PAMR.

Györfi et al. [2006] introduced a framework of nonparametric learning strategies based on nonparametric prediction techniques [Györfi and Schäfer 2003]. On each trading period, the class of strategies searches over historical price relatives and identifies a set of price relatives, whose previous price relatives (in a window) are similar to recent price relatives and then obtains an optimal CRP portfolio based on these similar price relatives. With this framework, *Nonparametric kernel-based moving window* (B^K) [Györfi et al. 2006] strategy measures similarity using Euclidean distance. To improve the computational efficiency, Györfi et al. [2007] proposed another variant called *Nonparametric Kernel-Based Semi-Log-Optimal* strategy, which is an approximation of the B^K strategy. Replacing the utility function from log utility by a Markowitz-type utility function, Ottucsák and Vajda [2007] proposed *Nonparametric Kernel-based Markowitz-type* strategy, which connects return and risk (mean and variance) with nonparametric learning strategy. Following the same framework, *Nonparametric Nearest Neighbor learning* (B^{NN}) [Györfi et al. 2008] aims to search for ℓ nearest neighbors in historical sequence, which has been empirically shown to be robust. Recently, Li et al. [2011a] proposed *Correlation-driven Nonparametric learning* (CORN) strategy to search for similar price relatives via correlation coefficient, and considerably boosted the empirical performance of the nonparametric learning approach. In addition, Györfi and Vajda [2008] and Györfi et al. [2012, Chapter 3] studied the nonparametric learning strategies in cases of transaction costs.

In addition, aggregating algorithms [Vovk 1990] have also been investigated for online portfolio selection. Singer [1997] proposed *Switching Portfolio* (SP), which switches among a set of underlying strategies according to a prior distribution. Levina and Shafer [2008] introduced *Gaussian Random Walk* (GRW), which applies the aggregating algorithm and switches according to a Gaussian distribution. Sequential prediction techniques, example *Add-beta* [Borodin et al. 2000] prediction method (*T0* & *M0* algorithm), can also be applied to tackle this task.

Last, we note that our work is very different from another large body of existing work in the literature [Borodin et al. 2000; Cao and Tay 2003; Kimoto et al. 1993; Lu et al. 2009; Tay and Cao 2001], which attempted to make financial time series forecasting and stock price predictions by applying machine learning techniques, such as neural networks [Kimoto et al. 1993], decision trees [Tsang et al. 2004], and support vector machines (SVM) [Cao and Tay 2003; Lu et al. 2009; Tay and Cao 2001], and so on. The key difference between these existing works and ours is that their learning goal is to make explicit predictions of future prices/trends while our learning goal is to directly optimize portfolio selection without explicitly predicting prices.

3.4. Analysis of Existing Work

Some existing strategies (EG and ONS) adopt the "trend-following" trading idea, which assumes that price relative follows the same trend as last price relative, that is, winning stocks tend to win again in the following trading period. Despite being popular and easy to understand, trend following is generally hard to implement effectively. In

Table I. Summary of Pros and Cons of Existing Algorithms and the Proposed Algorithms

Trading Ideas	Algorithms	Pros	Cons
Trend Following	EG & ONS	· Easy to understand (Intuitive) · Universal	· Only first order information · Poor empirical performance
Mean Reversion	UP	· Fits the markets (Counterintuitive) · Universal	· Passive exploitation of MR · Only first order information · Poor empirical performance
	Anticor	· Fits the markets (Counterintuitive) · Good empirical performance	· Heuristical exploitation of MR · Only first order information · Nonuniversal
Pattern Matching	B^K, B^{NN} & CORN	· Universal[1] · Good empirical performance	· Mixes TF and MR · Only first order information
Mean Reversion	CWMR	· Fits the markets (Counterintuitive) · Active exploitation of MR · Both first and second order information · Good empirical performance	· Nonuniversal[2]

[1]Li et al. [2011a] does not prove CORN's universality.
[2]We further propose CWMR mixture extension, which is universal, in Section 4.5.

addition, in many short-term trading situations, stock price relatives may not follow previous trends, as empirically evidenced by Jegadeesh [1990] and Lo and MacKinlay [1990].

Contrary to the trend-following trading idea, the "mean reversion" trading principle assumes that if a stock performs worse than others, it tends to perform better in the next trading period. Thus, a mean reversion strategy tends to purchase poor-performing assets and sell good-performing assets. Some strategies (CRP, UP, and Anticor) adopt this idea. Empirically, CRP and UP strategy, which passively revert to the mean, often perform worse than Anticor, which actively reverts to the mean and thus can better exploit the fluctuation of asset prices [Borodin et al. 2004]. On the other hand, since Anticor heuristically transfers proportions within a portfolio, based on statistical correlations, it often produces suboptimal results. A new strategy to actively exploit the mean reversion property with a powerful learning method is necessary.

Besides trend following and mean reversion, pattern matching based algorithms, including nonparametric learning algorithms (B^K, B^{NN}, and CORN), achieve excellent performance in the back tests. Algorithms in this category can flexibly identify many market conditions, including both mean reversion and trend following. However, in certain cases, the pattern matching-based algorithms may locate both mean reverting and trend following price relatives, whose patterns are essentially contradictory, thus weakening the following maximization of the conditional expected logarithmic cumulative wealth.

Finally, all existing algorithms only consider the first order information of a portfolio vector, while the second order information (volatility of a portfolio vector) could provide useful volatility information, which can facilitate the portfolio selection task.

Table I summarizes the pros and cons of the existing algorithms and the proposed Confidence Weighted Mean Reversion (CWMR) strategy.

4. CONFIDENCE WEIGHTED MEAN REVERSION STRATEGY

4.1. Motivation

The proposed method is based on the mean reversion trading idea, which in the context of portfolio, or multiple stocks, implies that well-performing stocks tend to perform worse than others in the subsequent trading periods, and poor-performing stocks are inclined to perform better. Thus if we want to maximize next portfolio return, we could

Table II. Summary of Mean Reversion Statistics in Real Markets

Dataset	$\overline{P}(B)$	$\overline{G}(B)$	$\overline{P}(C)$	$\overline{G}(C)$	$\overline{P}(D)$	$\overline{G}(A)$	$\overline{G}(Market)$	$\overline{G}_{day}(A)$
DJA	46.36%	1.001315	32.24%	0.998749	21.46%	1.000220	0.999982	1.187180
NDX	49.18%	1.001323	33.88%	0.998784	16.94%	1.000255	0.999979	1.220029
TSE	42.89%	1.022370	41.63%	0.978395	15.48%	1.000598	1.000405	1.000598
MSCI	54.19%	1.015737	45.05%	0.984046	0.76%	1.001107	1.000053	1.001107
NYSE (O)	43.43%	1.021599	39.86%	0.981949	16.71%	1.002523	1.000620	1.002523
NYSE (N)	47.87%	1.019624	43.19%	0.982050	8.93%	1.001644	1.000610	1.001644
W-NYSE (O)	53.31%	1.034946	46.14%	0.973532	0.55%	1.007108	1.003054	1.001418
W-NYSE (N)	54.71%	1.036923	45.01%	0.968286	0.28%	1.007158	1.002933	1.001428

Notes 1: DJA and NDX are high frequency data; NYSE (O), NYSE (N), TSE, and MSCI are daily frequency data; W-NYSE (O) and W-NYSE (N) are weekly frequency data. Detail can be found in Table V in Section 5.1.
Notes 2: We empirically choose δ depending on the price relatives' average drift such that all sets (A, B, C, and D) are nonempty. In particular, we set $\delta = 0.998$ in high frequency datasets, $\delta = 0.985$ in daily datasets, and $\delta = 0.985$ in weekly datasets. Our test have indicated that the statistics with other thresholds also reach the same conclusions.

minimize the expected return with respect to today's price relative, since next price relative tends to revert. This seems somewhat counterintuitive, but according to Lo and MacKinlay [1990], the effectiveness of mean reversion is due to the positive cross-autocovariances across assets.

Now let us empirically analyze real market data to show that mean reversion does exist in real markets[3]. In general, to test mean reversion, the actual trading frequency is one key parameter[4]. Although researchers in finance often test on weekly data [Bondt and Thaler 1985, 1987; Chaudhuri and Wu 2003; Jegadeesh 1991; Poterba and Summers 1988], we expand our test into three types of trading period, that is, high frequency, daily, and weekly. Since our portfolio is long-only[5], we focus on whether it is possible to obtain a higher return than the market by investing in poorly-performing stocks[6]. With a threshold δ, let A_t be the set of poorly-performing stocks ($x_{t,i}<\delta$), B_t be the set of mean reversion (MR) stocks ($x_{t,i}<\delta$ && $x_{t+1,i}>1$), C_t be the set of non-mean reversion (non-MR) stocks ($x_{t,i}<\delta$ && $x_{t+1,i}<1$), and D_t be the set of remaining stocks ($x_{t,i}<\delta$ && $x_{t+1,i}=1$). In each period t, let us denote the percentage of a set U (U can be either A, B, C, or D) as $P_t(U) = |U_t|/|A_t|$, where $|\cdot|$ denotes the cardinality of the set, and the gain of uniformly investing in the set as $G_t(U) = \sum_{i \in U_t} x_{ti}/|U_t|$. For a total of n trading periods, we can calculate their average values as, $\overline{P}(U) = \frac{1}{n-1}\sum_{t=1}^{n-1} P_t(U)$ and $\overline{G}(U) = \frac{1}{n-1}\sum_{t=1}^{n-1} G_t(U)$, respectively. In particular, we refer to the percentage of mean reversion stocks as $\overline{P}(B)$, and the gain of mean reversion stocks as $\overline{G}(B)$. To show whether buying poorly-performing stocks result in profit, we calculate the gain of uniformly investing in poorly-performing stocks, denoted as $\overline{G}(A)$, and the gain of uniformly investing in the whole market, denoted as $\overline{G}(Market)$. To compare the mean reversion property in different frequencies, we convert $\overline{G}(A)$ to a daily basis, denoted as $\overline{G}_{day}(A)$.[7]

[3]The test program and datasets will be available at `http://www.cais.ntu.edu.sg/~libin/portfolios`.
[4]The econometric detail is beyond the scope of this article, and one may refer to related econometric articles.
[5]Long-only means if something is considered undervalued, managers would invest, while if something is considered overvalued, managers would avoid it.
[6]If short is allowed, we can also show whether shorting the well-performing stocks provides a higher return.
[7]For the US markets, we assume one trading day has 780×30 seconds and one trading week has 5 trading days.

Table II gives the statistics on eight real datasets with different intervals. Clearly, mean reversion does exist in the real markets ($\overline{P}(B) > \overline{P}(C)$), and uniformly investing in poorly-performing stocks provides a larger profit than market ($\overline{G}(A) > \overline{G}(Market)$). By comparing $\overline{G}_{day}(A)$ on NYSE (O/N) daily and weekly datasets, we can see that mean reversion is stronger in daily data than in weekly data. Moreover, it seems that mean reversion is strongest in high frequency data, though the comparisons are based on different assets and periods. Finally, both mean reversion stocks (B) and non-mean reversion stocks (C) are important. However, the following proposed algorithm is mainly based on mean reversion stocks, which one can easily extend to non-mean reversion stocks.

Moreover, all state-of-the-art approaches only exploit the first order information of a portfolio vector, while the second order volatility information may also benefit the portfolio selection task. Empirical studies [Chopra and Ziemba 1993] show that, in portfolio selection, errors in variance have about 5% impact on the objective value, as do errors in mean. We do not consider the covariances among the portfolio vectors, since they has much smaller impact on the final objective value [Chopra and Ziemba 1993]. To take advantage of both first and second order information, we adopt Confidence Weighted (CW) learning [Crammer et al. 2008; Dredze et al. 2008], which was originally proposed for classification. The basic idea of CW is to maintain a Gaussian distribution for the classifier, and sequentially update the classifier distribution according to Passive Aggressive (PA) learning [Crammer et al. 2006]. Thus, CW learning can take advantage of both first and second order information for the solution.

To address these concerns with existing work, in this article, we present a novel online portfolio selection method named Confidence Weighted Mean Reversion (CWMR). In order to exploit the first and second order information of a portfolio vector, we model the portfolio vector as a Gaussian distribution, which can satisfy our motivations and is probably the most widely studied distribution. We do not consider higher orders and other distributions because of their complexities. Then, we sequentially update the distribution according to the mean reversion idea. On the one hand, we keep the previous distribution if the portfolio is mean reversion profitable. On the other hand, we move to a new distribution such that the new distribution is expected to profit, while keeping it close to the previous distribution. Differently from CRP and Anticor, CWMR actively exploits the mean reversion property of financial markets with a powerful learning method. Moreover, compared with traditional online portfolio selection algorithms, which only consider the first order information, the proposed CWMR algorithm exploits both the first and second order information.

4.2. Formulation

We model $\mathbf{b}$ as a diagonal Gaussian distribution with mean $\mu \in \mathbb{R}^m$ and diagonal covariance matrix $\Sigma \in \mathbb{R}^{m \times m}$ with nonzero diagonal elements and zero for off-diagonal elements. The ith element of μ represents the proportion of the ith element. The ith diagonal term of Σ stands for the confidence on the ith proportion.

At the beginning of the tth period, we figure out a $\mathbf{b}$ based on the distribution $\mathcal{N}(\mu, \Sigma)$, i.e., $\mathbf{b} \sim \mathcal{N}(\mu, \Sigma)$. Then, after $\mathbf{x}_t$ reveals, the wealth increases by a factor of $\mathbf{b}^\top \mathbf{x}_t$. It is straightforward that the return $\mathrm{D} = \mathbf{b}^\top \mathbf{x}_t$ can be viewed as a random variable of the following univariate Gaussian distribution

$$\mathrm{D} \sim \mathcal{N}\left(\mu^\top \mathbf{x}_t, \mathbf{x}_t^\top \Sigma \mathbf{x}_t\right).$$

The distribution mean is the return of the mean vector and the variance is proportional to the length of the projection of $\mathbf{x}_t$ on Σ.

According to the mean reversion idea, the probability of a profitable $\mathbf{b}$ with respect to a predefined mean reversion threshold ϵ is defined as

$$\Pr_{\mathbf{b}\sim\mathcal{N}(\mu,\Sigma)}\left[\mathrm{D} \le \epsilon\right] = \Pr_{\mathbf{b}\sim\mathcal{N}(\mu,\Sigma)}\left[\mathbf{b}^\top \mathbf{x}_t \le \epsilon\right].$$

For simplicity, we write $\Pr\left[\mathbf{b}^\top \mathbf{x}_t \le \epsilon\right]$ instead. The parameter ϵ can be chosen empirically, we will discuss it in Section 4.3 and empirically evaluate its effect in Section 5.4.3. Note that we are considering the mean reversion profitability in a portfolio consisting of multiple stocks, thus this definition is equivalent to the motivating idea of buying poorly-performing stocks.

The algorithm adjusts the distribution to ensure that the probability of a mean reversion profitable $\mathbf{b}$ is higher than a confidence level parameter $\theta \in [0, 1]$

$$\Pr\left[\mathbf{b}^\top \mathbf{x}_t \le \epsilon\right] \ge \theta.$$

This is a bit counterintuitive but reasonable with respect to the mean reversion idea. If the portfolio return $\mathbf{b}^\top \mathbf{x}_t$ is less than a threshold with a high probability, the next return tends to be higher in a high probability since $\mathbf{x}_{t+1}$ will revert.

Then, following the intuition underlying PA algorithms [Crammer et al. 2006], our algorithm chooses the distribution closest to the current distribution $\mathcal{N}(\mu_t, \Sigma_t)$ in the Kullback-Leibler (KL) [Kullback and Leibler 1951] divergence sense. This ensures that if the current distribution satisfies the constraint, that is, it is mean reversion profitable with a high probability, we retain the current distribution. As a result, at the end of the tth period, the algorithm sets the parameters of the distribution by solving the following optimization problem.

The Raw Optimization Problem CWMR.

$$\begin{aligned}
\left(\mu_{t+1}, \Sigma_{t+1}\right) = \arg\min \quad & \mathrm{D}_{\mathrm{KL}}\left(\mathcal{N}(\mu, \Sigma) \,\|\, \mathcal{N}(\mu_t, \Sigma_t)\right) \\
\text{s. t.} \quad & \Pr\left[\mathbf{b}^\top \mathbf{x}_t \le \epsilon\right] \ge \theta \\
& \mu \in \Delta_m.
\end{aligned} \tag{1}$$

The optimization problem (1) clearly reflects our motivation. On the one hand, if current μ_t is mean reversion profitable, that is, the first constraint is satisfied, CWMR chooses the same distribution, resulting in a passive CRP strategy. On the other hand, if μ_t does not satisfy the constraint, then CWMR tries to figure out a new distribution, which is expected to profit and is not far from the current distribution.

Now let us reformulate both objective and constraint for the optimization problem, following Boyd and Vandenberghe [2004]. For the objective part, the KL divergence between the two Gaussian distributions is given as follows.

$$\mathrm{D}_{\mathrm{KL}}(\mathcal{N}(\mu, \Sigma) \,\|\, \mathcal{N}(\mu_t, \Sigma_t)) = \frac{1}{2}\left(\log\left(\frac{\det\Sigma_t}{\det\Sigma}\right) + \mathrm{Tr}\left(\Sigma_t^{-1}\Sigma\right) + (\mu_t - \mu)^\top \Sigma_t^{-1} (\mu_t - \mu) - d\right).$$

For the constraint part, since $\mathbf{b} \sim \mathcal{N}(\mu, \Sigma)$, $\mathbf{b}^\top \mathbf{x}_t$ has a univariate Gaussian distribution with mean $\mu_D = \mu^\top \mathbf{x}_t$ and variance $\sigma_D^2 = \mathbf{x}_t^\top \Sigma \mathbf{x}_t$. Thus the probability of a return less than ϵ is

$$\Pr\left[\mathrm{D} \le \epsilon\right] = \Pr\left[\frac{\mathrm{D} - \mu_D}{\sigma_D} \le \frac{\epsilon - \mu_D}{\sigma_D}\right].$$

In the preceding equation, $\frac{\mathrm{D}-\mu_D}{\sigma_D}$ is a normally distributed random variable, the probability equals $\Phi\left(\frac{\epsilon-\mu_D}{\sigma_D}\right)$, where Φ is the cumulative distribution function of the Gaussian distribution. As a result, we can rewrite the constraint as, $\frac{\epsilon-\mu_D}{\sigma_D} \geq \Phi^{-1}(\theta)$. Substituting μ_D and σ_D by their definitions and rearranging terms we can obtain

$$\epsilon - \mu^\top \mathbf{x}_t \geq \phi\sqrt{\mathbf{x}_t^\top \Sigma \mathbf{x}_t},$$

where $\phi = \Phi^{-1}(\theta)$. Clearly, we require that the weighted summation of return and standard deviation is less than the threshold. Now we can rewrite the preceding optimization problem as the following.

The Revised Optimization Problem CWMR.

$$\begin{aligned}(\mu_{t+1}, \Sigma_{t+1}) = \arg\min \quad & \frac{1}{2}\left(\log\left(\frac{\det\Sigma_t}{\det\Sigma}\right) + \mathrm{Tr}\left(\Sigma_t^{-1}\Sigma\right) + (\mu_t - \mu)^\top \Sigma_t^{-1}(\mu_t - \mu)\right)\\ \text{such that} \quad & \epsilon - \mu^\top \mathbf{x}_t \geq \phi\sqrt{\mathbf{x}_t^\top \Sigma \mathbf{x}_t}\\ & \mu^\top \mathbf{1} = 1, \quad \mu \succeq 0.\end{aligned} \tag{2}$$

Note that the short version [Li et al. 2011b] assumes log utility [Bernoulli 1954; Latané 1959] on $\mu^\top \mathbf{x}_t$ and is slightly different from this version. Since both ϵ and ϕ are adjustable, they have the similar effect on μ. Assuming other parameters constant except μ, as $\mu^\top \mathbf{x}_t > \log \mu^\top \mathbf{x}_t$, current linear form can move μ towards the mean reversion profitable portfolio more than log form can. However, log form in this constraint causes another convexity issue besides the standard deviation on the right-hand side. To solve the optimization problem with a log, Li et al. [2011b] chose to replace the log term by its linear approximation, which may converge to a different solution. Thus, we adopt return without log, which has no convexity issues concerning the log and its linear approximation.

For optimization Problem (2), the first constraint is not convex in Σ, therefore we have two ways to handle it. The first way [Dredze et al. 2008] is to linearize it by omitting the square root, that is, $\epsilon - \mu^\top \mathbf{x}_t \geq \phi \mathbf{x}_t^\top \Sigma \mathbf{x}_t$. Thus we can have the first final optimization problem, named CWMR-Var.

The Final Optimization Problem 1 CWMR-Var.

$$\begin{aligned}(\mu_{t+1}, \Sigma_{t+1}) = \arg\min \quad & \frac{1}{2}\left(\log\left(\frac{\det\Sigma_t}{\det\Sigma}\right) + \mathrm{Tr}\left(\Sigma_t^{-1}\Sigma\right) + (\mu_t - \mu)^\top \Sigma_t^{-1}(\mu_t - \mu)\right)\\ \text{such that} \quad & \epsilon - \mu^\top \mathbf{x}_t \geq \phi \mathbf{x}_t^\top \Sigma \mathbf{x}_t\\ & \mu^\top \mathbf{1} = 1, \quad \mu \succeq 0.\end{aligned} \tag{3}$$

The second reformulation [Crammer et al. 2008] is to decompose Σ since it is positive semidefinite (PSD), that is, $\Sigma = \Upsilon^2$ with $\Upsilon = \mathrm{Q}\mathrm{diag}\left(\lambda_1^{1/2}, \ldots, \lambda_m^{1/2}\right)\mathrm{Q}^\top$, where Q is orthonormal and $\lambda_1, \ldots, \lambda_m$ are the eigenvalues of Σ and thus Υ is also PSD. This reformulation yields the second final optimization problem, named CWMR-Stdev.

The Final Optimization Problem 2 CWMR-Stdev.

$$\begin{aligned}(\mu_{t+1}, \Upsilon_{t+1}) = \arg\min \quad & \frac{1}{2}\left(\log\left(\frac{\det\Upsilon_t^2}{\det\Upsilon^2}\right) + \mathrm{Tr}\left(\Upsilon_t^{-2}\Upsilon^2\right) + (\mu_t - \mu)^\top \Upsilon_t^{-2}(\mu_t - \mu)\right)\\ \text{such that} \quad & \epsilon - \mu^\top \mathbf{x}_t \geq \phi\left\|\Upsilon \mathbf{x}_t\right\|, \quad \Upsilon \text{ is PSD}\\ & \mu^\top \mathbf{1} = 1, \quad \mu \succeq 0.\end{aligned} \tag{4}$$

Clearly, revised optimization Problem (2), is equivalent to raw optimization Problem (1). From the revised problem, we proposed two final optimization Problems (3) and (4), which are convex, and thus can be efficiently solved by convex optimization [Boyd and Vandenberghe 2004]. The first variation, CWMR-Var, linearizes the constraint, thus it results in an approximate solution for the revised optimization problem and the raw optimization problem. While the second variation, CWMR-Stdev, is equivalent to revised optimization Problem (2), and results in an exact solution for both the revised and raw optimization problems.

4.3. Algorithm

Now, let us generate the proposed algorithms based on the solutions of the two optimization problems. The solutions for optimization Problems (3) and (4) are shown in Proposition 4.1 and Proposition 4.2, respectively. The proofs are presented in Appendix 6 and Appendix 6, respectively.

PROPOSITION 4.1. *The solution to final optimization Problem* (3) *(CWMR-Var) is expressed as*

$$\mu_{t+1} = \mu_t - \lambda_{t+1} \Sigma_t \left(\mathbf{x}_t - \bar{x}_t \mathbf{1} \right), \quad \Sigma_{t+1}^{-1} = \Sigma_t^{-1} + 2\lambda_{t+1} \phi \mathbf{x}_t \mathbf{x}_t^{\top},$$

where λ_{t+1} *corresponds to the Lagrangian multiplier calculated according to Equation* (11) *in Appendix A and* $\bar{x}_t = \frac{\mathbf{1}^{\top} \Sigma_t \mathbf{x}_t}{\mathbf{1}^{\top} \Sigma_t \mathbf{1}}$ *denotes the confidence weighted average of* $\mathbf{x}_t$*.*

PROPOSITION 4.2. *The solution to final optimization Problem* (4) *(CWMR-Stdev) is expressed as*

$$\mu_{t+1} = \mu_t - \lambda_{t+1} \Sigma_t \left(\mathbf{x}_t - \bar{x}_t \mathbf{1} \right), \quad \Sigma_{t+1}^{-1} = \Sigma_t^{-1} + \lambda_{t+1} \phi \frac{\mathbf{x}_t \mathbf{x}_t^{\top}}{\sqrt{U_t}},$$

where λ_{t+1} *denotes the Lagrangian multiplier calculated according to Equation* (15) *in Appendix B,* $\bar{x}_t = \frac{\mathbf{1}^{\top} \Sigma_t \mathbf{x}_t}{\mathbf{1}^{\top} \Sigma_t \mathbf{1}}$ *represents the confidence weighted average of* $\mathbf{x}_t$*, and* $V_t = \mathbf{x}_t^{\top} \Sigma_t \mathbf{x}_t$ *and* $\sqrt{U_t} = \frac{-\lambda_{t+1} \phi V_t + \sqrt{\lambda_{t+1}^2 \phi^2 V_t^2 + 4V_t}}{2}$ *denote the return variances for the* t^{th} *and* $t{+}1^{\text{th}}$ *period, respectively.*

Initially, with no information available for the task, we simply initialize μ_1 to uniform, and each diagonal element of the covariance matrix Σ_1 to variance $\frac{1}{m^2}$, or equivalent standard deviation $\frac{1}{m}$. It is worth noting that we solve the optimization problems by ignoring the non-negativity constraint ($\mu \succeq 0$) for its complexity, which is a typical way to reduce the complexity as in existing work [Agarwal et al. 2006; Helmbold et al. 1997, 1998]. To solve the issue that μ can be negative, we simply project the resulting μ to the simplex domain to ensure the simplex constraint [Agarwal et al. 2006]. The projection can be efficiently implemented in linear time [Duchi et al. 2008] with respect to the dimension of a vector. In the context of investment, this means that we first allow shorting, and later lower the leverage with the projection. Another remaining issue is that although the covariance matrix is nonsingular in theory, in real computation, the covariance matrix Σ sometimes may be singular due to computer precision. To avoid this problem and be consistent with the projection of μ, we rescale Σ by normalizing its summation value to $\frac{1}{m}$, which equals the sum of elements in μ_1. Note that we arbitrarily chose $\frac{1}{m}$, while one can chose other values, which generally do not affect the empirical performance too much. The final CWMR algorithms are

Algorithm 2: Confidence Weighted Mean Reversion: CWMR $\left(\phi, \epsilon, \left(\mu_t, \Sigma_t\right), \mathbf{x}_1^t, t\right)$.

Input: ϕ: Confidence parameter; $\epsilon \in [0, 1]$: Mean reversion parameter; (μ_t, Σ_t): Current portfolio distribution; $\mathbf{x}_1^t$: Historical market sequence; t: Index of current trading period

Output: $\left(\mu_{t+1}, \Sigma_{t+1}\right)$: Next portfolio distribution

1 Calculate the following variables:

$$M_t = \mu_t^\top \mathbf{x}_t, \quad V_t = \mathbf{x}_t^\top \Sigma_t \mathbf{x}_t, \quad W_t = \mathbf{x}_t^\top \Sigma_t \mathbf{1}, \quad \bar{x}_t = \frac{\mathbf{1}^\top \Sigma_t \mathbf{x}_t}{\mathbf{1}^\top \Sigma_t \mathbf{1}}$$

2 Update the portfolio distribution:

$$\text{CWMR-Var} \begin{cases} \lambda_{t+1} \text{ as in Eq. (11) in Appendix 6} \\ \mu_{t+1} = \mu_t - \lambda_{t+1} \Sigma_t \left(\mathbf{x}_t - \bar{x}_t \mathbf{1}\right) \\ \Sigma_{t+1} = \left(\Sigma_t^{-1} + 2\lambda_{t+1}\phi \operatorname{diag}^2\left(\mathbf{x}_t\right)\right)^{-1} \end{cases}$$

$$\text{CWMR-Stdev} \begin{cases} \lambda_{t+1} \text{ as in Eq. (15) in Appendix 6} \\ \sqrt{U}_t = \frac{-\lambda_{t+1}\phi V_t + \sqrt{\lambda_{t+1}^2 \phi^2 V_t^2 + 4V_t}}{2} \\ \mu_{t+1} = \mu_t - \lambda_{t+1} \Sigma_t \left(\mathbf{x}_t - \bar{x}_t \mathbf{1}\right) \\ \Sigma_{t+1} = \left(\Sigma_t^{-1} + \lambda_{t+1} \frac{\phi}{\sqrt{U_t}} \operatorname{diag}^2\left(\mathbf{x}_t\right)\right)^{-1} \end{cases}$$

3 Normalize μ_{t+1} and Σ_{t+1}:

$$\mu_{t+1} = \operatorname*{arg\,min}_{\mu \in \Delta_m} \left\| \mu - \mu_{t+1} \right\|^2, \ \Sigma_{t+1} = \frac{\Sigma_{t+1}}{m \operatorname{Tr}\left(\Sigma_{t+1}\right)}$$

presented in Algorithm 2, and the online portfolio selection, with both deterministic and stochastic CWMR algorithms, is illustrated in Algorithm 3.

The algorithms have two possible parameters, that is, the confidence parameter ϕ, and the mean reversion parameter ϵ. Typically, the first parameter, ϕ, can be 1.28, 1.64, 1.95, or 2.57, with corresponding θ values 80%, 90%, 95%, or 99%. As we have found, ϕ does not overly affect the final performance. On the contrary, the second parameter, ϵ, has significant impact on the final performance. As our model is long-only, we put more weight on the poorly-performing stocks, thus, ϵ is often in the range of $[0, 1]$. On the one hand, if the value is too large, such as $\epsilon \geq 1.2$, the last portfolio distribution can always satisfy the constraint and no update is required. With initial uniform portfolio, CWMR will degrade to uniform CRP. On the other hand, if the value is too small, such as $\epsilon \leq 0.5$, the constraint cannot always be satisfied and then the distribution has to be frequently updated to satisfy the constraint. In between, CWMR updates the distribution when the last distribution cannot satisfy the constraint. We will further validate this analysis by evaluating the parameter effect in Section 5.4.3.

4.4. Analysis and Interpretation

In this section, we give some analysis and interpretations of the proposed algorithms. First, we compare CWMR algorithms with Confidence Weight (CW) learning [Crammer et al. 2008; Dredze et al. 2008]. Then, we analyze CWMR's update schemes, that is, μ and Σ, with running examples. Further, we analyze the behavior of the

Algorithm 3: Online Portfolio Selection with CWMR.

Input: $\phi = \Phi^{-1}(\theta)$: Confidence parameter; $\epsilon \in [0, 1]$: Mean reversion parameter; $\mathbf{x}_1^n$: Historical market sequence

Output: $\mathbf{S}_n$: Final cumulative wealth

Initialization: $t = 1$, $\mu_1 = \frac{1}{m}\mathbf{1}$, $\Sigma_1 = \frac{1}{m^2}\mathbf{I}$, $\mathbf{S}_0 = 1$

for $t = 1, \ldots, n$ **do**

Draw a portfolio $\mathbf{b}_t$ from $\mathcal{N}(\mu_t, \Sigma_t)$:

$$\text{Determinstic CWMR}: \quad \mathbf{b}_t = \mu_t$$

$$\text{Stochastic CWMR}: \quad \tilde{\mathbf{b}}_t \sim \mathcal{N}(\mu_t, \Sigma_t), \quad \mathbf{b}_t = \underset{\mathbf{b}\in\Delta_m}{\arg\min} \left\| \mathbf{b} - \tilde{\mathbf{b}}_t \right\|^2$$

Receive stock price relatives: $\mathbf{x}_t = (x_{t1}, \ldots, x_{tm})$;

Calculate the daily return and cumulative return: $\mathbf{S}_t = \mathbf{S}_{t-1} \times \left(\mathbf{b}_t^\top \mathbf{x}_t\right)$;

Update the portfolio distribution: $\left(\mu_{t+1}, \Sigma_{t+1}\right) = \text{CWMR}\left(\phi, \epsilon, (\mu_t, \Sigma_t), \mathbf{x}_1^t, t\right)$;

end

stochastic version. Finally, we show the computational time complexity and compare it to existing work.

The CWMR algorithms are partially motivated by CW learning, thus their formulations and subsequent derivations are similar. However, they address different problems, since CWMR aims to handle online portfolio selection, while CW focuses on classification. Although both objectives adopt KL divergence to measure the closeness between two distributions, their constraints reflect that they are oriented to different problems. To be specific, CW's constraint is the probability of a correct prediction, while CWMR's constraints are the probability of an underperforming portfolio in the current period plus the simplex constraint. If there is mean reversion, the portfolio should be profitable, in the next period. The formulations' differences result in different derivations.

Now we provide a preliminary analysis of the update behavior of mean μ, which is the main concern for CWMR, to reflect its underlying mean reversion idea. Both CWMR-Var and CWMR-Stdev have the same update on μ, that is, $\mu_{t+1} = \mu_t - \lambda_{t+1}\Sigma_t(\mathbf{x}_t - \bar{x}_t\mathbf{1})$. Obviously, λ_{t+1} is non-negative and Σ_t is PSD. The term $\mathbf{x}_t - \bar{x}_t\mathbf{1}$ denotes excess return vector for the t^{th} period, where $\bar{x}_t$ is confidence weighted average of $\mathbf{x}_t$. Holding other terms constant, the mean μ_{t+1} tends to move towards μ_t, while the magnitude is negatively related to the previous excess return, which is in effect, the mean reversion idea. Meanwhile, these movements are dynamically adjusted by optimal λ_{t+1}, previous covariance matrix Σ_t, and mean μ_t, which catch both first- and second-order information. To the best of our knowledge, none of previous online portfolio selection algorithms have explicitly exploited the second order information of $\mathbf{b}$, while the second order information could contribute to the success of the proposed algorithms.

Let us continue to analyze the update of the covariance matrix Σ. With only non-zero diagonal elements, we can write the update of the i^{th} variance as $\sigma^2 = \sigma_i^2 / \left(1 + \lambda_{t+1}\phi' x_{ti}^2 \sigma_i^2\right)$, where $\phi' = 2\phi$ for CWMR-Var and $\phi' = \frac{\phi}{\sqrt{U_t}}$ for CWMR-Stdev. Since both λ_{t+1} and ϕ' are positive, poorly-performing stocks with lower values of x_{ti} have higher variance terms than that of well-performing stocks with higher x_{ti}. Note here that Σ denotes the covariance matrix of $\mathbf{b}$ rather than $\mathbf{x}$. Thus, a higher value

Table III. Running Example of CWMR-Stdev on Cover's Game

t	$\mathbf{x}_t$	$\mathbf{b}_t$	$\mathbf{b}_t^\top \mathbf{x}_t$	λ_t	$\mathbf{x}_t - \bar{x}_t\mathbf{1}$	diag (Σ_t)	μ_t
0						(0.25, 0.25)	(0.5, 0.5)
1	(1.0, 0.5)	(0.5, 0.5)	0.75	40.78	(0.25, -0.25)	(0.10, 0.40)	(0.0, 1.0)
2	(1.0, 2.0)	(0.0, 1.0)	2.00	61.61	(-0.80, 0.20)	(0.40, 0.10)	(1.0, 0.0)
3	(1.0, 0.5)	(1.0, 0.0)	1.00	75.56	(0.10, -0.40)	(0.10, 0.40)	(0.0, 1.0)
4	(1.0, 2.0)	(0.0, 1.0)	2.00	31.61	(-0.80, 0.20)	(0.40, 0.10)	(1.0, 0.0)
5	(1.0, 0.5)	(1.0, 0.0)	1.00	75.56	(0.10, -0.40)	(0.10, 0.40)	(0.0, 1.0)
$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$

means that the corresponding mean is more volatile than others. Since we move the weights from well-performing stocks to poorly-performing ones, the latter will change more significantly than the former, that is, the latter has higher volatility. In the next update of μ, stocks with higher volatility would magnify the movement magnitude, and the direction would be determined by the excess return vector.

To better illustrate the updates, we give running updates based on the classic example by Cover and Gluss [1986]. Let a portfolio consist of cash and one volatility asset, and the sequence of $\mathbf{x}$ is $\left(1, \frac{1}{2}\right), (1, 2), \left(1, \frac{1}{2}\right), \ldots$. Obviously, market strategy can gain nothing since no asset grows in the long run. The best CRP strategy, with $\mathbf{b} = \left(\frac{1}{2}, \frac{1}{2}\right)$, grows to $\left(\frac{9}{8}\right)^{\frac{n}{2}}$ at the end of the n^{th} period. However, starting with $\mu_0 = \left(\frac{1}{2}, \frac{1}{2}\right)$, the CWMR strategy can grow to $\frac{3}{4} \times 2^{\frac{n-1}{2}}$ at the end of the n^{th} period. The running details for the first 5 periods are shown in Table III, and further details can be easily derived. Let us continue the preceding analysis of mean μ. In each period t, the mean moves toward the previous mean and also moves far away from the excess return vector ($\mathbf{x}_t - \bar{x}_t\mathbf{1}$), and the movement magnitude is determined by both λ_t and Σ_t. Note that in this example μ before projection is out of the simplex and is made sparse via normalization, which is not usual case in real tests. In summary, both the first and second order information contribute to the success of the strategy, according to the preceding analysis.

Then, let us compare the deterministic CWMR with the stochastic version (Line 3 in Algorithm 3), which includes a covariance matrix besides the mean to draw a portfolio. Interestingly, we find that Σ negatively affects CWMR's performance in the following aspects. First, the stochastic $\mathbf{b}$ drawn from the distribution is always different from the optimal mean μ, which obviously causes performance divergences. Given that Σ converges to the zero matrix (see the recursive updates in the two propositions), the distribution of $\mathbf{b}$ conditioning on the data converges to the point mass at the mean parameter value $\mu = \lim_t \mu_t$. Thus, it is clear that drawing weights $\mathbf{b}$ from the distribution (the stochastic version) is suboptimal, since we already have an estimate of μ. It is better to choose $\mathbf{b}$ as either the mode or mean (incidentally the same for the Gaussian case), which is actually the deterministic version. Another effect caused by the stochastic behavior is the additional projection, since sometimes the stochastic $\mathbf{b}$ may be out of the simplex domain. To better understand these two aspects, let us continue Cover's game in Table III. For the first case, assuming we are at the beginning of the first period, we have $\mu = (0.5, 0.5)$ and diag $(\Sigma) = (0.25, 0.25)$. We draw stochastic $\mathbf{b}$ for 10000 times, and the average $\mathbf{b}$ after projection is $(0.5038, 0.4962)$ (the value before projection is $(0.5070, 0.4993)$), which is slightly deviated from the optimal mean and will result in different performance. For the second case, assuming at the beginning of the second period, we have $\mu = (0, 1)$ and diag $(\Sigma) = (0.1, 0.4)$. We average 10,000

Table IV. Summary of Time Complexity Analysis for Online Portfolio Selection Algorithms

Methods	Time Complexity	Methods	Time Complexity
UP	$O(n^m)$ [Cover 1991]	ONS	$O\left(m^3 n\right)$
	$O\left(m^7 n^8\right)$ [Kalai and Vempala 2002]	Anticor	$O\left(N^3 m^2 n\right)$
EG/SP/GRW/M0	$O(mn)$	B^K/B^{NN}/CORN	$O\left(N^2 mn^2\right)+O\left(Nmn^2\right)$
CWMR	$O(mn)$		

Sources: All time complexities are acquired from their respective studies.
Notes 1: m denotes the number of stocks; n is the number of trading periods; N denotes the number of experts.
Notes 2: Nonparametric learning approaches (B^K, B^{NN}, and CORN) require a nonlinear optimization step each period, that is, $\mathbf{b}_{t+1} = \arg\max_{\mathbf{b}\in\Delta_m} \prod_i \left(\mathbf{b}^\top \mathbf{x}_i\right)$, whose time complexity is generally high. To produce an approximate solution, batch gradient projection algorithms [Helmbold et al. 1997] take $O(mn)$, while batch convex Newton method [Agarwal et al. 2006] take $O(m^3 n)$. In the table, we set step $O(mn)$ time complexity. In our implementation, we adopt the Matlab optimization toolbox (function *fmincon* with *active-set*) to obtain the exact solution.

stochastic b's after projection, and get an average $\mathbf{b} = (0.1391, 0.8609)$, which is far from the optimal mean $(0, 1)$. In both cases, stochastic $\mathbf{b}$ tends to deviate from the optimal mean, especially in the second case, and thus underperforms the deterministic CWMR, which is clearly shown in the experiments (see Table VII for a review).

Since computational time is of crucial importance for certain trading scenarios, such as high frequency trading, which can occur in fractions of a second, we finally show CWMR's time complexity with m stocks and n periods, where n is typically much larger than m. In the CWMR implementation, we only consider the diagonal elements of Σ, thus the inverse can be computed in linear time. Moreover, the projection (Line 3 in Algorithm 2) can be implemented[8] in $O(m)$ time [Duchi et al. 2008]. Thus, in total, CWMR algorithms (Algorithm 2) take $O(m)$ time per period. Straightforwardly, online portfolio selection with the CWMR (Algorithm 3) takes $O(mn)$ time. Table IV compares the time complexity of CWMR with that of existing strategies. Clearly, the proposed CWMR algorithms take no more time than any others.

4.5. CWMR-Mixture Algorithm

The proposed CWMR algorithms have two possible parameters, namely, ϕ and ϵ. Though CWMR empirically performs robustly with respect to the parameters (c.f. parameter sensitivity evaluation in Section 5.4.3), their existence limits CWMR's potential applicability. Moreover, although the proposed CWMR algorithm works well on real markets (c.f. Section 5 for a review), the lack of a traditional regret bound would reduce the confidence in its practical applicability. In this section, we address these two drawbacks by creating a CWMR mixture algorithm, which mixes the proposed CWMR algorithms with different parameter settings and other regret-bounded algorithms.

Since the mean reversion trading idea is counterintuitive, it is difficult to provide a traditional regret bound[9]. Alternatively, we treat each CWMR algorithm with a specified parameter setting as one expert in a setting of multiple experts, which consists of at least one universal strategy (such as UP, EG, ONS, etc.). Then we adopt no-regret learning algorithms [Cesa-Bianchi and Lugosi 2006] to bound the whole system. In this article, we use on idea similar to the buy and hold idea of Cover [1991], Akcoglu

[8]In the short version [Li et al. 2011b], we solved it using the Matlab optimization toolbox, which costs much more time.
[9]Borodin et al. [2004] failed to provide a regret bound for the Anticor strategy, which also exploits the mean reversion idea.

et al. [2005] and Borodin et al. [2004], that is, we uniformly distribute the wealth among N experts, then let them run, and finally pool them together. It is worth noting that rather than pooling experts of the same class (Cover [1991] pools the CRP class and Borodin et al. [2004] pools the Anticor class), we allow experts from different classes. One can optionally use other expert learning algorithms, such as online gradient update and online Newton update [Das and Banerjee 2011].

To begin with, let us define a set of N experts $\mathcal{Q} = \left\{Q^1, \ldots, Q^N\right\}$, including CWMR experts and at least one universal strategy. Initially, each expert is assigned to equal wealth, that is, for convenience, $\mathbf{S}_0\left(Q^j, \mathbf{x}^n\right) = 1, j = 1, \ldots, N$. At the beginning of the t^{th} period, each expert j generates his/her portfolio $\mathbf{b}_t^j, j = 1, \ldots, N$. Then, the mixture algorithm weights all experts' portfolios according to their historical performance, that is, $\mathbf{b}_t = \frac{\sum_{j=1}^N \mathbf{b}_t^j \mathbf{S}_{t-1}\left(Q^j, \mathbf{x}^n\right)}{\sum_{j=1}^N \mathbf{S}_{t-1}(Q^j, \mathbf{x}^n)}$. An individual expert with high historical performance has a high impact on $\mathbf{b}_t$. After $\mathbf{x}_t$ is revealed, the mixture algorithm can update the cumulative wealth $\mathbf{S}_t$ and an individual expert can update its performance $\mathbf{S}_t\left(Q^j, \mathbf{x}^n\right)$. In summary, Algorithm 4 illustrates the general procedure of the proposed CWMR mixture algorithm.

Clearly, the total wealth achieved by the mixture algorithm after n trading periods is equivalent to the uniform weighted wealth of all experts (since each expert is assigned equal initial wealth), that is,

$$\mathbf{S}_n = \frac{1}{N} \sum_{j=1}^{N} \mathbf{S}_n\left(Q^j, \mathbf{x}^n\right). \tag{5}$$

Thus, the final cumulative wealth is affected by all experts, and expert j's contribution is determined by its final performance $\mathbf{S}_n\left(Q^j, \mathbf{x}^n\right)$.

Ideally, indexed by (ϕ, ϵ), we can choose CWMR experts such that they cover all possible parameter settings, thus eliminating their effects. However, the cost of combining all possible parameter settings is inhibitively high. To boost the computational efficiency, we can choose finite discrete dimensions of the parameters, that is, a specified number of (ϕ, ϵ) combinations. On the other hand, choosing other experts can be arbitrary. Typically, one can choose algorithms with different trading ideas, such that the mixture algorithm can face different market scenarios.

The selection of experts also trades off an individual expert's performance and its computational time. First, Equation (5) clearly shows that each expert contributes to the final cumulative wealth by its performance, thus, choosing a worse expert may lower the final performance. Second, the mixture's computation time is generally the summation of all experts' individual times. In other words, choosing an expert with long running time may affect the practical scalability.

We present the theoretical regret bound of the proposed mixture algorithm in Theorem 4.3 and further declare its universal property in Corollary 4.4. The proofs are presented in Appendix 6 and Appendix 6, respectively.

THEOREM 4.3. *Assume that the CWMR mixture algorithm P competes against a finite class of N experts $\mathcal{Q} = \left\{Q^1, \ldots, Q^N\right\}$, which contains at least one regret-bounded algorithm (assume Cover's UP here). Then the worst-case logarithmic wealth ratio with respect to best constant rebalanced portfolio is bounded as*

$$\sup_{\mathbf{x}^n} \sup_{B \in \Delta_m} \ln \frac{\mathbf{S}_n\left(B, \mathbf{x}^n\right)}{\mathbf{S}_n\left(P, \mathbf{x}^n\right)} \leq (m-1) \ln (n-1) + \ln N,$$

Algorithm 4: The proposed CWMR Mixture (CWMR-Mix) framework.

Input: $\mathcal{Q} = \{Q^1, \ldots, Q^N\}$: N specified experts; $\mathbf{x}_1^n$: Historical market sequence
Output: $\mathbf{S}_n$: Final cumulative wealth

Initialization: $\mathbf{S}_0 = 1$, $\mathbf{S}_0 \left(Q^j, \mathbf{x}^n\right) = 1, j = 1, \ldots, N$
for $t = 1, 2, \ldots, n$ **do**
  Generate portfolio $\mathbf{b}_t^j$ from expert $Q^j, j = 1, \ldots, N$;
  Combine experts' portfolios for the final portfolio: $\mathbf{b}_t = \frac{\sum_j \mathbf{S}_{t-1}\left(Q^j, \mathbf{x}^n\right)\mathbf{b}_t^j}{\sum_j \mathbf{S}_{t-1}(Q^j, \mathbf{x}^n)}$;
  Receive stock price relatives: $\mathbf{x}_t = (x_{t1}, \ldots, x_{tm})$;
  Update mixture's cumulative return: $\mathbf{S}_t = \mathbf{S}_{t-1} \times \left(\mathbf{b}_t^\top \mathbf{x}_t\right)$;
  Update experts' cumulative return: $\mathbf{S}_t \left(Q^j, \mathbf{x}^n\right) = \mathbf{S}_{t-1} \left(Q^j, \mathbf{x}^n\right) \times \left(\mathbf{b}_t^j \cdot \mathbf{x}_t\right)$;
end

where m denotes the number of stocks, n is the number of trading periods, and N denotes the number of experts.

COROLLARY 4.4. *The proposed CWMR mixture algorithm is a universal portfolio selection algorithm.*

Remark. Theorem 4.3 shows that the mixture algorithm's worst-case logarithmic wealth ratio is bounded with respect to any CRP strategy, or regret in the online portfolio selection community. Compared with UP's worst-case logarithmic wealth ratio, the mixture has the additional term, $\ln N$. However, since N is usually finite, it will not affect the mixture's universal property, as Corollary 4.4 shows. Since other regret-bounded algorithms (e.g., EG and ONS) have the same regret bound as UP, changing the regret-bounded algorithm does not affect the current regret bound. Nevertheless, the preceding regret bound is easy to understand, and provides a theoretical guarantee for the mixture algorithm, which asymptotically approaches the BCRP strategy. It is worth noting that although the CMWR mixture algorithm is universal, the universal property of CWMR itself is still an open question.

5. NUMERICAL EXPERIMENTS

5.1. Experimental Testbed on Real Data

In our empirical study, we focus on historical data in stock markets, which are easy to obtain, and hence available for other researchers. Data from other markets, such as currency and commodity markets, are either expensive or hard to obtain and process, and thus reduce the experimental reproducibility. In our empirical experiments, we employ eight real and diverse datasets[10] from stock markets and index markets as summarized in Table V.

The first dataset is the well-known NYSE dataset pioneered by Cover [1991] and followed by most subsequent researchers in the field of online portfolio selection. This dataset contains 5651 daily price relatives of 36 stocks[11] in the New York Stock Exchange (NYSE) for a 22-year period, ranging from July 3rd 1962 to December 31st 1984. We refer to this dataset as "NYSE (O)." For consistency, we further collected their latest data from January 1st 1985 to June 30th 2010, which last for 6431 trading days. We

[10] All datasets and their compositions are available at `http://www.cais.ntu.edu.sg/~libin/portfolios`.
[11] According to Helmbold et al. [1998], the dataset was originally collected by Hal Stern, and we do not know the criteria for choosing these stocks.

Table V. Summary of the Eight Real Datasets in our Numerical Experiments

Dataset	Market	Region	Time frame	Frequency	# Data Points	# Assets
NYSE(O)	Stock	US	Jul. 3rd 1962 - Dec. 31st 1984	Daily	5651	36
NYSE(N)	Stock	US	Jan. 1st 1985 - Jun. 30th 2010	Daily	6431	23
TSE	Stock	CA	Jan. 4th 1994 - Dec. 31st 1998	Daily	1259	88
MSCI	Index	Global	Apr. 1st 2006 - Mar. 31st 2010	Daily	1043	24
DJA	Stock	US	Aug. 1st 2011 - Aug. 5th 2011	Half-Minute	3900	30
NDX	Stock	US	Aug. 1st 2011 - Aug. 5th 2011	Half-Minute	3900	100
W-NYSE(O)	Stock	US	Jul. 3rd 1962 - Dec. 31st 1984	Weekly	1130	36
W-NYSE(N)	Stock	US	Jan. 1st 1985 - Jun. 30th 2010	Weekly	1286	23

denote the new dataset as "NYSE (N)".[12] This dataset consists of 23 stocks rather than the previous 36 stocks, owing to the amalgamation and bankruptcy of certain stocks. All price relatives are adjusted for splits and dividends, which is consistent with the NYSE (O) dataset.

The third dataset is the "TSE" dataset used by Borodin et al. [2004], which consists of 88 stocks[13] from the Toronto Stock Exchange (TSE) containing price relatives of 1259 trading days, ranging from January 4th 1994 to December 31st 1998. The fourth self-collected dataset, "MSCI", is a collection of global equity indices, which are the constituents of the MSCI World Index[14]. It contains 24 indices, which represent the equity markets of 24 countries across the world, and totally consists of 1043 trading days, ranging from April 1st 2006 to March 31st 2010.

The next the two datasets are high frequency intraday data collected from Interactive Brokers[15]. The fifth dataset, "DJA", contains 30 index composites from the Dow Jones Industrial Average, while the sixth dataset "NDX" contains 100 index composites from the NASDAQ-100. With time intervals of 30 seconds, both datasets contain 3900 data points, or totally 5 trading days, ranging from August 1st 2011 to August 5th 2011.

The final two datasets are derived from the preceding two NYSE datasets to represent weekly frequency. We calculate each entry of weekly data by multiplying five daily entries in the NYSE data, and name them "W-NSYE (O)" and "W-NYSE (N)". With the same number of assets, the weekly datasets contain 1130 and 1286 data points, respectively.

Unlike previous studies, the preceding testbed covers a much longer trading period from 1962 to 2011 and contains a much larger number of assets, from 23 to 100, with different markets, which enables us to examine how the proposed CWMR strategy performs under various market situations. The first three datasets are chosen to test CWMR's capability with stocks, while the MSCI dataset aims to test the proposed CWMR on global indices, which may be potentially applicable to "Fund on Fund" (FOF)[16]. The two high frequency datasets enable us to evaluate various algorithms on the field of high frequency trading. To the best of our knowledge, we are the first to evaluate online portfolio selection algorithms with high frequency data. Though the weekly datasets, which are artificially calculated, may deviate from real data, they

---

[12]The dataset before 2007 was collected by Gábor Gelencsér (`http://www.cs.bme.hu/~oti/portfolio`), we collected the remaining data starting from 2007 to 2010.

[13]This dataset was collected by Borodin et al. [2004] and we do not know the criteria for selecting these stocks.

[14]The constituents of MSCI World Index can be found on MSCI Barra (`http://www.mscibarra.com`), accessed on May 28, 2010.

[15]The dataset is collected based on IB Student Trading Lab (`http://www.interactivebrokers.com`). We consider only intraday data and ignore the close-open gaps.

[16]Though many indices are tradable through exchange traded funds (ETFs), not every index is.

can test the algorithms over longer trading intervals. Finally, as a remark, although numerically tested on stock markets, the proposed CWMR could be generally applied to any type of financial market.

5.2. Experimental Setup and Metrics

In our experiments, we implemented the proposed CWMR approaches, namely, CWMR-Var and CWMR-Stdev. We denote the deterministic versions by CWMR-Var and CWMR-Stdev and the stochastic ones by CWMR-Var-s and CWMR-Stdev-s. For the latter, we repeat tests 100 times and report the average values.

Regarding the parameter settings, there are two key parameters in the proposed CWMR algorithms. One is confidence parameter ϕ and the other is sensitivity parameter ϵ. Roughly speaking, the best parameters are often dataset dependent. In the experiments, we simply set them empirically without tuning. In particular, we set the sensitivity parameter ϵ to 0.5, and set the confidence parameter ϕ to 2.0, or equivalently 95% confidence level, in both CWMR-Var(-s) and CWMR-Stdev(-s). As we will examine the parameter sensitivity in Section 5.4.3, the proposed CWMR algorithm is robust with respect to different parameter settings and our choices are not always the best.

Moreover, we also implemented the proposed CWMR Mixture approach. Ideally, the proposed CWMR Mixture approach does not contain any parameter, and tends to be robust. To make it computational efficient, we implemented the discrete version instead. To be specific, we chose two parameters, that is, ϕ and ϵ, in a lattice. For confidence level parameter ϕ, we chose 1.28, 1.64, 1.95, and 2.57, or equivalently in confidence level, 80%, 90%, 95%, and 99%. For mean reversion parameter ϵ, we chose from 0 to 0.8 in an interval of 0.2. We chose EG as the additional regret-bounded algorithm for its computational time (there is no essential difference arising from which universal algorithm is chosen), which is required for regret guarantee. Thus, we have totally 21 experts, including 20 CWMR experts each parameterized with one knot of the lattice plus EG algorithm. We refer to the mixture approaches as CWMR-Var-m and CWMR-Stdev-m.

The general criterion for evaluating a trading strategy measure is its investment return and risk. To measure investment return, we adopt the most common metric, *cumulative wealth* at the end of n trading periods, that is, $\mathbf{S}_n$. Another equivalent criterion is *annualized percentage yield* (APY)[17]. The higher the cumulative wealth or annualized percentage yield of a trading strategy, the better is the absolute return. In addition to absolute return, we are also interested in a strategy's *risk* and corresponding *risk-adjusted return*. Thus, we adopt *annualized standard deviation* (STD) of daily excess returns and *annualized Sharpe ratio* (SR)[18] [Sharpe 1963, 1994] to compare strategies' volatility risk and volatility risk-adjusted return, respectively. Besides volatility risk, we further measure a strategy's drawdown risk and drawdown risk-adjusted return by comparing *maximum drawdown* (MDD) and *Calmar Ratio* (CR) [Magdon-Ismail and Atiya 2004]. The lower the *annualized standard deviation* or *maximum drawdown* of a strategy, the less the risk. In a summary, the higher the *annualized Sharpe Ratio* or *Calmar Ratio* of a strategy, the better the risk-adjusted performance. The performance metrics are summarized in Table VI and details can be found in the online appendix.

[17]APY $= \sqrt[y]{\mathbf{S}_n} - 1$, where y represents the number of years corresponding to n trading periods, and we assume 252 trading days in one year.

[18]SR $=$ (APY $- R_f)/\sigma_p$, where R_f is the risk-free return (typically the return of Treasury bills, fixed at 4% in this work), and σ_p is the annualized standard deviation of daily excess returns.

Table VI. Summary of the Performance Metrics Used in the Numerical Experiments

Criteria	Performance Metrics	
Absolute return	Cumulative wealth ($\mathbf{S}_n$)*	Annualized percentage yield (APY)
Risk	Annualized standard deviation (STD)	Maximum drawdown (MDD)
Risk-adjusted return	Annualized Sharpe ratio (SR)	Calmar ratio (CR)

*This is the primary metric across the empirical evaluations.

Finally, due to the various frequencies of these datasets, CWMR algorithms exhibit different behaviors across the datasets. For example, the annualized Sharpe ratios of CWMR on the high frequency datasets are astronomically high, which makes the comparison inconsistent; on the other hand, with such a high frequency, the transaction cost becomes crucially important for trading, as all methods approach zeros with the existing transaction cost model [Blum and Kalai 1999; Borodin et al. 2004]. Moreover, since we have a testbed of eight datasets, which is the largest as far as we know, for concise presentation, we move some similar results on high frequency and weekly datasets to an online appendix. Thus, to compare consistently and present clearly, we mainly focus on daily datasets, and we only provide some main results for high frequency and weekly data.

5.3. Comparison Methods

In our experiments, we compare the proposed algorithms with a number of existing strategies as described in Section 3. Here we summarize these algorithms, whose parameters are set according to the recommendations from their respective studies[19].

(1) Market. Market (uniform BAH) strategy.
(2) Best-stock. Best stock in a market, which is obviously a hindsight strategy.
(3) BCRP. Best constant rebalanced portfolios strategy in hindsight.
(4) UP. Cover's Universal Portfolios implemented according to Kalai and Vempala [2002].
(5) EG. Exponential Gradient with parameter $\eta = 0.05$, suggested by Helmbold et al. [1998].
(6) ONS. Online Newton Step with parameter setting as suggested by Agarwal et al. [2006], that is, $\eta = 0$, $\beta = 1$, and $\gamma = 1/8$.
(7) SP. Switching Portfolios with parameter $\gamma = 1/4$, as suggested by Singer [1997].
(8) GRW. Gaussian Random Walk strategy with parameter $\sigma = 0.00005$, recommended by Levina and Shafer [2008].
(9) M0. A prediction-based algorithm with parameter $\beta = 0.5$, suggested by Borodin et al. [2000].
(10) Anticor. BAH_{30}(Anticor(Anticor))[20] as a variant of Anticor to smooth the performance volatility, which is the best solution proposed in Borodin et al. [2004].
(11) B^{K}. Nonparametric kernel-based moving window strategy with $W = 5$, $L = 10$, and $c = 1.0$ for daily datasets, which has the best empirical performance according to Györfi et al. [2006] and $c = 0.01$ for high frequency datasets, because of its low volatility.
(12) B^{NN}. Nonparametric nearest neighbor-based strategy with parameter $W = 5$, $L = 10$, and $p_\ell = 0.02 + 0.5\frac{\ell-1}{L-1}$, as suggested by Györfi et al. [2008].
(13) CORN. Correlation-driven nonparametric learning approach with parameter $W = 5$ and $\rho = 0.1$, as suggested by Li et al. [2011a].

[19]We can tune the parameters of competitors for better performance, but that is beyond the scope of this article.
[20]In the short version [Li et al. 2011b], we use BAH_{30}(Anticor), which performs slightly worse.

Table VII. Cumulative Wealth Achieved by Various Trading Strategies on the Eight Datasets. The Top Two Achievements in Each Dataset are Highlighted in Bold

Methods	NYSE (O)	NYSE (N)	TSE	MSCI	DJA	NDX	W-NYSE (O)	W-NYSE (N)
Market	14.50	18.06	1.61	0.91	0.93	0.92	14.43	18.17
Best-stock	54.14	83.51	6.28	1.50	0.98	1.01	53.80	83.09
BCRP	250.60	120.32	6.78	1.51	0.98	1.01	125.20	98.30
UP	26.68	31.49	1.60	0.92	0.93	0.92	23.95	27.83
EG	27.09	31.00	1.59	0.93	0.93	0.92	23.61	27.46
ONS	109.19	21.59	1.62	0.86	0.93	0.95	101.36	22.88
SP	27.08	31.55	1.60	0.93	0.93	0.92	23.89	27.89
GRW	27.73	30.45	1.61	0.93	0.93	0.92	23.86	27.37
M0	113.50	40.94	1.26	0.92	0.90	0.90	48.95	30.45
Anticor	2.41E+08	**6.21E+06**	39.36	3.22	1.07	1.19	1.93E+03	1.61E+03
B^K	1.08E+09	4.64E+03	1.62	2.64	0.97	0.93	28.87	31.05
B^{NN}	3.35E+11	6.80E+04	2.27	13.47	1.24	1.21	166.11	35.04
CORN	1.48E+13	5.37E+05	3.56	**26.10**	1.36	1.75	102.22	246.14
CWMR-Var	6.51E+15	1.44E+06	328.61	17.27	**2.08**	**3.83**	**3.49E+05**	**2.16E+04**
CWMR-Stdev	6.49E+15	1.41E+06	332.62	**17.28**	**2.09**	**3.82**	**3.48E+05**	**2.14E+04**
CWMR-Var-s	2.19E+15	7.47E+05	281.26	12.90	1.97	3.60	1.42E+05	1.29E+04
CWMR-Stdev-s	2.13E+15	7.38E+05	290.08	12.70	1.96	3.61	1.43E+05	1.17E+04
CWMR-Var-m	**8.00E+15**	**2.03E+06**	**358.12**	15.92	2.03	3.70	2.36E+05	1.65E+04
CWMR-Stdev-m	**7.95E+15**	2.01E+06	**360.70**	15.93	2.04	3.69	2.39E+05	1.69E+04

5.4. Experimental Results

5.4.1. Experiment 1. Evaluation of Cumulative Wealth. The main experiment is to evaluate the cumulative wealth at the end of trading periods, and the results are illustrated in Table VII.

From the results, we have some observations. First, the cumulative wealth achieved by CWMR-Var and CWMR-Stdev are similar, since they are two different solutions for the same optimization problem. Meanwhile, the performance achieved by the stochastic version is always smaller than that of the deterministic version, which validates the analysis in Section 4.4. Second, in most datasets, the cumulative wealth achieved by the proposed CWMR algorithms significantly surpasses all existing competitors, including the Anticor algorithm, which adopts the same mean reversion trading idea. This verifies that the mean reversion trading idea does exist in various financial markets, and the learning idea borrowed from CW can effectively exploit such information. Moreover, the proposed CWMR mixture versions, equipped with a regret bound, always achieve good results. To the best of our knowledge, no previous work has ever claimed such a high cumulative wealth, especially on the benchmark dataset "NYSE (O)."

It is worth discussing the results with different frequencies, as we first introduce the high frequency data in the evaluation. On the two high frequency datasets consisting of five trading days, although the markets suffered from a significant a drop, the proposed CWMR still accumulated high return. In fact, on a daily basis, the returns on high frequency data is much higher than those on low frequency data, that is, daily and weekly datasets. In a real market with commissions, ordinary people cannot obtain such a high return; while professional institutions may be able to obtain these returns. Note that we only consider the mid-prices of the bid ask spread, while in the high frequency scenario, the spread will have a considerable impact that is much higher than the brokerage commission. The results are still interesting to us, as they reflect that the short term market follows mean reversion and CWMR can efficiently exploit such movements. On the other hand, though the market returns on the two weekly NYSE datasets and their corresponding daily datasets are almost the same,

Table VIII. Statistical t-test of the Performance Achieved by the Proposed CWMR (CWMR-Stdev) Algorithm on the Eight Datasets

Statistics	NYSE(O)	NYSE(N)	TSE	MSCI	DJA	NDX	W-NYSE(O)	W-NYSE(N)
Size	5651	6431	1259	1043	3900	3900	1130	1286
MER (CWMR)	0.0070	0.0027	0.0057	0.0030	0.0002	0.0003	0.0127	0.0103
MER (Market)	0.0005	0.0005	0.0004	0.0000	-0.0000	-0.0000	0.0026	0.0025
Winning Ratio	56.17%	52.08%	56.00%	59.44%	63.49%	64.49%	59.29%	56.45%
α	0.0064	0.0021	0.0051	0.0030	0.0002	0.0004	0.0096	0.0072
β	1.2139	1.1325	1.5139	1.1161	1.1453	1.0502	1.2088	1.2876
t-statistics	15.9510	5.9496	3.9190	6.4078	17.8418	18.2656	7.5089	3.7249
p-value	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0001

Notes 1: MER denotes mean excess return. Excess return equals daily return of a strategy minus daily risk-free return.

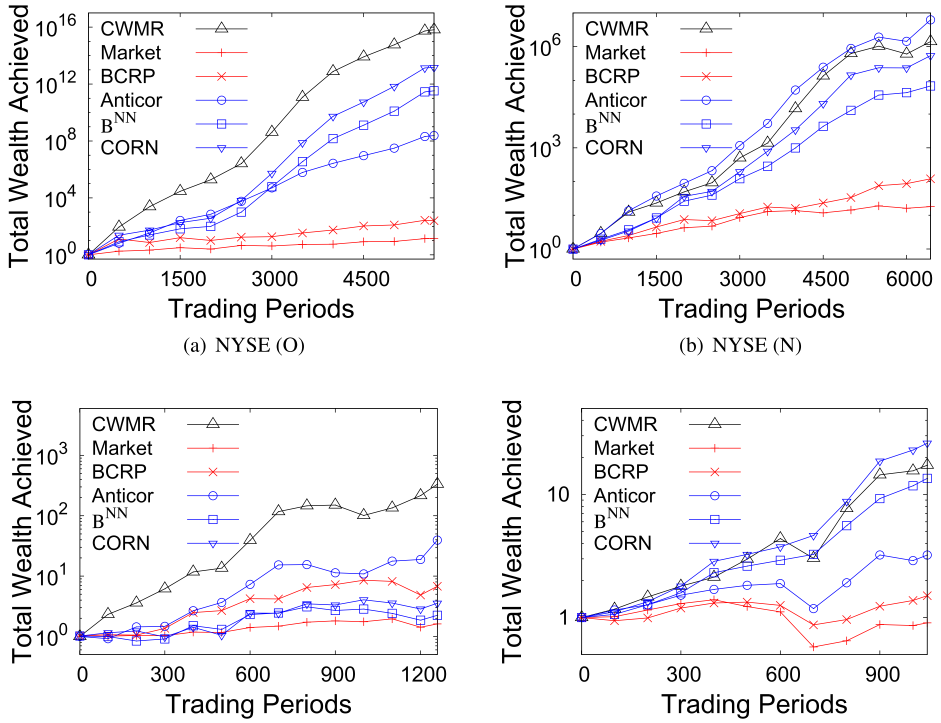
online portfolio selection algorithms perform divergently. Some state-of-the-art, strategies such as the three nonparametric learning algorithms, decrease to normal. CWMR algorithms also decrease, but the results are still high compared to the benchmarks and all other competitors. Finally, as we have analyzed in Section 4.1, the trading frequency seems to dominate the results of CMWR, that is, empirically the higher the trading frequency, the higher the return. Interestingly, Hazan and Kale [2009] come to a similar conclusion for the ONS algorithm, "one expects to see improving performance of our algorithm as the trading frequency increases."

Even though the results are somehow beyond imagination, we are also interested in whether the results can be generated by simple luck. To check this possibility, we conducted widely accepted statistical tests among practitioners [Grinold and Kahn 1999] (also refer to the online appendix). Table VIII summarizes the statistical test results. Briefly speaking, there exists almost no chance that the amazing cumulative wealth is generated by simple luck. Here, we claim again, even in the theoretical "perfect market" without transaction costs and other practical issues, no existing work has ever declared such high performance.

Another virtue of the proposed approach is its persistence during the entire period. To better see this, we plot the wealth curves achieved by the proposed CWMR algorithm (CWMR-Stdev), state-of-the-art algorithms (Anticor, B^{NN}, and CORN), plus two benchmarks (Market and BCRP). As other versions (CWMR-Var, CWMR-Var-Mix, and CWMR-Stdev-Mix) perform similarly to CWMR-Stdev, we ignore them in the figures. As shown in Figure 1, the proposed CWMR algorithms consistently get ahead of the state-of-the-art techniques, not to mention the benchmarks. This consistency again confirms the efficacy of the proposed CWMR algorithms.

5.4.2. Experiment 2. Evaluation of Risk-Adjusted Return. Besides the cumulative wealth metric, we also conduct experiments to evaluate the risk-adjusted return (both volatility risk-adjusted and drawdown risk-adjusted) achieved by the proposed algorithms.

Figure 2(a) and Figure 2(b) demonstrate the volatility risk measured by annualized standard deviation and corresponding risk-adjusted return measured by annualized Sharpe ratio (SR). Along with the results achieved by the proposed CWMR, we plot the results of the state-of-the-art algorithms (Anticor, B^{NN}, and CORN) and two benchmarks (Market and BCRP). Figure 2(a) shows that higher return often associates with higher risk, that is, volatility risks achieved by CWMR on the four datasets are higher than the benchmarks and competitors. However, Figure 2(b) clearly shows that in most cases the volatility risk-adjusted returns achieved by the proposed CWMR approach are higher than that of the benchmarks. This experiment again validates that the proposed CWMR approach is preferable in terms not only of the cumulative return metric, but also the volatility risk-adjusted return metric.



(a) NYSE (O)

(b) NYSE (N)

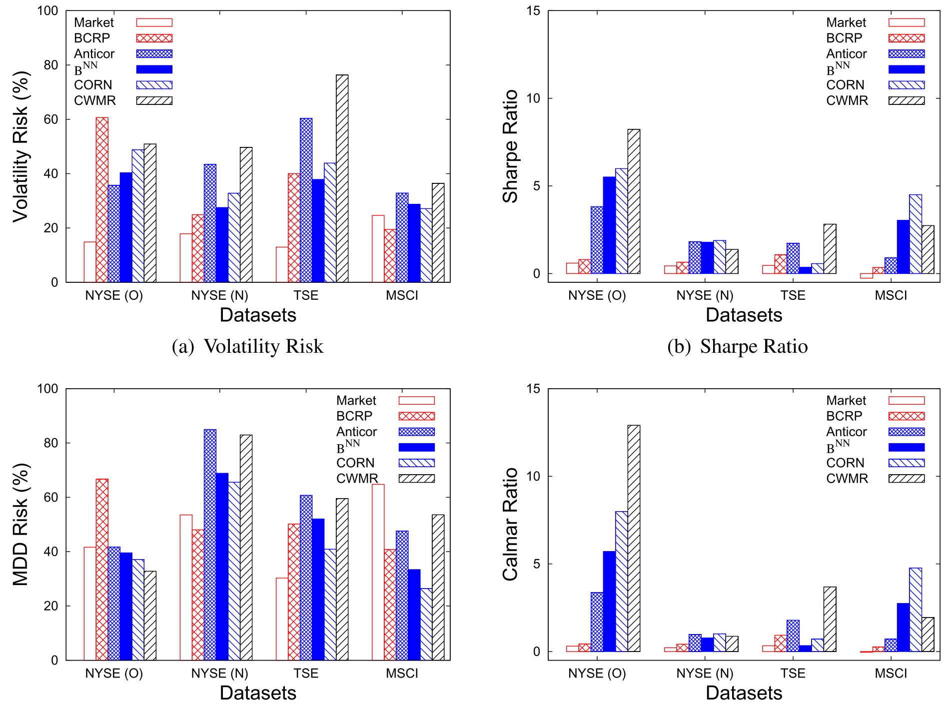
(c) TSE

(d) MSCI

Fig. 1. Trend of cumulative wealth achieved by various strategies during the entire period on the four daily datasets.

Similarly, Figure 2(c) and Figure 2(d) show the results on the drawdown risk measured by maximum drawdown (MDD) and drawdown risk-adjusted return in terms of Calmar ratio (CR) associated with some benchmarks and competitors. In Figure 2(c), the MDDs achieved by the proposed approach are not always the best, but are generally modest. For example, in the NYSE (O) datasets, CWMR almost achieves the least MDDs, which indicates the least drawdown risk. The drawdown risk-adjusted returns shown in Figure 2(d) clearly show the superiority of the proposed CWMR compared with the plotted benchmarks. These experiments validate superiority with respect to the drawdown risk-adjusted return metric, and again, they corroborate the effectiveness of the proposed CWMR approach.

5.4.3. Experiment 3. Evaluation of Parameter Sensitivity. The proposed CWMR approaches contain two parameters, that is, the confidence parameter ϕ and the mean reversion sensitivity parameter ϵ. Throughout the algorithm, the mean reversion sensitivity parameter decisively influences the final performance, that is, the smaller the mean reversion sensitivity, the better the final cumulative wealth. Figure 3 depicts its robustness with respect to the mean reversion sensitivity parameter, plus the final cumulative wealth achieved by Market and BCRP. The results first verify our preceding suspicion about the effect of the mean reversion sensitivity, that is, the final cumulative wealth increases as the sensitivity parameter decreases and becomes stable after crossing a data-dependent threshold. The results again verify that the mean reversion trading idea works in financial markets and the proposed CWMR algorithm can successfully exploit it, which generates significant final cumulative wealth, outperforming



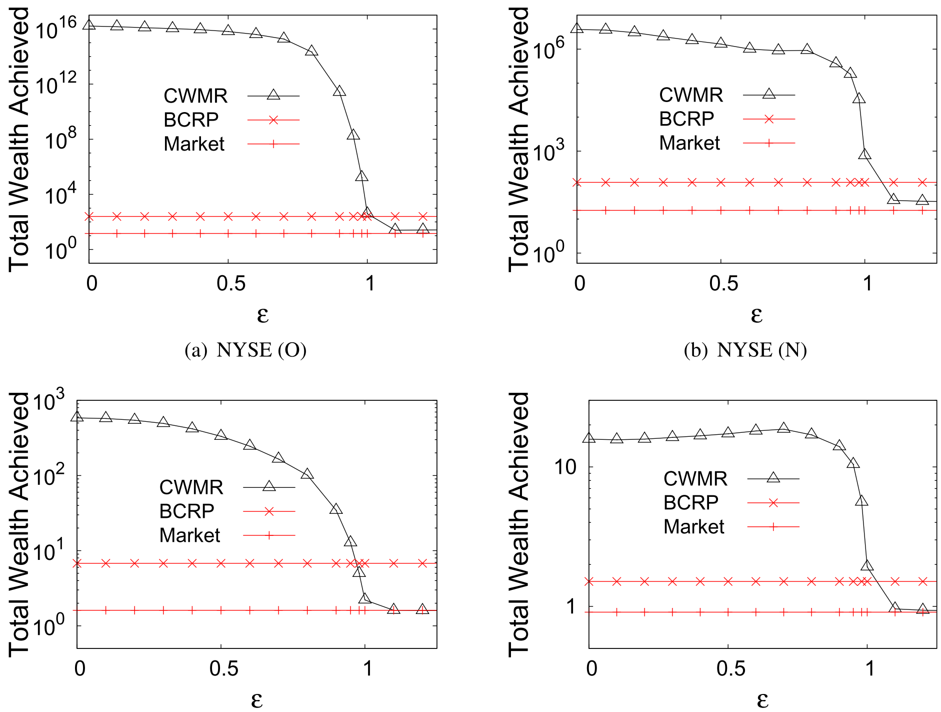
(a) Volatility Risk (b) Sharpe Ratio

(c) Maximum Drawdown Risk (d) Calmar Ratio

Fig. 2. Volatility analysis (Standard Deviation and Sharpe Ratio) and drawdown analysis (Maximum Drawdown and Calmar Ratio) of various strategies on the daily stock datasets. The results of CWMR are the rightmost bar on each dataset.

Market and BCRP strategy. Moreover, as analyzed in Section 4.4, CWMR degrades to a uniform CRP strategy when ϵ is larger than 1. Finally, it seems that after ϵ falls below certain critical values, the final wealth would be stable, which means the mean reversion trading idea has been completely exploited. Needless to say, our empirical parameter setting of $\epsilon = 0.5$ is not the best one. However, even under this setting, the proposed CWMR still significantly surpasses existing approaches.

5.4.4. Experiment 4. Evaluation of Practical Issues. In reality, an important and unavoidable issue is *transaction cost*. Generally, there are two ways to handle the transaction cost. The first way, adopted in this study, is to include transaction costs during portfolio rebalancing, while the portfolio selection process doesn't take it into consideration. The second way is to include the transaction costs in the portfolio selection process [Györfi and Vajda 2008; Györfi et al. 2012]. In this work, we adopt the *proportional transaction costs* model proposed by Blum and Kalai [1999] and Borodin et al. [2004]. With this model, rebalancing a portfolio incurs a transaction cost on every buy and sell operation, based upon a transaction cost rate $\gamma \in (0, 1)$. At the beginning of the t^{th} trading period, portfolio manager rebalances the portfolio from the previous closing price-adjusted portfolio $\hat{\mathbf{b}}_{t-1}$ to a new portfolio $\mathbf{b}_t$, incurring a transaction cost of $\frac{\gamma}{2} \times \sum_i |b_{t,i} - \hat{b}_{t-1,i}|$, where the initial portfolio $\hat{\mathbf{b}}_0$ is $(0, \dots, 0)$. To the best of our knowledge, this model cannot work for high frequency data, since even a small rate will cause all methods to approach zero.



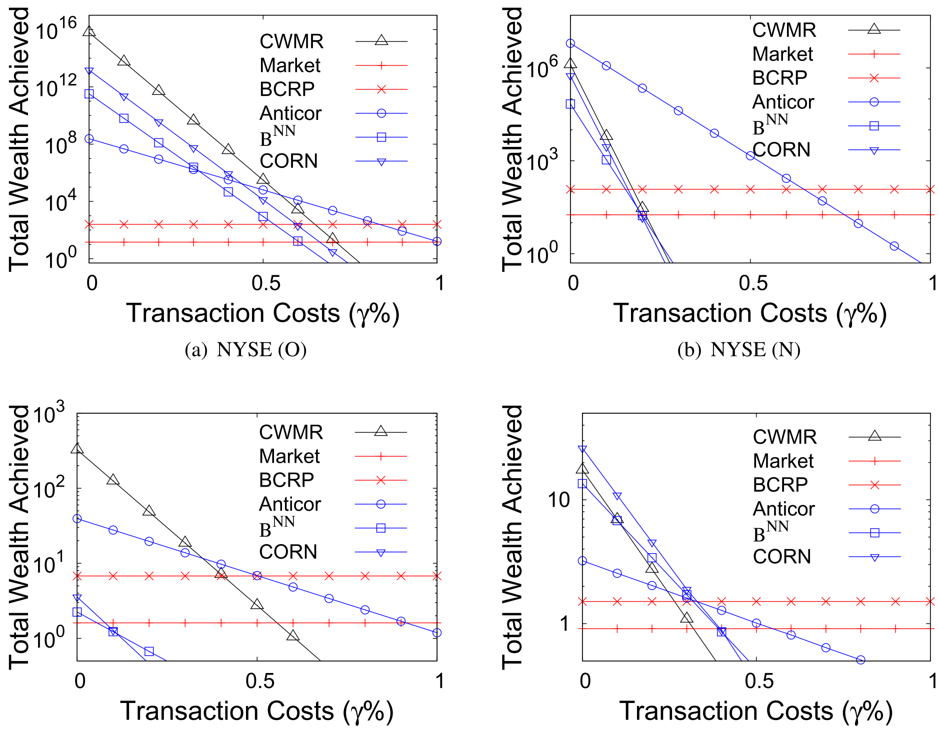
(a) NYSE (O) (b) NYSE (N) (c) TSE (d) MSCI

Fig. 3. Parameter Sensitivity of the total wealth achieved by CWMR-Stdev with respect to the mean reversion sensitivity parameter ϵ on the four daily datasets.

Figure 4 shows the results on the four daily datasets with varying transaction costs from 0% to 1%, plus the cumulative wealth achieved by two benchmarks (Market and BCRP) and the state-of-the-art strategies (Anticor, B^{NN}, and CORN). We observe that the performance with transaction costs is market dependent. In most cases, especially with small rates, CWMR can outperform the state-of-the-art algorithms. In other cases, though both are powered by mean reversion, CWMR underperforms Anticor, showing that aggressiveness results in more transaction costs. Nevertheless, compared with the benchmarks, the results clearly demonstrate that on all datasets, the performance is considerably robust with respect to the transaction costs[21], where the break-even rates range from 0.2% to 0.7%. Thus, the proposed CWMR can withstand moderate transaction costs even though we do not explicitly tackle them during the portfolio selection process.

Another practical issue in portfolio selection is *margin buying*, which allows the portfolio managers to buy securities with cash borrowed from security brokers. Following previous studies [Agarwal et al. 2006; Cover 1991; Helmbold et al. 1998], here the margin setting is assumed to be 50% down and 50% loan, at an annual interest rate of 6%, or at a daily interest rate of $c = 0.000238$. Thus, for each asset i, a new asset named *Margin Component* is generated with its price relative equal to $2{*}x_{ti}-1-c$. By adding Margin Component, we magnify both the potential profit and loss of a trading

[21] For example, for US equities and options, Interactive Brokers (`https://www.interactivebrokers.com`) charges US$ 0.005 per share. Since the average price of 30 composite stocks in Dow Jones Industrial Average was around US$50 at the end of August 2011, the commission is around 0.01% of trade value.



(a) NYSE (O)

(b) NYSE (N)

(c) TSE

(d) MSCI

Fig. 4. Robustness of the total wealth achieved by CWMR with respect to transaction cost rate (γ).

strategy on the ith asset. Table IX depicts the cumulative wealth achieved by various strategies when margin buying is allowed. We do not list CWMR-Var and its mixture version since their performance is similar to CWMR-Stdev and corresponding mixture version. As expected, the performance with margin buying in most cases, is significantly improved. Moreover, in the case of margin buying, the proposed approaches still surpass the state-of-the-art algorithms in most cases. Note that although leveraging the capital (margin buying) does improve the cumulative returns here, it does not necessarily improve the risk adjusted performance like the Sharpe ratio (when leveraging is free), or may actually decrease their risk adjusted performance, as a price is paid for leveraging. Nevertheless, this validates the efficacy of the proposed algorithms in the case of margin buying.

5.4.5. Experiment 5. Evaluation of Computational Time. The proposed algorithm achieves significant improvement over existing approaches, it is also computational efficient, as we have analyzed in Section 4.4. Table X shows the total computational time of the proposed CWMR (CWMR-Stdev and CWMR-Stdev-m) and four state of the art strategies (Anticor, B^K, B^{NN}, and CORN), whose performance is comparable to the proposed approach, on the eight datasets. Even the time costs per trading day of the competitors are acceptable on daily and week datasets, their costs in high frequency datasets are generally too expensive. On the contrary, CWMR computes much more efficiently than its competitors, especially in the domain of high frequency trading [Aldridge 2010], where transactions may occur in fractions of a second. Therefore, the computational efficiency confirms the real-world large-scale applicability of the proposed algorithms.

Table IX. Cumulative Wealth Achieved by Various Strategies on the Daily Stock Datasets with/without Margin Loans (ML)

Algorithm	NYSE (O)		NYSE (N)		TSE		MSCI	
	No ML	with ML	No ML	with ML	No ML	with ML	No ML	with ML
Market	14.5	15.75	18.06	17.68	1.61	1.71	0.91	0.69
Best-stock	54.14	54.14	83.51	173.18	6.28	10.53	1.50	1.50
BCRP	250.6	3755.09	120.32	893.63	6.78	21.23	1.51	1.54
UP	26.68	62.99	31.49	57.03	1.60	1.69	0.92	0.71
EG	27.09	63.28	31.00	55.55	1.59	1.68	0.93	0.72
ONS	09.19	517.21	21.59	228.37	1.62	0.88	0.86	0.33
Anticor	2.41E+08	1.05E+15	**6.21E+06**	**5.41E+09**	39.36	18.69	3.22	3.40
B^K	1.08E+09	6.29E+15	4.64E+03	3.72E+06	1.62	1.53	2.64	6.56
B^{NN}	3.35E+11	3.17E+20	6.80E+04	5.58E+07	2.27	2.17	14.47	**150.49**
CORN	1.48E+13	1.10E+22	5.37E+05	**1.72E+09**	3.56	5.00	**26.10**	**853.08**
CWMR-Stdev	**6.49E+15**	**6.59E+25**	1.41E+06	7.31E+07	**332.62**	**172.36**	**17.28**	76.29
CWMR-Stdev-m	**6.68E+15**	**1.73E+27**	**1.69E+06**	5.16E+08	**303.34**	**306.47**	13.69	65.19

Top two achievements on each datasets are highlighted.

Table X. Computational Time Costs on the Real Datasets (Seconds)

Time	NYSE(O)	NYSE(N)	TSE	MSCI	DJA	NDX	W-NYSE(O)	W-NYSE(N)
Anticor	2.57E+03	1.93E+03	2.15E+03	306	1.44E+03	8.64E+03	494	363
B^K	7.89E+04	5.78E+04	6.35E+03	2.60E+03	2.51E+04	1.39E+05	1.33E+03	1.31E+03
B^{NN}	4.93E+04	3.39E+04	1.32E+03	2.55E+03	6.64E+04	1.20E+06	3.33E+03	2.62E+03
CORN	8.78E+03	1.03E+04	1.59E+03	457	1.08E+04	9.36E+04	550	567
CWMR	12	11	3	1	6	19	1	1
CWMR-m	67	54	54	8	35	211	11	9

5.5. Discussion and Thread of Validity

5.5.1. On Model Assumptions. Any statement about such encouraging empirical results would be incomplete without acknowledging the simplified assumptions made in Section 2. To recall, we had made several assumptions regarding transaction cost, market liquidity and market impact, which would affect the practical deployment of the proposed algorithms.

The first assumption is that no transaction cost exists. In Section 5.4.4, we examined the effect of varying transaction costs, and the results show that the proposed algorithm can withstand moderate transaction costs. Currently, with the wide-spread adoption of electronic communication networks (ECNs) and multilateral trading facilities (MTFs) on financial markets, various online trading brokers charge very small transaction costs, especially for large institutional investors. They also use a flat-rate[22], based on the volume threshold one reaches. Such measures can facilitate portfolio managers to lower their transaction costs.

The second assumption is that the financial market is liquid and one can buy and sell any quantity at the quoted price. In practice, low market liquidity results in a large *bid-ask spread*—the gap between prices quoted for an immediate buy and an immediate sell. As a result, the execution of orders may incur a discrepancy between the prices sent by an algorithm and the prices actually executed. Moreover, stocks are often traded in multiples of a *lot*, which is a standard trading unit containing a number of stock shares. In this situation, the quantity of stocks may not be arbitrary divisible. In the experiments, we have tried to minimize the effect of market liquidity

[22] For example, for US equities and options, E*Trade (`http://www.etrade.com`, accessed on 16 March 2011.) charges only \$9.99 for \$50000+ or 30+ stocks per quarter.

by choosing stocks that have large market capitalization, which usually have small bid-ask spreads and discrepancies, and thus have a high market liquidity.

The final assumption is that the portfolio strategy would have no impact on the market, that is, the stock market will not be affected by any trading algorithm. In practice, the impact can be neglected if the cumulative wealth is not too large. However, as the experimental results show, the return generated by CWMR increases astronomically, which would inevitably impact the market. One simple way to handle this issue is to scale back the portfolio, as done by many quantitative funds. Moreover, the emerging algorithmic trading techniques, which slice a big order into multiple smaller orders and schedule these smaller orders to minimize the market impact, can significantly decrease the market impact of an algorithm.

Here, we emphasize again that this study assumes a perfect market, which is consistent with previous studies in the literature. It is important to note that even in such a perfect financial market, no algorithm has ever claimed such a high performance. Though it is common investment knowledge that past performance may not be a reliable indicator of future performance, such high performance does provide us confidence that the proposed CWMR algorithms may work well in unseen future markets.

5.5.2. On Back-Tests. Back-tests in the historical markets may suffer from "data-snooping bias." One common data-snooping bias is the dataset selection issue. On the one hand, we selected the two datasets, that is, NYSE (O) and TSE, based on previous studies, without consideration to the proposed approach. On the other hand, we developed CWMR algorithms based solely on the NYSE (O) dataset, and the other six datasets (NYSE (N), MSCI, DJA, NDX, W-NYSE(O) and W-NYSE(N) datasets) were obtained after the algorithm was fully developed. However, even though we are cautious about this dataset selection issue, it may still appear in the experiments, especially for the two datasets with relatively long histories, that is, NYSE (O) and NYSE (N). The NYSE (O) dataset, pioneered by Cover [1991] and followed by other researchers, has become one standard dataset in the online portfolio selection community. Since it contains 36 large cap NYSE stocks that survived in hindsight 22 years, this dataset suffers from extreme survival bias. Nevertheless, it is still useful to compare the performance among algorithms as done by all previous studies. The NYSE (N) dataset, as a continuation of NYSE (O), contains 23 assets that have survived from the previous 36 stocks for another 25 years. Therefore, it becomes even worse than the NYSE (O) dataset in terms of survival bias. In a word, even though the experimental results on these datasets clearly show the effectiveness of the proposed CWMR algorithms, their benefits on datasets that do not exhibit survival bias may be tempered.

Besides the survival bias, the lengths of the datasets also challenges the applicability of the proposed method. In finance and econometrics [Bondt and Thaler 1985, 1987; Jegadeesh 1991; Poterba and Summers 1988], it has been observed and demonstrated that the mean reverting phenomenon does exist in a long term, e.g., several years or decades, and thus may have no value to common investors. Our experiments validate the existence of the phenomenon, and the durations of our daily/weekly data are quite long, which is consistent with existing finance studies. However, the hypothesis of mean reverting may not persist all the time in such long durations, and as a result, the applicability of the proposed method may be challenged.

Another common bias is the asset selection issue. Three of the eight datasets (NYSE (O), W-NYSE (O), and TSE) are collected by others, and to the best of our knowledge, their assets are mainly the largest blue-chip stocks in their respective markets. We collected NYSE (N) (also W-NYSE (N)) ourselves as a continuation of NYSE (O) datasets, which again contain several of the largest survival stocks in NYSE (O). The remaining three self-collected datasets (MSCI, DJA, and NDX) were chosen according to the market indices in their respective markets. We tried to avoid the asset selection bias by arbitrarily choosing the representative stocks in their respective markets, which usually have large capitalization and high liquidity. Moreover, investing in these largest assets may reduce the market impact caused by any strategy.

6. CONCLUSION

In this article, we propose a novel online portfolio selection strategy named Confidence Weighted Mean Reversion (CWMR), which effectively learns portfolios by exploiting the mean reversion property in financial markets. The update schemes for the proposed algorithms are obtained by solving two optimization problems, taking into account the first and second order information of a portfolio vector, which goes beyond the existing approaches that usually only consider the first order information. The extended mixture version has a theoretic regret bound and is a universal portfolio selection method. Empirically, the proposed approach beats a number of competing state-of-the-art approaches on various up-to-date datasets collected from real markets.

In the future, we plan to study in detail the cause behind the existence of the mean reversion property in the financial markets. This will help us to further understand the nature of these markets. Second, we will develop more effective algorithms to improve the performance in the presence of high transaction costs. We also intend to explore the possibility of combining both the trend following and mean reversion principles to provide a more practically effective solution for the online portfolio selection tasks. Finally, we note that an interesting future direction is to extend our analysis for long-short only portfolios.

APPENDIXES

A. PROOF OF PROPOSITION 4.1

PROOF. Since considering the non-negativity constraint introduces too much complexity, we first relax the optimization problem without considering it, and later we will project the solution to the simplex domain to obtain the required vector.

The Lagrangian for optimization Problem (3) is,

$$\begin{aligned}\mathcal{L} =& \frac{1}{2}\left(\log\left(\frac{\det\Sigma_t}{\det\Sigma}\right) + \mathrm{Tr}(\Sigma_t^{-1}\Sigma) + (\mu_t - \mu)^\top \Sigma_t^{-1} (\mu_t - \mu)\right) \\ &+\lambda(\phi \mathbf{x}_t^\top \Sigma \mathbf{x}_t + \mu^\top \mathbf{x}_t - \epsilon) + \eta(\mu^\top \mathbf{1} - 1).\end{aligned}$$

Taking the derivative of the Lagrangian with respect to μ and setting it to zero, we can get the update of μ_{t+1},

$$0 = \frac{\partial \mathcal{L}}{\partial \mu} = \Sigma_t^{-1}(\mu - \mu_t) + \lambda \mathbf{x}_t + \eta \mathbf{1} \quad \Longrightarrow \quad \mu_{t+1} = \mu_t - \Sigma_t (\lambda \mathbf{x}_t + \eta \mathbf{1}), \tag{6}$$

where Σ_t is assumed to be nonsingular. Multiplying both sides of the update by $\mathbf{1}^\top$, we can get η,

$$1 = 1 - \mathbf{1}^\top \Sigma_t (\lambda \mathbf{x}_t + \eta \mathbf{1}) \quad \Longrightarrow \quad \eta = -\lambda \bar{x}_t, \tag{7}$$

where $\bar{x}_t = \frac{\mathbf{1}^\top \Sigma_t \mathbf{x}_t}{\mathbf{1}^\top \Sigma_t \mathbf{1}}$ denotes the confidence weighted average of the tth price relative. Plugging Eq. (7) into Eq. (6), we can get

$$\mu_{t+1} = \mu_t - \lambda \Sigma_t (\mathbf{x}_t - \bar{x}_t \mathbf{1}). \tag{8}$$

Moreover, taking the derivative of the Lagrangian with respect to Σ and setting it to zero, we can also have the update of Σ_{t+1},

$$0 = \frac{\partial \mathcal{L}}{\partial \Sigma} = -\frac{1}{2}\Sigma^{-1} + \frac{1}{2}\Sigma_t^{-1} + \lambda\phi \mathbf{x}_t \mathbf{x}_t^\top \quad \Longrightarrow \quad \Sigma_{t+1}^{-1} = \Sigma_t^{-1} + 2\lambda\phi \mathbf{x}_t \mathbf{x}_t^\top. \tag{9}$$

Now let us solve the Lagrange multiplier λ_{t+1} using KKT conditions. First following Dredze et al. [2008], we can compute the inverse using Woodbury identity [Golub and Van Loan 1996] as follows.

$$\Sigma_{t+1} = \left(\Sigma_t^{-1} + 2\lambda\phi \mathbf{x}_t \mathbf{x}_t^\top\right)^{-1} = \Sigma_t - \Sigma_t \mathbf{x}_t \frac{2\lambda\phi}{1 + 2\lambda\phi \mathbf{x}_t^\top \Sigma_t \mathbf{x}_t} \mathbf{x}_t^\top \Sigma_t. \tag{10}$$

The KKT conditions imply that either $\lambda = 0$, and no update is needed, or the constraint in optimization Problem (3) is an equality after the update. Taking Equation (8) and Equation (10) for the equality version of the first constraint, we get

$$\epsilon - (\mu_t - \lambda \Sigma_t (\mathbf{x}_t - \bar{x}_t \mathbf{1})) \cdot \mathbf{x}_t = \phi \left(\mathbf{x}_t^\top \left(\Sigma_t - \Sigma_t \mathbf{x}_t \frac{2\lambda\phi}{1 + 2\lambda\phi \mathbf{x}_t^\top \Sigma_t \mathbf{x}_t} \mathbf{x}_t^\top \Sigma_t \right) \mathbf{x}_t \right).$$

Now let $M_t = \mu_t^\top \mathbf{x}_t$ be the return mean, $V_t = \mathbf{x}_t^\top \Sigma_t \mathbf{x}_t$ be the return variance of the t^{th} trading period before updating, and $W_t = \mathbf{x}_t^\top \Sigma_t \mathbf{1}$ be the return variance of the t^{th} price relative with cash. We can simplify the preceding equation to

$$\lambda^2 (2\phi V_t^2 - 2\phi \bar{x}_t V_t W_t) + \lambda\, (2\phi\epsilon V_t - 2\phi V_t M_t + V_t - \bar{x}_t W_t) + (\epsilon - M_t - \phi V_t) = 0. \tag{11}$$

Let us define $a = 2\phi V_t^2 - 2\phi \bar{x}_t V_t W_t$, $b = 2\phi\epsilon V_t - 2\phi V_t M_t + V_t - \bar{x}_t W_t$, and $c = \epsilon - M_t - \phi V_t$. It is worth nothing that this quadratic form equation may have two/one/zero real roots. We can calculate its real roots (two real roots case: γ_{t1} and γ_{t2}; one real root case: γ_{t3}) as follows.

$$\gamma_{t1} = \frac{-b + \sqrt{b^2 - 4ac}}{2a}, \quad \gamma_{t2} = \frac{-b - \sqrt{b^2 - 4ac}}{2a}, \quad \text{or} \quad \gamma_{t3} = -\frac{c}{b}.$$

To ensure the non-negativity of the Lagrangian multiplier, we can project its value to $[0, +\infty)$,

$$\lambda = \max\{\gamma_{t1}, \gamma_{t2}, 0\}, \quad \text{or} \quad \lambda = \max\{\gamma_{t3}, 0\}, \quad \text{or} \quad \lambda = 0.$$

Note that these equations respectively correspond to three cases of real roots (two, one, or zero).

In practical computation, as we only adopt the diagonal elements of the covariance matrix, it is equivalent to compute λ from Eq. (11) but to update the covariance matrix with the following rule instead of Eq. (9).

$$\Sigma_{t+1}^{-1} = \Sigma_t^{-1} + 2\lambda\phi \operatorname{diag}^2(\mathbf{x}_t),$$

where $\operatorname{diag}(\mathbf{x}_t)$ denotes the diagonal matrix with the elements of $\mathbf{x}_t$ on the main diagonal. □

B. PROOF OF PROPOSITION 4.2

PROOF. Similar to the proof of Proposition 4.1, we relax the optimization problem without the non-negativity constraint, and project the solution to the simplex domain to obtain the required vector.

The Lagrangian for optimization Problem (4) is

$$\mathcal{L} = \frac{1}{2}\left(\log\left(\frac{\det\Upsilon_t^2}{\det\Upsilon^2}\right) + \mathrm{Tr}(\Upsilon_t^{-2}\Upsilon^2) + (\mu_t - \mu)^\top \Upsilon_t^{-2}(\mu_t - \mu)\right)$$
$$+ \lambda(\phi\,\|\Upsilon\mathbf{x}_t\| + \mu^\top\mathbf{x}_t - \epsilon) + \eta(\mu^\top\mathbf{1} - 1).$$

Taking the derivative of the Lagrangian with respect to μ and setting it to zero, we can get the update of μ_{t+1}.

$$0 = \frac{\partial\mathcal{L}}{\partial\mu} = \Upsilon_t^{-2}(\mu - \mu_t) + \lambda\mathbf{x}_t + \eta\mathbf{1} \quad\Longrightarrow\quad \mu_{t+1} = \mu_t - \Upsilon_t^2(\lambda\mathbf{x}_t + \eta\mathbf{1}),$$

where Υ_t is nonsingular. Multiplying both sides by $\mathbf{1}^\top$, we can get

$$\mathbf{1} = \mathbf{1} - \mathbf{1}^\top\Upsilon_t^2(\lambda\mathbf{x}_t + \eta\mathbf{1}) \quad\Longrightarrow\quad \eta = -\lambda\bar{x}_t,$$

where $\bar{x}_t = \frac{\mathbf{1}^\top\Upsilon_t^2\mathbf{x}_t}{\mathbf{1}^\top\Upsilon_t^2\mathbf{1}}$ is the confidence weighted average of the tth price relative. Plugging it into the update scheme of μ_{t+1}, we can get

$$\mu_{t+1} = \mu_t - \lambda\Upsilon_t^2(\mathbf{x}_t - \bar{x}_t\mathbf{1}).$$

Moreover, taking the derivative of the Lagrangian with respect to Υ and setting it to zero, we have,

$$0 = \frac{\partial\mathcal{L}}{\partial\Upsilon} = -\Upsilon^{-1} + \frac{1}{2}\Upsilon_t^{-2}\Upsilon + \frac{1}{2}\Upsilon\Upsilon_t^{-2} + \lambda\phi\frac{\mathbf{x}_t\mathbf{x}_t^\top\Upsilon}{2\sqrt{\mathbf{x}_t^\top\Upsilon^2\mathbf{x}_t}} + \lambda\phi\frac{\Upsilon\mathbf{x}_t\mathbf{x}_t^\top}{2\sqrt{\mathbf{x}_t^\top\Upsilon^2\mathbf{x}_t}}.$$

We can solve the preceding equation to obtain Υ^{-2},

$$\Upsilon_{t+1}^{-2} = \Upsilon_t^{-2} + \lambda\phi\frac{\mathbf{x}_t\mathbf{x}_t^\top}{\sqrt{\mathbf{x}_t^\top\Upsilon_{t+1}^2\mathbf{x}_t}}.$$

The preceding two updates can be expressed in terms of the covariance matrix as follows.

$$\mu_{t+1} = \mu_t - \lambda\Sigma_t(\mathbf{x}_t - \bar{x}_t\mathbf{1}), \quad \Sigma_{t+1}^{-1} = \Sigma_t^{-1} + \lambda\phi\frac{\mathbf{x}_t\mathbf{x}_t^\top}{\sqrt{\mathbf{x}_t^\top\Sigma_{t+1}\mathbf{x}_t}}. \tag{12}$$

Here, Σ_{t+1} is PSD and nonsingular.

Now let us solve the Lagrangian multiplier using its KKT condition. Following Crammer et al. [2008], we compute the inverse using the Woodbury identity [Golub and Van Loan 1996].

$$\Sigma_{t+1} = \Sigma_t - \Sigma_t\mathbf{x}_t\left(\frac{\lambda\phi}{\sqrt{\mathbf{x}_t^\top\Sigma_{t+1}\mathbf{x}_t} + \lambda\phi\mathbf{x}_t^\top\Sigma_t\mathbf{x}_t}\right)\mathbf{x}_t^\top\Sigma_t. \tag{13}$$

Similar to the proof of Proposition 4.1, let us set $M_t = \mu_t^\top\mathbf{x}_t$, $V_t = \mathbf{x}_t^\top\Sigma_t\mathbf{x}_t$, $W_t = \mathbf{x}_t^\top\Sigma_t\mathbf{1}$, and $U_t = \mathbf{x}_t^\top\Sigma_{t+1}\mathbf{x}_t$. Multiplying the preceding equation by $\mathbf{x}_t^\top$ (left) and $\mathbf{x}_t$ (right), we get $U_t = V_t - V_t\left(\frac{\lambda\phi}{\sqrt{U_t}+\lambda\phi V_t}\right)V_t$, which can be solved for U_t,

$$\sqrt{U_t} = \frac{-\lambda\phi V_t + \sqrt{\lambda^2\phi^2V_t^2 + 4V_t}}{2}. \tag{14}$$

The KKT condition implies that either $\lambda = 0$, and no update is needed, or the constraint in the optimization Problem (4) is an equality after the update. Substituting Eq. (12) and Eq. (14) into the equality version of the constraint, after rearranging in terms of λ, we get

$$\lambda^2 \left(\left(V_t - \bar{x}_t W_t + \frac{\phi^2 V_t}{2} \right)^2 - \frac{\phi^4 V_t^2}{4} \right) + 2\lambda \left(\epsilon - M_t \right) \left(V_t - \bar{x}_t W_t + \frac{\phi^2 V_t}{2} \right) + \left(\epsilon - M_t \right)^2 - \phi^2 V_t = 0. \tag{15}$$

Let $a = \left(V_t - \bar{x}_t W_t + \frac{\phi^2 V_t}{2} \right)^2 - \frac{\phi^4 V_t^2}{4}$, $b = 2 \left(\epsilon - M_t \right) \left(V_t - \bar{x}_t W_t + \frac{\phi^2 V_t}{2} \right)$, and $c = \left(\epsilon - M_t \right)^2 - \phi^2 V_t$. Note here we only consider real roots of this quadratic form equation. Thus, we can obtain γ_t as its roots (two real roots case: γ_{t1} and γ_{t2}; one real root case: γ_{t3}),

$$\gamma_{t1} = \frac{-b + \sqrt{b^2 - 4ac}}{2a}, \quad \gamma_{t2} = \frac{-b - \sqrt{b^2 - 4ac}}{2a} \quad \text{or} \quad \gamma_{t3} = -\frac{c}{b}.$$

To ensure the non-negativity of the Lagrangian multiplier, we project the roots to $[0, +\infty)$.

$$\lambda = \max \left\{ \gamma_{t1}, \gamma_{t2}, 0 \right\}, \quad \text{or} \quad \lambda = \max \left\{ \gamma_{t3}, 0 \right\}, \quad \text{or} \quad \lambda = 0,$$

which corresponds to three possible cases (two, one or zero real roots).

Following the proof of Proposition 4.1, we can update the diagonal covariance matrix as follows.

$$\Sigma_{t+1}^{-1} = \Sigma_t^{-1} + \lambda \frac{\phi}{\sqrt{U_t}} \operatorname{diag}^2 \left(\mathbf{x}_t \right),$$

where $\operatorname{diag}(\mathbf{x}_t)$ denotes the diagonal matrix with the elements of $\mathbf{x}_t$ on the main diagonal. □

C. PROOF OF THEOREM 4.3

PROOF. Before presenting our theorem, we first introduce two lemmas, which are given in Cesa-Bianchi and Lugosi [2006] (Example 10.3 and Theorem 10.3).

LEMMA C.1. *Assume that the investor competes against a finite class $\mathcal{Q} = \left\{ Q^1, \ldots, Q^N \right\}$ of investment strategies. A strategy P divides the initial wealth into N equal parts and invests each part according to experts Q^j. Then the total wealth of the strategy is $\mathbf{S}_n \left(P, \mathbf{x}^n \right) = \frac{1}{N} \sum_{j=1}^{N} \mathbf{S}_n \left(Q^j, \mathbf{x}^n \right)$, and the worst-case logarithmic wealth ratio is bounded as*

$$\sup_{\mathbf{x}^n} \sup_{Q \in \mathcal{Q}} \ln \frac{\mathbf{S}_n \left(Q, \mathbf{x}^n \right)}{\mathbf{S}_n \left(P, \mathbf{x}^n \right)} \leq \ln N.$$

LEMMA C.2. *If μ is the uniform density on the probability simplex Δ^m, then the wealth achieved by Cover's Universal Portfolio (UP) algorithm satisfies*

$$\sup_{\mathbf{x}^n} \sup_{B \in \Delta_m} \ln \frac{\mathbf{S}_n \left(B, \mathbf{x}^n \right)}{\mathbf{S}_n \left(UP, \mathbf{x}^n \right)} \leq (m-1) \ln (n+1).$$

First of all, it is not difficult to derive the following.

$$\sup_{\mathbf{x}^n} \sup_{B\in\Delta_m} \ln \frac{\mathbf{S}_n\left(B,\mathbf{x}^n\right)}{\mathbf{S}_n\left(P,\mathbf{x}^n\right)} = \sup_{\mathbf{x}^n} \ln \frac{\sup_{B\in\Delta_m}\mathbf{S}_n\left(B,\mathbf{x}^n\right)}{\mathbf{S}_n\left(UP,\mathbf{x}^n\right)} \frac{\mathbf{S}_n\left(UP,\mathbf{x}^n\right)}{\mathbf{S}_n\left(P,\mathbf{x}^n\right)}$$
$$\leq \sup_{\mathbf{x}^n} \ln \frac{\sup_{B\in\Delta_m}\mathbf{S}_n\left(B,\mathbf{x}^n\right)}{\mathbf{S}_n\left(UP,\mathbf{x}^n\right)} + \sup_{\mathbf{x}^n} \ln \frac{\mathbf{S}_n\left(UP,\mathbf{x}^n\right)}{\mathbf{S}_n\left(P,\mathbf{x}^n\right)}$$
$$\leq \sup_{\mathbf{x}^n} \ln \frac{\sup_{B\in\Delta_m}\mathbf{S}_n\left(B,\mathbf{x}^n\right)}{\mathbf{S}_n\left(UP,\mathbf{x}^n\right)} + \sup_{\mathbf{x}^n} \sup_{Q\in\mathcal{Q}} \ln \frac{\mathbf{S}_n\left(Q,\mathbf{x}^n\right)}{\mathbf{S}_n\left(P,\mathbf{x}^n\right)}.$$

Since we we buy and hold the experts with equal initial wealth $S_0\left(Q^j,\mathbf{x}^n\right) = 1, j = 1,\ldots,N$, we can apply Lemma C.1 to bound the worst-case regret bound with respect the best experts in the expert pool, that is, the second term in the last equation. Further, we apply Lemma C.2 to bound the worst-case regret bound with respect to the BCRP strategy, that is, the first term of the last equation. Combining the two lemmas, we can achieve the conclusion stated in the theorem. □

D. PROOF OF COROLLARY 4.4

PROOF. According to the result of Theorem 4.3 and the definition of universal property in Section 3.3, we have

$$\frac{1}{n} \sup_{\mathbf{x}^n} \sup_{B\in\Delta_m} \ln \frac{\mathbf{S}_n\left(B,\mathbf{x}^n\right)}{\mathbf{S}_n\left(P,\mathbf{x}^n\right)} \leq \frac{(m-1)\ln(n-1)+\ln N}{n} \xrightarrow{n\to\infty} 0.$$

According to the definition, the proposed CWMR mixture algorithm is universal. □

REFERENCES

Agarwal, A. and Hazan, E. 2005. New algorithms for repeated play and universal portfolio management. Tech. rep., Princeton University.

Agarwal, A., Hazan, E., Kale, S., and Schapire, R. E. 2006. Algorithms for portfolio management based on the Newton method. In *Proceedings of the 23rd International Conference on Machine Learning*. 9–16.

Akcoglu, K., Drineas, P., and Kao, M.-Y. 2005. Fast universalization of investment strategies. *SIAM J. Comput. 34,* 1, 1–22.

Aldridge, I. 2010. *High-Frequency Trading: A Practical Guide to Algorithmic Strategies and Trading Systems*. Wiley, Hoboken, N.J.

Belentepe, C. Y. 2005. A statistical view of universal portfolios. Ph.D. thesis, University of Pennsylvania.

Bernoulli, D. 1954. Exposition of a new theory on the measurement of risk. *Econometrica 23*, 23–36.

Blum, A. and Kalai, A. 1999. Universal portfolios with and without transaction costs. *Mach. Learn. 35,* 3, 193–205.

Bondt, W. F. M. D. and Thaler, R. 1985. Does the stock market overreact? *J. Finance 40*, 3, 793–805.

Bondt, W. F. M. D. and Thaler, R. 1987. Further evidence on investor overreaction and stock market seasonality. *J. Finance 42*, 3, 557–581.

Borodin, A., El-Yaniv, R., and Gogan, V. 2000. On the competitive theory and practice of portfolio selection (extended abstract). In *Proceedings of the 4th Latin American Symposium on Theoretical Informatics.* Springer-Verlag, Berlin, 173–196.

Borodin, A., El-Yaniv, R., and Gogan, V. 2004. Can we learn to beat the best stock. *J. Artif. Intell. Res. 21,* 579–594.

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

Breiman, L. 1961. Optimal gambling systems for favorable games. In *Proceedings of the Berkeley Symposium on Mathematical Statistics and Probability 1*, 65–78.

Cao, L. J. and Tay, F. E. H. 2003. Support vector machine with adaptive parameters in financial time series forecasting. *IEEE Trans. Neural Netw. 14*, 6, 1506–1518.

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

Cesa-Bianchi, N., Conconi, A., and Gentile, C. 2004. On the generalization ability of online learning algorithms. *IEEE Trans. Inf. Theory 50*, 9, 2050–2057.

Chaudhuri, K. and Wu, Y. 2003. Mean reversion in stock prices: Evidence from emerging markets. *Managerial Finance 29*, 22–37.

Chopra, V. K. and Ziemba, W. T. 1993. The effect of errors in means, variances, and covariances on optimal portfolio choice. *J. Portfolio Manage. 19*, 6–11.

Cover, T. and Ordentlich, E. 1998. Universal portfolios with short sales and margin. In *Proceedings of the Annual IEEE International Symposium on Information Theory*.

Cover, T. M. 1991. Universal portfolios. *Math. Finance 1*, 1, 1–29.

Cover, T. M. and Gluss, D. H. 1986. Empirical Bayes stock market portfolios. *Advances Appl. Math. 7,* 2, 170–181.

Cover, T. M. and Ordentlich, E. 1996. Universal portfolios with side information. *IEEE Trans. Inf. Theory* 42, 2, 348–363.

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

Crammer, K. and Singer, Y. 2003. Ultraconservative online algorithms for multiclass problems. *J. Mach. Learn. Res. 3*, 951–991.

Crammer, K., Dekel, O., Keshet, J., Shalev-Shwartz, S., and Singer, Y. 2006. Online passive-aggressive algorithms. *J. Mach. Learn. Res. 7*, 551–585.

Crammer, K., Dredze, M., and Pereira, F. 2008. Exact convex confidence-weighted learning. In *Advances in Neural Information Processing Systems.*

Crammer, K., Dredze, M., and Kulesza, A. 2009. Multi-class confidence weighted algorithms. In *Proceedings of the Conference on Empirical Methods in Natural Language*. Association for Computational Linguistics, Stroudsburg, PA. 496–504.

Cross, J. E. and Barron, A. R. 2003. Efficient universal portfolios for past-dependent target classes. *Math. Finance 13*, 2, 245–276.

Das, P. and Banerjee, A. 2011. Meta optimization and its application to portfolio selection. In *Proceedings of the 17th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining.* ACM, New York, NY, 1163–1171.

Dredze, M., Crammer, K., and Pereira, F. 2008. Confidence-weighted linear classification. In *Proceedings of the 25th International Conference on Machine Learning.* ACM, New York, NY, 264–271.

Duchi, J., Shalev-Shwartz, S., Singer, Y., and Chandra, T. 2008. Efficient projections onto the l_1-ball for learning in high dimensions. In *Proceedings of the 25th International Conference on Machine Learning*. ACM, New York, NY, 272–279.

Elton, E. J., Gruber, M. J., Brown, S. J., and Goetzmann, W. N. 1995. *Modern Portfolio Theory and Investment Analysis.* Wiley, New York.

Freund, Y. and Schapire, R. E. 1999. Large margin classification using the perceptron algorithm. *Mach. Learn. 37*, 3, 277–296.

Gaivoronski, A. A. and Stella, F. 2000. Stochastic nonstationary optimization for finding universal portfolios. *Annals Oper. Res. 100*, 165–188.

Gaivoronski, A. A. and Stella, F. 2003. On-line portfolio selection using stochastic programming. *Econ. Dynamics Contr. 27*, 6, 1013–1043.

Gentile, C. 2001. A new approximate maximal margin classification algorithm. *J. Mach. Learn. Res. 2,* 213–242.

Golub, G. H. and Van Loan, C. F. 1996. *Matrix Computations*. Johns Hopkins University Press, Baltimore, MD.

Grinold, R. and Kahn, R. 1999. *Active Portfolio Management: A Quantitative Approach for Producing Superior Returns and Controlling Risk*. McGraw-Hill, New York.

Györfi, L. and Schafer, D. 2003. Nonparametric prediction. In *Proceedings of Advances in Learning Theory: Methods, Models and Applications*. IOS Press, The Netherlands.

Györfi, L., and Vajda, I. 2008. Growth optimal investment with transaction costs. In *Proceedings of the 19th International Conference on Algorithmic Learning Theory*. Springer-Verlag, Berlin, 108–122.

Györfi, L., Lugosi, G., and Udina, F. 2006. Nonparametric kernel-based sequential investment strategies. *Math. Finance 16*, 2, 337–357.

Györfi, L., Urbán, A., and Vajda, I . 2007. Kernel-based semi-log-optimal empirical portfolio selection strategies. *Int. J. Theoret. Appl. Finance 10*, 3, 505–516.

Györfi, L., Udina, F., and Walk, H. 2008. Nonparametric nearest neighbor based empirical portfolio selection strategies. *Statist. Dec. 26*, 2, 145–157.

Györfi, L., Ottucsak, G., and Walk, H. 2012. *Machine Learning for Financial Engineering*. World Scientific, Singapore.

Hazan, E. and Kale, S. 2009. On stochastic and worst-case models for investing. In *Proceedings of Advances in Neural Information Processing Systems*.

Hazan, E. and Seshadhri, C. 2009. Efficient learning algorithms for changing environments. In *Proceedings of the 26th Annual International Conference on Machine Learning.* ACM, New York, NY. 393–400.

Helmbold, D. P., Schapire, R. E., Singer, Y., and Warmuth, M. K. 1997. A comparison of new and old algorithms for a mixture estimation problem. *Mach. Learn. 27,* 1, 97–119.

Helmbold, D. P., Schapire, R. E., Singer, Y., and Warmuth, M. K. 1998. On-line portfolio selection using multiplicative updates. *Math. Finance 8*, 4, 325–347.

Jegadeesh, N. 1990. Evidence of predictable behavior of security returns. *J. Finance 45*, 3, 881–898.

Jegadeesh, N. 1991. Seasonality in stock price mean reversion: Evidence from the U.S. and the U.K. *J. Finance 46*, 4, 1427–1444.

Kalai, A. and Vempala, S. 2002. Efficient algorithms for universal portfolios. *J. Mach. Learn. Res. 3,* 423–440.

Kelly, J., J. 1956. A new interpretation of information rate. *Bell Syst. Tech. J. 35,* 917–926.

Kimoto, T., Asakawa, K., Yoda, M., and Takeoka, M. 1993. Stock market prediction system with modular neural networks. *Neural Netw. Finance Invest.* 343–357.

Kivinen, J., Smola, A. J., and Williamson, R. C. 2004. Online learning with kernels. *IEEE Trans. Signal Process 52*, 8, 2165–2176.

Kullback, S. and Leibler, R. 1951. On information and sufficiency. *Annal Math. Stat. 22*, 79–86.

Latané, H. A. 1959. Criteria for choice among risky ventures. *J. Polit. Econ. 67*, 2, 144–155.

Levina, T. and Shafer, G. 2008. Portfolio selection and online learning. *Int. J. Uncertainty, Fuzziness Knowl.-Based Syst. 16*, 4, 437–473.

Li, B. and Hoi, S. C. H. 2012. On-line portfolio selection with moving average reversion. In *Proceedings of the 29th Annual International Conference on Machine Learning*. ACM, New York, NY.

Li, B., Hoi, S. C., and Gopalkrishnan, V. 2011a. Corn: Correlation-driven nonparametric learning approach for portfolio selection. *ACM Trans. Intell. Syst. Technol. 2*, 3, 21–29.

Li, B., Hoi, S. C. H., Zhao, P., and Gopalkrishnan, V. 2011b. Confidence weighted mean reversion strategy for on-line portfolio selection. *J. Mach. Learn. Res.—Proceedings Track 15*, 434–442.

Li, B., Zhao, P., Hoi, S., and Gopalkrishnan, V. 2012. Pamr: Passive aggressive mean reversion strategy for portfolio selection. *Mach. Learn. 87*, 2, 221–258.

Li, Y. and Long, P. M. 1999. The relaxed online maximum margin algorithm. In *Proceedings of Advances in Neural Information Processing Systems*.

Lo, A. W. and MacKinlay, A. C. 1990. When are contrarian profits due to stock market overreaction? *Rev. Financial Studies 3*, 2, 175–205.

Lu, C.-J., Lee, T.-S., and Chiu, C.-C. 2009. Financial time series forecasting using independent component analysis and support vector regression. *Decision Supp. Syst. 47*, 115–125.

Magdon-Ismail, M. and Atiya, A. 2004. Maximum drawdown. *Risk Mag. 10*, 99–102.

Markowitz, H. 1952. Portfolio selection. *J. Finance 7,* 1, 77–91.

Ordentlich, E. and Cover, T. M. 1996. On-line portfolio selection. In *Proceedings of the 9th Annual Conference on Computational Learning Theory*. ACM, New York, NY, 310–313.

Ottucsák, G. and Vajda, I. 2007. An asymptotic analysis of the mean-variance portfolio selection. *Statist. Decisions 25*, 63–88.

Poterba, J. M. and Summers, L. H. 1988. Mean reversion in stock prices: Evidence and implications. *J. Financial Econ. 22,* 1, 27–59.

Rosenblatt, F. 1958. The perceptron: A probabilistic model for information storage and organization in the brain. *Psychol. Rev. 65*, 386–407.

Sharpe, W. F. 1963. A simplified model for portfolio analysis. *Manage. Sci. 9*, 277–293.

Sharpe, W. F. 1964. Capital asset prices: A theory of market equilibrium under conditions of risk. *J. Finance 19*, 3, 425–442.

Sharpe, W. F. 1994. The Sharpe ratio. *J. Portfolio Manage. 21*, 1, 49–58.

Singer, Y. 1997. Switching portfolios. *Int. J. Neural Syst. 8,* 4, 488–495.

Tay, F. E. H. and Cao, L. 2001. Application of support vector machines in financial time series forecasting. *Omega 29*, 4, 309–317.

Thorp, E. O. 1971. Portfolio choice and the Kelly criterion. In *Proceedings of Business and Economics Section of the American Statistical Association*.

Tsang, E., Yung, P., and Li, J. 2004. Eddie-automation, a decision support tool for financial forecasting. *Decision Support Syst. 37*, 559–565.

Vovk, V. and Watkins, C. 1998. Universal portfolio selection. In *Proceedings of the 11th Annual Conference on Computational Learning Theory*. ACM, New York, NY, 12–23.

Vovk, V. G. 1990. Aggregating strategies. In *Proceedings of the 3rd Annual Workshop on Computational Learning Theory*. Morgan Kaufmann Publishers Inc., San Francisco, CA, 371–386.

Wang, J., Zhao, P., and Hoi, S. C. H. 2012. Exact soft confidence-weighted learning. In *Proceedings of the 29th Annual International Conference on Machine Learning*. ACM, New York, NY.

Zhao, P., Hoi, S. C. H., and Jin, R. 2011. Double updating online learning. *J. Mach. Learn. Res. 12*, 1587–1615.