European Journal of Operational Research 314 (2024) 1143–1158

Contents lists available at ScienceDirect

European Journal of Operational Research

journal homepage: www.elsevier.com/locate/eor

Innovative applications of O.R.

Distributed mean reversion online portfolio strategy with stock network

Yannan Zhong [a,c], Weijun Xu [a,c,*], Hongyi Li [b], Weiwei Zhong [a,c]

[a] School of Business Administration, South China University of Technology, Guangzhou, Guangdong, China
[b] Business School, Chinese University of Hong Kong, Shatin, New Territories, Hong Kong
[c] Guangzhou Financial Services Innovation and Risk Management Research Base, Guangzhou, Guangdong, China

ARTICLE INFO

Keywords:
Portfolio optimization
Online portfolio
Stock network
Mean reversion
Distributed optimization

ABSTRACT

Online portfolio selection is a practical problem in financial engineering and quantitative trading. Many empirical studies show that stock performance in the market is likely to follow mean reversion, and strategies based on mean reversion show better return performance than the market average. However, the existing mean reversion strategies are not universal and short selling is not allowed, which is unsuitable for real-time investment. In this paper, we propose a distributed mean reversion online portfolio strategy through a stock correlation sub-network to solve these problems. Theoretical analysis shows that our strategy is universal and the convergence rate is calculated. The empirical results show that our strategy is better than the existing universal strategies in terms of return performance, nor is it sensitive to transaction cost.

1. Introduction

Portfolio selection is a practical engineering task in finance that aims to minimize portfolio risk or maximize its return by allocating wealth in different assets. The research for portfolio strategy can be divided into two major schools. The first being the traditional portfolio strategy based on mean–variance (MV) theory (Markowitz, 1952). This static portfolio strategy mainly focuses on single-period (batch) portfolio selection problems. It measures the expected return and risk of a portfolio through the mean and variance or covariance matrix of prices, respectively, and weights the expected return and risk through utility functions to determine the best portfolio strategy in the investment period. Recently, MV theory has also been applied in the multi-period setting (e.g., Marc, 2001). The second is the online portfolio strategy (OLPS) based on capital growth theory (CGT) (Kelly, 1956). This dynamic portfolio strategy focuses on multi-period or sequential portfolio selection problems. It continuously adjusts the wealth proportion in different assets in each trading period to maximize the final log-cumulative wealth or log-cumulative wealth growth rate.

Recently, the rise of fintech such as online investment and quantitative trading has stimulated strong investment demand to buy, hold, and dynamically manage financial products such as funds, stocks, and futures in online portfolio platforms. The features of dynamically adjusted assets for OLPS make it a good fit for online investment and it has received significant attention from academics and investors.[1] Therefore, how to analyze stock performance and design a more realistic OLPS is the main purpose in this paper.

In a portfolio selection problem, investors and researchers are mainly concerned with return and risk. High return is the ultimate optimization goal of any OLPS, which is different from MV theory. The creativity of MV theory is that the variance or covariance of stock prices is first used to quantify the risk analysis, which opens a new avenue for asset allocation and risk management. With the development of complex network theory in recent years, traditional portfolio selection strategies have been strengthened to analyze the correlation of stock prices. Empirical studies show that a stock market is a complex network. As such, it is natural to combine MV theory and complex network theory in the portfolio selection problem. Existing studies (e.g., Eduard and Jochen (2017), Peralta and Zareei (2016)) have shown that stocks can influence each other and spread risks among them through a "correlation network". For example, nodes with higher centrality in a stock correlation network correspond to riskier

* Corresponding author at: School of Business Administration, South China University of Technology, Guangzhou, Guangdong, China.
E-mail addresses: yannanzhong@outlook.com (Y. Zhong), xuwj@scut.edu.cn (W. Xu), hongyili@cuhk.edu.hk (H. Li), bmweiweizhong@mail.scut.edu.cn (W. Zhong).

[1] According to the financial engineering team of CITIC Securities, as of Q4, 2021, in securities private placement, the scale of quantitative products is close to 1,610 billion RMB yuan, accounting for 24.83% of the proportion of private securities products. The total size of the quantitative private equity managers is estimated to be about 480 billion RMB yuan. Publicly offered quantitative funds reached 294.1 billion RMB yuan, and index funds reached 1.3 trillion RMB yuan. At the same time, in 2021, the daily trading volume of A-shares frequently exceeded one trillion RMB yuan, and various market entities are estimated to account for about 20%–30%.

https://doi.org/10.1016/j.ejor.2023.11.021
Received 20 October 2022; Accepted 13 November 2023
Available online 17 November 2023
0377-2217/

stocks (Peralta & Zareei, 2016). Using the network feature of stocks to improve the risk measurement of a MV portfolio strategy can improve the overall return performance of the strategy (Li et al., 2019; Vyrost et al., 2019).

However, few studies regarding OLPS analyze the correlation of stocks from the perspective of a stock network. This is because OLPS based on CGT mainly focuses on maximizing the cumulative wealth and does not specifically assume which distribution the assets follow. This makes it difficult for OLPS to analyze the correlation network between stocks. Moreover, OLPS generally assumes that short selling is not allowed, which assumption does not consider more realistic investment needs. Next, we briefly review two representative OLPSs to further explain the problems discussed above: the "follow the winner" strategy (also known as the universal strategy) and the "follow the loser" strategy (also known as the mean reversion strategy).

The follow the winner strategy mainly refers to the selection of "winner" assets in the past to determine the portfolio strategy in the next trading period, and it aims to be a universal strategy (Cover, 1991). A universal strategy can provide investors with a return, and is a best constant rebalanced portfolio (BCRP) strategy in theory. Thus, a universal strategy is the first choice for investors in most situations. However, a universal strategy usually requires that the relative prices of stocks are independent and active (Cover, 1991) (meaning short selling is not allowed). Notice that the universal strategy does not take stock correlation into account and does not satisfy the short selling need of investors.

Mean reversion strategies are designed based on the mean reversion feature of stocks (Jegadeesh, 1990). This strategy focuses on predicting stock returns based on the mean reversion feature of stock performance, thus a strategy model is constructed. In recent years, researchers have done a lot of work to improve the stock return prediction ability of mean reversion strategies (Huang et al., 2016; Li et al., 2015). However, these works do not deal with the prediction errors (Guo et al., 2021) and fail to consider the correlations between stocks, which means that the prediction errors among stocks are uncorrelated. As a result, the overall prediction errors of stocks in a mean reversion strategy diverge as the number of trading periods increases, thus the mean reversion strategy is not universal. Moreover, short selling is also not allowed in these strategies.

From the above analysis we find that OLPS faces several practical problems: (1) short selling is not allowed; (2) the relation between stocks is not considered—existing OLPS usually assume that the relative prices among stocks are independent and identically distributed, and they only consider the historical performance of each stock individually; and (3) the mean reversion strategy is not universal—it cannot provide the theoretical return guarantee for investors.

In light of the above problems, this paper aims to design a mean reversion online portfolio strategy that is universal from the perspective of a stock correlation network. Short selling is allowed in our strategy, making it more practical. To achieve this goal, we construct a dynamic stock correlation network to analyze the correlation of stocks and update our strategy in a distributed approach. In particular, (1) we analyze the stock correlations and construct a dynamic stock sub-network using stock centrality. In each trading period, we determine the updated rules for the next portfolio strategy based on this stock sub-network structure. (2) We propose a distributed mean reversion (DMR) strategy and short selling is allowed. The DMR is weighted by a series of portfolio strategies generated by distributed online trading machines. Specifically, each trading machine communicates with the others according to the dynamic stock sub-network structure and updates its next strategy. (3) Theoretical analysis shows that the DMR strategy is a universal strategy, and the convergence rate of the strategy is analyzed.

Lastly, we analyze the performance of the DMR and other universal strategies through empirical experiments. The results show that the DMR has a better return performance than other universal strategies, and that it is less sensitive to the transaction cost and the parameters of the strategy.

The rest of this paper is organized as follows. Section 2 explains the basic notations in this paper. Section 3 discusses the literature and related works. Section 4 provides the strategy model and analyzes the convergence of the DMR. Section 5 presents empirical results using real data. Section 6 gives the usage recommendations of the DMR strategy, as well as areas of future study.

2. Preliminaries

Before we present our work, we first introduce some mathematical notations. For the following variables in the mathematical context, lowercase letters denote constants such as a, and lowercase bold letters denote vectors such as $\boldsymbol{x}$. $\mathbf{1}$ is an m-dimensional column vector whose elements are all 1. Matrices and sets are denoted by capital letters such as matrix A and set Δ. $\boldsymbol{x}^T$ is the transpose of $\boldsymbol{x}$, R^m is the m-dimensional Euclidean space, R^m_+ is the nonnegative subset of the Euclidean space, and $\boldsymbol{x} \leq a$ means each element of $\boldsymbol{x}$ is less than a.

We consider a portfolio that contains m stocks and trades for T periods. $V_s = \{v_1, \dots, v_m\}$ is the index set of these stocks. In the kth period, we denote the closing price vector of stocks as $\boldsymbol{p}(k) = (p_1(k), \dots, p_m(k))^T$, where $p_i(k)$ is the closing price of stock v_i. The relative price vector (also called the market vector) is $\boldsymbol{x}(k) = (x_1(k), \dots, x_m(k))^T$, where $x_i(k) = \frac{p_i(k)}{p_i(k-1)}$ is the relative price of stock v_i, and $\bar{x}(k) = \frac{1}{m}\boldsymbol{x}(k)^T\mathbf{1}$ is the mean of $\boldsymbol{x}(k)$. Denote the portfolio strategy in period k as $\boldsymbol{b}(k) = (b_1(k), \dots, b_m(k))^T$. We call $\boldsymbol{b}$ a portfolio strategy with short selling not allowed if $\boldsymbol{b} \in \Delta^1_m = \{\boldsymbol{b} \in R^m | \boldsymbol{b}^T\boldsymbol{x}1 = 1, 0 \leq \boldsymbol{b} \leq 1\}$, where Δ^1_m is the feasible set of the portfolio strategy with short selling not allowed. If $\boldsymbol{b} \in \Delta^2_m = \{\boldsymbol{b} \in R^m | \boldsymbol{b}^T\mathbf{1} = 1\}$, then $\boldsymbol{b}$ is a portfolio strategy with short selling allowed. The cumulative wealth at the end of period k is $S(k) = S(0)\prod_{h=1}^{k}\boldsymbol{b}(k)^T\boldsymbol{x}(k)$, the final cumulative wealth is $S(T) = S(0)\prod_{k=1}^{T}\boldsymbol{b}(k)^T\boldsymbol{x}(k)$, where $S(0)$ is the initial capital, and we set $S(0) = 1$. $\boldsymbol{x}(i,j)$ represents the relative price from period i to period j, that is, $\boldsymbol{x}(i,j) = \{\boldsymbol{x}(i), \boldsymbol{x}(i+1), \dots, \boldsymbol{x}(j)\}$, and $\boldsymbol{x}_h(i,j) = (x_h(i), x_h(i+1), \dots, x_h(j))$ is the vector constructed by the hth element of all vectors in $\boldsymbol{x}(i,j)$. We denote the projection of a vector $\boldsymbol{x}$ on the set Ω as $P_\Omega[\boldsymbol{x}]$, i.e., $P_\Omega[\boldsymbol{x}] = argmin_{\boldsymbol{y}\in\Omega}\|\boldsymbol{y}-\boldsymbol{x}\|^2$. We use $dist(\boldsymbol{x}; B)$ for the Euclidean distance of a vector $\boldsymbol{x}$ from a set B, i.e., $dist(\boldsymbol{x}; B) = \min_{\boldsymbol{y}\in B}\|\boldsymbol{y}-\boldsymbol{x}\|$. Denote $\partial f(\boldsymbol{x})$ as the sub-differential set of function $f(\boldsymbol{x})$ at point $\boldsymbol{x}$, that is, $\partial f(\boldsymbol{x}) = \{\boldsymbol{d} \in R^m | \boldsymbol{d}^T(\boldsymbol{y}-\boldsymbol{x}) \leq f(\boldsymbol{y}) - f(\boldsymbol{x}), \forall \boldsymbol{y} \in R^m\}$, where f is a convex map from $R^m \to R$. $\nabla f(\boldsymbol{x})$ is the gradient of f at point $\boldsymbol{x}$. $\mathbb{E}[\boldsymbol{z}|\mathcal{F}]$ is the conditional expectation of random variable $\boldsymbol{z}$ at σ-algebra $\mathcal{F}$.

In Section 4, we set up multiple trading machines to carry out the updating rules of the DMR. The structure of the communication network among the trading machines is determined by the structure of the stock correlation sub-network. Before we introduce the construction of the stock correlation sub-network and its dynamic matching rules with the trading machines, we first give the topological structure and symbols of the communication network between the trading machines. Let $G(k) = (V, E(k))$ be the directed communication network between the n trading machines ($n \leq m$) in the kth period, where $V = \{1, 2, \dots, n\}$ is the node set of these machines. $E(k) = \{(i,j)|\text{if } j \text{ will send a strategy message to } i, \ \forall i,j \in V\}$ is the edge set of $G(k)$. We set the matrix $A(k) = \begin{cases} a_{ij}(k) > 0, & (i,j) \in E(k) \\ a_{ij}(k) = 0, & (i,j) \notin E(k) \end{cases}$ to be the weighted adjacency matrix of $G(k)$. From adjacency matrix $A(k)$, we can infer whether communication between machines i and j exists ($a_{ij}(k), a_{ji}(k) > 0$) and the importance ($a_{ij}(k)$) of the strategy message for each other. If communication between i and j is bidirectional, meaning that $(i,j) \in E(k)$ implies $(j,i) \in E(k)$, then we call the edge (i,j) undirected. If all edges in the network $G(k)$ are undirected, we call the network $G(k)$ an undirected network.

3. Literature review and related work

In this section, we introduce relevant research, including online portfolio strategies and complex networks of stocks. In addition, the DMR uses a distributed strategy update framework, therefore, the distributed optimization algorithm is also briefly introduced.

3.1. Online portfolio strategies

We first introduce a few common benchmark strategies, and then introduce four main OLPSs: "follow the winner", "follow the loser", "pattern matching", and "meta learning". Li and Steven's (2014) survey is a good reference for more information about OLPS.

3.1.1. Benchmarks

Buy and hold (BAH): The BAH strategy is a common investment strategy which allocates the initial capital $S(0)$ to m stocks in the initial trading period, and it does not adjust in the subsequent trading periods, i.e., $\boldsymbol{b}(0) = (a_1, \dots, a_m)$, where a_i is the share of asset v_i in initial capital $S(0)$ and $\sum_{i=1}^{m} a_i = S(0)$. A special BAH strategy is the uniform buy and hold (UBAH) strategy, also known as the market strategy. The UBAH strategy allocates the initial capital $S(0)$ equally among all assets, i.e., $\boldsymbol{b}(0) = (\frac{1}{m}S(0), \dots, \frac{1}{m}S(0))$.

Best stock: This off-line strategy assumes that the market vector $\{\boldsymbol{x}(t)\}_{t=1}^{T}$ is known at the beginning. The strategy puts all initial capital $S(0)$ into one stock that can achieve the maximum final cumulative return, i.e., $\boldsymbol{b}(0) = \boldsymbol{e}^i$, where $i = argmax_{i=1,\dots,m} \prod_{t=1}^{T} x_i(t)$, and $\boldsymbol{e}^i$ is a vector whose ith element is 1 and the rest of the elements are 0s.

Constant rebalanced portfolio (CRP): The CRP strategy can be regarded as a dynamic BAH strategy and determines the initial shares of assets as $(a_1, \dots, a_m)$, and adjusts the asset holding in subsequent trading periods according to the initial shares, that is, $\boldsymbol{b}(t) = (a_1, \dots, a_m), t = 1, 2, \dots, T$, where $a_1, \dots, a_m$ are constants satisfying $a_i \geq 0, \sum_{i=1}^{m} a_i = S(0)$.

Best constant rebalanced portfolio (BCRP): The BCRP strategy is a posterior optimal CRP strategy often used for theoretical analysis. BCRP assumes that the market vector $\{\boldsymbol{x}(t)\}_{t=1}^{T}$ is known in advance, and then it maximizes the final cumulative wealth, i.e., $\boldsymbol{b}(t) = argmax_{\boldsymbol{b}\in\Delta_m^1} \prod_{k=1}^{T} \boldsymbol{b}^T \boldsymbol{x}(k),\ t = 1, \dots, T$. Denoting $B^* = \{\boldsymbol{b} | \boldsymbol{b} = argmax_{\boldsymbol{b}\in\Delta_m^1} \prod_{k=1}^{T} \boldsymbol{b}^T \boldsymbol{x}(k)\}$, we know that $\forall \boldsymbol{b}^* \in B^*$ is a feasible BCRP strategy. For a given relative price sequence $\{\boldsymbol{x}(1), \dots, \boldsymbol{x}(T), \dots\}$, if we assume that the assets in the portfolio will not retreat and crash, then the set B^* is obviously non-empty. Cover (1991) proved that in an independent and identically distributed (i.i.d.) stock market (i.e., $\{\boldsymbol{x}(t)\}_{t=1}^{+\infty}$ is i.i.d.), the final cumulative wealth $S(T)^* = \prod_{t=1}^{T} \boldsymbol{b}^{*T} \boldsymbol{x}(t)$ of BCRP is theoretically no lower than that of any BAH, best stock, or CRP strategies.

Universal strategy: For an OLPS $\{\boldsymbol{b}(t)\}_{t=1}^{T}$, if its log-cumulative wealth growth rate $\frac{1}{T}\ln S(T)$ can asymptotically converge to that of a BCRP strategy with the increase of trading period T, i.e.,

$$\frac{1}{T}\left[\ln S(T)^* - \ln S(T)\right] \to 0 \quad (T \to +\infty),$$

then the strategy is universal. A universal strategy can theoretically provide investors with a lower bound of cumulative returns, namely $S(T)^*$. Being universal is important to measure the performance of OLPS in theory.

3.1.2. Follow the winner

The basic idea of the follow the winner strategy is to increase the weights of good stocks (high returns). Cover (1991) proposed the representative universal portfolio (UP) strategy, which laid the theoretical foundation for the above property. The UP strategy traverses all the feasible strategies in Δ_m^1 and weights these strategies according to their current cumulative wealth to obtain the next portfolio strategy. The strategy update rules of UP are as follows:

$$\boldsymbol{b}(t+1) = \frac{\int_{\Delta_m^1} \boldsymbol{b} S(t) d\boldsymbol{b}}{\int_{\Delta_m^1} S(t) d\boldsymbol{b}},\ \boldsymbol{b}(0) = (\frac{1}{m}, \dots, \frac{1}{m})^T.$$

Note that the UP strategy is defined without making any assumptions about the distribution of the asset prices, see Cover (1991). Later, Cover and Ordentlich (2006) introduced side information into the framework of UP, Kalai and Vempala (2002) improved the computation method of UP and reduced its computation time.

Helmbold et al. (1998) introduced the relative entropy measure $D_{RE}(\boldsymbol{b}(t+1) \| \boldsymbol{b}(t))$ as a penalty term for the maximum expected return and proposed the exponential gradient (EG) strategy, which is universal. The strategy of EG is:

$$\boldsymbol{b}(t+1) = \arg\max_{\boldsymbol{b}\in\Delta_m^2} \ \eta \ln \boldsymbol{b}(t+1)^T \boldsymbol{x}(t) - D_{RE}(\boldsymbol{b}(t+1) \| \boldsymbol{b}(t)),$$

where $\eta > 0$ is a learning rate, $D_{RE}(\boldsymbol{b}(t+1) \| \boldsymbol{b}(t)) = \sum_{i=1}^{m} b_i(t+1) \log \frac{b_i(t+1)}{b_i(t)}$. Solving this optimization problem, we can obtain the strategy update rule of EG as follows:

$$b_i(t+1) = \frac{b_i(t) \exp\left(\eta x_i(t) / \boldsymbol{b}(t)^T \boldsymbol{x}(t)\right)}{\sum_{j=1}^{m} b_j(t) \exp\left(\eta x_j(t) / \boldsymbol{b}^T(t) \boldsymbol{x}(t)\right)}.$$

Yang, He, et al. (2019) introduced side information into the framework of EG and proposed the EG with side-information (EGS) strategy, which has better performance than the EG strategy in terms of return. The strategy is:

$$b_i(t+1)|_{y_{t+1}} = \frac{b_i(s_{t+1}) \exp\left(\eta|_{y_{t+1}} x_i(s_{t+1}) / \boldsymbol{b}(s_{t+1})^T \boldsymbol{x}(s_{t+1})\right)}{\sum_{j=1}^{m} b_j(s_{t+1}) \exp\left(\eta|_{y_{t+1}} x_j(s_{t+1}) / \boldsymbol{b}^T(s_{t+1}) \boldsymbol{x}(s_{t+1})\right)},$$

where y_{t+1} is the side information in period t+1 and $\eta|_{y_{t+1}}$ is the learning rate determined by y_{t+1}, $s_{t+1} = \max\{\tau : \tau \leq t, t_\tau = y_{t+1}\}$. If s_{t+1} is empty, then $b_i(t+1)|_{y_{t+1}} = \frac{1}{m}$.

3.1.3. Follow the loser

The follow the loser strategy, also known as the mean reversion strategy, holds that the performance of a stock in the market follows the mean reversion features (Jegadeesh, 1990; Poterba & Summers, 1988), that is, stocks which perform poorly/better in the current or past period will perform better/worse in the next period. Investors can profit by increasing capital to worse stocks and decreasing capital to better stocks. From the existing research results on mean reversion, stock "performance" is mainly defined from the stock relative prices, and the measure of stock "performance" can be divided into single-period stock performance and multi-period stock performance.

Borodin et al. (2004) considered the multi-period mean reversion features from the growth rate of stocks (measured by the product of relative prices during a window) and proposed an anti-correlation (Anticor) strategy. Anticor assumes that the growth rate of any two stocks have a positive correlation in two consecutive time windows and each stock follows the mean reversion at the same time, that is, the inter-temporal growth rate of the two stocks rises or falls together and the growth rate of each stock reverses. Li et al. (2012) assumed that stock relative price follows single period mean reversion and proposed a passive aggressive mean reversion (PAMR) strategy. The PAMR strategy determines whether to buy or sell a stock in the next period by observing whether the stock's single-period return exceeds a predetermined mean reversion threshold ε. The PAMR strategy is as follows:

$$\boldsymbol{b}(t+1) = \arg\min_{\boldsymbol{b}\in\Delta_m^1} \frac{1}{2} \|\boldsymbol{b} - \boldsymbol{b}(t)\|^2 \quad s.t.\ l_\varepsilon(\boldsymbol{b}; \boldsymbol{x}(t)) = 0,$$

where $l_\varepsilon(\boldsymbol{b}; \boldsymbol{x}(t)) = \max\{0, \boldsymbol{b}(t)^T \boldsymbol{x}(t) - \varepsilon\}$. Solving this optimization problem, we can obtain the update rule of PAMR:

$$\boldsymbol{b}(t+1) = \boldsymbol{b}(t) - \lambda(t) \left(\boldsymbol{x}(t) - \bar{x}(t)\boldsymbol{1}\right),$$

where $\lambda(t) = \max\{0, \frac{\boldsymbol{b}^T \boldsymbol{x}(t) - \varepsilon}{\|\boldsymbol{x}(t) - \bar{x}(t)\boldsymbol{1}\|^2}\}$. Gao and Zhang (2013) improved the threshold function $l_\varepsilon(\boldsymbol{b}; \boldsymbol{x}(t))$ of PAMR and proposed the weighted moving average mean reversion (WMAMR) strategy.

However, the single-period mean reversion strategy only uses the stock information from a single trading period, which makes its final cumulative wealth fluctuate greatly and be far lower than that of the market strategy (Li et al., 2012) in some markets with insufficient mean reversion features. Li et al. (2015) assumed that the stock relative prices follow multi-period mean reversion and proposed the online moving average reversion (OLMAR) strategy. OLMAR captures the mean reversion feature from multiple historical trading periods and contains two methods to predict the next relative price $\tilde{\boldsymbol{x}}(t+1)$: the sample moving average (SMA) method based on w historical data and the exponential moving average (EMA) method based on the entire historical trading period.

The OLMAR strategy is:

$$\boldsymbol{b}(t+1) = \arg\min_{\boldsymbol{b}\in\Delta_m^1} \frac{1}{2}\|\boldsymbol{b}-\boldsymbol{b}(t)\|^2 \quad s.t.\ \boldsymbol{b}^T\tilde{\boldsymbol{x}}(t+1) \geq \varepsilon.$$

Huang et al. (2016) introduced the L_1-median to improve the robustness of the stock price prediction of OLMAR and proposed the robust median reversion (RMR) strategy. Lai et al. (2020) adopted an independent weight factor for each stock to improve the stock price prediction ability of OLMAR and proposed the reweighted price relative tracking (RPRT) strategy. Guo et al. (2021) proposed an adaptive online moving average method (AOLMA) to improve the accuracy of the return prediction of OLMAR. Li et al. (2018) optimized the transaction cost constraints of PAMR and OLMAR (by using a more realistic transaction cost calculation method) and proposed the transaction cost optimization (TCO) strategy. The empirical study shows that the multi-period mean reversion strategies can better capture the mean reversion feature of stocks, and its comprehensive performance in different markets is better than that of the single-period mean reversion strategy.

3.1.4. Pattern matching

The pattern matching strategy is different from the above two strategies. It filters the historical data, similar to the recent stock performance from the stock relative prices in the current trading period, and maximizes the utility function on each filtered historical relative price sequence of the stocks to generate the "experts" (the optimal strategies for maximizing the utility function). According to the performance of each expert, the strategy integrates them in a non-parametric form to determine the next portfolio strategy. The execution process of the pattern matching strategy can be divided into three parts: (1) compute the historical similarity set $C(\varepsilon)$, where ε is the threshold of similarity; (2) maximize the utility function $U(\boldsymbol{b}, C(\varepsilon))$ on the similarity set $C(\varepsilon)$ to obtain the corresponding experts $\boldsymbol{b}(\varepsilon)$; and (3) combine experts with different parameter values non-parametrically. The pattern matching strategy tries to maximize its expected log return in terms of $C(\varepsilon)$, which is consistent with CGT and results in an optimal fixed fraction portfolio.[2] Thus, this strategy is universal in theory.

3.1.5. Meta learning

The meta learning strategy takes several basic OLPSs as the experts and weights them to generate the next portfolio strategy. The experts selected can be any one of the above three types of strategy. The aggregation algorithm first proposed by Vovk and Watkins (1998) solved the problem of integrating multiple experts and generalized the worst-case bound of UP. Das and Banerjee (2011) and Agarwal et al. (2006) proposed the online gradient update (OGU) strategy and online Newton update (ONU) strategy, respectively. Hazan and Seshadhri (2009) proposed the follow the leader history (FLH) strategy, where a group of foundation experts are constantly updated from a different start point in history and make a forecast of future prices. He and Yang (2020) proved that the aggregation strategy of multiple universal strategies is still universal if we use the weak aggregation algorithm to deal with the weight problem between experts. Yang et al. (2022) considered a number of EG strategies in different parameter η as experts and proposed the weak aggregating exponential gradient (WAEG) strategy, which uses the weak aggregate algorithm to aggregate the experts' advice in each period. The WAEG strategy is:

$$\boldsymbol{b}(t+1) = \sum_{\eta\in H}^{k} \boldsymbol{b}^{\eta}(t+1) w^{\eta}(t+1),$$

where $b_i^{\eta}(t+1) = \frac{b_i^{\eta}(t)\exp(\eta\cdot x_i(t)/\boldsymbol{b}^{\eta}(t)^T\boldsymbol{x}(t))}{\sum_{i=1}^{m} b_i^{\eta}(t)\exp(\eta\cdot x_i(t)/\boldsymbol{b}^{\eta}(t)^T\boldsymbol{x}(t))}$ is the ith element of $\boldsymbol{b}^{\eta}(t+1)$, $w^{\eta}(t+1)$ is the aggregation weight generated by the weak aggregation algorithm, and H is the parameter set. They theoretically proved that the WAEG strategy is universal. Zhang et al. (2022) proposed a moving-window-based adaptive exponential gradient (MAEG) strategy which is suitable for nonstationary financial data sets. They set a number of strategies EG(η) (EG with learning rate η) as experts and calculated the cumulative return of all the experts during period $[t-w+1, t]$ to find the best expert. Then, they used the η^* which corresponds to the best expert to formulate the next portfolio. The strategy is:

$$b_i(t+1) = \frac{b_i(t)\exp\left(\eta_{t+1}(w)x_i(t)/\boldsymbol{b}(t)^T\boldsymbol{x}(t)\right)}{\sum_{j=1}^{m} b_j(t)\exp\left(\eta_{t+1}(w)x_j(t)/\boldsymbol{b}^T(t)\boldsymbol{x}(t)\right)},$$

where $\eta_{t+1}(w) = argmax_{\eta\in H} S(t, EG(\eta)) - S(t-w+1, EG(\eta))$, $S(t, EG(\eta))$ means the cumulative wealth in period t of the EG(η) strategy, and H is the parameter set. Additional meta learning strategies can be found in Akcoglu et al. (2005, 2020).

3.2. Complex network of stocks

A complex network is a way to study the complex relationships of a system based on nodes and edges. It is widely used in automatic control (Ram et al., 2010), financial analysis (Battiston et al., 2016), and so on. Stock correlation network research is the most intuitive application of complex networks in finance, where the relevant literature accounts for a large proportion of financial network research and scholars use complex network theory in the innovative interpretation of various financial phenomena. In recent years, many scholars have applied complex network theory in the field of portfolio selection. By analyzing the characteristics of a stock correlation network, they have carried out stock optimization and portfolio model improvement. Analysis of the characteristics of a stock correlation network is mainly divided into two steps: (1) construction of the stock correlation network, where the correlations of stocks are described by the correlations between the returns of stocks; and (2) analysis of the stock correlation network. Therefore, we introduce some basic concepts and characteristics about stock correlation networks.

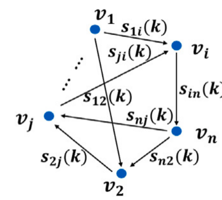


Fig. 1. An illustration of a stock correlation network.

(1) Construction of a stock correlation network

As shown in Fig. 1, the stock correlation network considers each stock as a node, and uses the correlations between stock returns to measure the mutual influence between the nodes. The correlations between stocks represent the degrees of mutual influence between the nodes. Assume that the stock correlation network of m stocks from V_s in period k is $G_s(k) = (V_s, E_s(k))$, where the edge set is $E_s(k) = \{(v_i, v_j)|v_i, v_j \in$

[2] It is worth noting that, the BCRP strategy is a special CGT optimal in an i.i.d. market.

$V_s\}$. Edge (v_i, v_j) means that the performance of stock v_j will influence that of stock v_i. The weighted adjacency matrix of network $G_s(k)$ is $S(k) = \begin{bmatrix} s_{11}(k) & \cdots & s_{1m}(k) \\ \vdots & \ddots & \vdots \\ s_{m1}(k) & \cdots & s_{mm}(k) \end{bmatrix}$, where $s_{ij}(k)$ is the weight of edge (v_i, v_j) which equals the correlation between stocks v_i and v_j. Assume the return series of stocks v_i and v_j are $\boldsymbol{x}$ and $\boldsymbol{y}$, respectively. Next, we introduce several measurements of the correlation between vectors $\boldsymbol{x}$ and $\boldsymbol{y}$.

Pearson correlation: This correlation describes the linear correlation between them, and the correlation is applicable to the case that $\boldsymbol{x}$ and $\boldsymbol{y}$ follow normal or unimodal distributions. The Pearson correlation is calculated as $\frac{Cov(\boldsymbol{x},\boldsymbol{y})}{\sigma_{\boldsymbol{x}}\cdot\sigma_{\boldsymbol{y}}}$, where $\sigma_{\boldsymbol{x}}$ and $\sigma_{\boldsymbol{y}}$ are the standard derivations of $\boldsymbol{x}$ and $\boldsymbol{y}$.

Spearman correlation: This correlation also describes the linear correlation between $\boldsymbol{x}$ and $\boldsymbol{y}$, but it does not require the overall distribution and size of data samples. The Spearman correlation is calculated as $\frac{(\boldsymbol{x}-\mu_{\boldsymbol{x}})^T(\boldsymbol{y}-\mu_{\boldsymbol{y}})}{\|\boldsymbol{x}-\mu_{\boldsymbol{x}}\|\|\boldsymbol{y}-\mu_{\boldsymbol{y}}\|}$, where x_h and y_h are the hth elements of $\boldsymbol{x}$ and $\boldsymbol{y}$, respectively, and $\mu_{\boldsymbol{x}}$ and $\mu_{\boldsymbol{y}}$ are the mean vectors of $\boldsymbol{x}$ and $\boldsymbol{y}$, respectively.

Partial correlation: This correlation measures the correlation between $\boldsymbol{x}$ and $\boldsymbol{y}$ when the influence of common factors is removed. Assume that vector $\boldsymbol{o}$ is taken as the common factor, the partial correlation is defined as $\frac{Pearson_{xy}-Pearson_{xo}\cdot Pearson_{yo}}{\sqrt{(1-Pearson_{xo}^2)(1-Pearson_{yo}^2)}}$, where the symbol $Pearson_{\boldsymbol{xy}}$ denotes the Pearson correlation of $\boldsymbol{x}$ and $\boldsymbol{y}$, and the other symbols have similar meanings.

Mutual information: The mutual information is calculated as $\sum_{\boldsymbol{x},\boldsymbol{y}} p(\boldsymbol{x},\boldsymbol{y}) \log \frac{p(\boldsymbol{x},\boldsymbol{y})}{p_{\boldsymbol{X}}(\boldsymbol{x})p_{\boldsymbol{Y}}(\boldsymbol{y})}$, and it measures the amount of information shared between returns and ranges in $[0, +\infty)$, where $p_{\boldsymbol{X}}(\boldsymbol{x})$ and $p_{\boldsymbol{Y}}(\boldsymbol{y})$ are the marginal probabilities of $\boldsymbol{x}$ and $\boldsymbol{y}$, respectively. $p(\boldsymbol{x},\boldsymbol{y})$ is the joint probability of $\boldsymbol{x}$ and $\boldsymbol{y}$.

The correlation coefficients presented above can characterize the correlations between nodes, but if the correlation coefficients are too small, they will somehow degrade the correlation properties of the network and are considered noise that needs to be eliminated. In addition, the correlations are usually non-zero. If the correlations are not filtered, the stock network is a fully connected network (any two nodes are correlated), which also will increase the noise of the network and reduce the effectiveness of the network structure. Therefore, after determining the measurements of stock correlations, we usually need to use certain noise filtering methods (e.g., a minimum spanning tree (MST) (Mantegna, 1999), planar maximum filter graph (PMFG) (Tumminello et al., 2005), or threshold (Billo et al., 2012)) to filter the noise in the correlations and thus determine the edge weights to construct the stock network. We will not discuss these in detail here.

(2) Analysis of a stock correlation network

Analyzing the characteristics of a stock correlation network can help us understand the relationship or influence between stocks. Based on existing work, we introduce several common centrality characteristics of a stock network: degree centrality, between centrality, closeness centrality, and eigenvector centrality. However, clustering coefficient, modularity, and other network characteristics (Nanda & Panda, 2014; Pai & Michel, 2009; Tola et al., 2008) will not be discussed here.

Degree centrality DC_{v_i}: DC_{v_i} is the degree of stock v_i in the stock correlation network, i.e., $DC_{v_i} = \sum_{v_j \neq v_i} I_{v_i v_j}$, where $I_{v_i v_j} = \begin{cases} 0, & (v_i, v_j) \notin E_s(k) \\ 1, & (v_i, v_j) \in E_s(k) \end{cases}$ is the weighted adjacency matrix of the correlation network. (This is the counterpart of the weighted adjacency matrix $S(k)$.) A high DC_{v_i} means a strong association between stock v_i and other stocks, which can reflect the influence of the stock in the whole network.

Between centrality BC_{v_i}: $BC_{v_i} = \sum_{v_k \neq v_i, v_k \neq v_j} \frac{l_{v_k v_j}(v_i)}{l_{v_k v_j}}$, where $l_{v_k v_j}$ is the number of the shortest path from v_k to v_j, $l_{v_k v_j}(v_i)$ is the number of the shortest path from v_k to v_j through v_i.

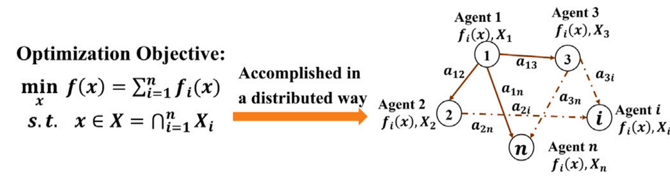


Fig. 2. Multi-agent system with its optimization objective and communication network.

Closeness centrality CC_{v_i}: CC_{v_i} counts the sum of the distances of stock v_i to other stocks in the network, i.e., $CC_{v_i} = \frac{1}{\sum_{v_j \neq v_i} l_{v_i v_j}}$. A big CC_{v_i} means a small overall distance between stock v_i and other stocks, which can also reflect the centrality of stock v_i in the network.

Eigenvector centrality EC_{v_i}: $EC_{v_i} = \lambda^{-1}(k) \sum_{j \neq i} s_{ij}(k) EC_{v_j}$, where $\lambda(k)$ is the largest eigenvalue of $S(k)$ and $(EC_{v_i})_{\{v_1,\ldots,v_m\}}$ is the eigenvector corresponding to $\lambda(k)$.

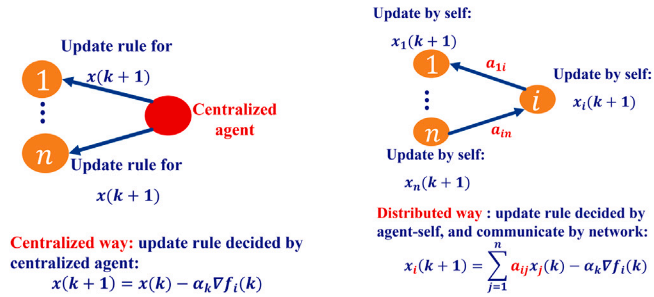
At present, numerous studies have discussed how to use the above stock centralities to comprehensively analyze stock performance. Li et al. (2019) used the above four network centralities to filter stocks and found that the stock portfolio with high centrality is more likely to diversify risk and improve yields. This means we can filter the stocks through the network centralities. Heiberger (2014) found that the network of the S&P500 index has a more centralized topology during crisis periods. In addition, a number of studies have looked at the topological characteristics of stock markets from the perspective of network centrality theory. For example, Wang et al. (2020) found that there is a weak momentum effect and a strong reversion effect in the Chinese stock market.

3.3. Distributed optimization for a multi-agent system

Distributed optimization is important in multi-agent system optimization. It is gradually replacing traditional centralized optimization and is widely used in big data cloud computing, distributed collaborative control, and logistics deployment. A multi-agent system is composed of intelligent agents with certain computing, communication, and sensing capabilities connected through a communication network (Fig. 2). Without a loss of generality, we can assume that the agent and its communication network are the n trading machines in V and the communication network $G(k) = (V, E(k))$ that we set up in the preliminaries, respectively. Each machine in the system has a local cost function $f_i(\boldsymbol{x})$ known only by itself and a set of local constraints X_i. The global objective function is $f(\boldsymbol{x}) = \sum_{i=1}^{n} f_i(\boldsymbol{x})$ and the overall constraint set is $X = \cap_{i=1}^{n} X_i$. Each machine i can exchange information (e.g., a locally computed state) with others through the communication network $G(k)$ in time k.

In the research, optimization algorithms solving multi-agent systems are mainly centralized. A central agent processes and distributes the information of each component, guides other agents to update and iterate the algorithm, and finally minimizes the overall objective function together. For example, if the overall optimization goal of the multi-agent system is to minimize a convex function $f(\boldsymbol{x})$, then the common gradient descent method $\boldsymbol{x}(k+1) = \boldsymbol{x}(k) - \alpha \nabla f(\boldsymbol{x}(k))$ is a special centralized optimization algorithm with $n = 1$ (Fig. 3(a)).

However, a centralized algorithm creates a series of practical problems, such as low utilization rates of agents' computing power, a high requirement of the central agent's computing power, long running times, and the failure of the central agent being equal to the failure of the whole system. To solve these problems, the optimization algorithms of multi-agent systems are gradually transitioning into distributed algorithms. Distributed algorithms do not need a central agent that is responsible for unified scheduling. Each agent updates its state only according to its own objective function and minimizes the objective function jointly and consistently (Fig. 3(b)). For more details on distributed optimization, please refer to the research review by Yang, Yi, et al. (2019).



(a) Centralized gradient descent. (b) Distributed gradient descent.

Fig. 3. The centralized gradient descent approach versus the distributed gradient descent approach.

4. Model setup

To solve the problems mentioned in the introduction, in this section, we construct our DMR strategy model. We first present the construction process of the dynamic stock correlation sub-network and combine it with the distributed algorithm. Then, we propose our DMR strategy to maximize the final cumulative wealth. Finally, we analyze the DMR strategy in terms of the theoretical and empirical aspects. Specifically, we discuss the following:

(1) In Section 4.1, we discuss how to measure and filter the correlations between stocks and how to build the dynamic stock correlation network in each trading period. Furthermore, the most influential stocks are selected to construct a stock correlation sub-network to represent the overall correlation of the stocks.

(2) In Section 4.2, under the assumption of short-term mean reversion and with short selling allowed, we explain how to construct the distributed optimal model of DMR by using a dynamic stock correlation sub-network and how to maximize the final cumulative wealth of the portfolio.

(3) In Section 4.3, we explain how to analyze the universal property of DMR and its theoretical convergence rate. The empirical results are presented in Section 5.

4.1. Dynamic stock correlation sub-network for trading machines

Existing studies show that the nodes with higher centralities correspond to riskier stocks (Peralta & Zareei, 2016) and that high risk may come with high expected returns. This means we can focus on the stocks with high centralities and represent the entire correlation of the portfolio by these stocks in each trading period. Here, we introduce the construction process of our dynamic stock correlation sub-network.

(1) In the kth trading period, for any two stocks v_i and v_j, we use the Pearson correlation of stock relative prices $\boldsymbol{x}_i(k-w+1,k)$ and $\boldsymbol{x}_j(k-w+1,k)$ to measure the correlation $Pearson_{v_iv_j}$ of stocks v_i and v_j, i.e., $Pearson_{v_iv_j} = \frac{Cov(\boldsymbol{x}_i(k-w+1,k),\boldsymbol{x}_j(k-w+1,k))}{\sigma_{\boldsymbol{x}_i(k-w+1,k)}\cdot\sigma_{\boldsymbol{x}_j(k-w+1,k)}}$, where w is the preset length of the historical window period.

(2) We use the threshold $\eta_k = 0$ to filter the edge (v_i, v_j) corresponding to $Pearson_{v_iv_j} \le \eta_k$ and construct the initial stock correlation network $G_s(k) = (V_s, E_s(k))$, where the stock edge set $E_s(k) = \{(v_i, v_j)|\text{if } Pearson_{v_iv_j} > \eta_k, v_i, v_j \in V_s\}$. Denoting the adjacency matrix of $G_s(k)$ as $S(k) = \begin{bmatrix} s_{11}(k) & \cdots & s_{1m}(k) \\ \vdots & \ddots & \vdots \\ s_{m1}(k) & \cdots & s_{mm}(k) \end{bmatrix}$, we know $s_{ij}(k) = \max\{\eta_k Pearson_{v_iv_j}\}$.

(3) We calculate the $DC_{v_i}(k)$ of each stock v_i, sort the m stock nodes in descending order according to $DC_{v_i}(k)$, and filter the top n stock nodes. Without a loss of generality, we can arrange these n points in descending order according to their centralities and assume that they form the top n node set $V_{top}(k) = \{v_1(k), \dots, v_n(k)\}$. Let the machines in V inherit the stock network structure in $V_{top}(k)$ as their communication network structure $E(k)$, i.e., $E(k) = \{(i,j)|(v_i(k), v_j(k)) \in E_s(k), v_i(k), v_j(k) \in V_{top}(k)\}$.

(4) According to $S(k)$ and the definition of $E(k)$, we know that the weighted adjacency matrix of $E(k)$ now is $\begin{bmatrix} s_{11}(k) & \cdots & s_{1n}(k) \\ \vdots & \ddots & \vdots \\ s_{n1}(k) & \cdots & s_{nn}(k) \end{bmatrix}$. We scale it to be a doubly stochastic matrix[3] and denote it as $A(k) = \begin{bmatrix} a_{11}(k) & \cdots & a_{1n}(k) \\ \vdots & \ddots & \vdots \\ a_{n1}(k) & \cdots & a_{nn}(k) \end{bmatrix}$.[4] Finally, we treat $A(k)$ as the weighted adjacency matrix of $E(k)$ and obtain the final stock correlation sub-network $G(k) = (V, E(k))$, which is also the communication network of trading machines.

4.2. Distributed mean reversion online portfolio strategy

To simplify the DMR strategy, we make the following assumptions.

Assumption 4.1. The first difference of closing price $\boldsymbol{p}(t)$ in period t follows the m-dimensional Brownian motion with zero drift, *i.e.*, $\boldsymbol{p}(t) - \boldsymbol{p}(t-1) \sim N_m(\boldsymbol{0}, \Lambda_t)$, where $\boldsymbol{0}$ denotes a 0 vector, Λ_t is the covariance matrix of $\boldsymbol{p}(t-w+1,t)$. So, the relative price in period t follows $\boldsymbol{x}(t) \sim N_m(\boldsymbol{1}, D_t^{-1}\Lambda_t D_t^{-1})$, where $D_t = diag(\boldsymbol{p}(t-1))$.

In Assumption 4.1, we assume that the stock relative price follows the multi-period mean reversion feature. Existing mean reversion strategies are based on the single-period or multiple-period inversion of relative prices to capture the mean reversion of stocks. For example, the PAMR strategy, representing the single-period mean reversion strategy, assumes that the next period relative price $\tilde{\boldsymbol{x}}(t+1) = 1/\boldsymbol{x}(t)$ (see the analysis in the first paragraph of Section 4.1 on Li et al. (2015)). This means that the next period relative price $\boldsymbol{x}(t+1)$ will revert to the previous historical relative price $\frac{1}{\boldsymbol{x}(t)}$. The OLMAR strategy, typical of the multi-period mean reversion strategy, estimates the relative stock price $\tilde{\boldsymbol{x}}(t+1) = (1/w)(1 + 1/\boldsymbol{x}(t) + \cdots + 1/\otimes_{i=0}^{w-2}\boldsymbol{x}(t-i))$. This indicates that the stock relative price $\boldsymbol{x}(t+1)$ in the next period will revert to the average performance of the stock relative price over the past period. However, the price fluctuations of the higher-value stocks may affect the relative price forecast $\tilde{\boldsymbol{x}}(t+1)$ more significantly. This is because the change between the relative price deviation $\Delta\boldsymbol{x}(t)$ and its inverse $1/\Delta\boldsymbol{x}(t)$ is not linear, and the further the relative price $\boldsymbol{x}(t)$ moves away from 1, the more dramatic the inverse $1/\boldsymbol{x}(t)$. For example, the stock prices of Goldman Sachs Group Inc. and Amazon are USD276 and USD98, respectively, which are very different. If stocks in a market fall together, the relative price of a high-priced stock will move further away from 1 compared to a low-value stock, and its inverse will be smaller than that of a low-priced stock. Since the mean reversion strategy treats this inversion as the expected future return, the high-value stock will receive less investment than the low-value stock. We should not abandon high-priced stocks like this, even though we are under the assumption of mean reversion. Note that the prediction of both relative prices of PAMR and OLMAR implies a special case, where the expected relative price of a stock is equal to 1 (let μ be the expected return, we know that $\mu = 1/\mu$ and $\mu = (1/w)[1 + 1/\mu + 1/\mu^2 + +1/\mu^{w-1}]$ have the common solution $\mu = 1$). Therefore, we present Assumption 4.1 and treat 1 as a uniform benchmark to characterize the mean reversion of all stocks. In addition, since the relative prices of stocks are maintained around

[3] A nonnegative matrix is called doubly stochastic if and only if its entries sum up to 1 in each of its rows and columns.

[4] The transformation of the matrix $S(k)$ into a doubly stochastic matrix $A(k)$ is as follows: assume that $S_i(k)$ is the ith column vector of $S(k)$ and $S_i'(k)$ is the ith row vector of $S(k)$, where $i = 1, \dots, m$. Let $\varrho = \max\{\max\{\boldsymbol{1}^T S_i(k)\}_{i=1}^n, \max\{S_j'(k)\boldsymbol{1}\}_{j=1}^n\}$, $\iota = \varrho n - \sum_{i=1}^n\sum_{j=1}^n s_{ij}(k)$, then the element $a_{ij}(k) = \varrho^{-1}s_{ij}(k) + (\varrho\iota)^{-1}(\varrho - S_i'(k)\boldsymbol{1})(\varrho - \boldsymbol{1}^T S_j(k))$.

Algorithm 1: Online portfolio strategy DMR.

Input: window size: $w \geq 1$, market sequence: $\boldsymbol{x}_1, \dots, \boldsymbol{x}_n$, machine number n, step size $\{\alpha_k\}$.
Output: final cumulative wealth: S_n.

initialize: $\boldsymbol{b}(1) = \frac{1}{m}\mathbf{1}$, $S(0) = 1$;
randomly generate the initial strategy of trading machines: $\hat{\boldsymbol{b}}_i(0) \in \Delta_m^2$, $i = 1, \cdots, n$;
for $k = 1, 2, \dots$ **do**
  calculate the weighted adjacency matrix $A(k)$;
  calculate the daily returns and cumulative returns: $S(k) = S(k-1) \times \boldsymbol{b}(k)^T \boldsymbol{x}(k)$;
  for $i = 1, 2, \cdots, n$: **do**
    the trading machine i aggregates strategies from its neighbors: $\hat{\boldsymbol{v}}_i(k) = \sum_{j=1}^{n} a_{ij}(k)\hat{\boldsymbol{b}}_j(k)$;
    the summed strategy $\hat{\boldsymbol{v}}_i(k)$ descends along the subgradient direction: $\hat{\boldsymbol{y}}_i(k) = \hat{\boldsymbol{v}}_i(k) - \alpha_k \hat{\boldsymbol{d}}_i^k(\hat{\boldsymbol{v}}_i(k))$;
    update the next strategy of machine i: $\hat{\boldsymbol{b}}_i(k+1) = P_{\Delta_m^2}[\hat{\boldsymbol{y}}_i(k)]$;
  end
  update DMR strategy: $\boldsymbol{b}(k+1) = \sum_{i=1}^{n} \frac{1}{n}\hat{\boldsymbol{b}}_i(k+1)$;
end

vector 1, it is also possible to use the relative price series of stocks that satisfy this assumption to uniformly characterize the correlations among stocks. This assumption is inspired by Ha and Zhang (2020) who presented the same assumption to study the impact of liquidity risk on OLPSs. Thus, they can restrict the stock relative price in the next period to a random variable with mean 1 and add multiple intra-day trading. The result shows that the impact of market liquidity risk and trading costs on OLPSs has been reduced.

Assumption 4.2. The sequence $\{\alpha_k\}_{k=1}^{+\infty}$ is non-negative mean-square convergent, i.e., $\sum_{k=1}^{+\infty} \alpha_k = +\infty$, $\sum_{k=1}^{+\infty} \alpha_k^2 < +\infty$.

Assumption 4.3. There exists an integer $B \geq 1$ such that the graph $(V, \bigcup_{l=0}^{B-1} E(k+l))$ is strongly connected for all $k \geq 0$.

Assumption 4.3 means that there must be at least one effective connection (non-zero correlation) between any two trading machines during the whole investment period, and this also suggests that one stock will influence another stock at least once in the future.

Assumption 4.4. The stock relative price of the kth period is bounded, *i.e.*, $\exists C_k, \varepsilon_k > 0$, $s.t.\ \varepsilon_k < \|\boldsymbol{x}(k)\| \leq C_k$.

In Assumption 4.4, the upper bound could also be uniform $C = \sup\{C_k\}_{k=1}^{+\infty}$. It guarantees that the wealth will not be infinite, *i.e.*, $\|\mathbb{E}[\ln \boldsymbol{b}^T \boldsymbol{x}(t)]\| \leq +\infty$. This assumption implies the existence of an $L_k > 0, s.t. \|\boldsymbol{d}_j(k)\| \leq L_k$, where $\boldsymbol{d}_j(k)$ is the subgradient of $\ln \boldsymbol{b}^T \boldsymbol{x}(k)$ at point $\boldsymbol{b}$. This upper bound also can be replaced by a uniform bound $L = \sup\{L_k\}_{k=1}^{+\infty}$ for each period k.

Next, we introduce our DMR strategy. The overall goal of DMR is to maximize the final cumulative wealth, i.e.,

$$\min_{\boldsymbol{b} \in \Delta_m^2} - \sum_{t=1}^{T} \ln \boldsymbol{b}^T \boldsymbol{x}(t). \tag{1}$$

For such an investment problem, Cover (1991) pointed out that if the market vector $\{\boldsymbol{x}(1), \dots, \boldsymbol{x}(T)\}$ is known before trading and short selling is not allowed ($\boldsymbol{b} \in \Delta_m^1$), then there theoretically exists a BCRP strategy that can maximize the final cumulative wealth, that is, $\boldsymbol{b}^* = argmin_{\boldsymbol{b} \in \Delta_m^1} - \sum_{k=1}^{T} \ln \boldsymbol{b}^T \boldsymbol{x}(k)$. Denote that $f(\boldsymbol{b}) = -\sum_{t=1}^{T} \ln \boldsymbol{b}^T \boldsymbol{x}(t)$, we know f is a convex function and $\boldsymbol{b}^*$ can be reached by the gradient

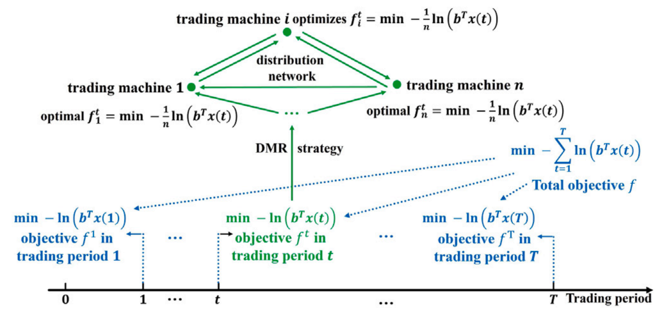


Fig. 4. The objective function of each period split into n parts.

descent projection method, namely $\boldsymbol{b}^*(k+1) = P_{\Delta_m^2}[\boldsymbol{b}(k)^* - \alpha_k \nabla f(\boldsymbol{b}(k)^*)]$, where $\boldsymbol{b}(k)^*$ and α_k are the strategy of the kth iteration and the iteration step, respectively. The initial strategy $\boldsymbol{b}(0)^*$ can be a random point in Δ_m^2.

However, we cannot know the complete market vector in advance, so ∇f is unknown in practice. In addition, allowing for short selling is a more realistic requirement for investors. To achieve the optimization goal with short selling allowed, we adopt the idea of a distributed optimization algorithm and divide problem (1) into several single optimization problems. Then, we use our trading machines to calculate the optimal strategy of each single problem in a distributed approach and propose our DMR strategy (see Fig. 4). Specifically:

First, we allow the n trading machines to simultaneously invest in the current portfolio in the initial trading period, and denote the strategy of each trading machine in period k as $\hat{\boldsymbol{b}}_1(k), \dots, \hat{\boldsymbol{b}}_n(k)$. We split $\min_{\boldsymbol{b} \in \Delta_m^2} - \sum_{t=1}^{T} \ln \boldsymbol{b}^T \boldsymbol{x}(t)$ into T single problems $D_1, \dots, D_T$ and treat $D_k, k = 1, \dots, T$ as the objective in the kth trading period, where D_k is as follows:

$$\begin{aligned} &\min \quad -\ln \boldsymbol{b}^T \boldsymbol{x}(k) \\ &s.t. \quad \boldsymbol{b} \in \Delta_m^2. \end{aligned} \tag{2}$$

Denote $f^k(\boldsymbol{b}) = -\ln(\boldsymbol{b}^T \boldsymbol{x}(k))$ as the objective function of D_k, where $f_i^k(\boldsymbol{b}) = -\frac{1}{n}\ln(\boldsymbol{b}^T \boldsymbol{x}(k))$ is the objective function of machine i in the kth period, and $\hat{\boldsymbol{d}}_i^k(\boldsymbol{b})$ is the subgradient (or gradient) of $f_i^k(\boldsymbol{b})$ at point $\boldsymbol{b}$. For example, we can set $\hat{\boldsymbol{d}}_i^k(\boldsymbol{b}) = \nabla f_i^k(\boldsymbol{b}) = -\frac{1}{n}\frac{\boldsymbol{x}(k)}{\boldsymbol{b}^T \boldsymbol{x}(k)}$.

Second, based on the communication network structure for the trading machines in Section 4.1, the trading machines can jointly carry out distributed strategy updates. The update rule for trading machine i is as follows:

$$\hat{\boldsymbol{v}}_i(k) = \sum_{j=1}^{n} a_{ij}(k)\hat{\boldsymbol{b}}_j(k), \tag{3}$$

$$\hat{\boldsymbol{y}}_i(k) = \hat{\boldsymbol{v}}_i(k) - \alpha_k \hat{\boldsymbol{d}}_i^k(\hat{\boldsymbol{v}}_i(k)), \tag{4}$$

$$\hat{\boldsymbol{b}}_i(k+1) = P_{\Delta_m^2}[\hat{\boldsymbol{y}}_i(k)], \tag{5}$$

where $\alpha_k > 0$ is the descent step size, $P_{\Delta_m^2}[\hat{\boldsymbol{y}}_i(k)]$ is the projection of $\hat{\boldsymbol{y}}_i(k)$ on Δ_m^2, $\hat{\boldsymbol{v}}_i(k)$ is the integration of the strategies of trading machine i's neighbors, and $\hat{\boldsymbol{b}}_i(k+1)$ is the strategy of machine i in the next period.

In Eqs. (3)–(5), we update the portfolio strategies among the trading machines. Each trading machine communicates its own strategy with its neighbors and evaluates the mutual influence of its own strategy with others through network $G(k)$. Specifically, machine i measures the influence of its neighbors' portfolio strategies on its future returns according to the weighted adjacency matrix $A(k)$, and then determines the next portfolio strategy $\hat{\boldsymbol{b}}_i(k+1)$.

Finally, all the next portfolio strategies $\{\boldsymbol{b}_i(k+1)\}_{i=1}^{n}$ of the trading machines are weighted to obtain the DMR strategy $\boldsymbol{b}(k+1)$, that is, $\boldsymbol{b}(k+1) = \sum_{i=1}^{n} \frac{1}{n}\hat{\boldsymbol{b}}_i(k+1)$, where the initial strategy $\boldsymbol{b}(0) = \frac{1}{m}\mathbf{1}$. The detailed process is shown in Algorithm 1, and for the initial strategies of n trading machines, we make the following assumptions:

Assumption 4.5. The initial strategies $\{\hat{\boldsymbol{b}}_1(0), \dots, \hat{\boldsymbol{b}}_n(0)\} \subset \Delta_m^2$ of trading machines are i.i.d.

Note that in the real world, we do not really need n trading machines to run DMR because most investors do not really need to make large-scale trades, and the computing resources of individual investors are usually limited. We can use virtual machines or parallel computing to set up multiple trading machines. For example, our empirical tests in this paper are built using virtual machine technology. The calculation and updating of all trading machines are completed by the simulation of only one computer. In addition, a suitable application scenario for the DMR strategy is a group investment scenario. The DMR strategy with n trading machines can easily be adapted to the case where n investors invest simultaneously in the same assets pool. Specifically, each investor i independently executes the investment strategy $\hat{\boldsymbol{b}}_i(k)$ on trading machine i in period k. The convergence analysis of DMR in the following shows that the n investors can theoretically obtain the same final cumulative wealth.

4.3. Convergence analysis

In this part, we analyze the theoretical properties of DMR. First, we demonstrate that the strategies of all trading machines are feasible in Theorem 4.1. Second, we prove that the strategy of each trading machine will converge to a BCRP strategy in Theorem 4.2. Third, in Theorem 4.3, we show that the DMR strategy and the strategy of each trading machine are universal. Finally, we provide the convergence rate from the DMR strategy to the BCRP strategy in Proposition 4.1.

Theorem 4.1. *For any machine i, the sequence $\{\hat{\boldsymbol{v}}_i(k)\}$ generated by Eqs. (3)–(5) converges, and $\sum_{k=0}^{+\infty} dist^2(\hat{\boldsymbol{v}}_i(k); \Delta_m^2) < +\infty$ a.s.*

Proof. According to Lemma 7.1 in the Appendix, or similar to Lemma 3.1 in Ram et al. (2010), for $\forall\ \bar{\boldsymbol{b}} \in \Delta_m^2,\ \boldsymbol{z} \in R^m$, in trading period k, we have

$$\begin{aligned}\|\hat{\boldsymbol{b}}_i(k+1) - \bar{\boldsymbol{b}}\|^2 &= \|P_{\Delta_m^2}[\hat{\boldsymbol{y}}_i(k)] - \bar{\boldsymbol{b}}\|^2\\ &\le (1+2\alpha_k^2 L^2)\|\hat{\boldsymbol{v}}_i(k) - \bar{\boldsymbol{b}}\|^2 - 2\alpha_k(f_i^k(\boldsymbol{z}) - f_i^k(\bar{\boldsymbol{b}}))\\ &\quad + r_\eta \alpha_k^2 L^2 + (\frac{1}{4\eta} + 2\alpha_k L)\|\hat{\boldsymbol{v}}_i(k) - \boldsymbol{z}\|^2 - \|\hat{\boldsymbol{b}}_i(k+1) - \hat{\boldsymbol{v}}_i(k)\|^2, \end{aligned} \tag{6}$$

where $\gamma_\eta = 4 + 16\eta,\ \eta$ is a positive scalar.

Let $\boldsymbol{z} = \bar{\boldsymbol{b}} = P_{\Delta_m^2}[\hat{\boldsymbol{v}}_i(k)]$, by the definition of the distance from point to set, we get:

$$\begin{aligned}dist^2(\hat{\boldsymbol{b}}_i(k+1); \Delta_m^2) &\le \|\hat{\boldsymbol{b}}_i(k+1) - P_{\Delta_m^2}[\hat{\boldsymbol{v}}_i(k)]\|^2\\ &\le (1+2\alpha_k^2 L^2) dist^2(\hat{\boldsymbol{v}}_i(k); \Delta_m^2) - \|\hat{\boldsymbol{b}}_i(k+1) - \hat{\boldsymbol{v}}_i(k)\|^2\\ &\quad + r_\eta \alpha_k^2 L^2 + (\frac{1}{4\eta} + 2\alpha_k L) dist^2(\hat{\boldsymbol{v}}_i(k); \Delta_m^2). \end{aligned} \tag{7}$$

Denote $\mathcal{F}_k = \{\hat{\boldsymbol{b}}_i(s); i \in V, 0 \le s \le k-1\}$ as the σ-algebra induced by the entire history of the algorithm up to period $k-1$. Therefore, given $\mathcal{F}_k$, the collection $\hat{\boldsymbol{b}}_i(0), \dots, \hat{\boldsymbol{b}}_i(k)$ and $\hat{\boldsymbol{v}}_i(0), \dots, \hat{\boldsymbol{v}}_i(k)$ generated by Eqs. (3)–(5) are fully determined. Take the conditional expectation of the above equation with respect to $\mathcal{F}_k$, we obtain

$$\begin{aligned}&\mathbb{E}[dist^2(\hat{\boldsymbol{b}}_i(k+1); \Delta_m^2)|\mathcal{F}_k]\\ &\le (1+2\alpha_k^2 L^2) dist^2(\hat{\boldsymbol{v}}_i(k); \Delta_m^2) - \|\hat{\boldsymbol{b}}_i(k+1) - \hat{\boldsymbol{v}}_i(k)\|^2\\ &+ r_\eta \alpha_k^2 L^2 + (\frac{1}{4\eta} + 2\alpha_k L) dist^2(\hat{\boldsymbol{v}}_i(k); \Delta_m^2)\\ &\le (1+2\alpha_k^2 L^2)\sum_{j=1}^n a_{ij}(k) dist^2(\hat{\boldsymbol{b}}_j(k); \Delta_m^2) + r_\eta \alpha_k^2 L^2\\ &+ (\frac{1}{4\eta} + 2\alpha_k L) dist^2(\hat{\boldsymbol{v}}_i(k); \Delta_m^2)\\ &- \|\hat{\boldsymbol{b}}_i(k+1) - \hat{\boldsymbol{v}}_i(k)\|^2. \end{aligned} \tag{8}$$

For $dist^2(\hat{\boldsymbol{v}}_i(k); \Delta_m^2) = dist^2(\sum_{j=1}^n a_{ij}(k)\hat{\boldsymbol{b}}_j(k); \Delta_m^2) \le \sum_{j=1}^n a_{ij}(k) dist^2(\hat{\boldsymbol{b}}_j(k); \Delta_m^2)$, and $\|\hat{\boldsymbol{b}}_i(k+1) - \hat{\boldsymbol{v}}_i(k)\|^2 \ge dist^2(\hat{\boldsymbol{v}}_i(k); \Delta_m^2)$, we get

$$\begin{aligned}\mathbb{E}[dist^2(\hat{\boldsymbol{b}}_i(k+1); \Delta_m^2)|\mathcal{F}_k] \le &(1+2\alpha_k^2 L^2)\sum_{j=1}^n a_{ij}(k) dist^2(\hat{\boldsymbol{b}}_j(k); \Delta_m^2) + r_\eta \alpha_k^2 L^2\\ &- (\frac{3}{4\eta} - 2\alpha_k L) dist^2(\hat{\boldsymbol{v}}_i(k); \Delta_m^2). \end{aligned} \tag{9}$$

Sum the distance squares from $i = 1$ to n, then we have

$$\begin{aligned}\mathbb{E}[\sum_{i=1}^n dist^2(\hat{\boldsymbol{b}}_i(k+1); \Delta_m^2)|\mathcal{F}_k] \le &(1+2\alpha_k^2 L^2)\sum_{i=1}^n dist^2(\hat{\boldsymbol{b}}_i(k); \Delta_m^2) + n r_\eta \alpha_k^2 L^2\\ &- (\frac{3}{4\eta} - 2\alpha_k L)\sum_{i=1}^n dist^2(\hat{\boldsymbol{v}}_i(k); \Delta_m^2). \end{aligned} \tag{10}$$

According to Assumption 4.2, we know $\alpha_k^2 \to 0$, further $\alpha_k \to 0$. This means that, $\exists k_0 > 0, s.t.\ \forall k > k_0$, $\frac{3}{4\eta} - 2\alpha_k L > 0$. Therefore, Eq. (10) satisfies the Sup-Martingale convergence theorem (Ram et al., 2010). We infer the following: (1) the sequence $\{\hat{\boldsymbol{b}}_i(k)\}$ converges, and (2) $\sum_{k=0}^{+\infty} dist^2(\hat{\boldsymbol{v}}_i(k); \Delta_m^2) < +\infty$ *a.s.*, which means that $\boldsymbol{v}_i(k)$ converges to Δ_m^2.

Theorem 4.1 means that the portfolio strategies $\{\hat{\boldsymbol{b}}_i(k)\}_{i=1}^n$ of all trading machines will eventually gather inside the feasible set Δ_m^2, but it is not enough to explain the relationship between $\{\hat{\boldsymbol{b}}_i(k)\}_{i=1}^n$ and the universal property of the DMR strategy. Next, we first prove that $\hat{\boldsymbol{b}}_i(k)$ converges to the BCRP strategy $\boldsymbol{b}^*$ in Theorem 4.2. Then, the return growth rate of strategy $\hat{\boldsymbol{b}}_i(k)$ will naturally converge to the return growth rate of this BCRP strategy $\boldsymbol{b}^*$, which will be addressed in Theorem 4.3.

Theorem 4.2. *If the optimal set B^* is nonempty, then $\forall i \in V$, $\exists \boldsymbol{b}^* \in B^*$, s.t. $\lim_{k\to\infty} \hat{\boldsymbol{b}}_i(k) = \boldsymbol{b}^*$.*

Proof. See the detailed proof in Appendix A.

This theorem means that the portfolio strategy of each trading machine will converge to the BCRP strategy $\boldsymbol{b}^*$. It is natural that the DMR strategy $\boldsymbol{b}(k)$ will also converge to $\boldsymbol{b}^*$. However, this does not mean that they are universal strategies. Consider that $\boldsymbol{b}(k)$ is obtained by integrating the strategy $\hat{\boldsymbol{b}}_i(k)$ from n trading machines, therefore, we only need to prove that the strategy $\hat{\boldsymbol{b}}_i(k)$ is universal. We address this problem using the following theorem.

Theorem 4.3. *For any trading machine $i \in V$, its strategy $\{\hat{\boldsymbol{b}}_i(k)\}_{k=1}^T$ is universal, i.e., $\frac{1}{T}\left[\ln(S_T^*) - \ln(S_T^i)\right] \to 0\ (T \to +\infty)$, where $S_T^* = \prod_{t=1}^T \boldsymbol{b}^{*T}\boldsymbol{x}(t)$ is the final cumulative wealth of a BCRP strategy $\boldsymbol{b}^*$, $S_T^i = \prod_{t=1}^T \hat{\boldsymbol{b}}_i(t)^T\boldsymbol{x}(t)$ is the final cumulative wealth of trading machine i.*

Proof. We have

$$\ln(S_T^*) - \ln(S_T^i) = \sum_{t=1}^T \left[\ln(\boldsymbol{b}^{*T}\boldsymbol{x}(t)) - \ln(\hat{\boldsymbol{b}}_i(t)^T\boldsymbol{x}(t))\right].$$

Take the Taylor expansion of $\ln(\hat{\boldsymbol{b}}_i(t)^T\boldsymbol{x}(t))$ at $\boldsymbol{b}^*$, we get $\ln(\hat{\boldsymbol{b}}_i(t)^T\boldsymbol{x}(t)) \ge \ln(\boldsymbol{b}^{*T}\boldsymbol{x}(t)) + \frac{(\hat{\boldsymbol{b}}_i(t)-\boldsymbol{b}^*)^T\boldsymbol{x}(t)}{\boldsymbol{b}^{*T}\boldsymbol{x}(t)}$. Substitute it into the above equation, then

$$\ln(S_T^*) - \ln(S_T^i) \le \sum_{t=1}^T \left[\frac{(\boldsymbol{b}^* - \hat{\boldsymbol{b}}_i(t))^T\boldsymbol{x}(t)}{\boldsymbol{b}^{*T}\boldsymbol{x}(t)}\right].$$

According to Assumption 4.4, $\boldsymbol{x}(t)$, $\boldsymbol{b}^*$, and $\hat{\boldsymbol{b}}_i(t)$ are bounded, then $\inf\{\boldsymbol{b}^{*T}\boldsymbol{x}(t)\}_{t=1}^{+\infty}$ and $\max\{\boldsymbol{b}^{*T}\boldsymbol{x}(t)\}_{t=1}^{+\infty}$ are bounded too. By Theorem 4.2, $\lim_{k\to+\infty}\hat{\boldsymbol{b}}_i(k) = \boldsymbol{b}^*$ *a.s.*, since $C = \sup\{C_k\}_{k=1}^{+\infty}$ and $\lim_{T\to+\infty}\frac{1}{T} = 0$, then for $\forall\varepsilon > 0$, $\exists\ N_1, N_2 > 0$, let $N = \max\{N_1, N_2\}$, when $t > N$, we have $\|\hat{\boldsymbol{b}}_i(t) - \boldsymbol{b}^*(t)\| \le \varepsilon \cdot \frac{\inf\{\boldsymbol{b}^{*T}\boldsymbol{x}(t)\}_{t=1}^{+\infty}}{C}$, $\frac{1}{T} \le \varepsilon \cdot \frac{\inf\{\boldsymbol{b}^{*T}\boldsymbol{x}(t)\}_{t=1}^{+\infty}}{\max\{\|\boldsymbol{b}^* - \hat{\boldsymbol{b}}_i(t)\|\}_{t=1}^N \cdot C}$, respectively. Then, we obtain

$$\frac{1}{T}\left[\ln(S_T^*)-\ln(S_T^i)\right] \le \frac{1}{T}\sum_{t=1}^{N}\left[\frac{(\boldsymbol{b}^*-\hat{\boldsymbol{b}}_i(t))^T\boldsymbol{x}(t)}{\boldsymbol{b}^{*T}\boldsymbol{x}(t)}\right]+\frac{1}{T}\sum_{t=N+1}^{T}\left[\frac{(\boldsymbol{b}^*-\hat{\boldsymbol{b}}_i(t))^T\boldsymbol{x}(t)}{\boldsymbol{b}^{*T}\boldsymbol{x}(t)}\right]$$
$$\le \frac{1}{T}\sum_{t=1}^{N}\frac{\|\boldsymbol{b}^*-\hat{\boldsymbol{b}}_i(t)\|\cdot\|\boldsymbol{x}(t)\|}{\boldsymbol{b}^{*T}\boldsymbol{x}(t)}+\frac{1}{T}\sum_{t=N+1}^{T}\frac{\|\boldsymbol{b}^*-\hat{\boldsymbol{b}}_i(t)\|\cdot\|\boldsymbol{x}(t)\|}{\boldsymbol{b}^{*T}\boldsymbol{x}(t)}$$
$$\le \frac{1}{T}\sum_{t=1}^{N}\frac{\|\boldsymbol{b}^*-\hat{\boldsymbol{b}}_i(t)\|\cdot C}{\inf\{\boldsymbol{b}^{*T}\boldsymbol{x}(t)\}_{t=1}^{N}}+\frac{1}{T}\sum_{t=N+1}^{T}\frac{\|\boldsymbol{b}^*-\hat{\boldsymbol{b}}_i(t)\|\cdot C}{\inf\{\boldsymbol{b}^{*T}\boldsymbol{x}(t)\}_{t=1}^{+\infty}}$$
$$\le \frac{N}{T}\cdot\frac{\max\{\|\boldsymbol{b}^*-\hat{\boldsymbol{b}}_i(t)\|\}_{t=1}^{N}\cdot C}{\inf\{\boldsymbol{b}^{*T}\boldsymbol{x}(t)\}_{t=1}^{+\infty}}+\frac{T-N}{T}\cdot\varepsilon$$
$$\le \frac{N}{T}\cdot\varepsilon+\varepsilon-\frac{N}{T}\cdot\varepsilon=\varepsilon.$$

This means $\frac{1}{T}\left[\ln(S_T^*)-\ln(S_T^i)\right]\to 0$ $(T\to+\infty)$, so the strategy of any trading machine i is universal.

By the convexity of the norm and the theorem above, it is self-evident that the DMR strategy is universal. Before we set up the experiments using real market data, we first analyze the theoretical convergence rate of the DMR strategy in Proposition 4.1.

Proposition 4.1. *Assume $\alpha_k=\frac{1}{(k+a)^q}$, where $a\ge 0, q\in(0,1]$. When $q=1$, the convergence rate of $f(\tilde{\boldsymbol{b}}(k))$ to f^* is $O(\frac{1}{\ln k})$; when $q\in(0,1)$, the convergence rate of that is $O(\frac{1}{k^{1-q}})$.*

Proof. By Theorem 4.2,

$$\sum_{i=1}^{n}\|\hat{\boldsymbol{b}}_i(k+1)-\boldsymbol{b}^*\|^2 \le (1+2\alpha_k^2L^2)\sum_{i=1}^{n}\|\hat{\boldsymbol{b}}_i(k)-\boldsymbol{b}^*\|^2-2\alpha_k(f(\bar{\boldsymbol{z}}(k))-f^*)$$
$$+4\alpha_kL\sum_{i=1}^{n}\max_{i\in V}\{\|\boldsymbol{v}_i(k)-\bar{\boldsymbol{v}}(k)\|\}$$
$$+nr_\eta\alpha_k^2L^2+(2\alpha_kL-\frac{3}{4\eta})\sum_{i=1}^{n}dist^2(\boldsymbol{v}_i(k);\Delta_m^2). \tag{11}$$

Iterating it back to the initial time, we get

$$\sum_{i=1}^{n}\|\hat{\boldsymbol{b}}_i(k+1)-\boldsymbol{b}^*\|^2 \le (1+2\alpha_k^2L^2)\sum_{i=1}^{n}\|\hat{\boldsymbol{b}}_i(k)-\boldsymbol{b}^*\|^2-2\alpha_k(f(\bar{\boldsymbol{z}}(k))-f^*)$$
$$+4\alpha_kL\sum_{i=1}^{n}\max_{i\in V}\{\|\boldsymbol{v}_i(k)-\bar{\boldsymbol{v}}(k)\|\}$$
$$+nr_\eta\alpha_k^2L^2+(2\alpha_kL-\frac{3}{4\eta})\sum_{i=1}^{n}dist^2(\boldsymbol{v}_i(k);\Delta_m^2)$$
$$\le\sum_{\tau=0}^{k}(1+2\alpha_\tau^2L^2)\sum_{i=1}^{n}\|\hat{\boldsymbol{b}}_i(\tau)-\boldsymbol{b}^*\|^2+\sum_{\tau=0}^{k}nr_\eta\alpha_\tau^2L^2$$
$$+\sum_{\tau=0}^{k}4\alpha_\tau L^2\sum_{i=1}^{n}\max_{i\in V}\{\|\boldsymbol{v}_i(\tau)-\bar{\boldsymbol{v}}(\tau)\|\}$$
$$-\sum_{\tau=0}^{k}2\alpha_\tau(f(\bar{\boldsymbol{z}}(\tau))-f^*)$$
$$+\sum_{\tau=0}^{k}(2\alpha_\tau L-\frac{3}{4\eta})\sum_{i=1}^{n}dist^2(\boldsymbol{v}_i(\tau);\Delta_m^2). \tag{12}$$

By the convex property of f, we know $\frac{\sum_{\tau=0}^{k}\alpha_\tau f(\bar{\boldsymbol{z}}(\tau))}{\sum_{\tau=0}^{k}\alpha_\tau}-f^*\le\frac{\rho_1(k)+\rho_2(k)+\rho_3(k)}{2\sum_{\tau=0}^{k}\alpha_\tau}$, where

$$\rho_1(k)=\sum_{\tau=0}^{k}(1+2\alpha_\tau^2L^2)\sum_{i=1}^{n}\|\hat{\boldsymbol{b}}_i(\tau)-\boldsymbol{b}^*\|^2+\sum_{\tau=0}^{k}nr_\eta\alpha_\tau^2L^2,$$
$$\rho_2(k)=\sum_{\tau=0}^{k}4\alpha_\tau L^2\sum_{i=1}^{n}\max_{i\in V}\{\|\boldsymbol{v}_i(\tau)-\bar{\boldsymbol{v}}(\tau)\|\},$$
$$\rho_3(k)=\sum_{\tau=0}^{k}(2\alpha_\tau L-\frac{3}{4\eta})\sum_{i=1}^{n}dist^2(\boldsymbol{v}_i(\tau);\Delta_m^2).$$

Let $\tilde{\boldsymbol{b}}_i(k)=\frac{\sum_{\tau=0}^{k}\alpha_\tau\bar{\boldsymbol{z}}(\tau)}{\sum_{\tau=0}^{k}\alpha_\tau}$, then $f(\tilde{\boldsymbol{b}}(k))-f^*\le\frac{\rho_1(k)+\rho_2(k)+\rho_3(k)}{2\sum_{\tau=0}^{k}\alpha_\tau}$. We set $\alpha_k=\frac{1}{(k+a)^q}, a\ge 0, q\in(0,1]$. Then

$$\sum_{\tau=0}^{k}\frac{1}{(\tau+a)^q}\le\int_0^{k+1}\frac{ds}{(s+a)^q}=\begin{cases}\frac{1}{1-q}[(k+1+a)^{1-q}-a^{1-q}], & 0<q<1\\ \ln\frac{k+a+1}{a}, & q=1.\end{cases}$$

According to the boundedness of a convergent sequence, $\forall k>0$, $\exists M>0$, $s.t.$ $M\ge\frac{\rho_1(k)+\rho_2(k)+\rho_3(k)}{2}$ and we obtain

$$f(\tilde{\boldsymbol{b}}(k))-f^*\le\begin{cases}(1-q)M/[(k+1+a)^{1-q}-a^{1-q}], & 0<q<1\\ M/\ln\frac{k+a+1}{a}, & q=1.\end{cases}$$

So, if $q=1$, the convergence rate of $f(\tilde{\boldsymbol{b}}(k))$ to f^* is $O(\frac{1}{\ln k})$; if $q\in(0,1)$, the convergence rate of that is $O(\frac{1}{k^{1-q}})$.

5. Algorithmic trading empirical analysis using real datasets

In this section, we report the empirical studies for DMR and other OLPSs, including the experimental platform and equipment, dataset descriptions, experimental design, and detailed experimental results and analyses. The codes and data can be found in our repository,[5] and more basic codes and descriptions can be found in Li et al. (2015).

5.1. Dataset descriptions

All the datasets are constructed of daily stock closing prices from indexes such as MSCI, NYSE, and FS100. All price data come from Choice (the financial terminal of the Oriental Fortune website). Detailed information on the datasets are shown in Table 1, where "Dataset Volatility" is defined as $\frac{1}{m}\sum_{i=1}^{m}\sqrt{\frac{\sum_{t=1}^{T}(p_i(t)-\mu_i)^2}{T}}$, and μ_i is the mean of stock i. Table 2 presents a more detailed data structure of the datasets.

We build four new datasets for empirical study: FS100, DJIA01, NYSE_Onew, and NYSE_Nnew. For the six classic online portfolio datasets (MSCI, SP500, TSE, DJIA, NYSE_O, and NYSE_N), we do not present their empirical results in the main text because the time horizons of these datasets are very long and the contents have changed to some extent (e.g., some stocks have been delisted or renamed and re-listed). The first dataset FS100 is composed of stocks from the FTSE 100 index in the UK. The second dataset DJIA01 is updated from the original DJIA dataset, which now contains 29 stocks instead of 30 due to the delisting of one stock in the DJIA index. For the last two datasets, NYSE_Nnew and NYSE_Onew, we construct them with 37 and 36 randomly selected stocks from the top 200 well-known stocks of the NYSE NASDAQ Market and the OTC system, respectively, and treat them as extensions of the classical datasets NYSE_N and NYSE_O.

5.2. Algorithm setup and equipment for empirical testing

We set the number of trading machines $n=10$ for the DMR strategy. For any trading machine i, its initial portfolio strategy $\hat{\boldsymbol{b}}_i(0)$ is randomly generated from Δ_m^1, and the historical window size $w=4$. The computer equipment we used has an AMD Ryzen 5 3500X 3.6ghz CPU and the code is built using Python 3.8. In addition, the default parameters of all comparison strategies are used as follows:

- DMR ($w=4$, $n=10$, $\alpha_k=\frac{1}{k+1000}$): DMR strategy.
- DMR (UBAH) ($w=4$, $n=10$, $\alpha_k=\frac{1}{k+1000}$): DMR strategy, where the initial strategy of each trading machine is the UBAH strategy.
- UP (Cover, 1991): Universal portfolio strategy.
- EG ($\eta=0.05$) (Helmbold et al., 1998): Exponential gradient strategy.

5 https://github.com/jasminesor/DMR

Table 1
Dataset descriptions.

Dataset	Time	Stocks	Description	Days	Dataset Volatility
FS100	20191023–20211015	101	Britain's FTSE 100 index	502	51.50
DJIA01	20150102–20171229	29	Dow Jones Composite Stocks	755	3.42
NYSE_Nnew	20061002–20211015	37	NYSE NASDAQ Stock Market Top 200	3787	37.23
NYSE_Onew	20061002–20211015	36	NYSE OTC Counter System Top 200	3787	10.92

Table 2
The data structure of datasets.

Dataset	Min mean	Max mean	Dataset mean	Min volatility	Max volatility	Dataset volatility
FS100	41.1953	11 993.1952	2017.5375	117.4135	6841111.139	51.50
DJIA01	27.7086	198.5133	91.7978	2.8590	2264.2091	3.42
NYSE_Nnew	9.69507	557.4604	64.2579	13.1422	145 542.3366	37.23
NYSE_Onew	4.41144	1 031.8696	141.2763	26.9534	938 429.3044	10.92

Table 3
Selection of evaluation index.

Criteria	Performance metrics	
Absolute return	Cumulative wealth (CW)	Annualized return (AR)
Risk	Maximum drawdown (MDD)	
Risk-adjusted return	Sharpe ratio (Sharpe)	Calmar ratio (CR)

Table 4
Runtime (in seconds) comparison in average investment period.

Runtime	FS100	DJIA01	NYSE_Nnew	NYSE_Onew
UP	0.0011	0.0004	0.0004	0.0004
EG	0.0004	0.0004	0.0004	0.0004
EGS	0.0007	0.0006	0.0006	0.0006
WAEG	0.0040	0.0039	0.0041	0.0041
MAEG	0.0027	0.0028	0.0031	0.0030
DMR (UBAH)	0.0091	0.0047	0.0053	0.0051
DMR	0.0091	0.0048	0.0053	0.0051

- EGS ($\eta = 0.05$) (Yang, He, et al., 2019): Exponential gradient strategy with side-information.
- WAEG ($\eta = 0.05$, $\eta_k = \{0.01, 0.02, \dots, 0.2\}$) (Yang et al., 2022): Weakly aggregating exponential gradient strategy.
- MAEG ($w = 30$, $\eta_k = \{0.01, 0.02, \dots, 0.2\}$) (Zhang et al., 2022): Moving-window-based adaptive exponential gradient strategy.

Next, we evaluate the performance of DMR and other universal strategies from the perspective of return ability, risk tolerance, and risk-adjusted return ability. The specific evaluation indicators are shown in Table 3. In addition, Table 4 lists the average runtime of all tested strategies in a single trading period on each dataset. The runtime of DMR on all datasets is within 1 second,[6] meaning that DMR ensures that investors can complete their online trading tasks almost instantly.

There are some practical problems that investors need to face in the real world. Therefore, before conducting our empirical study, we need to consider these practical problems to facilitate the subsequent empirical tests.

The first problem is transaction costs, which are commissions and taxes that investors pay to exchanges and governments when they buy or sell certain shares of an asset. The trading markets in different countries and regions may have different specific regulations on transaction costs, but for investors, such fees are always unavoidable. Therefore, in the study of OLPS, it is usually assumed that there is no transaction cost in order to conduct a "fair" analysis of the strategy. As such, if it is not explicitly stated, the transaction cost charged by the market will be assumed to be 0 in the subsequent empirical analyses. We discuss the impact of transaction cost on strategy returns separately (see Fig. 7). We consider the proportional transaction cost model (Blum & Kalai, 1999), that is, it is assumed that investors need to use a fixed proportion of the trading amount to pay for the transaction cost in every trading period. Specifically, the holding shares of asset i in period t is $\tilde{b}_i(t) = \frac{b_i(t)\cdot x_i(t)}{\boldsymbol{b}(t)^T \boldsymbol{x}(t)}$. Assume that the portfolio strategy in the next period is $\boldsymbol{b}(t+1)$ and the fixed transaction cost ratio is γ, the adjusted amount of asset i in the next period is $b_i(t+1) - \tilde{b}_i(t)$. Then, after deducting the transaction cost, the cumulative wealth in period n is $S_n^\gamma = S_0 \prod_{t=1}^n \left[\boldsymbol{b}(t)^T \boldsymbol{x}(t) \times (1 - \frac{\gamma}{2} \times \sum_{i=1}^m |b_i(t+1) - \tilde{b}_i(t)|)\right]$.

Another problem is the restrictions on the short selling of assets. Short selling is allowed in DMR, but it is not allowed in the existing OLPSs. Therefore, in the subsequent empirical studies, the comparison results of strategies are obtained on the premise that short selling is not allowed, except in DMR and DMR(UBAH). We also compare the return and risk result of the DMR(UBAH) strategy when short selling is allowed (SSA) and when short selling is not allowed (SSNA). Detailed results can be found in Table 7. In addition, there are other practical problems, such as margin trading (Edirisinghe et al., 2021) and limit order book (LOB) (Matthias, 2022), which are not considered in this paper.

Table 5
Cumulative wealth comparison across different datasets.

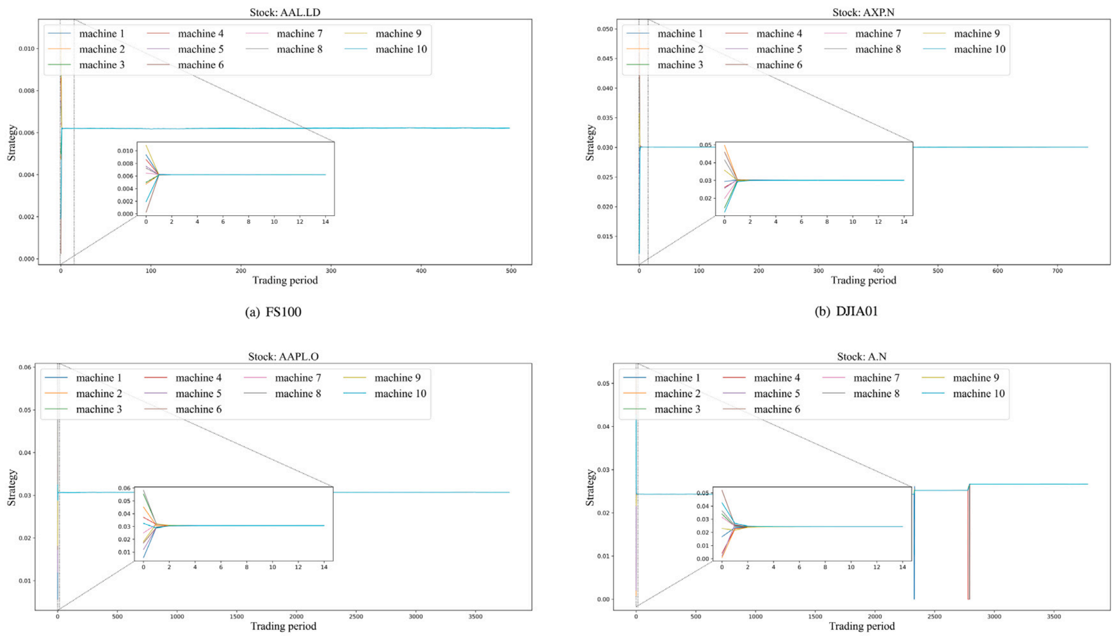
CW	FS100	DJIA01	NYSE_Nnew	NYSE_Onew
UP	1.21974	1.28464	4.210701	14.38129
EG	1.21741	1.28586	1.529821	2.483124
EGS	1.22527	1.28840	4.152956	13.72941
WAEG	1.21405	1.28533	1.510368	2.461698
MAEG	1.21955	1.27910	1.546756	2.511206
DMR (UBAH)	1.21961	1.28613	4.280618	14.60321
DMR	1.22559	1.30600	4.305088	15.09632

5.3. Experimental results

5.3.1. Cumulative wealth

Table 5 shows the results of final cumulative wealth (CW) for all the strategies on each dataset. The results show that DMR outperforms other universal strategies and shows a higher return advantage. When the initial strategy of each trading machine is UBAH, although the final cumulative wealth of DMR(UBAH) decreases to a certain extent, its numerical performance is not different from that of DMR. The reason being that the UBAH strategy usually needs to hold assets in a longer term and cannot differentiate itself in the short term (window size $w = 4$) from DMR. This result also shows that the return of DMR is not sensitive to the initial strategy.

[6] We used virtual technology in the experiment. Although we assume that the number of trading machines is 10, the computation tasks of these 10 trading machines are actually completed by the same computer. Therefore, the runtime of DMR or DMR(UBAH) in Table 4 is the average runtime of a single trading period of all machines divided by 10.



(a) FS100 (b) DJIA01

(c) NYSE_Onew (d) NYSE_Nnew

Fig. 5. Convergence process of each trading machine's portfolio strategy.

5.3.2. *Convergence process of the machine strategy*

We randomly select one stock from each dataset as the observation object and plot the holding process of n trading machines on it in Fig. 5. The results show that the strategies of all trading machines gradually converge to a consistent strategy with the increase of the trading period. Theorem 4.3 tells us that this consistent strategy is, in fact, a universal strategy.

5.3.3. *Risk-adjusted return*

The results in Fig. 6 illustrate the ability of risk tolerance and the risk-adjusted return of each strategy. The results in Fig. 6(a) show that the annualized return of DMR is higher than those of the other strategies, except the EGS strategy on dataset **FS100**. If the initial strategy of the trading machine is UBAH, its annual return changes slightly but is not significant. This preliminarily finding indicates that DMR is not sensitive to the initial value of trading machines and the return is relatively stable. On one hand, in Fig. 6(b), the maximum drawdown (MDD) of DMR is lower than those of the others and has a further reduction than that of DMR(UBAH) on four datasets. This means that DMR has a higher risk tolerance and the selection of the initial strategy can affect the MDD of it to a certain extent. On the other hand, the results in Fig. 6(c) and Fig. 6(d) show that the return of DMR under unit risk (Sharpe ratio) and unit retreat risk (Calmar ratio) are both higher than those of the others on dataset **NYSE_Onew**, **NYSE_Nnew**, and **FS100**. For dataset **DJIA01**, the DMR has a similar $Sharpe$ performance to other strategies. Considering the performance in the above MDD comparison, it shows that the ability of DMR to generate unit return for investors is more stable than the others. In addition, the MDD result also indicates that if a group of investors needs to use the DMR strategy to achieve the same expected return, they should try to use different initial portfolio strategies, which would expose them to less drawdown risk.

To further support the above findings, we refer to the method in Li et al. (2015) to verify the robustness of our strategy's performance and test the α factor of DMR. For the α factor, we regress the daily returns $S_{s,t}$ of DMR on the daily returns $S_{m,t}$ of the market (CRP strategy with $\boldsymbol{b}(t) = (\frac{1}{m}, \dots, \frac{1}{m})$) by the regression equation $S_{s,t} - r_f = \alpha + \beta(S_{m,t} - r_f) + \varepsilon_t$, where r_f is the daily returns of risk-free assets (here, we simply set it to 0). Assuming that α is 0, the strategy does not get excess return. The t-test is used to observe the significance of α different from 0 in each dataset. The results in Table 6 show that the DMR strategy has similar α significance results as the existing universal strategies. They are insignificant on **FS100** and **DJIA01** and significant on **NYSE_Onew**. On the other hand, the α factor of the DMR strategy using the randomly generated initial feasible strategy is significant on **NYSE_Nnew**, and the α factor of the DMR(UBAH) strategy using the UBAH strategy as the initial strategy is insignificant except on **NYSE_Onew**. Note that the initial strategies of both the DMR(UBAH) strategy and the market strategy are the UBAH strategies. This indicates that UBAH, as the initial strategy, reduces the significance of the α factor of the DMR strategy, or we can say that the DMR(UBAH) follows the market strategy on these datasets. We believe that DMR has a higher probability for achieving excess returns and, in practice, the performance of the DMR strategy can be improved by selecting an appropriate initial strategy.

5.4. *Parameter sensitivity*

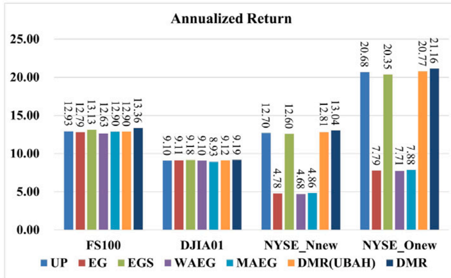
In this part, we analyze the sensitivity of DMR to realistic factors and strategy parameters. The section contains the transaction cost, short selling restriction, number of trading machines, and step size of a strategy.

5.4.1. *Transaction cost*

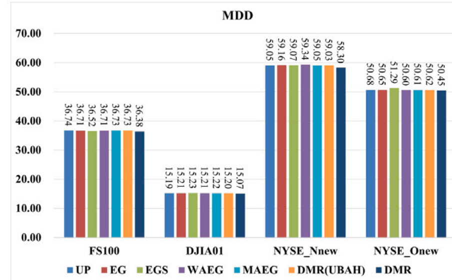
Fig. 7 shows that the cumulative wealth of all universal strategies does not decrease much as the transaction costs increase, and the rate of decline is relatively flat. This means that the universal strategies are not sensitive to transaction costs. Considering that short selling is allowed for DMR, it means that DMR has a better practical application than the existing universal strategies.

5.4.2. *Short selling allowed or not allowed*

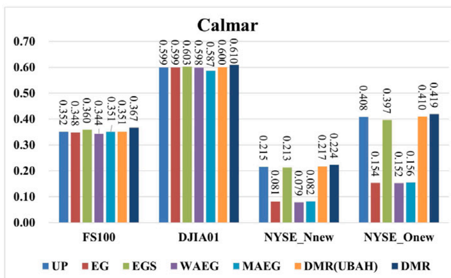
Now, we discuss the influence of short selling allowed (SSA) or short selling not allowed (SSNA) for DMR. The results in Table 7 show that removing the restriction on short selling can improve the cumulative wealth of DMR, but this improvement is not too much in terms of numerical value. This means that if we let all trading machines start trading from the UBAH strategy, SSA or SSNA does not shift the DMR strategy much in the feasible set. DMR can maintain a



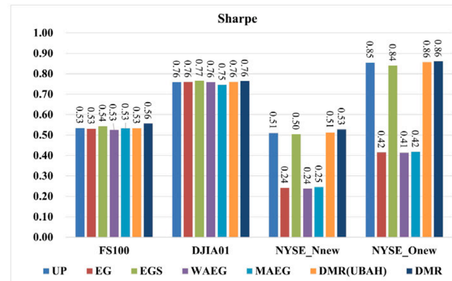
(a) Absolute return: AR



(b) Risk: MDD



(c) Risk-adjusted return: Calmar



(d) Risk-adjusted return: Sharpe

Fig. 6. Risk and risk-adjusted return performance of all strategies.

Table 6
Statistical t-test of the performance of DMR on the stock datasets.

	FS100				DJIA01			
	α	β	t-statistic	p-value	α	β	t-statistic	p-value
UP	0.0055	0.9995	1.0096	0.3132	−0.0172	0.9995	−1.0788	0.2810
EG	−0.0829	0.9960	0.8194	0.4130	−0.0060	0.9997	−1.0467	0.2956
EGS	0.2233	0.9984	1.8169	0.0698	0.0610	0.9996	−0.8919	0.3727
WAEG	−0.1913	0.9918	0.5449	0.5860	−0.0179	0.9995	−1.1083	0.2681
MAEG	−0.0175	0.9992	0.9615	0.3368	−0.1763	0.9990	−1.2377	0.2162
DMR (UBAH)	−0.0156	0.9993	0.9623	0.3364	−0.0008	1.0000	−1.0025	0.3164
DMR	0.5409	0.9916	2.0914	0.0370	0.0475	1.0025	−1.1983	0.2312
	NYSE_Nnew				NYSE_Onew			
	α	β	t-statistic	p-value	α	β	t-statistic	p-value
UP	−0.0568	0.9979	−16.3380	0.0000	−0.0723	0.9987	−22.7284	0.0000
EG	−3.2132	0.6249	−69.9039	0.0000	−4.6954	0.6006	−73.7063	0.0000
EGS	−0.1563	0.9975	−37.2205	0.0000	−0.4088	0.9991	−95.3162	0.0000
WAEG	−3.2865	0.6232	−77.8829	0.0000	−4.6977	0.5971	−76.6110	0.0000
MAEG	−3.1504	0.6265	−62.6113	0.0000	−4.6676	0.6036	−69.2592	0.0000
DMR (UBAH)	0.0292	0.9997	0.2825	0.7776	−0.0050	0.9998	−2.7861	0.0054
DMR	0.4063	0.9880	28.3610	0.0000	0.1366	1.0118	18.1043	0.0000

Table 7
Performance comparison of DMA (UBAH) in SSA and SSNA.

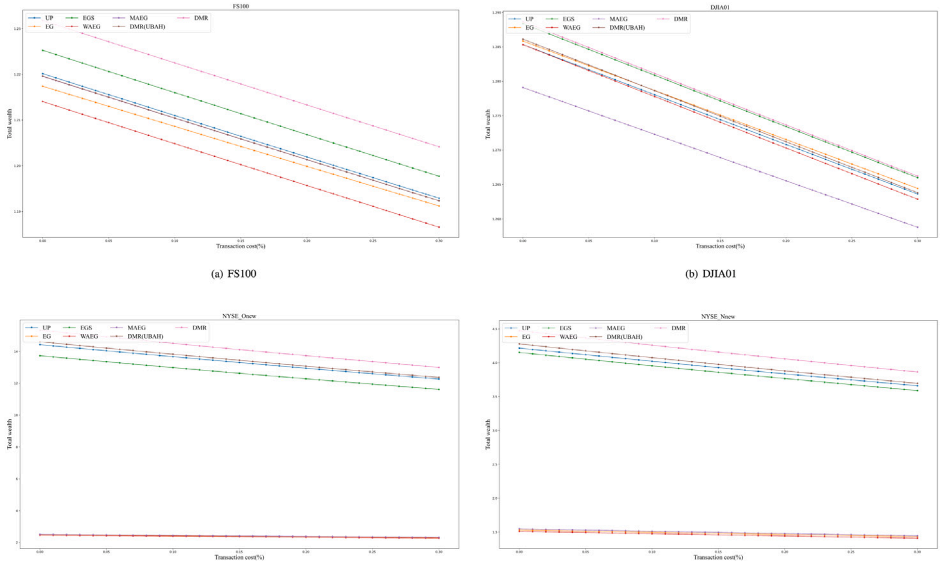
Criteria	FS100		DJIA01		NYSE_Nnew		NYSE_Onew	
	SSA	SSNA	SSA	SSNA	SSA	SSNA	SSA	SSNA
CW	1.22	1.22	1.286	1.286	4.44	4.28	14.6	14.6
Annualized volatility	24.2	24.0	11.99	11.99	25.00	25.02	24.2	24.2
Annualized turnover	3.29	3.29	1.644	1.644	3.16	3.21	3.67	3.67

similar cumulative wealth under both SSA and SSNA. Thus, we turn our analysis to the risk indicators, namely the annualized turnover rate and annualized volatility. The results in Table 7 illustrate that the turnover rate and volatility of DMR in SSA are higher than those in SSNA, which means that the DMR strategy adjusts assets more frequently in SSA and generates more risk.

5.4.3. Machine number n and step size α_k

Now, we analyze the impact of machine number n and step size α_k on the cumulative wealth of DMR. We set $\alpha_k \in \{\frac{1}{k+1}, \frac{1}{k+10}, \frac{1}{k+100}, \frac{1}{k+500}, \frac{1}{k+1000}, \frac{1}{k+10000}, \frac{1}{k+20000}\}$, $n \in \{1, 5, 7, 10, 15, 20, 23\}$, and obtain the cumulative wealth of DMR under various possible combinations of (α_k, n).

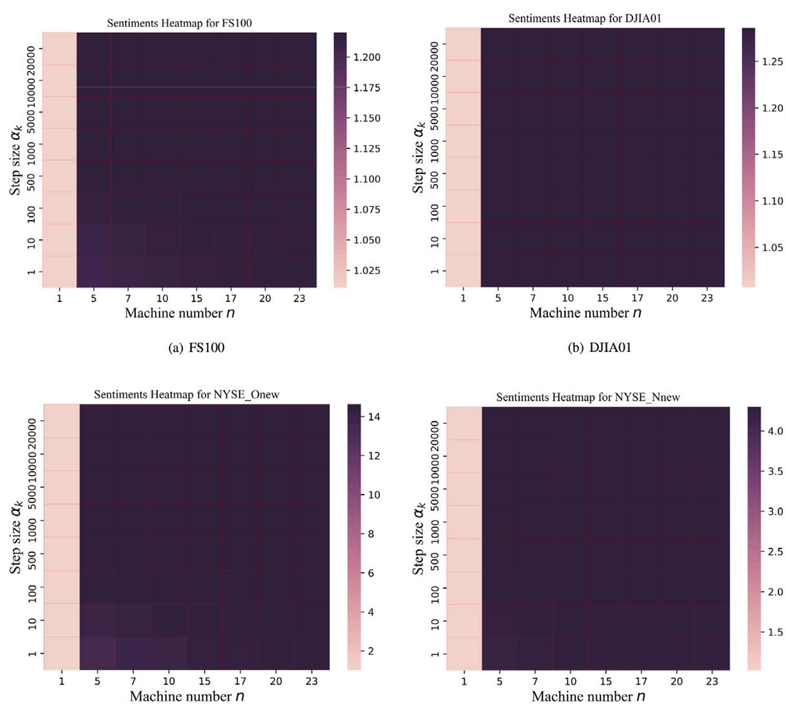
In Section 4, we set the stock correlation sub-network as the communication network structure of trading machines. This is because we believe that stocks with high centrality generate higher returns, and we hope that the structure of the sub-network filtered with centrality can help the trading machines calculate the strategies that can gain higher returns. The greater the centrality of a stock, the more stocks are associated with it. We can consider the sub-network as a condensation of the correlations in the complete stock network. It is natural to wonder whether the larger the sub-network size, the better the DMR



(a) FS100 (b) DJIA01

(c) NYSE_Onew (d) NYSE_Nnew

Fig. 7. Comparison of cumulative wealth under different transaction costs γ.



(a) FS100 (b) DJIA01

(c) NYSE_Onew (d) NYSE_Nnew

Fig. 8. Heat maps of the cumulative wealth of DMR (UBAH) in different step size α_k and machine number n.

strategy. Moreover, since different combinations of assets may yield different results, what is the impact of the size of the sub-network on the DMR strategy when faced with different asset portfolios? In the DMR strategy, the sub-network size and the number of trading machines are equal. Therefore, by varying the number of trading machines, we can analyze our above question under different sizes of stock sub-networks. The results in Fig. 8 show that:

(1) The effect of α_k is relatively intuitive. When $n = 1$, different values of α_k produce almost equal returns for DMR. This result is expected because α_k most directly affects the convergence rate of DMR and the optimal strategy for each trading machine. When $n = 1$, the DMR strategy is a centralized strategy that does not distinguish between trading machines and does not utilize any stock network information. At this point, α_k is the only factor to affect the speed of convergence of DMR to the BCRP. However, the trading periods in our datasets are generally more than two years or even more than 10 years, which could be a long-term investment period. Therefore, when the step size α_k changes, the convergence rate of DMR barely changes and the final cumulative wealth is the same.

When $n > 1$, the returns of DMR gradually increase as α_k decreases. Note that the DMR strategy has been transformed into a distributed strategy at this point. Under the influence of the stock sub-network structure, a distinction gradually arises between the optimal strategies of trading machines. The increase of step size α_k further enhances the

Table 8
Performance of DMR (UBAH) in different machine number n with fixed $\alpha_k = \frac{1}{k+1000}$.

Machine number n		5	7	10	15	17	20	23
FS100	CW	1.219098369	1.219388283	1.2196057	1.2197749	1.2198147	1.219859	1.219893
	Annualized volatility	24.16578744	24.17590732	24.183507	24.189423	24.190816	24.19238	24.19354
	MDD	36.72741052	36.72925353	36.730636	36.731711	36.731964	36.73225	36.73246
DJIA01	CW	1.286096455	1.286115916	1.2861305	1.2861419	1.2861445	1.286148	1.28615
	Annualized volatility	11.98601542	11.98605457	11.986084	11.986107	11.986112	11.98612	11.98612
	MDD	15.20181093	15.20153952	15.201336	15.201178	15.20114	15.2011	15.20107

sparsity between these strategies of the trading machines. However, Theorem 4.2 shows that the strategies of each trading machine will eventually converge to the same BCRP strategy. Therefore, the smaller the α_k, the faster the optimal strategies of all trading machines converge, and the larger the cumulative return of DMR. It is important to state that we do not recommend setting α_k too small because the optimization theory suggests that if the step size α_k is too small, the optimal strategies of some trading machines may still be "stuck" in the neighborhood of the previous strategy after a certain trading period. If a trading machine's strategy is "not good enough" for a given trading period, this may have a negative impact on the DMR strategy.

(2) The effect of the number of trading machines n on the DMR strategy is more complex. The cumulative return of DMR is minimized when $n = 1$. The stock sub-network corresponding to each trading period contains only one asset with the highest centrality, and the DMR strategy has actually become a centralized algorithm ($n = 1$) without using any information from the stock sub-network. Once the DMR strategy transforms into the distributed algorithm ($n > 1$), the returns of DMR increase significantly. When $n > 1$, the return of DMR increases as the number of trading machines (stock sub-network size) increases. This indicates that the DMR strategy exploits the information in the sub-network of stocks and obtains higher returns than the centralized algorithm. It is feasible to use centrality to filter stock sub-networks. When $n > 1$, the results in Fig. 8 reveal two points worth discussing:

(1) For the most part, the results in Fig. 8 seem to suggest that the bigger the n, the larger the returns. This is counter intuitive as different datasets contain different assets and have different effects on the DMR strategy. Also, n obviously cannot exceed the number of assets m contained in the dataset; for example, **FS100**, **DJIA01**, **NYSE_Nnew**, and **NYSE_Onew** contain 101, 29, 37, and 36 assets, respectively. Therefore, we set $\alpha_k = \frac{1}{k+1000}$ and conduct another experiment to further examine the effect of n. First, we set $n = \{101, 29, 37, 36\}$ for the above four datasets. We find that the strategy returns are 1.22, 1.286154, 4.259, and 14.6198, respectively. Compared to the results in Fig. 8, the return of DMR increases on **FS100** and **DJIA01** and decreases on **NYSE_Nnew** and **NYSE_Onew**. This suggests that more trading machines (stock sub-network size) is not necessarily better. For different asset portfolios, we should not simply set n equal to the number of assets m. If n increases, the sub-network will "absorb" more low-centrality stocks, which may lead to lower returns. Second, we observe the relationship between other performance indicators of the DMR strategy and the variation of n in dataset **DJIA01** and **FS100**. The results in Table 8 show that the annualized volatility, MDD, and cumulative wealth of the DMR strategy gradually increase as n increases. This indicates that increasing the number of trading machines n also makes the stock sub-network absorb more low-centrality stocks, which exposes the strategy to more market volatility risk.

(2) In the **NYSE_Nnew** dataset, when $\alpha_k \leq \frac{1}{k+500}$, the return of DMR decreases slowly as n increases, and the standard deviation of the change in returns does not exceed 0.01. In the **NYSE_Onew** dataset, when $\alpha_k \leq \frac{1}{k+5000}$, the return of DMR increases first, then decreases and increases again, but the standard deviation of the change in returns does not exceed 0.025. This means that DMR in these two datasets is more sensitive to the decrease in step size α_k. The smaller step size α_k makes the trading machine iterate the investment strategy very slowly in each period, "hindering" the convergence of the DMR strategy. We consider this effect to be acceptable because the fluctuations of the returns are small (in Fig. 8 the change is not obvious). In practice, if we expect to set a small α_k, we can provide a small n to obtain a "relatively good" return performance.

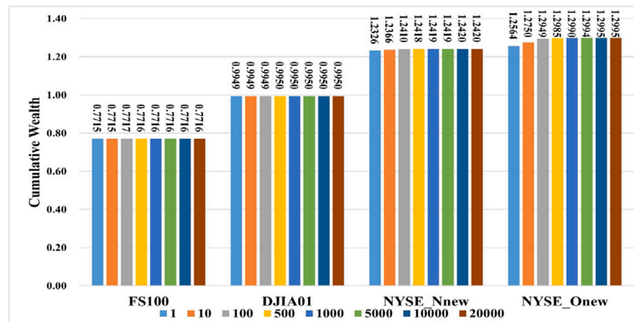


Fig. 9. Histograms of the cumulative wealth of the strategy under different α_k, where total trading periods is 100.

To further verify the effect of α_k in a short-term trading period, we limit the trading periods in each dataset to 100 days, and again obtain the change of the cumulative wealth of DMR for different values of α_k. We set $n = 10$ and $\alpha_k \in \{\frac{1}{k+1}, \frac{1}{k+10}, \frac{1}{k+100}, \frac{1}{k+500}, \frac{1}{k+1000}, \frac{1}{k+10000}, \frac{1}{k+20000}\}$. The results are summarized in 9.

The results in 9 show that when $\alpha_k < \frac{1}{k+100}$, the change in returns is monotonic across the datasets as α_k decreases. When $\alpha_k \geq \frac{1}{k+100}$, the returns appear to rise and then fall on the **FS100** dataset, and fall and then rise on the **DJIA01** dataset. This indicates that a large step size α_k tends to affect the strategy adjustment of the trading machine in the pre-trading period which, in turn, affects the returns of DMR in the pre-trading period. However, as trading progresses, the portfolio strategies of the trading machines gradually converge so that the change of returns of the DMR strategy becomes stable. Therefore, setting α_k in a small range can avoid unexpected fluctuations of the DMR strategy in the early trading periods, e.g., a large α_k results in trading machines being more likely to leapover the best strategy in the feasible set when updating its own strategy, making the DMR strategy take a longer trading period to eliminate this effect, thereby affecting the strategy return.

6. Conclusions

In this paper, we proposed a DMR strategy based on a stock correlation sub-network and a distributed algorithm. DMR is weighted by a group of portfolio strategies generated by a series of online trading machines connected by the stock correlation sub-network. Each trading machine communicates with the other machines according to the dynamic stock correlation sub-network structure to realize the exchange of strategies and update its own strategies for the next trading period. This strategy is suitable for intelligent investment situations such as large-scale investment and automated trading. Furthermore, short selling is allowed in our strategy, thereby satisfying more realistic investment needs.

The theoretical analysis shows that DMR and the strategies of all online trading machines are universal, which can provide a theoretical income guarantee for investors. The empirical test results show that, on one hand, the performance of DMR is better than those of existing universal strategies. DMR is also less sensitive to transaction cost. On the other hand, DMR is less sensitive to the number of trading machines and the step size numerically; however, determining an appropriate number of trading machines and step size at the beginning of investment can improve the performance of DMR. From the experimental results of this paper, "avoiding setting the number of trading machines and step size to extremely large or small" is a reasonable investment recommendation. In addition, "how to analyze the optimal size of the stock correlation sub-network" and "how to find the optimal step size" will be the focus of our future research.

CRediT authorship contribution statement

Yannan Zhong: Conceptualization, Methodology, Software, Writing – original draft, Visualization. **Weijun Xu:** Formal analysis, Resources, Writing – review & editing, Funding acquisition. **Hongyi Li:** Formal analysis, Resources, Writing – review & editing, Funding acquisition. **Weiwei Zhong:** Formal analysis, Resources, Writing – review & editing, Funding acquisition.

Acknowledgments

This work was, in part, supported by the National Natural Science Foundation of China (72271095, 71771091), the Guangdong Basic and Applied Basic Research Foundation, China (2022A1515010224), the Fundamental Research Funds for the Central Universities, China (PTJS202204), and the National Project Cultivation Plan of School of Business Administration of South China University of Technology, China.

We would like to thank Professor Bin Li and Professor László Györfi for their helpful comments and suggestions. We also want to thank the associate editor and anonymous reviewers for their helpful comments and suggestions.

Appendix A. Supplementary data

Supplementary material related to this article can be found online at https://doi.org/10.1016/j.ejor.2023.11.021.

References

Agarwal, A., Hazan, E., Kale, S., & Schapire, R. E. (2006). Algorithms for portfolio management based on the newton method. In *Proceedings of the 23rd international conference on machine learning* (pp. 9–16). New York, NY: ACM, http://dx.doi.org/10.1145/1143844.1143846.

Akcoglu, K., Drineas, P., & Kao, M. (2005). Fast universalization of investment strategies. *SIAM Journal on Computing*, *34*(1), 1–22. http://dx.doi.org/10.1137/S0097539702405619.

Akcoglu, K., Drineas, P., & Kao, M. (2020). Fast universalization of investment strategies with provably good relative returns. In *Proceedings of the 29th international colloquium on automata, languages, and programming* (pp. 888–900). Berlin, German: Springer, http://dx.doi.org/10.1007/3-540-45465-9_76.

Battiston, S., Farmer, J. D., Flache, A., Garlaschelli, D., Haldane, A. G., Heesterbeek, H., Hommes, C., Jaeger, C., May, R., & Scheffer, M. (2016). Complexity theory and financial regulation. *Science*, *351*(6725), 818–819. http://dx.doi.org/10.1126/science.aad0299.

Billo, M., Getmansky, M., Lo, A. W., & Pelizzon, L. (2012). Econometric measures of connectedness and systemic risk in the finance and insurance sectors. *Journal of Financial Economics*, *104*(3), 535–559. http://dx.doi.org/10.1016/j.jfineco.2011.12.010.

Blum, A., & Kalai, A. (1999). Universal portfolios with and without transaction costs. *Machine Learning*, *35*(3), 193–205. http://dx.doi.org/10.1023/A:1007530728748.

Borodin, A., Ran, E. Y., & Gogan, V. (2004). Can we learn to beat the best stock. *Journal of Artificial Intelligence Research*, *21*(1), 579–594. http://dx.doi.org/10.1613/jair.1336.

Cover, T. (1991). Universal portfolio. *Mathematical Finance*, *1*(1), 1–29. http://dx.doi.org/10.1111/j.1467-9965.1991.tb00002.x.

Cover, T., & Ordentlich, E. (2006). Universal portfolios with side information. *IEEE Transactions on Information Theory*, *42*(2), 348–363. http://dx.doi.org/10.1109/18.485708.

Das, P., & Banerjee, A. (2011). Meta optimization and its application to portfolio selection. In *Proceedings of international conference on knowledge discovery and data mining* (pp. 1163–1171). New York, NY: ACM, http://dx.doi.org/10.1145/2020408.2020588.

Edirisinghe, C., Jeong, J., & N., Chen J. (2021). Optimal portfolio deleveraging under market impact and margin restrictions. *European Journal of Operational Research*, *294*(2), 746–759. http://dx.doi.org/10.1016/j.ejor.2021.02.016.

Eduard, B., & Jochen, P. (2017). Interconnectedness risk and active portfolio management: The information-theoretic perspective. *SSRN Electronic Journal*, *3*(4), 25–47. http://dx.doi.org/10.2139/ssrn.2909839.

Gao, L., & Zhang, W. (2013). Weighted moving average passive aggressive algorithm for online portfolio selection. In *Proceedings of the 2013 5th international conference on intelligent human–machine systems and cybernetics* (pp. 327–330). New York, NY, USA: IEEE Computer Society, http://dx.doi.org/10.1109/IHMSC.2013.84.

Guo, Sini, Gu, J. W., & W.K., Ching (2021). Adaptive online portfolio selection with transaction costs. *European Journal of Operational Research*, *295*(3), 1074–1086. http://dx.doi.org/10.1016/j.ejor.2021.03.023.

Ha, Y., & Zhang, H. (2020). Algorithmic trading for online portfolio selection under limited market liquidity. *European Journal of Operational Research*, *286*(3), 1033–1051. http://dx.doi.org/10.1016/j.ejor.2020.03.050.

Hazan, E., & Seshadhri, C. (2009). Efficient learning algorithms for changing environments. In *Proceedings of the 26th annual international conference on machine learning* (pp. 393–400). New York, NY: ACM, http://dx.doi.org/10.1145/1553374.1553425.

He, J., & Yang, X. (2020). Universal portfolio selection strategy by aggregating online expert advice. *Optimization and Engineering*, *23*(1), 85–109. http://dx.doi.org/10.1007/s11081-020-09555-2.

Heiberger, R. H. (2014). Stock network stability in times of crisis. *Physica A. Statistical Mechanics and its Applications*, *393*(1), 376–381. http://dx.doi.org/10.1016/j.physa.2013.08.053.

Helmbold, D. P., Schapire, R. E., Singer, Y., & Warmuth, M. K. (1998). On-line portfolio selection using multiplicative updates. *Mathematical Finance*, *8*(4), 325–347. http://dx.doi.org/10.1111/1467-9965.00058.

Huang, D., Zhou, J., Li, B., Steven, C. H. H., & Zhou, S. (2016). Robust median reversion strategy for online portfolio selection. *IEEE Transactions on Knowledge and Data Engineering*, *28*(9), http://dx.doi.org/10.1109/TKDE.2016.2563433.

Jegadeesh, N. (1990). Evidence of predictable behavior of security returns. *The Journal of Finance*, *45*(3), 881–898. http://dx.doi.org/10.1111/j.1540-6261.1990.tb05110.x.

Kalai, A., & Vempala, S. (2002). Efficient algorithms for universal portfolios. *Journal of Machine Learning Research*, *3*(1), 423–440. http://dx.doi.org/10.5555/944919.944942.

Kelly, J. L. (1956). A new interpretation of information rate. *IEEE Transactions on Information Theory*, *2*(3), 185–189. http://dx.doi.org/10.1109/TIT.1956.1056803.

Lai, Z., Yang, P., Fang, L., & Wu, X. (2020). Reweighted price relative tracking system for automatic portfolio optimization. *IEEE Transactions on Systems, Man, and Cybernetics: Systems*, *50*(11), 4349–4361. http://dx.doi.org/10.1109/TSMC.2018.2852651.

Li, Y., Jiang, X., Tian, Y., Li, S., & Zheng, B. (2019). Portfolio optimization based on network topology. *Physica A. Statistical Mechanics and its Applications*, *515*(3), 671–681. http://dx.doi.org/10.1016/j.physa.2018.10.014.

Li, B., & Steven, C. H. H. (2014). Online portfolio selection: A survey. *ACM Computing Surveys*, *46*(3), 1–36. http://dx.doi.org/10.1145/2512962.

Li, B., Steven, C. H., Sahoo, D., & Liu, Z. (2015). Moving average reversion strategy for on-line portfolio selection. *Artificial Intelligence*, *222*(1), 104–123. http://dx.doi.org/10.1016/j.artint.2015.01.006.

Li, B., Wang, J., Huang, D., & Steven, C. H. H. (2018). Transaction cost optimization for online portfolio selection. *Quantitative Finance*, *18*(8), 1411–1424. http://dx.doi.org/10.1080/14697688.2017.1357831.

Li, B., Zhao, P., Steven, C. H. H., & Gopalkrishnan, V. (2012). Pamr: passive aggressive mean reversion strategy for portfolio selection. *Machine Learning*, *87*(2), 221–258. http://dx.doi.org/10.1007/s10994-012-5281-z.

Mantegna, R. N. (1999). Hierarchical structure in financial markets. *The European Physical Journal B - Condensed Matter and Complex Systems*, *11*(1), 193–197. http://dx.doi.org/10.1016/S0010-4655(99)00302-1.

Marc, C. S. (2001). Markowitz revisited: Mean–variance models in financial portfolio analysis. *SIAM Review*, *43*, 31–85. http://dx.doi.org/10.1137/S0036144500376650.

Markowitz, H. M. (1952). Portfolio selection. *The Journal of Finance*, *7*(1), 77–91. http://dx.doi.org/10.1111/j.1540-6261.1952.tb01525.x.

Matthias, Schnaubelt (2022). Deep reinforcement learning for the optimal placement of cryptocurrency limit orders. *European Journal of Operational Research*, *296*(3), 993–1006. http://dx.doi.org/10.1016/j.ejor.2021.04.050.

Nanda, S. J., & Panda, G. (2014). A survey on nature inspired metaheuristic algorithms for partitional clustering. *Swarm and Evolutionary Computation*, *16*(1), 1–18. http://dx.doi.org/10.1016/j.swevo.2013.11.003.

Pai, G. A. V., & Michel, T. (2009). Evolutionary optimization of constrained k-means clustered assets for diversification in small portfolios. *IEEE Transactions on Evolutionary Computation*, *13*(5), 1030–1053. http://dx.doi.org/10.1109/TEVC.2009.2014360.

Peralta, G., & Zareei, A. (2016). A network approach to portfolio selection. *Journal of Empirical Finance*, *38*(1), 157–180. http://dx.doi.org/10.1016/j.jempfin.2016.06.003.

Poterba, J. M., & Summers, L. H. (1988). Mean reversion in stock prices: Evidence and implications. *Journal of Financial Economics*, *22*(1), 27–59. http://dx.doi.org/10.1016/0304-405X(88)90021-9.

Ram, S. S., Nedic, A., & Veeravalli, V. V. (2010). Distributed stochastic subgradient projection algorithms for convex optimization. *Journal of Optimization Theory and Applications*, *147*(3), 516–545. http://dx.doi.org/10.1007/s10957-010-9737-7.

Tola, V., Lillo, F., Gallegati, M., & Mantegna, R. N. (2008). Cluster analysis for portfolio optimization. *Journal of Economic Dynamics & Control*, *32*(1), 235–258. http://dx.doi.org/10.1016/j.jedc.2007.01.034.

Tumminello, M., Aste, T., Matteo, T. D., & Mantegna, R. N. (2005). A tool for filtering information in complex systems. *Applyed Physical Sciences*, *102*(30), 10421–10426. http://dx.doi.org/10.1073/pnas.0500298102.

Vovk, V., & Watkins, C. (1998). Universal portfolio selection. In *Proceedings of the 11th annual conference on computational learning theory* (pp. 12–23). New York, NY: ACM, http://dx.doi.org/10.1145/279943.279947.

Vyrost, T., Lyocsa, S., & Eduard, B. (2019). Network-based asset allocation strategies. *The North American Journal of Economics and Finance*, *47*(1), 516–536. http://dx.doi.org/10.1016/j.najef.2018.06.008.

Wang, Z., Gao, X., An, H., Tang, R., & Sun, Q. (2020). Identifying influential energy stocks based on spillover network. *International Review of Financial Analysis*, *68*(11), 1–12. http://dx.doi.org/10.1016/j.irfa.2018.11.004.

Yang, X., He, J., Lin, H., & Zhang, Yong (2022). Boosting exponential gradient strategy for online portfolio selection: An aggregating experts' advice method. *Computational Economics*, *55*, 231–251. http://dx.doi.org/10.1007/s10614-019-09890-2.

Yang, X., He, J., Zhang, Y., & Zhang, W. (2019). Exponential gradient strategy for online portfolio with side information. *System Engineering Theory and Practice*, *39*(1), 60–69. http://dx.doi.org/10.12011/1000-6788-2017-1669-10.

Yang, T., Yi, X., Wu, J., Yuan, Y., Wu, D., Meng, Z., Hong, Y., Wang, H., Lin, Z., & Johansson, K. H. (2019). A survey of distributed optimization. *Annual Reviews in Control*, *47*(1), 278–305. http://dx.doi.org/10.1016/j.arcontrol.2019.05.006.

Zhang, Yong, Hong, Lin, Lina, Zheng, & Xingyu, Yang (2022). Adaptive online portfolio strategy based on exponential gradient updates. *Journal of Combinatorial Optimization*, *43*, 672–696. http://dx.doi.org/10.1007/s10878-021-00800-7.