Quantitative Finance

ISSN: (Print) (Online) Journal homepage: www.tandfonline.com/journals/rquf20

Adaptive online mean-variance portfolio selection with transaction costs

Sini Guo, Jia-Wen Gu, Wai-Ki Ching & Benmeng Lyu

Quantitative Finance, 2024
Vol. 24, No. 1, 59–82, https://doi.org/10.1080/14697688.2023.2287134

Adaptive online mean-variance portfolio selection with transaction costs

SINI GUO†, JIA-WEN GU‡, WAI-KI CHING§¶* and BENMENG LYU‡§

†School of Management and Economics, Beijing Institute of Technology, Beijing, People's Republic of China
‡Department of Mathematics, Southern University of Science and Technology, Shenzhen, People's Republic of China
§Advanced Modeling and Applied Computing Laboratory, Department of Mathematics, The University of Hong Kong, Pokfulam Road, Hong Kong
¶Hughes Hall, Wollaston Road, Cambridge, UK

(*Received 6 March 2023; accepted 18 November 2023; published online 19 December 2023*)

Online portfolio selection is attracting increasing attention in both artificial intelligence and finance communities due to its efficiency and practicability in deriving optimal investment strategies in real investment activities where the market information is constantly renewed every second. The key issues in online portfolio selection include predicting the future returns of risky assets accurately given historical data and providing optimal investment strategies for investors in a short time. In the existing online portfolio selection studies, the historical return data of one risky asset is used to estimate its future return. In this paper, we incorporate the peer impact into the return prediction where the predicted return of one risky asset not only depends on its past return data but also the other risky assets in the financial market, which gives a more accurate prediction. An adaptive moving average method with peer impact (AOLPI) is proposed, in which the decaying factors can be adjusted automatically in the investment process. In addition, the adaptive mean-variance (AMV) model is firstly applied in online portfolio selection where the variance is employed to measure the investment risk and the covariance matrix can be linearly updated in the investment process. The adaptive online moving average mean-variance (AOLPIMV) algorithm is designed to provide flexible investment strategies for investors with different risk preferences. Finally, numerical experiments are presented to validate the effectiveness and advantages of AOLPIMV.

Keywords: Online portfolio selection; Adaptive moving average method; Peer impact; Mean-variance model; Quadratic programming

1. Introduction

Financial portfolio selection has been studied for decades by academic researchers and industrial practitioners since Markovitz published the seminal work *portfolio selection* (Markowitz 1952). The mean-variance model was firstly proposed and intended to achieve a nice trade-off between maximizing the investment return and minimizing the investment risk. Since then, a large number of extensions were studied on the basis of the mean-variance model. For example, from the perspective of risk measure, semi-variance was studied by focusing on the volatility of returns less than the expected return (Markowitz 1959). The absolute deviation was proposed in Konno and Yamazaki (1991) to measure the investment risk of the portfolio and to overcome the limitation of the large computation burden of variance in real applications. In terms of investment horizon, multi-period portfolio selection was studied where the allocations of risky assets in different periods can be readjusted according to the change of market environment. For example, the multi-period model in Li and Ng (2000), the continuous time model in Zhou and Li (2000), the incomplete market model in Lim (2004), the nonnegative terminal state-constrained model in Bielecki *et al.* (2005), the semimartingale model in Xia (2005), the generalized mean-variance model in Gu *et al.* (2020), etc. González-Díaz *et al.* (2021) tackled the portfolio selection problem with transaction costs by using bilevel programming, where the broker-dealer controlled the charging fees of risky assets and the investor determined the capital allocation strategy. Kocuk and Cornuéjols (2020) modeled the returns of risky assets as a mixture of normal random variables and proposed the second-order cone representable approximation

*Corresponding author. E-mail: wching@hku.hk

to compute the Conditional Value-at-Risk of the whole portfolio. More studies on portfolio selection can be found in Balter *et al.* (2021), Katsikis *et al.* (2022), Li *et al.* (2022), Ling *et al.* (2020), Liesiö *et al.* (2020), Maghsoodi (2023), Staino and Russo (2020), Van Staden *et al.* (2021), and Zhou *et al.* (2021).

Most of the portfolio selection studies usually assume that the return of the risky asset is a random variable, of which its distribution function can be estimated by analyzing the historical return data. However, in practice, it is usually difficult to precisely estimate these distribution functions since the performance of risky asset is affected by multiple market factors such as constantly changing market environment and financial policies. In addition, large volume of financial data is generated every second in the 'Big data' era and the investors inevitably adjust the investment strategy timely when new financial data arrives, which makes it impractical to constantly estimate the distribution functions of risky assets in a practical decision-making process. With the rapid development of artificial intelligence and machine learning techniques, online portfolio selection is receiving increasing attention. Online portfolio selection is a typical sequential decision-making and optimization problem with the aim of maximizing the cumulative return over the whole investment horizon. At the beginning of each period, the investor determines the optimal strategy in a short time by utilizing all the available updated return data. Different from the existing traditional portfolio selection studies, online portfolio selection determines the optimal strategy of each period by analyzing the latest historical return data without specific statistical assumption of the market behavior, which exhibits strong practicability in real applications.

Several benchmark strategies are proposed in online portfolio selection studies. One popular benchmark is the buy-and-hold (BAH) strategy (Li and Hoi 2015). In this strategy, the investor determines the allocations of risky assets at the beginning of the investment and holds these risky assets until the end of the investment horizon, representing that the investment strategy is not rebalanced in all subsequent periods. We note that the allocation proportion of each risky asset alters in subsequent periods due to the price of asset changes frequently. In particular, when the initial capital is uniformly allocated to all the risky assets, the uniform buy-and-hold (UBAH) strategy is derived (Li and Hoi 2015), which is also called the Market strategy and can be used to reflect the trend of the financial market. Among all the BAH strategies, one can invest all the initial capital to the best risky asset in hindsight, which is called the Best Stock strategy. Another widely adopted benchmark is the constant rebalanced portfolios (CRP) strategy (Kelly 1956), which requires the investor to rebalance the portfolio constantly to ensure the allocation proportion of each risky asset is the same for all periods. When the capital is uniformly allocated in all periods, one can derive the uniform constant rebalanced portfolios (UCRP) strategy (Li and Hoi 2015). One can also select the best CRP (BCRP) strategy in hindsight (Cover 1991), which can only be used when full market information is available.

Except for the above benchmark strategies, various online portfolio selection algorithms are studied based on different market principles. The momentum principle assumes that risky assets achieving good performance in the past will continue performing well in the future. It is preferable to buy over-performing assets and sell under-performing assets. Cover (1991) proposed the universal portfolios (UP) algorithm, which laid the foundation of online portfolio selection theory. In the UP algorithm, the initial capital is allocated to various portfolio managers where each manger employs a certain rebalanced strategy. The cumulative return of UP is derived by calculating the weighted average of returns of all the rebalanced strategies. Helmbold *et al.* (1998) firstly proposed the exponential gradient (EG) strategy by tracking the best risky asset in the previous period. A regularization term was incorporated in the objective to keep the portfolio of this period not far away from the previous one. Agarwal *et al.* (2006) studied the online Newton step (ONS) method by computing the gradient and Hessian matrix of the log function of return in each period, which is verified to be efficient in practical implementation.

Another popular class of strategies is based on the mean-reversion principle which assumes that the current good-performing risky assets will perform poorly in the next period. It encourages buying the under-performing assets with the potential of making large profits in the future. This is called the 'follow the loser' strategy. Borodin *et al.* (2004) proposed the Anti-correlation (Anticor) strategy by increasing the proportions of risky assets suffering loss and reducing the proportions of assets making profits, where the proportion adjustment is conducted based on the cross-correlation matrix of returns of risky assets. This may seem countering common sense but is verified to be effective in empirical experiments. Li *et al.* (2012) embedded the passive aggressive learning technique into online portfolio selection and proposed the passive-aggressive mean-reversion (PAMR) strategy. The update scheme of PAMR is able to obtain a good trade-off between cumulative return and investment risk. Li and Hoi (2012, 2015) proposed simple and exponential moving average methods to predict future returns of risky assets by calculating the weighted average of the multiple periods of historical return data. The online moving average reversion (OLMAR) algorithm is then designed by minimizing the Euclidean distance of the target portfolio and previous portfolio with constraint that the expected return is more than a given threshold. Guo *et al.* (2021) improved the exponential moving average method by introducing an adaptive decaying factor which could be adjusted automatically according to the performance of risky assets. The adaptive online net profit maximization (AOLNPM) algorithm is proposed to tackle the online portfolio selection problem with transaction costs. Huang *et al.* (2016) exploited the mean-reversion phenomenon in online portfolio selection by introducing the robust L_1-median estimator, which is robust to noisy data and outliers. The robust median reversion (RMR) strategy is designed to overcome the poor performance of previous online portfolio selection algorithms when noisy financial data is involved.

Besides the above online portfolio selection studies following the momentum principle and the mean-reversion principle, the pattern matching strategies are also studied by combining the 'follow the winner' and 'follow the loser' strategies together. Usually, this type of strategy consists of

Table 1. Online portfolio selection algorithms.

Type	Algorithms	References
Benchmark algorithms	Buy-and-hold (BAH)	Li and Hoi (2015)
	Uniform buy-and-hold (UBAH)	Li and Hoi (2015)
	Constant rebalanced portfolios (CRP)	Kelly (1956)
	Uniform constant rebalanced portfolios (UCRP)	Li and Hoi (2015)
	Best constant rebalanced portfolios (BCRP)	Cover (1991)
Momentum principle based algorithms	Universal portfolios (UP)	Cover (1991)
	Exponential gradient (EG)	Helmbold *et al.* (1998)
	Online Newton step (ONS)	Agarwal *et al.* (2006)
	Variable rebalanced portfolios (VRP)	Gaivoronski and Stella (2000)
	Adaptive portfolio selection (APS)	Gaivoronski and Stella (2003)
Mean reversion principle based algorithms	Anticorrelation (Anticor)	Borodin *et al.* (2004)
	Passive aggressive mean-reversion (PAMR)	Li *et al.* (2012)
	Online moving average reversion (OLMAR)	Li and Hoi (2012, 2015)
	Robust median reversion (RMR)	Huang *et al.* (2016)
	Adaptive online net profit maximization (AOLNPM)	Guo *et al.* (2021)
	Confidence weighted mean-reversion (CWMR)	Li *et al.* (2013, 2011b)
Pattern matching based algorithms	Nonparametric histogram-based strategy	Györfi and Schäfer (2003)
	Nonparametric kernel-based strategy	Györfi *et al.* (2006)
	Nonparametric nearest neighbor-based strategy	Györfi *et al.* (2008)
	Correlation-driven nonparametric learning (CORN)	Li *et al.* (2011a)
	Local adaptive online portfolio (LOAD)	Guan and An (2019)
Meta-learning algorithms	Aggregating algorithm (AA)	Vovk and Watkins (1998)
	Online expert aggregation (OEA)	He and Yang (2022), He and Peng (2023), Yang *et al.* (2022)
	Follow the leading history (FLH)	Hazan and Seshadhri (2009)
	Online gradient update (OGU)	Das and Banerjee (2011)
	Online Newton update (ONU)	Das and Banerjee (2011)

two procedures: sample selection and portfolio optimization. The sample selection process is used to select a similar historical return data sequence of risky asset to predict the return of the next period. The portfolio optimization is then conducted by solving a certain utility function based on the predicted returns. Györfi and Schäfer (2003) employed the nonparametric histogram-based sample selection method by partitioning the latest market window sequence and historical market window sequence, and selecting the historical return whose preceding market window sequence is similar to the latest market window sequence as the future return estimation. Later, the nonparametric kernel-based sample selection (Györfi *et al.* 2006) and nonparametric nearest neighbor-based sample selection (Györfi *et al.* 2008) methods are employed to detect a similar market window sequence. Li *et al.* (2011a) detected the similar market window sequence by using correlation coefficient and proposed the correlation-driven nonparametric learning (CORN) strategy, of which the universal consistency was proved.

Another type of widely adopted online portfolio selection strategy is meta-learning (MA) algorithm. Similar to UP strategy, MA algorithm distributes the capital to various base experts. The difference lies in that the base expert in MA can be equipped with strategies from different classes. One portfolio is generated by each base expert and all the portfolios from these base experts are integrated into the final output. Vovk and Watkins (1998) solved the online portfolio selection problem by using aggregating algorithm (AA), which is an extension of the UP strategy. He and Yang (2022) proposed the online expert aggregation (OEA) strategy by aggregating the expert advice with the weak aggregating algorithm. Hazan and Seshadhri (2009) studied the follow the leading history (FLH) algorithm where the set of experts is dynamically adjusted according to the performances of these experts. Das and Banerjee (2011) proposed the online gradient update (OGU) and online Newton update (ONU) algorithms on the basis of EG and ONS, and achieved superior heuristic performance and the universality property. A summary of online portfolio selection algorithms is shown in table 1.

The present research focuses on the online portfolio selection problem, with a particular emphasis on the role of peer impact in enhancing asset return prediction accuracy. The concept of peer impact is based on the observation that risky assets in the financial market typically exhibit correlations with one another (Treynor and Black 1973). In this regard, price fluctuations in one asset can induce corresponding changes in other assets, thereby giving rise to contagion effects (Azizpour *et al.* 2018). Previous studies have investigated the multi-dimensional jump-diffusion model, wherein asset prices are assumed to follow mutually-exciting jump processes (Aït-Sahalia and Hurd 2016), as well as the mean-variance portfolio selection model in contagious financial markets, where price jumps are driven by a multivariate Hawkes process with mutual-excitation effect (Shen and Zou 2022). Veldkamp (2006) has also posited that asset price comovement can arise when investors acquire asset-payoff-relevant information from an information market, and pricing multiple assets based on common information may cause their prices to covary. Additionally, it is worth noting that the returns of risky assets are influenced by market/economy states, which in turn affect investment decisions and behaviors, and ultimately impact the performance of risky assets. For instance, Honda (2003) has proposed that the mean return of a risky asset is contingent upon an unobservable economic regime, and has employed a continuous-time Markov chain to model the regime-switching mechanism.

In AOLNPM algorithm (Guo *et al.* 2021), the adaptive online moving average (AOLMA) method is applied to predict the future return of a risky asset based on its past historical return data. Comparing with the exponential moving average method in Li and Hoi (2015), AOLMA employs an adaptive decaying factor which is adjusted gradually as the investment goes on and improves the accuracy of return prediction significantly. Although AOLMA is able to achieve higher accuracy than the existing methods, there are still some limitations constraining the prediction performance. The first one is that AOLMA predicts the future return of asset i by using the historical return data of risky asset i only without considering the impact of peer assets. In a real financial market, the performance of one asset may be affected by that of other assets and the financial market. For example, stocks in the same category may have similar price fluctuation, and usually the prices of stocks have a high probability of moving up in a Bull market and moving down in a Bear market. Secondly, in AOLMA, different assets share the same decaying factor in each period. However, the price fluctuation of different assets may differ a lot and it is better to employ different decaying factors for different assets. In addition, AOLNPM algorithm derives the optimal investment strategy by solving the net profit maximization (NPM) model, where no risk constraint is considered. In real applications, risk-neural and risk-averse investors treat controlling investment risk as one important objective. It is necessary and important to incorporate the investment risk into the objective of online portfolio selection. To overcome the above limitations, we propose the adaptive online moving average with peer impact (AOLPI) method where the impact of peer assets is considered in the return prediction. The adaptive mean-variance (AMV) model is constructed where the risk constraint is incorporated and can be applied by investors with different risk preferences. The contributions of this paper are summarized below:

- The peer impact of risky assets is incorporated into the AOLMA method and the decaying factors of risky assets can be adjusted differently.
- The AMV model is constructed where the adaptive covariance matrix is employed to measure the investment risk of the portfolio and can be linearly updated in each iteration. The AOLPIMV algorithm is proposed by integrating AOLPI and AMV together.
- The AMV model is transformed into a standard quadratic programming problem by using a change of variables. Theoretical proof is given to guarantee the existence of the optimal solution.
- Multiple numerical experiments are conducted to validate the effectiveness of AOLPIMV.

The remainder of this paper is organized as follows. Section 2 briefly introduces the background of the online portfolio selection problem. Section 3 introduces the AOLPI method and AMV model. Section 4 provides numerical experiments to validate the effectiveness of AOLPIMV and section 5 concludes the paper and discusses some future research issues.

2. Problem setting

In this section, we briefly introduce the decision-making process of online portfolio selection with transaction costs, which is a typical sequential optimization problem.

Assume that the investor plans to invest an initial capital W_0 into n assets within the investment horizon of T periods. At the beginning of each period, the investment strategy $\mathbf{x}_t = (x_{t1}, x_{t2}, \ldots, x_{tn})$ should be determined to maximize the net return of the whole portfolio, where x_{ti} refers to the proportion of capital allocated to asset i, $t = 1, 2, \ldots, T$, $i = 1, 2, \ldots, n$, and held until the end of this period. Different from the traditional multi-period portfolio selection problem where the return distributions in different periods are estimated before the investment starts, online portfolio selection does not require to derive the precise return distributions. The reason lies in the fact that the real return data of each period can be obtained immediately when this period ends and it is inefficient and inaccurate to estimate the distribution functions frequently. Suppose that the return vector of all the assets at period t is $\mathbf{r}_t = (r_{t1}, r_{t2}, \ldots, r_{tn})$, where r_{ti} is the return of asset i, then the final cumulative return at the end of period T can be expressed as follows:

$$W_T = W_0 \prod_{t=1}^{T} \left(\mathbf{r}_t \mathbf{x}_t^\top - \delta \parallel \mathbf{x}_t - \tilde{\mathbf{x}}_{t-1} \parallel_1 \right), \tag{1}$$

where $\mathbf{r}_t \mathbf{x}_t^\top$ is the total return of the whole portfolio at period t, and $\delta \parallel \mathbf{x}_t - \tilde{\mathbf{x}}_{t-1} \parallel_1$ is the transaction cost incurred by the change of asset investment proportions between period $(t-1)$ and period t. Here δ is the unit transaction cost and $\tilde{\mathbf{x}}_{t-1} = (\tilde{x}_{(t-1)1}, \tilde{x}_{(t-1)2}, \ldots, \tilde{x}_{(t-1)n})$ is the investment proportion of all the assets at the end of period $(t-1)$. We remark that $\tilde{\mathbf{x}}_{t-1}$ is not equal to $\mathbf{x}_{t-1}$ since the asset price at the beginning of period $(t-1)$ is different from the price at the end of period $(t-1)$, then the corresponding proportions of the total capital are changed. For each asset i, we have

$$\tilde{x}_{(t-1)i} = \frac{r_{(t-1)i} x_{(t-1)i}}{\mathbf{r}_{t-1} \mathbf{x}_{t-1}^\top}, \quad i = 1, 2, \ldots, n, \tag{2}$$

and

$$\parallel \mathbf{x}_t - \tilde{\mathbf{x}}_{t-1} \parallel_1 = \sum_{i=1}^{n} |x_{ti} - \tilde{x}_{(t-1)i}|. \tag{3}$$

The return of Asset i at period t is derived by $r_{ti} = p_{ti}/p_{(t-1)i}$ where p_{ti} is the opening/closing price at period t and p_{0i} is the initial price before the investment starts. In practical online portfolio selection decision-making process, r_t cannot be obtained until the end of period t. Usually the investor will first predict r_t by using the historical return data, then construct the return maximization model which is given as follows:

$$\begin{cases} \max & W_T \\ \text{s.t.} & \mathbf{e}_n \mathbf{x}_k^\top = 1, k = 1, 2, \ldots, T, \\ & \mathbf{0} \le \mathbf{x}_k \le \mathbf{e}_n, k = 1, 2, \ldots, T, \end{cases} \tag{4}$$

where $\mathbf{e}_n$ is a $1 \times n$ vector of all ones, i.e. $\mathbf{e}_n = (1, 1, \ldots, 1)$, the first constraint refers to that the sum of all the investment proportions in each period is equal to 1, and the second constraint gives the lower and upper bounds for each proportion. The above model was studied by Guo *et al.* (2021) whose objective is to maximize the net cumulative return over the whole investment horizon with transaction costs.

3. Online mean-variance model with transaction costs

This section proposes the method of solving online portfolio selection problem considering transaction costs by using mean-variance model, where the variance is adopted to measure the total investment risk. We first incorporate the peer impact into the online moving average method to improve the estimation accuracy of future returns of assets. Then the adaptive mean-variance model is adopted in online portfolio selection such that the return and risk can be taken into account simultaneously.

3.1. Adaptive online moving average with peer impact method

One key issue in online portfolio selection is the estimation of the future return $\mathbf{r}_t$ at the beginning of period t. An accurate estimation is useful to assist the investor to avert the poor performing assets and put large allocation of capital on assets with high potential of making profits in the future. Li and Hoi (2012, 2015) firstly proposed the moving average methods including simple moving average (SMA) and exponential moving average (EMA) methods to predict the future returns of assets by calculating the arithmetical average of certain recent historical returns. Inspired by Li and Hoi, Guo *et al.* (2021) proposed the adaptive online moving average (AOLMA) method by introducing the adaptive decaying factor which can be automatically adjusted according to the performances of assets. In the following, we introduce the adaptive online moving average with peer impact (AOLPI) method on the basis of AOLMA.

Suppose that the investor has the past historical price data $\mathbf{p}_1, \mathbf{p}_2, \ldots, \mathbf{p}_{t-1}$ at the beginning of period t, where $\mathbf{p}_k = (p_{k1}, p_{k2}, \ldots, p_{kn})$, $k = 1, 2, \ldots, t-1$. According to the SMA method with truncated window size w, the estimated price of asset i at period t is

$$\hat{p}_{ti} = \frac{1}{w} \sum_{k=t-w}^{t-1} p_{ki},$$

and the estimation of r_{ti} is calculated by

$$\begin{aligned}\hat{r}_{ti} &= \frac{\hat{p}_{ti}}{p_{(t-1)i}} \\ &= \frac{1}{w}\left(1 + \frac{1}{r_{(t-1)i}} + \frac{1}{r_{(t-1)i}\cdot r_{(t-2)i}} + \cdots + \frac{1}{\prod_{i=0}^{w-2} r_{(t-1-k)i}}\right)\end{aligned}$$

which is arithmetical average of truncated historical prices and returns. According to the EMA method, the estimated price and return at period t can be expressed as follows:

$$\begin{aligned}\hat{p}_{ti} &= (1-\theta)^{t-2}p_{1i} + \theta(1-\theta)^{t-3}p_{2i} + \cdots \\ &\quad + \theta(1-\theta)p_{(t-1)i} + \theta p_{(t-1)i} = \theta p_{(t-1)i} + (1-\theta)\hat{p}_{(t-1)i},\end{aligned}$$

and

$$\hat{r}_{ti} = \theta + (1-\theta)\cdot\frac{\hat{r}_{(t-1)i}}{r_{(t-1)i}},$$

where θ is the given decaying factor. In AOLMA, the estimations of p_{ti} and r_{ti} are given, respectively, by

$$\hat{p}_{ti} = \theta_{ti}p_{(t-1)i} + (1-\theta_{ti})\hat{p}_{(t-1)i}$$

and

$$\hat{r}_{ti} = \theta_{ti} + (1-\theta_{ti})\cdot\frac{\hat{r}_{(t-1)i}}{r_{(t-1)i}}.$$

Here, the decaying factor θ_{ti} is adjusted as the investment goes on.

It can be seen clearly from the above methods that the prediction of $\hat{r}_{ti}$ only depends on the past historical data sequence $\mathbf{D}_{ti} = [r_{1i}, r_{2i}, \ldots, r_{(t-1)i}]$. In this paper, we introduce the *peer impact* into the AOLMA method. For asset i, denote the average return data of all the other assets as $\bar{\mathbf{D}}_{ti}$, which can be expressed as

$$\begin{aligned}\bar{\mathbf{D}}_{ti} &= [R_{1i}, R_{2i}, \ldots, R_{(t-1)i}] \\ &= \left[\frac{\sum_{k\neq i} r_{1k}}{n-1}, \frac{\sum_{k\neq i} r_{2k}}{n-1}, \ldots, \frac{\sum_{k\neq i} r_{(t-1)k}}{n-1}\right]. \end{aligned} \tag{5}$$

Then $\bar{\mathbf{D}}_{ti}$ can be used to measure the performances of the other assets in the financial market. We note that $\mathbf{D}_{ti}$ and $\bar{\mathbf{D}}_{ti}$ can be updated at each period when new return data is available. The return r_{ti} can be estimated by considering the two data sequences simultaneously. Firstly, we apply the AOLMA method to the data sequence $\mathbf{D}_{ti}$ and derive

$$u_{ti} = \theta_{ti}^{(1)} + (1-\theta_{ti}^{(1)})\cdot\frac{\hat{r}_{(t-1)i}}{r_{(t-1)i}}. \tag{6}$$

Similarly, for data sequence $\bar{\mathbf{D}}_{ti}$, we derive the corresponding moving average return by

$$v_{ti} = \theta_{ti}^{(2)} + (1-\theta_{ti}^{(2)})\cdot\frac{\hat{R}_{(t-1)i}}{R_{(t-1)i}}, \tag{7}$$

where $\theta_{ti}^{(1)}$ and $\theta_{ti}^{(2)}$ are the decaying factors of asset i at period t for $\mathbf{D}_{ti}$ and $\bar{\mathbf{D}}_{ti}$, respectively. Then, we estimate the future return r_{ti} based on u_{ti} and v_{ti}, which can be expressed as follows:

$$\hat{r}_{ti} = \alpha_i u_{ti} + \beta_i v_{ti}, \tag{8}$$

where α_i and β_i represent the weighting factors of u_{ti} and v_{ti}, implying that the estimation of $\hat{r}_{ti}$ is both affected by the historical data $\mathbf{D}_{ti}$ and $\bar{\mathbf{D}}_{ti}$. Here α_i and β_i can be determined by using the least square method. At the beginning of period t, all the historical data $\mathbf{D}_{ti}$ and $\bar{\mathbf{D}}_{ti}$ are available and the values of u_{ti} and v_{ti} can be derived by equations (6) and (7), respectively. Then the sum of mean squared errors over the first $(t-1)$ periods is

$$MSE_i = \sum_{k=1}^{t-1} (\alpha_i u_{ki} + \beta_i v_{ki} - r_{ki})^2.$$

To minimize the sum of mean squared errors, we set the derivatives with respect to α_i and β_i and obtain

$$\frac{\mathrm{d}MSE_i}{\mathrm{d}\alpha_i} = 2\sum_{k=1}^{t-1} (\alpha_i u_{ki} + \beta_i v_{ki} - r_{ki}) u_{ki} = 0 \tag{9}$$

and

$$\frac{\mathrm{d}MSE_i}{\mathrm{d}\beta_i} = 2\sum_{k=1}^{t-1} (\alpha_i u_{ki} + \beta_i v_{ki} - r_{ki}) v_{ki} = 0. \tag{10}$$

By solving equations (9) and (10), we derive the estimations for α_i and β_i:

$$\hat{\alpha}_{ti} = \frac{(\sum_{k=1}^{t-1} r_{ki}u_{ki})(\sum_{k=1}^{t-1} v_{ki}^2) - (\sum_{k=1}^{t-1} u_{ki}v_{ki})(\sum_{k=1}^{t-1} r_{ki}v_{ki})}{(\sum_{k=1}^{t-1} u_{ki}^2)(\sum_{k=1}^{t-1} v_{ki}^2) - (\sum_{k=1}^{t-1} u_{ki}v_{ki})^2}$$

and

$$\hat{\beta}_{ti} = \frac{(\sum_{k=1}^{t-1} r_{ki}v_{ki})(\sum_{k=1}^{t-1} u_{ki}^2) - (\sum_{k=1}^{t-1} u_{ki}v_{ki})(\sum_{k=1}^{t-1} u_{ki}r_{ki})}{(\sum_{k=1}^{t-1} u_{ki}^2)(\sum_{k=1}^{t-1} v_{ki}^2) - (\sum_{k=1}^{t-1} u_{ki}v_{ki})^2},$$

respectively. Here the estimations can be updated at each period when new return data is available. To achieve a better estimation accuracy, we employ the adaptive decaying factors following from AOLMA. It follows from equations (6), (7) and (8) that

$$\hat{r}_{ti} = \hat{\alpha}_{ti}\left[\theta_{ti}^{(1)} + (1-\theta_{ti}^{(1)})\cdot\frac{\hat{r}_{(t-1)i}}{r_{(t-1)i}}\right] + \hat{\beta}_{ti}\left[\theta_{ti}^{(2)} + (1-\theta_{ti}^{(2)})\cdot\frac{\hat{R}_{(t-1)i}}{R_{(t-1)i}}\right],$$

and the corresponding estimation error is

$$r_{ti} - \hat{r}_{ti} = r_{ti} - \hat{\alpha}_{ti}\left(\frac{\theta_{ti}^{(1)} r_{(t-1)i} + (1-\theta_{ti}^{(1)})\hat{r}_{(t-1)i}}{r_{(t-1)i}}\right) - \hat{\beta}_{ti}\left(\frac{\theta_{ti}^{(2)} R_{(t-1)i} + (1-\theta_{ti}^{(2)})\hat{R}_{(t-1)i}}{R_{(t-1)i}}\right),$$

which implies that the estimation error is a linear function of decaying factors $\theta_{ti}^{(1)}$ and $\theta_{ti}^{(2)}$. Here we elaborate on the iteration process of $\theta_{ti}^{(1)}$. The derivative of $(r_{ti} - \hat{r}_{ti})$ with respect to $\theta_{ti}^{(1)}$ is

$$\frac{\mathrm{d}(r_{ti} - \hat{r}_{ti})}{\mathrm{d}\theta_{ti}^{(1)}} = -\hat{\alpha}_{ti}\frac{r_{(t-1)i} - \hat{r}_{(t-1)i}}{r_{(t-1)i}}.$$

We note that r_{ti} can be obtained at the end of period t, then $\theta_{ti}^{(1)}$ can be updated to potentially reduce the estimation error of the next period. The iteration mechanism is summarized in table 2.

Table 2. Decaying factor iteration table.

	$\hat{\alpha}_{ti}(r_{(t-1)i} - \hat{r}_{(t-1)i}) \geq 0$	$\hat{\alpha}_{ti}(r_{(t-1)i} - \hat{r}_{(t-1)i}) < 0$
$r_{ti} \geq \hat{r}_{ti}$	$\theta_{ti}^{(1)} = \theta_{ti}^{(1)} + \gamma$	$\theta_{ti}^{(1)} = \theta_{ti}^{(1)} - \gamma$
$r_{ti} < \hat{r}_{ti}$	$\theta_{ti}^{(1)} = \theta_{ti}^{(1)} - \gamma$	$\theta_{ti}^{(1)} = \theta_{ti}^{(1)} + \gamma$

In case that $r_{ti} \geq \hat{r}_{ti}$ and $\hat{\alpha}_{ti}(r_{(t-1)i} - \hat{r}_{(t-1)i}) \geq 0$, the decaying factor is updated by $\theta_{ti}^{(1)} = \theta_{ti}^{(1)} + \gamma$. This is due to the fact that when $\hat{\alpha}_{ti}(r_{ti} - \hat{r}_{ti})$ is larger than 0 and $r_{(t-1)i} \geq \hat{r}_{(t-1)i}$, the derivative $\frac{\mathrm{d}(r_{ti}-\hat{r}_{ti})}{\mathrm{d}\theta_{ti}^{(1)}}$ is less than 0. Then the estimation error can be decreased by increasing the value of $\theta_{ti}^{(1)}$. Similarly, in case that $r_{ti} \geq \hat{r}_{ti}$ and $\hat{\alpha}_{ti}(r_{(t-1)i} - \hat{r}_{(t-1)i}) < 0$, the estimation error can be decreased by updating the decaying factor with $\theta_{ti}^{(1)} = \theta_{ti}^{(1)} - \gamma$. We note that the initial value of the decaying factor before the investment starts is set to be 0.5. In the iteration process, when $\theta_{ti}^{(1)} > 1$ or $\theta_{ti}^{(1)} < 0$, we reset the decaying factor as 0.5. This decaying factor iteration mechanism is effective and has been validated by Guo *et al.* (2021). For the factor $\theta_{ti}^{(2)}$, we employ a similar iteration mechanism.

Remark 3.1 The AOLPI method proposed in this study utilizes linear regression to predict future returns of risky assets, motivated by two main factors. Firstly, linear regression models have been extensively applied in the fields of risky asset return prediction and portfolio selection, as evidenced by studies such as Wang *et al.*'s (2021) time-dependent weighted least squares approach for stock return prediction. Further supporting literature can be found in references such as Guo *et al.* (2023), Li and Hoi (2012), and Li and Hoi (2015). Secondly, compared to more complex nonlinear return prediction methods like neural networks and reinforcement learning, established linear regression methods are easier to implement when data is continually updated, allowing for a linear update of predicted returns based on previous period predictions and the most recently updated return data. This feature is particularly advantageous for online decision-making scenarios where investors must make decisions within tight time constraints.

Example 3.1 We apply the proposed AOLPI method to a real data set DJIA and compare the results with the previous AOLMA method. This data set is selected from Li and Hoi (2015) and contains 507 trading days' returns of 30 stocks ranging from 14 January 2001 to 14 January 2003, i.e. $T = 507$ and $n = 30$. We set the step size as $\gamma = 0.0004$ in both of the methods. For each stock i, the average relative estimation error can be calculated by

$$AER_i = \frac{1}{T}\sum_{k=1}^{T}\frac{|\hat{r}_{ki} - r_{ki}|}{r_{ki}} \times 100\%.$$

The average relative errors of these stocks by using AOLMA and AOLPI are shown in figure 1, and the corresponding error differences are shown in figure 2. For example, for Stock 27, the average relative errors of AOLMA and AOLPI are 2.14% and 1.71%, respectively. It is clear that AOLPI achieves a smaller average relative error for each stock, which validates its effectiveness. This is because AOLPI incorporates

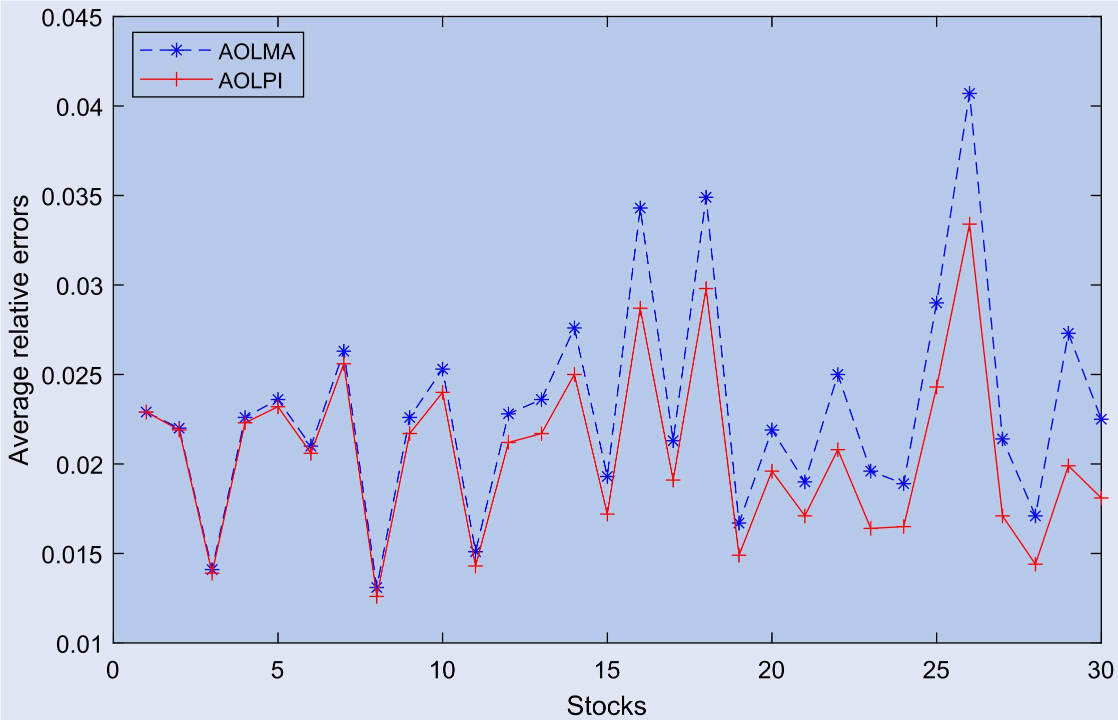

Figure 1. A comparison of average relative errors.

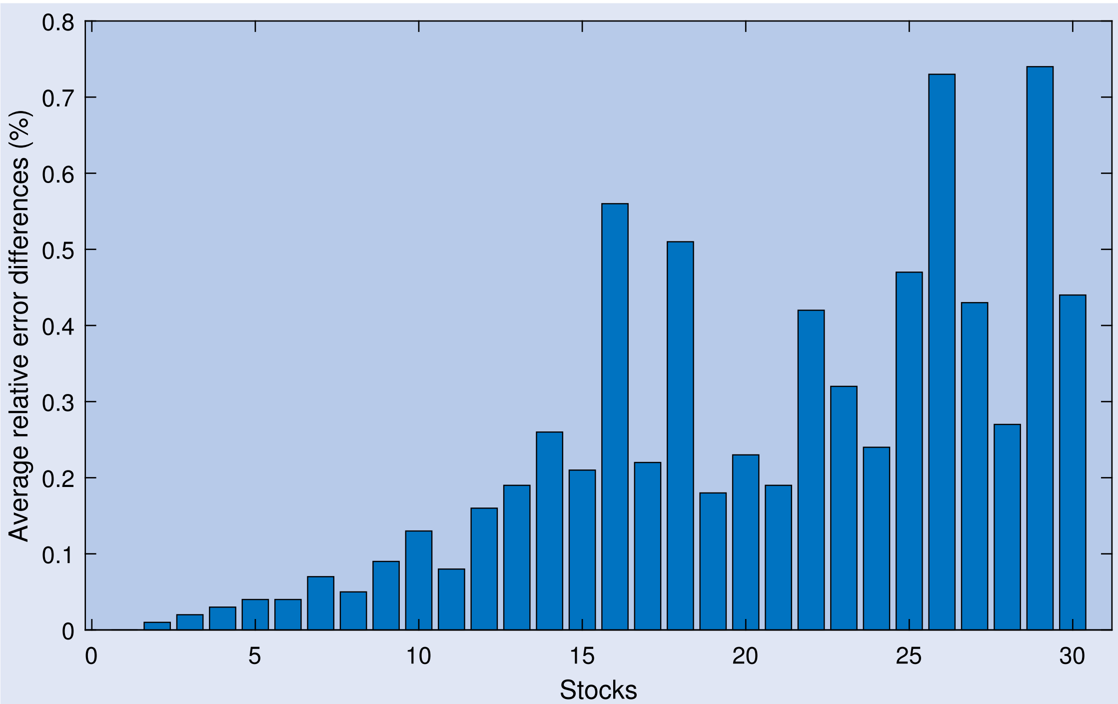

Figure 2. Average relative error difference.

more market information than AOLMA by considering the historical data $\mathbf{D}_{ti}$ and $\bar{\mathbf{D}}_{ti}$ simultaneously.

The idea of incorporating peer impact into the online moving average method is motivated by the fact that the price of a risky asset is not only related to its own historical prices but also the prices of the other assets in the market. For example, Treynor and Black (1973) pointed out that various empirical studies demonstrated that portfolios of several randomly selected securities tend to correlate very highly with the market portfolio. Secondly, the contagion effect exists in financial market as it is illustrated in section 1. The price jump of one asset may cause the price jumps of the other assets in the financial market (See Aït-Sahalia and Hurd 2016, Shen and Zou 2022). According to Leary and Roberts (2014), smaller firms tend to adjust their financial policies in response to actions taken by larger firms, which can have implications for changes in prices. Furthermore, the performance and returns of risky assets are also influenced by market/economy states, as highlighted by Honda (2003). Based on these factors, we propose incorporating peer impact into future return prediction for risky assets. The results of example 3.1 demonstrate that considering peer impact can lead to improvements in return prediction accuracy.

3.2. *Adaptive mean-variance model*

In this section, we propose the adaptive mean-variance (AMV) model to solve the practical online portfolio selection problem. In net profit maximum (NPM) model (Guo *et al.* 2021), the objective in each period is to maximize the total net return after excluding the transaction cost. Although the NPM model achieves good performance in the numerical experiments, its practicability is limited since it is actually intended for risk-appetite investors who aim to maximize the total investment return without caring about the investment risk. Usually investment risk is not considered in the NPM model, one reason is that it is inefficient and time-consuming to constantly update the risk measures as the return data is updated in each period and there are usually a large number of assets. Guo *et al.* (2021) employed the absolute deviation to measure the investment risk of each period in the comparison experiment, only the historical returns of the most recent w periods are used, which is much smaller than T and cannot fully reflect the return fluctuation of assets. Here the variance of risky asset is employed as a measure for the investment risk. We construct the adaptive online mean-variance model (AMV), where the covariance matrix can be linearly updated in each period when new return data arrives. The AMV model is given as follows:

$$\begin{cases} \max & \mathbf{r}_t\mathbf{x}_t^\top - \eta\mathbf{x}_t\boldsymbol{\Sigma}_t\mathbf{x}_t^\top - \delta \parallel \mathbf{x}_t - \tilde{\mathbf{x}}_{t-1} \parallel_1 \\ \text{s.t.} & \mathbf{e}_n\mathbf{x}_t^\top = 1, \\ & \mathbf{0} \le \mathbf{x}_t \le \mathbf{e}_n \end{cases} \tag{11}$$

where $\boldsymbol{\Sigma}_t$ is the covariance matrix of all the assets estimated at the beginning of period t, and η is the weighting factor ($\eta > 0$). We note that model (11) is constructed when the past returns $\mathbf{r}_1, \mathbf{r}_2, \ldots, \mathbf{r}_{t-1}$ are already known. The key issue in constructing model (11) lies in the updating process of $\boldsymbol{\Sigma}_t$, which can be expressed as follows:

$$\boldsymbol{\Sigma}_t = \begin{pmatrix} \sigma_{11}^{(t)} & \sigma_{12}^{(t)} & \sigma_{13}^{(t)} & \cdots & \sigma_{1n}^{(t)} \\ \sigma_{21}^{(t)} & \sigma_{22}^{(t)} & \sigma_{23}^{(t)} & \cdots & \sigma_{2n}^{(t)} \\ \sigma_{31}^{(t)} & \sigma_{32}^{(t)} & \sigma_{33}^{(t)} & \cdots & \sigma_{3n}^{(t)} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ \sigma_{n1}^{(t)} & \sigma_{n2}^{(t)} & \sigma_{n3}^{(t)} & \cdots & \sigma_{nn}^{(t)} \end{pmatrix},$$

where each $\sigma_{ij}^{(t)}$ is the covariance of assets i and j at the beginning of period t. For each asset i, we denote the average value of the past returns $r_{1i}, r_{2i}, \ldots, r_{(t-1)i}$ as μ_{ti}, where

$$\mu_{ti} = \frac{\sum_{k=1}^{t-1} r_{ki}}{t-1}.$$

Then the iteration formula relating μ_{ti} and $\mu_{(t+1)i}$ can be derived as follows:

$$\mu_{(t+1)i} = \frac{\sum_{k=1}^{t} r_{ki}}{t} = \frac{(t-1)\mu_{ti} + r_{ti}}{t} = \left(\frac{t-1}{t}\right)\mu_{ti} + \frac{1}{t}r_{ti}. \tag{12}$$

The variance of asset i can be calculated by

$$\sigma_{ii}^{(t)} = \frac{1}{t-2}\sum_{k=1}^{t-1}(r_{ki} - \mu_{ti})^2,$$

and the iteration formula relating $\sigma_{ii}^{(t)}$ and $\sigma_{ii}^{(t+1)}$ is given as follows:

$$\sigma_{ii}^{(t+1)} = \frac{t-2}{t-1}\sigma_{ii}^{(t)} + \frac{1}{t}(r_{ti} - \mu_{ti})^2. \tag{13}$$

For any two assets i and j, we have

$$\sigma_{ij}^{(t)} = \frac{1}{t-2}\sum_{k=1}^{t-1}(r_{ki} - \mu_{ti})(r_{kj} - \mu_{tj}). \tag{14}$$

Then the iteration formula relating $\sigma_{ij}^{(t)}$ and $\sigma_{ij}^{(t+1)}$ can be derived as follows:

$$\sigma_{ij}^{(t+1)} = \frac{t-2}{t-1}\sigma_{ij}^{(t)} + \frac{1}{t}(r_{ti} - \mu_{ti})(r_{tj} - \mu_{tj}). \tag{15}$$

The technical details for obtaining equations (13) and (15) can be seen in appendix 1. It can be seen clearly that equation (15) coincides with equation (13) when $i = j$. According to equations (12), (13) and (15), the covariance matrix can be linearly updated when new return data is obtained at the end of period t, which is given by

$$\boldsymbol{\Sigma}_{t+1} = \left(\frac{t-2}{t-1}\right)\boldsymbol{\Sigma}_t + \frac{1}{t}\boldsymbol{M}_t, \tag{16}$$

where $\boldsymbol{M}_t$ is an $n \times n$ matrix with its (ij)th entry being given by $(r_{ti} - \mu_{ti})(r_{tj} - \mu_{tj})$.

We then combine the AOLPI method with the AMV model together in solving the following optimization problem:

$$\begin{cases} \max & \hat{\mathbf{r}}_t\mathbf{x}_t^\top - \eta\mathbf{x}_t\boldsymbol{\Sigma}_t\mathbf{x}_t^\top - \delta \parallel \mathbf{x}_t - \tilde{\mathbf{x}}_{t-1} \parallel_1 \\ \text{s.t.} & \mathbf{e}_n\mathbf{x}_t^\top = 1, \\ & \mathbf{0} \le \mathbf{x}_t \le \mathbf{e}_n, \end{cases} \tag{17}$$

where the first two terms $\hat{\mathbf{r}}_t\mathbf{x}_t - \eta\mathbf{x}_t\boldsymbol{\Sigma}_t\mathbf{x}_t^\top$ in the objective function is a standard quadratic programming. The tricky one is the nonlinear transaction cost $\delta \parallel \mathbf{x}_t - \tilde{\mathbf{x}}_{t-1} \parallel_1$ which is $\sum_{i=1}^{n} |x_{ti} - \tilde{x}_{(t-1)i}|$ where

$$\tilde{x}_{(t-1)i} = \frac{r_{(t-1)i}x_{(t-1)i}}{\mathbf{r}_{t-1}\mathbf{x}_{t-1}^\top}$$

according to equations (2) and (3). For the convenience of computations, we employ the method of change of variables to transform the nonlinear transaction cost into a linear one.

For each term $|x_{ti} - \tilde{x}_{(t-1)i}|$, suppose that there are nonnegative variables y_{ti} and z_{ti} such that

$$\begin{pmatrix} |x_{ti} - \tilde{x}_{(t-1)i}| \\ x_{ti} - \tilde{x}_{(t-1)i} \end{pmatrix} = \begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix} \begin{pmatrix} y_{ti} \\ z_{ti} \end{pmatrix}, \quad i = 1, 2, \ldots, n. \tag{18}$$

We can always solve the values of y_{ti} and z_{ti} since the determinant of the matrix

$$\begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix}$$

is nonzero. It can be derived from equation (18) that $x_{ti} = \tilde{x}_{(t-1)i} + y_{ti} - z_{ti}$. Taking all x_{ti} into consideration, we have

$$\mathbf{x}_t = (\mathbf{y}_t, \mathbf{z}_t)N + \tilde{\mathbf{x}}_{t-1}, \tag{19}$$

where $N = (\mathbf{I}_n, -\mathbf{I}_n)^\top$ and $\mathbf{I}_n$ is an identity matrix of size $n \times n$. The transaction cost term can be transformed into

$$\delta \parallel \mathbf{x}_t - \tilde{\mathbf{x}}_{t-1} \parallel_1 = \delta \sum_{i=1}^{n} (y_{ti} + z_{ti}) = \delta \mathbf{e}_{2n} (\mathbf{y}_t, \mathbf{z}_t)^\top,$$

where $\mathbf{e}_{2n}$ is the $1 \times 2n$ row vector of all ones. For the objective function in model (17), we derive the following transformation:

$$\begin{aligned}
&\hat{\mathbf{r}}_t \mathbf{x}_t^\top - \eta \mathbf{x}_t \boldsymbol{\Sigma}_t \mathbf{x}_t^\top - \delta \parallel \mathbf{x}_t - \tilde{\mathbf{x}}_{t-1} \parallel_1 \\
&\quad = \hat{\mathbf{r}}_t ((\mathbf{y}_t, \mathbf{z}_t)N + \tilde{\mathbf{x}}_{t-1})^\top \\
&\qquad - \eta ((\mathbf{y}_t, \mathbf{z}_t)N + \tilde{\mathbf{x}}_{t-1}) \boldsymbol{\Sigma}_t ((\mathbf{y}_t, \mathbf{z}_t)N + \tilde{\mathbf{x}}_{t-1})^\top \\
&\qquad - \delta \mathbf{e}_{2n} (\mathbf{y}_t, \mathbf{z}_t)^\top \\
&\quad = -\eta (\mathbf{y}_t, \mathbf{z}_t) N \Sigma_t N^\top (\mathbf{y}_t, \mathbf{z}_t)^\top \\
&\qquad + (\hat{\mathbf{r}}_t N^\top - \delta \mathbf{e}_{2n} - 2\eta \tilde{\mathbf{x}}_{t-1} \boldsymbol{\Sigma}_t N^\top)(\mathbf{y}_t, \mathbf{z}_t)^\top \\
&\qquad + \hat{\mathbf{r}}_t \tilde{\mathbf{x}}_{t-1}^\top - \eta \tilde{\mathbf{x}}_{t-1} \boldsymbol{\Sigma}_t \tilde{\mathbf{x}}_{t-1}^\top.
\end{aligned}$$

For the first constraint $\mathbf{e}_n \mathbf{x}_t^\top = 1$ and the second constraint $\mathbf{0} \leq \mathbf{x}_t \leq \mathbf{e}_n$, by replacing $\mathbf{x}_t$ with $\mathbf{y}_t$ and $\mathbf{z}_t$ according to equation (19), we obtain

$$\mathbf{e}_n (\mathbf{y}_t - \mathbf{z}_t)^\top = 0 \quad \text{and} \quad -\tilde{\mathbf{x}}_{t-1} \leq (\mathbf{y}_t, \mathbf{z}_t)N \leq \mathbf{e}_n - \tilde{\mathbf{x}}_{t-1},$$

respectively. Then model (17) can be re-formulated as follows:

$$\begin{cases} \max & \boldsymbol{F}_t (\mathbf{y}_t, \mathbf{z}_t)^\top - \eta (\mathbf{y}_t, \mathbf{z}_t) \boldsymbol{H}_t (\mathbf{y}_t, \mathbf{z}_t)^\top + \boldsymbol{C}_t \\ \text{s.t.} & \mathbf{e}_n (\mathbf{y}_t - \mathbf{z}_t)^\top = 0, \\ & -\tilde{\mathbf{x}}_{t-1} \leq (\mathbf{y}_t, \mathbf{z}_t)N \leq \mathbf{e}_n - \tilde{\mathbf{x}}_{t-1}, \\ & \mathbf{0} \leq \mathbf{y}_t, \mathbf{0} \leq \mathbf{z}_t, \end{cases} \tag{20}$$

where $\boldsymbol{F}_t = \hat{\mathbf{r}}_t N^\top - \delta \mathbf{e}_{2n} - 2\eta \tilde{\mathbf{x}}_{t-1} \boldsymbol{\Sigma}_t N^\top$, $\boldsymbol{H}_t = N \boldsymbol{\Sigma}_t N^\top$ and $\boldsymbol{C}_t = \hat{\mathbf{r}}_t \tilde{\mathbf{x}}_{t-1}^\top - \eta \tilde{\mathbf{x}}_{t-1} \boldsymbol{\Sigma}_t \tilde{\mathbf{x}}_{t-1}^\top$. Here $\boldsymbol{H}_t$ satisfies the property in the following theorem.

THEOREM 3.1 *The matrix $\boldsymbol{H}_t$ in model* (20) *is semi-positive definite for $t = 3, 4, \ldots, T$.*

The proof is given in appendix 2. Based on this, we can derive the following theorem.

THEOREM 3.2 *There is at least one optimal solution for model* (20) *if the feasible region is not empty.*

The proof is given in appendix 3.

REMARK 3.2 Theorem 3.2 states that there is an optimal solution for model (20) when its feasible region is not empty. In fact, as it is shown in equation (18), for any given feasible solution $\mathbf{x}_t$ of model (17), we can always find a feasible solution $(\mathbf{y}_t, \mathbf{z}_t)$ of model (20). Due to the fact that the convex feasible region of model (17) contains various feasible solutions, the feasible region of model (20) is not empty.

In addition, we can also derive the following theorem.

THEOREM 3.3 *Model* (20) *achieves the optimal solution $\mathbf{y}_t^* = (y_{t1}^*, y_{t2}^*, \ldots, y_{tn}^*)$ and $\mathbf{z}_t^* = (z_{t1}^*, z_{t2}^*, \ldots, z_{tn}^*)$ if and only if model* (17) *achieves the optimal solution $\mathbf{x}_t^* = (x_{t1}^*, x_{t2}^*, \ldots, x_{tn}^*)$.*

The proof is given in appendix 4.

Model (20) can be solved directly by using the *quadratic* function in Matlab. After deriving the solutions of $\mathbf{y}_t$ and $\mathbf{z}_t$, the solution $\mathbf{x}_t$ can be determined according to equation (19). This method is called AOLPIMV algorithm by integrating AOLPI and AMV together.

REMARK 3.3 The initial investment proportions are uniformly distributed, which means $\mathbf{x}_1 = (\frac{1}{n}, \frac{1}{n}, \ldots, \frac{1}{n})$. This coincides with Guo *et al.* (2021) and Li and Hoi (2012, 2015).

The key of solving online portfolio selection by model (20) lies in the updating process of $\boldsymbol{\Sigma}_t$, which is the advantage of our proposed model over the work by Guo *et al.* (2021). The significant difference between online portfolio selection and traditional multi-period portfolio selection is that the historical return data can be updated constantly as new return data can be obtained in each period. Online portfolio selection concerns more on whether the investment strategy can be derived timely since each time period is so short that the investors have to make quick decisions. It is usually inefficient and impractical to estimate the covariance matrix constantly to measure the investment risk. Owing to this limitation, Guo *et al.* (2021) do not consider the investment risk in their online portfolio selection model. This work overcomes the difficulty of measuring the risk in online portfolio selection by employing the covariance matrix which can be updated linearly in each period (see equation (16)), and an adaptive online mean-variance model is constructed. Thus our proposed online portfolio selection algorithm can be adopted by different types of investors including risk-averse investors.

4. Numerical experiments

In this section, we conduct numerical experiments to validate the effectiveness and advantages of our proposed AOLPIMV algorithm over some other state-of-the-art online portfolio selection algorithms. Some real benchmark data sets are employed which are selected from previous online portfolio selection studies, including MSCI, NYSE-O and TSE (Li and Hoi 2015). In addition, we also select some stocks from the American financial market and construct a new data set

NASTDA. The results reveal that our proposed AOLPIMV algorithm dominates other online portfolio selection algorithms in cumulative return, Sharpe ratio, Information ratio and Calmar ratio, which shows its outstanding effectiveness and practicability.

4.1. Experimental setup

This example provides multiple numerical experiments on different data sets, including MSCI, NYSE-O and TSE (Li and Hoi 2015). MSCI is a benchmark data set obtained from global equity indices that constitute the MSCI World Index. It contains the historical return data of 24 stocks in 1043 trading days ranging from 1 April 2006 to 31 March 2010. NYSE-O is a benchmark data set pioneered by Cover (1991) which contains the historical returns of 36 American stocks in 5651 trading days ranging from 3 June 1962 to 31 December 1984. NYSE-N is an updated data set containing the return data of 23 stocks during the period from 1 January 1985 to 30 June 2010. TSE contains the return data of 88 Canadian stocks in 1259 trading days from 4 January 1994 to 31 December 1998. In addition to the above three data sets, we collect the historical return data of another 20 stocks in the American stock market ranging from 3 January 2006 to 7 October 2010, covering 1119 trading days, which are contained in the dataset NASTDA. The corresponding stock codes are ADBE, ALB, AMZN, AXP, BAX, CMCSA, CSCO, FDX, GS, GT, HA, HMC, INTC, KO, MSFT, NFLX, NKE, NUAN, NVDA, and PG, respectively.

We employ our proposed AOLPIMV algorithm to solve the online portfolio selection problems using these four data sets. As a comparison, we select 19 popular online portfolio selection algorithms from previous studies including AOLNPM (Guo *et al.* 2021), Market strategy (Li and Hoi 2015), UCRP (Li and Hoi 2015), BCRP (Cover 1991), Anticor-1 (Borodin *et al.* 2004), Anticor-2 (Borodin *et al.* 2004), CWMR-V (Li *et al.* 2013), CWMR-S (Li *et al.* 2013), CORN (Li *et al.* 2011a), EG (Helmbold *et al.* 1998), ONS (Agarwal *et al.* 2006), OLMAR-1 (Li and Hoi 2012, 2015), OLMAR-2 (Li and Hoi 2012, 2015), PAMR (Li *et al.* 2012), PAMR-1 (Li and Hoi 2015), PAMR-2 (Li and Hoi 2015), TCO-1 (Li *et al.* 2018), TCO-2 (Li *et al.* 2018) and UP (Li and Hoi 2015, Li *et al.* 2018). The parameters of these algorithms are set as follows. Among all the 20 algorithms, AOLPIMV and AOMNPM employ the adaptive decaying factors which can be adjusted in the investment process, and the corresponding iteration step size γ is set as 0.00040, 0.00045, 0.00100 0.00065 and 0.00010 for MSCI, NYSE-O, NYSE-N, TSE and NASTDA, respectively. It is stated in Guo *et al.* (2021) that the step size should be set separately for different data sets since their return magnitudes differ a lot. Similarly, the risk weighting factor η in model (17) is set as 0.6, 0.05, 0.1, 0.7 and 0.005, respectively. The window size w is set to be 6, which is consistent with the work of Guo *et al.* (2021). The transaction cost δ is set to be 0.0005. For TCO-1 and TCO-2, the parameters are set as the values adopted in Li *et al.* (2018) (That is, $\eta = 10$ and $\lambda = 10$. We remark that the η in TCO-1 and TCO-2 refers to a smoothing parameter used in the updating of the portfolio, which is different from the risk weighting factor in this work). For the AOLNPM algorithm, the parameter window size is selected from Guo *et al.* (2021) which is $w = 6$, and the step size is the same as that of AOLPIMV in this work for a better comparison. For the Market strategy, the weight for each asset is $1/n$. For the other 15 algorithms, the parameters are set as the default values in Li and Hoi (2015) (Specifically, Anticor-1: maximal window size $W = 30$, Anticor-2: maximal window size $W = 30$, CWMR-V: confidence parameter $\phi = 2$ and mean-reversion threshold $\epsilon = 0.5$, CWMR-S: confidence parameter $\phi = 2$ and mean-reversion threshold $\epsilon = 0.5$, CORN: window size $w = 5$ and correlation threshold $c = 0.1$, EG: learning rate $\eta = 0.05$, ONS: mixture parameter $\eta = 0$, trade-off parameter $\beta = 1$ and heuristic tuning parameter $\delta = 1/8$, PAMR: mean-reversion threshold $\epsilon = 0.5$, PAMR-1: mean-reversion threshold $\epsilon = 0.5$ and aggressive parameter $C = 500$, PAMR-2: mean-reversion threshold $\epsilon = 0.5$ and aggressive parameter $C = 500$, OLMAR-1: mean-reversion threshold $\epsilon = 10$ and window size $W = 10$, OLMAR-2: mean-reversion threshold $\epsilon = 10$ and decaying factor $\alpha = 0.5$. BCRP, UCRP and UP do not need any parameters except the transaction cost rate. Remark that the transaction cost rate for all these 15 algorithms is set as $\lambda = 2\delta = 0.001$ since the transaction cost is calculated by $\frac{\lambda}{2} \parallel \mathbf{x}_t - \tilde{\mathbf{x}}_{t-1} \parallel_1$ in Li and Hoi (2015). We then employ these 20 algorithms to solve the same online portfolio selection problems, and compare their performances in terms of cumulative return, Sharpe ratio, Information ratio and Calmar ratio.

4.2. Cumulative return

Cumulative return refers to the total return from the beginning to the end of the investment period, which is an important objective that the investors concern and can be used to reflect the performance of the algorithm in making profits. (The formula of the cumulative return W_T can be seen in equation (1) .) Note that the cumulative returns in different data sets exhibit considerable disparities. For example, the cumulative returns obtained on MSCI are no more than 15. While some cumulative returns obtained on NYSE-O are larger than 1×10^{16}. For a better comparison of the diverse portfolio strategies within and across individual data sets, we employ the following normalization method to obtain the normalized cumulative returns (NW_T):

$$NW_T = \frac{1}{1 + e^{-0.3 \log_{10} W_T}}, \tag{21}$$

which is able to transform the original cumulative return into a value ranging from 0 to 1, and guarantees that a larger cumulative return gives a larger normalized cumulative return (NW_T increases monotonically with respect to W_T). For example, the cumulative return of AOLPIMV on NYSE-O is 1.3991×10^{16}, its normalized cumulative return derived by equation (21) is 0.9922. The normalized cumulative returns of these 20 algorithms on MSCI, NYSE-O, NYSE-N, TSE and NASTDA are shown in table 3.

It can be seen from table 3 that AOLPIMV gains the largest cumulative returns on MSCI, NYSE-O, NYSE-N and TSE, and second largest return on NASTDA. As a comparison,

Table 3. NW_T on MSCI, NYSE-O, NYSE-N, TSE and NASTDA.

Method	MSCI	NYSE-O	NYSE-N	TSE	NASTDA
AOLPIMV	**0.5819**	**0.9922**	**0.8902**	**0.6817**	**0.5528**
AOLNPM	0.5781	0.9910	0.8651	0.6791	0.5288
Anticor-1	0.5271	0.8816	0.8011	0.5982	0.5342
Anticor-2	0.5305	0.9089	0.8606	0.6066	0.5410
BCRP	0.5133	0.6707	0.6500	0.5618	**0.5563**
CWMR-V	0.5624	0.9841	0.7596	0.6527	0.5234
CWMR-S	0.5624	0.9841	0.7591	0.6530	0.5234
CORN	0.5477	0.9509	0.7044	0.5372	0.5000
EG	0.4973	0.6048	0.6089	0.5149	0.5161
Market	0.4968	0.5862	0.5931	0.5156	0.5155
ONS	0.4938	0.6438	0.5947	0.5137	0.5350
OLMAR-1	0.5649	0.9897	0.8872	0.6049	0.5467
OLMAR-2	0.5745	0.9920	0.8786	0.6752	0.5279
PAMR	0.5583	0.9835	0.7550	0.6459	0.5204
PAMR-1	0.5589	0.9835	0.7550	0.6459	0.5204
PAMR-2	0.5621	0.9835	0.7588	0.6443	0.5222
TCO-1	0.5546	0.9806	0.8602	0.6325	0.5478
TCO-2	0.5471	0.9753	0.8723	0.6421	0.5489
UCRP	0.4974	0.6048	0.6093	0.5149	0.5187
UP	0.4970	0.6032	0.6081	0.5143	0.5187

AOLNPM gains the second largest return on MSCI and TSE. In general, AOLPIMV and AOLNPM dominate most of the other 18 algorithms in cumulative return, which is consistent with the results in Guo *et al.* (2021). The difference between AOLPIMV and AOLNPM lies in the fact that AOLPIMV introduces the peer impact in the prediction of future return and considers the risk constraint in the objective function of net profit maximization model. It seems countering common sense that the cumulative return with risk constraint is larger than the model without risk constraint. Since the aggressive investment strategy obtained by AOLNPM algorithm aims to achieve the largest return without considering investment risk, which has the possibility of losing a large amount of money in real investment process. The strategy derived by AOLPIMV is relatively conservative and may achieve a larger return than the aggressive strategy. In addition, as is shown in the Experimental Setup, the investment horizons of different data sets differ a lot. For example, MSCI contains the return data of 5 years, while NYSE-N contains the return data of 26 years. For a better comparison, we derive the average annualized cumulative return (AACR) by calculating the cumulative returns of each year (see table 4). Here, the AACRs are not normalized since they can provide more valuable return information (the AACRs of the Market strategy on these five data sets are 1.0385, 1.1391, 1.1277, 1.1094 and 1.1605, indicating that the corresponding annualized net return rates are 3.85%, 13.91%, 12.77%, 10.94% and 16.05%, respectively). Remark that the largest cumulative return does not necessarily lead to the largest AACR since cumulative return refers to the total return of the whole investment horizon while AACR is the average value of cumulative returns of each single year. It can be seen clearly from table 4 that AOLPIMV achieves good performances on all these five data sets. By comprehensively analyzing the cumulative return and average annualized cumulative return, AOLPIMV performs best among all these 20 online portfolio selection algorithms.

Furthermore, we study the cumulative returns of AOLPIMV under different transaction cost rates and make comparisons with the other algorithms. Take MSCI as an example, we set the transaction cost rate as 0.0005, 0.00075, 0.001 and 0.00125, respectively. Due to the fact that the Market strategy does not require any capital reallocation in the online decision-making process, we only focus on the other 19 algorithms (see table 5). It is clear that the NW_T of AOLPIMV is still the largest for all different transaction cost rates, and AOLNPM performs best in ACCR. For each algorithm, the NW_T and ACCR decrease when the transaction cost rate increases, which is consistent with the results in Guo *et al.* (2021), Li and Hoi (2015), and Li *et al.* (2018). This is because a larger transaction cost rate will increase the trading cost and decrease the cumulative return. In addition, we calculate the average turnover of each algorithm by

$$AT = \frac{1}{2T}\sum_{t=1}^{T} \| \mathbf{x}_t - \tilde{\mathbf{x}}_{t-1} \|_1 . \qquad (22)$$

Remark that equation (22) is originated from Li *et al.* (2018), and there is no transaction remainder factor which is used in Li *et al.* (2018) since we calculate the transaction cost by $\delta \| \mathbf{x}_t - \tilde{\mathbf{x}}_{t-1} \|_1$. The turnovers of all the algorithms are shown in figure 3. We observe that BCRP and UCRP achieve the smallest turnovers since these two algorithms are often used as benchmarks. The turnovers of ONS and UP are very small, which indicates the strategies are adjusted very rarely, and their cumulative returns are both less than 1 ($NW_T < 0.5$). CORN achieves the largest turnover, but its cumulative return is only 4.3459 (the corresponding NW_T is 0.5477 shown in table 3). By way of comparison, the AOLPIMV turnover ranks in a middling position, yet it generates the largest cumulative return. This suggests that excessively frequent strategy adjustments may result in high transaction costs and reduced cumulative returns, while infrequent adjustments could entail missed profit-making opportunities. Therefore, striking a balance between transaction cost, return, and risk is considered optimal. This explains why AOLPIMV performs the best in cumulative return.

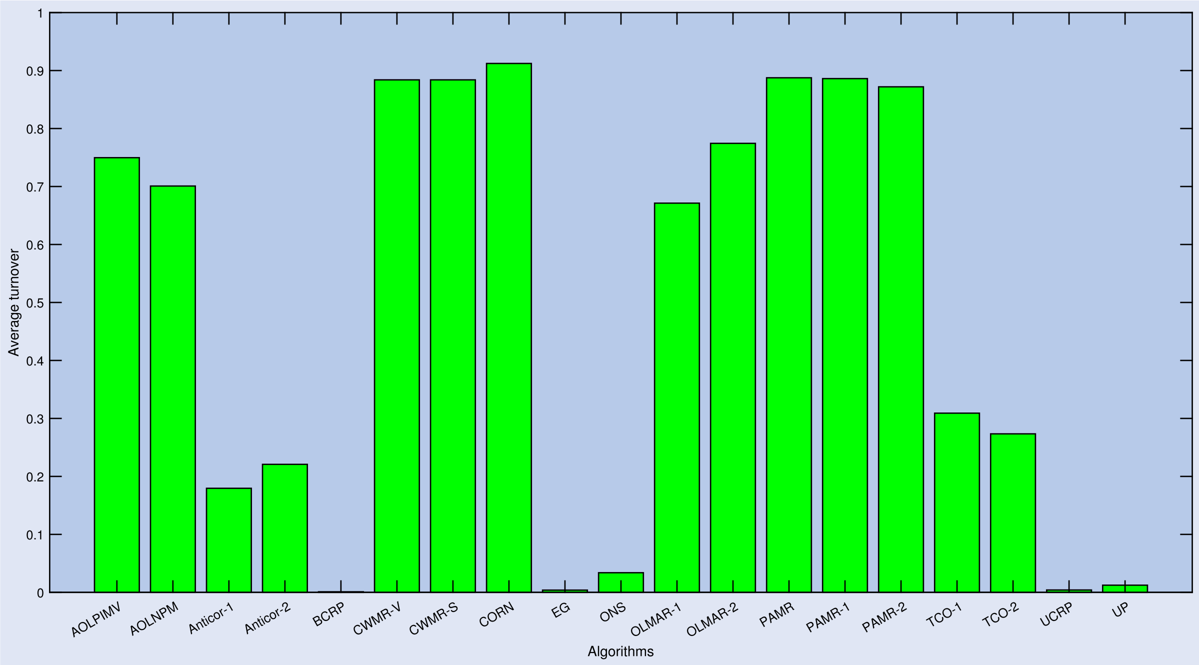

Figure 3. Turnovers under different algorithms on MSCI.

Table 4. AACR on MSCI, NYSE-O, NYSE-N, TSE and NASTDA.

Method	MSCI	NYSE-O	NYSE-N	TSE	NASTDA
AOLPIMV	**1.7688**	**7.3908**	**2.2004**	**4.8028**	**1.5412**
AOLNPM	**1.7842**	7.6885	2.0367	4.2890	1.3721
Anticor-1	1.2174	2.0660	1.5667	2.2226	1.2899
Anticor-2	1.2436	2.3202	1.8085	2.5480	1.3717
BCRP	1.1207	1.3301	1.1515	1.5398	1.5258
CWMR-V	1.5411	6.2649	1.5459	3.6514	1.1921
CWMR-S	1.5413	6.2593	1.5454	3.6521	1.1925
CORN	1.4324	3.4144	1.3717	1.4693	1.0747
EG	1.0416	1.1658	1.1507	1.1052	1.1673
Market	1.0385	1.1391	1.1277	1.1094	1.1605
ONS	1.0347	1.2416	1.1654	1.1017	1.3712
OLMAR-1	1.5445	6.4067	2.1970	3.5399	1.5245
OLMAR-2	1.6723	**8.0230**	2.1344	**4.9246**	1.3374
PAMR	1.5033	6.3362	1.5240	3.4066	1.1736
PAMR-1	1.5091	6.3362	1.5239	3.4066	1.1732
PAMR-2	1.5372	6.2942	1.5322	3.3627	1.1889
TCO-1	1.4381	5.4445	1.9156	3.1525	1.4835
TCO-2	1.3760	4.3226	2.0043	3.7739	**1.5823**
UCRP	1.0418	1.1654	1.2214	1.1054	1.1865
UP	1.0394	1.1628	1.1497	1.1016	1.1866

4.3. Sharpe ratio

Sharpe ratio is a widely applied risk-adjusted return measure firstly proposed by Nobel laureate Sharpe (1966). Apart from investment return, investors also concern the asset price fluctuation which is an investment risk. For portfolios with equal or similar investment return, the investors prefer the one with low price volatility. For convenience of comparing portfolios with different investment return and risk, Sharpe ratio is proposed by calculating the ratio between average return earned in excess of the risk-free rate and the volatility, which can be expressed as follows:

$$Sharpe\ ratio = \frac{\bar{r} - r_f}{\sigma}. \tag{23}$$

Here $\bar{r}$ is the average return over all the investment horizon, r_f is the risk-free return which is set as 1 (this is reasonable since each investment period we consider here is very short), and σ refers to the return volatility of the whole portfolio.

Generally, the investors prefer a portfolio with larger Sharpe ratio. We obtain the Sharpe ratios of returns derived by different algorithms (see table 6). It is clear that AOLPIMV performs the best in MSCI and TSE, achieves the second largest ratio on NYSE-O, the fourth largest ratio on NYSE-N, and third largest ratio on NASTDA, while TCO-1 performs the best in NYSE-O and BCRP performs the best in NASTDA, and Anticor-2 performs the best on NYSE-N. For AOLNPM, although it also performs well in MSCI, NYSE-O, NYSE-N and TSE, its performance in NASTDA is quite poor,

Table 5. NW_T and AACR under different transaction cost rates.

	$\delta = 0.0005$		$\delta = 0.00075$		$\delta = 0.001$		$\delta = 0.00125$	
Method	NW_T	AACR	NW_T	AACR	NW_T	AACR	NW_T	AACR
AOLPIMV	**0.5819**	**1.7688**	**0.5685**	**1.6127**	**0.5589**	**1.5078**	**0.5534**	**1.4526**
AOLNPM	0.5781	**1.7842**	0.5641	**1.6206**	0.5570	**1.5219**	0.5503	**1.4611**
Anticor-1	0.5357	1.2174	0.5271	1.2174	0.5210	1.1720	0.5179	1.1501
Anticor-2	0.5305	1.2436	0.5268	1.2147	0.5230	1.1865	0.5193	1.1591
BCRP	0.5133	1.1207	0.5133	1.1205	0.5132	1.1204	0.5132	1.1202
CWMR-V	0.5624	1.5411	0.5475	1.3979	0.5325	1.2693	0.5175	1.1538
CWMR-S	0.5624	1.5413	0.5475	1.3981	0.5326	1.2695	0.5175	1.1539
CORN	0.5477	1.4324	0.5319	1.2945	0.5164	1.1731	0.5008	1.0644
EG	0.4973	1.0416	0.4973	1.0412	0.4972	1.0407	0.4971	1.0403
ONS	0.4938	1.0347	0.4932	1.0311	0.4926	1.0276	0.4920	1.0240
OLMAR-1	0.5649	1.5445	0.5537	1.4336	0.5422	1.3314	0.5308	1.2372
OLMAR-2	0.5745	1.6723	0.5616	1.5330	0.5485	1.4064	0.5354	1.2911
PAMR	0.5583	1.5033	0.5433	1.3636	0.5283	1.2381	0.5131	1.1254
PAMR-1	0.5589	1.5091	0.5440	1.3689	0.5289	1.2430	0.5139	1.1299
PAMR-2	0.5621	1.5372	0.5474	1.3962	0.5326	1.2694	0.5178	1.1553
TCO-1	0.5546	1.4381	0.5470	1.3679	0.5395	1.3021	0.5332	1.2514
TCO-2	0.5471	1.3760	0.5425	1.3365	0.5379	1.2983	0.5333	1.2613
UCRP	0.4974	1.0418	0.4973	1.0413	0.4972	1.0408	0.4971	1.0403
UP	0.4970	1.0394	0.4969	1.0389	0.4966	1.0371	0.4963	1.0356

of which the ratio is lower than that of AOLPIMV, BCRP, Anticor-1, Anticor-2, ONS, OLMAR-1, TCO-1 and TCO-2. Consider datasets MSCI, NYSE-O, NYSE-N, TSE and NASTDA, the Sharpe ratios of AOLPIMV are 7.83%, 6.55%, 8.89%, 2.49%, and 44.50%, respectively, larger than AOLNPM. This shows that AOLPIMV achieves relatively good and steady performance while the performance of AOLNPM differs a lot on different data sets.

We then employ the bootstrap-based Sharpe ratio difference significance test to check whether the Sharpe ratio derived by AOLPIMV is significantly different from that of the other algorithms. This method is proposed by Ledoit and Wolf (2008) who used the studentized time series bootstrap confidence interval to check the difference of two Sharpe ratios. The p-values for the test between the Sharpe ratio of AOLPIMV and that of the other algorithms are shown in table 7. For each data set, there are two columns containing Diff and p-value, where Diff represents the Sharpe ratio difference between AOLPIMV and another selected algorithm. For example, by comparing AOLPIMV and Anticor-1 on MSCI, the Sharpe ratio difference is 0.0602 and the p-value is 0.0002. According to table 7, most of the p-values are less than 0.05, which is able to support that the Sharpe ratios are significantly different. It should be noted that the p-values of AOLNPM, OLMAR-1 and OLMAR-2 are larger than 0.05 in most of the data sets. This may be due to the fact that these three algorithms also perform well in the Sharpe ratio with their performance similar to AOLPIMV, causing the difference not to be very significant. In addition, for NASTDA, there are eight p-values larger than 0.05, which may be explained by the following reasons. Firstly, the Sharpe ratios of all the algorithms in NASTDA are much closer compared with the other four data sets (see table 6), with the largest ratio 0.0627 and smallest ratio 0.0155. Secondly, the selection of the significance test method also affects the final p-values. As is validated in Ledoit and Wolf (2008), for the same testing data, different methods may lead to significantly distinct p-values (see table 3 in Ledoit and Wolf 2008). Finally, randomness exists in the bootstrap-based Sharpe ratio difference significance test. For the same parameter set and data set, different results may be obtained in multiple trials. In general, AOLPIMV is significantly different from the majority of the other algorithms in Sharpe ratios.

4.4. *Information ratio*

Different from the Sharpe ratio measuring the excess return over risk-free return rate per unit of risk, Information ratio is a risk-adjusted return measuring the excess return over a benchmark such as market index (Guan and An 2019). On the one hand, Information ratio can be used to reflect the portfolio manager's ability of generating excess return. On the other hand, the tracking error, representing the standard deviation of the difference between the return of the portfolio and the benchmark return, is incorporated to measure the consistency of the performance. Here the tracking error represents the consistency level of the portfolio tracking the benchmark. For the investors, the portfolio with smaller tracking error is better since it means that the portfolio can beat the benchmark consistently over time. The Information ratio is calculated by the following formula:

$$\textit{Information ratio} = \frac{\bar{r} - \bar{r}^{\star}}{\sigma(r - r^{\star})}, \tag{24}$$

where $\bar{r}^{\star}$ is the average return of the benchmark and $\sigma(r - r^{\star})$ represents the tracking error. In this paper, the benchmark is set as the return of the Market strategy.

Table 8 provides the Information ratios of all the 20 algorithms on different data sets. It is clear that AOLPIMV achieves the largest Information ratios on TSE, and second largest ratio on MSCI and NYSE-O, and the fifth largest ratio on NYSE-N. TCO-1 performs the best on MSCI and NYSE-O. There are several Information ratios which are less than 0. For example, the Information ratio of ONS on dataset MSCI is -0.0027, which means that the average

Table 6. Sharpe ratios on MSCI, NYSE-O, NYSE-N, TSE and NASTDA.

Method	MSCI	NYSE-O	NYSE-N	TSE	NASTDA
AOLPIMV	**0.1115**	**0.2032**	0.0870	**0.1072**	0.0552
AOLNPM	0.1034	0.1907	0.0799	0.1046	0.0382
Anticor-1	0.0513	0.1583	0.0862	0.0982	0.0472
Anticor-2	0.0538	0.1502	**0.0929**	0.0882	0.0498
BCRP	0.0381	0.0597	0.0166	0.0725	**0.0627**
CWMR-V	0.0920	0.1907	0.0594	0.1020	0.0344
CWMR-S	0.0921	0.1907	0.0591	0.1023	0.0344
CORN	0.0821	0.1383	0.0573	0.0428	0.0155
EG	0.0030	0.0722	0.0501	0.0485	0.0321
Market	0.0017	0.0552	0.0457	0.0505	0.0317
ONS	0.0002	0.0767	0.0305	0.0264	0.0492
OLMAR-1	0.0897	0.1913	0.0863	0.0714	0.0503
OLMAR-2	0.1003	0.2014	0.0840	0.1027	0.0376
PAMR	0.0866	0.1886	0.0589	0.1016	0.0322
PAMR-1	0.0874	0.1886	0.0589	0.1016	0.0322
PAMR-2	0.0922	0.1901	0.0600	0.1008	0.0335
TCO-1	0.0893	**0.2119**	0.0902	0.0899	0.0549
TCO-2	0.0768	0.1945	0.0887	0.0929	0.0556
UCRP	0.0031	0.0725	0.0501	0.0485	0.0359
UP	0.0023	0.0715	0.0496	0.0467	0.0358

Table 7. p-Values with significance level 5%.

Method	MSCI		NYSE-O		NYSE-N		TSE		NASTDA	
AOLPIMV	Diff	p-value	Diff	p-value	Diff	p-value	Diff	p-value	Diff	p-value
AOLNPM	0.0081	**0.0934**	0.0125	0.0026	0.0071	**0.9950**	0.0026	**0.1628**	0.0071	**0.1104**
Anticor-1	0.0602	0.0002	0.0449	0.0002	0.0009	0.0002	0.0090	0.0002	0.0009	0.0002
Anticor-2	0.0577	0.0004	0.0530	0.0002	−0.0059	0.0026	0.0189	0.0002	−0.0059	0.0002
BCRP	0.0734	0.0002	0.1435	0.0002	0.0704	0.0002	0.0347	0.0002	0.0704	0.0042
CWMR-V	0.0195	0.0004	0.0125	0.0002	0.0277	0.0002	0.0051	**0.0880**	0.0277	**0.1318**
CWMR-S	0.0194	0.0006	0.0125	0.0002	0.0280	0.0002	0.0049	**0.0870**	0.0280	**0.1292**
CORN	0.0294	0.0004	0.0648	0.0012	0.0298	0.0002	0.0644	0.0002	0.0298	0.0128
EG	0.1085	0.0002	0.1310	0.0002	0.0369	0.0002	0.0587	0.0002	0.0369	0.0002
Market	0.1098	0.0002	0.1480	0.0002	0.0413	0.0002	0.0567	0.0002	0.0413	0.0002
ONS	0.1113	0.0002	0.1265	0.0002	0.0565	0.0002	0.0807	0.0002	0.0565	0.0002
OLMAR-1	0.0218	**0.3462**	0.0119	**0.5364**	0.0007	**0.4806**	0.0357	**0.2474**	0.0007	**0.1056**
OLMAR-2	0.0112	**0.0782**	0.0018	**0.6748**	0.0030	**0.2412**	0.0045	**0.1040**	0.0030	**0.1164**
PAMR	0.0249	0.0012	0.0146	0.0002	0.0281	0.0002	0.0056	0.0014	0.0281	**0.2014**
PAMR-1	0.0241	0.0016	0.0146	0.0002	0.0281	0.0002	0.0056	0.0020	0.0281	**0.2142**
PAMR-2	0.0192	0.0002	0.0131	0.0002	0.0271	0.0002	0.0064	0.0012	0.0271	**0.0560**
TCO-1	0.0222	0.0002	−0.0088	0.0002	−0.0032	0.0002	0.0173	0.0468	−0.0032	0.0002
TCO-2	0.0347	0.0002	0.0087	0.0002	−0.0016	0.0002	0.0143	0.0490	−0.0016	0.0002
UCRP	0.1084	0.0002	0.1307	0.0002	0.0369	0.0002	0.0587	0.0002	0.0369	0.0002
UP	0.1092	0.0002	0.1317	0.0002	0.0374	0.0002	0.0605	0.0002	0.0374	0.0002

return of ONS is less than the Market strategy. This coincides with the result that ONS obtains smaller cumulative return compared with the Market strategy. However, a negative Information ratio does not always guarantee that the cumulative return of the algorithm is worse than that of the Market strategy. Take BCRP as an example, its Information ratio on NYSE-N is −0.0057, but its cumulative return ($W_T = 115.6646$) is significantly larger than that of the Market strategy ($W_T = 18.0566$), which justifies that the Information ratio and cumulative return are intrinsically different performance metrics. Although the AOLNPM algorithm proposed in Guo *et al.* (2021) also performs well on these data sets, our proposed AOLPIMV further improves the Information ratios of returns compared with AOLNPM, which reflects its strong ability of making excess returns and high consistency level.

4.5. Calmar ratio

Calmar ratio is another widely adopted risk-adjusted return measure which was firstly proposed by Young (1991). One significant contribution of Calmar ratio lies in the introduction of maximum drawdown to measure investment risk instead of variance. When measuring risk by variance, high return is treated equally to low return and may also lead to a large variance. As a comparison, the maximum drawdown focuses on the maximal loss from the peak of the portfolio to the trough before the next peak is reached, which is a downside risk measure and is preferred by investors. Calmar ratio is calculated by

$$\mathit{Calmar\ ratio} = \frac{\bar{r} - 1}{\mathit{MDD}(r)}, \tag{25}$$

Table 8. Information ratios on MSCI, NYSE-O, NYSE-N, TSE, and NASTDA.

Method	MSCI	NYSE-O	NYSE-N	TSE	NASTDA
AOLPIMV	**0.1584**	**0.2001**	0.0770	**0.1022**	0.0495
AOLNPM	0.1522	0.1871	0.0701	0.0998	0.0291
Anticor-1	0.1235	0.1576	0.0765	0.0903	0.0459
Anticor-2	0.1057	0.1447	**0.0837**	0.0802	0.0477
BCRP	0.0359	0.0386	−0.0057	0.0617	**0.0562**
CWMR-V	0.1375	0.1863	0.0469	0.0963	0.0241
CWMR-S	0.1375	0.1863	0.0466	0.0965	0.0241
CORN	0.1161	0.1302	0.0399	0.0331	−0.0029
EG	0.0281	0.0345	0.0242	−0.0082	0.0096
ONS	−0.0027	0.0394	0.0121	0.0069	0.0515
OLMAR-1	0.1297	0.1870	0.0771	0.0659	0.0445
OLMAR-2	0.1466	0.1982	0.0745	0.0976	0.0284
PAMR	0.1291	0.1839	0.0462	0.0956	0.0214
PAMR-1	0.1305	0.1839	0.0462	0.0956	0.0213
PAMR-2	0.1400	0.1856	0.0473	0.0948	0.0229
TCO-1	**0.1665**	**0.2123**	0.0797	0.0838	0.0535
TCO-2	0.1410	0.1940	0.0790	0.0872	0.0542
UCRP	0.0277	0.0337	0.0238	−0.0075	0.0503
UP	0.0128	0.0306	0.0221	−0.0139	0.0378

where $MDD(r) = \sqrt{\frac{1}{T}\sum_{t=1}^{T}\min\{r_t - 1, 0\}^2}$. The Calmar ratios of these 20 algorithms are shown in table 9. It can be found that all the obtained Calmar ratios are larger than 0, representing that the average returns of all the algorithms in different data sets are larger than 1 according to equation (25). Particularly, our proposed AOLPIMV performs best in MSCI, and achieves the second largest ratio in NYSE-O. For MSCI, the Calmar ratio derived from AOLPIMV is larger than the second largest ratio by 6.96%. For NYSE-N, TSE and NASTDA, AOLPIMV also achieves satisfactory performance which beats 75% of the algorithms. The good performance of AOLPIMV can be attributed to the consideration of risk constraint in online decision-making of each period, which is beneficial to averting the risky assets with large return volatility. As a comparison, AOLNPM aims to maximize the net return of each period without considering investment risk and achieves smaller Calmar ratio than AOLPIMV on each of the dataset. Although TCO-1 performs the best on NYSE-O, its performance in MSCI and TSE is poor. In general, AOLPIMV performs the best among all the algorithms considering the cumulative return, Sharpe ratio, Information ratio and Calmar ratio.

4.6. Influence of the peer impact and risk constraint

To further study the influence of introducing peer impact on the online portfolio selection decision-making, we use the adaptive moving average mean-variance (AOLMAMV) algorithm, where the AOLMA is used to predict the future returns of risky assets without considering the peer impact. By setting $\eta = 0.6$ and $\gamma = 0.00040$, we derive the cumulative returns of AOLPIMV and AOLMAMV (see figure 4). It is clear that AOLPIMV achieves a larger return. The Sharpe ratio, Information ratio and Calmar ratio of AOLMAMV are 0.1040, 0.1532 and 0.1616, respectively, which are all smaller than the values of AOLPIMV (see tables 6, 8 and 9). The coefficients $\hat{\alpha}_{ti}$ and $\hat{\beta}_{ti}$ can be adjusted automatically when new return data is available in each period. Take asset 14 as an example, the values of $\hat{\alpha}_{t(14)}$ and $\hat{\beta}_{t(14)}$ are updated in each period (see figure 5). At the beginning of the investment, $\hat{\alpha}_{t(14)}$ is nearly equal to 1 and $\hat{\beta}_{t(14)}$ is almost 0. As the investment process goes on, the value of $\hat{\beta}_{t(14)}$ increases gradually, meaning that the performance of the other assets in the market contributes increasingly to the estimation of the return of this asset. Although $\hat{\beta}_{t(14)}$ is much smaller than $\hat{\alpha}_{t(14)}$ in the whole process, it cannot be neglected, and considering the peer impact in online portfolio selection can significantly improve the performance of algorithm.

To study the impact of risk constraint on the performance of portfolio, we select the dataset MSCI and set the weighting factor η in the objective function of model (20) varying from 0.6 to 200.6. The cumulative return is shown in figure 6. It can be seen clearly that the cumulative return decreases when the weighting factor increases, which is consistent with the real investment behaviors since the investors are getting more and more conservative as the weighting factor becomes larger. The investment decision is more aggressive when the weighting factor is smaller. Take for example, for the cumulative return with $\eta = 0.6$, it is larger than the return with $\eta = 90.6$ from period 500 to period 600. However, this return increases significantly from period 500 to period 600, and then is encountered with a large loss from period 600 to period 700, which exhibits a large volatility. As a comparison, the cumulative return with $\eta = 90.6$ increases all the time consistently and steadily though the net profit is not large, which is preferred for risk-neural and risk-averse investors. We also derive the NW_T, ACCR, Sharpe ratio, Information ratio and Calmar ratio with different weighting factors (see table 10). The value of ACCR decreases as the weighting factor becomes larger, which is consistent with the cumulative return. It is clear that the Sharpe ratio is the largest when $\eta = 90.6$, Information ratio is the largest when $\eta = 40.6$, Calmar ratio is the largest when $\eta = 100.6$. These three ratios are quite small when $\eta = 0.6$ though the cumulative return is very large. This reflects the necessity and importance of considering risk constraint in a real investment process. There are three

Table 9. Calmar ratios on MSCI, NYSE-O, NYSE-N, TSE, and NASTDA.

Method	MSCI	NYSE-O	NYSE-N	TSE	NASTDA
AOLPIMV	**0.1721**	**0.4062**	0.1427	0.1898	0.0851
AOLNPM	0.1609	0.3724	0.1277	0.1810	0.0578
Anticor-1	0.0751	0.2862	0.1368	0.1635	0.0731
Anticor-2	0.0797	0.2726	**0.1541**	0.1452	0.0790
BCRP	0.0520	0.0941	0.0244	0.1199	**0.1001**
CWMR-V	0.1377	0.3853	0.0960	**0.1905**	0.0523
CWMR-S	0.1378	0.3853	0.0959	**0.1910**	0.0524
CORN	0.1289	0.2607	0.0916	0.0696	0.0226
EG	0.0041	0.1106	0.0704	0.0649	0.0457
Market	0.0023	0.0835	0.0637	0.0675	0.0448
ONS	0.0002	0.1252	0.0457	0.0406	0.0744
OLMAR-1	0.1365	0.3737	0.1420	0.1233	0.0786
OLMAR-2	0.1549	0.4001	0.1380	0.1788	0.0572
PAMR	0.1281	0.3798	0.0946	0.1828	0.0491
PAMR-1	0.1294	0.3798	0.0946	0.1828	0.0490
PAMR-2	0.1370	0.3842	0.0965	0.1814	0.0511
TCO-1	0.1359	**0.4443**	0.1484	0.1646	0.0878
TCO-2	0.1165	0.3850	0.1478	0.1788	0.0880
UCRP	0.0042	0.1113	0.0704	0.0650	0.0513
UP	0.0032	0.1096	0.0697	0.0626	0.0512

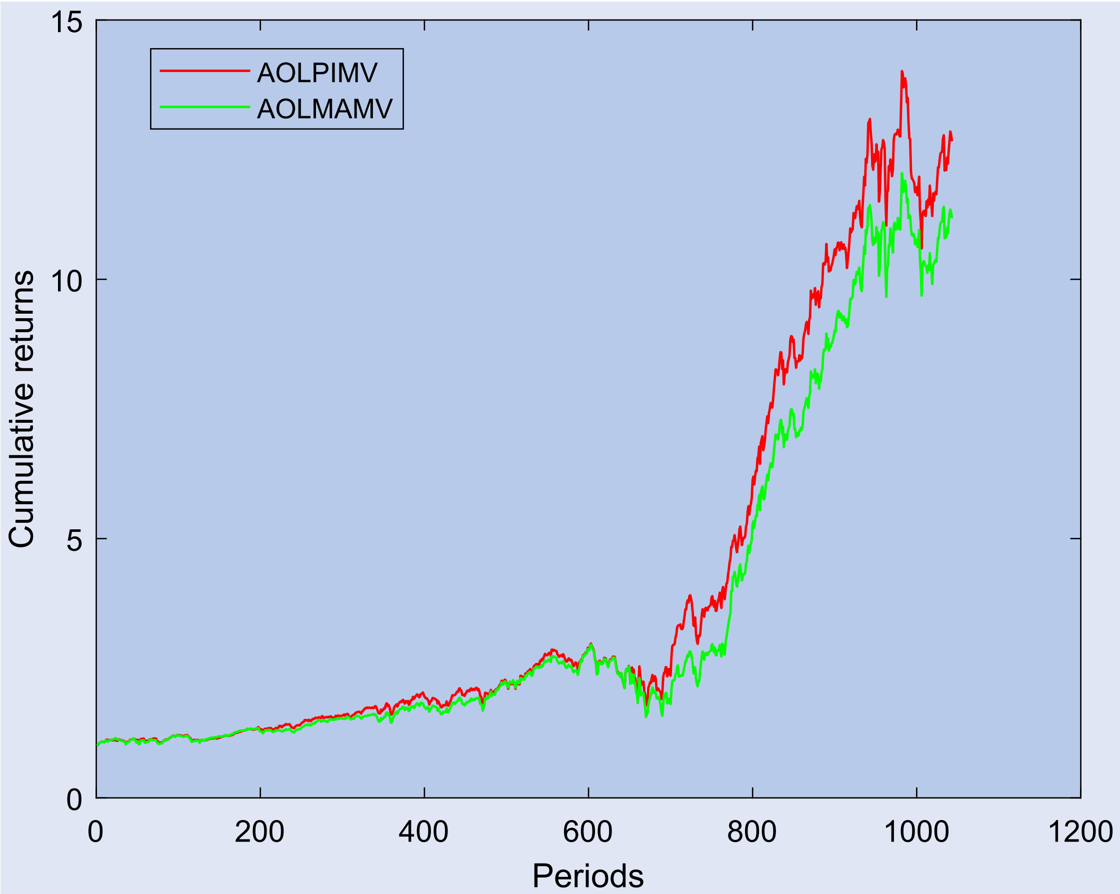

Figure 4. Cumulative returns of AOLPIMV and AOLMAMV.

types of investors: risk-appetite, risk-neural and risk averse. The AOLNPM algorithm proposed in Guo *et al.* (2021) is actually intended for risk-appetite investors which aims to maximize the net profit in each period without considering the investment risk. Then the risk-neural and risk-averse investors would not take the risk of employing this algorithm in the real investment. As a comparison, the AOLPIMV algorithm simultaneously considers the return and risk in each period and can provide flexible investment strategies for different types of investors.

In addition, we analyze the capital allocation of the optimal investment strategies in each period. By setting $\eta = 90.6$, we employ the AOLPIMV algorithm to derive the final investment strategies over the 1043 periods. Since the

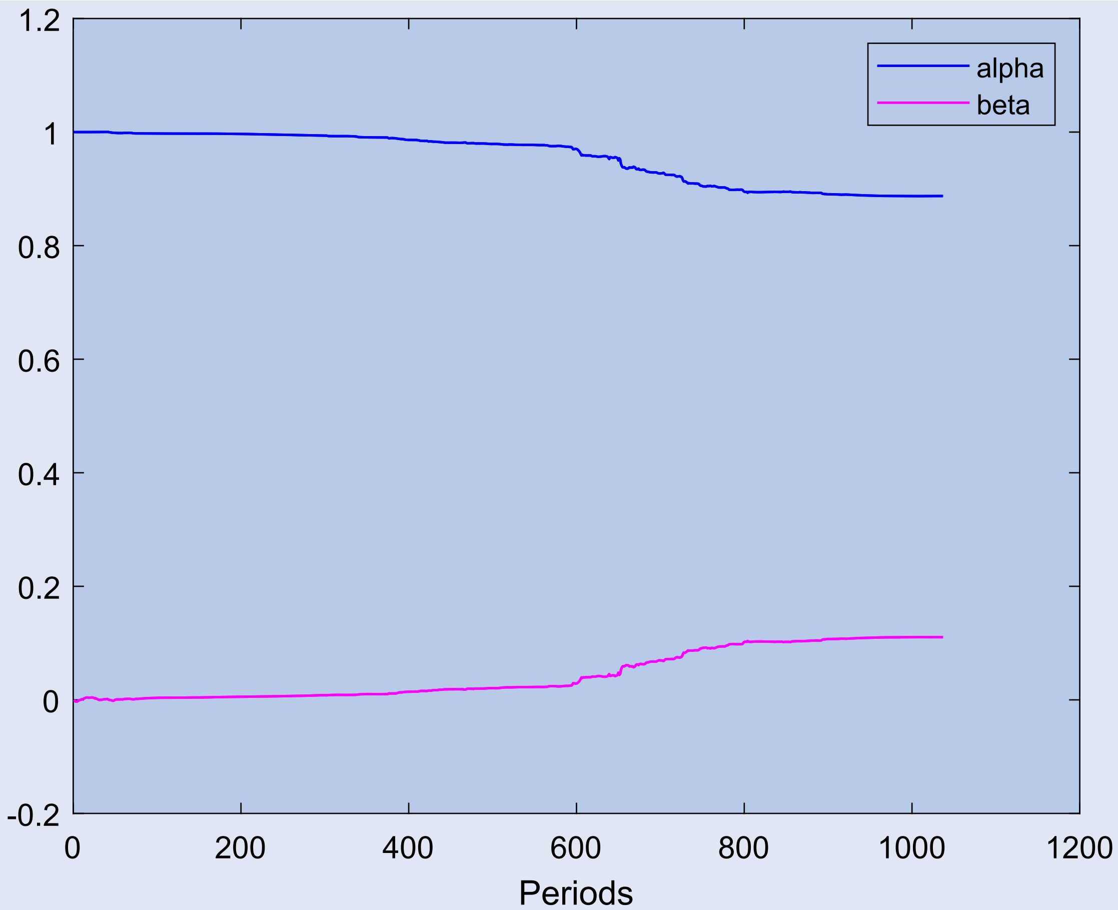

Figure 5. $\hat{\alpha}_{t(14)}$ and $\hat{\beta}_{t(14)}$ iteration process for asset 14.

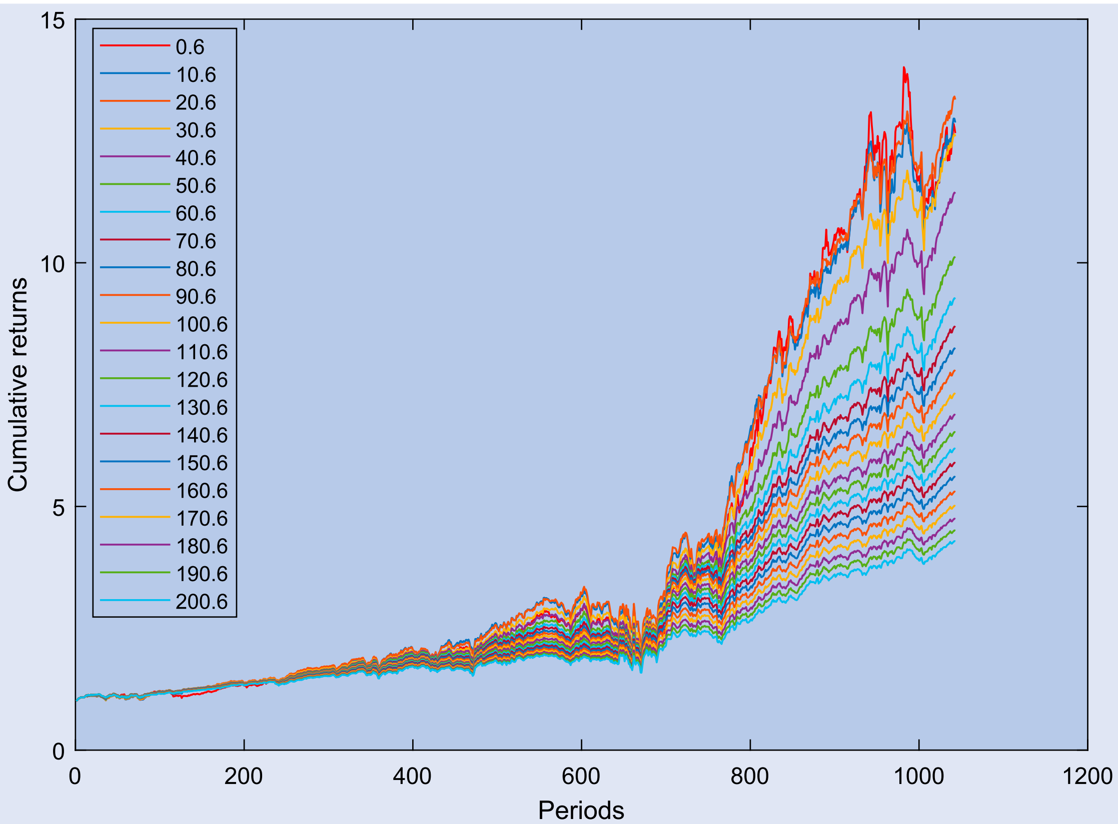

Figure 6. Cumulative returns with different weighting factors.

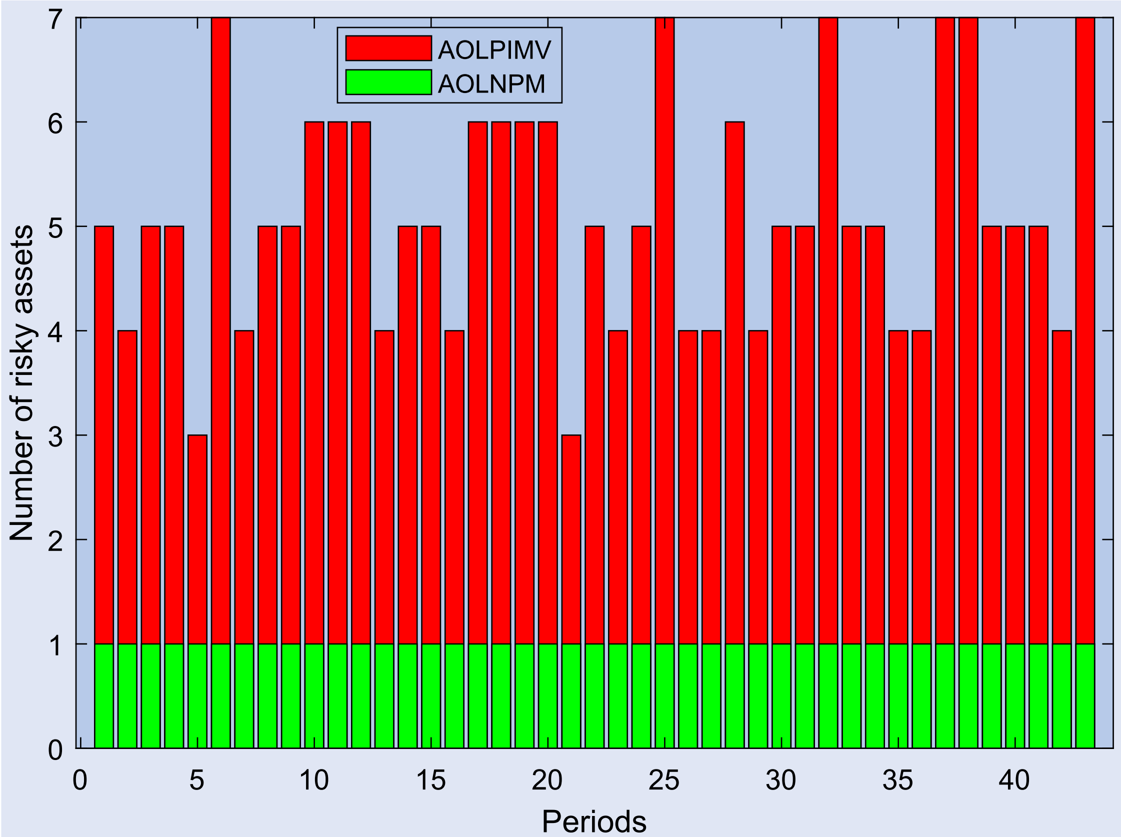

Figure 7. Numbers of assets in the last 43 periods.

whole investment horizon is very long, we focus on the last 43 periods and the statistics of the number of assets which are selected to invest in each period. It is shown in figure 7, at least three different assets are selected for investment in each period (the red bar in figure 7). In periods 6, 25, 32, 37, 38 and 43, the optimal strategy allocates the available capital to seven assets. The average number of assets invested in each period is 5. As a comparison, we employ the AOLNPM algorithm to obtain the optimal investment strategies and find that only one asset is selected in each of the last 43 periods (the green bar in figure 7). This is due to the fact that AOLPIMV algorithm solves a quadratic programming problem where the investment risk term is considered in the objective function while AOLNPM solves a linear programming problem. This reflects that our proposed AOLPIMV algorithm is able to obtain a more distributive investment strategy than AOLNPM, which is consistent with the higher Sharpe ratio, Information ratio and Calmar ratio of AOLPIMV.

4.7. Sensitivity analysis

There are several parameters involved in our proposed AOLPIMV algorithm, mainly including the step size γ and risk weighting factor η. To study the performance of the algorithm in different parameters, we follow the method of Li *et al.* (2018) and observe the cumulative performances of the algorithm under various combinations of γ and η. The value of γ is set to range from 0.00001 to 0.00100. The maximum value is selected as 0.00100 just due to that there are more than 1000 periods for each data set. Theoretically, for the most extreme situation where the step size keeps increasing (decreasing) 0.00100 in every period (It is impossible in real applications), the final step size may be larger than 1 (less than 0) after 1000 periods. The value of η is set to vary from 0.1 to 1 since we have tested the influence of risk constraint in figure 6 by setting different risk weighting factors which are larger than 1.

The heat maps of the cumulative returns under different parameters are shown in figure 8. Firstly, it can be seen that the cumulative returns vary with different parameters. For example, on datasets NYSE-O, NYSE-N and TSE, for a given step size γ, the cumulative return decreases as the risk weighting factor η becomes larger, which is consistent with the market rule 'high return, high risk'. This is reasonable since a larger risk weighting factor represents that the investor concerns more on averting the investment risk instead of taking risk to pursue high profit. For a given risk weighting factor η, the cumulative returns also differ when the step sizes are different. On MSCI and NYSE-O, the best step size lies in the middle region. On NYSE-N and TSE, the best step size lies in the right-hand region. For NASTDA, the best step size lies in the left-hand region. Secondly, for all the combinations of the parameters, the worst cumulative return still dominates most of the other online portfolio selection algorithms. For example, for MSCI, the worst cumulative return is 10.3304 ($NW_T = 0.5755$), which is still larger than that of the other 18 algorithms in table 3. In fact, the parameters we choose in table 3 may not be optimal compared with all the combinations of parameters in figure 8, yet their performance is still satisfactory. This means that there is a wide range of

Table 10. NW_T, AACR, Sharpe ratio, Information ratio and Calmar ratio with different η.

η	NW_T	AACR	Sharpe ratio	Information ratio	Calmar ratio
0.6	0.5819	**1.5727**	0.1115	0.1584	0.1721
10.6	0.5825	1.5315	0.1199	0.1740	0.1872
20.6	**0.5836**	1.5477	0.1286	0.1902	0.2034
30.6	0.5818	1.5194	0.1313	0.1970	0.2085
40.6	0.5787	1.4750	0.1316	**0.1980**	0.2094
50.6	0.5748	1.4272	0.1298	0.1949	0.2070
60.6	0.5720	1.3948	0.1297	0.1953	0.2073
70.6	0.5699	1.3728	0.1305	0.1965	0.2098
80.6	0.5682	1.3563	0.1319	0.1976	0.2132
90.6	0.5664	1.3392	**0.1325**	0.1970	0.2154
100.6	0.5644	1.3208	0.1320	0.1950	**0.2158**
110.6	0.5625	1.3033	0.1311	0.1918	0.2150
120.6	0.5608	1.2886	0.1302	0.1884	0.2143
130.6	0.5590	1.2741	0.1290	0.1846	0.2126
140.6	0.5575	1.2613	0.1277	0.1808	0.2108
150.6	0.5559	1.2484	0.1261	0.1769	0.2082
160.6	0.5541	1.2344	0.1241	0.1724	0.2043
170.6	0.5523	1.2206	0.1219	0.1676	0.1998
180.6	0.5505	1.2079	0.1196	0.1629	0.1951
190.6	0.5488	1.1960	0.1173	0.1583	0.1904
200.6	0.5472	1.1850	0.1150	0.1538	0.1855

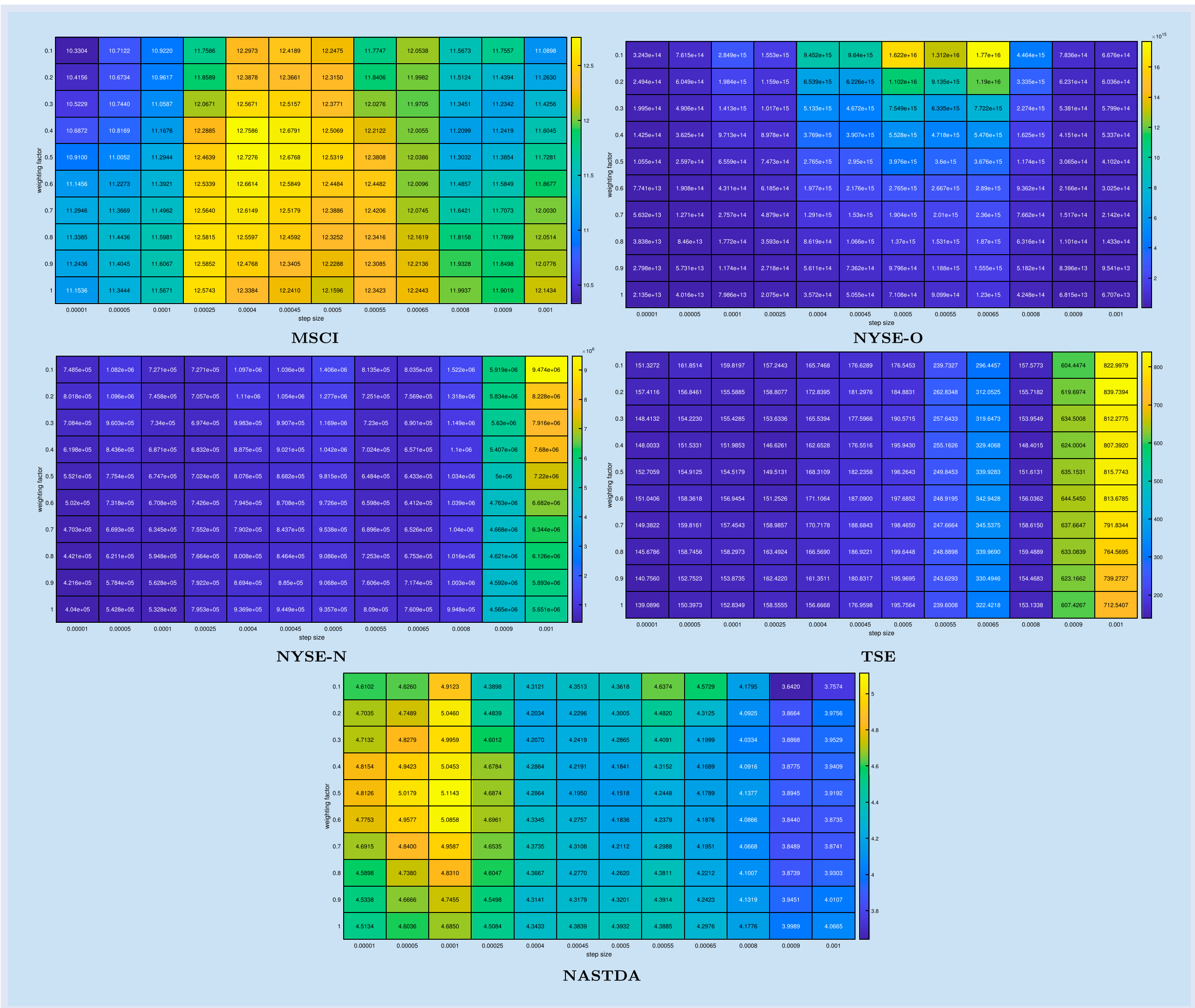

Figure 8. Cumulative returns under different combinations of γ and η. (a) MSCI, (b) NYSE-O, (c) NYSE-N, (d) TSE and (e) NASTDA.

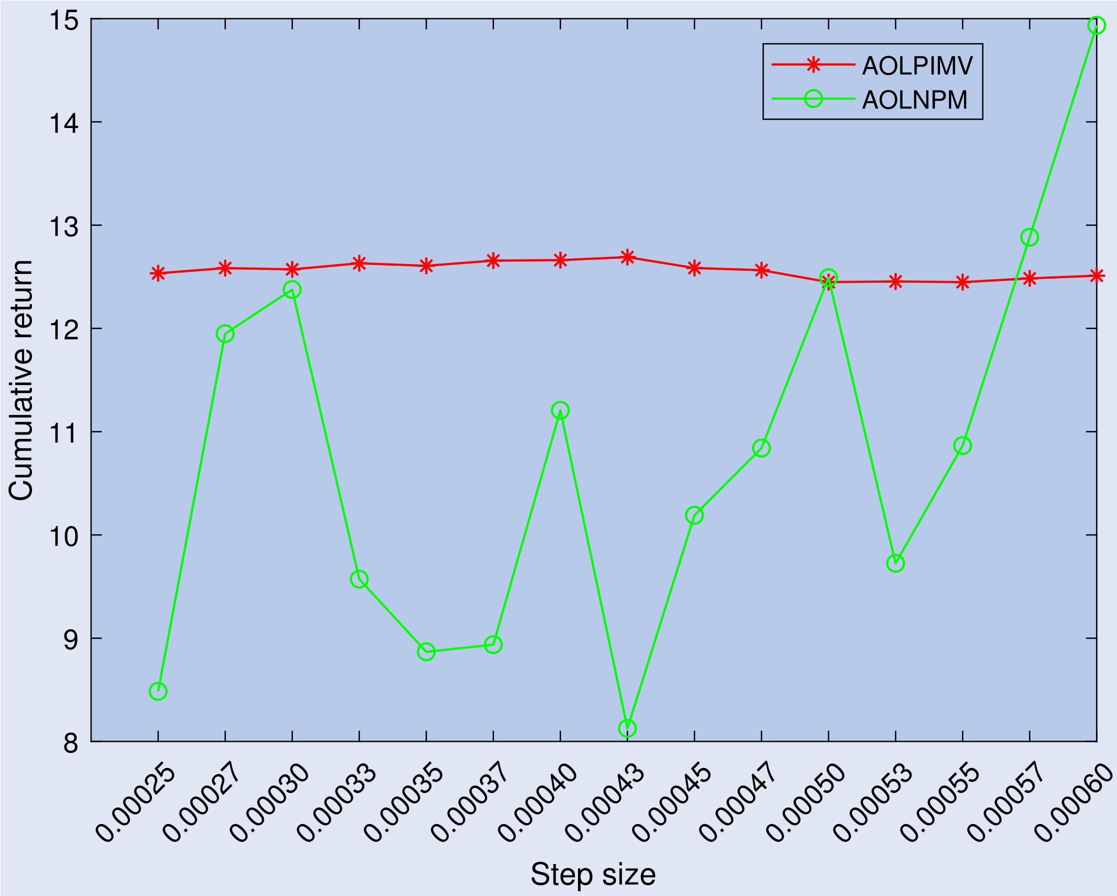

Figure 9. Cumulative returns of AOLPIMV and AOLNPM with different step sizes.

parameters that are able to achieve satisfactory performance, which coincides with the conclusions of Li *et al.* (2018). It should be noted that accurately determining an investor's true risk preference remains a challenge at present. The selection of the best step size and risk weighting factor needs to be studied further in the future.

In addition, the steadiness of the AOLPIMV is much better than AOLNPM. Take MSCI as an example, we set the step size ranging from 0.00025 to 0.00060 and observe the cumulative returns of AOLPIMV and AOLNPM, respectively (see figure 9). It can be seen clearly that the cumulative returns of AOLPIMV under different parameters fluctuate around 12.5000, with a maximum value of 12.6906 and the minimum value of 12.4482. As a comparison, the steadiness of the AOLNPM is much worse, with a maximum value of 14.9357 and the minimum value of 8.1266. This is due to that AOLNPM only considers maximizing the return of each period, which highly depends on the accuracy of return prediction. Although Guo *et al.* (2021) have validated that the predicted return is robust to the step size, the robustness for the cumulative return cannot be guaranteed since a small fluctuation in return prediction may lead to a totally different strategy. AOLPIMV adds the risk constraint to the online portfolio decision-making to achieve a more diversified strategy (see figure 7), which is able to improve its steadiness.

Regarding the parameter window size, denoting the number of available returns for each risky asset at the start of online investment (Guo *et al.* 2021), we vary w between 2 and 9 to evaluate AOLPIMV's performance across these five data sets. The results can be seen in table 11. It can be seen clearly that the performance of AOLPIMV is quite satisfactory under various window sizes. Compared with the results in table 11, the obtained normalized cumulative return in table 3 is not optimal. For example, the best normalized cumulative return for MSCI in table 11 is 0.5846 with $w = 9$ (the corresponding cumulative return $W_T = 13.7846$), which is larger than 0.5819 in table 3 ($W_T = 12.6614$). Similar conclusions can be obtained by observing the values of ACCR. We set $w = 6$ in our experiment just for consistency with the parameter setting in Guo *et al.* (2021). Generally, the performance will be better for a larger window size w. This implies that more available return information before decision-making is beneficial to improving the performance of AOLPIMV.

5. Conclusions

This paper studies the online portfolio selection with transaction costs. To accurately predict the future returns of assets, we propose the AOLPI method which not only considers the historical returns of assets but also the peer impact of other assets. The decaying factors of assets can be adjusted automatically in each period according to the performances of past return predictions. The numerical experiments reveal that our proposed AOLPI method performs better than the AOLMA method. We construct the AMV model where the

Table 11. NW_T and AACR of AOLPIMV under different window sizes w.

Window size w	MSCI NW_T	MSCI AACR	NYSE-O NW_T	NYSE-O AACR	NYSE-N NW_T	NYSE-N AACR	TSE NW_T	TSE AACR	NASTDA NW_T	NASTDA AACR
2	0.5809	1.7662	0.9915	6.9785	0.8847	2.1071	0.6768	4.6221	0.5463	1.4688
3	0.5805	1.7607	0.9921	7.2543	0.8850	2.1644	0.6782	4.6677	0.5484	1.4864
4	0.5825	1.7759	0.9918	7.1338	0.8827	2.0900	0.6770	4.6091	0.5509	1.5155
5	0.5835	1.7843	0.9920	7.2988	0.8922	2.2005	0.6808	4.7996	0.5515	1.5217
6	0.5819	1.7688	**0.9922**	**7.3904**	0.8902	2.2004	0.6817	4.8028	0.5528	1.5412
7	0.5825	1.7757	0.9920	7.2444	0.8910	2.1899	0.6832	4.8455	0.5483	1.4964
8	0.5844	1.7935	0.9919	7.2407	**0.8964**	**2.2439**	0.6816	4.7592	0.5534	1.5459
9	**0.5846**	**1.7956**	0.9920	7.2975	0.8845	2.0846	**0.6892**	**5.1988**	**0.5553**	**1.5780**

investment return and risk are considered simultaneously in the decision-making process. To reduce the computational burden of introducing the risk constraint, we employ the covariance matrix which can be linearly updated when new return data arrives in each period. We integrate the AOLPI and AMV together and propose the AOLPIMV algorithm to solve practical online portfolio selection issues. Finally, numerical experiments are provided to verify the effectiveness of AOLPIMV algorithm. The results reveal that our proposed AOLPIMV algorithm performs better than the other 19 popular online portfolio selection algorithms in cumulative return, Sharpe ratio, Information ratio and Calmar ratio. In addition, we study the impact of risk constraint on the performance of portfolio by setting different weighting factors. It shows that the AOLPIMV algorithm can provide reasonable investment strategies for investors with different risk preferences.

Some further research issues can be studied in the future. Firstly, the current online portfolio selection studies usually assume that the assets can be bought/sold at any time. However, real asset trading is conducted in a limit order book. It is beneficial to incorporate the liquidity of assets, trading volume and bid-ask spreads into the online portfolio selection decision-making. Secondly, more indicators can be considered such as the skewness and kurtosis of the portfolio which are widely applied in traditional portfolio selection studies. Thirdly, currently, no capital is withdrawn or added to the online portfolio selection process. In real applications, investors can adjust the amount of capital according to different market situations. For example, additional capital can be introduced when the financial market is booming, and capital may be withdrawn when the market is in recession. In addition, the stock market may be affected by some external factors such as geopolitical events, and the obtained return data may be incomplete. As an interesting future research issue, one can focus on the online portfolio selection by considering more realistic external factors. Finally, determining optimal step size and risk weighting factor is an important direction for further exploration.

Acknowledgments

The authors would like to thank the editor and the anonymous referees for their helpful suggestions and comments.

Disclosure statement

No potential conflict of interest was reported by the author(s).

Funding

This research work was supported by National Natural Science Foundation of China [grant numbers 72201033 and 72002201], Hong Kong Research Grants Council [grant number 17309522], Seed Funding Research Grant of HKU-TCL Joint Research Centre for Artificial Intelligence, The University of Hong Kong, Introduction Project of China Postdoctoral International Exchange Program [grant number YJ20220045], China Postdoctoral Science Foundation [grant number 2022M720426], Beijing Institute of Technology Research Fund Program for Young Scholars, Guangdong Basic and Applied Basic Research Foundation [grant number 2023A1515030197], and Shenzhen Humanities & Social Sciences Key Research Bases.

References

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

Aït-Sahalia, Y. and Hurd, T.R., Portfolio choice in markets with contagion. *J. Financ. Econ.*, 2016, **14**(1), 1–28.

Azizpour, S., Giesecke, K., Kale, S. and Schwenkler, G., Exploring the sources of default clustering. *J. Financ. Econ.*, 2018, **129**(1), 154–183.

Balter, A.G., Mahayni, A. and Schweizer, N., Time-consistency of optimal investment under smooth ambiguity. *Eur. J. Oper. Res.*, 2021, **293**(2), 643–657.

Best, M.J., *Quadratic Programming with Computer Programs*, 2017 (CRC Press: New York).

Bielecki, T.R., Jin, H., Pliska, S. and Zhou, X., Continuous-time mean-variance portfolio selection with bankruptcy prohibition. *Math. Finance*, 2005, **15**(2), 213–244.

Borodin, A., El-Yaniv, R. and Gogan, V., Can we learn to beat the best stock? *J. Artif. Intell. Res.*, 2004, **21**(1), 579–594.

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

Das, P. and Banerjee, A., Meta optimization and its application to portfolio selection. In *Proceedings of International Conference on Knowledge Discovery and Data Mining*, San Diego, California, USA, pp. 1163–1171, 2011.

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

Gaivoronski, A.A. and Stella, F., On-line portfolio selection using stochastic programming. *J. Econ. Dyn. Control*, 2003, **27**(6), 1013–1043.

González-Díaz, J., González-Rodríguez, B., Leal, M. and Puerto, J., Global optimization for bilevel portfolio design: Economic insights from the Dow Jones index. *Omega*, 2021, **102**, 102353.

Gu, J.W., Si, S.J. and Zheng, H., Constrained utility deviation-risk optimization and time-consistent HJB equation. *SIAM J. Control Optim.*, 2020, **58**, 866–894.

Guan, H. and An, Z.Y., A local adaptive learning system for online portfolio selection. *Knowl. Based. Syst.*, 2019, **186**, 104958.

Guo, S.N., Gu, J.W. and Ching, W.K., Adaptive online portfolio selection with transaction costs. *Eur. J. Oper. Res.*, 2021, **295**(3), 1074–1086.

Guo, S.N., Gu, J.W., Fok, C.H. and Ching, W.K., Online portfolio selection with state-dependent price estimators and transaction costs. *Eur. J. Oper. Res.*, 2023, **311**(1), 333–353.

Györfi, L. and Schäfer, D., Nonparametric prediction. In *Advances in Learning Theory: Methods, Models and Applications*, pp. 339–354, 2003 (IOS Press: Amsterdam).

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

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

Hazan, E. and Seshadhri, C., Efficient learning algorithms for changing environments. In *Proceedings of the International Conference on Machine Learning*, Montreal, Quebec, Canada, pp. 393–400, 2009.

He, J.A. and Peng, F.P., Adaptive moment estimation for universal portfolio selection strategy. *Optim. Eng.,* 2023, **24**(4), 2357–2385.

He, J.A. and Yang, X.Y., Universal portfolio selection strategy by aggregating online expert advice. *Optim. Eng.*, 2022, **23**(1), 85–109. https://doi.org/10.1007/s11081-020-09555-2

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

Honda, T., Optimal portfolio choice for unobservable and regime-switching mean returns. *J. Econ. Dyn. Control*, 2003, **28**, 45–78.

Huang, D., Zhou, J., Li, B., Hoi, S.C. and Zhou, S., Robust median reversion strategy for online portfolio selection. *IEEE. Trans. Knowl. Data. Eng.*, 2016, **28**(9), 2480–2493.

Katsikis, V.N., Mourtas, S.D., Stanimirović, P.S., Li, S. and Cao, X.W., Time-varying mean-variance portfolio selection problem solving via LVI-PDNN. *Comput. Oper. Res.*, 2022, **138**, 105582.

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

Kocuk, B. and Cornuéjols, G., Incorporating Black-Litterman views in portfolio construction when stock returns are a mixture of normals. *Omega*, 2020, **91**, 102008.

Konno, H. and Yamazaki, H., Mean-absolute deviation portfolio optimization model and its applications to Tokyo stock market. *Manage. Sci.*, 1991, **37**(5), 519–531.

Leary, M.T. and Roberts, M.R., Do peer firms affect corporate financial policy? *J. Finance*, 2014, **69**(1), 139–178.

Ledoit , O. and Wolf, M., Robust performance hypothesis testing with the Sharpe ratio. *J. Empir. Finance*, 2008, **15**, 850–859.

Li, B. and Hoi, S.C., On-line portfolio selection with moving average reversion. In *Proceedings of International Conference on Machine Learning*, Edinburgh, Scotland, pp. 563–570, 2012.

Li, B. and Hoi, S.C., *Online Portfolio Selection: Principles and Algorithms*, 2015 (CRC Press: Boca Raton).

Li, D. and Ng, W., Optimal dynamic portfolio selection: Multiperiod mean-variance formulation. *Math. Finance*, 2000, **10**(3), 387–406.

Li, B., Hoi, S.C. and Gopalkrishnan, V., CORN: Correlation-driven nonparametric learning approach for portfolio selection. *ACM. Trans. Intell. Syst. Technol.*, 2011a, **2**(3), 1–29. https://doi.org/10.1145/1961189.1961193

Li, B., Hoi, S.C., Zhao, P. and Gopalkrishnan, V., Confidence weighted mean reversion strategy for on-line portfolio selection. In *Proceedings of the International Conference on Artificial Intelligence and Statistics*, Fort Lauderdale, FL, USA, pp. 434–442, 2011b.

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

Li, B., Hoi, S.C., Zhao, P. and Gopalkrishnan, V., Confidence weighted mean reversion strategy for online portfolio selection. *ACM. Trans. Knowl. Discov. Data.*, 2013, **7**(1), Article No: 4.

Li, B., Wang, J.L., Huang, D.J. and Hoi, S.C., Transaction cost optimization for online portfolio selection. *Quant. Finance*, 2018, **18**(8), 1411–1424.

Li, Z.X., Jiang, H., Chen, Z.P. and Wong, W.K., A mental account-based portfolio selection model with an application for data with smaller dimensions. *Comput. Oper. Res.*, 2022, **144**, 105801.

Liesiö, J., Xu, P. and Kuosmanen, T., Portfolio diversification based on stochastic dominance under incomplete probability information. *Eur. J. Oper. Res.*, 2020, **286**(2), 755–768.

Lim, A.E.B., Quadratic hedging and mean-variance portfolio selection with random parameters in an incomplete market. *Math. Oper. Res.*, 2004, **29**(1), 132–161.

Ling, A.F., Sun, J. and Wang, M.H., Robust multi-period portfolio selection based on downside risk with asymmetrically distributed uncertainty set. *Eur. J. Oper. Res.*, 2020, **285**(1), 81–95.

Maghsoodi, A.I., Cryptocurrency portfolio allocation using a novel hybrid and predictive big data decision support syste. *Omega*, 2023, **115**, 102787.

Markowitz, H., Portfolio selection. *J. Finance*, 1952, **3**(1), 77–91.

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

Sharpe, W.F., Mutual fund performance. *J. Bus.*, 1966, **39**(1), 119–138.

Shen, Y. and Zou, B., Mean-Variance portfolio selection in contagious markets. *SIAM J. Financ. Math.*, 2022, **13**(2), 391–425.

Staino, A. and Russo, E., Nested conditional value-at-risk portfolio selection: A model with temporal dependence driven by market-index volatility. *Eur. J. Oper. Res.*, 2020, **280**(2), 741–753.

Treynor, J.L. and Black, F., How to use security analysis to improve portfolio selection. *J. Bus.*, 1973, **46**(1), 66–86.

Van Staden, P.M., Dang, D.M. and Forsyth, P.A., The surprising robustness of dynamic mean-variance portfolio optimization to model misspecification errors. *Eur. J. Oper. Res.*, 2021, **289**(2), 774–792.

Veldkamp, L.L., Information markets and the comovement of asset prices. *Rev. Econ. Stud.*, 2006, **73**(3), 823–845.

Vovk, V.G. and Watkins, C., Universal portfolio selection. In *Proceedings of the Annual Conference on Learning Theory*, Madison, Wisconsin, USA, pp. 12–23, 1998.

Wang, Y.D., Hao, X.F. and Wu, C.F., Forecasting stock returns: A time-dependent weighted least squares approach. *J. Financ. Mark.*, 2021, **53**, 100568.

Xia, J.M., Mean-variance portfolio choice: Quadratic partial hedging. *Math. Finance*, 2005, **15**, 533–538.

Yang, X.Y., He, J.A. and Zhang, Y., Aggregating exponential gradient expert advice for online portfolio selection. *J. Oper. Res. Soc.*, 2022, **73**(3), 587–597.

Young, T.W., Calmar ratio: A smoother tool. *Futures*, 1991, **20**(1), 40.

Zhou, X. and Li, D., Continuous-time mean-variance portfolio selection: A stochastic LQ framework. *Appl. Math. Optim.*, 2000, **42**, 19–33.

Zhou, Z.B., Gao, M., Xiao, H.L., Wang, R. and Liu, W.B., Big data and portfolio optimization: A novel approach integrating DEA with multiple data sources. *Omega*, 2021, **104**, 102479.

Appendices

This section gives the technical details for obtaining equations (13) and (15), and provides the proofs of theorems 3.1, 3.2 and 3.3, respectively.

Appendix 1. Technical details for obtaining equations (13) and (15)

Due to the variance of asset i at the beginning of period t is calculated by

$$\sigma_{ii}^{(t)} = \frac{1}{t-2}\sum_{k=1}^{t-1}(r_{ki}-\mu_{ti})^2,$$

then the variance of asset i at the beginning of period $t+1$ is given by

$$\begin{aligned}\sigma_{ii}^{(t+1)} &= \frac{\sum_{k=1}^{t-1}(r_{ki}-\mu_{ti}+\mu_{ti}-\mu_{(t+1)i})^2+(r_{ti}-\mu_{(t+1)i})^2}{t-1}\\ &= \frac{\sum_{k=1}^{t-1}[(r_{ki}-\mu_{ti})^2+(\mu_{ti}-\mu_{(t+1)i})^2]+(r_{ti}-\mu_{(t+1)i})^2}{t-1}\\ &= \frac{(t-2)\sigma_{ii}^{(t)}+\frac{t-1}{t^2}(\mu_{ti}-r_{ti})^2+\frac{(t-1)^2}{t^2}(r_{ti}-\mu_{ti})^2}{t-1}\\ &= \frac{t-2}{t-1}\sigma_{ii}^{(t)}+\frac{1}{t}(r_{ti}-\mu_{ti})^2,\end{aligned}$$

which proves equation (13). Similarly, for assets i and j, $\sigma_{ij}^{(t)}$ is given by

$$\sigma_{ij}^{(t)} = \frac{1}{t-2}\sum_{k=1}^{t-1}(r_{ki}-\mu_{ti})(r_{kj}-\mu_{tj}).$$

Then at the beginning of period $t+1$, $\sigma_{ij}^{(t+1)}$ can be calculated by

$$\begin{aligned}\sigma_{ij}^{(t+1)} &= \frac{1}{t-1}\sum_{k=1}^{t}(r_{ki}-\mu_{(t+1)i})(r_{kj}-\mu_{(t+1)j})\\ &= \frac{\begin{array}{c}\sum_{k=1}^{t-1}(r_{ki}-\mu_{ti}+\mu_{ti}-\mu_{(t+1)i})(r_{kj}-\mu_{tj}+\mu_{tj}-\mu_{(t+1)j})\\ +(r_{ti}-\mu_{(t+1)i})(r_{tj}-\mu_{(t+1)j})\end{array}}{t-1}\\ &= \frac{\begin{array}{l}\sum_{k=1}^{t-1}[(r_{ki}-\mu_{ti})(r_{kj}-\mu_{tj})+(\mu_{tj}-\mu_{(t+1)j})\\ (\mu_{ti}-\mu_{(t+1)i})]+(r_{ti}-\mu_{(t+1)i})(r_{tj}-\mu_{(t+1)j})\end{array}}{t-1}\\ &= \frac{\begin{array}{c}(t-2)\sigma_{ij}^{(t)}+\frac{t-1}{t^2}(\mu_{tj}-r_{tj})(\mu_{ti}-r_{ti})\\ +\frac{(t-1)^2}{t^2}(r_{ti}-\mu_{ti})(r_{tj}-\mu_{tj})\end{array}}{t-1}\\ &= \frac{t-2}{t-1}\sigma_{ij}^{(t)}+\frac{1}{t}(r_{ti}-\mu_{ti})(r_{tj}-\mu_{tj}),\end{aligned}$$

which gives rise to equation (15).

Appendix 2. Proof of theorem 3.1

Proof To prove that $\boldsymbol{H}_t$ is semi-positive definite, we show that $\boldsymbol{\Sigma}_t$ is semi-positive definite considering that $\boldsymbol{H}_t = \boldsymbol{N}\boldsymbol{\Sigma}_t\boldsymbol{N}^\top$, where $t \geq 3$ according to equation (14). We apply principle of mathematical induction method to prove that $\boldsymbol{\Sigma}_t$ is a semi-positive definite matrix. For $t=3$, the matrix $\boldsymbol{\Sigma}_3$ is

$$\boldsymbol{\Sigma}_3 = \begin{pmatrix}\sigma_{11}^{(3)} & \sigma_{12}^{(3)} & \sigma_{13}^{(3)} & \cdots & \sigma_{1n}^{(3)}\\ \sigma_{21}^{(3)} & \sigma_{22}^{(3)} & \sigma_{23}^{(3)} & \cdots & \sigma_{2n}^{(3)}\\ \sigma_{31}^{(3)} & \sigma_{32}^{(3)} & \sigma_{33}^{(3)} & \cdots & \sigma_{3n}^{(3)}\\ \vdots & \vdots & \vdots & \ddots & \vdots\\ \sigma_{n1}^{(3)} & \sigma_{n2}^{(3)} & \sigma_{n3}^{(3)} & \cdots & \sigma_{nn}^{(3)}\end{pmatrix},$$

where each

$$\sigma_{ij}^{(3)} = (r_{1i}-\mu_{3i})(r_{1j}-\mu_{3j})+(r_{2i}-\mu_{3i})(r_{2j}-\mu_{3j}).$$

Set $\sigma_{ij}^{(3,1)} = (r_{1i}-\mu_{3i})(r_{1j}-\mu_{3j})$ and $\sigma_{ij}^{(3,2)} = (r_{2i}-\mu_{3i})(r_{2j}-\mu_{3j})$. Then $\boldsymbol{\Sigma}_3$ can be expressed as

$$\boldsymbol{\Sigma}_3 = \boldsymbol{\Sigma}_{3,1}+\boldsymbol{\Sigma}_{3,2},$$

where

$$\boldsymbol{\Sigma}_{3,1} = \begin{pmatrix}\sigma_{11}^{(3,1)} & \sigma_{12}^{(3,1)} & \sigma_{13}^{(3,1)} & \cdots & \sigma_{1n}^{(3,1)}\\ \sigma_{21}^{(3,1)} & \sigma_{22}^{(3,1)} & \sigma_{23}^{(3,1)} & \cdots & \sigma_{2n}^{(3,1)}\\ \sigma_{31}^{(3,1)} & \sigma_{32}^{(3,1)} & \sigma_{33}^{(3,1)} & \cdots & \sigma_{3n}^{(3,1)}\\ \vdots & \vdots & \vdots & \ddots & \vdots\\ \sigma_{n1}^{(3,1)} & \sigma_{n2}^{(3,1)} & \sigma_{n3}^{(3,1)} & \cdots & \sigma_{nn}^{(3,1)}\end{pmatrix},$$

and

$$\boldsymbol{\Sigma}_{3,2} = \begin{pmatrix}\sigma_{11}^{(3,2)} & \sigma_{12}^{(3,2)} & \sigma_{13}^{(3,2)} & \cdots & \sigma_{1n}^{(3,2)}\\ \sigma_{21}^{(3,2)} & \sigma_{22}^{(3,2)} & \sigma_{23}^{(3,2)} & \cdots & \sigma_{2n}^{(3,2)}\\ \sigma_{31}^{(3,2)} & \sigma_{3,2}^{(32)} & \sigma_{3,3}^{(32)} & \cdots & \sigma_{3n}^{(3,2)}\\ \vdots & \vdots & \vdots & \ddots & \vdots\\ \sigma_{n1}^{(3,2)} & \sigma_{n2}^{(3,2)} & \sigma_{n3}^{(3,2)} & \cdots & \sigma_{nn}^{(3,2)}\end{pmatrix}.$$

For any nonzero vector $\mathbf{x} = (x_1, x_2, \ldots, x_n)$, we have

$$\mathbf{x}\boldsymbol{\Sigma}_3\mathbf{x}^\top = \mathbf{x}\boldsymbol{\Sigma}_{3,1}\mathbf{x}^\top+\mathbf{x}\boldsymbol{\Sigma}_{3,2}\mathbf{x}^\top = \left(\sum_{i=1}^{n}(r_{1i}-\mu_{3i})x_i\right)^2 + \left(\sum_{i=1}^{n}(r_{2i}-\mu_{3i})x_i\right)^2 \geq 0.$$

Then $\boldsymbol{\Sigma}_3$ is a semi-positive definite matrix. Suppose that $\boldsymbol{\Sigma}_t$ is semi-positive definite, which means that $\mathbf{x}\boldsymbol{\Sigma}_t\mathbf{x}^\top \geq 0$ for any nonzero vector $\mathbf{x}$. For $\boldsymbol{\Sigma}_{t+1}$, we have

$$\mathbf{x}\boldsymbol{\Sigma}_{t+1}\mathbf{x}^\top = \left(\frac{t-2}{t-1}\right)\mathbf{x}\boldsymbol{\Sigma}_t\mathbf{x}^\top+\frac{1}{t}\mathbf{x}\boldsymbol{M}_t\mathbf{x}^\top$$

where

$$\boldsymbol{M}_t = \begin{pmatrix}(r_{t1}-\mu_{t1})^2 & (r_{t1}-\mu_{t1})(r_{t2}-\mu_{t2}) & \cdots & (r_{t1}-\mu_{t1})(r_{tn}-\mu_{tn})\\ (r_{t2}-\mu_{t2})(r_{t1}-\mu_{t1}) & (r_{t2}-\mu_{t2})^2 & \cdots & (r_{t2}-\mu_{t2})(r_{tn}-\mu_{tn})\\ (r_{t3}-\mu_{t3})(r_{t1}-\mu_{t1}) & (r_{t3}-\mu_{t3})(r_{t2}-\mu_{t2}) & \cdots & (r_{t3}-\mu_{t3})(r_{tn}-\mu_{tn})\\ \vdots & \vdots & \ddots & \vdots\\ (r_{tn}-\mu_{tn})(r_{t1}-\mu_{t1}) & (r_{tn}-\mu_{tn})(r_{t2}-\mu_{t2}) & \cdots & (r_{tn}-\mu_{tn})^2\end{pmatrix}.$$

Since

$$\mathbf{x}\boldsymbol{M}_t\mathbf{x}^\top = \left(\sum_{i=1}^{n}(r_{ti}-\mu_{ti})x_i\right)^2 \geq 0$$

then we have $\mathbf{x}\boldsymbol{\Sigma}_{t+1}\mathbf{x}^\top \geq 0$. Therefore, $\boldsymbol{\Sigma}_t$ is a semi-positive definite matrix. Then for any nonzero vector $(\mathbf{y}, \mathbf{z})$, we have

$$(\mathbf{y},\mathbf{z})\boldsymbol{H}_t(\mathbf{y},\mathbf{z})^\top = ((\mathbf{y},\mathbf{z})N)\boldsymbol{\Sigma}_t((\mathbf{y},\mathbf{z})N)^\top \geq 0.$$

Therefore, the matrix $\boldsymbol{H}_t$ is semi-positive definite for $t = 3, 4, \ldots, m$. The proof is then completed. ■

Appendix 3. Proof of theorem 3.2

Proof The objective function of model (20) can be transformed into a minimization programming which can be expressed as follows:

$$\min \quad \eta(\mathbf{y}_t,\mathbf{z}_t)\boldsymbol{H}_t(\mathbf{y}_t,\mathbf{z}_t)^\top - \boldsymbol{F}_t(\mathbf{y}_t,\mathbf{z}_t)^\top - \boldsymbol{C}_t.$$

According to theorem 3.1, for any feasible solution $(\mathbf{y}_t,\mathbf{z}_t)$, $\eta(\mathbf{y}_t,\mathbf{z}_t)\boldsymbol{H}_t(\mathbf{y}_t,\mathbf{z}_t)^\top \geq 0$ since $\eta > 0$ and $\boldsymbol{H}_t$ is semi-positive definite. The second term $-\boldsymbol{F}_t(\mathbf{y}_t,\mathbf{z}_t)^\top$ is equal to

$$-(\hat{\mathbf{r}}_t - 2\eta\tilde{\mathbf{x}}_{t-1}\boldsymbol{\Sigma}_t)[(\mathbf{y}_t,\mathbf{z}_t)N]^\top + \delta\mathbf{e}_{2n}(\mathbf{y}_t,\mathbf{z}_t)^\top.$$

According to the second constraint of model (20), $-\tilde{\mathbf{x}}_{t-1} \leq (\mathbf{y}_t,\mathbf{z}_t)\boldsymbol{N} \leq \mathbf{e}_n - \tilde{\mathbf{x}}_{t-1}$, representing that $(\mathbf{y}_t,\mathbf{z}_t)\boldsymbol{N}$ lies in a n-dimensional bounded convex polytope. Then the minimal value of $-(\hat{\mathbf{r}}_t - 2\eta\tilde{\mathbf{x}}_{t-1}\boldsymbol{\Sigma}_t)[(\mathbf{y}_t,\mathbf{z}_t)\boldsymbol{N}]^\top$ can be achieved when $(\mathbf{y}_t,\mathbf{z}_t)\boldsymbol{N}$ lies in one of the vertices of the polytope. Denote S the set of vertices, $S = \{\boldsymbol{\vartheta}_1,\boldsymbol{\vartheta}_2,\ldots,\boldsymbol{\vartheta}_\varsigma\}$ where $\varsigma = |S|$. Then $-(\hat{\mathbf{r}}_t - 2\eta\tilde{\mathbf{x}}_{t-1}\boldsymbol{\Sigma}_t)[(\mathbf{y}_t,\mathbf{z}_t)\boldsymbol{N}]^\top$ is bounded below by

$$\min_{\boldsymbol{\vartheta}_i\in S}\{-(\hat{\mathbf{r}}_t - 2\eta\tilde{\mathbf{x}}_{t-1}\boldsymbol{\Sigma}_t)\boldsymbol{\vartheta}_i^\top\}.$$

Owing to $\mathbf{0} \leq \mathbf{y}_t, \mathbf{0} \leq \mathbf{z}_t$, we have $\delta\mathbf{e}_{2n}(\mathbf{y}_t,\mathbf{z}_t)^\top \geq 0$. Therefore, for any feasible solution $(\mathbf{y}_t - \mathbf{z}_t)$, the objective function

$$\eta(\mathbf{y}_t,\mathbf{z}_t)\boldsymbol{H}_t(\mathbf{y}_t,\mathbf{z}_t)^\top - \boldsymbol{F}_t(\mathbf{y}_t,\mathbf{z}_t)^\top - \boldsymbol{C}_t$$

is bounded below by

$$\min_{\boldsymbol{\vartheta}_i\in S}\{-(\hat{\mathbf{r}}_t - 2\eta\tilde{\mathbf{x}}_{t-1}\boldsymbol{\Sigma}_t)\boldsymbol{\vartheta}_i^\top\} - \boldsymbol{C}_t.$$

It is proved in Best (2017) that there is at least one optimal solution if the feasible region of the quadratic programming is nonempty and the objective value is bounded below. The proof is then completed. ∎

Appendix 4. Proof of theorem 3.3

Proof Firstly, we show that model (20) is not equivalent to model (17). It is shown in the above discussion, for any feasible solution $\mathbf{x}_t$ of model (17), we can always find the corresponding solution $\mathbf{y}_t$ and $\mathbf{z}_t$ satisfying the constraints of model (20), where

$$|x_{ti} - \tilde{x}_{(t-1)i}| = y_{ti} + z_{ti}, \quad x_{ti} - \tilde{x}_{(t-1)i} = y_{ti} - z_{ti}, \quad i = 1,2,\ldots,n. \tag{A1}$$

However, for any feasible $\mathbf{y}_t$ and $\mathbf{z}_t$ of model (20), we cannot guarantee a feasible solution $\mathbf{x}_t$ satisfying the constraints of model (17). For each $i = 1,2,\ldots,n$, there are four cases for the values of y_{ti} and z_{ti}: (i): $y_{ti} = 0$ and $z_{ti} = 0$; (ii): $y_{ti} = 0$ and $z_{ti} > 0$; (iii): $y_{ti} > 0$ and $z_{ti} = 0$; (iv): $y_{ti} > 0$ and $z_{ti} > 0$. For any y_{ti} and z_{ti} in Cases (i), (ii) and (iii), we can always find a x_{ti} by solving equation (19). For case (iv), there is no x_{ti} satisfying equation (19). Therefore, model (20) is not equivalent to model (17).

For cases (i), (ii) and (iii), it is clear that there is a one-to-one correspondence between the feasible solutions of models (17) and (20). To prove that model (20) achieves the optimal solution if and only if model (17) achieves the optimal solution, it suffices to show that (y^*_{ti}, z^*_{ti}) does not belong to case (iv) for each $i = 1,2,\ldots,n$. Without loss of generality, we assume that there is (y^*_{tk}, z^*_{tk}) belonging to case (iv), $k = 1,2,\ldots,n$, and satisfying $y^*_{tk} > z^*_{tk} > 0$. According to equation (20), the corresponding objective value of $(\mathbf{y}^*_t, \mathbf{z}^*_t)$ is

$$\begin{aligned}
&\sum_{i=1}^n \hat{r}_{ti}\tilde{x}_{(t-1)i} + \sum_{i=1}^n \tilde{x}_{(t-1)i}(y^*_{ti} - z^*_{ti}) \\
&\quad - \eta\sum_{i=1}^n\sum_{j=1}^n \sigma^{(t)}_{ij}(\tilde{x}_{(t-1)i} + y^*_{ti} - z^*_{ti})(\tilde{x}_{(t-1)j} + y^*_{tj} - z^*_{tj}) \\
&\quad - \sum_{i=1}^n \delta(y^*_{ti} + z^*_{ti}).
\end{aligned}$$

Then we consider another solution $(\mathbf{y}^\star_t, \mathbf{z}^\star_t) = (\mathbf{y}^*_t - z^*_{tk}\boldsymbol{\psi}_k, \mathbf{z}^*_t - z^*_{tk}\boldsymbol{\psi}_k)$ where $\boldsymbol{\psi}_k$ is the basis vector of which the k-th element is 1 and all the other elements are zero. Then the corresponding objective value is

$$\begin{aligned}
&\sum_{i=1}^n \hat{r}_{ti}\tilde{x}_{(t-1)i} + \sum_{i=1}^n \tilde{x}_{(t-1)i}(y^*_{ti} - z^*_{ti}) \\
&\quad - \eta\sum_{i=1}^n\sum_{j=1}^n \sigma^{(t)}_{ij}(\tilde{x}_{(t-1)i} + y^*_{ti} - z^*_{ti})(\tilde{x}_{(t-1)j} + y^*_{tj} - z^*_{tj}) \\
&\quad - \sum_{i\neq k} \delta(y^*_{ti} + z^*_{ti}) - \delta(y^*_{tk} - z^*_{tk}),
\end{aligned}$$

which is larger than the function value of $(\mathbf{y}^*_t, \mathbf{z}^*_t)$ since $\delta > 0$. This contradicts to the assumption that $(\mathbf{y}^*_t, \mathbf{z}^*_t)$ is the optimal solution of model (20). Then all the (y^*_{ti}, z^*_{ti}) in $(\mathbf{y}^*_t, \mathbf{z}^*_t)$ cannot belong to case (iv). Model (20) achieves the optimal solution $\mathbf{y}^*_t = (y^*_{t1}, y^*_{t2}, \ldots, y^*_{tn})$ and $\mathbf{z}^*_t = (z^*_{t1}, z^*_{t2}, \ldots, z^*_{tn})$ if and only if model (17) achieves the optimal solution $\mathbf{x}^*_t = (x^*_{t1}, x^*_{t2}, \ldots, x^*_{tn})$. The proof is then completed. ∎