Cardinality-constrained Portfolio Selection via Two-timescale Duplex Neurodynamic Optimization

Man-Fai Leung[a], Jun Wang[b,*], Hangjun Che[c]

[a]School of Computing and Information Science, Faculty of Science and Engineering, Anglia Ruskin University, Cambridge, U.K.
[b]Department of Computer Science and School of Data Science, City University of Hong Kong, Kowloon, Hong Kong
[c]Chongqing Key Laboratory of Nonlinear Circuits and Intelligent Information Processing, College of Electronic and Information Engineering, Southwest University, Chongqing, China

Abstract

This paper addresses portfolio selection based on neurodynamic optimization. The portfolio selection problem is formulated as a biconvex optimization problem with a variable weight in the Markowitz risk-return framework. In addition, the cardinality-constrained portfolio selection problem is formulated as a mixed-integer optimization problem and reformulated as a biconvex optimization problem. A two-timescale duplex neurodynamic approach is customized and applied for solving the reformulated portfolio optimization problem. In the two-timescale duplex neurodynamic approach, two recurrent neural networks operating at two timescales are employed for local searches, and their neuronal states are reinitialized upon local convergence using a particle swarm optimization rule to escape from local optima toward global ones. Experimental results on four datasets of world stock markets are elaborated to demonstrate the superior performance of the neurodynamic optimization approach to three baselines in terms of two major risk-adjusted performance criteria and portfolio returns.

Keywords: Neurodynamic optimization, mean-variance portfolio selection, cardinality constraints

1. Introduction

The modern portfolio theory began with the groundbreaking work of Nobel laureate Markowitz on the mean-variance analysis (Markowitz (1952)). It is based on i) the quantification of the risk of a portfolio using statistical measures; ii) the diversification of assets to be invested for reducing the portfolio risk, and iii) the optimization of trade-offs between risk and return (Kolm et al. (2014)). As a major task in investment management, portfolio selection is to decide the proportions of invested stocks and bonds for asset allocation.

In the classic work, portfolio selection is handled by treating one of the objectives as a constraint (Markowitz (1952)) or combining both objectives into one (e.g., Sharpe ratio (Sharpe (1964))). The former strategy is to achieve the highest expected return subject to a given level of risk or to attain the lowest risk subject to a given level of expected return. The latter strategy is to optimize a scalarized objective function for simultaneous return maximization and risk minimization. In view of the two objectives, portfolio selection is also made by optimizing the risk and return to obtain a set of Pareto-optimal solutions (Ponsich et al. (2013)). A natural way is to optimize both objectives explicitly via scalarization (Ponsich et al. (2013)) or maximization of utility functions (Kroll et al. (1984), Sharpe (2007)) (e.g., the von Neumann-Morgenstern utility function (Morgenstern & Von Neumann (1953)) to characterize a set of Pareto-optimal solutions for decision makers to choose. Both methods have their limitations: A set of predefined weights is needed for scalarization, and the distribution of the resulting Pareto-optimal solutions depends on the weights (Steuer (1986)). Investors' prior preference information is required for maximizing utility functions (Kroll et al. (1984)). These issues are tackled in many studies (Ponsich et al. (2013), Kolm et al. (2014), Mansini et al. (2014), Zopounidis et al. (2015), Ertenlice & Kalayci (2018)).

With the recent advances in artificial intelligence, it is highly desirable or advantageous to develop computationally intelligent approaches to portfolio optimization. Specifically, neurodynamic optimization approaches based on recurrent neural networks (RNNs) are competent for portfolio selection due to the nature of parallel and distribution in information processing. As the counterparts of

*This work is supported in part by the Research Grants Council of the Hong Kong Special Administrative Region of China (Grant 11202019); in part by the Laboratory for AI-Powered Financial Technologies in Hong Kong; in part by the Fundamental Research Funds for the Central Universities (Grant no. SWU020006), in part by the National Natural Science Foundation of China(Grant No. 62003281); and in part by the Natural Science Foundation of Chongqing, China (Grant no. cstc2021jcyj-msxmX1169).

*Corresponding author

Email addresses: man-fai.leung@aru.ac.uk (Man-Fai Leung), jwang.cs@cityu.edu.hk (Jun Wang), hjche123@swu.edu.cn (Hangjun Che)

biological neural systems, neurodynamic approaches can function as computational models for solving various optimization problems in parallel (Hopfield & Tank (1986), Tank & Hopfield (1986)). Since the pioneering work by Hopfield and Tanks (Hopfield & Tank (1986), Tank & Hopfield (1986)), numerous globally convergent neurodynamic approaches are developed for solving various optimization problems, such as convex and pseudoconvex optimization problems with real-valued and complex-valued variables (Wang (1994), Xia et al. (2008), Guo et al. (2011), Liu & Wang (2011), Liu et al. (2012), Liu & Wang (2013), Zhang et al. (2015), Hosseini (2016), Bian et al. (2018), Liu & Qin (2019), Liu et al. (2020b,a), Xu et al. (2020), Wen et al. (2021), Liu et al. (2022), Zhao et al. (2022)), distributed optimization problems (Liu et al. (2017), Xia et al. (2021), Jiang et al. (2022), Xia et al. (2022)), multiple-objective optimization problems (Leung & Wang, 2018, Yang et al., 2018), and global and combinatorial optimization problems (Yan et al., 2014, 2017, Che & Wang, 2019a,b). In particular, a collaborative neurodynamic optimization (CNO) approach is developed for robust portfolio selection based on a minimax and bi-objective problem formulation (Leung & Wang (2021)), where multiple neural networks are employed to characterize the Pareto front. Recently, cardinality-constrained portfolio selection is reformulated as a mixed-integer optimization problem and a CNO-based approach is developed for solving it (Leung & Wang (2022)). By using a population of RNNs to search Pareto-optimal solutions by optimizing a weighted objective function and a meta-heuristic rule to optimize the weight, the CNO-based approach is able to generate very good Pareto-optimal solutions.

Based on our previous works on neurodynamic optimization, this paper presents a timescale duplex neurodynamic approach to portfolio optimization. The Markowitz mean-variance portfolio selection problem is reformulated as a biconvex optimization problem with conditional value at risk. In the proposed method, two RNNs timescales are employed operating at two to search for optimal solutions and a meta-heuristic is used to reinitialize neuronal states to escape local minima. The novelties and contributions of this work are summarized as follows.

i. The reformulated portfolio selection problem with cardinality constraints enables to optimize the conditional Sharpe ratio while selecting a subset of stocks with a given cardinality.

ii. The customized duplex neurodynamic system consists of two RNNs only with significantly reduced spatial complexity compared to the existing CNO approach.

iii. Experimental results on four datasets show that the neurodynamic approach outperforms three baselines in terms of Sharpe ratio, conditional Sharpe ratio, cumulative return, and annualized return.

The remainder of this paper is organized as follows: Section 2 introduces the preliminaries on portfolio optimization, two existing neurodynamic models, and collaborative neurodynamic optimization. Section 3 describes the problem reformulations of the portfolio optimization with and without cardinality constraints. Section 4 delineates the two-timescale duplex neurodynamic method for cardinality-constrained portfolio selection. Section 5 elaborates the experimental results. Section 6 concludes the paper.

2. Preliminaries

2.1. Biconvex optimization

The following definitions are some basic concepts of biconvex optimization.

Definition 1 (Gorski et al. (2007)): The set $\mathscr{Z} \subset \mathscr{X} \times \mathscr{Y}$ is called a biconvex set on $\mathscr{X} \times \mathscr{Y}$ if $\mathscr{Z}_x$ is convex for every $x \in \mathscr{X}$ and $\mathscr{Z}_y$ is convex for every $y \in \mathscr{Y}$, where $\mathscr{X} \subseteq \Re^m$ and $\mathscr{Y} \subseteq \Re^n$ are two nonempty convex sets, $\mathscr{Z}_x$ and $\mathscr{Z}_y$ are two sections of $\mathscr{Z}$ defined as follows:

$$\mathscr{Z}_x = \{(x, y) \in \mathscr{Z} | y \in \mathscr{Y}\},\ \mathscr{Z}_y = \{(x, y) \in \mathscr{Z} | x \in \mathscr{X}\}.$$

Definition 2 (Gorski et al. (2007)): A function $f(x, y) : \mathscr{Z} \to \Re$ is called a biconvex function on $\mathscr{Z} \subseteq \mathscr{X} \times \mathscr{Y}$ if $f(x, \cdot) : \mathscr{Z}_x \to \Re$ is a convex function on $\mathscr{Z}_x$ for every fixed $x \in \mathscr{X}$ and $f(\cdot, y) : \mathscr{Z}_y \to \Re$ is a convex function on $\mathscr{Z}_y$ for every fixed $y \in \mathscr{Y}$.

Definition 3 (Gorski et al. (2007)): A biconvex optimization problem is defined as follows:

$$\min_{x \in \mathscr{X}, y \in \mathscr{Y}} \ f(x, y) \tag{1}$$

where $f(x, y)$ is biconvex with respect to x and y on $\mathscr{X} \times \mathscr{Y}$.

2.2. Mean-variance portfolio selection

The mean-variance (MV) framework suggests that investors should quantify the risk and return of an asset and then allocate funds to the assets based on the risk-return trade-off. The proportion of each asset invested among the set constitutes a portfolio. For simplicity, it is assumed that no short-selling is allowed in this paper. Let $y \in \mathbb{Y} = [0, 1]^n$ be the proportions of wealth invested to n assets, $\mu \in \Re^n$ be the mean returns, and V be the covariance matrix. $\mu^T y$ and $y^T V y$ are the expected return and variance of the portfolio, respectively. The mean-variance portfolio selection can be reformulated as follows:

The MV framework aims to minimize risk or maximize the return of a portfolio (Markowitz (1952)):

$$\begin{aligned} \min_{y}\ & y^T V y \\ \text{s.t.}\ & \mu^T y \geq \mu_{\min}, \\ & e^T y = 1, \\ & y \geq 0; \end{aligned} \tag{2}$$

or

$$\begin{aligned}\max_{y}\ & \mu^T y\\ \text{s.t. } & y^T V y \leq \sigma_{\max},\\ & e^T y = 1,\\ & y \geq 0;\end{aligned} \tag{3}$$

where $\mu_{\min}$ is the minimum allowable portfolio return in problem (2), $\sigma_{\max}$ is the maximum allowable variance in problem (3), e is the vector of ones, and $e^T y = 1$ is the budget constraint. However, it is pointed out that problems (2) and (3) are sensitive to estimation errors. To overcome such a limitation, a robust approach within the mean-variance frameworksuch as minimax portfolio selection can be adopted (Young (1998), Polak et al. (2010)). The approach aims to maximize the worst expected returns of portfolios (Deng et al. (2005), Leung & Wang (2021)):

$$\begin{aligned}\min_{x}\max_{y}\ & (1-\beta)x^T y - \beta y^T V y\\ \text{s.t. } & e^T y = 1,\\ & y \geq 0,\end{aligned} \tag{4}$$

where $\beta \in (0,1)$ is the risk-aversion parameter, $x \in \mathbb{X} = [\underline{x}, \bar{x}]^n$ is the expected rate of returns of n assets, $\underline{x}$ and $\bar{x}$ are respectively the lower and upper bound vectors of x obtained from historical data (Deng et al. (2005)). The smaller value of β is, the higher the resulting investment risk is. The larger the value of β is, the more conservative the portfolio is. The minimax problem formulation in (4) results in robust portfolios most suitable for short-term investment in turbulent markets. Nevertheless, it usually tends to be conservative and results in underperforming portfolio returns for long-term investments in efficient markets.

2.3. Cardinality-constrained portfolio selection

In the MV framework, such as (2) or (3), a constructed portfolio is supposed to be selected from all available assets in a frictionless market. Due to various forms of market friction, investors tend to invest a limited number of assets. In particular, cardinality constraints are widely adopted in portfolio selection due to various needs, such as the reduction of transaction costs and the increase in execution efficiency (Ruiz-Torrubiano & Suárez (2010)). As a result, a limited number of risky securities are selected to construct a portfolio, which leads to the introduction of cardinality constraints, and the complexity of the portfolio selection problem increases significantly (Chang et al. (2000)). The optimization problem (2) with cardinality constraints is formulated as follows:

$$\begin{aligned}\min_{y}\ & y^T V y\\ \text{s.t. } & \mu^T y \geq \mu_{\min},\\ & e^T y = 1,\\ & ||y||_0 \leq k,\\ & y \geq 0,\end{aligned} \tag{5}$$

where $||y||_0 \leq k$ is the cardinality constraint. However, the inclusion of the cardinality constraint in the problem formulation leads to global or mixed-integer optimization problems (Woodside-Oriakhi et al. (2011), Gao & Li (2013), Hardoroudi et al. (2017), Kalayci et al. (2020)).

2.4. Conditional Value-at-Risk

As one of the objective functions in the Markowitz mean-variance framework, variance cannot fairly characterize market volatility (Ang & Chen (2002)). A popular alternative risk measure is value-at-risk (Morgan (1994)). Let $\xi \in \Re^n$ be random returns, VaR is defined as:

$$\mathrm{VaR}_\theta(y) = \min\{\rho \in \Re : \mathbb{P}(-\xi^T y \leq \rho) \geq \theta\},$$

where $0 < \theta < 1$ (Rockafellar & Uryasev (2000)). It should be noted that $\mathrm{VaR}_\theta(y)$ is nonconvex with respect to y (Artzner et al. (1999)). Based on VaR, conditional value-at-risk (CVaR) is defined as the expectation of the upper bound of VaR (Rockafellar & Uryasev (2000)):

$$\mathrm{CVaR}_\theta(y) = \mathbb{E}\{-\xi^T y | -\xi^T y \geq \mathrm{VaR}_\theta(y)\} \tag{6}$$

where $\mathbb{E}(\cdot)$ is the expectation operator.

Parametric and sampling methods are two major approaches to calculating CVaR in (6) (Gaivoronski & Pflug (2005)). If the distribution of asset returns is known, the parametric approach can be used. On the other hand, the sampling approach computes CVaR based on actual historical data. Let N return observations be $\xi_1, \xi_2, \ldots, \xi_N$. The CVaR risk measure is approximated as follows (Rockafellar & Uryasev (2000)):

$$\mathrm{CVaR}_\theta(y) \approx \rho + \frac{1}{N(1-\theta)}\sum_{j=1}^{N}\max(0, -\xi_j^T y - \rho).$$

Based on the sampling approximation of CVaR, a mean-CVaR bicriteria portfolio optimization problem is formulated as follows:

$$\begin{aligned}&\min_{y}\ -\mu^T y\\ &\min_{\sigma,\rho}\ \rho + \frac{1}{N(1-\theta)}\sum_{j=1}^{N}\sigma_j\\ &\text{s.t. } \sigma_j \geq -\xi_j^T y - \rho,\ \sigma_j \geq 0,\ j = 1, 2, \ldots, N;\\ &\quad e^T y = 1;\\ &\quad y \geq 0;\end{aligned} \tag{7}$$

where $\sigma_j = \max(0, -\xi_j^T y - \rho)$ for all j.

2.5. Sharpe Ratio and conditional Sharpe Ratio

Sharpe ratio (SR), proposed by Nobel laureate William Sharpe, is a well-known risk-adjusted performance criterion for evaluating portfolios (Sharpe (1994)). The ratio standardizes the excess return of a portfolio over the risk-free rate by the standard deviation (Christiansen et al.

(2007)). It is also used as an objective function for portfolio optimization (e.g., Liu et al. (2012, 2013)) as follows:

$$\begin{aligned} \max_y \ & \frac{\mu^T y - r_f}{\sqrt{y^T V y}} \\ \text{s.t. } & e^T y = 1, \\ & y \geq 0, \end{aligned} \tag{8}$$

where r_f is the risk-free rate of return.

In analogy to (8), the conditional Sharpe ratio (CSR) is defined by replacing variance with CVaR (Eling & Schuhmacher (2007)) and used for portfolio optimization:

$$\begin{aligned} \max_y \ & \frac{\mu^T y - r_f}{\text{CVaR}_\theta(y)}, \\ \text{s.t. } & e^T y = 1, \\ & y \geq 0. \end{aligned} \tag{9}$$

2.6. Selected neurodynamic models

Consider the following constrained optimization problem:

$$\begin{aligned} \min_{y \in \mathscr{Y}} \ & \psi(y) \\ \text{s.t. } & g(y) \leq 0 \end{aligned} \tag{10}$$

where $\psi : \Re^n \to \Re$, $g : \Re^n \to \Re^m$ denotes the i-th inequality constraint with $g_i(y)(i = 1, \ldots, m)$, and both $\psi(y)$ and $g(y)$ are assumed to be twice differentiable.

The Lagrangian function for optimization problem (10) is

$$L(y) = \psi(y) + \alpha^T g(y) \tag{11}$$

where $\alpha \in \Re^m$ is the Lagrangian multiplier. Based on (11), a neurodynamic model for solving (10) is described as follows (Xia et al. (2008)):

$$\begin{cases} \epsilon \dfrac{dy}{dt} = -y + (y)^+ - \nabla\psi((y)^+) - \nabla g((y)^+)(\alpha)^+, \\ \epsilon \dfrac{d\alpha}{dt} = -\alpha + (\alpha)^+ - g((y)^+) \end{cases} \tag{12}$$

where ϵ is a positive time constant, $\nabla\psi(\cdot)$ denotes the gradient of ψ, and $(\cdot)^+$ is the piecewise linear activation function which is defined as follows:

$$(y_i)^+ = \begin{cases} 0, & y_i < 0; \\ y_i, & y_i \geq 0. \end{cases}$$

If $\psi(y)$ is nonsmooth, a globally convergent neurodynamic model for solving (10) is described as follows (Li et al. (2015)):

$$\epsilon \frac{dy}{dt} \in -\nabla\psi(y) - \lambda\partial \sum_i \max\{0, g_i(y)\} \tag{13}$$

where λ is a penalty parameter, $\partial(\cdot)$ denotes Clarke's generalized gradient (Liu & Wang (2011)) and

$$\partial \max\{0, g_i(y)\} = \begin{cases} \nabla g_i(y), & g_i(y) > 0 \\ [0,1]\nabla g_i(y), & g_i(y) = 0. \\ 0, & g_i(y) < 0 \end{cases}$$

If the optimization problem (10) consists of an equality constraint such as $h(y) = 0$, the constraint can be equivalently replaced with two inequality constraints $h(y) \leq 0$ and $-h(y) \leq 0$ (Yan et al. (2017)).

A generic form of the neurodynamic system for solving (10) is described as follows:

$$\epsilon \frac{dy}{dt} \in \phi(\nabla\psi(y), \mathscr{Y}) \tag{14}$$

where $\phi(\cdot)$ is a function of $\nabla\psi(y)$ and $\mathscr{Y}$.

2.7. Collaborative neurodynamic optimization

It is challenging to solve global optimization problems with nonconvex objective functions using an individual neurodynamic model. To overcome the difficulty, various CNO approaches with multiple neurodynamic models are proposed recently (e.g., Yan et al. (2014, 2017), Che & Wang (2019a,b), Che & Wang (2021)). In a CNO approach, multiple neurodynamic models are employed collaboratively to seek global optimal solutions. The initial states of the models are updated by using meta-heuristics such as particle swarm optimization (PSO) (Clerc & Kennedy (2002)) with its update rule defined as follows:

$$\begin{aligned} v_i(j+1) =& c_0 v_i(j) + c_1 r_1(\tilde{y}_i(j) - y_i(j)) \\ & + c_2 r_2(\hat{y} - y_i(j)), \end{aligned} \tag{15}$$

$$y_i(j+1) = y_i(j) + v_i(j+1) \tag{16}$$

where $y_i(j) = (y_{i1}(k), \ldots, y_{in}(j))^T$ and $v_i(j) = (v_{i1}(j), \ldots, v_{in}(j))^T$ is the position and velocity of the i-th particle at the j-th iteration, c_0 is inertia weight; c_1 and c_2 are weighting parameters; r_1 and r_2 are random values generated in [0, 1], $\tilde{y}_i(j) = (\tilde{y}_{i1}(j), \ldots, \tilde{y}_{in}(j))^T$ is the previous best solution for the i-th particle at the j-th iteration; $\hat{y} = (\hat{y}_1, \ldots, \hat{y}_n)^T$ is the best solution of the swarm.

To enhance the exploratory capability, wavelet mutation is sometimes adopted (Ling et al. (2008), Fan & Wang (2017)). Let η be defined by a wavelet function; i.e.,

$$\eta = \frac{1}{\sqrt{a}} \exp(-\frac{\varrho}{2a}) \cos(\frac{5\varrho}{a}),$$

$a = \exp(10(j/j_{\max}))$, $\boldsymbol{e}$ be the Euler number, $j_{\max}$ be the maximum number of iterations, and ϱ be a uniformly distributed number generated within $(-2.5a, 2.5a)$ randomly (Ling et al. (2008)), the wavelet mutation is defined as follows:

$$y_i(k+1) = \begin{cases} y_i(j) + \eta(1 - y_i(j)), & \eta > 0; \\ y_i(j) + \eta(y_i(j)), & \eta < 0. \end{cases} \tag{17}$$

Let $\tau > 0$ be a threshold, the wavelet mutation is performed if $(\tau > \delta)$. δ is defined by a diversity measure (Juang (2004))

$$\delta = \frac{1}{M} \sum_{i=1}^{M} ||y_i(j+1) - \hat{y}||_2$$

where M is the number of particles in a group.

With the use of multiple RNNs repeatedly reinitialized using a meta-heuristic rule, it is proven that the CNO approaches are almost surely convergent to global optimal solutions of the optimization problems (**??**Che & Wang (2019b)).

As a special CNO approach with a pair of RNNs, a two-timescale duplex neurodynamic system based on generic model (14) (Che & Wang (2019b)) for solving (1) is defined in the following coupled differential equations:

$$\begin{cases} \epsilon_x \dfrac{dx}{dt} \in F(\nabla f(x,y), x), \\ \epsilon_y \dfrac{dy}{dt} \in F(\nabla f(x,y), y) \end{cases} \tag{18}$$

where $F(\cdot)$ is a function of $\nabla f(x,y)$ and x or y, $\nabla f(x,y)$ denotes the gradient of $f(x,y)$, ϵ_x and ϵ_y are two different time constants.

If $\mathscr{X} = \{(x,y)|g_i(x,y) \leq 0, i = 1, \dots, m\}$, system (18) based on model (13) is described as follows:

$$\begin{cases} \epsilon_x \dfrac{dx}{dt} \in -\nabla f(x,y) - \lambda \partial \displaystyle\sum_i \max\{0, g_i(x,y)\}, \\ \epsilon_y \dfrac{dy}{dt} \in -\nabla f(x,y) - \lambda \partial \displaystyle\sum_i \max\{0, g_i(x,y)\}. \end{cases} \tag{19}$$

In particular, two two-timescale duplex neurodynamic systems are proposed for biconvex optimization (Che & Wang (2019b)) and mixed-integer optimization (Che & Wang (2021)). Each system consists of two RNNs operating at two different timescales. It is proven that the duplex neurodynamic systems are almost surely convergent to global optimal solutions (Che & Wang (2019b), Che & Wang (2021)).

As a computationally intelligent optimization technique, CNO has been applied to portfolio optimization with promising results. In (Leung & Wang (2019)), a bi-objective portfolio optimization problem is formulated based on (2) and (3). The problem is then solved by a CNO approach and a set of solutions is generated. In addition to bi-objective portfolio selection, a CNO approach is developed for minimax portfolio optimization such as (4) in (Leung & Wang (2021)). In (Leung et al. (2022)), decentralized robust portfolio optimization problems are formulated based on the MV framework and they are solved by neurodynamic-based systems.

3. Problem Formulations

Using the conditional Sharpe ratio as the objective function, a cardinality-constrained portfolio optimization problem is formulated as follows:

$$\begin{aligned} \max_y \quad & \frac{\mu^T y - r_f}{\mathrm{CVaR}_\theta(y)} \\ \text{s.t.} \quad & e^T y = 1, \\ & ||y||_0 \leq k, \\ & y \geq 0. \end{aligned} \tag{20}$$

Problem (20) is equivalently reformulated as follows:

$$\begin{aligned} \min_y \quad & \frac{\mathrm{CVaR}_\theta(y)}{\mu^T y - r_f} \\ \text{s.t.} \quad & e^T y = 1, \\ & ||y||_0 \leq k, \\ & y \geq 0. \end{aligned} \tag{21}$$

Problems (20) and (21) are mixed-integer optimization problems due to the discontinuity of the l_0-norm in the cardinality constraint. In addition, their fractional objective functions incur some difficulties in optimization (Wang et al. (2021)).

To obviate the use of a fractional function, we introduce a variable weight γ and reformulate problem (21) by minimizing the weighted numerator and maximizing the weighted denominator of its factional objective function as follows:

$$\begin{aligned} \min_{y,\gamma} \quad & \frac{\gamma^2}{2}\mathrm{CVaR}_\theta(y)^2 - \gamma(\mu^T y - r_f) \\ \text{s.t.} \quad & e^T y = 1, \\ & ||y||_0 \leq k, \\ & y \geq 0, \\ & \gamma \geq 0. \end{aligned} \tag{22}$$

For long-term investments, it is reasonable to assume that the expected return is not less than the risk-free return (i.e., $\mu^T y \geq r_f$). Based on the above assumption and in view of minimization, the nonnegativity constraint on γ becomes redundant and can be dropped.

Once the nonnegativity constraint of γ is dropped, the optimal solution of γ for any given y may be analytically derived by zeroing the partial derivative of the surrogate function in (22) with respect to γ as follows:

$$\gamma = \frac{\mu^T y - r_f}{\mathrm{CVaR}_\theta(y)^2}. \tag{23}$$

Substituting (23) into the surrogate objective function in (22), we have

$$\min_y \quad -\frac{1}{2}\frac{(\mu^T y - r_f)^2}{\mathrm{CVaR}_\theta(y)^2}. \tag{24}$$

Clearly, it is equivalent to the surrogate objective function in (20) and (21) in terms of optimal solutions.

Although the objective function in problem (22) is no longer fractional, directly solving it is still nontrivial due to the discontinuity of the l_0-norm in the cardinality constraint. By introducing a binary vector $z \in \{0,1\}^n$ and using the sample approximation of CVaR in (7), the cardinality constrained portfolio optimization problem in (22) is further reformulated as follows:

$$\begin{aligned}
\min_{\gamma,\rho,\sigma,y,z} \ & \frac{\gamma^2}{2}\Big(\rho + \frac{1}{N(1-\theta)}\sum_{J=1}^{N}\sigma_J\Big)^2 - \gamma(\mu^T y - r_f) \\
\text{s.t. } & \sigma_i \geq -\xi_i^T y - \rho, \ \sigma_i \geq 0, \ i = 1,2,\ldots,N; \\
& e^T y = 1; \\
& e^T z \leq k; \\
& 0 \leq y \leq z; \\
& z \in \{0,1\}^n;
\end{aligned} \tag{25}$$

where $z \in \{0,1\}^n$, $e^T z \leq k$ is the cardinality constraint. When $z_i = 0$, y_i is zero because of the constraint $0 \leq y \leq z$, indicating that the i-th stock is not selected in the portfolio.

In view that mixed-integer optimization problem (25) is difficult to solve, as in (Che et al., 2022), the binary constraint $z \in \{0,1\}^n$ is replaced by a set of bilinear and linear equality constraints as follows:

$$z \circ \zeta = 0, \quad z + \zeta - e = 0 \tag{26}$$

where $z = (z_1, \ldots, z_n)^T \in \Re^n$ and $\zeta = (\zeta_1, \ldots, \zeta_n)^T \in \Re^n$, $\circ$ denotes the Hadamard product operator of two vectors. The equality constraints are satisfied if only if $z_i = 1$ or 0 and $\zeta_i = 0$ or 1 for all i. Based on (26), problem (25) is finally reformulated as follows:

$$\begin{aligned}
\min_{\gamma,\rho,\sigma,y,z,\zeta} \ & \frac{\gamma^2}{2}\Big(\rho + \frac{1}{N(1-\theta)}\sum_{J=1}^{N}\sigma_J\Big)^2 - \gamma(\mu^T y - r_f) \\
\text{s.t. } & \sigma_J \geq -\xi_J^T y - \rho, \ \sigma_J \geq 0, \ J = 1,2,\ldots,N; \\
& e^T y = 1; \\
& e^T z \leq k; \\
& 0 \leq y \leq z; \\
& z \circ \zeta = 0; \\
& z + \zeta - e = 0.
\end{aligned} \tag{27}$$

For fixed γ and ζ, problem (27) is convex and for fixed ρ, σ, y and z, problem (27) is also convex. According to Definitions 1 and 2, problem (27) is biconvex.

4. Two-Timescale Duplex Neurodynamic System

To solve the biconvex portfolio optimization problem in (27), a neurodynamic model is customized based on RNN (19) as follows:

$$\begin{cases}
\epsilon_1 \dfrac{d\gamma}{dt} \in -\nabla f_c(\gamma,\rho,\sigma,y) - \lambda\partial \displaystyle\sum_i \max\{0, \boldsymbol{c}_i(\rho,\sigma,y,z,\zeta)\} \\
\epsilon_2 \dfrac{d\rho}{dt} \in -\nabla f_c(\gamma,\rho,\sigma,y) - \lambda\partial \displaystyle\sum_i \max\{0, \boldsymbol{c}_i(\rho,\sigma,y,z,\zeta)\} \\
\epsilon_2 \dfrac{d\sigma}{dt} \in -\nabla f_c(\gamma,\rho,\sigma,y) - \lambda\partial \displaystyle\sum_i \max\{0, \boldsymbol{c}_i(\rho,\sigma,y,z,\zeta)\} \\
\epsilon_2 \dfrac{dy}{dt} \in -\nabla f_c(\gamma,\rho,\sigma,y) - \lambda\partial \displaystyle\sum_i \max\{0, \boldsymbol{c}_i(\rho,\sigma,y,z,\zeta)\} \\
\epsilon_2 \dfrac{dz}{dt} \in -\lambda\partial \displaystyle\sum_i \max\{0, \boldsymbol{c}_i(\rho,\sigma,y,z,\zeta)\} \\
\epsilon_1 \dfrac{d\zeta}{dt} \in -\lambda\partial \displaystyle\sum_i \max\{0, \boldsymbol{c}_i(\rho,\sigma,y,z,\zeta)\}
\end{cases} \tag{28}$$

where $f_c(\gamma,\rho,\sigma,y)$ is the objective function in (27) and $c(\rho,\sigma,y,z,\zeta)$ is the vector-valued inequality constraints. As shown in (28), the neural network model consists of six layers. The dynamics of the first four layers minimize the objective function in (27) with various constraints and the states move toward the feasible region, while the dynamics of the fifth and sixth layers are to realize the binary constraint z by satisfying a set of the bilinear and linear equality constraints (26). In RNN (28), there are $2n+N+2$ neurons.

As the portfolio optimization problem (27) is biconvex, the use of a single RNN may not be able to converge to the global optimum. As in (Che & Wang (2021), Che et al. (2022)), by making use of two RNNs operating at different timescales (i.e., $\epsilon_1 > \epsilon_2$ for RNN1, and $\epsilon_2 < \epsilon_1$ for RNN2), a two-timescale duplex neurodynamic approach is developed for solving the biconvex portfolio optimization (27). Besides, the PSO rule (Clerc & Kennedy (2002)) is adopted to reinitialize the states of RNN (28).

Algorithm 1 details the two-timescale duplex neurodynamic approach to cardinality-constrained portfolio selection. In Step 1, the states of RNN1 and RNN2 are initialized. In Steps 2-5, the termination threshold ε is set. The individual minima $p_I(0)$ and the minimum of the system $p^*(0)$ are set. In Steps 8-13, $p_I(j)$ is obtained as the best solution among the steady states of the two RNNs up to the j-th iteration. In Steps 14-18, the group minimum p^* is updated. In Steps 19-22, PSO is adopted to reinitialize the searching process of the two RNNs. The optimization process continues until meeting the termination criterion $||p^*(j+1) - p^*(j)|| \leq \varepsilon$. The spatial complexity of the algorithm depends dominantly on the number of neurons. As the algorithm employs two RNNs and there are $2n+N+2$ neurons in each RNN, the spatial complexity of the algorithm is $4n+2N+4$.

As optimization problem (27) is biconvex, the two-timescale duplex neurodynamic optimization approach in Algorithm 1, with different initial states and sufficiently different time constants in RNN1 and RNN2, is almost surely convergent to the global optimum of problem (27) (Che & Wang (2019b), Che & Wang (2021)).

Algorithm 1: Two-Timescale Duplex Neurodynamic Optimization for Biconvex Portfolio Optimization

---

Initialize the states of RNN1, RNN2 randomly: $(\gamma_1(0), \rho_1(0), \sigma_1(0), y_1(0), z_1(0), \zeta_1(0))$, $(\gamma_2(0), \rho_2(0), \sigma_2(0), y_2(0), z_2(0), \zeta_2(0))$, and set the error tolerance ε;
for $I = 1:2$ **do**
  $p_I(0) = (\gamma_I(0), \rho_I(0), \sigma_I(0), y_I(0), z_I(0), \zeta_I(0))$;
  $p^*(0) = \arg\min(f_c(\gamma_I(0), \rho_I(0), \sigma_I(0), y_I(0)))$;
end
$j \leftarrow 1$;
while $||p^*(j+1) - p^*(j)|| \geq \varepsilon$ **do**
  Compute steady states $(\bar{\gamma}_1(j), \bar{\rho}_1(j), \bar{\sigma}_1(j), \bar{y}_1(j), \bar{z}_1(j), \bar{\zeta}_1(j))$ and $(\bar{\gamma}_2(j), \bar{\rho}_2(j), \bar{\sigma}_2(j), \bar{y}_2(j), \bar{z}_2(j), \bar{\zeta}_2(j))$ by (28);
  if $f_c(\bar{\gamma}_I(j), \bar{\rho}_I(j), \bar{\sigma}_I(j), \bar{y}_I(j)) < f_c(p_I(j))$ **then**
    $p_I(j+1) = (\bar{\gamma}_I(j), \bar{\rho}_I(j), \bar{\sigma}_I(j), \bar{y}_I(j), \bar{z}_I(j), \bar{\zeta}_I(j))$;
  else
    $p_I(j+1) = p_I(j)$;
  end
  if $f_c(p_I(j+1)) < f_c(p^*(j))$ **then**
    $p^*(j+1) = p_I(j+1)$;
  else
    $p^*(j+1) = p^*(j)$;
  end
  Compute $(\gamma_I(j+1), \rho_I(j+1), \sigma_I(j+1), y_I(j+1), z_I(j+1), \zeta_I(j+1))$ by (15) and (16);
  if $(\tau > \delta)$ **then**
    Perform the wavelet mutation using (17);
  end
  $j \leftarrow j+1$;
end

---

5. Experimental Results

5.1. Setups

As in (Leung & Wang (2021, 2022)), the experiments are based on four datasets: HDAX (Deutsche Borse), FTSE (London Stock Exchange), HSCI (Hong Kong Stock Exchange), and SP500 (New York Stock Exchange and Nasdaq Stock Market), constructed based on the 938 weekly adjusted closing prices of stocks from January 3, 2000 to December 29, 2017. According to the common practice, suspended and newly enlisted stocks within the period are excluded (Chang et al. (2000), Woodside-Oriakhi et al. (2011), Guastaroba & Speranza (2012)). Therefore, datasets HDAX, FTSE, HSCI, and SP500 consist of 49, 56, 77, and 356 stocks, respectively. In the experiments, the datasets are divided for in-sample learning and out-of-sample testing in two ways: first one-third for in-sample pre-training and rest two-thirds for out-of-sample testing, half and half. During the out-of-sample testing, the problem parameter learning continues based on all available historical return data from the beginning week to the week before next portfolio rebalancing. That is, the portfolios are optimized with the problem parameters updated periodically based on the pricing data in a sequentially prolonged time window.

In addition, k in (27) is set to different values (i.e., $k = 44, 34, 24, 14$, and 4 on HDAX; $k = 50, 39, 28, 16$, and 5 on FTSE; $k = 69, 53, 38, 23$, and 7 on HSCI and $k = 320, 249, 178, 106$, and 35 on SP500), as in (Leung & Wang (2022)).

The risk-free rate r_f is determined based on the annualized return rates of the US Treasury three-month T-bill r_{yearly} and converted to weekly rates according to $(1 + r_{weekly})^{938/18} - 1 = r_{yearly}$, $r_f = r_{weekly} = (1 + r_{yearly})^{18/938} - 1$ (Hodoshima (2018)). As r_f is a simple return rate, simple return rates are used in all experiments.

To evaluate the performance of the proposed neurodynamic approach to portfolio optimization, three strong competitors are used for comparison: 1) a collaborative neurodynamic approach (denoted as CNO) with 20 neurodynamic models Leung & Wang (2022), 2) an equally-weighted approach for portfolio selection (denoted as EW) (DeMiguel et al. (2009)), and 3) market index (denoted as MI).

In the CVaR estimation, θ is set to 0.95, and N is set as the number of all available historical data at the decision time. In the two-timescale duplex neurodynamic model, $\epsilon_1/\epsilon_2 = 10$ in RNN1, $\epsilon_2/\epsilon_1 = 0.1$ in RNN2. In the PSO rule, c_1 and c_2 are set to 1.49. In the algorithm, termination threshold $\varepsilon = 10^{-3}$, the diversity threshold $\tau = 0.1$ as in (Fan & Wang (2017)), $j_{\max} = 50$, r_1, r_2, and the initial states of $\gamma, \rho, \sigma, y, z$, and ζ are randomly generated within 0 and 1.

5.2. Results

Figures 1 to 4 depict the transient behaviors of y, z, and ζ of the two-timescale duplex neurodynamic system in (28) on the four datasets. The subplots in the first row of these figures show that the state vector y converges within seven iterations and the subsets of stocks are selected for cardinality-constraints portfolios. Besides, the subplots in the second and third rows of the figures show that the state vectors z and ζ converge to zero or one, indicating whether a stock is selected or not in cardinality-constrained portfolios.

Table 1 records annualized SR, CSR, and returns of the resulting portfolios on the four datasets, where DNO denotes the proposed duplex neurodynamic optimization, CNO denotes the CNO-based method for cardinality-constrained portfolio selection (Leung & Wang (2022)), EW denotes the equally-weighted portfolio, MI denotes the market in-

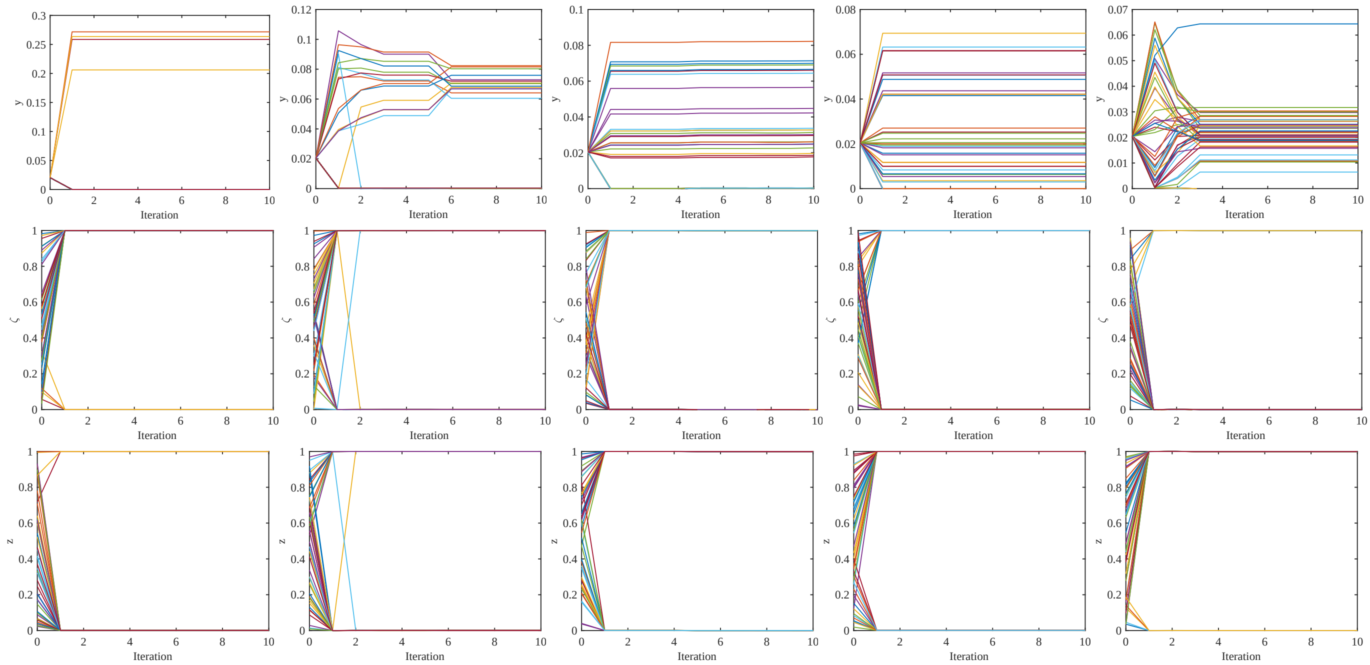

Fig. 1. Transient states y, ζ, and z of the two-timescale duplex neurodynamic model (28) for solving portfolio optimization problem (27) with $k = 4, 14, 24, 34$, and 44 (the subplots of the columns from left to right) on HDAX.

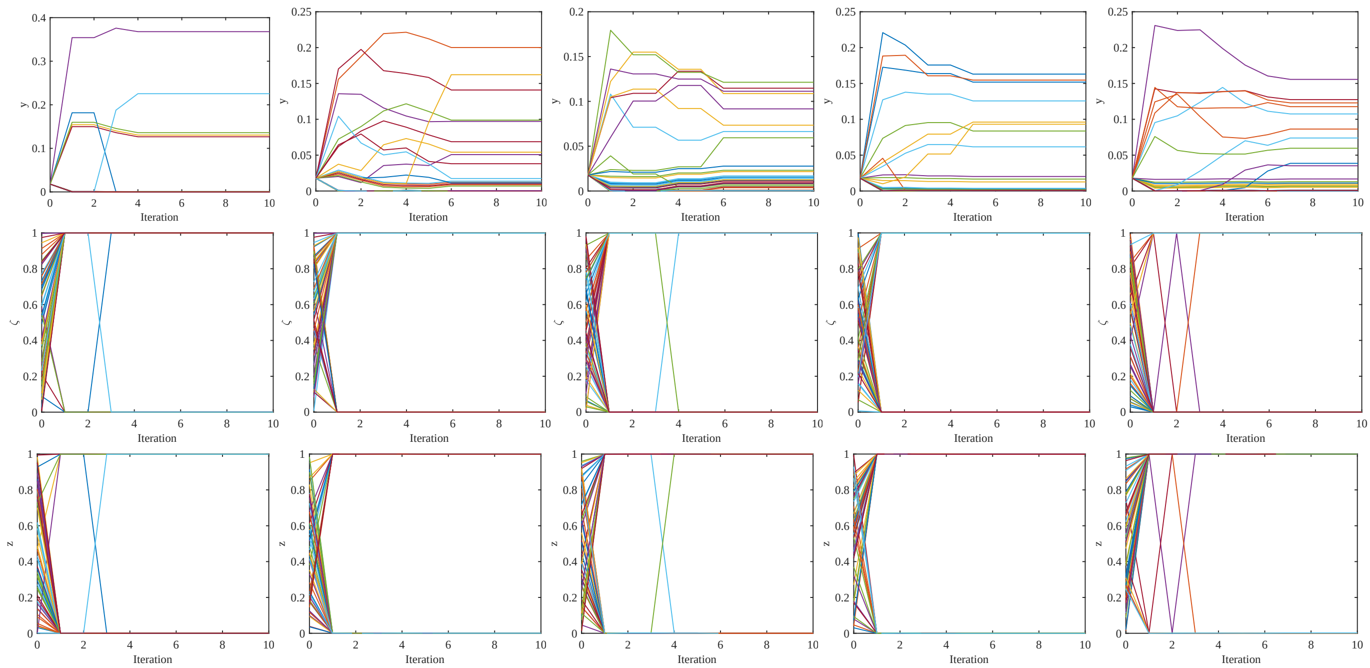

Fig. 2. Transient states y, ζ, and z of the two-timescale duplex neurodynamic model (28) for solving portfolio optimization problem (27) with $k = 5, 16, 28, 39$, and 50 (the subplots of the columns from left to right) on FTSE.

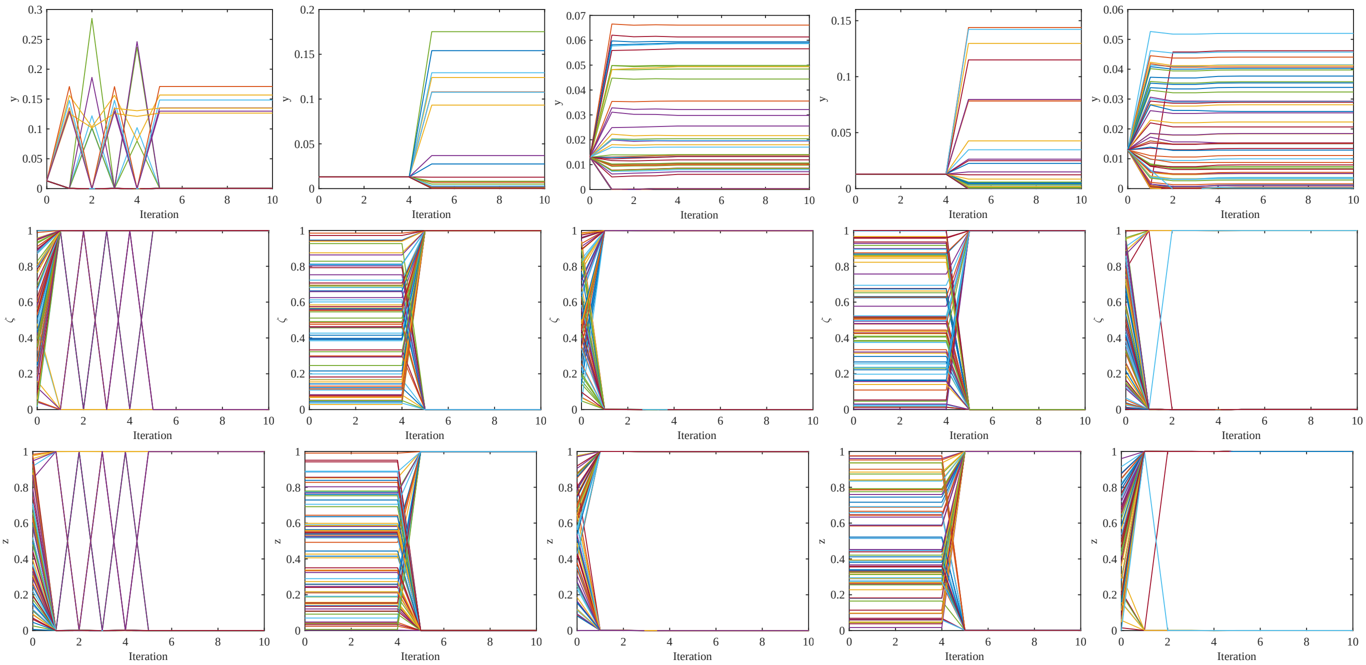

Fig. 3. Transient states y, ζ, and z of the two-timescale duplex neurodynamic model (28) for solving portfolio optimization problem (27) with $k = 7, 23, 38, 53$, and 69 (the subplots of the columns from left to right) on HSCI.

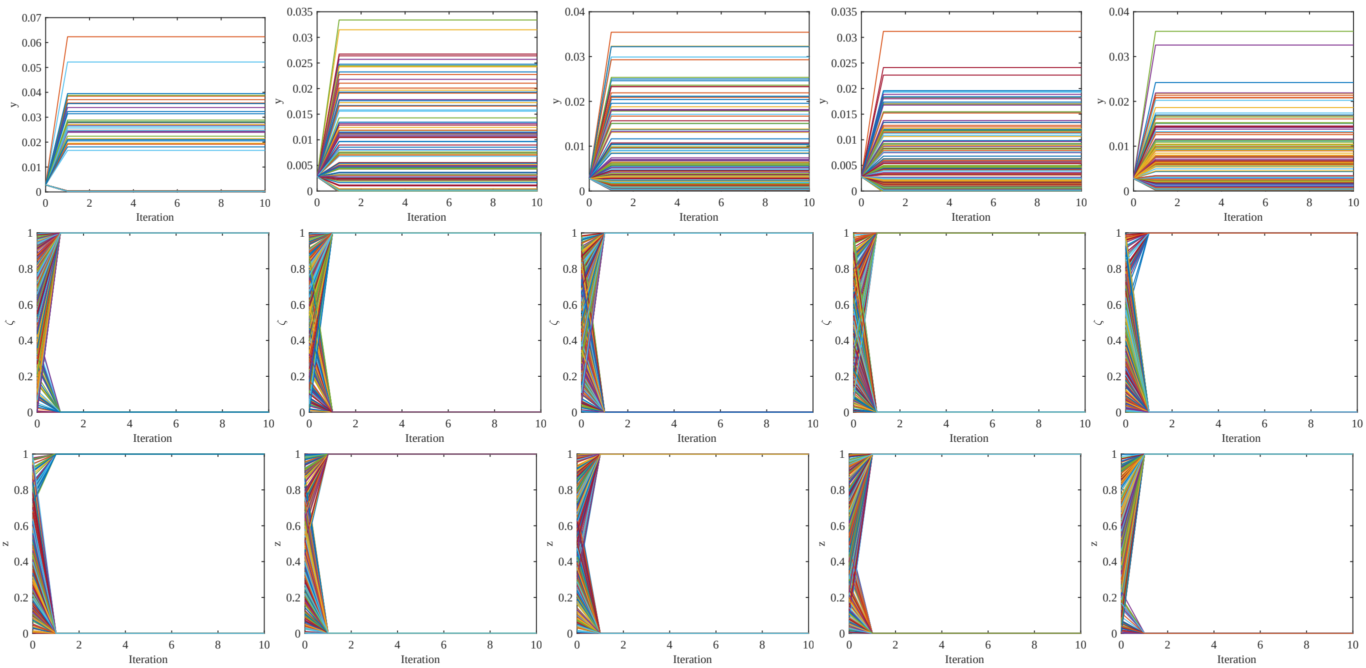

Fig. 4. Transient states y, ζ, and z of the two-timescale duplex neurodynamic model (28) for solving portfolio optimization problem (27) with $k = 35, 106, 178, 249$, and 320 (the subplots of the columns from left to right) on SP500.

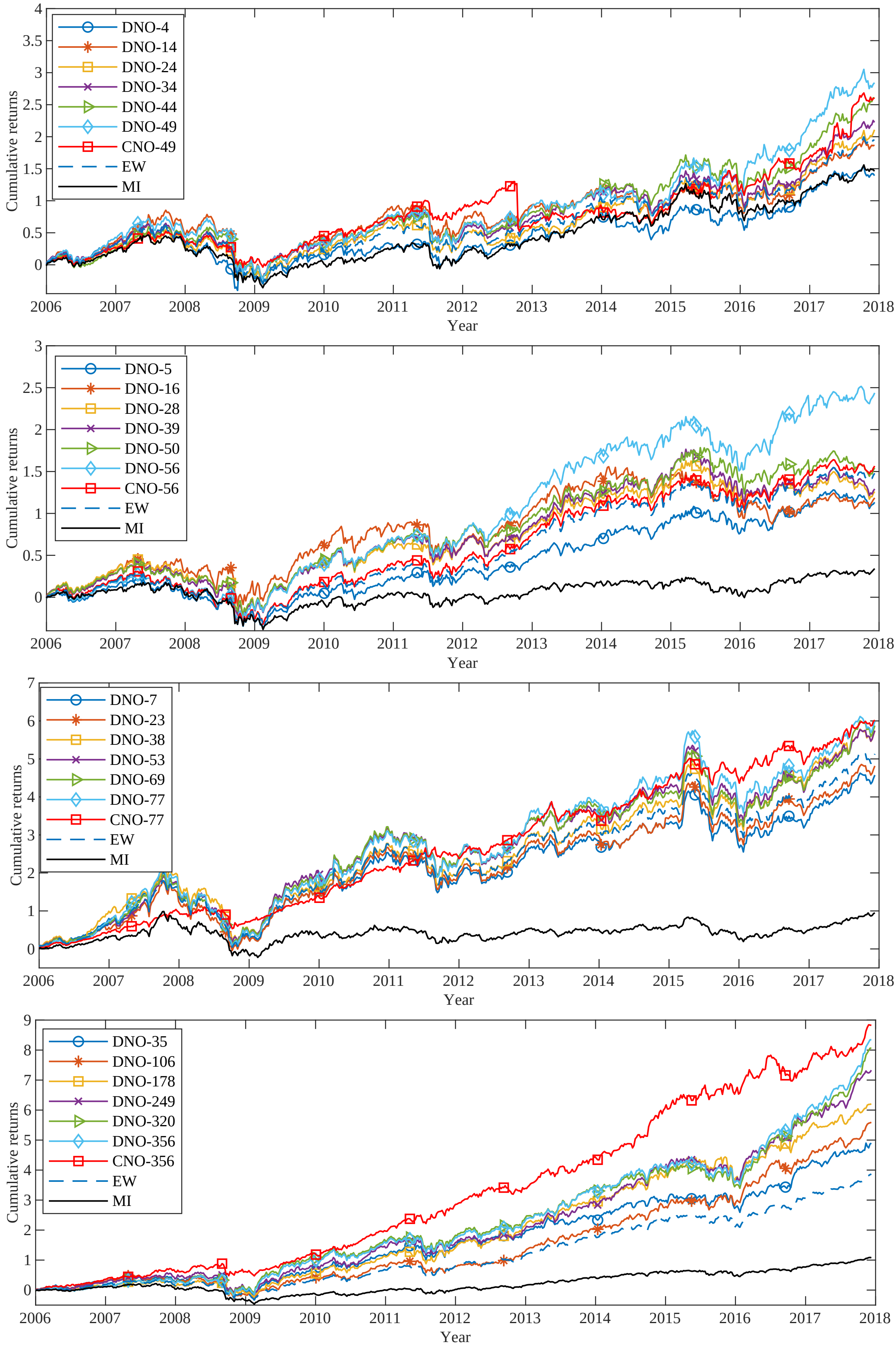

Fig. 5. Cumulative returns of different portfolios based on datasets from HDAX (the first subplot), FTSE (the second subplot), HSCI (the third subplot), and SP500 (the last subplot) based on 1/3-2/3 partitioned datasets.

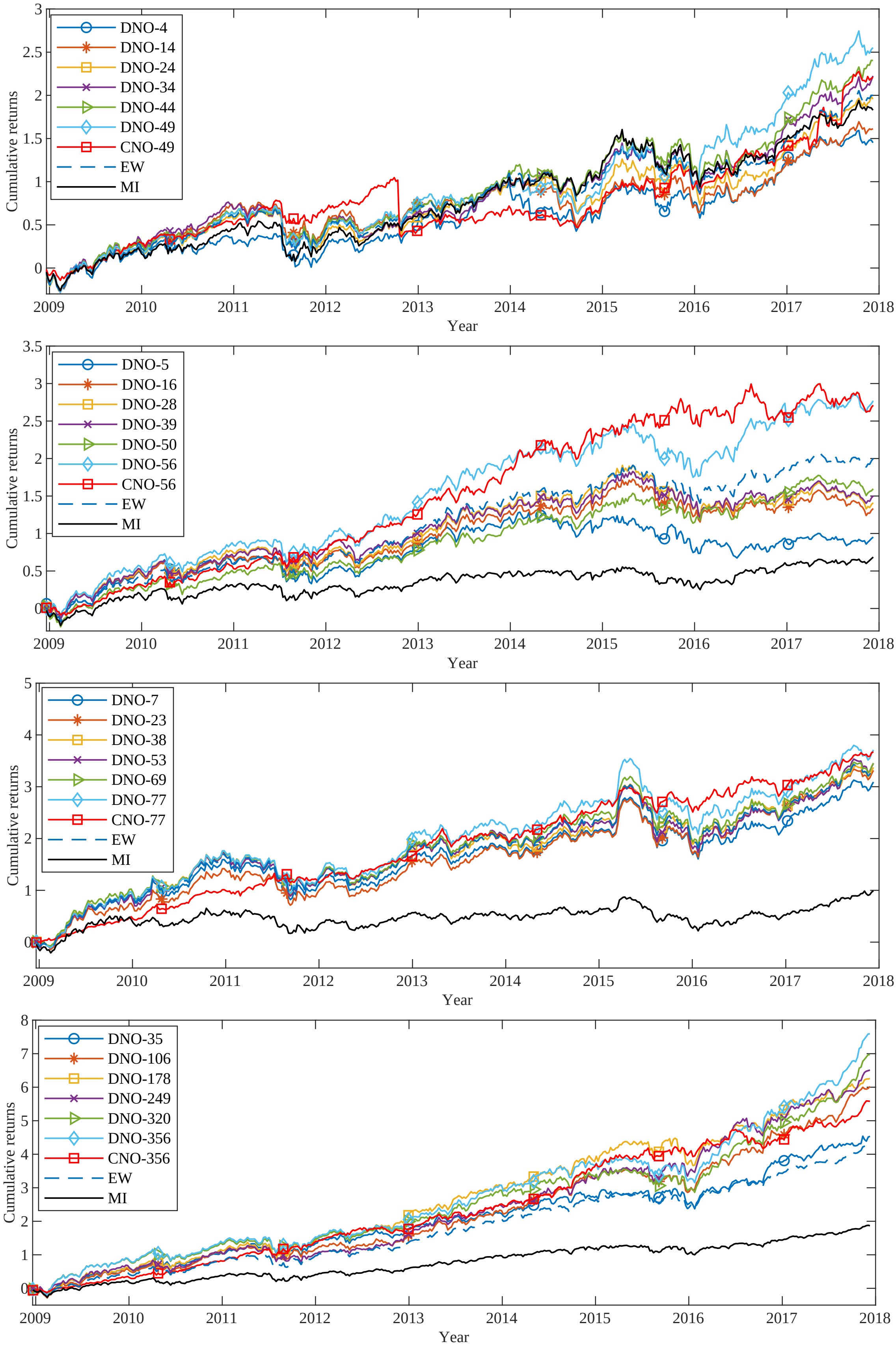

Fig. 6. Cumulative returns of different portfolios based on datasets from HDAX (the first subplot), FTSE (the second subplot), HSCI (the third subplot) and SP500 (the last subplot) based on half-and-half partitioned datasets.

Table 1. Resulting annualized Sharpe ratios, conditional Sharpe ratios, and returns based on 1/3-2/3 partitioned datasets.

Dataset	n	k	Sharpe ratio				Conditional Sharpe ratio				Annualized return (%)			
			DNO	CNO	EW	MI	DNO	CNO	EW	MI	DNO	CNO	EW	MI
HDAX	49	4	0.3463	0.3071	0.4715	0.4087	0.1701	0.1639	0.2044	0.1859	7.5389	8.3862	9.4248	7.7961
		14	0.4535	0.4244			0.1990	0.1828			9.1808	9.4128		
		24	0.4857	0.4627			0.2001	0.1909			9.9109	9.8893		
		34	0.4947	0.4600			0.2059	0.1938			10.2469	10.0069		
		44	0.5497	0.5396			0.2097	0.2043			11.2796	11.0843		
		49	0.5588	0.5415			0.2123	0.2082			11.8719	11.2828		
FTSE	56	5	0.3713	0.3040	0.4470	0.1705	0.1474	0.1156	0.1764	0.0808	6.5139	6.0002	7.8591	2.4769
		16	0.3774	0.3240			0.1496	0.1244			6.7312	6.6281		
		28	0.4002	0.3338			0.1674	0.1288			7.0314	7.1953		
		39	0.4075	0.3817			0.1709	0.1413			7.1560	7.4920		
		50	0.4319	0.4221			0.1807	0.1451			8.0872	7.6369		
		56	0.5617	0.5399			0.2332	0.2218			10.8381	8.3538		
HSCI	77	7	0.7044	0.6347	0.7518	0.3174	0.2839	0.2915	0.2357	0.1057	15.3946	15.8822	16.3132	5.7223
		23	0.7186	0.6599			0.3006	0.2981			15.8024	15.9351		
		38	0.7712	0.6704			0.3094	0.3042			17.2347	16.2815		
		53	0.7730	0.6923			0.3125	0.3103			17.2441	16.6426		
		69	0.7882	0.7074			0.3193	0.3140			17.4298	17.0163		
		77	0.7903	0.7386			0.3241	0.3203			17.5837	17.6290		
SP500	356	35	0.7112	0.6748	0.7229	0.3810	0.3108	0.3009	0.2429	0.1253	15.9359	18.9221	14.0971	6.2927
		106	0.7231	0.7169			0.3195	0.3121			17.0170	17.5123		
		178	0.7294	0.7207			0.3309	0.3224			17.8873	18.9221		
		249	0.7313	0.7230			0.3416	0.3377			19.3203	19.5123		
		320	0.7392	0.7309			0.3508	0.3457			20.1652	20.2054		
		356	0.7403	0.7337			0.3520	0.3506			20.4722	20.9749		

Table 2. Resulting annualized Sharpe ratio, conditional Sharpe ratio and returns based on half-and-half partitioned datasets.

Dataset	n	k	Sharpe ratio				Conditional Sharpe ratio				Annualized return (%)			
			DNO	CNO	EW	MI	DNO	CNO	EW	MI	DNO	CNO	EW	MI
HDAX	49	4	0.5486	0.5231	0.6873	0.6726	0.2475	0.2349	0.2734	0.3447	10.4758	7.9459	12.9857	12.2463
		14	0.5975	0.5689			0.2863	0.2487			11.2641	11.1892		
		24	0.6687	0.6357			0.2963	0.2532			12.7997	12.1576		
		34	0.7271	0.6848			0.3401	0.2688			13.8972	13.4195		
		44	0.7574	0.7119			0.3658	0.2899			14.5806	13.6594		
		49	0.7582	0.7565			0.3836	0.3129			15.1259	13.8340		
FTSE	56	5	0.4844	0.4349	0.8303	0.4364	0.2053	0.1936	0.3659	0.2374	7.7156	7.9548	12.9811	5.9705
		16	0.6338	0.6187			0.2963	0.2311			9.8699	8.2565		
		28	0.6600	0.6302			0.3051	0.2412			10.2688	8.4036		
		39	0.6755	0.6554			0.3159	0.2631			10.7203	9.1148		
		50	0.7214	0.7201			0.3562	0.2824			11.1810	12.4485		
		56	0.8756	0.8542			0.3965	0.3554			15.8881	15.6841		
HSCI	77	7	0.9315	0.8645	0.9791	0.4782	0.4353	0.4086	0.3459	0.1710	16.9461	16.9073	17.8610	7.9394
		23	0.9411	0.8790			0.4420	0.4185			17.6416	17.1799		
		38	0.9551	0.8941			0.4487	0.4214			17.8137	17.4809		
		53	0.9640	0.9074			0.4521	0.4267			18.0446	17.7833		
		69	1.9881	0.9186			0.4566	0.4332			18.0600	18.0823		
		77	1.0147	0.9369			0.4635	0.4471			18.7928	18.6832		
SP500	356	35	1.1314	1.1145	1.1599	0.8255	0.5019	0.4147	0.4120	0.2801	20.9448	19.7990	20.3834	12.4253
		106	1.1834	1.1440			0.5633	0.4417			24.1746	20.3223		
		178	1.2083	1.1748			0.5852	0.5189			24.6381	20.9244		
		249	1.2141	1.1904			0.6197	0.5435			25.1157	21.5870		
		320	1.2187	1.2126			0.6362	0.5679			25.9424	22.3755		
		356	1.2231	1.2175			0.6539	0.6488			26.9975	23.2873		

dex. The resulting DNO portfolios achieve the highest annualized SR, CSR values, and returns as $k = 24, 34$, and 44; though DNO underperforms the EW portfolio that is not constrained by cardinality on HDAX when $k = 4$ and 14. For the FTSE dataset, the DNO portfolios result in higher annualized SR values than the baselines as $k = 56$; though DNO underperforms the EW portfolio that is not constrained by cardinality when $k = 5, 16, 28, 39$ and 50. Similarly, the DNO portfolios result in higher annualized CSR values than the baselines as $k = 50$ and 56; though DNO underperforms EW in terms of annualized CSR and returns when $k = 5, 16, 28$ and 39. The DNO portfolios result in higher annualized SR values than the baselines as $k = 38, 53, 69$, and 77; though DNO underperforms EW in terms of annualized SR and returns when $k = 7$ and 23 on HSCI. Besides, the resulting DNO portfolios outperform the three baselines in terms of annualized CSR in the HSCI dataset. The DNO portfolios result in higher annualized SR values than the baselines in the SP500 dataset. Similarly, the DNO portfolios result in higher annualized CSR values and returns than the baselines. Besides, it can be seen that the SR and CSR values of the DNO portfolios increase when the value of k increases. Table 2 shows similar results, where the data of the first one-half periods are used for in-sample learning. By comparing the results in Tables 1 and 2, it can be seen that all results in Table 2 are better than those in Table 1, because more samples are used in the former than the latter for in-sample learning.

Figures 5 and 6 depict the cumulative returns of various portfolios rebalanced weekly based on the 1/3-2/3 and 1/2-1/2 partitioned datasets, respectively. It can be seen that the cumulative returns increase as the value of k increases in general. Specifically, Figure 5 shows that the cumulative returns of the DNO portfolios are the highest as $k = 44$, $k = 50$, $k = 69$ and $k = 320$ on the four datasets. However, DNO-4 and DNO-14 on HDAX, DNO-5 to DNO-39 on FTSE, and DNO-7 and DNO-23 on HSCI underperform the EW portfolios, as the EW portfolios are not constrained by any cardinality. Figure 6 shows similar results, where the data of the first one-half periods are used for in-sample pre-training. The figure shows that the cumulative returns of the DNO portfolios are the highest as $k = 44$, $k = 69$ and $k = 320$ on the HDAX, HSCI and SP500 datasets, respectively. Nevertheless, DNO-4, DNO-14, DNO-24 on HDAX, DNO-5 to DNO-50 on FTSE, and DNO-7, DNO-23 on HSCI underperform the EW portfolios, as the EW portfolios are unconstrained by any cardinality.

6. Concluding Remarks

In this paper, a neurodynamic approach is developed for cardinality-constrained portfolio selection in Markowitz's risk-return framework. A mixed-integer optimization problem is formulated to implement the cardinality constraint and a biconvex optimization problem is reformulated to optimize the subsets of stocks with given cardinalities for portfolio selection. The two-timescale duplex neurodynamic approach consists of a pair of neurodynamic optimization models operating in parallel. Experimental results on the benchmark datasets in four world markets show that the proposed method is globally convergent and outperforms three baselines in terms of two commonly used risk-adjusted criteria and investment returns. The superior performances result from the combined use of the optimally weighted biconvex problem formulation for maximizing conditional Sharpe ratio and the effective optimizer driven by two-timescale duplex neurodynamics. Further investigations include developing more efficient neurodynamic approaches for high-frequency trading and multi-period portfolio selection.

References

Ang, A., & Chen, J. (2002). Asymmetric correlations of equity portfolios. *Journal of Financial Economics*, *63*, 443–494.

Artzner, P., Delbaen, F., Eber, J.-M., & Heath, D. (1999). Coherent measures of risk. *Mathematical Finance*, *9*, 203–228.

Bian, W., Ma, L., Qin, S., & Xue, X. (2018). Neural network for nonsmooth pseudoconvex optimization with general convex constraints. *Neural Networks*, *101*, 1–14.

Chang, T.-J., Meade, N., Beasley, J. E., & Sharaiha, Y. M. (2000). Heuristics for cardinality constrained portfolio optimization. *Computers & Operations Research*, *27*, 1271–1302.

Che, H., & Wang, J. (2019a). A collaborative neurodynamic approach to global and combinatorial optimization. *Neural Networks*, *114*, 15 – 27.

Che, H., & Wang, J. (2019b). A two-timescale duplex neurodynamic approach to biconvex optimization. *IEEE Transactions on Neural Networks and Learning Systems*, *30*, 2503–2514.

Che, H., & Wang, J. (2021). A two-timescale duplex neurodynamic approach to mixed-integer optimization. *IEEE Transactions on Neural Networks and Learning Systems*, *32*, 36–48.

Che, H., Wang, J., & Cichocki, A. (2022). Bicriteria sparse nonnegative matrix factorization via two-timescale duplex neurodynamic optimization. *IEEE Transactions on Neural Networks and Learning Systems*, . In press.

Christiansen, C., Joensen, J. S., & Nielsen, H. S. (2007). The risk-return trade-off in human capital investment. *Labour Economics*, *14*, 971–986.

Clerc, M., & Kennedy, J. (2002). The particle swarm-explosion, stability, and convergence in a multidimensional complex space. *IEEE Transactions on Evolutionary Computation*, *6*, 58–73.

DeMiguel, V., Garlappi, L., & Uppal, R. (2009). Optimal versus naive diversification: How inefficient is the 1/N portfolio strategy? *The Review of Financial Studies*, *22*, 1915–1953.

Deng, X.-T., Li, Z.-F., & Wang, S.-Y. (2005). A minimax portfolio selection strategy with equilibrium. *European Journal of Operational Research*, *166*, 278–292.

Eling, M., & Schuhmacher, F. (2007). Does the choice of performance measure influence the evaluation of hedge funds? *The Journal of Banking and Finance*, *31*, 2632–2647.

Ertenlice, O., & Kalayci, C. B. (2018). A survey of swarm intelligence for portfolio optimization: Algorithms and applications. *Swarm Evol. Comput.*, *39*, 36–52.

Fan, J., & Wang, J. (2017). A collective neurodynamic optimization approach to nonnegative matrix factorization. *IEEE Transactions on Neural Networks and Learning Systems*, *28*, 2344–2356.

Gaivoronski, A. A., & Pflug, G. (2005). Value-at-risk in portfolio optimization: properties and computational approach. *Journal of Risk*, *7*, 1–31.

Gao, J., & Li, D. (2013). Optimal cardinality constrained portfolio selection. *Oper. Res.*, *61*, 745–761.

Gorski, J., Pfeuffer, F., & Klamroth, K. (2007). Biconvex sets and optimization with biconvex functions: A survey and extensions. *Math. Meth. Oper. Res.*, *66*, 373–407.

Guastaroba, G., & Speranza, M. G. (2012). Kernel search: An application to the index tracking problem. *European Journal of Operational Research*, *217*, 54–68.

Guo, Z., Liu, Q., & Wang, J. (2011). A one-layer recurrent neural network for pseudoconvex optimization subject to linear equality constraints. *IEEE Transactions on Neural Networks*, *22*, 1892–1900.

Hardoroudi, N. D., Keshvari, A., Kallio, M., & Korhonen, P. (2017). Solving cardinality constrained mean-variance portfolio problems via MILP. *Annals of Operations Research*, *254*, 47–59.

Hodoshima, J. (2018). Stock performance by utility indifference pricing and the Sharpe ratio. *Quant. Finance*, *19*, 1–12.

Hopfield, J. J., & Tank, D. W. (1986). Computing with neural circuits - A model. *Science*, *233*, 625–633.

Hosseini, A. (2016). A non-penalty recurrent neural network for solving a class of constrained optimization problems. *Neural Networks*, *73*, 10–25.

Jiang, X., Qin, S., Xue, X., & Liu, X. (2022). A second-order accelerated neurodynamic approach for distributed convex optimization. *Neural Networks*, *146*, 161–173.

Juang, C.-F. (2004). A hybrid of genetic algorithm and particle swarm optimization for recurrent network design. *IEEE Transactions on Systems, Man, and Cybernetics, Part B (Cybernetics)*, *34*, 997–1006.

Kalayci, C. B., Polat, O., & Akbay, M. A. (2020). An efficient hybrid metaheuristic algorithm for cardinality constrained portfolio optimization. *Swarm and Evolutionary Computation*, *54*, 100662.

Kolm, P. N., Tütüncü, R., & Fabozzi, F. J. (2014). 60 years of portfolio optimization: Practical challenges and current trends. *European Journal of Operational Research*, *234*, 356–371.

Kroll, Y., Levy, H., & Markowitz, H. M. (1984). Mean-variance versus direct utility maximization. *The Journal of Finance*, *39*, 47–61.

Leung, M.-F., & Wang, J. (2018). A collaborative neurodynamic approach to multiobjective optimization. *IEEE Transactions on Neural Networks and Learning Systems*, *29*, 5738 – 5748.

Leung, M.-F., & Wang, J. (2019). A collaborative neurodynamic optimization approach to bicriteria portfolio selection. In H. Lu, H. Tang, & Z. Wang (Eds.), *Advances in Neural Networks – ISNN 2019* (pp. 318–327). Cham: Springer International Publishing.

Leung, M.-F., & Wang, J. (2021). Minimax and biobjective portfolio selection based on collaborative neurodynamic optimization. *IEEE Transactions on Neural Networks and Learning Systems*, *32*, 2825–2836.

Leung, M.-F., & Wang, J. (2022). Cardinality-constrained portfolio selection based on collaborative neurodynamic optimization. *Neural Networks*, *145*, 68–79.

Leung, M.-F., Wang, J., & Li, D. (2022). Decentralized robust portfolio optimization based on cooperative-competitive multiagent systems. *IEEE Transactions on Cybernetics*, (pp. 1–10).

Li, G., Yan, Z., & Wang, J. (2015). A one-layer recurrent neural network for constrained nonconvex optimization. *Neural Networks*, *61*, 10–21.

Ling, S.-H., Iu, H. H., Chan, K. Y., Lam, H.-K., Yeung, B. C., & Leung, F. H. (2008). Hybrid particle swarm optimization with wavelet mutation and its industrial applications. *IEEE Transactions on Systems, Man, and Cybernetics, Part B (Cybernetics)*, *38*, 743–763.

Liu, N., & Qin, S. (2019). A neurodynamic approach to nonlinear optimization problems with affine equality and convex inequality constraints. *Neural Networks*, *109*, 147–158.

Liu, N., Wang, J., & Qin, S. (2022). A one-layer recurrent neural network for nonsmooth pseudoconvex optimization with quasiconvex inequality and affine equality constraints. *Neural Networks*, *147*, 1–9.

Liu, Q., Dang, C., & Huang, T. (2013). A one-layer recurrent neural network for real-time portfolio optimization with probability criterion. *IEEE Transactions on Cybernetics*, *43*, 14–23.

Liu, Q., Guo, Z., & Wang, J. (2012). A one-layer recurrent neural network for constrained pseudoconvex optimization and its application for dynamic portfolio optimization. *Neural Networks*, *26*, 99–109.

Liu, Q., & Wang, J. (2011). A one-layer recurrent neural network for constrained nonsmooth optimization. *IEEE Trans. Systems, Man and Cybernetics - Part B: Cybernetics*, *40*, 1323–1333.

Liu, Q., & Wang, J. (2013). A one-layer projection neural network for nonsmooth optimization subject to linear equalities and bound constraints. *IEEE Transactions on Neural Networks and Learning Systems*, *24*, 812–824.

Liu, Q., Yang, S., & Wang, J. (2017). A collective neurodynamic approach to distributed constrained optimization. *IEEE Transactions on Neural Networks and Learning Systems*, *28*, 1747–1758.

Liu, S., Jiang, H., Zhang, L., & Mei, X. (2020a). A neurodynamic optimization approach for complex-variables programming problem. *Neural Networks*, *129*, 280–287.

Liu, Y., Zheng, Y., Lu, J., Cao, J., & Rutkowski, L. (2020b). Constrained quaternion-variable convex optimization: a quaternion-valued recurrent neural network approach. *IEEE Transactions on Neural Networks and Learning Systems*, *31*, 1022–1035.

Mansini, R., Ogryczak, W., & Speranza, M. G. (2014). Twenty years of linear programming based portfolio optimization. *European Journal of Operational Research*, *234*, 518–535.

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

Morgan, J. (1994). *Riskmetrics Technical Document*. New York: Morgan Guaranty Trust Company.

Morgenstern, O., & Von Neumann, J. (1953). *Theory of Games and Economic Behavior*. Princeton University Press.

Polak, G. G., Rogers, D. F., & Sweeney, D. J. (2010). Risk management strategies via minimax portfolio optimization. *European Journal of Operational Research*, *207*, 409–419.

Ponsich, A., Jaimes, A. L., & Coello, C. A. C. (2013). A survey on multiobjective evolutionary algorithms for the solution of the portfolio optimization problem and other finance and economics applications. *IEEE Transactions on Evolutionary Computation*, *17*, 321–344.

Rockafellar, R. T., & Uryasev, S. (2000). Optimization of conditional value-at-risk. *Journal of Risk*, *2*, 21–42.

Ruiz-Torrubiano, R., & Suárez, A. (2010). Hybrid approaches and dimensionality reduction for portfolio selection with cardinality constraints. *IEEE Computational Intelligence Magazine*, *5*, 92–107.

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

Sharpe, W. F. (1994). The Sharpe ratio. *J. Portfolio Manag.*, *21*, 49–58.

Sharpe, W. F. (2007). Expected utility asset allocation. *Financial Analysts Journal*, *63*, 18–30.

Steuer, R. E. (1986). *Multiple Criteria Optimization: Theory, Computation, and Applications*. Wiley.

Tank, D., & Hopfield, J. (1986). Simple 'neural' optimization networks: An A/D converter, signal decision circuit, and a linear programming circuit. *IEEE Transactions on Circuits and Systems*, *33*, 533–541.

Wang, J. (1994). A deterministic annealing neural network for convex programming. *Neural Networks*, *7*, 629–641.

Wang, Y., Wang, J., & Che, H. (2021). Two-timescale neurodynamic approaches to supervised feature selection based on alternative problem formulations. *Neural Networks*, *142*, 180–191.

Wen, X., Luan, L., & Qin, S. (2021). A continuous-time neurodynamic approach and its discretization for distributed convex optimization over multi-agent systems. *Neural Networks*, *143*, 52–61.

Woodside-Oriakhi, M., Lucas, C., & Beasley, J. E. (2011). Heuristic algorithms for the cardinality constrained efficient frontier. *European Journal of Operational Research*, *213*, 538–550.

Xia, Y., Feng, G., & Wang, J. (2008). A novel neural network for solving nonlinear optimization problems with inequality con-

straints. *IEEE Transactions on Neural Networks*, *19*, 1340–1353.
Xia, Z., Liu, Y., Kou, K. I., & Wang, J. (2022). Clifford-valued distributed optimization based on recurrent neural networks. *IEEE Transactions on Neural Networks and Learning Systems*, (pp. 1–12).
Xia, Z., Liu, Y., Lu, J., Cao, J., & Rutkowski, L. (2021). Penalty method for constrained distributed quaternion-variable optimization. *IEEE Transactions on Cybernetics*, *51*, 5631–5636.
Xu, C., Chai, Y., Qin, S., Wang, Z., & Feng, J. (2020). A neurodynamic approach to nonsmooth constrained pseudoconvex optimization problem. *Neural Networks*, *124*, 180–192.
Yan, Z., Fan, J., & Wang, J. (2017). A collective neurodynamic approach to constrained global optimization. *IEEE Transactions on Neural Networks and Learning Systems*, *28*, 1206–1215.
Yan, Z., Wang, J., & Li, G. (2014). A collective neurodynamic optimization approach to bound-constrained nonconvex optimization. *Neural Networks*, *55*, 20–29.
Yang, S., Liu, Q., & Wang, J. (2018). A collaborative neurodynamic approach to multiple-objective distributed optimization. *IEEE Transactions on Neural Networks and Learning Systems*, *29*, 981–992.
Young, M. R. (1998). A minimax portfolio selection rule with linear programming solution. *Manag. Sci.*, *44*, 673–683.
Zhang, S., Xia, Y. S., & Zheng, W. (2015). A complex-valued neural dynamical optimization approach and its stability analysis. *Neural Networks*, *61*, 59–67.
Zhao, Y., Liao, X., & He, X. (2022). Novel projection neurodynamic approaches for constrained convex optimization. *Neural Networks*, *150*, 336–349.
Zopounidis, C., Galariotis, E., Doumpos, M., Sarri, S., & AndriosopouloS, K. (2015). Multiple criteria decision aiding for finance: An updated bibliographic survey. *European Journal of Operational Research*, *247*, 339–348.