IEEE Access
Multidisciplinary : Rapid Review : Open Access Journal

Received December 8, 2020, accepted December 22, 2020, date of publication December 30, 2020, date of current version January 7, 2021.

Digital Object Identifier 10.1109/ACCESS.2020.3047967

A New Data-Driven Distributionally Robust Portfolio Optimization Method Based on Wasserstein Ambiguity Set

NINGNING DU[1], YANKUI LIU[2], AND YING LIU[1]
[1]Risk Management and Financial Engineering Laboratory, College of Mathematics and Information Science, Hebei University, Baoding 071002, China
[2]Key Laboratory of Machine Learning and Computational Intelligence, College of Mathematics and Information Science, Hebei University, Baoding 071002, China

Corresponding author: Ying Liu (yingliu@hbu.edu.cn)

This work was supported in part by the National Natural Science Foundation of China under Grant 61773150 and Grant 61374184, and in part by the Advanced Talents Incubation Program of the Hebei University under Grant 801260201209.

ABSTRACT Since optimal portfolio strategy depends heavily on the distribution of uncertain returns, this article proposes a new method for the portfolio optimization problem with respect to distribution uncertainty. When the distributional information of the uncertain return rate is only observable through a finite sample dataset, we model the portfolio selection problem with a robust optimization method from the data-driven perspective. We first develop an ambiguous mean-CVaR portfolio optimization model, where the ambiguous distribution set employed in the distributionally robust model is a Wasserstein ball centered within the empirical distribution. In addition, the computationally tractable equivalent model of the worst-case expectation under the uncertainty set of a cone is derived, and some theoretical conclusions of the box, budget and ellipsoid uncertainty set are obtained. Finally, to demonstrate the effectiveness of our mean-CVaR portfolio optimization method, two practical examples concerning the Chinese stock market and United States stock market are considered. Furthermore, some numerical experiments are carried out under different uncertainty sets. The proposed data-driven distributionally robust portfolio optimization method offers some advantages over the ambiguity-free stochastic optimization method. The numerical experiments illustrate that the new method is effective.

INDEX TERMS Portfolio optimization, data-driven, mean-CVaR, Wasserstein metric, distributionally robust.

I. INTRODUCTION

The portfolio optimization problem was studied by Markowitz [1] in 1952, who first proposed a systematic method of mean-variance to quantify portfolio return and risk. The mean-variance methodology has become the most popular way to solve the problem of portfolio optimization. Moreover, it was subsequently expanded into a pioneering book [2]. For more recent researches on portfolio optimization, the reader may refer to [3]–[5], which provides a useful reference for handling portfolio selection problems for both researchers and practitioners. The core of modern portfolio theory involves balancing return and risk, determining an effective portfolio strategy, and allocating capital to multiple available assets, to minimize risk while maximizing returns [6]. In the case of maximizing the portfolio return while minimizing risk, the optimal portfolio strategy depends heavily on the underlying hypothesis on the uncertainty of asset return and the choice of risk measure.

In general, the expectation is used to describe return, and a widely accepted measure of risk is value at risk (VaR) [7], [8]. However, to overcome the limitations of VaR, such as the lack of subadditivity, Rockafellar and Uryasev [9] proposed a modified version of conditional value at risk (CVaR), which was defined as the mean of the tail distribution exceeding VaR. Further, Alexander and Baptista [10] demonstrated that a CVaR constraint is tighter than a VaR constraint if the CVaR and VaR bounds coincide for a given confidence. Ben-Tal *et al.* [11] illustrated that the CVaR approximation of the probability constraint is the

The associate editor coordinating the review of this manuscript and approving it for publication was Tachun Lin.

best among the approximations yielded by our generating-function-based approximation scheme. CVaR is currently a popular risk measure suggested by theoreticians and market practitioners [12], [13].

To accommodate the uncertainty of return in the portfolio optimization model, we usually use mathematical methods to analyze the historical data in the market and attempt to establish an optimization model. However, in the context of massive data in the financial market, identifying how to use historical data reasonably and effectively is the key to solving the portfolio optimization problem [14]. Decision makers hope to obtain available information about the future return of stocks through historical data and to describe the distribution information of uncertain return as much as possible. To avoid the loss caused by unsuccessful investment, several optimization methods have been proposed. According to the way in which the uncertainty is described, these methods can primarily be classified into two groups: stochastic optimization and robust optimization. Stochastic optimization takes advantage of known distribution information, which governs the random return. However, it is often impossible to obtain the full knowledge of future returns and determine the true distributions in reality: instead, they should be estimated from historical data. Moreover, the distribution of forecast errors is inferred from the historical data, and therefore, sufficient historical data are needed but are nearly always unavailable [11]. Such difficult unpredictability provides a strong incentive for an investor to adopt distributionally robust optimization.

Distributionally robust optimization is a modeling program, adopting the worst-case approach, where the worst case is chosen from a proscribed ambiguity distribution set. Distributionally robust optimization research was first proposed by Scarf [15], and much progress has been made in recent years [16]–[19]. Among these researchers, Chang *et al.* [16] investigated a distributionally robust scheduling problem on identical parallel machines based on a distributional set, specified by the support and estimated moment information. Delage and Ye [17] proposed a model that described uncertainty in both the distribution form (discrete, Gaussian, exponential, etc.) and as moments (mean and covariance matrix). Xu *et al.* [19] proposed two discretization schemes for solving the distributionally robust optimization: one for the dual approach and the other directly through the ambiguity set. More recently, the distributionally robust optimization method has also become increasingly popular in the portfolio field with respect to ambiguous sets. Liu and Liu [20] developed a novel parametric credibilistic optimization method for the project portfolio selection problem. Kang *et al.* [14] presented a computationally tractable optimization method for a robust mean-CVaR portfolio selection model under the condition of distribution ambiguity. Rujeerapaiboon *et al.* [21] designed fixed-mix strategies that offered similar performance guarantees as the growth-optimal portfolio by using methods from distributionally robust optimization. Van Parys *et al.* [22] discussed a stock portfolio pricing problem, where the distributional information was limited to second-order moment information in conjunction with structural information.

The key ingredient is to choose the ambiguity distribution set for the distributionally robust optimization model. A good ambiguous set should better facilitate a tractable representation of the distributionally robust optimization model as a structured mathematical problem that can be solved with optimization software. Scholars usually construct a distribution set according to large-scale historical data. It is necessary for a good ambiguous set to contain the true data-generating distribution with high probability guarantee.

The data-driven method has been subsequently proposed. For the time being, there are several data-driven approaches for distributionally robust optimization to construct ambiguous sets: see, for example, [23], [24] and the references therein. Sample average approximation is a widely used method for data-driven decision making under the condition of uncertainty [25]. For further understanding regarding the data-driven method of distributionally robust optimization, interested readers can refer to the following literature. Çetinkaya and Thiele [26] investigated an iterative data-driven approximation of a problem where the investor sought to maximize the expected return of his or her portfolio, subject to a quantile constraint, given historical realizations of stock returns. Fernandes *et al.* [27] provided a new perspective on robust portfolio optimization, wherein they imposed an intuitive loss constraint for the optimal portfolio considering asset returns in a data-driven polyhedral uncertainty set. Chi *et al.* [28] proposed a data-driven robust model of portfolio optimization with relative entropy constraints based on an instance-based credit risk assessment framework.

Data-driven studies concern data analysis and various statistical methods to construct ambiguous sets. Bertsimas *et al.* [29] presented a new, systematic schema for constructing uncertainty sets from data using statistical hypothesis tests, including the χ^2-test, *G*-test, and Kolmogorov-Smirnov test. Wang *et al.* [30] adopted likelihood estimation to structure the ambiguous set. Gupta [31] proposed a Bayesian framework for assessing the relative strengths of data-driven ambiguity sets in distributionally robust optimization. Jiang and Guan [32] and Zhao and Zhang [33] formulated data-driven optimization models with phi-divergence.

The distributionally robust optimization program with the Wasserstein ambiguity set can often be reformulated as a finite convex program. The Wasserstein ambiguity set contains all distributions that satisfy certain convex constraints. It is close to a nominal or most likely distribution with respect to the prescribed probability metric. The Wasserstein ambiguity set was first used by Pflug and Wozabal [34] in the context of portfolio selection optimization. Recently, research on the Wasserstein metric has become increasingly popular. Hanasusanto and Kuhn [35] demonstrated that the distributionally robust linear program was equivalent to a co-positive program or that it could be approximated arbitrarily closely by a sequence of co-positive programs when the ambiguity set was composed of a 2-Wasserstein ball centered at a discrete distribution. Carlsson *et al.* [36] considered a distributionally robust version of the Euclidean traveling salesman problem, in which the authors computed the worst-case spatial distribution of demand against all distributions whose Wasserstein distances to an observed demand distribution were bounded from above. Duan *et al.* [37] proposed an optimal power flow model that minimized the expectation of the quadratic cost function with respect to the worst-case probability distribution and guaranteed that the chance constraints were satisfied for any distribution in the ambiguity set with the Wasserstein ball centered at the empirical distribution.

To some extent, our work in this article has a relationship with the research [39], which supposes a polytope uncertainty set, while our paper develops a more general form of uncertainty set, i.e., the uncertainty set of random return is cone. There is no discussion about the Wasserstein ambiguity set under a cone representation in the existing literature, which motivates us to address this issue in this article. Due to the difficulty of obtaining the optimal portfolio strategy, the semi-infinite portfolio optimization problem is transformed into a finite convex programming problem under the convexity assumption, in which the support set of random return is a cone. It is worth noting that the equivalent models derived under various uncertainty sets are also different. As a result, three kinds of equivalent models are obtained under the conditions in which the uncertainty set is a box, budget or ellipsoid. Furthermore, some practical cases pertaining to the Chinese stock market are considered to illustrate our new modeling ideas.

In comparison with the exiting literature, the contributions to the area of distributionally robust portfolio optimization and management include the following summarized aspects.

1) A new data-driven mean-CVaR portfolio optimization framework for the distributionally robust linear optimization model is developed, in which the ambiguous set employed in the distributionally robust model is the Wasserstein ball centered at the empirical distribution and the empirical distribution is based on a sample dataset.
2) In our new model, the worst-case expectation reduces to the optimal value of a linear convex program when the cone representation is used in the proposed Wasserstein ambiguous set. To obtain the computationally tractable linear convex programming model, some theoretical results are discussed in detail. Specifically, three equivalent models under the cones of box, budget and ellipsoid are derived.
3) To demonstrate the effectiveness of our proposed model, we address some practical cases pertaining to the Chinese stock market and analyze the parameters of the model through the computational results. To demonstrate the advantage of our new method, we compare the proposed data-driven distributionally robust portfolio optimization method with the traditional ambiguity-free stochastic optimization method via numerical experiments.
4) A general modeling method for portfolio optimization problem is presented in this article. The model is also applicable to stock markets in other countries, except for the Chinese stock market. Further, some numerical experiments based on the United States stock market are performed to demonstrate the generality of the model.

This article is structured as follows. Section II presents mean-CVaR portfolio optimization for the distributionally robust linear optimization model. Moreover, we develop the computationally tractable equivalent representation of the robust mean-CVaR portfolio model with respect to the Wasserstein ambiguous set. In Section III, a small-scale practical case study is addressed to demonstrate the validity of proposed equivalent models, and the models are based on the box and budget support sets. The computational results of the two models under different parameters are analyzed and discussed. Furthermore, a large-scale numerical experiment is presented in Section IV. The comparison of data-driven distributionally robust portfolio optimization method and the ambiguity-free stochastic optimization method is included in this section. Section V concludes the paper.

For simplicity, the following notations and variables are employed throughout this article to develop the data-driven mean-CVaR portfolio optimization model. $\boldsymbol{\xi} \in R^m$ represents the random return vector of m stocks. $\boldsymbol{x} \in R^m$ denotes the portfolio allocation proportion in m stocks. $r(\cdot,\cdot)$ is the function of portfolio return. $loss(\cdot)$ is the expectation of loss function. $\mathrm{P} \in \mathcal{P}$ is the distribution of random vector $\boldsymbol{\xi}$, where $\mathcal{P}$ is a set of family distributions. The support set of distribution family $\mathcal{P}$ is Ξ, which is the available information set for describing the distribution family. $\hat{\Xi}_N$ denotes the dataset comprising N independent samples of stock returns. $\hat{\mathrm{P}}_N$ is a discrete empirical probability distribution supported by $\hat{\Xi}_N$. $d_w(\cdot,\cdot)$ refers to the Wasserstein metric. ϵ denotes the distance of Wasserstein metric, that is, the radius of Wasserstein ball. $\boldsymbol{K}$ is a closed convex pointed cone with a nonempty interior, and $\boldsymbol{K}_*$ is its dual cone. The semidefinite cone of size n, denoted by $\boldsymbol{S}_+^n$. $\boldsymbol{H}_+^n$ denotes the Hermitian semidefinite cone of size n. A is a data matrix of dimension $n \times m$, B is a data matrix of dimension $n \times k$, and $\boldsymbol{c}$ is an n-dimensional vector. $Diag(\cdot)$ represents a diagonal matrix with the elements of the vector on the main diagonal. Given a norm $||\boldsymbol{\xi}||_p = (\sum_{i=1}^m |\xi_i|^p)^{\frac{1}{p}}$ on R^m, the dual norm is defined by $||\boldsymbol{\xi}||_q = (\sum_{i=1}^m |\xi_i|^q)^{\frac{1}{q}}$, where $p, q \neq 0$ and $1/p+1/q = 1$. $\hat{\boldsymbol{\xi}}_i$ is a m-dimensional vector, which represents the i-th realized value of random return vector for $i = 1, \ldots, N$. The conjugate of $h(\boldsymbol{\xi})$ is defined as $h^*(\boldsymbol{y}) = \sup_{\boldsymbol{\xi} \in R^m} \boldsymbol{y}^{\mathrm{T}}\boldsymbol{\xi} - h(\boldsymbol{\xi})$. δ_Ξ denotes the conjugate of the characteristic function of support Ξ. φ_i and ψ_i are $m+1$ dimensional decision vectors for $i = 1, \ldots, N$. In addition, Λ, Γ and Ω are the parameters of uncertainty set of box, budget and ellipsoid respectively. α is referred to as the risk level coefficient of CVaR. η denotes the weight coefficient.

II. PORTFOLIO OPTIMIZATION PROBLEM UNDER WASSERSTEIN METRIC

The portfolio optimization problem pertains to allocating capital over a number of available stocks in order to maximize the return on the investment while minimizing the risk. In this section, we will study the portfolio optimization problem with respect to the Wasserstein metric. A new robust mean-CVaR portfolio optimization model is established firstly, and then, the computationally tractable equivalent form is derived with respect to the Wasserstein metric.

A. ROBUST MEAN-CVaR PORTFOLIO OPTIMIZATION MODEL

Investors select m stocks from the Chinese stock market whose returns are captured by the random vector $\boldsymbol{\xi} = (\xi_1, \ldots, \xi_m)^{\mathrm{T}}$. The vector of percentage weights $\boldsymbol{x} = (x_1, \ldots, x_m)^{\mathrm{T}}$ denotes the amounts of investment in m stocks, and

$$\boldsymbol{x} \in D = \{\sum_{j=1}^{m} x_j = 1,\ \boldsymbol{x} \geq 0\}.$$

When portfolio $\boldsymbol{x}$ invests the percentage of available capital x_j in stock j for $j = 1, \ldots, m$, its return equals

$$r(\boldsymbol{x}, \boldsymbol{\xi}) = \boldsymbol{x}^{\mathrm{T}}\boldsymbol{\xi}.$$

Moreover, $-r(\boldsymbol{x}, \boldsymbol{\xi})$ is called a loss function.

Investors always attempt to distribute their capital to different stocks in the hope of maximizing returns. However, the return of a stock is affected by various factors, such as economics, politics, law and culture. Under the influence of such multitudinous factors, investors usually employ the expected value of return to measure the quality of portfolio allocation strategies. A natural approach to generate optimal values r is to approximate the distribution P of random vector $\boldsymbol{\xi}$ with a discrete empirical probability distribution, which is also a data-driven view [39].

The expected value of the loss function is

$$loss(\boldsymbol{\xi}) = \mathrm{E}^{\mathrm{P}}[-r(\boldsymbol{x}, \boldsymbol{\xi})], \tag{1}$$

where P is the distribution of the stock return. When the distribution P of random vector $\boldsymbol{\xi}$ is obtained, the optimal approximation of Eq. (1) is closely related to the CVaR. It is defined as

$$\mathrm{CVaR}_\alpha(-r) = \min_{e \in R}\{e + \frac{1}{\alpha}\mathrm{E}^{\mathrm{P}}\{\max[-r(\boldsymbol{x}, \boldsymbol{\xi}) - e, 0]\}, \tag{2}$$

the min in the right-hand side of Eq. (2) is attained, and

$$\Pr\{-r(\boldsymbol{x}, \boldsymbol{\xi}) > \mathrm{CVaR}_\alpha(-r)\} \leq \alpha,$$

where $\alpha \in (0, 1)$ is referred to as the risk level of the CVaR with respect to the distribution P. The distribution information of random vector $\boldsymbol{\xi}$ is always described by historical data in parctice. However, the estimated CVaR may contain considerable estimation errors with limited historical data. It is very difficult for investor to deal with the worst-case situation over the set of probability distributions, which is defined by the limited information available [14]. In order to effectively deal with the problem, the worst-case CVaR risk measure with distribution ambiguity is provided in this article. When the parameter of risk level $\alpha \in (0, 1)$ is given, the worst-case CVaR of portfolio $\boldsymbol{x} = (x_1, \ldots, x_m)^{\mathrm{T}}$ is defined by

$$\max_{\mathrm{P} \in \mathcal{P}}\ e + \frac{1}{\alpha}\mathrm{E}^{\mathrm{P}}\{\max[-r(\boldsymbol{x}, \boldsymbol{\xi}) - e, 0]\},$$

the distribution of random vector $\boldsymbol{\xi}$ belongs to an ambiguity distribution set $\mathcal{P}$. Thus, we obtain the following distributionally robust bi-objective portfolio optimization model:

$$\begin{aligned} &\min_{\boldsymbol{x}}\ \max_{\mathrm{P} \in \mathcal{P}}\ \mathrm{E}^{\mathrm{P}}[-r(\boldsymbol{x}, \boldsymbol{\xi})] \\ &\min_{\boldsymbol{x}, e}\ \max_{\mathrm{P} \in \mathcal{P}}\ e + \frac{1}{\alpha}\mathrm{E}^{\mathrm{P}}\{\max[-r(\boldsymbol{x}, \boldsymbol{\xi}) - e, 0]\} \\ &\text{s.t. } \boldsymbol{x} \in D, \quad e \in R. \end{aligned} \tag{3}$$

To find the optimal portfolio strategy or the optimal value of return, we first solve the bi-objective optimization model (3) by using a weighted sum method. The two objective functions of model (3) can then be combined into the following single objective function with weight $\eta \in [0, 1]$,

$$\begin{aligned} &\min_{\boldsymbol{x} \in D, e \in R}\ \max_{\mathrm{P} \in \mathcal{P}}\ \{\eta \mathrm{E}^{\mathrm{P}}[-r(\boldsymbol{x}, \boldsymbol{\xi})] + (1 - \eta)(e \\ &\qquad + \frac{1}{\alpha}\mathrm{E}^{\mathrm{P}}\{\max[-r(\boldsymbol{x}, \boldsymbol{\xi}) - e, 0]\})\} \\ &= \min_{\boldsymbol{x} \in D, e \in R}\ \max_{\mathrm{P} \in \mathcal{P}}\ \{\mathrm{E}^{\mathrm{P}}[\max\{-(\frac{1 - \eta}{\alpha} + \eta)\boldsymbol{x}^{\mathrm{T}}\boldsymbol{\xi} \\ &\qquad + e(1 - \eta)(1 - \frac{1}{\alpha}), -\eta\boldsymbol{x}^{\mathrm{T}}\boldsymbol{\xi} + e(1 - \eta)\}]\}. \end{aligned}$$

While the weight coefficient η transforms the dual-objective function model (3) into a single-objective model, it expresses an attitude of investors towards loss and risk. Put it differently, the values of η indicate whether investors pay more attention to loss or risk. The weight coefficient η is also called risk preference parameter, which quantifies the investor's risk-aversion [39]. Therefore, the robust mean-CVaR portfolio optimization model can be reexpressed as

$$\begin{aligned} &\min_{\boldsymbol{x}, e}\ \max_{\mathrm{P} \in \mathcal{P}}\ \mathrm{E}^{\mathrm{P}}[f(\boldsymbol{x}, \boldsymbol{\xi})] \\ &\text{s.t. } \boldsymbol{x} \in D, \quad e \in R, \end{aligned} \tag{4}$$

where

$$\begin{aligned} f(\boldsymbol{x}, \boldsymbol{\xi}) = \max\{&-(\frac{1 - \eta}{\alpha} + \eta)\boldsymbol{x}^{\mathrm{T}}\boldsymbol{\xi} + e(1 - \eta)(1 - \frac{1}{\alpha}), \\ &-\eta\boldsymbol{x}^{\mathrm{T}}\boldsymbol{\xi} + e(1 - \eta)\}, \end{aligned}$$

and the distribution family $\mathcal{P}$ is supported on Ξ.

From the model (4), we can observe that it has a finite number of decision variables, and the inner maximization over probability distribution of the worst-case expectation program is infinite dimensional. In order to better understand model (4), we introduce an artificial variable t and rewrite the model into the following equivalent form:

$$\begin{aligned}&\min_{\boldsymbol{x},e}\ t\\ &\text{s.t. } \mathrm{E}^{\mathrm{P}}[f(\boldsymbol{x},\boldsymbol{\xi})]\le t,\quad \forall \mathrm{P}\in\mathcal{P},\\ &\quad \boldsymbol{x}\in D,\quad e\in R.\end{aligned}\tag{5}$$

It is obvious that the model (5) is a semi-infinite programming with a finite number of variables and an infinite number of constraints. In optimization theory, many optimization problems involve some set of variables and some set of constraints. If the optimization problem has a finite number of variables and an infinite number of constraints, or an infinite number of variables and a finite number of constraints, then the optimization problem is called semi-infinite programming [38]. Therefore, the robust mean-CVaR portfolio optimization model (4) is equivalent to a semi-infinite programming. It is necessary to determine the structure of the ambiguous distribution set before solving it. In the existing literature, there are many ambiguous sets based on different measures, such as ϕ-divergence [31] and Kullback-Leibler divergence [32]. In recent years, some researchers ([17] and [39]) try to solve the problem by means of the Wasserstein metric. In the following subsection, we will develop a more general form of the uncertainty set.

B. EQUIVALENT MODEL UNDER WASSERSTEIN METRIC

In this subsection, the equivalent form of the robust mean-CVaR portfolio optimization model will be derived with respect to the Wasserstein metric. In practice, the exact distribution information of random variables must be determined according to a large amount of historical data. Due to the noise and outliers in the data or the imperfect data of newly listed assets, determining the exact distributions of future returns is often a difficult assignment.

This is not to suggest that there are no successful solutions. Note that the distribution P is not precisely known and is often partially observable through a finite set of N independent samples of past realizations of the random vector $\boldsymbol{\xi}$. The dataset comprising these samples is denoted by $\hat{\Xi}_N = \{\hat{\boldsymbol{\xi}}_i\}_{i\le N} \subseteq \Xi$. To generate data-driven solutions, we approximate the distribution P with a discrete empirical probability distribution $\hat{\mathrm{P}}_N$ supported by $\hat{\Xi}_N$. Thus,

$$\min_{\boldsymbol{x}\in D,e\in R}\{\mathrm{E}^{\hat{\mathrm{P}}_N}[f(\boldsymbol{x},\boldsymbol{\xi})]=\frac{1}{N}\sum_{i=1}^{N}f(\boldsymbol{x},\hat{\boldsymbol{\xi}}_i)\}.\tag{6}$$

This formulation is a widely popular method for data-driven decision making under the condition of uncertainty [29].

Next, model (4) will be addressed by an alternative method of distributionally robust optimization that can resist the uncertainty of distribution. Specifically, we use $\hat{\Xi}_N$ to design an ambiguity set $\mathcal{P}_N$, which contains all distributions generated by these samples with high probability level, i.e.,

$$\mathcal{P}_N(\mathrm{P})=\{\mathrm{P}\text{ is supported on }\Xi: d_w(\mathrm{P},\hat{\mathrm{P}}_N)\le\epsilon\}.\tag{7}$$

Kantorovich and Rubinstein [40] originally established the following result for distributions with bounded support

$$d_w(\mathrm{P},\hat{\mathrm{P}}_N)=\sup_g\{\int_\Xi g(\xi)\mathrm{P}(d\xi)-\int_\Xi g(\xi)\hat{\mathrm{P}}_N(d\xi)\},$$

where g indicates Lipschitz functions with $|g(\xi^1)-g(\xi^2)|\le ||\xi^1-\xi^2||$ for all $\xi^1,\xi^2\in\Xi$.

The ambiguity set $\mathcal{P}_N(\mathrm{P})$ can be viewed as a Wasserstein ball of radius ϵ centered at the empirical distribution $\hat{\mathrm{P}}_N$, named the Wasserstein ambiguous set. The Wasserstein metric is a widely used probability metric, which is represented by distance functions on the space of probability distributions. Moreover, it enables us to define the following data-driven distributionally robust optimization problem that minimizes the worst-case expected value:

$$\min_{\boldsymbol{x}\in D,e\in R}\ \max_{\mathrm{P}\in\mathcal{P}_N(\mathrm{P})}\ \mathrm{E}^{\mathrm{P}}[f(\boldsymbol{x},\boldsymbol{\xi})].\tag{8}$$

Obviously, the worst-case expectation program constitutes an infinite-dimensional optimization problem over probability distributions, and it appears to be intractable. To help investors obtain the optimal portfolio allocation proportions, it is necessary to effectively address the infinite-dimensional optimization problem. How on earth should we solve this problem? In the existing literature, the problem of infinite dimensions can be transformed into finite convex programming under the certain convexity assumption [39], in which the conclusion of convex reduction inspires us to seek other computationally tractable forms for data-driven distributionally robust portfolio selection optimization models. In the following, we will demonstrate that the data-driven distributionally robust portfolio optimization model (8) can be represented as a finite-dimensional convex problem, which can be solved via existing commercial optimization software. A major result of computational tractability of the data-driven distributionally robust portfolio optimization program is provided in the following theorem.

Theorem 1: Assume that the uncertainty set is a cone representation, that is, $\Xi=\{\boldsymbol{\xi}\in R^m:\exists\boldsymbol{u}\in R^k: A\boldsymbol{\xi}+B\boldsymbol{u}+\boldsymbol{c}\in\boldsymbol{K}\}$. $\boldsymbol{K}$ *is a closed convex pointed cone with a nonempty interior,* A, B *are fixed matrices and* $\boldsymbol{c}$ *is a fixed vector. Moreover, consider the function* $f(\boldsymbol{x},\boldsymbol{\xi})=\max\{h(\boldsymbol{\xi}),l(\boldsymbol{\xi})\}$, *where*

$$\begin{aligned}h(\boldsymbol{\xi})&=-(\frac{1-\eta}{\alpha}+\eta)\boldsymbol{x}^{\mathrm{T}}\boldsymbol{\xi}+e(1-\eta)(1-\frac{1}{\alpha}),\\ l(\boldsymbol{\xi})&=-\eta\boldsymbol{x}^{\mathrm{T}}\boldsymbol{\xi}+e(1-\eta).\end{aligned}$$

Then the worse-case expectation of (8) equals to

$$\begin{aligned}&\min_{\boldsymbol{x},e,\lambda,s_i,\varphi_i,\psi_i}\ \lambda\epsilon+\frac{1}{N}\sum_{i=1}^{N}s_i\\ &s.t.\ -(\frac{1-\eta}{\alpha}+\eta)\boldsymbol{x}^{\mathrm{T}}\hat{\boldsymbol{\xi}}_i+e(1-\eta)(1-\frac{1}{\alpha})\\ &\quad+\varphi_i^{\mathrm{T}}(\boldsymbol{c}+A\hat{\boldsymbol{\xi}}_i)\le s_i\quad\forall i\le N\\ &\quad-\eta\boldsymbol{x}^{\mathrm{T}}\hat{\boldsymbol{\xi}}_i+e(1-\eta)+\psi_i^{\mathrm{T}}(\boldsymbol{c}+A\hat{\boldsymbol{\xi}}_i)\\ &\quad\le s_i\quad\forall i\le N\end{aligned}$$

$$\begin{aligned}
& ||-A^{\mathrm{T}}\varphi_i+(\frac{1-\eta}{\alpha}+\eta)\boldsymbol{x}||_* \leq \lambda \quad \forall i \leq N \\
& ||-A^{\mathrm{T}}\psi_i+\eta\boldsymbol{x}||_* \leq \lambda \quad \forall i \leq N \\
& B^{\mathrm{T}}\varphi_i=0, B^{\mathrm{T}}\psi_i=0 \quad \forall i \leq N \\
& \varphi_i \in \boldsymbol{K}_*, \psi_i \in \boldsymbol{K}_* \quad \forall i \leq N \\
& \boldsymbol{x} \in D, \quad e \in R, \qquad (9)
\end{aligned}$$

where $||\cdot||_*$ *is the dual norm of* $||\cdot||$.

Proof: According to Theorem 4.2 [39] about the result of convex reduction, we have

$$\begin{aligned}
\min_{\boldsymbol{x},e,\lambda,s_i,y_i,z_i,a_i,b_i} \ & \lambda\epsilon+\frac{1}{N}\sum_{i=1}^{N}s_i, \\
\text{s.t. } & [-h(\boldsymbol{\xi})]^*(y_i-a_i)+\delta_\Xi(a_i)-y_i^{\mathrm{T}}\hat{\boldsymbol{\xi}}_i \leq s_i, \\
& [-l(\boldsymbol{\xi})]^*(z_i-b_i)+\delta_\Xi(b_i)-z_i^{\mathrm{T}}\hat{\boldsymbol{\xi}}_i \leq s_i, \\
& ||y_i||_* \leq \lambda,\ ||z_i||_* \leq \lambda,\ \forall i \leq N, \qquad (10)
\end{aligned}$$

where $[-h(\boldsymbol{\xi})]^*(y_i-a_i)$ is the conjugate of $-h(\boldsymbol{\xi})$ evaluated at (y_i-a_i), $[-l(\boldsymbol{\xi})]^*(z_i-b_i)$ is the conjugate of $-l(\boldsymbol{\xi})$ evaluated at (z_i-b_i), and δ_Ξ represents the conjugate of the characteristic function of support Ξ.

By the definition of the conjugate operator, we have

$$[-h(\boldsymbol{\xi})]^*(y)=\sup_{\boldsymbol{\xi}} y^{\mathrm{T}}\boldsymbol{\xi}+(\frac{1-\eta}{\alpha}+\eta)\boldsymbol{x}^{\mathrm{T}}\boldsymbol{\xi}-e(1-\eta)(1-\frac{1}{\alpha}),$$

when $[-h(\boldsymbol{\xi})]^*$ equals to 0 if y is $-(\frac{1-\eta}{\alpha}+\eta)\boldsymbol{x}$, and equals to ∞ otherwise. Moreover,

$$[-l(\boldsymbol{\xi})]^*(z)=\sup_{\boldsymbol{\xi}} z^{\mathrm{T}}\boldsymbol{\xi}+\eta\boldsymbol{x}^{\mathrm{T}}\boldsymbol{\xi}-e(1-\eta)$$

and $[-l(\boldsymbol{\xi})]^*$ equals to 0 if z is $-\eta\boldsymbol{x}$, and equals to ∞ otherwise.

In addition,

$$\begin{aligned}
\delta_\Xi(v) &= \sup_{\boldsymbol{\xi},v}\{v^{\mathrm{T}}\boldsymbol{\xi}: A\boldsymbol{\xi}+B\boldsymbol{u}+\boldsymbol{c} \in \boldsymbol{K}\} \\
&= \inf_{\gamma}\{\boldsymbol{c}^{\mathrm{T}}\gamma: B^{\mathrm{T}}\gamma=0, A^{\mathrm{T}}\gamma=-v, \gamma \in \boldsymbol{K}_*\},
\end{aligned}$$

where the second equality follows from the conic duality theorem, which holds as the uncertainty set is non-empty. As a consequence, model (9) follows by substituting the above expressions into to model (10). □

Remark 1: Theorem 1 presents an equivalent form of the loss function's worst-case expectation under uncertainty set Ξ *of a cone* $\boldsymbol{K}$*, where* $\boldsymbol{K}$ *is a given convex cone. It considers a rather general case when the uncertainty set* Ξ *is given by a conic representation* $\Xi=\{\boldsymbol{\xi}\in R^m: \exists\boldsymbol{u}\in R^k: A\boldsymbol{\xi}+B\boldsymbol{u}+\boldsymbol{c}\in \boldsymbol{K}\}$*. Theorem 1 indicates that as long as the uncertainty set can be expressed as a cone, an equivalent tractable form of a robust portfolio optimization model can be obtained on this basis. The conclusion presented in Theorem 1 is general. To the best of our knowledge, there are four common choices of the conic representation* $\boldsymbol{K}$*, including the simplest one-dimensional cone, Lorentz cone, semidefinite cone, and Hermitian semidefinite cone. Specifically, we characterize different cones in the following four cases.*

(i) When the cone $\boldsymbol{K}$ *is the simplest one-dimensional cone, that is,* $\boldsymbol{K}=\{q\in R^n: q_i\geq 0,\ i=1,\ldots,n\}$*. Thus, we can obtain the uncertainty set*

$$\Xi=\{\boldsymbol{\xi}\in R^m: \exists\boldsymbol{u}\in R^k: q=A\boldsymbol{\xi}+B\boldsymbol{u}+\boldsymbol{c}: \quad q_i\geq 0,\ i=1,\ldots,n\}.$$

The simplest one-dimensional cone is regular and self-dual, so its dual conic $\boldsymbol{K}_*=\boldsymbol{K}$*. In addition, if* $A\boldsymbol{\xi}=[\boldsymbol{\xi};0]$ *and* $\boldsymbol{c}=[0_{m\times 1};C_1]$*, then the uncertainty set* Ξ *becomes*

$$\Xi=\{[\boldsymbol{\xi};C_1]\in R^m\times R:\ \xi_j\geq 0,\ C_1\geq 0,\ j=1,\ldots,m\}.$$

If $A\boldsymbol{\xi}=[\Sigma^{-1}\boldsymbol{\xi};0]$ *with* $\Sigma=Diag\{\Delta_1',\ldots,\Delta_m'\}$*,* $\boldsymbol{c}=[0_{m\times 1};C_2]$*, then the uncertainty set* Ξ *can be reformulated*

$$\Xi=\{[\Sigma^{-1}\boldsymbol{\xi};C_2]\in R^m\times R:\ \frac{\xi_j}{\Delta_j'}\geq 0,\ C_2\geq 0,\ j=1,\ldots,m\}.$$

(ii) When the cone $\boldsymbol{K}$ *is Lorentz cone, that is,* $\boldsymbol{K}=\{[q;t]\in R^n\times R: t\geq\sqrt{\sum_{i=1}^n q_i^2}\}$*. Therefore, the uncertainty set* Ξ *is as follows*

$$\Xi=\{\boldsymbol{\xi}\in R^m: \exists\boldsymbol{u}\in R^k: [q;t]=A\boldsymbol{\xi}+B\boldsymbol{u}+\boldsymbol{c}: \quad t\geq\sqrt{\sum_{i=1}^n q_i^2}\}.$$

Note that the Lorentz cone is regular and self-dual, its dual conic $\boldsymbol{K}_*=\boldsymbol{K}$*. If* $A\boldsymbol{\xi}=[\boldsymbol{\xi};0]$ *and* $\boldsymbol{c}=[0_{m\times 1};C_1]$*, we can obtain the uncertainty set*

$$\Xi=\{[\boldsymbol{\xi};C_1]\in R^m\times R:\ C_1\geq\sqrt{\sum_{j=1}^m \xi_j^2}\}.$$

Moreover, if $A\boldsymbol{\xi}=[\Sigma^{-1}\boldsymbol{\xi};0]$ *with* $\Sigma=Diag\{\Delta_1',\ldots,\Delta_m'\}$*,* $\boldsymbol{c}=[0_{m\times 1};C_2]$*, then the uncertainty set* Ξ *can be reexpressed*

$$\Xi=\{[\Sigma^{-1}\boldsymbol{\xi};C_2]\in R^m\times R:\ C_2\geq\sqrt{\sum_{j=1}^m(\frac{\xi_j}{\Delta_j'})^2}\}.$$

(iii) When the cone $\boldsymbol{K}$ *is semidefinite cone, that is,* $\boldsymbol{K}=\{[q;t]\in \boldsymbol{S}^n\times R: tI_n-q\in \boldsymbol{S}_+^n\}$*. Thus, we have*

$$\Xi=\{\boldsymbol{\xi}\in R^m: \exists\boldsymbol{u}\in R^k: [q;t]=A\boldsymbol{\xi}+B\boldsymbol{u}+\boldsymbol{c}: \quad tI_n-q\in \boldsymbol{S}_+^n\}.$$

The semidefinite cone is regular and self-dual, its dual conic $\boldsymbol{K}_*=\boldsymbol{K}$*. Further, the uncertainty set* Ξ *can be rewritten as*

$$\Xi=\{[\boldsymbol{\xi};C_1]\in \boldsymbol{S}^m\times R:\ C_1I_m-\boldsymbol{\xi}\in \boldsymbol{S}_+^m\},$$

where $A\boldsymbol{\xi}=[\boldsymbol{\xi}_{m\times m};0]$ *and* $\boldsymbol{c}=[0_{m\times m};C_1]$*. Moreover, if* $A\boldsymbol{\xi}=[\Sigma^{-1}\boldsymbol{\xi}_{m\times m};0]$ *with* $\Sigma=Diag\{\Delta_1',\ldots,\Delta_m'\}$*,* $\boldsymbol{c}=[0_{m\times m};C_2]$*, then the uncertainty set* Ξ *can be reexpressed*

$$\Xi=\{[\Sigma^{-1}\boldsymbol{\xi};C_2]\in \boldsymbol{S}^m\times R:\ C_2I_m-\Sigma^{-1}\boldsymbol{\xi}\in \boldsymbol{S}_+^m\}.$$

(iv) When the cone $\boldsymbol{K}$ is Hermitian semidefinite cone, that is, $\boldsymbol{K} = \{[q; t] \in \boldsymbol{H}^n \times R : tI_n - q \in \boldsymbol{H}^n_+\}$. Thus, the uncertainty set is

$$\Xi = \{\boldsymbol{\xi} \in R^m : \exists \boldsymbol{u} \in R^k : [q; t] = A\boldsymbol{\xi} + B\boldsymbol{u} + \boldsymbol{c} : tI_n - q \in \boldsymbol{H}^n_+\}.$$

Same as the one-dimensional cone, Lorentz cone and semidefinite cone, the Hermitian semidefinite cone is regular and self-dual, its dual conic $\boldsymbol{K}_* = \boldsymbol{K}$. Further, when $A\boldsymbol{\xi} = [\boldsymbol{\xi}_{m\times m}; 0]$ and $\boldsymbol{c} = [0_{m\times m}; C_1]$, the uncertainty set Ξ can be rewritten as

$$\Xi = \{[\boldsymbol{\xi}; C_1] \in \boldsymbol{H}^m \times R : \ C_1 I_m - \boldsymbol{\xi} \in \boldsymbol{H}^m_+\}.$$

Moreover, the uncertainty set Ξ can be reformulated

$$\Xi = \{[\Sigma^{-1}\boldsymbol{\xi}; C_2] \in \boldsymbol{H}^m \times R : \ C_2 I_m - \Sigma^{-1}\boldsymbol{\xi} \in \boldsymbol{H}^m_+\},$$

where $A\boldsymbol{\xi} = [\Sigma^{-1}\boldsymbol{\xi}_{m\times m}; 0]$ with $\Sigma = Diag\{\Delta'_1, \ldots, \Delta'_m\}$, $\boldsymbol{c} = [0_{m\times m}; C_2]$.

In the following, we will build tractable reformulations of the portfolio selection model (8) based on Theorem 1. Three commonly used uncertainty sets for cone representations, i.e., box, budget, and ellipsoid uncertainty sets, are discussed in detail. The equivalent models under these uncertainty sets are presented, and some effective numerical experiments are carried out in the ensuing sections.

Setting $\varphi_i = [\upsilon_i^1; \tau_i^1]$, $\psi_i = [\upsilon_i^2; \tau_i^2]$ with m-dimensional υ_i and one-dimensional τ_i for each $i \leq N$. The equivalent model with respect to the box uncertainty set is as follows in Corollary 1.

Corollary 1: Let the uncertainty set Ξ be a box, that is, $\Xi = \{\boldsymbol{\xi} \in R^m : |\xi_j| \leq \Lambda,\ j = 1, 2, \ldots, m\}$. The model (9) can be reduced to

$$\begin{aligned}
\min_{\boldsymbol{x}, e, \lambda, s_i, \upsilon_i, \tau_i} \quad & \lambda\epsilon + \frac{1}{N}\sum_{i=1}^{N} s_i \\
s.t. \quad & -(\frac{1-\eta}{\alpha} + \eta)\boldsymbol{x}^{\mathrm{T}}\hat{\boldsymbol{\xi}}_i + e(1-\eta)(1-\frac{1}{\alpha}) \\
& + (\upsilon_i^1)^{\mathrm{T}}\hat{\boldsymbol{\xi}}_i + \Lambda\tau_i^1 \leq s_i \quad \forall i \leq N \\
& -\eta\boldsymbol{x}^{\mathrm{T}}\hat{\boldsymbol{\xi}}_i + e(1-\eta) + (\upsilon_i^2)^{\mathrm{T}}\hat{\boldsymbol{\xi}}_i + \Lambda\tau_i^2 \\
& \leq s_i \quad \forall i \leq N \\
& || - \upsilon_i^1 + (\frac{1-\eta}{\alpha} + \eta)\boldsymbol{x}||_* \leq \lambda \quad \forall i \leq N \\
& || - \upsilon_i^2 + \eta\boldsymbol{x}||_* \leq \lambda \quad \forall i \leq N \\
& \tau_i^1 \geq ||\upsilon_i^1||_1,\ \tau_i^2 \geq ||\upsilon_i^2||_1 \quad \forall i \leq N \\
& \boldsymbol{x} \in D, \quad e \in R.
\end{aligned} \tag{11}$$

Proof: For the box uncertainty set, its cone representation is as follows:

$$\Xi = \{\boldsymbol{\xi} \in R^m : A\boldsymbol{\xi} + \boldsymbol{c} \in \boldsymbol{K}\},$$

where

$A\boldsymbol{\xi} = [\boldsymbol{\xi}; 0], \boldsymbol{c} = [0_{m\times 1}; \Lambda]$, and $\boldsymbol{K} = \{[\omega; \pi] \in R^m \times R : \pi \geq ||\omega||_\infty\}$, its dual conic $\boldsymbol{K}_* = \{[\omega; \pi] \in R^m \times R : \pi \geq ||\omega||_1\}$.

By substituting the above expression into model (9), we can get model (11). □

Furthermore, the equivalent forms of distributionally robust portfolio optimization model (9) are derived when the cone $\boldsymbol{K}$ is assumed to be a budget uncertainty set. The theoretical result is summarized as the following Corollary 2.

Corollary 2: Let the uncertainty set be budget, i.e., $\Xi = \{\boldsymbol{\xi} \in R^m : \sum_{j=1}^m |\frac{\xi_j}{\Delta_j}| \leq \Gamma\}$. The equivalent representation of model (9) is

$$\begin{aligned}
\min_{\boldsymbol{x}, e, \lambda, s_i, \upsilon_i, \tau_i} \quad & \lambda\epsilon + \frac{1}{N}\sum_{i=1}^{N} s_i \\
s.t. \quad & -(\frac{1-\eta}{\alpha} + \eta)\boldsymbol{x}^{\mathrm{T}}\hat{\boldsymbol{\xi}}_i + e(1-\eta)(1-\frac{1}{\alpha}) \\
& + (\upsilon_i^1)^{\mathrm{T}}Q^{-1}\hat{\boldsymbol{\xi}}_i + \Gamma\tau_i^1 \leq s_i \quad \forall i \leq N \\
& -\eta\boldsymbol{x}^{\mathrm{T}}\hat{\boldsymbol{\xi}}_i + e(1-\eta) + (\upsilon_i^2)^{\mathrm{T}}Q^{-1}\hat{\boldsymbol{\xi}}_i + \Gamma\tau_i^2 \\
& \leq s_i \quad \forall i \leq N \\
& || - Q^{-1}\upsilon_i^1 + (\frac{1-\eta}{\alpha} + \eta)\boldsymbol{x}||_* \leq \lambda \quad \forall i \leq N \\
& || - Q^{-1}\upsilon_i^2 + \eta\boldsymbol{x}||_* \leq \lambda \quad \forall i \leq N \\
& \tau_i^1 \geq ||\upsilon_i^1||_\infty,\ \tau_i^2 \geq ||\upsilon_i^2||_\infty \quad \forall i \leq N \\
& \boldsymbol{x} \in D, \quad e \in R.
\end{aligned} \tag{12}$$

Proof: Budget uncertainty set Ξ has the conic representation as follows:

$$\Xi = \{\boldsymbol{\xi} \in R^m : A\boldsymbol{\xi} + \boldsymbol{c} \in \boldsymbol{K}\},$$

where, $A\boldsymbol{\xi} = [Q^{-1}\boldsymbol{\xi}; 0]$ with $Q = Diag\{\Delta_1, \ldots, \Delta_m\}$, $\boldsymbol{c} = [0_{m\times 1}; \Gamma]$ and $\boldsymbol{K} = \{[\omega; \pi] \in R^m \times R : \pi \geq ||\omega||_1\}$, whence $\boldsymbol{K}_* = \{[\omega; \pi] \in R^m \times R : \pi \geq ||\omega||_\infty\}$.

Model (12) can be obtained by substituting the above expression into model (9). □

The ellipsoid uncertainty set is significantly different for the box and budget uncertainty sets since its constraints are nonlinear. Therefore, according to Theorem 1, the distributionally robust portfolio optimization model (9) is equivalent to a conic quadratic optimization model. It is presented in Corollary 3.

Corollary 3: Let Ξ be an ellipsoid uncertainty set, that is, $\Xi = \{\boldsymbol{\xi} \in R^m : \sqrt{\sum_{j=1}^m (\frac{\xi_j}{\Delta'_j})^2} \leq \Omega\}$. The model (9) can be reduced to

$$\begin{aligned}
\min_{\boldsymbol{x}, e, \lambda, s_i, \upsilon_i, \tau_i} \quad & \lambda\epsilon + \frac{1}{N}\sum_{i=1}^{N} s_i \\
\text{s.t.} \quad & -(\frac{1-\eta}{\alpha} + \eta)\boldsymbol{x}^{\mathrm{T}}\hat{\boldsymbol{\xi}}_i + e(1-\eta)(1-\frac{1}{\alpha}) \\
& + (\upsilon_i^1)^{\mathrm{T}}\Sigma^{-1}\hat{\boldsymbol{\xi}}_i + \Omega\tau_i^1 \leq s_i \quad \forall i \leq N \\
& -\eta\boldsymbol{x}^{\mathrm{T}}\hat{\boldsymbol{\xi}}_i + e(1-\eta) + (\upsilon_i^2)^{\mathrm{T}}\Sigma^{-1}\hat{\boldsymbol{\xi}}_i + \Omega\tau_i^2 \\
& \leq s_i \quad \forall i \leq N \\
& || - \Sigma^{-1}\upsilon_i^1 + (\frac{1-\eta}{\alpha} + \eta)\boldsymbol{x}||_* \leq \lambda \quad \forall i \leq N \\
& || - \Sigma^{-1}\upsilon_i^2 + \eta\boldsymbol{x}||_* \leq \lambda \quad \forall i \leq N \\
& \tau_i^1 \geq ||\upsilon_i^1||_2,\ \tau_i^2 \geq ||\upsilon_i^2||_2 \quad \forall i \leq N \\
& \boldsymbol{x} \in D, \quad e \in R.
\end{aligned} \tag{13}$$

TABLE 1. The assets from Chinese stock market.

Stock code	Return	Stock code	Return	Stock code	Return	Stock code	Return	Stock code	Return
000039	ξ_1	000685	ξ_5	002001	ξ_9	600028	ξ_{13}	600422	ξ_{17}
000151	ξ_2	000789	ξ_6	002024	ξ_{10}	600176	ξ_{14}	600519	ξ_{18}
000488	ξ_3	000886	ξ_7	600000	ξ_{11}	600276	ξ_{15}	600624	ξ_{19}
000521	ξ_4	000966	ξ_8	600016	ξ_{12}	600340	ξ_{16}	600737	ξ_{20}

TABLE 2. The daily returns of stock 000966 from March 1, 2019 to April 30, 2019.

Date	Return	Date	Return	Date	Return	Date	Return	Date	Return	Date	Return
2019-3-01	0.9905	2019-3-04	0.9976	2019-3-05	1.0120	2019-3-06	1.0143	2019-3-07	0.9953	2019-3-08	0.9341
2019-3-11	1.0302	2019-3-12	1.0367	2019-3-13	1.0024	2019-3-14	0.9741	2019-3-15	1.0942	2019-3-18	0.9691
2019-3-19	0.9886	2019-3-20	0.9931	2019-3-21	1.0116	2019-3-22	1.0023	2019-3-25	0.9863	2019-3-26	0.9675
2019-3-27	0.9904	2019-3-28	0.9806	2019-3-29	1.0198	2019-4-01	1.0387	2019-4-02	0.9977	2019-4-03	1.0047
2019-4-04	1.0140	2019-4-08	1.0413	2019-4-09	1.0396	2019-4-10	1.0106	2019-4-11	1.0419	2019-4-12	1.1006
2019-4-15	1.0073	2019-4-16	0.9945	2019-4-17	1.0294	2019-4-18	1.0993	2019-4-19	1.0226	2019-4-22	0.9448
2019-4-23	0.9666	2019-4-24	1.0190	2019-4-25	0.9407	2019-4-26	1.0252	2019-4-29	0.9684	2019-4-30	1.0018

Proof: For the ellipsoid uncertainty set, it also has the following conic representation:

$$\Xi = \{\boldsymbol{\xi} \in R^m : A\boldsymbol{\xi} + \boldsymbol{c} \in \boldsymbol{K}\}.$$

Moreover, $A\boldsymbol{\xi} = [\Sigma^{-1}\boldsymbol{\xi}; 0]$ with $\Sigma = Diag\{\Delta'_1, \ldots, \Delta'_m\}$, $\boldsymbol{c} = [0_{m\times 1}; \Omega]$ and $\boldsymbol{K} = \{[\omega; \pi] \in R^m \times R : \pi \geq ||\omega||_2\}$, whence $\boldsymbol{K}_* = \boldsymbol{K}$.

By substituting the above expression into model (9), the model (13) is obtained. □

III. SMALL-SCALE EXPERIMENTS BASED ON DAILY RETURN IN 2019

In this section, we conduct some small-scale experiments based on real market data of the daily returns of twenty stocks over two months by utilizing the distributionally robust portfolio optimization model. All of the numerical experiments in this article were performed on a personal computer with an Inter(R) Core(TM)i5-4200U 1.60GHz CPU and 4.00GB of RAM. To better demonstrate the feasibility and effectiveness of the proposed models, the LINGO 11.0.0.20 solver is employed to find the robust optimal portfolio strategy.

A. DATA AND METHODOLOGY

To perform the experiments, twenty stocks from the Chinese stock market were selected by the *Straight Flush* stock software. The stock codes are summarized in Table 1, and the stock return $\boldsymbol{\xi} = (\xi_1, \ldots, \xi_{20})$ is a twenty-dimensional vector. It is generally known that the stock market is an extremely unstable market. Stock returns are affected by various factors, for instance, macro factors, corporate factors, and market factors. Among them, macro factors include economic, political and legal factors; corporate factors refer to the impacts of the operations of listed companies on stock prices; and market factors refer to various stock market operations that can affect stock market prices. Most importantly, the economic or noneconomic factors will all cause the drift of stock returns, and so the return $\xi_j, j = 1, \ldots, 20$ of the j-th stock is uncertain. Fig. 1 depicts the changes in the daily return of Stock 000966. It is easily deduced that stock returns

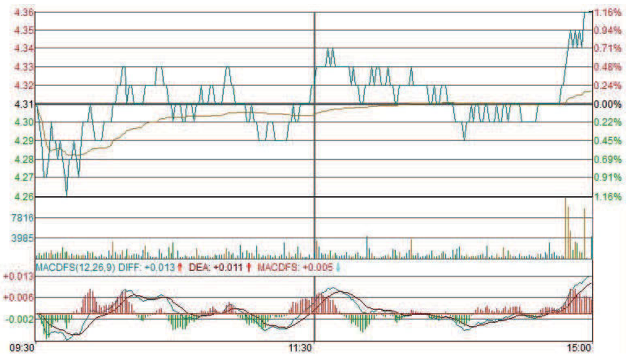


FIGURE 1. The return of stock 000966 on March 21, 2019.

nearly always change between opening quotation and closing quotation.

The daily returns of twenty stocks are recorded as the dataset for numerical experiments. The sample period of these historical return values is March 1 to April 30, 2019, for 42 observations in total, i.e., $N = 42$, as shown in Table 2. The daily returns of twenty investments over two months are set according to the closing quotation. For example, for stock 000966, ¥1 invested at 9:30 a.m. on March 5, 2019, was worth ¥1.0120 by 15:00 p.m., and therefore, its daily return was ¥1.0120.

Because of the limited number of samples, it is impossible to determine the true distribution of stock return without ambiguity. In this article, the ambiguous distribution of stock return is estimated via some observations from datasets. More specifically, the central distribution of Wasserstein ambiguity sets is a discrete empirical distribution $\hat{\mathbb{P}}_N$, and the radius ϵ of the Wasserstein ambiguity set represents the inaccuracy of the distribution. We derive some theoretical results in Section II under the proposed Wasserstein ambiguity set. Moreover, to obtain the optimal portfolio strategy, models (11) and (12) are solved in this subsection.

B. COMPUTATIONAL RESULTS

Some key parameters, such as Λ, ϵ, η and α, are involved in model (11), and they exert different effects on the

TABLE 3. The optimal portfolio strategy under $\epsilon = 0.01$, $\eta = 0.5$, and $\alpha = 0.10$ (%).

Λ	1.2	4.8	6.0	7.2	8.4	9.6	10.8	12.0
x_1	2.976204	4.394472	4.943657	4.650317	4.858534	4.815788	4.913306	5.471090
x_2	2.965601	8.818534	4.704177	4.650317	4.771632	5.222669	4.851538	4.881853
x_3	2.965657	4.519900	4.586700	4.660966	6.647131	4.794253	4.851538	4.879159
x_4	3.340908	4.394472	5.159345	4.669246	4.729523	4.789746	4.838185	4.879904
x_5	3.197912	4.421565	5.011433	4.658084	4.728980	4.801079	4.838583	4.877491
x_6	3.001048	5.003542	5.137018	4.725399	5.731329	4.802767	4.838478	4.899426
x_7	3.024683	5.003542	4.676808	4.831958	4.858534	4.922057	4.838918	4.882799
x_8	2.940102	4.519960	4.622595	4.724456	4.771817	5.724801	4.897965	5.044488
x_9	3.919016	4.545974	4.636528	4.717379	4.728980	4.795661	7.092431	4.877491
x_{10}	2.967574	4.655389	4.585562	4.692910	4.858534	4.850743	4.875529	4.881367
x_{11}	3.765285	4.394472	4.679400	4.670955	4.728980	4.789746	4.848014	4.877491
x_{12}	3.026080	4.578533	4.601638	4.702595	4.839872	4.977108	4.838099	4.899426
x_{13}	3.605691	4.394472	4.877672	4.652348	4.728980	4.789746	5.016953	5.044829
x_{14}	2.965771	4.396468	6.909021	4.653174	4.776359	4.827693	4.909952	4.878434
x_{15}	2.988334	8.902303	7.609615	10.70304	4.858534	6.771708	4.874156	6.172648
x_{16}	3.811005	4.892129	4.626486	4.650317	4.728980	4.807466	4.839009	4.915638
x_{17}	3.330827	4.416168	4.714498	4.664371	6.168784	4.815788	4.865565	4.882776
x_{18}	3.735161	4.695019	4.615981	4.898420	4.771541	4.854810	5.219781	4.877835
x_{19}	38.51903	4.446156	4.685885	4.754577	4.737719	5.035916	4.859821	4.882749
x_{20}	2.954105	4.606932	4.615981	4.669167	4.975256	4.810453	4.892179	4.993107

computational results. In this subsection, the optimal portfolio strategy and optimal value of model (11) under different parameters are discussed.

1) THE INFLUENCE OF UNCERTAINTY SET PARAMETER Λ

As shown in Corollary 1, parameter Λ limits the possible values of stock returns, i.e., $|\xi_j| \leq \Lambda, j = 1, \ldots 20$ when the support set of uncertain return is a box. This means that Λ is larger than the maximum return of these twenty stocks. According to the data in Table 2, several values of Λ are selected for small-scale experiments: $\Lambda = 1.2, 4.8, 6.0, 7.2, 8.4, 9.6, 10.8$ and 12.0. The computational results of optimal portfolio allocation proportions under different values of Λ are reported in Table 3. To better illustrate the change of the optimal portfolio allocation under different Λ, Figs. 2 and 3 are depicted.

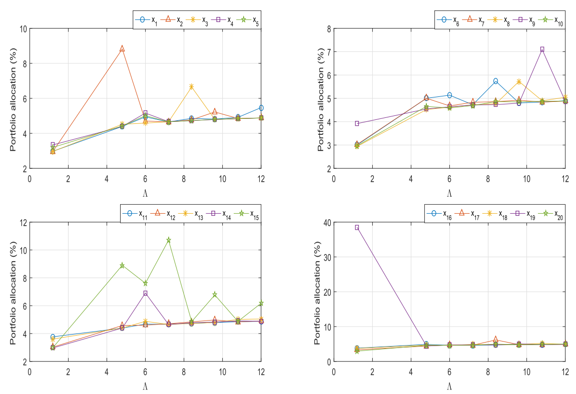


FIGURE 2. The variation of optimal portfolio strategy under $\epsilon = 0.01$, $\eta = 0.5$ and $\alpha = 0.10$.

Figs. 2 and 3 show the variation of portfolio allocation ratios of twenty stocks with respect to Λ. From Fig. 2, we can identify that the proportions of stocks 000039 (x_1), 000685 (x_5), 002024 (x_{10}), 600000 (x_{11}), 600016 (x_{12}), 600028 (x_{13}),

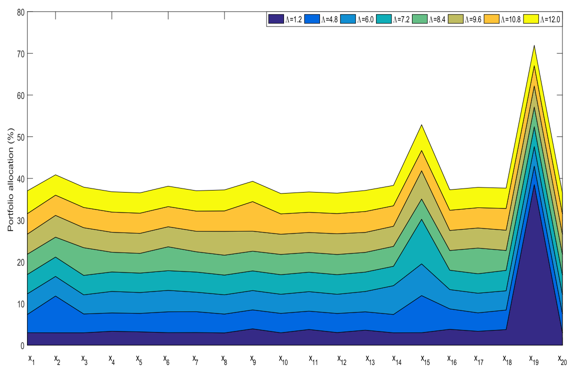


FIGURE 3. The variation of optimal portfolio strategy under $\epsilon = 0.01$, $\eta = 0.5$ and $\alpha = 0.10$.

600340 (x_{16}), 600519 (x_{18}) and 600737 (x_{20}) are gradually increasing and that the ratios of stock 600624 (x_{19}) are decreasing with respect to Λ. Fig. 3 illustrates that when $\Lambda = 1.2$, stock 600624 (x_{19}) has the largest proportion of portfolio allocation, while stock 600276 (x_{15}) has the largest proportion of portfolio allocation at $\Lambda = 4.8, 6.0, 7.2, 9.6$. When $\Lambda = 8.4$, stock 000488 (x_3) has the largest portfolio allocation proportion; if $\Lambda = 10.8$, stock 002001 (x_9) has the largest portfolio allocation proportion. Further, Fig. 4 is plotted to show the effect of parameter Λ on the optimal value of model (11). It is easy to find that the optimal value increases with the growth of the value of parameter Λ. In other words, as the value of parameter Λ becomes increasingly larger, the solution of the model (11) becomes more conservative with respect to the box uncertainty set.

To avoid duplicate conclusions, the box uncertainty set parameter is set to 1.2 (i.e., $\Lambda = 1.2$) in the following subsections.

2) THE INFLUENCE OF WASSERSTEIN RADIUS ϵ

Parameter ϵ denotes the radius of the Wasserstein ambiguity set, and its value determines the size of a Wasserstein ball.

TABLE 4. The optimal portfolio strategy under $\alpha = 0.10$, $\eta = 0.5$ (%).

ϵ	0.01	0.02	0.03	0.04	0.05	0.06	0.07	0.08
x_1	2.976204	3.951587	2.963765	2.939149	2.991845	3.309891	3.169332	3.515813
x_2	2.965601	2.992168	2.939149	4.203487	3.039914	2.942468	2.939149	6.890562
x_3	2.965657	3.288256	3.052666	2.944713	8.659118	3.037573	3.840007	11.83841
x_4	3.340908	4.042624	2.939149	4.097494	3.290864	3.120648	3.125103	3.012482
x_5	3.197912	3.011502	2.939149	2.973343	9.686530	3.500562	3.031927	4.254210
x_6	3.001048	4.399264	42.06525	2.939149	28.46627	3.456723	4.490958	3.852786
x_7	3.024683	3.238451	2.939149	2.939149	3.117229	2.992683	20.93686	3.199627
x_8	2.940102	3.317422	2.939149	3.584060	2.939405	3.423311	3.031927	3.607515
x_9	3.919016	31.14320	3.540777	3.285968	2.939149	3.160954	3.045047	22.10261
x_{10}	2.967574	2.992053	2.953170	2.939149	3.002092	3.651296	11.33026	3.338597
x_{11}	3.765285	2.959755	2.939149	4.676267	2.991781	35.07696	3.100161	4.017087
x_{12}	3.026080	3.030901	2.939149	2.939149	2.976742	3.039262	3.130043	4.082240
x_{13}	3.605691	3.378074	2.963765	3.045181	3.916631	4.172911	3.129617	4.044420
x_{14}	2.965771	3.394952	3.134436	2.939149	2.991859	5.256308	5.193533	3.076732
x_{15}	2.988334	3.339688	2.939597	37.62889	3.177529	3.494965	2.939306	3.305319
x_{16}	3.811005	9.031736	3.073066	3.284166	2.950766	3.338797	3.002374	2.939273
x_{17}	3.330827	3.289645	2.939149	2.939464	3.558539	3.549499	2.947500	3.649777
x_{18}	3.735161	3.063461	3.359929	3.249906	2.991815	3.052170	2.941810	3.106210
x_{19}	38.51903	2.985077	3.299465	3.513017	2.979821	3.263863	2.939333	3.103163
x_{20}	2.954105	3.150184	3.140925	2.939149	3.332096	3.159152	11.73575	3.063172

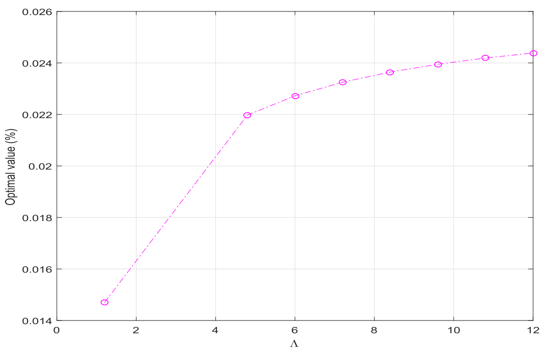


FIGURE 4. The optimal value under $\epsilon = 0.01$, $\eta = 0.5$ and $\alpha = 0.10$.

To put it differently, the value of the radius implies the number of distributions contained in Wasserstein ambiguity set. There are more distributions in the Wasserstein ambiguity set when ϵ is larger. To get a relatively appropriate Wasserstein radius, we refer to some related literature, such as the literature [39], to select the values of radius ϵ for numerical experiments. Esfahani and Kuhn [39] plotted a figure to visualize the corresponding optimal portfolio proportion as a function of the radius, where the value of radius belonged to some intervals ([0.001, 0.01], [0.01, 0.1], and [0.1, 1]). To better show the changes in the portfolio allocation proportions of different stocks under the radius ϵ, we focus on the situation when ϵ belongs to the interval [0.01, 0.1]. In the numerical experiment, we make ϵ equal to $0.01, 0.02, \ldots, 0.08$. The computational results imply that a slight change in ϵ exerts a great impact on the solution of the model. Therefore, model (11) is very sensitive to parameter ϵ. We refer to reference [39] for the selection of ϵ. The computational results of portfolio strategies with different values of ϵ are reported in Table 4. The results show that the optimal portfolio allocations vary according to the radius of the Wasserstein ambiguity set, and we obtain a relatively decentralized portfolio strategy.

Based on Table 4, we plot Fig. 5, which represents the variation of portfolio allocation proportion of each stock with respect to Wasserstein ball radius ϵ. In Fig. 5, the allocation proportions of stocks 000789 (x_6), 002001 (x_9), 600000 (x_{11}), 600276 (x_{15}) and 600624 (x_{19}) change obviously, but the variations in allocation proportions of the other stocks are not as obvious. Moreover, when $\epsilon = 0.01$, stock 600000 (x_{11}) accounts for the largest allocation proportion; 002001 (x_9) accounts for the largest allocation proportion at $\epsilon = 0.02$ and 0.08; 000789 (x_6) accounts for the largest allocation proportion at $\epsilon = 0.03$ and 0.05; 600276 (x_{15}) accounts for the largest allocation proportion at $\epsilon = 0.04$; 600000 (x_{11}) accounts for the largest allocation proportion at $\epsilon = 0.06$; and 000886 (x_7) accounts for the largest allocation proportion at $\epsilon = 0.07$.

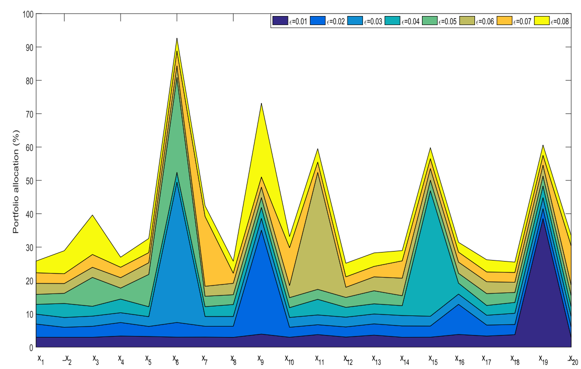


FIGURE 5. The variation of optimal portfolio strategy with different ϵ under $\eta = 0.5$, $\alpha = 0.10$.

Furthermore, Fig. 6 is depicted to illustrate the variation tendency of objective value. The optimal objective value is proportional to the radius of the Wasserstein ambiguity set, and its value is enlarged according to the increase of radius. As the radius of the Wasserstein ambiguity set becomes larger, the computational results become more conservative.

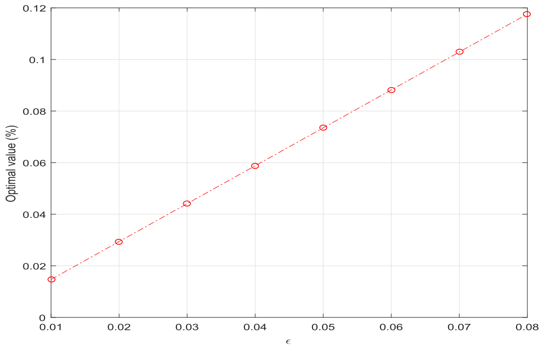


FIGURE 6. The optimal objective value under $\eta = 0.5$, $\alpha = 0.10$.

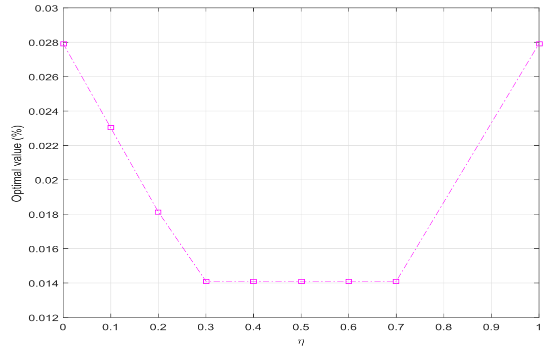


FIGURE 8. The optimal value at weight coefficient η under $\epsilon = 0.01$, $\alpha = 0.02$.

3) THE INFLUENCE OF WEIGHT COEFFICIENT η

Parameter η is the weight coefficient of the objective function, and its value indicates whether investors pay more attention to loss or risk. If $\eta > 0.5$, investors want to reduce investment losses as much as possible; if $\eta < 0.5$, investors want to reduce the risk of investment as much as possible. In addition, we consider two critical cases: risk aversion, i.e., $\eta = 1$, and risk neutrality i.e., $\eta = 0$.

To illustrate the variation of portfolio allocation proportions under different weight coefficients, Fig. 7 is depicted as follows. This figure shows that the portfolio allocation proportions of stocks are relatively decentralized under different parameter values of weight coefficient η. Investors will invest in twenty stocks in different proportions. More specifically, stock 000039 (x_1) accounts for the largest portfolio allocation proportion at $\eta = 0.1$; stock 000151 (x_2) accounts for the largest portfolio allocation proportion at $\eta = 0.5$; and stock 000488 (x_3) accounts for the largest portfolio allocation proportion at $\eta = 0.2$. When $\eta = 0$, stock 000886 (x_7) has the largest proportion of portfolio allocation, and when $\eta = 1.0$, stock 002024 (x_{10}) has the largest proportion of portfolio allocation. Further, the relationship between the optimal value and the parameter of weight coefficient η is shown in Fig. 8.

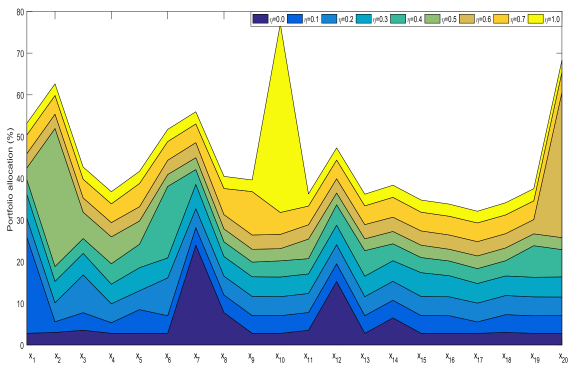


FIGURE 7. The optimal portfolio allocation proportion under $\epsilon = 0.01$, $\alpha = 0.02$.

4) THE INFLUENCE OF RISK LEVEL α

As defined in Eq. (2), the parameter $\alpha \in (0, 1)$ is referred to as the risk level of the CVaR. In Subsections III-B1 and III-B2, the results are reported at risk level $\alpha = 0.10$, while the results in Subsection III-B3 is computed at risk level $\alpha = 0.02$. Taking into account the effect of risk level on optimal portfolio allocation proportions and objective optimal value, some experiments in α are conducted. Table 5 presents the optimal portfolio allocation proportion of model (11) under different risk level α. To better demonstrate the variation of optimal portfolio strategy at different parameter values of risk level α, Fig. 9 is plotted. It illustrates that the change range of allocation proportions of stocks 000039 (x_1), 000685 (x_5), 002001 (x_9), 600016 (x_{12}), 600276 (x_{15}) and 600624 (x_{19}) is relatively large under different risk levels, while the change amplitudes of allocation proportions of the other stocks are relatively small. More specifically, stock 000151 (x_2) has the largest allocation proportion at $\alpha = 0.02$; stock 002001 (x_9) accounts for the largest allocation proportion at $\alpha = 0.04$; stock 600276 (x_{15}) accounts for the largest allocation proportion at $\alpha = 0.06, 0.16$; stock 600016 (x_{12}) accounts for the largest allocation proportion at $\alpha = 0.08$; stock 600624 (x_{19}) accounts for the largest allocation proportion at $\alpha = 0.10$; stock 000685 (x_5) accounts for the largest allocation proportion at $\alpha = 0.12$; and stock 000039 (x_1) accounts for the largest allocation proportion at $\alpha = 0.14$. Furthermore, Fig. 10 is plotted, which shows the optimal value of model (11) under different risk level α. When the risk level α increases, the optimal value of the model also increases. This means that the higher the risk level α is, the more conservative the solution of the model.

C. COMPARISONS BETWEEN BOX AND BUDGET UNCERTAINTY SET

In the previous section, the optimal portfolio allocation strategy and optimal value of the equivalent model are discussed with respect to the box uncertainty set with different parameters, including the coefficient of the box uncertainty set Λ, the radius ϵ of the Wasserstein ambiguity set, weight coefficient η and risk level α.

Without loss of generality, we set the parameter $\Delta_j = 1$, $j = 1, \ldots, m$ in model (12). Note that the constraints on the box and budget uncertainty set are both linear. The characteristic of the box uncertainty set is to limit the random return

TABLE 5. The optimal portfolio strategy under $\epsilon = 0.01$, $\eta = 0.5$ (%).

α	0.02	0.04	0.06	0.08	0.10	0.12	0.14	0.16
x_1	2.832709	2.876451	8.861273	2.908532	2.976204	2.970416	27.80199	6.739332
x_2	33.10918	2.929182	2.898184	2.912650	2.965601	3.363480	3.002356	8.272558
x_3	6.307053	2.902684	2.905716	2.910512	2.965657	2.970416	3.369332	3.119678
x_4	6.446282	2.873964	2.882727	2.908532	3.340908	3.158075	15.91269	3.098962
x_5	5.572005	2.849175	2.884507	2.908532	3.197912	42.309630	3.022422	3.365627
x_6	2.848075	2.896690	4.839665	2.908532	3.001048	2.970416	3.002356	9.865580
x_7	2.848075	2.972441	3.482349	2.936580	3.024683	3.183287	3.029955	3.837392
x_8	3.099534	2.883756	5.309054	2.961425	2.940102	2.970416	3.002356	3.073519
x_9	3.074734	45.02231	2.898184	3.005750	3.919016	2.970416	3.003618	3.035406
x_{10}	2.921174	2.928237	2.906018	3.095361	2.967574	2.970416	4.437327	3.095719
x_{11}	4.703948	2.877402	3.633882	3.095361	3.765285	2.970416	3.180531	3.128576
x_{12}	2.821858	2.936795	6.792914	42.77828	3.026080	2.970416	3.050273	3.034990
x_{13}	2.826561	2.886913	2.906056	3.586691	3.605691	3.058588	3.050273	3.038108
x_{14}	2.959979	2.849175	6.843275	3.584286	2.965771	2.970416	3.057987	3.299401
x_{15}	2.967767	2.927330	17.04878	2.935063	2.988334	2.970416	3.002356	19.89509
x_{16}	2.821305	2.863852	8.889146	2.928134	3.811005	2.970416	3.002356	4.815444
x_{17}	3.046383	2.857880	3.633882	2.910056	3.330827	3.032179	3.025451	3.462592
x_{18}	3.069726	2.950834	2.913649	2.908532	3.735161	2.970416	3.019825	4.455632
x_{19}	2.848075	2.851516	4.564762	2.908656	38.51903	3.279354	3.007252	3.060484
x_{20}	2.875580	2.863416	2.905980	2.908532	2.954105	2.970416	3.019299	4.305908

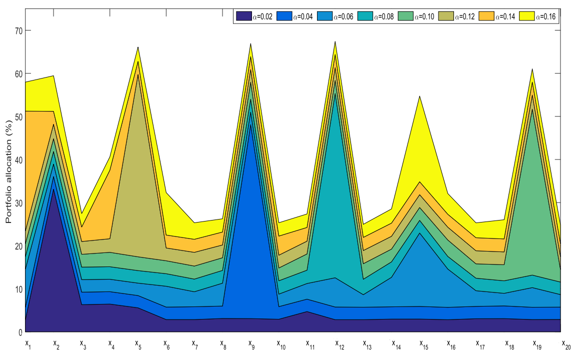


FIGURE 9. The variation of optimal portfolio strategy of model (11) under different risk level α.

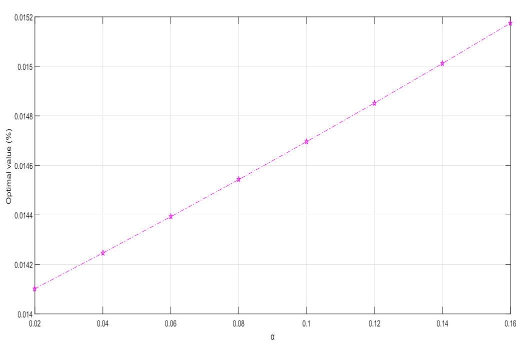


FIGURE 10. The optimal value of model (11) under different risk level α.

of each stock to within a fixed interval, i.e., $|\xi_j| < \Lambda, j = 1, \ldots, 20$. In contrast, the budget uncertainty set limits the total return of twenty stocks to within a certain interval, that is, $\sum_{j=1}^{20} \xi_j < \Gamma$. From this point of view, if $\Gamma < 20\Lambda$, then the computational results under the box uncertainty set are more conservative than those under the budget uncertainty set, while if $\Gamma > 20\Lambda$, then the computational results under the budget uncertainty set are more conservative than those under the box uncertainty set.

To confirm the above assertion, some numerical experiments are performed, and Fig. 11 is plotted. It shows the optimal objective values of model (11) and (12) at different risk levels with $\Delta_j = 1, j = 1, \ldots, 20$. It reveals that the objective value increases with the escalation of risk level. In other words, as the risk level α becomes increasingly higher, the solutions of model (11) and (12) become more conservative. Moreover, when $\Gamma = 10$ and $\Gamma = 20$, model (11) is more conservative than model (12); in contrast, when $\Gamma = 36$ and $\Gamma = 46$, model (12) is more conservative.

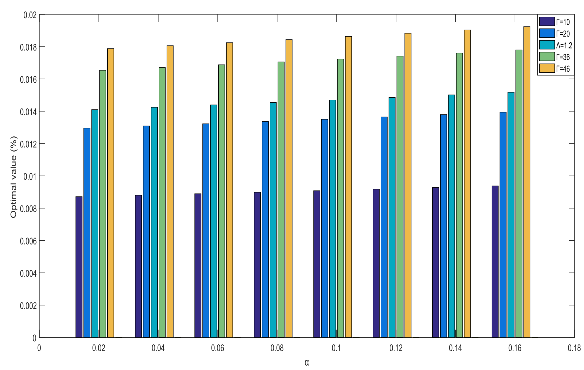


FIGURE 11. Comparisons of the objective values of models (11) and (12) under $\eta = 0.5$, $\epsilon = 0.01$.

IV. LARGE-SCALE EXPERIMENTS BASED ON MONTHLY RETURN OF TEN YEARS

To illustrate the proposed model and method in the data-driven context, this section addresses the portfolio optimization problem with respect to a large set of historical data. In large-scale experiments, the following discussion intends to apply the monthly returns of stocks over ten years for solving the proposed model with respect to the ellipsoid uncertainty set.

TABLE 6. The monthly returns of stock 000521.

Date	Return	Date	Return	Date	Return	Date	Return	Date	Return	Date	Return
2009-01	1.1308	2009-02	1.2121	2009-03	1.0841	2009-04	1.0524	2009-05	1.1614	2009-06	1.0480
2009-07	1.1293	2009-08	0.8116	2009-09	1.1107	2009-10	1.3939	2009-11	1.3114	2009-12	1.1548
2010-01	0.8964	2010-02	1.0263	2010-03	0.9421	2010-04	1.0545	2010-05	0.8792	2010-06	0.8284
2010-07	1.1979	2010-08	0.9819	2010-09	1.0428	2010-10	1.1045	2010-11	0.9561	2010-12	1.0124
2011-01	0.9485	2011-02	1.1389	2011-03	0.9321	2011-04	0.9359	2011-05	0.8639	2011-06	0.9839
2011-07	0.9946	2011-08	0.7722	2011-09	0.8539	2011-10	0.9850	2011-11	0.9005	2011-12	0.8296
2012-01	0.9774	2012-02	1.1547	2012-03	0.8960	2012-04	1.1518	2012-05	1.0058	2012-06	0.8690
2012-07	0.9180	2012-08	0.8237	2012-09	1.0205	2012-10	1.0460	2012-11	0.9148	2012-12	1.1381

A. DATA AND METHODOLOGY

Similarly, the real return data from the Chinese stock market are collected. Our dataset is also extracted from the stock software *Straight Flush*, including the monthly returns of twenty stocks. The sample period of these values of stock returns is January 2009 to December 2018 for 120 observations in total, i.e., $N = 120$. The monthly return for each of the twenty stocks is obtained for ten years, and it describes the fluctuation of the return value of each stock within a month. Here, the initial return value of a month is regarded as the opening price, and the end of a month is the closing price. We choose the closing price at the end of the month as the monthly return value of the stock. Fig. 12 shows the fluctuation in the monthly return of stock 000521 from January 2009 to December 2018. Moreover, some of the monthly returns of stock 000521 are shown in Table 6. This table provides the monthly returns of stock 000521 from January 2009 to December 2012. For instance, ¥1 invested in stock 000521 on January 1, 2009, was worth ¥1.1308 on January 30, 2009.

FIGURE 12. The monthly return of stock 000521 from 2009 to 2018.

To find the optimal portfolio policy, we first derive the equivalent form of the distributionally robust portfolio optimization model (9) under the ellipsoid uncertainty set. The distributionally robust portfolio optimization model is equivalent to a conic quadratic optimization model (13). Thus, the numerical experiments in this subsection are devoted to solving model (13).

B. COMPUTATIONAL RESULTS

Based on the historical return data of stocks collected above, the LINGO 11.0.0.20 solver is employed, and the equivalent model (13) is solved in large-scale numerical experiments. To test the accuracy of the model, some sensitivity analyses on different parameters in model (13) are applied. The optimal portfolio allocation proportion and optimal value of model (13) are presented in the following subsections.

1) THE INFLUENCE OF UNCERTAINTY SET PARAMETER Ω

The radius of the ellipsoid uncertainty set Ω reflects the size of the ellipsoid. Without losing generality, we set $\Delta'_j = 1, j = 1, \ldots, m$, and the constraint in the ellipsoid uncertainty set becomes $\sqrt{\sum_{j=1}^{m} \xi_j^2} \leq \Omega$, i.e., $\sqrt{20 \max_j \xi_j^2} \leq \Omega$. Several values of ellipsoid uncertainty parameter Ω are selected for the experiments, and the portfolio allocation proportions of model (13) under different Ω are reported in Table 7.

To better illustrate the variation of optimal portfolio allocation proportions in Table 7, we depict Fig. 13, which displays the change of portfolio ratios of twenty stocks with respect to Ω. More specifically, with the increase in parameter Ω, the portfolio allocation proportions of stocks 000488 (x_3), 002001 (x_9), 002040 (x_{10}), 600000 (x_{11}), 600016 (x_{12}), 600028 (x_{13}), 600276 (x_{15}) and 600519 (x_{18}) are gradually decreasing, while those of the other stocks are increasing. Further, Fig. 14 displays the change in the optimal value of model (13) with respect to Ω. The optimal value of model (13) escalates with increasing Ω, which means that the larger Ω is, the more conservative the solution.

To avoid duplicate conclusions, the ellipsoid uncertainty set parameter is set to 6.0 (i.e. $\Omega = 6.0$) in the following subsections.

2) THE INFLUENCE OF WASSERSTEIN RADIUS ϵ

In this subsection, we solve model (13) to determine the optimal portfolio strategy under different parameter values of Wasserstein's spherical radius and report the computational results in Table 8; furthermore, Fig. 15 is depicted. From this figure, we can determine the changing trend of portfolio allocation proportions at different values of Wasserstein's spherical radius ϵ. Specifically, the proportion of each stock portfolio allocation is obviously distinct under the same

TABLE 7. The optimal portfolio strategy under $\epsilon = 0.01$, $\eta = 0.5$, and $\alpha = 0.10$ (%).

Ω	6	24	30	36	42	48	54	60
x_1	4.819722	4.954902	4.963941	4.969954	4.974254	4.977462	4.979964	4.981972
x_2	4.874888	4.968694	4.974974	4.979149	4.982134	4.984356	4.986100	4.987489
x_3	5.059102	5.014747	5.011817	5.009851	5.008249	5.007393	5.006568	5.005910
x_4	4.938854	4.984686	4.987768	4.989840	4.991272	4.992350	4.993208	4.993885
x_5	4.947160	4.986762	4.989429	4.991194	4.992459	4.993392	4.994130	4.994716
x_6	4.949764	4.987413	4.989949	4.991628	4.992831	4.993718	4.994420	4.994976
x_7	4.904764	4.976148	4.980949	4.984090	4.986402	4.988094	4.989420	4.990479
x_8	5.021416	5.005326	5.004280	5.003570	5.003067	5.002675	5.002381	5.002142
x_9	5.036293	5.009045	5.007255	5.006049	5.005192	5.004540	5.004030	5.003630
x_{10}	5.021169	5.005191	5.004230	5.003529	5.003030	5.002647	5.002352	5.002116
x_{11}	5.187160	5.046762	5.037451	5.031171	5.026724	5.023401	5.020793	5.018716
x_{12}	5.145135	5.036899	5.029027	5.024236	5.020744	5.018150	5.016124	5.014512
x_{13}	5.135218	5.033776	5.027063	5.022516	5.019323	5.016901	5.015024	5.013522
x_{14}	4.998607	4.999623	4.999640	4.999750	4.999809	4.999829	4.999844	4.999860
x_{15}	5.150962	5.037712	5.030192	5.025025	5.021567	5.018865	5.016774	5.015096
x_{16}	4.949985	4.987475	4.990021	4.991690	4.992905	4.993748	4.994450	4.995000
x_{17}	4.899436	4.974830	4.979905	4.983267	4.985664	4.987426	4.988826	4.989943
x_{18}	5.180466	5.045088	5.036112	5.030152	5.025788	5.022565	5.020046	5.018046
x_{19}	4.866953	4.966714	4.973410	4.977847	4.981010	4.983369	4.985218	4.986696
x_{20}	4.912945	4.978208	4.982586	4.985492	4.987575	4.989117	4.990329	4.991294

TABLE 8. The optimal portfolio strategy of model (13) at $\alpha = 0.10$, $\eta = 0.5$ (%).

ϵ	0.01	0.02	0.03	0.04	0.05	0.06	0.07	0.08
x_1	4.819722	5.017196	4.823575	4.819690	4.819698	4.824805	4.830488	4.835194
x_2	4.874888	5.049394	4.906513	4.874885	4.874880	4.893740	4.914823	4.932462
x_3	5.059102	5.008832	5.068327	5.059101	5.059108	5.064343	5.070199	5.074701
x_4	4.938854	4.888525	4.947435	4.938853	4.938842	4.933902	4.928348	4.924245
x_5	4.947160	5.007457	5.020429	4.947190	4.947148	4.904790	4.857438	4.822615
x_6	4.949764	5.010525	4.904169	4.949795	4.949753	4.938490	4.925878	4.917168
x_7	4.904764	4.988027	5.033945	4.904762	4.904756	4.885079	4.863072	4.841455
x_8	5.021416	4.956831	5.023001	5.021415	5.021421	4.998966	4.973876	4.955611
x_9	5.036293	4.980831	5.040336	5.036291	5.036310	5.050703	5.066812	5.078063
x_{10}	5.021169	4.991940	4.871768	5.021167	5.021242	5.030572	5.041085	5.048017
x_{11}	5.187160	4.993328	5.214329	5.187169	5.187179	5.173677	5.158649	5.148352
x_{12}	5.145135	5.040305	5.149204	5.145134	5.145132	5.141004	5.136379	5.133094
x_{13}	5.135218	4.963133	5.144470	5.135217	5.135166	5.122253	5.107767	5.096481
x_{14}	4.998607	5.000210	5.003093	4.998605	4.998594	5.024806	5.054100	5.075939
x_{15}	5.150962	4.957459	5.003092	5.150960	5.150966	5.160804	5.171804	5.180210
x_{16}	4.949985	5.153689	5.007405	4.950010	4.950002	4.979582	5.012606	5.037460
x_{17}	4.899436	4.940945	4.895502	4.899431	4.899427	4.919690	4.942342	4.960230
x_{18}	5.180466	5.044726	5.224574	5.180433	5.180485	5.154459	5.125396	5.103489
x_{19}	4.866953	4.960684	4.846945	4.866952	4.866951	4.854918	4.841448	4.831741
x_{20}	4.912945	5.045963	4.871888	4.912943	4.912941	4.943417	4.977490	5.003472

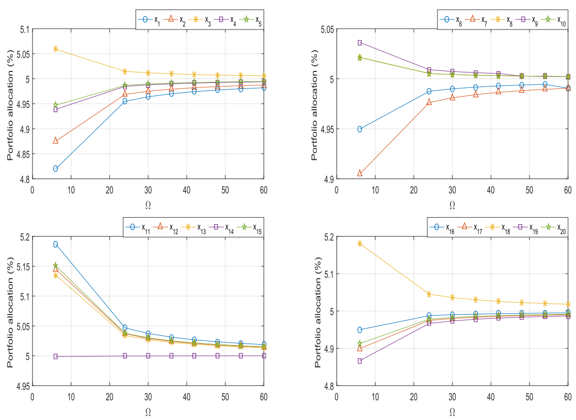


FIGURE 13. The variation of optimal portfolio strategy under $\epsilon = 0.01$, $\eta = 0.5$ and $\alpha = 0.10$.

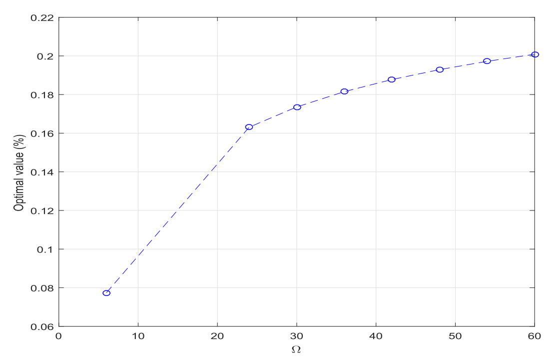


FIGURE 14. The optimal value of model (13) under $\epsilon = 0.01$, $\eta = 0.5$ and $\alpha = 0.10$.

parameter value. With ϵ approaching 0.10, the portfolio allocation ratio of each stock tends to be relatively stable.

Taking into account the difference in portfolio strategies under different ϵ, the relationship between optimal objective values and the values of Wasserstein's spherical radius ϵ is given in Fig. 16. It shows that the optimal objective value

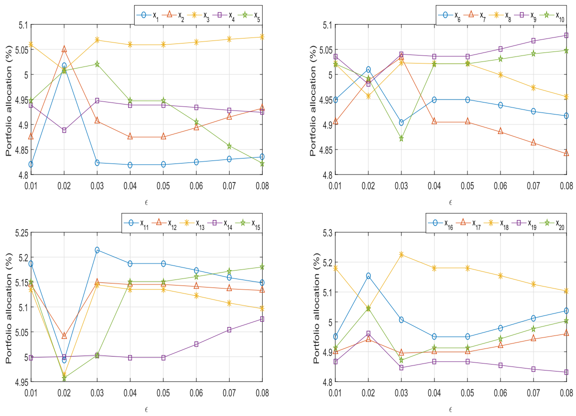


FIGURE 15. The variation of optimal portfolio strategy under different values of ϵ.

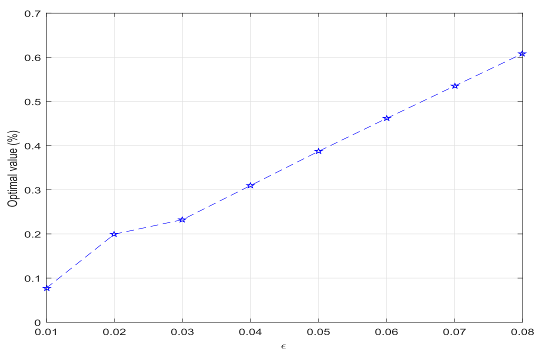


FIGURE 16. The optimal objective value of model (13) at different values ϵ.

enlarges with the increase in ϵ. Therefore, as the radius ϵ of the Wasserstein ambiguity set becomes increasingly larger, the solution of model (13) becomes more conservative.

3) THE INFLUENCE OF WEIGHT COEFFICIENT η

In order to test the impact of the weight coefficient η, Fig. 17 is plotted to show the optimal portfolio strategy of model (13). It is easy to determine that the portfolio allocation proportion fluctuates around 5%. Especially, the allocation proportions are closest to 5% at $\eta = 0$ or $\eta = 1.0$. This result is estimated from a large representation of historical data. Practice shows that investors often acquire a relatively uniform proportion of portfolio allocation when they try to estimate stock information based on historical data. From Fig. 17, we can find that when the weight parameter η is between 0.3 and 0.8, the portfolio allocation ratio of each stock is relatively different. In contrast, when $\eta < 0.3$ or $\eta > 0.8$, the portfolio allocation ratio of each stock is close to 5%. To better illustrate this point, Fig. 18 is plotted and displays the optimal value of the model under different values of weight coefficient η. From this figure we also can determine that when the value of weight coefficient η is small, the solution of model (13) is more conservative.

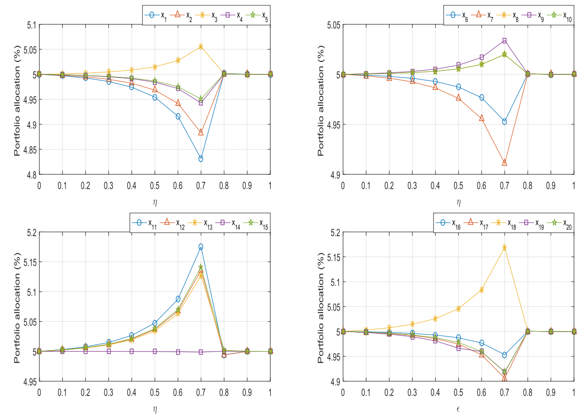


FIGURE 17. The variation of optimal portfolio strategy of model (13) under different weight coefficient η.

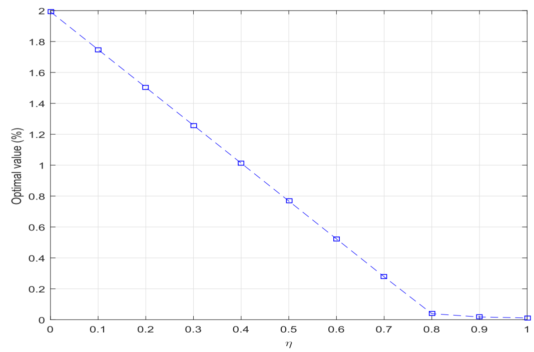


FIGURE 18. The optimal value of model (13) at different weight coefficient η under $\epsilon = 0.01$, $\alpha = 0.02$.

4) THE INFLUENCE OF RISK LEVEL α

In this subsection, the change in the optimal portfolio strategy with respect to different risk levels is shown in Fig. 19, which vividly displays the change of the portfolio allocation proportion of each stock with respect to risk level α. Specifically, with the increase in risk level α, the portfolio allocation proportions of stocks 000488 (x_3), 000966 (x_8), 002001 (x_9), 002040 (x_{10}), 600000 (x_{11}), 600016 (x_{12}), 600028 (x_{13}), 600276 (x_{15}) and 600519 (x_{18}) are gradually increasing, while those of the other stocks are decreasing. Particularly, the proportion of portfolio allocation is closer to 5% at $\alpha = 0.02$. Further, Fig. 20 depicts the optimal objective values of model (13) with respect to different risk levels. It is easy to see that the smaller the risk level is, the more conservative the solution of model (13).

C. COMPARISON WITH AMBIGUITY-FREE CASE

If the radius of the Wasserstein ambiguity set decreases to zero, i.e., $\epsilon = 0$, then the ambiguity set is reduced to a set containing only one distribution. In this case, the distributionally robust problem reduces to an ambiguity-free stochastic program [39]. Moreover, the optimal portfolio allocation decision and optimal value of the convex program shrink to the equivalent model under the discrete empirical probability distribution in the ambiguity-free limit.

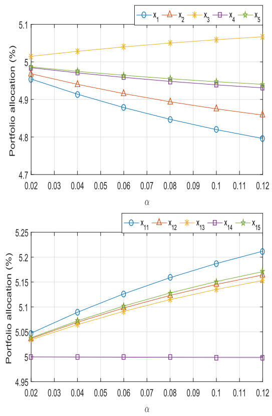


FIGURE 19. The optimal portfolio strategy of model (13) at risk level α under $\epsilon = 0.01$, $\eta = 0.5$.

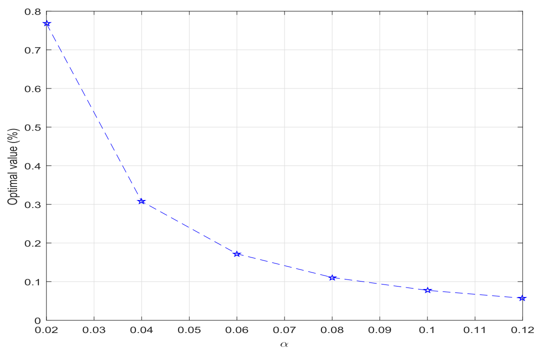


FIGURE 20. The optimal value of model (13) at risk level α under $\epsilon = 0.01$, $\eta = 0.5$.

For better comparison, some computational results under the ambiguity-free condition are represented. When the radius ϵ of the Wasserstein ambiguity set decreases to zero, Figs. 21- 24 are plotted. Among them, Figs. 21 and 22 present the optimal portfolio allocation strategy under different weight coefficients η under the ambiguity-free condition. Fig. 21 shows that when $\eta = 0, 0.1, 0.2, \ldots, 0.9$, the portfolio allocation proportion of each stock is relatively concentrated, and that they are primarity distributed between 3% and 6%. However, in the case of risk aversion, i.e., $\eta = 1$, the computational results show that investors should maximize investment in the first stock 000039. Further, Fig. 22 shows that weight coefficient has no significant effect on the results of distributionally robust portfolio optimization models with respect to the ellipsoid uncertainty set under the ambiguity-free condition, except in the case of risk aversion. Moreover, Figs. 23 and 24 are depicted, which present the impact of risk level α on optimal portfolio allocation strategy under the ambiguity-free condition. It is easy to determine that the change in risk level α exerts little effect on the optimal portfolio allocation proportion under the ambiguity-free condition and that the portfolio allocation ratio of each stock is concentrated between 4.5% and 5.3%. This also proves that the inaccuracy in the distribution of random returns cannot be ignored.

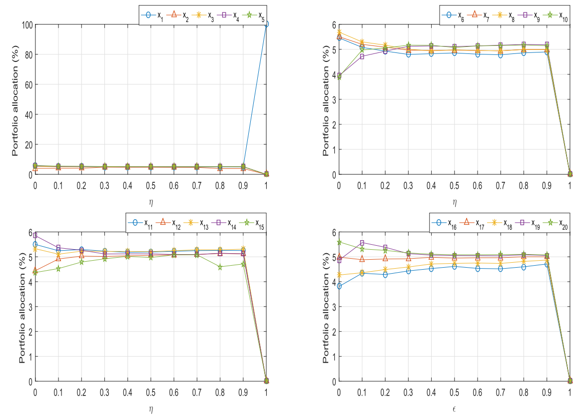


FIGURE 21. The variation of optimal portfolio strategy with different η under $\alpha = 0.01$, $\epsilon = 0$.

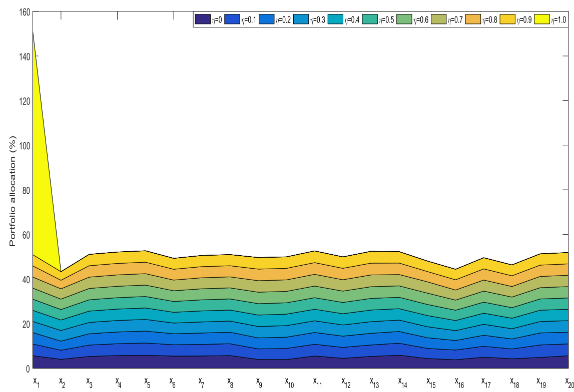


FIGURE 22. The variation of optimal portfolio strategy under $\alpha = 0.01$, $\epsilon = 0$.

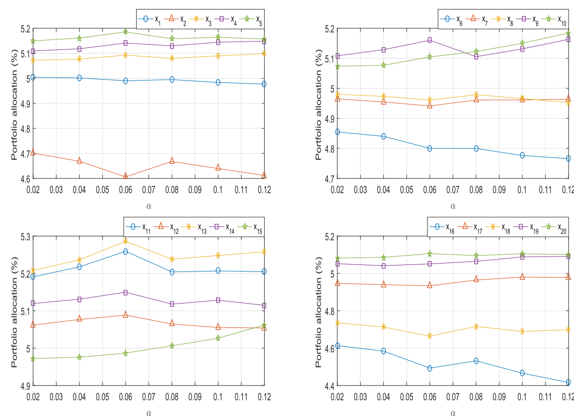


FIGURE 23. The variation of optimal portfolio strategy with different α under $\eta = 0.5$, $\epsilon = 0$.

Note that the objective value of all experiments conducted under the ambiguity-free condition is zero. It seems that the computational results of the portfolio optimization model under the ambiguity-free condition are more considerable. However, there is no doubt that the optimal portfolio strategy of the proposed distributionally robust portfolio optimization model has the ability to immunize against distribution uncertainty. Even if the radius of the Wasserstein ambiguity set has

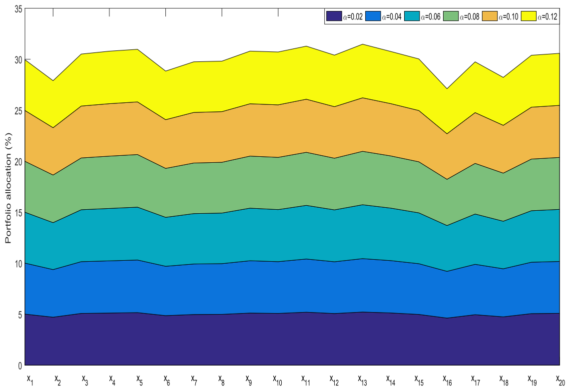


FIGURE 24. The variation of optimal portfolio strategy under $\eta = 0.5$, $\epsilon = 0$.

a small perturbation, its influence on the optimal portfolio strategy and expected loss cannot be neglected.

V. CONCLUSION

In this article, we study the mean-CVaR portfolio optimization problem under distribution uncertainty from a new perspective, where the uncertainty set is described by a cone representation. A distributionally robust linear optimization framework for the portfolio optimization problem is developed first. Moreover, we use the Wasserstein metric to construct a ball in the space of probability distributions centered at the empirical distribution of the sample dataset. Then, a linear convex program under a cone representation that performs best in view of the worst-case expectation within the Wasserstein ambiguous set is presented. To calibrate the levels of the Wasserstein ambiguous set, we introduce a data-driven standard to provide an essential modeling guide. Specifically, we develop a new method to describe the distribution information of stock returns by constructing Wasserstein ambiguity set, which is generated based on the historical data of stock. A new data-driven mean-CVaR portfolio optimization framework for the distributionally robust linear optimization model is proposed in this article. Moreover, the proposed portfolio model is a general model. It can be employed in different stock markets.

To demonstrate the effectiveness of our proposed method, we provide some practical cases pertaining to the Chinese stock market. Numerical experiments are carried out based on the respective daily and monthly returns of twenty stocks, and the influence of different model parameters on the computational results is analyzed. Further, we discuss three new equivalent models with respect to box, budget, and ellipsoid uncertainty sets. A small-scale experiment based on the daily returns of stocks is carried out with respect to the box uncertainty set. To show the advantage of our method, we compare the computational results of the portfolio optimization model under different uncertainty sets, i.e., box and budget. Under the ellipsoid uncertainty set, the large-scale experiment based on the monthly returns of twenty stocks is conducted. Moreover, the comparison between the proposed data-driven distributionally robust portfolio optimization method and the traditional ambiguity-free stochastic optimization method is presented via numerical experiments. In the appendix, we perform some numerical experiments based on the United States stock market to demonstrate the generality of the model.

APPENDIX
NUMERICAL EXPERIMENT BASED ON THE UNITED STATES STOCK MARKET

In order to further demonstrate the generality of the data-driven mean-CVaR portfolio optimization model proposed in this article, we conduct some new numerical experiments in this section based on the United States stock market.

A. DATA AND METHODOLOGY

The daily returns of twenty stocks from the United States stock market are recorded as a data set to perform the experiments. Similarly, the daily returns of twenty stocks from the United States stock market are set according to the closing quotation. The sample period of historical return values of twenty stocks is August 3 to September 30, 2020, for 42 observations in total, that is, $N = 42$. Table 9 summarizes the daily returns of these twenty stocks from the United States stock market on September 14. We would like to estimate the distribution of twenty stocks via the 42×20 observations in dataset. Owing to the limited historical data, it is impossible to determine the true distribution of stock return without ambiguity. Wasserstein ambiguity set with a discrete empirical distribution $\hat{\mathbb{P}}_N$ is constructed in this article, where $\hat{\mathbb{P}}_N$ is supported by 42×20 observations values of stock returns.

To find the optimal portfolio strategy, the numerical experiments in this subsection are devoted to solving model (11), which is the equivalent representation of data-driven mean-CVaR portfolio optimization model under box uncertainty set.

B. COMPUTATIONAL RESULTS

Similarity, some sensitivity analyses on different parameters in model (11) are presented in the following subsections to test the accuracy of the model.

1) THE INFLUENCE OF UNCERTAINTY SET PARAMETER Λ

Consistent with the Section III-B1, here we take the parameter Λ of box uncertainty set equals to 1.2, 4.8, 6.0, 7.2, 8.4, 9.6, 10.8 and 12.0. In order to illustrate the change of the optimal portfolio allocation proportion under different Λ, Fig. 25 is plotted, which shows the variation of portfolio allocation proportion of twenty stocks with respect to the parameter of box uncertainty set. From Fig. 25, we can identify that the stock AAPL (x_1) has the largest proportion of portfolio allocation at $\Lambda = 7.2$. When the parameter Λ of box uncertainty set equals to 1.2, 12, stock LBTYA (x_7) has the largest proportion of portfolio allocation, while stock GOOG (x_9) has the largest proportion of portfolio allocation

TABLE 9. **The assets from United States stock market.**

Stock code		Return	Stock code		Return	Stock code		Return	Stock code		Return
AAPL	ξ_1	1.0300	MSFT	ξ_6	1.0068	ADBE	ξ_{11}	1.0309	BIIB	ξ_{16}	1.0310
TSLA	ξ_2	1.1258	LBTYA	ξ_7	1.0051	ADI	ξ_{12}	1.0072	BKNG	ξ_{17}	1.0005
SBUX	ξ_3	1.0159	INTC	ξ_8	1.0026	AMD	ξ_{13}	1.0204	CDW	ξ_{18}	1.0164
PEP	ξ_4	1.0063	GOOG	ξ_9	0.9991	AMZN	ξ_{14}	0.9957	CHKP	ξ_{19}	1.0231
NVDA	ξ_5	1.0582	FB	ξ_{10}	0.9983	AVGO	ξ_{15}	1.0070	ZM	ξ_{20}	1.0536

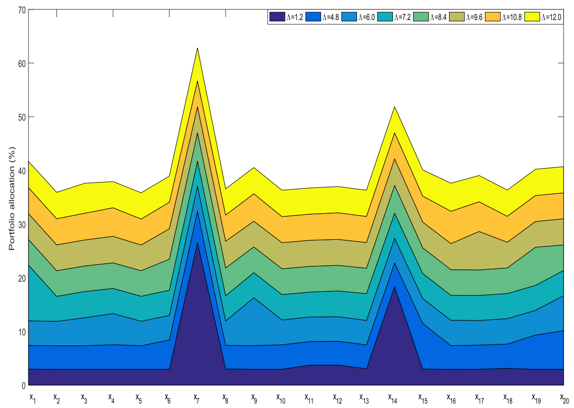


FIGURE 25. **The optimal portfolio strategy based on United States stocks under different Λ.**

at $\Lambda = 6.0$. If the parameter Λ equals to 4.8, 8.4, 9.6, 10.8, the largest proportion of portfolio allocation are stocks AVGO (x_{15}), CHKP (x_{19}), BKNG (x_{17}), and BIIB (x_{16}) respectively. It is not difficult to find that the influence of uncertainty set parameter on the optimal portfolio strategy is visible. Further, Fig. 26 is depicted to reveal the effect of uncertainty set parameter Λ on the optimal value. Fig. 26 illustrates that the optimal value of model (11) increases with the growth of the value of parameter Λ. Put it differently, as the value of Λ becomes increasingly larger, the solution results of the model (11) becomes more conservative with respect to the box uncertainty set.

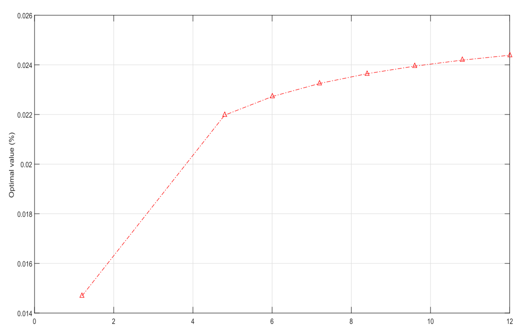


FIGURE 26. **The optimal value of model (11) based on United States stocks under different uncertainty set parameter Λ.**

2) THE INFLUENCE OF WASSERSTEIN RADIUS ϵ

In this subsection, the value of Wasserstein radius ϵ is primarily based on the result of literature [39], we make ϵ equal to 0.01, 0.02, 0.03, 0.04, 0.05, 0.06, 0.07, 0.08 in the numerical experiment. To better present the optimal portfolio strategy under different values of Wasserstein radius ϵ, Figs. 27 and 28 are depicted. From the two figures, it is easy to catch sight of the sensitivity of the optimal portfolio strategy to Wasserstein radius ϵ. In Fig. 27, the portfolio allocation proportions of stocks TSLA (x_2), PEP (x_4), NVDA (x_5), MSFT (x_6), LBTYA (x_7), GOOG (x_9), ADI (x_{12}), AMD (x_{13}), AMZN (x_{14}), BIIB (x_{16}), CDW (x_{18}), and CHKP (x_{19}) change obviously, the variations in allocation proportions of the other stocks are not as obvious. In addition, Fig. 28 illustrates that the stock LBTYA (x_7) accounts for the largest allocation proportion at $\epsilon = 0.01$; ADI (x_{12}) accounts for the largest allocation proportion at $\epsilon = 0.02$; BIIB (x_{16}) accounts for the largest allocation proportion at $\epsilon = 0.03, 0.06$; CDW (x_{18}) accounts for the largest allocation proportion at $\epsilon = 0.04$; AMD (x_{13}) has a largest allocation proportion at $\epsilon = 0.05, 0.07$; CHKP (x_{19}) has a largest allocation proportion at $\epsilon = 0.08$. Further, we plot Fig. 29, which illustrates the variation of optimal value of model (11) under different Wasserstein radius. The optimal objective value is enlarged according to the increase of Wasserstein radius. Put it differently, the computational results of model (11) become more conservative as the Wasserstein radius becomes larger.

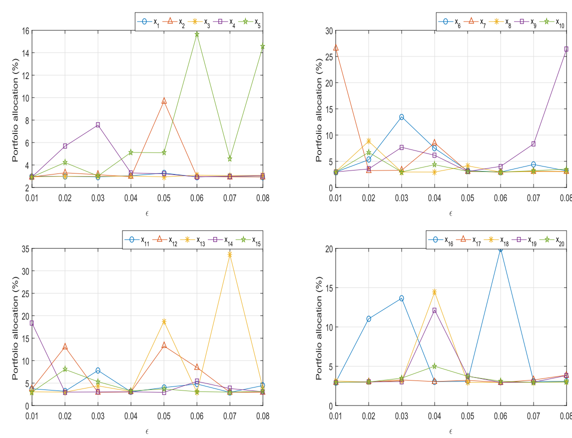


FIGURE 27. **The variation of optimal portfolio strategy based on United States stocks under different ϵ.**

3) THE INFLUENCE OF WEIGHT COEFFICIENT η

It is well known that weight coefficient η indicates whether investors pay more attention to loss or risk. Investors would like to reduce the portfolio losses as $\eta > 0.5$, and investors would like to reduce the risk of investment as $\eta < 0.5$.

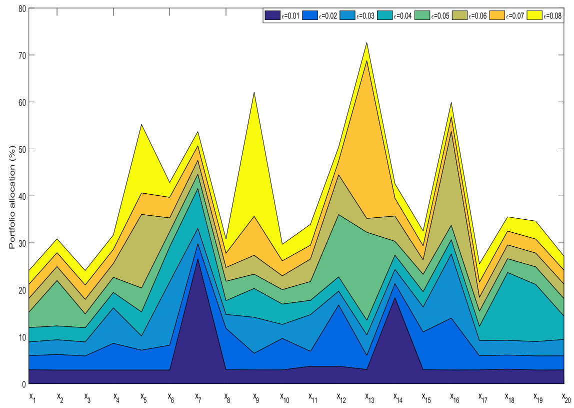


FIGURE 28. The optimal portfolio strategy based on United States stocks under different ϵ.

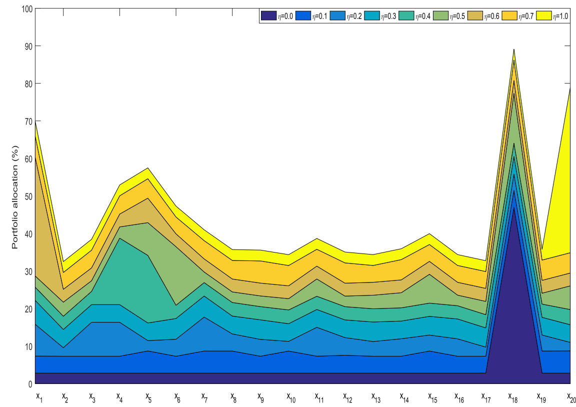


FIGURE 30. The optimal portfolio strategy based on United States stocks under different η.

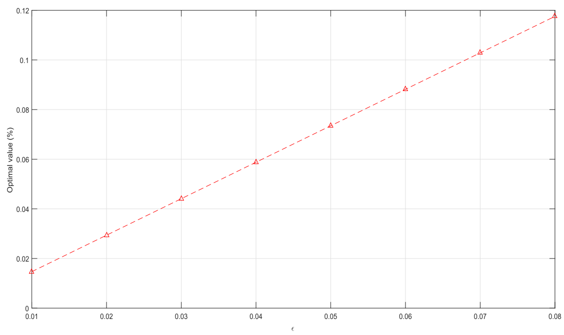


FIGURE 29. The optimal value of model (11) based on United States stocks under different Wasserstein radius ϵ.

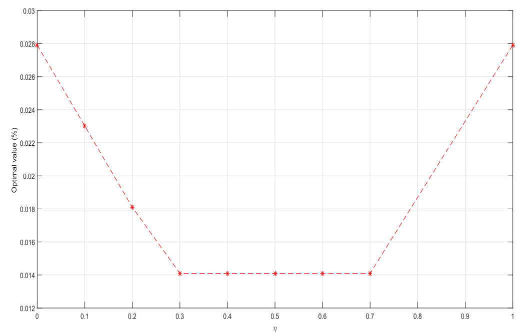


FIGURE 31. The optimal value of model (11) based on United States stocks under different weight coefficient η.

In this subsection, Fig. 30 is depicted, which illustrates the optimal portfolio allocation proportion under different weight coefficients. When the weight coefficient equals to 0, i.e., $\eta = 0$, the investor is risk neutrality. At $\eta = 0$, the investor invests 47% of his or her assets in stock CDW (x_{18}), which accounts for a largest portfolio allocation proportion. If the weight coefficient equals to 1, i.e., $\eta = 1$, the investor is risk aversion and invests 44% of his or her assets in stock ZW (x_{20}). ZW (x_{20}) has a largest portfolio allocation proportion at $\eta = 1$. In addition, stock NVDA (x_5) accounts for a largest portfolio allocation proportion at $\eta = 0.4$; stock MSFT (x_6) accounts for a largest portfolio allocation proportion at $\eta = 0.5$; stock AAPL (x_1) accounts for a largest portfolio allocation proportion at $\eta = 0.6$; and the portfolio allocation proportions are relatively decentralized at $\eta = 0.1, 0.2, 0.3$, and $\eta = 0.7$. Furthermore, the relationship between the optimal value of model (11) and weight coefficient is present in Fig. 31.

4) THE INFLUENCE OF RISK LEVEL α

Some numerical experiments with respect to risk level are conducted in this subsection to show the effect of risk level on optimal portfolio allocation and optimal value of model (11). In order to better illustrate the variation of optimal portfolio strategy at different risk level, Fig. 32 is depicted. In Fig. 32, the change amplitudes of portfolio allocation proportions of stocks SBUX (x_3), LBTYA (x_7), INTC (x_8), ADBE (x_{11}), AMZN (x_{14}), BIIB (x_{16}), and BKNG (x_{17}) is relatively large under different risk levels, while the change range of portfolio allocation proportions of the other stocks are relatively small. More specifically, stock MSFT (x_6) has a largest allocation proportion at $\alpha = 0.02$; stock BIIB (x_{16}) has a largest allocation proportion at $\alpha = 0.04$; stock SBUX (x_3) accounts for a largest allocation proportion at $\alpha = 0.06, 0.16$; stock INTC (x_8) has a largest allocation proportion at $\alpha = 0.08, 0.14$; stock LBTYA (x_7) accounts for a largest allocation proportion at $\alpha = 0.10$; stock ADBE (x_{11}) accounts for a largest allocation proportion at $\alpha = 0.12$. Further, we plot Fig 33, which shows the optimal value of model (11) under different parameters of risk level. It illustrates that the optimal value increases as the value of risk level increases. Put it differently, the higher the risk level α is, the more conservative the solution of model (11).

C. DISCUSSION

A new modelling method is developed to describe the distribution information of stock returns by constructing Wasserstein ambiguity set, which is generated based on the historical data of stock. The methodology developed by us is general and can naturally be used on the stock markets of other countries, except for the Chinese stock market. Therefore, we implement some numerical experiments based on United

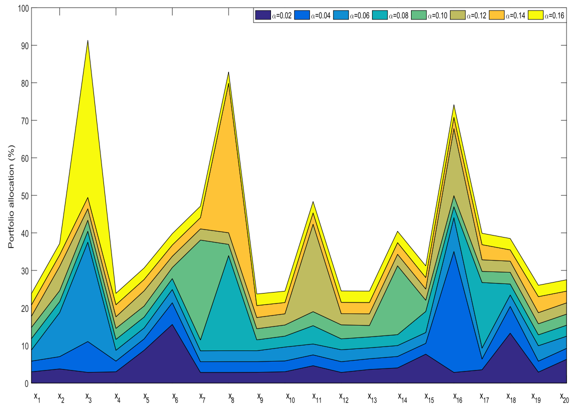


FIGURE 32. The optimal portfolio strategy of model (13) at risk level α under $\epsilon = 0.01$, $\eta = 0.5$.

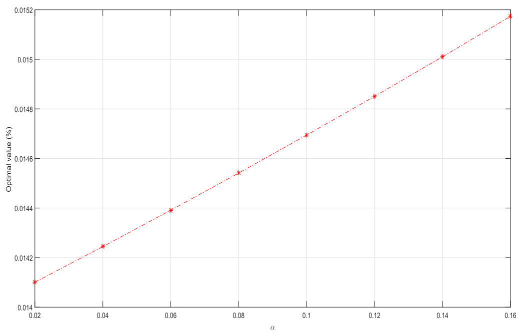


FIGURE 33. The optimal value of model (13) at risk level α under $\epsilon = 0.01$, $\eta = 0.5$.

States stock market in this section. The modification is to reconstruct Wasserstein ambiguity set based on the historical data of stock return from United States stock market.

When the model is used in a specific situation, historical data is needed to construct Wasserstein ambiguity set. Thus, Wasserstein ambiguity set contains different distribution information in comparison with the numerical experiments based on Chinese stock market. However, whether it is a numerical experiment based on the Chinese stock market or a numerical experiment based on United States stock market, the trend of the optimal objective value affected by the key parameters of box uncertainty parameter Λ, Wasserstein radius ϵ, weight coefficient η and risk level α is consistent. Specifically, the optimal value increases with the growth of the value of parameters Λ, ϵ and α respectively. To put it differently, as the values of Λ, ϵ and α become increasingly larger, the solution of model becomes more conservative with respect to the three parameters.

CONFLICTS OF INTEREST

No potential conflict of interest was reported by the authors.

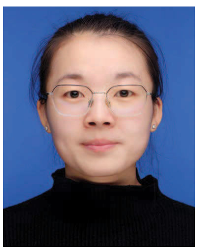
ACKNOWLEDGMENT

This authors would like to thank the editor and the anonymous referees for their useful comments and suggestions, which helped to improve the presentation of this article.

REFERENCES

[1] H. Markowitz, "Portfolio selection," *J. Finance*, vol. 7, no. 1, pp. 77–91, 1952.
[2] H. Markowitz, *Portfolio Selection: Efficient Diversfication of Investments*. New York, NY, USA: Wiley, 1959.
[3] Y. Y. Zhang, X. Li, and S. N. Guo, "Portfolio selection problems with Markowitz's mean–variance framework: A review of literature," *Fuzzy Optim. Decis. Making*, vol. 17, no. 2, pp. 125–158, 2018.
[4] M. Altinoz and O. T. Altinoz, "Systematic initialization approaches for portfolio optimization problems," *IEEE Access*, vol. 7, pp. 57779–57794, 2019.
[5] R. Gao, Y. Li, Y. Bai, and S. Hong, "Bayesian inference for optimal risk hedging strategy using put options with stock liquidity," *IEEE Access*, vol. 7, pp. 146046–146056, 2019.
[6] N. Du, Y.-K. Liu, and Y. Liu, "New safe approximation of ambiguous probabilistic constraints for financial optimization problem," *Discrete Dyn. Nature Soc.*, vol. 2019, pp. 1–19, May 2019.
[7] C. Francq and J.-M. Zakoïan, "Estimation risk for the VaR of portfolios driven by semi-parametric multivariate models," *J. Econometrics*, vol. 205, no. 2, pp. 381–401, Aug. 2018.
[8] K. T. Lwin, R. Qu, and B. L. Maccarthy, "Mean-VaR portfolio optimization: A nonparametric approach," *Eur. J. Oper. Res.*, vol. 260, no. 2, pp. 751–766, 2017.
[9] R. T. Rockafellar and S. Uryasev, "Optimization of conditional value-at-risk," *J. Risk*, vol. 2, no. 3, pp. 21–41, 2000.
[10] G. J. Alexander and A. M. Baptista, "A comparison of VaR and CVaR constraints on portfolio selection with the mean-variance model," *Manage. Sci.*, vol. 50, no. 9, pp. 1261–1273, Sep. 2004.
[11] A. Ben-Tal, L. El Ghaoui, and A. Nemirovski, *Robust Optimization*. Princeton, NJ, USA: Princeton Univ. Press, 2009.
[12] Y. Takano, K. Nanjo, N. Sukegawa, and S. Mizuno, "Cutting plane algorithms for mean-CVaR portfolio optimization with nonconvex transaction costs," *Comput. Manage. Sci.*, vol. 12, no. 2, pp. 319–340, Apr. 2015.
[13] Q. Xu, Y. Zhou, C. Jiang, K. Yu, and X. Niu, "A large CVaR-based portfolio selection model with weight constraints," *Econ. Model.*, vol. 59, pp. 436–447, Dec. 2016.
[14] Z. Kang, X. Li, Z. Li, and S. Zhu, "Data-driven robust mean-CVaR portfolio selection under distribution ambiguity," *Quant. Finance*, vol. 19, no. 1, pp. 105–121, Jan. 2019.
[15] H. E. Scarf, "A min-max solution of an inventory problem," *Stud. Math. Inventory Prod.*, vol. 25, no. 2, pp. 352–367, 1957.
[16] Z. Chang, J.-Y. Ding, and S. Song, "Distributionally robust scheduling on parallel machines under moment uncertainty," *Eur. J. Oper. Res.*, vol. 272, no. 3, pp. 832–846, Feb. 2019.
[17] E. Delage and Y. Ye, "Distributionally robust optimization under moment uncertainty with application to data-driven problems," *Oper. Res.*, vol. 58, no. 3, pp. 595–612, Jun. 2010.
[18] W. Wiesemann, D. Kuhn, and M. Sim, "Distributionally robust convex optimization," *Oper. Res.*, vol. 62, no. 6, pp. 1358–1376, Dec. 2014.
[19] H. Xu, Y. Liu, and H. Sun, "Distributionally robust optimization with matrix moment constraints: Lagrange duality and cutting plane methods," *Math. Program.*, vol. 169, no. 2, pp. 489–529, Jun. 2018.
[20] Y. Liu and Y.-K. Liu, "Distributionally robust fuzzy project portfolio optimization problem with interactive returns," *Appl. Soft Comput.*, vol. 56, pp. 655–668, Jul. 2017.
[21] N. Rujeerapaiboon, D. Kuhn, and W. Wiesemann, "Robust growth-optimal portfolios," *Manage. Sci.*, vol. 62, no. 7, pp. 2090–2109, Jul. 2016.
[22] B. P. G. Van Parys, P. J. Goulart, and M. Morari, "Distributionally robust expectation inequalities for structured distributions," *Math. Program.*, vol. 173, nos. 1–2, pp. 251–280, Jan. 2019.
[23] A. Ben-Tal, D. den Hertog, A. De Waegenaere, B. Melenberg, and G. Rennen, "Robust solutions of optimization problems affected by uncertain probabilities," *Manage. Sci.*, vol. 59, no. 2, pp. 341–357, Feb. 2013.
[24] D. Klabjan, D. Simchi-Levi, and M. Song, "Robust stochastic lot-sizing by means of histograms," *Prod. Oper. Manage.*, vol. 22, no. 3, pp. 691–710, May 2013.
[25] D. Bertsimas, V. Gupta, and N. Kallus, "Robust sample average approximation," *Math. Program.*, vol. 171, nos. 1–2, pp. 217–282, Sep. 2018.
[26] E. Çetinkaya and A. Thiele, "Data-driven portfolio management with quantile constraints," *OR Spectr.*, vol. 37, no. 3, pp. 761–786, Jul. 2015.
[27] B. Fernandes, A. Street, D. Valladão, and C. Fernandes, "An adaptive robust portfolio optimization model with loss constraints based on data-driven polyhedral uncertainty sets," *Eur. J. Oper. Res.*, vol. 255, no. 3, pp. 961–970, Dec. 2016.

[28] G. Chi, S. Ding, and X. Peng, "Data-driven robust credit portfolio optimization for investment decisions in P2P lending," *Math. Problems Eng.*, vol. 2019, pp. 1–10, Jan. 2019.
[29] D. Bertsimas, V. Gupta, and N. Kallus, "Data-driven robust optimization," *Math. Program.*, vol. 167, no. 2, pp. 235–292, Feb. 2018.
[30] Z. Wang, P. W. Glynn, and Y. Ye, "Likelihood robust optimization for data-driven problems," *Comput. Manage. Sci.*, vol. 13, no. 2, pp. 241–261, Apr. 2016.
[31] V. Gupta, "Near-optimal Bayesian ambiguity sets for distributionally robust optimization," *Manage. Sci.*, vol. 65, no. 9, pp. 4242–4260, Sep. 2019.
[32] R. Jiang and Y. Guan, "Data-driven chance constrained stochastic program," *Math. Program.*, vol. 158, nos. 1–2, pp. 291–327, Jul. 2016.
[33] S. Zhao and K. Zhang, "A distributionally robust optimization approach to reconstructing missing locations and paths using high-frequency trajectory data," *Transp. Res. C, Emerg. Technol.*, vol. 102, pp. 316–335, May 2019.
[34] G. Pflug and D. Wozabal, "Ambiguous in portfolio selection," *Quant. Finance*, vol. 7, no. 4, pp. 435–442, 2007.
[35] G. A. Hanasusanto and D. Kuhn, "Conic programming reformulations of two-stage distributionally robust linear programs over Wasserstein balls," *Oper. Res.*, vol. 66, no. 3, pp. 849–869, Jun. 2018.
[36] J. G. Carlsson, M. Behroozi, and K. Mihic, "Wasserstein distance and the distributionally robust TSP," *Oper. Res.*, vol. 66, no. 6, pp. 1603–1624, Nov. 2018.
[37] C. Duan, W. Fang, L. Jiang, L. Yao, and J. Liu, "Distributionally robust chance-constrained approximate AC-OPF with Wasserstein metric," *IEEE Trans. Power Syst.*, vol. 33, no. 5, pp. 4924–4936, Sep. 2018.
[38] R. Hettich and K. O. Kortanek, "Semi-infinite programming: Theory, methods, and applications," *SIAM Rev.*, vol. 35, no. 3, pp. 380–429, Sep. 1993.
[39] P. M. Esfahani and D. Kuhn, "Data-driven distributionally robust optimization using the Wasserstein metric: Performance guarantees and tractable reformulations," *Math. Program.*, vol. 171, nos. 1–2, pp. 115–166, Sep. 2018.
[40] L. V. Kantorovich and G. S. Rubinshtein, "On a space of completely additive functions," *Vestnik Leningrad Univ*, vol. 13, no. 7, pp. 52–59, 1958.



NINGNING DU received the B.S. and M.S. degrees in mathematics from the College of Mathematics and Information Science, Hebei University, Baoding, China, in 2017 and 2020, respectively. She is currently pursuing the Ph.D. degree with the School of Economics and Management, Beihang University. Her research interests include robust optimization method and its application in financial optimization problem, and supply chain financing.

YANKUI LIU received the B.S. and M.S. degrees from the Department of Mathematics, Hebei University, Baoding, China, in 1989 and 1992, respectively, and the Ph.D. degree in computational mathematics from the Department of Mathematical Science, Tsinghua University, Beijing, China, in 2003. He is currently a Professor with the College of Mathematics and Information Science, Hebei University. He is the author of 70 research articles and five monographs. His work ranges from theoretical/foundational work, including credibility measure theory and robust credibilistic optimization methods, to algorithmic analysis and design for optimization problems, such as credibilistic approximation approaches and their convergence, and to applications in various engineering and management problems. Based on the citations in Scopus database, he is featured among the Most Cited Chinese Researchers in the field of computer science since 2014.

YING LIU received the B.S. and M.S. degrees from the Department of Mathematics and Computer, Hebei University, Baoding, China, in 2003 and 2008, respectively, and the Ph.D. degree in management science and engineering from the Department of Management Science, Hebei University, in 2017. She is currently an Associate Professor with the College of Mathematics and Information Science, Hebei University. She is the author of 30 research articles. Her research interest includes fuzzy optimization methods and its applications in various engineering and management problems.