Yokohama Publishers
ISSN 1349-8169 ONLINE JOURNAL

Pacific Journal of Optimization
Vol. 18, No. 4, pp. 825–839, 2022

DISTRIBUTIONALLY ROBUST BILEVEL PROGRAMMING BASED ON WORST CONDITIONAL VALUE-AT-RISK*

Leiyan Xu, Rui Shen and Zhiqing Meng

Abstract: Distributionally robust bilevel programming based on worst conditional value at risk(WCVaR) is a decision model composed of upper and lower objective functions under WCVaR. First, a concave probability density distribution cluster describing random fluctuation is defined. Then the structure of distributionally robust bilevel programming based on WCVaR with the concave probability density distribution cluster is defined, where both the upper and the lower objective function of the model include WCVaR measures. It is proved that this distributionally robust bilevel programming can be approximately equivalently expressed as robust bilevel programming with parameters, which provides us with an approximate method to solve distributionally robust optimization based on WCVaR with the concave probability density distribution cluster.

Key words: *distributionally robust bilevel programming, conditional value at risk, concave probability density distribution function cluster*

Mathematics Subject Classification: *90C29, 91B30*

1 Introduction

The classical robust optimization method (also known as minimax robust optimization) solves optimization problems with a uncertain parameter set and finds out the equilibrium solution under worst case of the parameter set. The early robust optimization studied the robust problem where the uncertain set consists of some real parameter sets. Later, many people studied the stochastic robust optimization where uncertain parameters are composed of probability (density) distribution function with random variables in risk decision-making [18]. And this kind of stochastic robust optimization is also called distributionally robust optimization.

In many practical problems, uncertain factors (i.e. random variables) directly affect the accuracy of decision-making. As early as 1958, Scarf first used the single-layer distributionally robust optimization model to solve the inventory management problem where random variables affect the uncertainty of decision-making [5]. Most of the distributionally robust optimization models take the expected mean risk loss or return as the objective function. The expected risk mean measures the average risk fluctuation by the probability distribution function describing the risk factors. When risk fluctuation is small, the expected mean measures the risk, and the strategy obtained by robust optimization is relatively robust

*This work is supported by the National Natural Science Foundation of China under Grant No.12101319.

in application. However, when demand fluctuation is large, the obtained robust strategy becomes less robust if the same expected mean loss function is also used to describe the robust loss. Therefore, the distributionally robust optimization with the objective function composing of mean is not suitable for the robust decision when risk fluctuation is large and random. When the risk fluctuation is very large, especially when the tail risk of the probability distribution function is large, the expected mean risk loss decision model will fail [1]. In order to overcome the problem caused by large risk fluctuation, value at risk (VaR) measure is put forward, which depicts the maximum loss value at a given risk loss level. In 1999, Rockfeller and Uryasev revised VaR [16] and proposed a conditional value at risk (CVaR) to describe the risk loss with risk aversion. The research shows that CVaR is more suitable for solving the decision-making problem with large risk fluctuation, e.g. risk averse product ordering and inventory problem with large demand fluctuation. Therefore, using CVaR model to solve two-level robust decision-making problems with large risk fluctuation is more in line with the actual.

Early research on CVaR robustness was mainly on securities portfolio [23, 24, 10, 11]. Later, CVaR robust optimization model was applied to inventory management. For example, in 2014, Qiu established a CVaR robust optimization inventory model under the convex probability distribution cluster of disturbance parameters [13]. In 2016, a CVaR robust optimization model based on the convex probability distribution cluster is studied to solve the inventory management for direct chain enterprises [21]. Since 2017, the objective loss function has been applied in the bilevel risk decision-making model which takes CVaR risk into consideration. The main research results cover bilevel robust optimization problems in some practical fields, such as the two-level robust optimization decision-making problems of electric vehicle integrator participation in competition decision [9], green supply chain network design problem [8], intelligent distribution problem [2, 3], wind power generator supply problem [15], microgrid operation scheduling [6], hazardous waste management problem [17] and Dr aggregator scheduling [14] in the power market.

In the above studies, CVaR is included in the upper decision function of the distributionally robust bilevel optimization, but not in the lower level decision function. From the above literature it is understood that a special algorithm is needed for solving each distributionally robust bilevel optimization, i.e. there is no general theory on robust bilevel optimization under conditional risk value. Many studies show that CVaR is suitable for decision-making problems with large fluctuation [23, 24, 10, 11, 13, 21]. Therefore,it is necessary to make an in-depth theoretical research on robust bilevel optimization based on CVaR.The upper level and the lower level mean the decision order in a bilevel decision-making problem, which is different from the one in perfect competition. In the game of perfect competition, for decision-makers there is no order in decision-making. But in a bilevel decision-making problem, suppliers are the upper decision-makers who is the first in the supply chain to determine wholesale prices, and retailers are the lower level decision-makers who then determines the order quantity. Furthermore, the lower level demand risk often directly affects the upper level demand decision in supply chain, especially when demand shows "bullwhip effect" where demand fluctuates greatly, which makes the mean risk measure obviously not suitable for the actual decision. So, the upper and lower levels in supply chain management should both be taken to consider the influence of demand probability distribution. Thus, this paper studies some equivalent problems to solve distributionally robust bilevel programming based on WCVaR.

The remainder of this paper is organized as follows. In Section 2, a concave probability density distribution cluster describing random fluctuation is defined; the structure of distributionally robust bilevel programming based on WCVaR with the concave probability density distribution cluster is defined; the approximate equivalence of distributionally robust bilevel programming to robust bilevel probramming is proved; and a distributionally robust bilevel programming is formulated by an approximate linear bilevel programming. In Section 3, a conclusion is given.

2 Distributionally Robust Bilevel Programming based on WCVaR

In this section, distributionally robust bilevel programming based on WCVaR (or called bilevel robust CVaR model) is studied. There are an upper level decision maker and a lower level decision maker in this bilevel robust CVaR model. The upper level decision maker has his own decision variables, random perturbation factors, a constraint set and an objective function, while the lower level decision maker has all of his own. The bilevel robust decision model consists of the upper robust optimization problem and the lower one. The structure of the model is defined as follows.

For the upper robust optimization problem, let the upper loss function $F_{\boldsymbol{\tau}}(\boldsymbol{x},\boldsymbol{y},\boldsymbol{\xi}): R^n \times R^m \times R^r \to R^1$ be continuous on variables $(\boldsymbol{x},\boldsymbol{y}) \in R^n \times R^m$, where $\boldsymbol{x} \in R^n$ is a variable for the upper decision maker, $\boldsymbol{y} \in R^m$ is a variable for the lower decision maker, $\boldsymbol{\xi}$ is a continuous random variable, $\boldsymbol{\tau} \in T \subset R^{r_1}$ is a perturbation parameter and T is a convex perturbation set. Let the constraint set of all variables $(\boldsymbol{x},\boldsymbol{y})$ of the upper level problem be defined as

$$X = \{(\boldsymbol{x},\boldsymbol{y}) \mid f_i(\boldsymbol{x},\boldsymbol{y}) \le 0, i = 1,2,\ldots,I\},$$

where $f_i : R^n \times R^m \to R^1$ are continuous. Let $\Xi_u = \{p(\boldsymbol{\xi},\boldsymbol{\tau}) | \boldsymbol{\tau} \in T\}$ be the probability density distribution function cluster of the upper level problem, where $p(\boldsymbol{\xi},\boldsymbol{\tau})$ is a probability density distribution function for $\boldsymbol{\tau} \in T$.

For the lower robust optimization problem, let the lower loss function $G_{\boldsymbol{\kappa}}(\boldsymbol{x},\boldsymbol{y},\boldsymbol{\zeta}): R^n \times R^m \times R^s \to R^1$ be continuous on variables $(\boldsymbol{x},\boldsymbol{y}) \in R^n \times R^m$, where $\boldsymbol{x} \in R^n$ is a variable for the upper decision maker, $\boldsymbol{y} \in R^m$ is a variable for the lower decision maker, $\boldsymbol{\zeta}$ is a continuous random variable, $\boldsymbol{\kappa} \in K \subset R^{s_1}$ is a perturbation parameter and K is a convex perturbation set. Let the constraint feasible set of all variables $(\boldsymbol{x},\boldsymbol{y})$ of the lower level problem be defined as

$$Y = \{(\boldsymbol{x},\boldsymbol{y}) \mid g_j(\boldsymbol{x},\boldsymbol{y}) \le 0, j = 1,2,\ldots,J\},$$

where $g_j : R^n \times R^m \to R^1$ are continuous. Let $\Xi_l = \{q(\boldsymbol{\zeta},\boldsymbol{\kappa}) | \boldsymbol{\kappa} \in K\}$ be the probability density distribution function cluster of the lower level problem, where $q(\boldsymbol{\zeta},\boldsymbol{\kappa})$ is a probability density distribution function for $\boldsymbol{\kappa} \in K$.

Ξ_u is the concave probability density distribution function cluster of the upper level problem, if any probability density distribution function $p(\boldsymbol{\xi},\boldsymbol{\tau}_1), p(\boldsymbol{\xi},\boldsymbol{\tau}_2) \in \Xi_u$, for all $\boldsymbol{\tau}_1,\boldsymbol{\tau}_2 \in T$ and $\forall t \in [0,1]$ such that the corresponding probability density distribution function $p(\boldsymbol{\xi}, t\boldsymbol{\tau}_1 + (1-t)\boldsymbol{\tau}_2) \in \Xi_u$ satisfies

$$tp(\boldsymbol{\xi},\boldsymbol{\tau}_1) + (1-t)p(\boldsymbol{\xi},\boldsymbol{\tau}_2) \le p(\boldsymbol{\xi}, t\boldsymbol{\tau}_1 + (1-t)\boldsymbol{\tau}_2)$$

for $t\boldsymbol{\tau}_1 + (1-t)\boldsymbol{\tau}_2 \in T$.

Ξ_l is the concave probability density distribution function cluster of the lower level problem, if any probability density distribution function $q(\boldsymbol{\zeta},\boldsymbol{\kappa}_1), q(\boldsymbol{\zeta},\boldsymbol{\kappa}_2) \in \Xi_u$, for all $\boldsymbol{\kappa}_1,\boldsymbol{\kappa}_2 \in K$ and $\forall t \in [0,1]$ such that the corresponding probability density distribution function $q(\boldsymbol{\zeta}, t\boldsymbol{\kappa}_1 + (1-t)\boldsymbol{\kappa}_2) \in \Xi_l$ satisfies

$$tq(\boldsymbol{\zeta},\boldsymbol{\kappa}_1) + (1-t)q(\boldsymbol{\zeta},\boldsymbol{\kappa}_2) \le q(\boldsymbol{\zeta}, t\boldsymbol{\kappa}_1 + (1-t)\boldsymbol{\kappa}_2)$$

for $t\boldsymbol{\kappa}_1 + (1-t)\boldsymbol{\kappa}_2 \in K$.

The probability density distribution function cluster doesn't have to be Wasserstein ball [7, 22], because the definition of the probability density distribution function cluster is defined without Wasserstein metric.The following example shows that there are many concave probability density distribution function clusters.

Example 2.1. Let a triangular distributional cluster be given $\Xi = \{p(\xi,\tau) | \tau \in T = [-\epsilon, \epsilon]\}$ where $0 < \epsilon < 1$, and a probability density distribution function be

$$p(\xi,\tau) = \begin{cases} \frac{1}{2}(1-\tau^2\xi), & \forall \xi \in [-1,1], \\ 0, & \forall \xi \notin [-1,1], \end{cases}$$

where τ is a perturbation variable, and ξ a random variable at different values $\tau \in T$. It is clear that $tp(\xi,\tau_1)+(1-t)p(\xi,\tau_2) \le p(\xi, t\tau_1 + (1-t)\tau_2) \in \Xi$ for all $\tau_1, \tau_2 \in T$ and $\forall t \in (0,1)$. So, Ξ is a concave probability density distribution function cluster.

Example 2.2. Let N probability density distribution functions: $p_i(\xi), i = 1, 2, \ldots, N$. Let a set

$$\Xi = \{p(\boldsymbol{\xi}, \boldsymbol{\tau}) = \sum_{i=1}^{N} \lambda_i p_i(\xi) | \sum_{i=1}^{N} \lambda_i = 1, \lambda_i \ge 0, i = 1, 2, \ldots, N\}.$$

Let $T = \{\boldsymbol{\tau} = (\lambda_1, \lambda_2, \ldots, \lambda_N) | \sum_{i=1}^{N} \lambda_i = 1, \lambda_i \ge 0, i = 1, 2, \ldots, N\}$, where T is convex. It is clear that $tp(\boldsymbol{\xi}, \boldsymbol{\tau}_1) + (1-t)p(\boldsymbol{\xi}, \boldsymbol{\tau}_2) \le p(\boldsymbol{\xi}, t\boldsymbol{\tau}_1 + (1-t)\boldsymbol{\tau}_2) \in \Xi$ for all $p(\boldsymbol{\xi}, \boldsymbol{\tau}_1), p(\boldsymbol{\xi}, \boldsymbol{\tau}_2) \in \Xi$ and $\forall t \in (0,1)$. So, Ξ is a concave probability density distribution function cluster.

A cumulative probability distribution function $\Psi^U_{\boldsymbol{\tau}}(\boldsymbol{x}, \boldsymbol{y}, u)$ of $F_{\boldsymbol{\tau}}(\boldsymbol{x}, \boldsymbol{y}, \boldsymbol{\xi})$ is defined as

$$\Psi^U_{\boldsymbol{\tau}}(\boldsymbol{x}, \boldsymbol{y}, u) = P\{F_{\boldsymbol{\tau}}(\boldsymbol{x}, \boldsymbol{y}, \boldsymbol{\xi}) \le u\} = \int_{F_{\boldsymbol{\tau}}(\boldsymbol{x}, \boldsymbol{y}, \boldsymbol{z}) \le u} p(\boldsymbol{z}, \boldsymbol{\tau}) d\boldsymbol{z}, \boldsymbol{\tau} \in T. \tag{2.1}$$

A cumulative probability distribution function $\Psi^L_{\boldsymbol{\kappa}}(\boldsymbol{x}, \boldsymbol{y}, v)$ of $G_{\boldsymbol{\kappa}}(\boldsymbol{x}, \boldsymbol{y}, \boldsymbol{\zeta})$ is defined as

$$\Psi^L_{\boldsymbol{\kappa}}(\boldsymbol{x}, \boldsymbol{y}, v) = P\{G_{\boldsymbol{\kappa}}(\boldsymbol{x}, \boldsymbol{y}, \boldsymbol{\zeta}) \le v\} = \int_{G_{\boldsymbol{\kappa}}(\boldsymbol{x}, \boldsymbol{y}, \boldsymbol{w}) \le v} q(\boldsymbol{w}, \boldsymbol{\kappa}) d\boldsymbol{w}, \boldsymbol{\kappa} \in K. \tag{2.2}$$

VaR loss value is defined for the upper decision maker and the lower decision maker respectively as follows.

Definition 2.3. Let α denote the confidence level with $\alpha \in (0,1)$ and

$$u_{\boldsymbol{\tau}}(\boldsymbol{x}, \boldsymbol{y}) = \min\{u \mid \Psi^U_{\boldsymbol{\tau}}(\boldsymbol{x}, \boldsymbol{y}, u) \ge \alpha\}, \qquad \boldsymbol{\tau} \in T. \tag{2.3}$$

Then $u_{\boldsymbol{\tau}}(\boldsymbol{x}, \boldsymbol{y})$ is called a α-VaR loss value of the upper decision maker at $(\boldsymbol{x}, \boldsymbol{y})$ under a confidence level α and perturbation $\boldsymbol{\tau}$. $u_{\boldsymbol{\tau}}(\boldsymbol{x}, \boldsymbol{y})$ represents the minimum guaranteed loss at decision $(\boldsymbol{x}, \boldsymbol{y})$ under the confidence level α, total of which representing the VaR loss of the upper decision maker.

Let β denote the confidence level with $\beta \in (0,1)$ and

$$v_{\boldsymbol{\kappa}}(\boldsymbol{x}, \boldsymbol{y}) = \min\{v \mid \Psi^L_{\boldsymbol{\kappa}}(\boldsymbol{x}, \boldsymbol{y}, v) \ge \beta\}, \qquad \boldsymbol{\kappa} \in K. \tag{2.4}$$

Then $v_{\boldsymbol{\kappa}}(\boldsymbol{x}, \boldsymbol{y})$ is called a β-VaR loss value of the lower decision maker at $(\boldsymbol{x}, \boldsymbol{y})$ under a confidence level β and perturbation $\boldsymbol{\kappa}$. $v_{\boldsymbol{\kappa}}(\boldsymbol{x}, \boldsymbol{y})$ represents the minimum guaranteed loss at decision $(\boldsymbol{x}, \boldsymbol{y})$ under the confidence level β, total of which representing the VaR loss of the lower decision maker.

When T is a point, the α-VaR defined above is consistent with the literature(Rockafellar, Uryasev,2002).

Define functions

$$\Phi^U_{\boldsymbol{\tau},\alpha}(\boldsymbol{x},\boldsymbol{y},u) = (1-\alpha)^{-1}\int_{F_{\boldsymbol{\tau}}(\boldsymbol{x},\boldsymbol{y},\boldsymbol{z})\geq u} F_{\boldsymbol{\tau}}(\boldsymbol{x},\boldsymbol{y},\boldsymbol{z})p(\boldsymbol{z},\boldsymbol{\tau})d\boldsymbol{z}, \qquad \boldsymbol{\tau}\in T \tag{2.5}$$

and

$$\Phi^L_{\boldsymbol{\kappa},\beta}(\boldsymbol{x},\boldsymbol{y},v) = (1-\beta)^{-1}\int_{G_{\boldsymbol{\kappa}}(\boldsymbol{x},\boldsymbol{y},\boldsymbol{w})\geq v} G_{\boldsymbol{\kappa}}(\boldsymbol{x},\boldsymbol{y},\boldsymbol{w})q(\boldsymbol{w},\boldsymbol{\kappa})d\boldsymbol{w}, \qquad \boldsymbol{\kappa}\in K. \tag{2.6}$$

WCVaR values of the upper decision maker and of the lower decision maker is defined respectively as follows.

Definition 2.4. Let α,β denote the confidence levels with $\alpha,\beta\in(0,1)$. Define

$$UWCVaR^{\alpha}_T(\boldsymbol{x},\boldsymbol{y}) \quad = \quad \max_{\boldsymbol{\tau}\in T}\Phi^U_{\boldsymbol{\tau},\alpha}(\boldsymbol{x},\boldsymbol{y},u_{\boldsymbol{\tau}}(\boldsymbol{x},\boldsymbol{y})) \tag{2.7}$$

and

$$LWCVaR^{\beta}_K(\boldsymbol{x},\boldsymbol{y}) \quad = \quad \max_{\boldsymbol{\kappa}\in K}\Phi^L_{\boldsymbol{\kappa},\beta}(\boldsymbol{x},\boldsymbol{y},v_{\boldsymbol{\kappa}}(\boldsymbol{x},\boldsymbol{y})). \tag{2.8}$$

$UWCVaR^{\alpha}_T(\boldsymbol{x},\boldsymbol{y})$ is called α-WCVaR of the upper decision maker at $(\boldsymbol{x},\boldsymbol{y})$ under confidence level α. $LWCVaR^{\beta}_K(\boldsymbol{x},\boldsymbol{y})$ is called β-WCVaR of the lower decision maker at $(\boldsymbol{x},\boldsymbol{y})$ under confidence level β.

Based on WCVaR, a distributionally robust bilevel programming model is defined as

$$\begin{array}{lll}
\text{(BRCVaR)} & \min & UWCVaR^{\alpha}_T(\boldsymbol{x},\boldsymbol{y}) = \max\limits_{\boldsymbol{\tau}\in T}\Phi^U_{\boldsymbol{\tau},\alpha}(\boldsymbol{x},\boldsymbol{y},u_{\boldsymbol{\tau}}(\boldsymbol{x},\boldsymbol{y}))\\
& s.t. & (\boldsymbol{x},\boldsymbol{y})\in X,\\
& & where\ \ \boldsymbol{y}\ is\ an\ optimal\ solution\ to\ the\\
& & lower\ optimization\ problem\ for\ an\ \boldsymbol{x},\\
& & \min\ LWCVaR^{\beta}_K(\boldsymbol{x},\boldsymbol{y}) = \max\limits_{\boldsymbol{\kappa}\in K}\Phi^L_{\boldsymbol{\kappa},\beta}(\boldsymbol{x},\boldsymbol{y},v_{\boldsymbol{\kappa}}(\boldsymbol{x},\boldsymbol{y}))\\
& & s.t.\ (\boldsymbol{x},\boldsymbol{y})\in Y,
\end{array}$$

where $\boldsymbol{x}$ is a variable of the upper decision-maker and $\boldsymbol{y}$ is a variable of the lower decision-maker. When $\boldsymbol{x}$ is fixed, all optimal solutions $\boldsymbol{y}(\boldsymbol{x})$ to the lower optimization problem min $LWCVaR^{\beta}_K(\boldsymbol{x},\boldsymbol{y})$ s.t. $(\boldsymbol{x},\boldsymbol{y})\in Y$ are first solved. Then, an optimal solution to the upper optimization problem

$$\min\ LWCVaR^{\beta}_K(\boldsymbol{x},\boldsymbol{y}(\boldsymbol{x}))\ \text{ s.t. }\ (\boldsymbol{x},\boldsymbol{y}(\boldsymbol{x}))\in X$$

is solved. In fact, it is difficult to calculate $UWCVaR^{\alpha}_T(\boldsymbol{x},\boldsymbol{y})$ and $LWCVaR^{\beta}_K(\boldsymbol{x},\boldsymbol{y})$. Two other loss functions are defined by

$$\Theta^U_{\boldsymbol{\tau},\alpha}(\boldsymbol{x},\boldsymbol{y},u) = u + (1-\alpha)^{-1}\int_{\boldsymbol{z}\in R^r}[F_{\boldsymbol{\tau}}(\boldsymbol{x},\boldsymbol{y},\boldsymbol{z})-u]^+p(\boldsymbol{z},\boldsymbol{\tau})d\boldsymbol{z}, \qquad \boldsymbol{\tau}\in T \tag{2.9}$$

and

$$\Theta^L_{\boldsymbol{\kappa},\beta}(\boldsymbol{x},\boldsymbol{y},v) = v + (1-\beta)^{-1}\int_{\boldsymbol{w}\in R^s}[G_{\boldsymbol{\kappa}}(\boldsymbol{x},\boldsymbol{y},\boldsymbol{w})-v]^+q(\boldsymbol{w},\boldsymbol{\kappa})d\boldsymbol{w}, \qquad \boldsymbol{\kappa}\in K, \tag{2.10}$$

where $[F_{\boldsymbol{\tau}}(\boldsymbol{x},\boldsymbol{y},\boldsymbol{z}) - z]^{+} = \max\{F_{\boldsymbol{\tau}}(\boldsymbol{x},\boldsymbol{y},\boldsymbol{z}) - z, 0\}$ and $[G_{\boldsymbol{\kappa}}(\boldsymbol{x},\boldsymbol{y},\boldsymbol{w}) - v]^{+} = \max\{G_{\boldsymbol{\kappa}}(\boldsymbol{x},\boldsymbol{y},\boldsymbol{w}) - v, 0\}$.

Function $\Phi^{U}_{\boldsymbol{\tau},\alpha}(\boldsymbol{x},\boldsymbol{y},u)$ and function $\Theta^{U}_{\boldsymbol{\tau},\alpha}(\boldsymbol{x},\boldsymbol{y},u)$ have the following relation:

$$\Theta^{U}_{\boldsymbol{\tau},\alpha}(\boldsymbol{x},\boldsymbol{y},u) \quad = \quad \Phi^{U}_{\boldsymbol{\tau},\alpha}(\boldsymbol{x},\boldsymbol{y},u) + u(1-\alpha)^{-1}[\Psi^{U}_{\boldsymbol{\tau}}(\boldsymbol{x},\boldsymbol{y},u) - \alpha].$$

When $1-\alpha = P\{F_{\boldsymbol{\tau}}(\boldsymbol{x},\boldsymbol{y},\boldsymbol{z}) \geq u\}$, we have $\Theta^{U}_{\boldsymbol{\tau},\alpha}(\boldsymbol{x},\boldsymbol{y},u) = \Phi^{U}_{\boldsymbol{\tau},\alpha}(\boldsymbol{x},\boldsymbol{y},u)$.

Function $\Phi^{L}_{\boldsymbol{\kappa},\beta}(\boldsymbol{x},\boldsymbol{y},v)$ and function $\Theta^{L}_{\boldsymbol{\kappa},\beta}(\boldsymbol{x},\boldsymbol{y},v)$ have the following relation:

$$\Theta^{L}_{\boldsymbol{\kappa},\beta}(\boldsymbol{x},\boldsymbol{y},v) \quad = \quad \Phi^{L}_{\boldsymbol{\kappa},\beta}(\boldsymbol{x},\boldsymbol{y},v) + v(1-\beta)^{-1}[\Psi^{L}_{\boldsymbol{\kappa}}(\boldsymbol{x},\boldsymbol{y},v) - \beta].$$

When $1-\beta = P\{G_{\boldsymbol{\kappa}}(\boldsymbol{x},\boldsymbol{y},\boldsymbol{w}) \geq v\}$, we have $\Theta^{L}_{\boldsymbol{\kappa},\beta}(\boldsymbol{x},\boldsymbol{y},v) = \Phi^{L}_{\boldsymbol{\kappa},\beta}(\boldsymbol{x},\boldsymbol{y},v)$.

It is always assumed that the following condition holds:

$$P\{F_{\boldsymbol{\tau}}(\boldsymbol{x},\boldsymbol{y},\boldsymbol{z}) = u\} = \int_{F_{\boldsymbol{\tau}}(\boldsymbol{x},\boldsymbol{y},\boldsymbol{z})=u} p(\boldsymbol{z},\boldsymbol{\tau})d\boldsymbol{z} = 0, \quad u \in R^{1}, \quad \boldsymbol{\tau} \in T$$

and

$$P\{G_{\boldsymbol{\kappa}}(\boldsymbol{x},\boldsymbol{y},\boldsymbol{w}) = v\} = \int_{G_{\boldsymbol{\kappa}}(\boldsymbol{x},\boldsymbol{y},\boldsymbol{w})=v} q(\boldsymbol{w},\boldsymbol{\kappa})d\boldsymbol{w} = 0, \quad v \in R^{1}, \quad \boldsymbol{\kappa} \in K$$

According to literature [16], the following conclusion holds.

Lemma 2.5. *For a given* $(\boldsymbol{x},\boldsymbol{y})$, $\boldsymbol{\tau} \in T$ *and* $\boldsymbol{\kappa} \in K$, *if* $\min_{u\in R^1} \Theta^{U}_{\boldsymbol{\tau},\alpha}(\boldsymbol{x},\boldsymbol{y},u)$ *has only one optimal solution* $\bar{u}$ *and* $\min_{v\in R^1} \Theta^{L}_{\boldsymbol{\kappa},\beta}(\boldsymbol{x},\boldsymbol{y},v)$ *has only one optimal solution* $\bar{v}$, *then*

$$\Phi^{U}_{\boldsymbol{\tau},\alpha}(\boldsymbol{x},\boldsymbol{y},\bar{u}) = \min_{u\in R^1} \Theta^{U}_{\boldsymbol{\tau},\alpha}(\boldsymbol{x},\boldsymbol{y},u), \tag{2.11}$$

$$\Phi^{L}_{\boldsymbol{\kappa},\beta}(\boldsymbol{x},\boldsymbol{y},\bar{v}) = \min_{v\in R^1} \Theta^{L}_{\boldsymbol{\kappa},\beta}(\boldsymbol{x},\boldsymbol{y},v), \tag{2.12}$$

where $\bar{u} = u_{\boldsymbol{\tau}}(\boldsymbol{x},\boldsymbol{y})$ *and* $\bar{v} = v_{\boldsymbol{\kappa}}(\boldsymbol{x},\boldsymbol{y})$.

Let us first prove the following conclusion.

Lemma 2.6. *For a given* $(\boldsymbol{x},\boldsymbol{y})$, *suppose that* $F_{\boldsymbol{\tau}}(\boldsymbol{x},\boldsymbol{y},\boldsymbol{z})$ *is quasi-concave on* $\boldsymbol{\tau} \in T$ *and* $G_{\boldsymbol{\kappa}}(\boldsymbol{x},\boldsymbol{y},\boldsymbol{z})$ *is quasi-concave on* $\boldsymbol{\kappa} \in K$, *then*

$$UWCVaR^{\alpha}_{T}(\boldsymbol{x},\boldsymbol{y}) = \max_{\boldsymbol{\tau}\in T} \min_{u\in R^1} \Theta^{U}_{\boldsymbol{\tau},\alpha}(\boldsymbol{x},\boldsymbol{y},u) = \min_{u\in R^1} \max_{\boldsymbol{\tau}\in T} \Theta^{U}_{\boldsymbol{\tau},\alpha}(\boldsymbol{x},\boldsymbol{y},u), \tag{2.13}$$

$$LWCVaR^{\beta}_{K}(\boldsymbol{x},\boldsymbol{y}) = \max_{\boldsymbol{\kappa}\in K} \min_{v\in R^1} \Theta^{L}_{\boldsymbol{\kappa},\beta}(\boldsymbol{x},\boldsymbol{y},v) = \min_{v\in R^1} \max_{\boldsymbol{\kappa}\in K} \Theta^{L}_{\boldsymbol{\kappa},\beta}(\boldsymbol{x},\boldsymbol{y},v). \tag{2.14}$$

Proof. According to Lemma 2.5,we have

$$UWCVaR^{\alpha}_{T}(\boldsymbol{x},\boldsymbol{y}) = \max_{\boldsymbol{\tau}\in T} \min_{u\in R^1} \Theta^{U}_{\boldsymbol{\tau},\alpha}(\boldsymbol{x},\boldsymbol{y},u).$$

According to literature [16], $\Theta^{U}_{\boldsymbol{\tau},\alpha}(\boldsymbol{x},\boldsymbol{y},u)$ is convex on u. Let any $\boldsymbol{\tau}_1, \boldsymbol{\tau}_2 \in T$, $\forall t \in [0,1]$, then we have $t\boldsymbol{\tau}_1 + (1-t)\boldsymbol{\tau}_2 \in T$. So, the corresponding probability density distribution function at $t\boldsymbol{\tau}_1 + (1-t)\boldsymbol{\tau}_2$ is $p(\boldsymbol{z}, t\boldsymbol{\tau}_1 + (1-t)\boldsymbol{\tau}_2)$. By Equation (2.9), we have

$$\begin{aligned} & \Theta^{U}_{t\boldsymbol{\tau}_1+(1-t)\boldsymbol{\tau}_2,\alpha}(\boldsymbol{x},\boldsymbol{y},u) \\ = \quad & [u + (1-\alpha)^{-1} \int_{\boldsymbol{z}\in R^m} (F_{t\boldsymbol{\tau}_1+(1-t)\boldsymbol{\tau}_2}(\boldsymbol{x},\boldsymbol{y},\boldsymbol{z}) - u)^{+} p(\boldsymbol{z}, t\boldsymbol{\tau}_1 + (1-t)\boldsymbol{\tau}_2)d\boldsymbol{z}] \\ \geq \quad & t\Theta_{\boldsymbol{\tau}_1,\alpha}(\boldsymbol{x},\boldsymbol{y},u) + (1-t)\Theta_{\boldsymbol{\tau}_2,\alpha}(\boldsymbol{x},\boldsymbol{y},u), \end{aligned}$$

where $t\boldsymbol{\tau}_1 + (1-t)\boldsymbol{\tau}_2 \in T$. Therefore, $\Theta^U_{\boldsymbol{\tau},\alpha}(\boldsymbol{x}, \boldsymbol{y}, u)$ is concave on $\boldsymbol{\tau}$. According to literature [19], (2.13) is true. Similarly, (2.14) is true. □

By Lemma 2.6 the following conclusion is obtained.

Lemma 2.7. *Suppose that* $F_{\boldsymbol{\tau}}(\boldsymbol{x}, \boldsymbol{y}, \boldsymbol{z})$ *is quasi-concave on* $\boldsymbol{\tau} \in T$ *and* $G_{\boldsymbol{\kappa}}(\boldsymbol{x}, \boldsymbol{y}, \boldsymbol{z})$ *is quasi-concave on* $\boldsymbol{\kappa} \in K$*, then*

$$\min_{(\boldsymbol{x},\boldsymbol{y})\in X} UWCVaR^{\alpha}_{T}(\boldsymbol{x}, \boldsymbol{y}) = \min_{(\boldsymbol{x},\boldsymbol{y})\in X} \min_{u\in R^1} \max_{\boldsymbol{\tau}\in T} \Theta^U_{\boldsymbol{\tau},\alpha}(\boldsymbol{x}, \boldsymbol{y}, u), \tag{2.15}$$

$$\min_{(\boldsymbol{x},\boldsymbol{y})\in Y} LWCVaR^{\beta}_{K}(\boldsymbol{x}, \boldsymbol{y}) = \min_{(\boldsymbol{x},\boldsymbol{y})\in Y} \min_{v\in R^1} \max_{\boldsymbol{\kappa}\in K} \Theta^L_{\boldsymbol{\kappa},\beta}(\boldsymbol{x}, \boldsymbol{y}, v). \tag{2.16}$$

Therefore, the following conclusion is proved by Lemma 2.7.

Lemma 2.8. *Suppose that* $F_{\boldsymbol{\tau}}(\boldsymbol{x}, \boldsymbol{y}, \boldsymbol{z})$ *is quasi-concave on* $\boldsymbol{\tau} \in T$ *and* $G_{\boldsymbol{\kappa}}(\boldsymbol{x}, \boldsymbol{y}, \boldsymbol{z})$ *is quasi-concave on* $\boldsymbol{\kappa} \in K$*. Then*

(i) solving $\min_{(\boldsymbol{x},\boldsymbol{y})\in X} UWCVaR^{\alpha}_{T}(\boldsymbol{x}, \boldsymbol{y})$ *is equivalent to solving the following problem:*

$$\begin{array}{lll} \textit{(FUWCVaR)} & \min & \gamma \\ & s.t. & \gamma \geq \Theta^U_{\boldsymbol{\tau},\alpha}(\boldsymbol{x}, \boldsymbol{y}, u), \qquad \forall \boldsymbol{\tau} \in \boldsymbol{T} \\ & & (\boldsymbol{x}, \boldsymbol{y}) \in X, \gamma \in R^1, u \in R^1, \end{array}$$

where $(\boldsymbol{x}, \boldsymbol{y}, \gamma, u)$ *are decision variables.*

(ii) solving $\min_{(\boldsymbol{x},\boldsymbol{y})\in Y} LWCVaR^{\beta}_{K}(\boldsymbol{x}, \boldsymbol{y})$ *is equivalent to solving the following problem:*

$$\begin{array}{lll} \textit{(FLWCVaR)} & \min & \delta \\ & s.t. & \delta \geq \Theta^L_{\boldsymbol{\kappa},\beta}(\boldsymbol{x}, \boldsymbol{y}, v), \qquad \forall \boldsymbol{\kappa} \in K \\ & & (\boldsymbol{x}, \boldsymbol{y}) \in Y, \delta \in R^1, v \in R^1, \end{array}$$

where $(\boldsymbol{x}, \boldsymbol{y}, \delta, v)$ *are decision variables.*

Proof. (i) By Lemma 2.7, we only need to prove that solving (FUWCVaR) is equivalent to solving $\min_{(\boldsymbol{x},\boldsymbol{y})\in X} \min_{u\in R^1} \max_{\boldsymbol{\tau}\in T} \Theta^U_{\boldsymbol{\tau},\alpha}(\boldsymbol{x}, \boldsymbol{y}, u)$. Suppose that $(\bar{\boldsymbol{x}}, \bar{\boldsymbol{y}}, \bar{\gamma}, \bar{u})$ is an optimal solution to (FUWCVaR). If $(\bar{\boldsymbol{x}}, \bar{\boldsymbol{y}}, \bar{u})$ is not an optimal solution to $\min_{(\boldsymbol{x},\boldsymbol{y})\in X} \min_{u\in R^1} \max_{\boldsymbol{\tau}\in T} \Theta^U_{\boldsymbol{\tau},\alpha}(\boldsymbol{x}, \boldsymbol{y}, u)$, then there are an $(\boldsymbol{x}', \boldsymbol{y}') \in X$ and $u \in R^1$, such that

$$\max_{\boldsymbol{\tau}\in T} \Theta^U_{\boldsymbol{\tau},\alpha}(\boldsymbol{x}', \boldsymbol{y}', u') < \max_{\boldsymbol{\tau}\in T} \Theta^U_{\boldsymbol{\tau},\alpha}(\bar{\boldsymbol{x}}, \bar{\boldsymbol{y}}, \bar{u}).$$

Let $\gamma' = \max_{\boldsymbol{\tau}\in T} \Theta^U_{\boldsymbol{\tau},\alpha}(\boldsymbol{x}', \boldsymbol{y}', u')$. Obviously, $\bar{\gamma} = \max_{\boldsymbol{\tau}\in T} \Theta^U_{\boldsymbol{\tau},\alpha}(\bar{\boldsymbol{x}}, \bar{\boldsymbol{y}}, \bar{u})$ and

$$\gamma' \geq \Theta^U_{\boldsymbol{\tau},\alpha}(\boldsymbol{x}', \boldsymbol{y}', u'), \forall \boldsymbol{\tau} \in \boldsymbol{T}.$$

Therefore, $(\boldsymbol{x}', \boldsymbol{y}', \gamma', u')$ is a feasible solution to (FUWCVaR) and $\gamma' < \bar{\gamma}$. So, $(\bar{\boldsymbol{x}}, \bar{\boldsymbol{y}}, \bar{\gamma}, \bar{u})$ is not a optimal solution to (FUWCVaR). This is a contradiction.

In turn, suppose that $(\bar{\boldsymbol{x}}, \bar{\boldsymbol{y}}, \bar{u})$ is an optimal solution to $\min_{(\boldsymbol{x},\boldsymbol{y})\in X} \min_{u\in R^1} \max_{\boldsymbol{\tau}\in T} \Theta^U_{\boldsymbol{\tau},\alpha}(\boldsymbol{x}, \boldsymbol{y}, u)$. Let $\bar{\gamma} = \max_{\boldsymbol{\tau}\in T} \Theta^U_{\boldsymbol{\tau},\alpha}(\bar{\boldsymbol{x}}, \bar{\boldsymbol{y}}, \bar{u})$. If $(\bar{\boldsymbol{x}}, \bar{\boldsymbol{y}}, \bar{\gamma}, \bar{u})$ is not a optimal solution to (FUWCVaR), then there are an $(\boldsymbol{x}', \boldsymbol{y}') \in X$ and (γ', u') such that $\gamma' < \bar{\gamma}$, where

$$\gamma' \geq \Theta^U_{\boldsymbol{\tau},\alpha}(\boldsymbol{x}', \boldsymbol{y}', u'), \forall \boldsymbol{\tau} \in \boldsymbol{T}.$$

Specifically,

$$\max_{\boldsymbol{\tau}\in T} \Theta^U_{\boldsymbol{\tau},\alpha}(\boldsymbol{x}', \boldsymbol{y}', u') \leq \gamma' < \bar{\gamma}.$$

Clearly $(\bar{\boldsymbol{x}}, \bar{\boldsymbol{y}}, \bar{u})$ is not an optimal solution to$\min_{(\boldsymbol{x},\boldsymbol{y})\in X} \min_{u\in R^1} \max_{\boldsymbol{\tau}\in T} \Theta^U_{\boldsymbol{\tau},\alpha}(\boldsymbol{x}, \boldsymbol{y}, u)$. Thus, a contradiction occurs.

Similarly, (ii) is true. □

By Lemma 2.5 and 2.8, if $(\bar{\boldsymbol{x}}, \bar{\boldsymbol{y}}, \bar{\gamma}, \bar{u})$ is a optimal solution to (FUWCVaR), then there is a $\bar{\boldsymbol{\tau}} \in T$ such that $\bar{\gamma} = \Theta^U_{\boldsymbol{\tau},\alpha}(\bar{\boldsymbol{x}}, \bar{\boldsymbol{y}}, \bar{u})$, and $\bar{u} = u_{\bar{\boldsymbol{\tau}}}(\bar{\boldsymbol{x}}, \bar{\boldsymbol{y}})$ is the VaR risk loss value. If $(\bar{\boldsymbol{x}}, \bar{\boldsymbol{y}}, \bar{\delta}, \bar{v})$ is a optimal solution to (FLWCVaR), then there is a $\bar{\boldsymbol{\kappa}} \in K$ such that $\bar{\delta} = \Theta^L_{\boldsymbol{\kappa},\beta}(\bar{\boldsymbol{x}}, \bar{\boldsymbol{y}}, \bar{v})$, and $\bar{v} = v_{\bar{\boldsymbol{\kappa}}}(\bar{\boldsymbol{x}}, \bar{\boldsymbol{y}})$ is the VaR risk loss value.

So, by Lemma 2.8, the following theorem is obtained.

Theorem 2.9. *Suppose that $F_{\boldsymbol{\tau}}(\boldsymbol{x}, \boldsymbol{y}, \boldsymbol{z})$ is quasi-concave on $\boldsymbol{\tau} \in T$ and $G_{\boldsymbol{\kappa}}(\boldsymbol{x}, \boldsymbol{y}, \boldsymbol{z})$ is quasi-concave on $\boldsymbol{\kappa} \in K$. Then solving (BRCVaR) is equivalent to solving the following problem:*

$$
\begin{array}{lll}
\textit{(FBRCVaR)} & \min & \gamma \\
& s.t. & \gamma \geq \Theta^U_{\boldsymbol{\tau},\alpha}(\boldsymbol{x}, \boldsymbol{y}, u), \qquad \forall \boldsymbol{\tau} \in \boldsymbol{T} \\
& & (\boldsymbol{x}, \boldsymbol{y}) \in X, \gamma \in R^1, u \in R^1, \\
& & \textit{where } (\boldsymbol{y}, \delta, v) \textit{ is an optimal solution to the} \\
& & \textit{lower optimization problem at the tuple } (\boldsymbol{x}, \gamma, u) \\
& & \min \ \delta \\
& s.t. & \delta \geq \Theta^L_{\boldsymbol{\kappa},\beta}(\boldsymbol{x}, \boldsymbol{y}, v), \qquad \forall \boldsymbol{\kappa} \in K \\
& & (\boldsymbol{x}, \boldsymbol{y}) \in Y, \delta \in R^1, v \in R^1.
\end{array}
$$

Theorem 2.9 shows that (FBRCVaR) can be used to describe the actual robust bilevel decision model. But, given that (FBRCVaR) is a semi-infinite programming problem, it is very difficult to solve it directly. In practical applications, there is a finite number of probability density distribution functions and of the corresponding loss functions for bilevel decision problems.

For the upper optimization problem, assume that there exists a finite number of probability density distribution functions $\{p_k(\boldsymbol{\xi})|k = 1, 2, \ldots, N\}$ and the corresponding loss function $F_k(\boldsymbol{x}, \boldsymbol{y}, \boldsymbol{\xi})(k = 1, 2, \ldots, N)$. A mixed probability density distribution function cluster is defined by

$$
\Xi_N = \{p(\boldsymbol{\xi}, \boldsymbol{\tau}) = \sum_{k=1}^{N} \lambda_k p_k(\boldsymbol{\xi}) | \boldsymbol{\tau} \in T\}, \tag{2.17}
$$

where $T = \{\boldsymbol{\tau} = (\lambda_1, \lambda_2, \ldots, \lambda_N) | \sum_{k=1}^{N} \lambda_k = 1, \lambda_k \geq 0, k = 1, 2, \ldots, N\}$. Obviously, Ξ_N is a concave probability density distribution function cluster that is similar to that in literature(Zhu et al,2009). For $\boldsymbol{\tau} \in T$, the corresponding loss function is defined by

$$
F_{\boldsymbol{\tau}}(\boldsymbol{x}, \boldsymbol{y}, \boldsymbol{\xi}) = \{\sum_{k_1=1}^{N} \lambda_{k_1} F_{k_1}(\boldsymbol{x}, \boldsymbol{y}, \boldsymbol{\xi}) \mid \boldsymbol{\tau} \in T\}. \tag{2.18}
$$

It is clear that $F_{\boldsymbol{\tau}}(\boldsymbol{x}, \boldsymbol{y}, \boldsymbol{\xi})$ is quasi-concave on $\boldsymbol{\tau} \in T$. For $k, k_1 = 1, 2, \ldots, N$ and $\alpha \in (0, 1)$, define

$$
\Theta^U_{k,k_1,\alpha}(\boldsymbol{x}, \boldsymbol{y}, u) = u + (1 - \alpha)^{-1} \int_{\boldsymbol{z}\in R^r} (F_{k_1}(\boldsymbol{x}, \boldsymbol{y}, \boldsymbol{z}) - u)^+ p_k(\boldsymbol{z}) d\boldsymbol{z}. \tag{2.19}
$$

For the lower optimization problem, assume that there exists a finite number of probability density distribution functions $\{q_k(\boldsymbol{\zeta})|k = 1, 2, \ldots, M\}$ and the corresponding loss function $G_k(\boldsymbol{x}, \boldsymbol{y}, \boldsymbol{\zeta})(k = 1, 2, \ldots, M)$. A mixed probability density distribution function cluster is defined as follows:

$$\Xi_M = \{q(\boldsymbol{\zeta}, \boldsymbol{\kappa}) = \sum_{k=1}^{M} \mu_k q_k(\boldsymbol{\zeta})|\boldsymbol{\kappa} \in K\}, \tag{2.20}$$

where $K = \{\boldsymbol{\kappa} = (\mu_1, \mu_2, \ldots, \mu_N)| \sum_{k=1}^{N} \mu_k = 1, \mu_k \geq 0, k = 1, 2, \ldots, M\}$. Obviously, Ξ_M is a concave probability density distribution function cluster that is similar to that in literature(Zhu et al,2009). For $\boldsymbol{\kappa} \in K$, the corresponding loss function is defined by

$$G_{\boldsymbol{\kappa}}(\boldsymbol{x}, \boldsymbol{y}, \boldsymbol{\zeta}) = \{\sum_{k_1=1}^{M} \mu_{k_1} G_{k_1}(\boldsymbol{x}, \boldsymbol{y}, \boldsymbol{\zeta}) \mid \boldsymbol{\kappa} \in K\}. \tag{2.21}$$

It is clear that $G_{\boldsymbol{\kappa}}(\boldsymbol{x}, \boldsymbol{y}, \boldsymbol{\zeta})$ is quasi-concave on $\boldsymbol{\kappa} \in K$. For $k, k_1 = 1, 2, \ldots, M$ and $\beta \in (0, 1)$, define

$$\Theta^L_{k,k_1,\beta}(\boldsymbol{x}, \boldsymbol{y}, v) = v + (1 - \alpha)^{-1} \int_{\boldsymbol{w} \in R^s} (G_{k_1}(\boldsymbol{x}, \boldsymbol{y}, \boldsymbol{w}) - v)^+ q_k(\boldsymbol{w}) d\boldsymbol{w}. \tag{2.22}$$

We have the following results.

Lemma 2.10. *Suppose that (2.17),(2.18),. . .,(2.22) are true. Then*

(1) *solving*

$$\begin{array}{llll} \text{(FUWCVaR)} & \min & \gamma & \\ & s.t. & \gamma \geq \Theta^U_{\boldsymbol{\tau},\alpha}(\boldsymbol{x}, \boldsymbol{y}, u), & \forall \boldsymbol{\tau} \in \boldsymbol{T} \\ & & (\boldsymbol{x}, \boldsymbol{y}) \in X, \gamma \in R^1, u \in R^1, & \end{array}$$

is equivalent to solving the following problem:

$$\begin{array}{llll} \text{(FUWCVaR1)} & \min & \gamma & \\ & s.t. & \gamma \geq \Theta^U_{k,k_1,\alpha}(\boldsymbol{x}, \boldsymbol{y}, u), & k, k_1 = 1, 2, \ldots, N, \\ & & (\boldsymbol{x}, \boldsymbol{y}) \in X, \gamma \in R^1, u \in R^1, & \end{array}$$

where $(\boldsymbol{x}, \boldsymbol{y}, \gamma, u)$ are decision variables.

(2) *solving*

$$\begin{array}{llll} \text{(FLWCVaR)} & \min & \delta & \\ & s.t. & \delta \geq \Theta^L_{\boldsymbol{\kappa},\beta}(\boldsymbol{x}, \boldsymbol{y}, v), & \forall \boldsymbol{\kappa} \in K \\ & & (\boldsymbol{x}, \boldsymbol{y}) \in Y, \delta \in R^1, v \in R^1, & \end{array}$$

is equivalent to solving the following problem:

$$\begin{array}{llll} \text{(FLWCVaR1)} & \min & \delta & \\ & s.t. & \delta \geq \Theta^L_{k,k_1,\beta}(\boldsymbol{x}, \boldsymbol{y}, v), & k, k_1 = 1, 2, \ldots, M, \\ & & (\boldsymbol{x}, \boldsymbol{y}) \in Y, \delta \in R^1, v \in R^1, & \end{array}$$

where $(\boldsymbol{x}, \boldsymbol{y}, \delta, v)$ are decision variables.

Proof. If we are to prove solving (FUWCVaR) is equivalent to solving (FUWCVaR1), we only need to prove that (FUWCVaR) and (FUWCVaR1) have the same set of feasible solutions. Suppose that $(\boldsymbol{x}, \boldsymbol{y}, \gamma, u)$ is a feasible solution to (FUWCVaR1), then we have

$$\gamma \geq \Theta^U_{k,k_1,\alpha}(\boldsymbol{x}, \boldsymbol{y}, u), \qquad k, k_1 = 1, 2, \ldots, N.$$

Let us take $\lambda_{k_1}, \lambda_{k_1} \geq 0, k, k_1 = 1, 2, \ldots, N, \sum_{k=1}^N \lambda_k = 1, \sum_{k_1=1}^N \lambda_{k_1} = 1$; when the two sides of the above inequality are multiplied by λ_k and λ_{k_1} , they can be added together as follows:

$$\gamma \geq \sum_{k=1}^N \lambda_k \sum_{k_1=1}^N \lambda_{k_1} \Theta^U_{k,k_1,\alpha}(\boldsymbol{x}, \boldsymbol{y}, u).$$

By (2.9), for all $\boldsymbol{\tau} \in T$, we have

$$\begin{aligned}
\Theta^U_{\boldsymbol{\tau},\alpha}(\boldsymbol{x}, \boldsymbol{y}, u) &= u + (1-\alpha)^{-1} \int_{\boldsymbol{z} \in R^r} [F_{\boldsymbol{\tau}}(\boldsymbol{x}, \boldsymbol{y}, \boldsymbol{z}) - u]^+ p(\boldsymbol{z}, \boldsymbol{\tau}) d\boldsymbol{z} \\
&= u + (1-\alpha)^{-1} \int_{\boldsymbol{z} \in R^r} [\sum_{k_1=1}^N \lambda_{k_1} F_{k_1}(\boldsymbol{x}, \boldsymbol{y}, \boldsymbol{z}) - u]^+ \sum_{k=1}^N \lambda_k p_k(\boldsymbol{z}) d\boldsymbol{z} \\
&= \sum_{k=1}^N \lambda_k (u + (1-\alpha)^{-1} \int_{\boldsymbol{z} \in R^r} [\sum_{k_1=1}^N \lambda_{k_1} (F_{k_1}(\boldsymbol{x}, \boldsymbol{y}, \boldsymbol{z}) - u)]^+ p_k(\boldsymbol{z}) d\boldsymbol{z}) \\
&\leq \sum_{k=1}^N \lambda_k \sum_{k_1=1}^N \lambda_{k_1} \Theta^U_{k,k_1,\alpha}(\boldsymbol{x}, \boldsymbol{y}, u).
\end{aligned}$$

So, we have $\gamma \geq \Theta^U_{\boldsymbol{\tau},\alpha}(\boldsymbol{x}, \boldsymbol{y}, u)$. Hence, $(\boldsymbol{x}, \boldsymbol{y}, \gamma, u)$ is a feasible solution to (FUWCVaR).

In turn, let $(\boldsymbol{x}, \boldsymbol{y}, \gamma, u)$ be a feasible solution to (FUWCVaR), we have

$$\begin{aligned}
\gamma &\geq \Theta^U_{\boldsymbol{\tau},\alpha}(\boldsymbol{x}, \boldsymbol{y}, u) \\
&= \sum_{k=1}^N \lambda_k (u + (1-\alpha)^{-1} \int_{\boldsymbol{z} \in R^r} [\sum_{k_1=1}^N \lambda_{k_1} (F_{k_1}(\boldsymbol{x}, \boldsymbol{y}, \boldsymbol{z}) - u)]^+ p_k(\boldsymbol{z}) d\boldsymbol{z}).
\end{aligned}$$

Let $\lambda_i = 1, \lambda_j = 1$ where $i = k$ and $j = k_1$; $\lambda_i = 0, \lambda_j = 0$ where $i = 1, 2, \ldots, N$ and $j = 1, 2, \ldots, N$ but $i \neq k$ and $j \neq k_1$, then the above inequality becomes

$$\gamma \geq \Theta^U_{k,k_1,\alpha}(\boldsymbol{x}, \boldsymbol{y}, u).$$

Hence, $(\boldsymbol{x}, \boldsymbol{y}, \gamma, u)$ is a feasible solution to (FUWCVaR1) when $\lambda_i = 1$ and $\lambda_j = 1$, where i and j traverses $1, 2, \ldots, N$.

Now, by Lemma 2.10, we obtain the following conclusion. □

Theorem 2.11. *Suppose that (2.17),(2.18),...,(2.22) are true. Then solving (BRCVaR)* *is equivalent to solving the following problem:*

$$
\begin{array}{lll}
\textit{(FBRCVaR1)} & \min & \gamma \\
 & s.t. & \gamma \geq \Theta^{U}_{k,k_1,\alpha}(\boldsymbol{x},\boldsymbol{y},u), \qquad k,k_1=1,2,\ldots,N, \\
 & & (\boldsymbol{x},\boldsymbol{y}) \in X, \gamma \in R^1, u \in R^1, \\
 & & \textit{where } (\boldsymbol{y},\delta,v) \textit{ is an optimal solution to the} \\
 & & \textit{lower optimization problem at the tuple } (\boldsymbol{x},\gamma,u) \\
 & & \min \ \delta \\
 & s.t. & \delta \geq \Theta^{L}_{k,k_1,\beta}(\boldsymbol{x},\boldsymbol{y},v), \qquad k,k_1=1,2,\ldots,M, \\
 & & (\boldsymbol{x},\boldsymbol{y}) \in Y, \delta \in R^1, v \in R^1.
\end{array}
$$

Theorem 2.11 shows that if the finite probability (density) distribution functions and the corresponding loss functions are known, then their mixing distribution and finite distribution have the same loss WCVaR value. Therefore, we can compute (FBRCVaR1) approximately for a practical problem.

For the upper optimization problem, take an approximation of sample points $\{\boldsymbol{z}_{k,k_1,k_2} \mid k_2 = 1,2,\ldots,N_1;\ k,k_1 = 1,2,\ldots,N\}$,

$$
\begin{aligned}
\Theta^{U}_{k,k_1,\alpha}(\boldsymbol{x},\boldsymbol{y},u) &= u + (1-\alpha)^{-1}\int_{\boldsymbol{z}\in R^r} (F_{k_1}(\boldsymbol{x},\boldsymbol{y},\boldsymbol{z}) - u)^{+} p_k(\boldsymbol{z}) d\boldsymbol{z} \\
&\approx u + (1-\alpha)^{-1} N_1^{-1} \sum_{k_2=1}^{N_1} (F_{k_1}(\boldsymbol{x},\boldsymbol{y},\boldsymbol{z}_{k,k_1,k_2}) - u)^{+}.
\end{aligned}
$$

For the lower optimization problem, take an approximation of sample points $\{\boldsymbol{w}_{k,k_1,k_2} \mid k_2 = 1,2,\ldots,M_1;\ k,k_1 = 1,2,\ldots,M\}$,

$$
\begin{aligned}
\Theta^{L}_{k,k_1,\beta}(\boldsymbol{x},\boldsymbol{y},v) &= v + (1-\beta)^{-1}\int_{\boldsymbol{w}\in R^s} (G_{k_1}(\boldsymbol{x},\boldsymbol{y},\boldsymbol{w}) - v)^{+} q_k(\boldsymbol{w}) d\boldsymbol{w} \\
&\approx v + (1-\beta)^{-1} M_1^{-1} \sum_{k_2=1}^{M_1} (G_{k_1}(\boldsymbol{x},\boldsymbol{y},\boldsymbol{z}_{k,k_1,k_2}) - v)^{+}.
\end{aligned}
$$

So, (FBRCVaR1) can be approximately solved by the following problem:

$$
\begin{array}{lll}
\text{(FBRCVaR2)} & \min & \gamma \\
 & s.t. & \gamma \geq u + (1-\alpha)^{-1} N_1^{-1} \sum_{k_2=1}^{N_1} (F_{k_1}(\boldsymbol{x},\boldsymbol{y},\boldsymbol{z}_{k,k_1,k_2}) - u)^{+}, k, \\
 & & \qquad k_1 = 1,2,\ldots,N, \\
 & & (\boldsymbol{x},\boldsymbol{y}) \in X, \gamma \in R^1, u \in R^1, \\
 & & \textit{where } (\boldsymbol{y},\delta,v) \textit{ is an optimal solution to the} \\
 & & \textit{lower optimization problem at the tuple } (\boldsymbol{x},\gamma,u) \\
 & & \min \ \delta \\
 & s.t. & \delta \geq v + (1-\beta)^{-1} M_1^{-1} \sum_{k_2=1}^{M_1} (G_{k_1}(\boldsymbol{x},\boldsymbol{y},\boldsymbol{z}_{k,k_1,k_2}) - v)^{+}, k, \\
 & & \qquad k_1 = 1,2,\ldots,M, \\
 & & (\boldsymbol{x},\boldsymbol{y}) \in Y, \delta \in R^1, v \in R^1.
\end{array}
$$

Therefore, we can obtain the approximate robust equilibrium solution to the original problem (BRCVaR) by solving the nonlinear bilevel optimization problem (FBRCVaR2). When all the functions $F_{k_1}, G_{k_1}, f_i, g_j$ are linear, (FBRCVaR2) becomes a linear bilevel programming. When all the functions are convex, the exact penalty function method can be used to solve (FBRCVaR2). The penalty optimization problem of (FBRCVaR2) is defined by

$$
\begin{aligned}
\text{(FBRCVaR2)}(\rho,\varrho)\quad \min\quad & \tilde{F}(\boldsymbol{x},\boldsymbol{y},\gamma,u,\rho) = \gamma + \rho\sum_{i=1}^{I}\max\{f_i(\boldsymbol{x},\boldsymbol{y}),0\} \\
& +\rho\sum_{k,k_1=1,2,\dots,N}^{N}\max\{-\gamma+u+(1-\alpha)^{-1}N_1^{-1} \\
& \qquad\sum_{k_2=1}^{N_1}(F_{k_1}(\boldsymbol{x},\boldsymbol{y},\boldsymbol{z}_{k,k_1,k_2})-u)^+,0\} \\
s.t.\quad & (\boldsymbol{x},\boldsymbol{y})\in R^n\times R^m, \gamma\in R^1, u\in R^1, \\
& \textit{where } (\boldsymbol{y},\delta,v) \textit{ is an optimal solution to the} \\
& \textit{lower optimization problem at the tuple } (\boldsymbol{x},\gamma,u) \\
& \min\ \tilde{G}(\boldsymbol{x},\boldsymbol{y},\delta,v,\varrho) = \delta + \varrho\sum_{i=1}^{I}\max\{h_j(\boldsymbol{x},\boldsymbol{y}),0\} \\
& +\varrho\sum_{k,k_1=1,2,\dots,M}^{M}\max\{-\delta+v+(1-\beta)^{-1}M_1^{-1} \\
& \qquad\sum_{k_2=1}^{M_1}(G_{k_1}(\boldsymbol{x},\boldsymbol{y},\boldsymbol{z}_{k,k_1,k_2})-v)^+,0\} \\
s.t.\quad & (\boldsymbol{x},\boldsymbol{y})\in R^n\times R^m, \delta\in R^1, v\in R^1,
\end{aligned}
$$

where $(\rho,\varrho) > 0$ is penalty parameter. By using the smoothing function

$$
p_\epsilon(t) = \begin{cases} 0, & if \quad t\le 0, \\ \frac{t^3}{6\epsilon^2}, & if \quad 0\le t\le\epsilon, \\ t+\frac{\epsilon^2}{2t}-\frac{4\epsilon}{3}, & if \quad t\ge\epsilon, \end{cases}
$$

which is second-order continuously differentiable in [20] and $\lim_{\epsilon\to 0} p_\epsilon(t) = \max\{t,0\}$, the problem (FBRCVaR2)$(\rho,\varrho)$ is defined by

$$
\begin{aligned}
\text{(FBRCVaR2)}(\rho,\varrho)\quad \min\quad & \tilde{F}_\epsilon(\boldsymbol{x},\boldsymbol{y},\gamma,u,\rho)+\rho\|\nabla\tilde{G}_\epsilon(\boldsymbol{x},\boldsymbol{y},\delta,v,\varrho)\|^2 \\
s.t.\quad & (\boldsymbol{x},\boldsymbol{y})\in R^n\times R^m, \delta\in R^1, v\in R^1.\gamma\in R^1, u\in R^1,
\end{aligned}
$$

where the smoothing function $\tilde{F}$ and $\tilde{G}$ are defined respectively by

$$\begin{aligned}\tilde{F}_\epsilon(\boldsymbol{x},\boldsymbol{y},u,\rho) &= \gamma+\rho\sum_{i=1}^{I}p_\epsilon(f_i(\boldsymbol{x},\boldsymbol{y}))+\rho\sum_{k,k_1=1,2,\ldots,N}^{N}p_\epsilon(-\gamma\\ &\quad +u+(1-\alpha)^{-1}N_1^{-1}\sum_{k_2=1}^{N_1}p_\epsilon(F_{k_1}(\boldsymbol{x},\boldsymbol{y},\boldsymbol{z}_{k,k_1,k_2})-u)),\\ \tilde{G}(\boldsymbol{x},\boldsymbol{y},u,\rho) &= \delta+\varrho\sum_{i=1}^{I}p_\epsilon(h_j(\boldsymbol{x},\boldsymbol{y}))+\varrho\sum_{k,k_1=1,2,\ldots,M}^{M}\\ &\quad p_\epsilon(-\delta+v+(1-\beta)^{-1}M_1^{-1}\sum_{k_2=1}^{M_1}p_\epsilon(G_{k_1}(\boldsymbol{x},\boldsymbol{y},\boldsymbol{z}_{k,k_1,k_2})-v)).\end{aligned}$$

Approximate solution to (FBRCVaR2) is solved by (FBRCVaR2)(ρ,ϱ). As shown in [21], the numerical results are obtained by the penalty function method.

3 Conclusion

In this study, a concave probability density distribution function cluster with a convex perturbation set is defined. Then, VaR loss value and WCVaR value are defined respectively for the upper decision maker and the lower decision maker in a two-level decision system. A distributionally robust bilevel programming (BRCVaR) based on WCVaR with the concave probability density distribution cluster is defined. Next, it is proved that the distributionally robust bilevel programming (BRCVaR) based on WCVaR is equivalent to a robust bilevel programming with a perturbation parameter set. When the concave probability density distribution function cluster is a finite number of mixed probability density distribution functions, (BRCVaR) is proved equivalent to bilevel programming (FBRCVaR1) with inequality constraints. In particular, this model (BRCVaR) can be solved approximately by the nonlinear bilevel optimization problem (FBRCVaR2) when the finite sample data is known.

The future work will mainly include the following three aspects. (1) When the objective function and constraint functions are convex in (BRCVaR) model, its optimization condition and dual theory should be studied. (2) When the objective function and constraint functions are convex in (BRCVaR) model, its algorithm is worth studying. (3) The application of (BRCVaR) model in supply chain is worth studying. In addition, concepts such as distributionally robust constraints under Wasserstein distance in [4, 12] shall be introduced to the study of (BRCVaR) model.

Acknowledgements

The authors would like to express their gratitudes to anonymous referees' detailed comments and remarks that help us improve our presentation of this paper considerably.

References

[1] G.J.Alexander and A.M. Baptista, A comparison of VaR and CVaR constraints on portfolio selection with the mean-variance model, *Management Science* 50 (2004) 1261–1273.

[2] S.S.M. Bagher, M. Jamal and S.K. Miadreza, Risk-based bi-level model for simultaneous profit maximization of a smart distribution company and electric vehicle parking lot owner *Energies* 10 (2017)1–11.

[3] S.S.M. Bagher, M. Jamal and S.K. Miadreza, Bi-level model for operational scheduling of a distribution company that supplies electric vehicle parking lots, *Electric Power Systems Research*174 (2019) 1–11.

[4] J.G. Carlsson and Y. Wang, Distributions with maximum spread subject to Wasserstein distance constraints, *Journal of the Operations Research Society of China* 7 (2019) 69-106.

[5] E. Delage and Y. Ye, Distributional robust optimization under moment uncertainty with application to data-driven problems, *Operations Research* 58 (2010) 595–612.

[6] H. Golpira, S. Bahramara and S.A.R. Khan, Robust bi-level risk-based optimal scheduling of microgrid operation against uncertainty, *Rairo-Operations Research* 54 (2020) 993–1012.

[7] G.A. Hanasusanto and D. Kuhn, Conic programming reformulations of two-stage distributionally robust linear programs over Wasserstein balls, *Operations Research* 66 (2018) 849–869.

[8] G. Heris, N. Esmaeil and Z. Mostafa, Robust bi-level optimization for green opportunistic supply chain network design problem against uncertainty and environmental, *Computers & Industrial Engineering* 107 (2017) 301–312.

[9] R.K. Homa, V.D. Mostafa and N.H. Reza, A stochastic bi-level scheduling approach for the participation of EV aggregators in competitive electricity markets, *Applied Sciences-Basel* 7(2017) 1–16.

[10] D.S. Huang, S.S. Zhu and F.J. Fabozzi, Portfolio selection with uncertain exit time: a robust CVaR approach, *Journal of Economic Dynamics & Control* 32 (2008) 594-623.

[11] D.S. Huang, S.S. Zhu and F.J. Fabozzi, Portfolio selection under distributional uncertainty: a relative robust CVaR approach, *European Journal of Operational Research* 203 (2010) 185–194.

[12] Y. Mei, Z.P. Chen, B.B. Ji, Z.J. Xu and J. Liu, Data-driven stochastic programming with distributionally robust constraints under Wasserstein distance: asymptotic properties, *Journal of the Operations Research Society of China* 9 (2021) 525–542.

[13] R.Z. Qiu, J. Shang and X.Y. Huang, Robust inventory decision under distribution uncertainty: a CVaR-based optimization approach, *International Journal of Production Economics* 153 (2014) 13–23.

[14] K.H. Rashidizadeh and D.M. Vahedipour, A regret-based stochastic bi-level framework for scheduling of DR aggregator under uncertainties, *IEEE Transactions on Smart Grid* 11 (2020) 3171–3184.

[15] K.H. Rashidizadeh, D.M. Vahedipour and S.K. Miadreza, A bi-level risk-constrained offering strategy of a wind power producer considering demand side resources, *International Journal of Electrical Power & Energy Systems* 104 (2019) 562–574.

[16] R.T. Rockefeller and S. Uryasev, Optimization of conditional value-at-risk, *Journal of Risk* 2 (1999) 21–41.

[17] Z. Saeidi, R. Tavakkoli and M. Navabakhsh, A bi-level and robust optimization-based framework for a hazardous waste management problem: a real-world application *Journal of Cleaner Production* 252 (2020) 1–17.

[18] H. Scarf, A min-max solution of an inventory problem, in: *Studies in the Mathematical Theory of Inventory and Production*, K. Arrow, S. Karlin and H. Scarf (eds.), Stanford University Press, Silver, E.A., 1958, pp. 201–209.

[19] M. Sion, On general minimax theorems, *Pacific J. Mathematics* 8 (1958) 171–176.

[20] X.S. Xu, Z.Q. Meng, J.W. Sun, L.G. Huang and R. Shen, A second-order smooth penalty function algorithm for constrained optimization problems, *Computationaloptimization and Applications* 55 (2013)155–172.

[21] L.Y. Xu and Z.Q. Meng, Multi-loss WCVaR risk decision optimization based on weight for centralized supply problem of direct chain enterprises, *Asia-Pacific Journal of Operational Research* 36 (2019): 1940007.

[22] M.C. Yue, D. Kuhn and W. Wiesemann, On linear optimization over Wasserstein balls, *Mathematical Programming* 2021(Online).

[23] S.S. Zhu and M. Fukushima, Worst-Case Conditional Value-at-Risk with Application to Robust Portfolio Management, *Operations Research* 57 (2009) 1155–1168.

[24] S.S. Zhu, D. Li and S.Y. Wang, Robust Portfolio Selection under Downside Risk Measures, *Quantitative Finance* 9 (2009) 869–885.

Manuscript received 22 August 2021
revised 3 January 2022
accepted for publication 26 January 2022

Leiyan Xu
Institute Of Digital Business,
Nanjing Vocational College Of Information Technology
Nanjing,Jiangsu, 210023, China
E-mail address: Xuly@Njcit.Cn

Rui Shen
School Of Economics
Zhejiang University Of Technology
Hangzhou, Zhejiang, 310023, China
E-mail address: Mengzhiqing@Zjut.Edu.Cn

Zhiqing Meng
School of Management
Zhejiang University of Technology
Hangzhou, Zhejiang, 310023, China
E-mail address: shenrui126@126.com