On the Equivalence and Performance of Distributionally Robust Optimization and Robust Satisficing Models

Zhiyuan Wang, Lun Ran, Minglong Zhou*, Long He

Appendix. Proofs and Details

A. Proofs

Proof of Proposition 1 If the optimal solution to the problem $DRO(\theta)$ is $(\boldsymbol{x}^*, k^*, \tau^*)$, then fixing $\tau = \tau^*$, $(\boldsymbol{x}^*, k^*)$ will be the optimal solution to the problem $RS(\tau^*)$ since the constraints of both models share the same and their objective is to minimize k. □

Proof of Theorem 1 We begin by discussing the boundary cases separately. When $\tau_{rs} = Z_0$, we first show that the problem $RS(Z_0)$ and $DRO(0)$ have the same optimal solution $\boldsymbol{x}^*$, where $\boldsymbol{x}^* = \arg\min_{\boldsymbol{x}\in\mathcal{X}} \frac{1}{S}\sum_{s\in[S]} f(\boldsymbol{x}, \hat{\boldsymbol{z}}_s)$. The first constraint in the problem $RS(Z_0)$ becomes $\frac{1}{S}\sum_{s\in[S]} \sup_{z\in\mathcal{Z}}\{f(\boldsymbol{x},\boldsymbol{z}) - k\rho(\boldsymbol{z},\hat{\boldsymbol{z}}_s)\} \le Z_0$. There also exists the inequality $\frac{1}{S}\sum_{s\in[S]} \sup_{z\in\mathcal{Z}}\{f(\boldsymbol{x},\boldsymbol{z}) - k\rho(\boldsymbol{z},\hat{\boldsymbol{z}}_s)\} \ge \frac{1}{S}\sum_{s\in[S]} f(\boldsymbol{x},\hat{\boldsymbol{z}}_s) \ge Z_0$ where the equality holds only when $\boldsymbol{x} = \boldsymbol{x}^*$. Thus, $\boldsymbol{x}^*$ is the optimal solution to the problem $RS(Z_0)$. For the problem $DRO(0)$, it reduces to the empirical optimization, and, by definition, its optimal solution is also $\boldsymbol{x}^*$.

When $\tau_{rs} = \overline{\tau}$, we have $k^*(\tau_{rs}) = 0$ since all constraints could hold when $k(\tau_{rs}) = 0$. Simply let $\overline{\theta} = \frac{1}{S}\sum_{s\in[S]} \sup_{z\in\mathcal{Z}}\{\rho(\boldsymbol{z},\hat{\boldsymbol{z}}_s)\}$ which means there is no distance restriction on the distribution, then the problem $DRO(\overline{\theta})$ becomes $\min_{x\in\mathcal{X}} \sup_{z\in\mathcal{Z}} f(\boldsymbol{x},\boldsymbol{z})$ and the problem $RS(\overline{\tau})$ becomes

$$
\begin{aligned}
\min_{x\in\mathcal{X}} \quad & 0 \\
s.t. \quad & \sup_{z\in\mathcal{Z}} f(\boldsymbol{x},\boldsymbol{z}) \le \overline{\tau}
\end{aligned}
$$

Since $\overline{\tau} = \inf_{\boldsymbol{x}\in\mathcal{X}} \sup_{\boldsymbol{z}\in\mathcal{Z}} f(\boldsymbol{x},\boldsymbol{z})$, both problems have the same optimal solutions.

Next, we investigate the case for $\tau_{rs} \in \big(Z_0, \bar{\tau}\big)$. We first define $k^*(\tau)$ as the optimal value of the problem $RS(\tau)$ and $g(\tau,\theta) = \tau + k^*(\tau)\theta$. Then we can rewrite the problem (DRO) as: $\min_{\tau\ge Z_0}\ g(\tau,\theta)$. We provide the following lemma on the convexity of $k^*(\tau)$.

LEMMA 1. *$k^*(\tau)$ is convex if the function $f(\boldsymbol{x},\boldsymbol{z})$ is convex over $\boldsymbol{x}$.*

Proof Suppose that $(\boldsymbol{x}_1, k_1)$ and $(\boldsymbol{x}_2, k_2)$ are the minimizer of the problem $RS(\tau_1)$ and $RS(\tau_2)$ where $\tau_1 \ge Z_0, \tau_2 \ge Z_0$. We claim that $\big(\lambda\boldsymbol{x}_1 + (1-\lambda)\boldsymbol{x}_2, \lambda k_1 + (1-\lambda)k_2\big)$ is a feasible solution of the problem $RS(\lambda\tau_1 + (1-\lambda)\tau_2)$ for any $\lambda\in[0,1]$. To see this, we have

$$
\begin{aligned}
&\frac{1}{S}\sum_{s\in[S]}\sup_{z\in\mathcal{Z}}\{f\big(\lambda\boldsymbol{x}_1 + (1-\lambda)\boldsymbol{x}_2, \boldsymbol{z}\big) - \big(\lambda k_1 + (1-\lambda)k_2\big)\rho(\boldsymbol{z},\hat{\boldsymbol{z}}_s)\} \\
\le&\frac{1}{S}\sum_{s\in[S]}\sup_{z\in\mathcal{Z}}\{\lambda[f(\boldsymbol{x}_1,\boldsymbol{z}) - k_1\rho(\boldsymbol{z},\hat{\boldsymbol{z}}_s)] + (1-\lambda)[f(\boldsymbol{x}_2,\boldsymbol{z}) - k_2\rho(\boldsymbol{z},\hat{\boldsymbol{z}}_s)]\} \\
\le&\lambda\frac{1}{S}\sum_{s\in[S]}\sup_{z\in\mathcal{Z}}\{f(\boldsymbol{x}_1,\boldsymbol{z}) - k_1\rho(\boldsymbol{z},\hat{\boldsymbol{z}}_s)\} + (1-\lambda)\frac{1}{S}\sum_{s\in[S]}\sup_{z\in\mathcal{Z}}\{f(\boldsymbol{x}_2,\boldsymbol{z}) - k_2\rho(\boldsymbol{z},\hat{\boldsymbol{z}}_s)\} \\
\le&\lambda\tau_1 + (1-\lambda)\tau_2
\end{aligned}
$$

where the first inequality holds since the convexity of the function $f(\boldsymbol{x},\boldsymbol{z})$, and the third holds because of the assumption of $(\boldsymbol{x}_1,k_1)$ and $(\boldsymbol{x}_2,k_2)$. Moreover, due to the convexity of set $\mathcal{X}$, we also have $\lambda\boldsymbol{x}_1+(1-\lambda)\boldsymbol{x}_2\in\mathcal{X}$. Because $k^*\big(\lambda\tau_1+(1-\lambda)\tau_2\big)$ is an optimal solution, we have $k^*\big(\lambda\tau_1+(1-\lambda)\tau_2\big)\le\lambda k_1+(1-\lambda)k_2=\lambda k^*(\tau_1)+(1-\lambda)k^*(\tau_2)$. Thus $k^*(\tau)$ is convex. □

When $\tau_{rs}\in\big(Z_0,\bar\tau\big)$, from the above lemma, we know that $k^*(\tau_{rs})$ is finite that there exist feasible solutions of $RS(\tau_{rs})$ whenever $\tau_{rs}\ge Z_0$. In addition, $\tau_{rs}\in\text{int}(\text{dom }k^*)$, so we have that $\partial k^*(\tau_{rs})$ is nonempty and bounded.

Next, we show that $0\notin\partial k^*(\tau_{rs})$ for any $\tau_{rs}\in\big(Z_0,\bar\tau\big)$. If there exists $\hat\tau\in\big(Z_0,\bar\tau\big)$ such that $0\in\partial k^*(\hat\tau)$, then $\hat\tau$ is the minimizer of $k^*(\tau)$ and $k^*(\hat\tau)\le k^*(\bar\tau)$. Because $k^*(\tau)$ is a nonincreasing function, we have $k^*(\hat\tau)\ge k^*(\bar\tau)$. Thus $k^*(\hat\tau)=k^*(\bar\tau)=0$. Since $k^*(\hat\tau)=0$, the first constraint in the problem $RS(\hat\tau)$ indicates that $\inf_{\boldsymbol{x}\in\mathcal{X}}\sup_{\boldsymbol{z}\in\mathcal{Z}}f(\boldsymbol{x},\boldsymbol{z})\le\hat\tau$. So we have $\inf_{\boldsymbol{x}\in\mathcal{X}}\sup_{\boldsymbol{z}\in\mathcal{Z}}f(\boldsymbol{x},\boldsymbol{z})\le\hat\tau<\bar\tau$, which contradicts the assumption.

Hence we can get θ^* by letting $\theta^*=-\frac{1}{\alpha}$, where $\alpha\in\partial k^*(\tau_{rs})$. We claim that $\alpha<0$. Indeed, $k^*(\tau)\ge k^*(\tau_{rs})+\alpha(\tau-\tau_{rs})$ for any $\tau\ge Z_0$ by the definition of subgradient. Because $k^*(\tau)$ is a nonincreasing function, the inequality holds when $\tau>\tau_{rs}$ only if $\alpha\le0$. Added with $\alpha\ne0$, we get $\alpha<0$. Then $\theta^*>0$, which is meaningful. From the *Moreau-Rockafellar Theorem*[Rockafellar (1970),*Theorem 23.8*], it has $\partial_\tau g(\tau_{rs},\theta^*)=1+\theta^*\partial k^*(\tau_{rs})$ since $\tau_{rs}\in\text{int}(\text{dom }k^*)$. Therefore $0=1+\theta^*\alpha\in1+\theta^*\partial k^*(\tau_{rs})=\partial_\tau g(\tau_{rs},\theta^*)$, where $g(\tau,\theta^*)=\tau+k^*(\tau)\theta^*$. Thus, we find the θ^* to make $\min_{\tau\ge Z_0}\ g(\tau,\theta^*)$ reach the minimum at $\tau=\tau_{rs}$. That is, $DRO(\theta^*)$ and $RS(\tau_{rs})$ share the same optimal solution. □

Proof of Proposition 2 For any $\theta\ge\overline\theta$, the problem $DRO(\theta)$ is always equivalent to $\min_{x\in\mathcal{X}}\sup_{z\in\mathcal{Z}}f(\boldsymbol{x},\boldsymbol{z})$.

When $\tau>\overline\tau$, from latter part of the proof for Theorem 1, we already know that the problem $RS(\tau)$ presents

$$\begin{aligned}\min_{x\in\mathcal{X}}\quad&0\\ s.t.\quad&\sup_{z\in\mathcal{Z}}f(\boldsymbol{x},\boldsymbol{z})\le\tau\end{aligned}$$

The optimal solution $\boldsymbol{x}^*$ to the problem $DRO(\theta)$, $\boldsymbol{x}^*\in\arg\inf_{\boldsymbol{x}\in\mathcal{X}}\sup_{\boldsymbol{z}\in\mathcal{Z}}f(\boldsymbol{x},\boldsymbol{z})$, is also optimal to the problem $RS(\tau)$ since the constraint $\sup_{z\in\mathcal{Z}}f(\boldsymbol{x}^*,\boldsymbol{z})=\overline\tau\le\tau$ holds. Thus, derive the conclusion. □

Proof of Corollary 1 Because $f(\boldsymbol{x},\boldsymbol{z})$ is linear, and $\mathcal{Z}$ is a polyhedron, from Theorem 1 we just need to get α^* from the dual problem of $RS(\tau)$.

Let us reformulate $RS(\tau)$ as

$$\begin{aligned}\min\quad&k\\ s.t.\quad&\frac{1}{S}\sum_{s\in[S]}\beta_s\le\tau\\ &f(\boldsymbol{x},\boldsymbol{z})-k\rho(\boldsymbol{z},\hat{\boldsymbol{z}}_s)\le\beta_s\quad\forall z\in\mathcal{Z},s\in[S]\\ &x\in\mathcal{X}\end{aligned}$$

Substituting $f(\boldsymbol{x},\boldsymbol{z})=\boldsymbol{x}^T\boldsymbol{z}$ and $\rho(\boldsymbol{z},\hat{\boldsymbol{z}}_s)=||\boldsymbol{z}-\hat{\boldsymbol{z}}_s||_1$, $\mathcal{Z}=\{\boldsymbol{z}|A\boldsymbol{z}\ge\boldsymbol{b}\},\mathcal{X}=\{\boldsymbol{x}|C\boldsymbol{x}\le\boldsymbol{h}\}$, we get the robust counterpart of the second constraint.

$$\begin{aligned}&(\boldsymbol{x}^T+\boldsymbol{\rho}_{\boldsymbol{s}}^TA)\hat{\boldsymbol{z}}_s-\boldsymbol{\rho}_{\boldsymbol{s}}^T\boldsymbol{b}\le\beta_s\quad\forall s\in[S]\\ &||\boldsymbol{x}+A^T\boldsymbol{\rho}_{\boldsymbol{s}}||_\infty\le k\quad\forall s\in[S]\end{aligned}$$

where $\boldsymbol{\rho}_s, \forall s \in [S]$ are the dual variables for the support set $\mathcal{Z}$.

Then taking the dual variables $\boldsymbol{\alpha}, \boldsymbol{\gamma}, \boldsymbol{\phi}, \boldsymbol{\varphi}$ for the constraints, the dual problem is derived as follows,

$$\begin{aligned}
\max_{\alpha \le 0, (\boldsymbol{\gamma}, \boldsymbol{\phi^1}, \boldsymbol{\phi^2}, \boldsymbol{\varphi}) \ge 0} \quad & \alpha\tau - \boldsymbol{\varphi}^T \boldsymbol{h} \\
s.t. \quad & \sum_{s \in [S]} \mathbf{1}^T (\boldsymbol{\phi}_s^1 + \boldsymbol{\phi}_s^2) \le 1 \\
& \frac{\alpha}{S} + \gamma_s = 0 \quad \forall s \in [S] \\
& \sum_{s \in [S]} (\gamma_s \hat{\boldsymbol{z}}_s + \boldsymbol{\phi}_s^1 - \boldsymbol{\phi}_s^2) + C^T \boldsymbol{\varphi} = 0 \\
& A(\gamma_s \hat{\boldsymbol{z}}_s + \boldsymbol{\phi}_s^1 - \boldsymbol{\phi}_s^2) - \gamma_s \boldsymbol{b} \ge 0 \quad \forall s \in [S]
\end{aligned}$$

Then eliminating variables $\gamma_s, \forall s \in [S]$, the dual problem can be rewritten as

$$\begin{aligned}
\max_{\alpha \le 0, (\boldsymbol{\phi^1}, \boldsymbol{\phi^2}, \boldsymbol{\varphi}) \ge 0} \quad & \alpha\tau - \boldsymbol{\varphi}^T \boldsymbol{h} \\
s.t. \quad & \sum_{s \in [S]} \mathbf{1}^T (\boldsymbol{\phi}_s^1 + \boldsymbol{\phi}_s^2) \le 1 \\
& \sum_{s \in [S]} (\boldsymbol{\phi}_s^1 - \boldsymbol{\phi}_s^2 - \frac{\alpha}{S} \hat{\boldsymbol{z}}_s) + C^T \boldsymbol{\varphi} = 0 \\
& A(\boldsymbol{\phi}_s^1 - \boldsymbol{\phi}_s^2 - \frac{\alpha}{S} \hat{\boldsymbol{z}}_s) + \frac{\alpha}{S} \boldsymbol{b} \ge 0 \quad \forall s \in [S]
\end{aligned}$$

□

Proof of Theorem 2 First, for any θ, if $(\tau^*, k^*, x^*, y_1^*(z,u), \ldots, y_S^*(z,u))$ is the optimal solution to the problem *DRO-LDR*(θ), then $(k^*, x^*, y_1^*(z,u), \ldots, y_S^*(z,u))$ will also be the optimal solution to the problem *RS-LDR*(τ^*), since the constraints in both models are the same. Next, if we fix τ_{rs} for the problem *RS-LDR*(τ), we try to find a corresponding θ^* so that the two models share the same optimal solution. We can represent *DRO-LDR(θ)* as $\min_{\tau \ge \underline{\tau}} \ \tau + k^*(\tau)\theta$ where $k^*(\tau)$ represents the optimal value of the variable k in the problem *RS-LDR*(τ) and $\underline{\tau}$ denoting the smallest τ making the problem feasible.

$k^*(\tau)$ is convex. For any τ_1, τ_2, let $(k_1, x_1, y_1^1(z,u), \ldots, y_S^1(z,u))$ and $(k_2, x_2, y_1^2(z,u), \ldots, y_S^2(z,u))$ be the optimal solution to the problem *RS-LDR(τ_1)* and *RS-LDR(τ_2)* respectively. We claim that $(\lambda k_1 + (1-\lambda)k_2, \lambda x_1 + (1-\lambda)x_2, \lambda y_1^1(z,u) + (1-\lambda)y_1^2(z,u), \ldots, \lambda y_S^1(z,u) + (1-\lambda)y_S^2(z,u))$ where $\lambda \in [0,1]$ is a feasible solution of the problem *RS-LDR($\lambda\tau_1 + (1-\lambda)\tau_2$)*. Firstly, $\lambda x_1 + (1-\lambda)x_2 \in \mathcal{X}$ since $\mathcal{X}$ is convex and $\lambda y_s^1(z,u) + (1-\lambda)y_s^2(z,u) \in \mathcal{L}^{m+1,P}, \forall s \in [S]$ since linear operation didn't change its linearity. Secondly, we have

$$\begin{aligned}
& A(z)(\lambda x_1 + (1-\lambda)x_2) + B(\lambda y_s^1(z,u) + (1-\lambda)y_s^2(z,u)) \\
=& \lambda(A(z)x_1 + By_s^1(z,u)) + (1-\lambda)(A(z)x_2 + By_s^2(z,u)) \\
\ge& \lambda b(z) + (1-\lambda)b(z) = b(z) \quad \forall (\boldsymbol{z}, u) \in \bar{\mathcal{Z}}_s, s \in [S]
\end{aligned}$$

And last, we have

$$\begin{aligned}
& \frac{1}{S} \sum_{s \in [S]} \sup_{(z,u) \in \bar{\mathcal{Z}}_s} \{\boldsymbol{c}(\boldsymbol{z})^T(\lambda x_1 + (1-\lambda)x_2) + \boldsymbol{d}^T(\lambda y_s^1(z,u) + (1-\lambda)y_s^2(z,u)) - (\lambda k_1 + (1-\lambda)k_2)u\} \\
=& \frac{1}{S} \sum_{s \in [S]} \sup_{(z,u) \in \bar{\mathcal{Z}}_s} \{\lambda(\boldsymbol{c}(\boldsymbol{z})^T \boldsymbol{x}_1 + \boldsymbol{d}^T \boldsymbol{y}_s^1(\boldsymbol{z},u) - k_1 u) + (1-\lambda)(\boldsymbol{c}(\boldsymbol{z})^T \boldsymbol{x}_2 + \boldsymbol{d}^T \boldsymbol{y}_s^2(\boldsymbol{z},u) - k_2 u)\} \\
\le& \lambda \frac{1}{S} \sum_{s \in [S]} \sup_{(z,u) \in \bar{\mathcal{Z}}_s} \{\boldsymbol{c}(\boldsymbol{z})^T \boldsymbol{x}_1 + \boldsymbol{d}^T \boldsymbol{y}_s^1(\boldsymbol{z},u) - k_1 u\} + (1-\lambda)\frac{1}{S} \sum_{s \in [S]} \sup_{(z,u) \in \bar{\mathcal{Z}}_s} \{\boldsymbol{c}(\boldsymbol{z})^T \boldsymbol{x}_2 + \boldsymbol{d}^T \boldsymbol{y}_s^2(\boldsymbol{z},u) - k_2 u\} \\
\le& \lambda\tau_1 + (1-\lambda)\tau_2
\end{aligned}$$

Then we have that $k^*(\lambda\tau_1+(1-\lambda)\tau_2) \le \lambda k_1 + (1-\lambda)k_2 = \lambda k^*(\tau_1) + (1-\lambda)k^*(\tau_1)$, i.e. $k^*(\tau)$ is convex.

If $\tau_{rs} = \underline{\tau}$, we claim that *RS-LDR*$(\underline{\tau})$ share the same optimal solution with *DRO-LDR*(0). First, the optimal value of the problem *DRO-LDR*(0) is just $\underline{\tau}$ since *DRO-LDR*(0) takes the same formulation as the definition of $\underline{\tau}$. And we denote its optimal solution as $(\boldsymbol{x}^*, \boldsymbol{y}_1^*(\boldsymbol{z},u), \ldots, \boldsymbol{y}_S^*(\boldsymbol{z},u))$. Next, for the problem *RS-LDR*$(\underline{\tau})$, we have that for every $(\boldsymbol{x}, \boldsymbol{y}_1(\boldsymbol{z},u), \ldots, \boldsymbol{y}_S(\boldsymbol{z},u)) \in \bar{\mathcal{X}}$, the following inequality holds

$$\frac{1}{S}\sum_{s\in[S]} \sup_{(z,u)\in\bar{\mathcal{Z}}_s} \{\boldsymbol{c}^T\boldsymbol{x} + \boldsymbol{d}^T\boldsymbol{y}_s(\boldsymbol{z},u) - ku\} \ge \underline{\tau}$$

where the equality holds if and only if it takes the solution $(\boldsymbol{x}^*, \boldsymbol{y}_1^*(\boldsymbol{z},u), \ldots, \boldsymbol{y}_S^*(\boldsymbol{z},u))$ since the definition of $\underline{\tau}$. Thus $(\boldsymbol{x}^*, \boldsymbol{y}_1^*(\boldsymbol{z},u), \ldots, \boldsymbol{y}_S^*(\boldsymbol{z},u))$ is also the optimal solution to the problem *RS-LDR*$(\underline{\tau})$.

If $\tau_{rs} \in (\underline{\tau}, \overline{\tau})$, we know that $k^*(\tau_{rs})$ must be finite since there exist feasible solutions of *RS-LDR*(τ_{rs}) whenever $\tau_{rs} \ge \underline{\tau}$. In addition, $\tau_{rs} \in \operatorname{int}(\operatorname{dom} k^*)$, so we have that $\partial k^*(\tau_{rs})$ is nonempty and bounded from convex analysis.

Next we show that $0 \notin \partial k^*(\tau_{rs})$ if $\tau_{rs} \in (\underline{\tau}, \overline{\tau})$. If not, there exists $\hat{\tau} \in (\underline{\tau}, \overline{\tau})$ such that $0 \in \partial k^*(\hat{\tau})$. Then $\hat{\tau}$ is the minimizer of $k^*(\tau)$ and $k^*(\tau) = 0$ for every $\tau \ge \hat{\tau}$ since $k^*(\tau)$ is a nonincreasing function and $k^*(\overline{\tau}) = 0$. Then we have

$$\inf_{(x,y_1(z,u),\ldots,y_S(z,u))\in\bar{\mathcal{X}}} \frac{1}{S}\sum_{s\in[S]} \sup_{(z,u)\in\bar{\mathcal{Z}}_s} \{\boldsymbol{c}^T\boldsymbol{x} + \boldsymbol{d}^T\boldsymbol{y}_s(\boldsymbol{z},u)\} \le \hat{\tau} < \overline{\tau}$$

which is contradict to the assumption.

Hence when $\tau_{rs} \in (\underline{\tau}, \overline{\tau})$, we can get θ^* by letting $\theta^* = -\frac{1}{\alpha}$, where $\alpha \in \partial k^*(\tau_{rs})$. Therefore it has $0 = 1 + \theta^*\alpha \in 1 + \theta^*\partial k^*(\tau_{rs}) = \partial(\tau + k^*(\tau_{rs})\theta^*)$. Thus, we find the θ^* which makes the problem $\min_{\tau\ge Z_0}\ \tau + k^*(\tau)\theta^*$ take the optimal solution with $\tau = \tau_{rs}$.

If $\tau_{rs} = \overline{\tau}$, we have $k^*(\tau_{rs}) = 0$ since the constraints hold when $k(\tau_{rs}) = 0$. So the problem *RS*$(\overline{\tau})$ becomes

$$\begin{aligned} \min_{(x,y_1(z,u),\ldots,y_S(z,u))\in\bar{\mathcal{X}}} \quad & 0 \\ s.t. \quad & \frac{1}{S}\sum_{s\in[S]} \sup_{(z,u)\in\bar{\mathcal{Z}}_s} \{\boldsymbol{c}^T\boldsymbol{x} + \boldsymbol{d}^T\boldsymbol{y}_s(\boldsymbol{z},u)\} \le \overline{\tau} \end{aligned}$$

While the problem *DRO-LDR*$(\overline{\theta})$ becomes

$$\begin{aligned} \min \quad & \frac{1}{S}\sum_{s\in[S]} \sup_{(z,u)\in\bar{\mathcal{Z}}_s} \{\boldsymbol{c}^T\boldsymbol{x} + \boldsymbol{d}^T\boldsymbol{y}_s(\boldsymbol{z},u)\} \\ s.t. \quad & (x, y_1(z,u), \ldots, y_S(z,u)) \in \bar{\mathcal{X}} \end{aligned}$$

Both problems have the same optimal solutions, i.e.,

$$(\boldsymbol{x}^*, \boldsymbol{y_1(z,u)}, \ldots, \boldsymbol{y_S(z,u)}) \in \arg\inf_{(x,y_1(z,u),\ldots,y_S(z,u))\in\bar{\mathcal{X}}} \frac{1}{S}\sum_{s\in[S]} \sup_{(z,u)\in\bar{\mathcal{Z}}_s} \{\boldsymbol{c}^T\boldsymbol{x} + \boldsymbol{d}^T\boldsymbol{y}_s(\boldsymbol{z},u)\}.$$

□

Proof of Theorem 3 Suppose that the optimal solution of the GRS problem $GRS(\tau', \theta')$ is denoted as $(\boldsymbol{x}^*, k^*, \eta^*)$. We claim that there exists a solution where $\eta^* = k^*$. To prove this, we consider two cases:

1. If $\eta^* < k^*$, selecting a smaller value of k equal to η^* would not change the left-hand side value of the first constraint, so it still holds. However, the objective value would become smaller, contradicting the assumption of optimality.

2. If $\eta^* > k^*$, reducing η^* to k^* would decrease the left-hand side value of the first constraint. Despite this reduction, the constraint would still hold, and the optimal value would remain the same.

Therefore, we can replace η with k and reformulate $GRS(\tau', \theta')$ into the following problem:

$$\begin{aligned} \min \quad & k \\ s.t. \quad & \frac{1}{S}\sum_{s\in[S]}\sup_{z\in\mathcal{Z}}\{f(\boldsymbol{x},\boldsymbol{z}) - k\rho(\boldsymbol{z},\hat{\boldsymbol{z}}_s)\} \le \tau' - k\theta', \\ & x \in \mathcal{X}, \\ & k \ge 0. \end{aligned}$$

Suppose that its optimal solution is denoted as $(\boldsymbol{x}^*, k^*)$. We claim that the optimal value of $RS(\tau - k^*\theta')$ will be k^*. We can easily show that $(\boldsymbol{x}^*, k^*)$ is a feasible solution of $RS(\tau' - k^*\theta')$.

Now, let's assume that the optimal solution of $RS(\tau' - k^*\theta')$ is denoted as $(\bar{\boldsymbol{x}}, \bar{k})$. Then $\bar{k}$ is not smaller than k^*. Suppose the contrary, that $\bar{k} < k^*$, then we have $\frac{1}{S}\sum_{s\in[S]}\sup_{z\in\mathcal{Z}}\{f(\bar{\boldsymbol{x}},\boldsymbol{z}) - \bar{k}\rho(\boldsymbol{z},\hat{\boldsymbol{z}}_s)\} \le \tau' - k^*\theta'$. This inequality implies that $\frac{1}{S}\sum_{s\in[S]}\sup_{z\in\mathcal{Z}}\{f(\bar{\boldsymbol{x}},\boldsymbol{z}) - \bar{k}\rho(\boldsymbol{z},\hat{\boldsymbol{z}}_s)\} \le \tau' - \bar{k}\theta'$. Therefore, $\bar{k}$ would be a feasible solution of $GRS(\tau', \theta')$, which contradicts the assumption that k^* is the optimal solution of the GRS problem.

Hence, $(\boldsymbol{x}^*, k^*)$ is also the optimal solution of $RS(\tau' - k^*\theta')$, indicating that it has the same optimal solution as $GRS(\tau', \theta')$.

The other direction is trivial, since when $\theta' = 0$, problem $GRS(\tau, 0)$ reduces to $RS(\tau)$. □

Proof of Theorem 4 We first rewrite the two problems. Problem $DRO(\theta)$ can be rewritten as $\min_{\boldsymbol{x}\in\mathcal{X},k\ge 0}\frac{1}{S}\sum_{s\in[S]}\sup_{z\in\mathcal{Z}}\{f(\boldsymbol{x},\boldsymbol{z}) - k\rho(\boldsymbol{z},\hat{\boldsymbol{z}}_s)\} + k\theta$. Problem $GDRC(k', \theta')$ can be reformulated as $\min_{\boldsymbol{x}\in\mathcal{X},0\le k\le k'}\frac{1}{S}\sum_{s\in[S]}\sup_{z\in\mathcal{Z}}\{f(\boldsymbol{x},\boldsymbol{z}) - k\rho(\boldsymbol{z},\hat{\boldsymbol{z}}_s)\} + k\theta$. The two formulations are almost identical except for the range on k in GDRC. Define a function $h(\boldsymbol{x},k) := \frac{1}{S}\sum_{s\in[S]}\sup_{z\in\mathcal{Z}}\{f(\boldsymbol{x},\boldsymbol{z}) - k\rho(\boldsymbol{z},\hat{\boldsymbol{z}}_s)\}$, then Problem $DRO(\theta)$ is $\min_{\boldsymbol{x}\in\mathcal{X},k\ge 0} h(\boldsymbol{x},k) + k\theta$ and Problem $GDRC(k', \theta')$ is $\min_{\boldsymbol{x}\in\mathcal{X},0\le k\le k'} h(\boldsymbol{x},k) + k\theta$. Note that the function $h(\boldsymbol{x},k)$ is convex since it is the point-wise maximization on $f(\boldsymbol{x},\boldsymbol{z})$.

We start from the easy direction. For a fixed θ, suppose $(\boldsymbol{x}^*, k^*)$ is an optimal solution to Problem $DRO(\theta)$. Let $k' = k^*, \theta' = \theta$, then $(\boldsymbol{x}^*, k^*)$ is also a optimal solution to Problem $GDRC(k^*, \theta)$ since the constraint $k \le k^*$ allows k takes value k^*.

Let us prove the other direction. For a fixed (k', θ'), suppose $(\boldsymbol{x}^*, k^*)$ is an optimal solution to Problem $GDRC(k', \theta')$. If $k^* < k'$, then let $\theta = \theta'$ and we have that $(\boldsymbol{x}^*, k^*)$ is also an optimal solution to Problem $DRO(\theta')$. If not, we could find a smaller solution to Problem $GDRC(k', \theta')$ since it is convex, which is contradictory to the assumption. If $k^* = k'$, from the first order optimality condition, we have $0 \in \partial_{\boldsymbol{x}} h(\boldsymbol{x}^*, k') + \mathcal{N}_{\mathcal{X}}(\boldsymbol{x}^*)$, where $\mathcal{N}_{\mathcal{X}}(\boldsymbol{x}^*)$ is the normal cone to $\mathcal{X}$ at the point $\boldsymbol{x}^*$. Then take some $\theta \in -\partial_k h(\boldsymbol{x}^*, k')$, the solution $(\boldsymbol{x}^*, k')$ will be optimal to Problem $DRO(\theta)$. Since the Problem $DRO(\theta)$ is convex, we just need to prove that $0 \in \partial\{h(\boldsymbol{x}^*, k') + k'\theta\} + \big(\mathcal{N}_{\mathcal{X}}(\boldsymbol{x}^*), 0\big)$. Since the function $f(\boldsymbol{x},\boldsymbol{z})$ is convex on $\boldsymbol{x}$ and continuous on $\boldsymbol{z}$, and the metric ρ is continuous on $\boldsymbol{z}$, according to [Shapiro et al. (2021), *Theorem 7.26*], we have

$$\begin{aligned} \partial\{h(\boldsymbol{x},k) + k\theta\} =& \frac{1}{S}\sum_{s\in[S]} \operatorname{conv}\left\{\cup_{z\in\bar{\mathcal{Z}}_s(\boldsymbol{x},k)}\{\partial\{f(\boldsymbol{x},\boldsymbol{z}) - k\rho(\boldsymbol{z},\hat{\boldsymbol{z}}_s)\}\}\right\} + \partial(k\theta) \\ =& \frac{1}{S}\sum_{s\in[S]} \operatorname{conv}\left\{\cup_{z\in\bar{\mathcal{Z}}_s(\boldsymbol{x},k)}\{\big(\partial_{\boldsymbol{x}} f(\boldsymbol{x},\boldsymbol{z}), -\partial_k k\rho(\boldsymbol{z},\hat{\boldsymbol{z}}_s)\big)\}\right\} + (\boldsymbol{0}, \theta) \end{aligned}$$

where $\bar{\mathcal{Z}}_s(\boldsymbol{x},k) := \arg\max_{z\in\mathcal{Z}}\{f(\boldsymbol{x},\boldsymbol{z}) - k\rho(\boldsymbol{z},\hat{\boldsymbol{z}}_s)\}$ is the optimal solution set on $\boldsymbol{z}$. Similarly, we have $\partial_{\boldsymbol{x}} h(\boldsymbol{x},k) = \frac{1}{S}\sum_{s\in[S]}\mathrm{conv}\left\{\cup_{z\in\bar{\mathcal{Z}}_s(\boldsymbol{x},k)}\{\partial_{\boldsymbol{x}} f(\boldsymbol{x},\boldsymbol{z})\}\right\}$ and $\partial_k h(\boldsymbol{x},k) = \frac{1}{S}\sum_{s\in[S]}\mathrm{conv}\left\{\cup_{z\in\bar{\mathcal{Z}}_s(\boldsymbol{x},k)}\{-\partial_k k\rho(\boldsymbol{z}\}\right\}$. Thus, we derive that

$$\partial\{h(\boldsymbol{x}^*,k') + k'\theta\} = \big(\partial_{\boldsymbol{x}} h(\boldsymbol{x}^*,k'), \partial_k h(\boldsymbol{x}^*,k')\big) + \big(\boldsymbol{0},\theta\big) = \big(\partial_{\boldsymbol{x}} h(\boldsymbol{x}^*,k'), \partial_k h(\boldsymbol{x}^*,k') + \theta\big),$$

which indicates that $0 \in \partial\{h(\boldsymbol{x}^*,k') + k'\theta\} + \big(\mathcal{N}_{\mathcal{X}}(\boldsymbol{x}^*),0\big)$. □

Proof of Theorem 5 We start by proving one direction. For a fixed $\bar{k}$ and θ', suppose an optimal solution to the problem *GDRC*$(\bar{k},\theta')$ is $(\bar{\boldsymbol{x}},\bar{\tau},\bar{\eta})$. We claim that the optimal solution of problem *GRS*$(\bar{\tau},\theta')$ is also optimal to problem *GDRC*$(\bar{k},\theta')$.

First, $(\bar{\boldsymbol{x}},\bar{k},\bar{\eta})$ is a feasible solution to the problem *GRS*$(\bar{\tau},\theta')$. Suppose an optimal solution to problem *GRS*$(\bar{\tau},\theta')$ is $(\boldsymbol{x}^*,k^*,\eta^*)$. Then we have $k^* \le \bar{k}$.

If $k^* = \bar{k}$, then $(\boldsymbol{x}^*,\bar{\tau},\eta^*)$ is a feasible solution and thus an optimal solution to problem *GDRC*$(\bar{k},\theta')$.

If $k^* < \bar{k}$. Suppose the optimal value of the problem *GDRC*(k^*,θ') is τ_c^*, then $\tau_c^* \ge \bar{\tau}$ since the optimal value of *GDRC*(k,θ') is non-increasing in k. If $\tau_c^* = \bar{\tau}$, then $(\boldsymbol{x}^*,\tau_c^*,\eta^*)$ is an optimal solution to the problem *GDRC*(k^*,θ') since the constraint holds, and the optimal value is obtained. At the same time, $(\boldsymbol{x}^*,\tau_c^*,\eta^*)$ is also an optimal solution to the problem *GDRC*$(\bar{k},\theta')$, since the following inequality holds:

$$\begin{aligned}&\theta'\eta^* + \frac{1}{S}\sum_{s\in[S]}\sup_{z\in\mathcal{Z}}\{f(\boldsymbol{x}^*,\boldsymbol{z}) - \min\{\bar{k},\eta^*\}\rho(\boldsymbol{z},\hat{\boldsymbol{z}}_s)\}\\ \le&\theta'\eta^* + \frac{1}{S}\sum_{s\in[S]}\sup_{z\in\mathcal{Z}}\{f(\boldsymbol{x}^*,\boldsymbol{z}) - \min\{k^*,\eta^*\}\rho(\boldsymbol{z},\hat{\boldsymbol{z}}_s)\} \le \bar{\tau}\end{aligned}$$

Thus, both problems *GDRC*$(\bar{k},\theta')$ and *GRS*$(\bar{\tau},\theta')$ have an optimal solution $\boldsymbol{x}^*$. If problem *GDRC*$(\bar{k},\theta')$ has an unique optimal solution, we have $\boldsymbol{x}^* = \bar{\boldsymbol{x}}$.

If $\tau_c^* > \bar{\tau}$, then $(\boldsymbol{x}^*,\bar{\tau},\eta^*)$ is not a feasible solution to problem *GDRC*(k^*,θ'), and we have $\theta'\eta^* + \frac{1}{S}\sum_{s\in[S]}\sup_{z\in\mathcal{Z}}\{f(\boldsymbol{x}^*,\boldsymbol{z}) - \min\{k^*,\eta^*\}\rho(\boldsymbol{z},\hat{\boldsymbol{z}}_s)\} > \bar{\tau}$, which contradicts the assumption that $(\boldsymbol{x}^*,k^*,\eta^*)$ is an optimal solution to the problem *GRS*$(\bar{\tau},\theta')$.

Next, let's prove the other direction. For a fixed $\hat{\tau}$ and θ', suppose the problem *GRS*$(\hat{\tau},\theta')$ has an optimal solution $(\hat{\boldsymbol{x}},\hat{k},\hat{\eta})$. We claim that the optimal solution to problem *GDRC*$(\hat{k},\theta')$ is also optimal to problem *GRS*$(\hat{\tau},\theta')$.

First, it is easy to show that $(\hat{\boldsymbol{x}},\hat{\tau},\hat{\eta})$ is a feasible solution to the problem *GDRC*$(\hat{k},\theta')$. Suppose an optimal solution to problem *GDRC*$(\hat{k},\theta')$ is $(\boldsymbol{x}^*,\tau^*,\eta^*)$, then we have $\tau^* \le \hat{\tau}$.

If $\tau^* = \hat{\tau}$, then $(\boldsymbol{x}^*,\hat{k},\eta^*)$ is a feasible solution and thus an optimal solution to problem *GRS*$(\hat{\tau},\theta')$.

Suppose $\tau^* < \hat{\tau}$ and the optimal value of the problem *GRS*(τ^*,θ') is k_s^*, then $k_s^* \ge \hat{k}$ since the optimal value of *GRS*(τ,θ') is non-increasing in τ. If $k_s^* = \hat{k}$, $(\boldsymbol{x}^*,\hat{k},\eta^*)$ is an optimal solution to the problem *GRS*(τ^*,θ'). Since $(\boldsymbol{x}^*,\tau^*,\eta^*)$ is an optimal solution to the problem *GDRC*$(\hat{k},\theta')$, we have the following inequality: $\theta'\eta^* + \frac{1}{S}\sum_{s\in[S]}\sup_{z\in\mathcal{Z}}\{f(\boldsymbol{x}^*,\boldsymbol{z}) - \min\{\hat{k},\eta^*\}\rho(\boldsymbol{z},\hat{\boldsymbol{z}}_s)\} \le \tau^*$, which means that $(\boldsymbol{x}^*,\hat{k},\eta^*)$ is a feasible and optimal solution of the problem *GRS*(τ^*,θ'). Similarly, we have $(\boldsymbol{x}^*,\hat{k},\eta^*)$ is a feasible and optimal solution to the problem *GRS*$(\hat{\tau},\theta')$. Thus, both problems *GRS*$(\hat{\tau},\theta')$ and *GDRC*$(\hat{k},\theta')$ have the same optimal solution $\boldsymbol{x}^*$. If problem *GRS*$(\hat{\tau},\theta')$ has an unique optimal solution, we have $\boldsymbol{x}^* = \hat{\boldsymbol{x}}$.

If $k_s^* > \hat{k}$, then $(\boldsymbol{x}^*, \hat{k}, \eta^*)$ is not a feasible solution to problem $GRS(\tau^*, \theta')$, and we have $\theta'\eta^* + \frac{1}{S}\sum_{s\in[S]}\sup_{z\in\mathcal{Z}}\{f(\boldsymbol{x}^*, \boldsymbol{z}) - \min\{\hat{k}, \eta^*\}\rho(\boldsymbol{z}, \hat{\boldsymbol{z}}_s)\} > \tau^*$. Meanwhile, from the assumption that $(\boldsymbol{x}^*, \tau^*, \eta^*)$ is an optimal solution of the problem $GDRC(\hat{k}, \theta')$, we have $\theta'\eta^* + \frac{1}{S}\sum_{s\in[S]}\sup_{z\in\mathcal{Z}}\{f(\boldsymbol{x}^*, \boldsymbol{z}) - \min\{\hat{k}, \eta^*\}\rho(\boldsymbol{z}, \hat{\boldsymbol{z}}_s)\} \le \tau^*$, which leads to a contradiction. □

Proof of Proposition 3 Because of the nonincreasing and convexity of $k^*(\tau)$ and $0 \in \partial k^*(\tau_{rs})$, we know that $k^*(\tau) = k^*(\tau_{rs}), \forall \tau \ge \tau_{rs}$. Thus, for any $\tau \ge \tau_{rs}$, the problem $RS(\tau)$ has the same optimal solution. We just need to focus on the lowest τ_{rs} such that $0 \in \partial k^*(\tau_{rs})$. For the problem $DRO(\theta) = \min_{\tau \ge Z_0} k^*(\tau)\theta + \tau$, we try to prove that when $\theta \to \infty$, the optimal solution is approaching τ_{rs}. The optimal τ of the problem $DRO(\theta)$ must not larger than τ_{rs} since $k^*(\tau) + \tau = k^*(\tau_{rs}) + \tau > k^*(\tau_{rs}) + \tau_{rs}$ if $\tau > \tau_{rs}$. If the optimal $\tau \in [Z_0, \tau_{rs})$, just let $\theta > \frac{\tau_{rs} - \tau}{k^*(\tau) - k^*(\tau_{rs})}$. We have that $k^*(\tau)\theta + \tau > k^*(\tau_{rs})\theta + \tau_{rs}$ which is contradictory to the optimality of τ. So when θ is large enough, the optimal solution will approach τ_{rs}. □

B. Generalized p-Wasserstein distance

We consider a generalized p-Wasserstein distance for $p \ge 1$ as follows:

$$\begin{aligned} d_W(\mathbb{P}, \mathbb{P}^\dagger) = \inf_{(\tilde{\boldsymbol{z}}, \tilde{\boldsymbol{z}}^\dagger)\sim\mathbb{Q}} & \left(\mathbb{E}_{\mathbb{Q}}[\rho(\tilde{\boldsymbol{z}}, \tilde{\boldsymbol{z}}^\dagger)^p]\right)^{1/p} \\ s.t. \quad & \prod\nolimits_{\tilde{\boldsymbol{z}}} \mathbb{Q} = \mathbb{P} \\ & \prod\nolimits_{\tilde{\boldsymbol{z}}^\dagger} \mathbb{Q} = \mathbb{P}^\dagger. \end{aligned}$$

Thus, the ambiguity set for the DRO model presents the following equivalent formulation:

$$\mathcal{F}(\theta) := \left\{ \mathbb{P} \in \mathcal{P}(\mathbb{R}^m) \;\middle|\; \begin{array}{l} \tilde{\boldsymbol{z}} \sim \mathbb{P} \\ \mathbb{P}(\tilde{\boldsymbol{z}} \in \mathcal{Z}) = 1 \\ d_W(\mathbb{P}, \mathbb{P}^\dagger)^p \le \theta^p \end{array} \right\}.$$

Using a standard duality argument for DRO model, we obtain:

$$\min_{\boldsymbol{x}\in\mathcal{X}} \sup_{\mathbb{P}\in\mathcal{P}(\mathcal{Z})} \inf_{k\ge 0} \mathbb{E}_{\mathbb{P}}\left[f(\boldsymbol{x}, \tilde{\boldsymbol{z}})\right] - k d_W(\mathbb{P}, \mathbb{P}^\dagger)^p + k\theta^p.$$

If $f(x,z)$ is a measurable function in a Polish space $(\mathcal{Z}, \rho)$, the DRO model can be reformulated as the following problem by Gao and Kleywegt (2023):

$$\begin{aligned} \min \quad & \tau + k\theta^p \\ s.t. \quad & \frac{1}{S}\sum_{s\in[S]} \sup_{z\in\mathcal{Z}}\{f(\boldsymbol{x}, \boldsymbol{z}) - k\rho(\boldsymbol{z}, \hat{\boldsymbol{z}}_s)^p\} \le \tau \\ & x \in \mathcal{X},\; k \ge 0. \end{aligned} \tag{DRO-p}$$

The corresponding RS model with generalized p-Wasserstein distance is as follows:

$$\begin{aligned} \min \quad & k \\ s.t. \quad & \mathbb{E}_{\mathbb{P}}\left[f(\boldsymbol{x}, \tilde{\boldsymbol{z}})\right] - \tau \le k d_W(\mathbb{P}, \mathbb{P}^\dagger)^p \quad \forall \mathbb{P} \in \mathcal{P}(\mathcal{Z}) \\ & x \in \mathcal{X},\; k \ge 0. \end{aligned}$$

We can reformulate it as follows:

$$\begin{aligned} \kappa(\tau) := \min \quad & k \\ s.t. \quad & \frac{1}{S}\sum_{s\in[S]} \sup_{z\in\mathcal{Z}}\{f(\boldsymbol{x}, \boldsymbol{z}) - k\rho(\boldsymbol{z}, \hat{\boldsymbol{z}}_s)^p\} \le \tau \\ & x \in \mathcal{X},\; k \ge 0. \end{aligned} \tag{RS-p}$$

Thus, following the same proof procedure, Proposition 1 and Theorem 1 continue to hold for Problem (DRO-p) and Problem (RS-p) with parameters θ^p and τ respectively.

C. Example of a non-convex problem

Example 3 (Non-convex:A combinatorial optimization). Following the setup and results in Long et al. (2023), we give an example of the robust combinatorial model presenting in its equivalent reformulation:

$$\begin{aligned}
\min\ & -x_1 - x_2 + x_1(5.5-k)^+ + x_2(7-k)^+ + x_3(3-k)^+ + k\theta \\
s.t.\ & x_1 + x_2 + x_3 \geq 1 \\
& k \in \{0, 3, 5.5, 7\} \\
& x_1, x_2, x_3 \in \{0, 1\}
\end{aligned}$$

The corresponding combinatorial robust satisficing model would be given by:

$$\begin{aligned}
\min\ & k \\
s.t.\ & x_1 + x_2 + x_3 \geq 1 \\
& -x_1 - x_2 + x_1(5.5-k)^+ + x_2(7-k)^+ + x_3(3-k)^+ \leq \tau \\
& k \geq 0 \\
& x_1, x_2, x_3 \in \{0, 1\}
\end{aligned}$$

We will show that the two models do not share the same family of optimal solutions, due to the non-convexity of $\mathcal{X}$. In this case, the RS model has a broader solution family such that some optimal solution only obtained in RS model will not appear in the DRO model. It is easy to know that when $\theta \in [0, 0.5]$, the optimal solutions of the robust combinatorial model are $x^* = (1,1,1)$ and $x^* = (1,1,0)$. When $\theta > 0.5$, the optimal solution will be $x^* = (0,0,1)$. There is no other optimal solution in the solution family of the robust combinatorial model. However, the combinatorial robust satisficing model owns an additional minimizer $x^* = (1,0,1)$ when $\tau \in (-1, 0)$.

D. Explicit Formulation in Lot-sizing

After the lifted affine recourse adaptation technique, the corresponding formulation of *DRO-LDR*(θ) is as follows:

$$\begin{aligned}
\min\quad & \boldsymbol{c}^T \boldsymbol{x} + k\theta + \frac{1}{S}\sum_{s\in[S]} \beta_s \\
s.t.\quad & \sum_{i\in[N]} \boldsymbol{d}_i^T \boldsymbol{y}_i^{(s)}(\boldsymbol{z}, u) + \boldsymbol{l}^T \boldsymbol{w}^{(s)}(\boldsymbol{z}, u) - ku \leq \beta_s \quad \forall (\boldsymbol{z}, u) \in \bar{\mathcal{Z}}_s, s \in [S] \\
& x_i + w_i^{(s)}(\boldsymbol{z}, u) + \sum_{j\in[N]} y_{ji}^{(s)}(\boldsymbol{z}, u) - \sum_{j\in[N]} y_{ij}^{(s)}(\boldsymbol{z}, u) - z_i \geq 0, \quad \forall (\boldsymbol{z}, u) \in \bar{\mathcal{Z}}_s, s \in [S], i \in [N] \\
& \boldsymbol{y}^{(s)}(\boldsymbol{z}, u) \geq \boldsymbol{0}, \boldsymbol{w}^{(s)}(\boldsymbol{z}, u) \geq \boldsymbol{0} \quad \forall (\boldsymbol{z}, u) \in \bar{\mathcal{Z}}_s, s \in [S] \\
& \boldsymbol{0} \leq \boldsymbol{x} \leq \bar{\boldsymbol{\delta}} \\
& \boldsymbol{y}^{(s)} \in \mathcal{L}^{(N+1, N\times N)}, \boldsymbol{w}^{(s)} \in \mathcal{L}^{(N+1, N)} \quad \forall s \in [S],
\end{aligned}$$

where the lifted uncertainty support set under each empirical scenario $s \in [S]$ is defined as $\bar{\mathcal{Z}}_s := \{(\boldsymbol{z}, u) \in \mathcal{Z} \times R \,|\, ||\boldsymbol{z} - \hat{\boldsymbol{z}}_s|| \leq u\}$.

The corresponding formulation of *RS-LDR*(τ) is presented as follows,

$$\begin{aligned}
\min \quad & k \\
s.t. \quad & \boldsymbol{c}^T\boldsymbol{x} + \frac{1}{S}\sum_{s\in[S]}\beta_s - \tau \le 0 \\
& \sum_{i\in[N]} \boldsymbol{d}_i^T \boldsymbol{y}_i^{(s)}(\boldsymbol{z},u) + \boldsymbol{l}^T\boldsymbol{w}^{(s)}(\boldsymbol{z},u) - ku \le \beta_s \quad \forall(\boldsymbol{z},u)\in\bar{\mathcal{Z}}_s, s\in[S] \\
& x_i + w_i^{(s)}(\boldsymbol{z},u) + \sum_{j\in[N]} y_{ji}^{(s)}(\boldsymbol{z},u) - \sum_{j\in[N]} y_{ij}^{(s)}(\boldsymbol{z},u) - z_i \ge 0, \quad \forall(\boldsymbol{z},u)\in\bar{\mathcal{Z}}_s, s\in[S], i\in[N] \\
& \boldsymbol{y}^{(s)}(\boldsymbol{z},u)\ge\boldsymbol{0}, \boldsymbol{w}^{(s)}(\boldsymbol{z},u)\ge\boldsymbol{0} \quad \forall(\boldsymbol{z},u)\in\bar{\mathcal{Z}}_s, s\in[S] \\
& \boldsymbol{0}\le\boldsymbol{x}\le\bar{\boldsymbol{\delta}} \\
& \boldsymbol{y}^{(s)}\in\mathcal{L}^{(N+1,N\times N)}, \boldsymbol{w}^{(s)}\in\mathcal{L}^{(N+1,N)} \quad \forall s\in[S].
\end{aligned}$$

E. Reformulations and Specific Relationship in Portfolio

The DRO formulation is equivalent to:

$$\begin{aligned}
\inf \quad & k\theta + \frac{1}{S}\sum_{s\in[S]}\beta_s \\
s.t. \quad & \boldsymbol{C}\boldsymbol{x}\le\boldsymbol{h} \\
& \gamma t - \boldsymbol{x}^T\hat{\boldsymbol{z}}_s + \boldsymbol{\rho}_{s1}^T(\boldsymbol{A}\hat{\boldsymbol{z}}_s - \boldsymbol{b}) \le \beta_s \quad \forall s\in[S] \\
& (1-\frac{1}{\epsilon})\gamma t - (1+\frac{\gamma}{\epsilon})\boldsymbol{x}^T\hat{\boldsymbol{z}}_s + \rho_{s2}^T(\boldsymbol{A}\hat{\boldsymbol{z}}_s - \boldsymbol{b}) \le \beta_s \quad \forall s\in[S] \\
& ||\boldsymbol{A}^T\boldsymbol{\rho}_{s1} - \boldsymbol{x}||_2 \le k \quad \forall s\in[S] \\
& ||\boldsymbol{A}^T\boldsymbol{\rho}_{s2} - (1+\frac{\gamma}{\epsilon})\boldsymbol{x}||_2 \le k \quad \forall s\in[S].
\end{aligned}$$

The RS formulation is equivalent to:

$$\begin{aligned}
\inf \quad & k \\
s.t. \quad & \boldsymbol{C}\boldsymbol{x}\le\boldsymbol{h} \\
& \frac{1}{S}\sum_{s\in[S]}\beta_s \le \tau \\
& \gamma t - \boldsymbol{x}^T\hat{\boldsymbol{z}}_s + \boldsymbol{\rho}_{s1}^T(\boldsymbol{A}\hat{\boldsymbol{z}}_s - \boldsymbol{b}) \le \beta_s \quad \forall s\in[S] \\
& (1-\frac{1}{\epsilon})\gamma t - (1+\frac{\gamma}{\epsilon})\boldsymbol{x}^T\hat{\boldsymbol{z}}_s + \rho_{s2}^T(\boldsymbol{A}\hat{\boldsymbol{z}}_s - \boldsymbol{b}) \le \beta_s \quad \forall s\in[S] \\
& ||\boldsymbol{A}^T\boldsymbol{\rho}_{s1} - \boldsymbol{x}||_2 \le k \quad \forall s\in[S] \\
& ||\boldsymbol{A}^T\boldsymbol{\rho}_{s2} - (1+\frac{\gamma}{\epsilon})\boldsymbol{x}||_2 \le k \quad \forall s\in[S].
\end{aligned}$$

Introducing the dual variables $\alpha, \boldsymbol{\varphi}, \lambda_s^1, \lambda_s^2, \boldsymbol{\phi}_s^1, \phi_s^2, \boldsymbol{\phi}_s^3, \phi_s^4, \forall s \in [S]$, we get the dual problem of the RS as follows:

$$\begin{aligned}
\max_{\alpha \leq 0} \quad & -\boldsymbol{\varphi}^T h + \alpha\tau \\
s.t. \quad & \sum_{s\in[S]} (\phi_s^2 + \phi_s^4) \leq 1 \\
& \lambda_s^1 + \lambda_s^2 + \frac{\alpha}{S} = 0 \quad \forall s \in [S] \\
& \sum_{s\in[S]} [\lambda_s^1 + (1 - \frac{1}{\epsilon})\lambda_s^2] = 0 \\
& \boldsymbol{C}^T \boldsymbol{\varphi} + \sum_{s\in[S]} [-\lambda_s^1 \hat{\boldsymbol{z}}_s - (1 + \frac{\gamma}{\epsilon})\lambda_s^2 \hat{\boldsymbol{z}}_s + \boldsymbol{\phi}_s^1 + (1 + \frac{\gamma}{\epsilon})\boldsymbol{\phi}_s^3] \geq 0 \\
& \lambda_s^1(\boldsymbol{A}\hat{\boldsymbol{z}}_s - \boldsymbol{b}) - \boldsymbol{A}\boldsymbol{\phi}_s^1 \geq 0 \quad \forall s \in [S] \\
& \lambda_s^2(\boldsymbol{A}\hat{\boldsymbol{z}}_s - \boldsymbol{b}) - \boldsymbol{A}\boldsymbol{\phi}_s^3 \geq 0 \quad \forall s \in [S] \\
& ||\boldsymbol{\phi}_s^1||_2 \leq \phi_s^2 \quad \forall s \in [S] \\
& ||\boldsymbol{\phi}_s^3||_2 \leq \phi_s^4 \quad \forall s \in [S].
\end{aligned}$$

Therefore, based on Theorem 1, we can derive the corresponding $\theta^* = -\frac{1}{\alpha^*}$, where α^* is the optimal solution to the dual problem of the RS.