# The Terminator: An Integration of Inner and Outer Approximations for Solving Regular and Distributionally Robust Chance Constrained Programs via Variable Fixing

Nan Jiang
H. Milton Stewart School of Industrial & Systems Engineering, Georgia Institute of Technology, Atlanta, GA 30332, nanjiang@gatecht.edu

Weijun Xie
H. Milton Stewart School of Industrial & Systems Engineering, Georgia Institute of Technology, Atlanta, GA 30332, wxie@gatecht.edu

We present a novel approach aimed at enhancing the efficacy of solving both regular and distributionally robust chance constrained programs using an empirical reference distribution. In general, these programs can be reformulated as mixed-integer programs (MIPs) by introducing binary variables for each scenario, indicating whether a scenario should be satisfied. While existing methods have predominantly focused on either inner or outer approximations, this paper bridges this gap by studying a scheme that effectively combines these approximations via variable fixing. Through probing the restricted outer approximations and comparing them with the inner approximations, we derive optimality cuts that can notably reduce the number of binary variables by effectively setting them to either one or zero. We conduct a theoretical analysis of variable fixing techniques, deriving an asymptotic closed-form expression. This expression quantifies the proportion of binary variables that should be optimally fixed to zero. Our empirical results showcase the advantages of our approach, both in terms of computational efficiency and solution quality. Notably, we solve all the tested instances from literature to optimality, signifying the robustness and effectiveness of our proposed approach.

*Key words*: Chance Constraint; Distributionally Robust; Variable Fixing

---

## 1. Introduction

A Regular Chance Constrained Program (RCCP) that commits to a specific probability distribution, denoted as $\widehat{\mathbb{P}}$, takes on the following form:

$$v^* = \min_{\boldsymbol{x}\in\mathcal{X}} \left\{ \boldsymbol{c}^\top \boldsymbol{x} \colon \widehat{\mathbb{P}} \left\{ \tilde{\boldsymbol{\xi}} \colon \boldsymbol{a}_i(\boldsymbol{x})^\top \tilde{\boldsymbol{\xi}} \le b_i(\boldsymbol{x}), \forall i \in [I] \right\} \ge 1-\varepsilon \right\}, \tag{RCCP}$$

to minimize a linear objective function $\boldsymbol{c}^\top\boldsymbol{x}$ subject to a deterministic and compact set $\mathcal{X}$, an uncertain constraint system, defined by possibly multiple linear constraints $\boldsymbol{a}_i(\boldsymbol{x})^\top\tilde{\boldsymbol{\xi}} \le b_i(\boldsymbol{x})$ for all $i\in[I]$, must be satisfied with probability $1-\varepsilon$. In RCCP, the scalar $\varepsilon\in(0,1)$ is commonly referred to as the "risk parameter." The distribution $\widehat{\mathbb{P}}$ of random parameters $\tilde{\boldsymbol{\xi}}$ is an equiprobable empirical one generated from $N$ independent and identically distributed (i.i.d.) samples $\widehat{\boldsymbol{\xi}}^i_{\{i\in[N]\}}$, with $\widehat{\mathbb{P}}\{\tilde{\boldsymbol{\xi}} = \widehat{\boldsymbol{\xi}}^i\} = 1/N$. An RCCP can involve either a single uncertain constraint ($I=1$) or multiple uncertain constraints ($I>1$), leading to either a single RCCP or a joint RCCP, respectively. In both cases, the functions $\boldsymbol{a}_i(\boldsymbol{x})$ and $b_i(\boldsymbol{x})$ associated with each constraint $i\in[I]$ are affine, representing linear relationships between the decision variables $\boldsymbol{x}$ and the uncertain parameters $\tilde{\boldsymbol{\xi}}$.

When dealing with an RCCP with an arbitrary distribution, we often approximate the distribution using i.i.d. samples. By doing so, we can transform the problem into a Distributionally Robust Chance Constrained Program (DRCCP) with a carefully selected ambiguity set, which can achieve better out-of-sample performance guarantees. Formally, a DRCCP admits the form of

$$\min_{\boldsymbol{x}\in\mathcal{X}} \left\{ \boldsymbol{c}^\top \boldsymbol{x} \colon \inf_{\mathbb{P}\in\mathcal{P}} \mathbb{P} \left\{ \tilde{\boldsymbol{\xi}} \colon \boldsymbol{a}_i(\boldsymbol{x})^\top \tilde{\boldsymbol{\xi}} \le b_i(\boldsymbol{x}), \forall i \in [I] \right\} \ge 1-\varepsilon \right\}. \tag{DRCCP}$$

In DRCCP, the uncertain constraints are required to be satisfied with probability $1-\varepsilon$ for any probability distribution $\mathbb{P}$ from an ambiguity set $\mathcal{P}$, where set $\mathcal{P}$ is defined as a subset of probability distributions $\mathbb{P}$ from a measurable space $(\Omega, \mathcal{F})$ equipped with the sigma algebra $\mathcal{F}$ and induced by the random parameters $\tilde{\boldsymbol{\xi}}$ with support set $\Xi \subseteq \mathbb{R}^m$. It is worthy of mentioning that when $\mathcal{P}$ is a singleton, i.e., $\mathcal{P} = \{\widehat{\mathbb{P}}\}$, DRCCP reduces to RCCP.

## 1.1. Overview and Relevant Literature of Regular Chance Constrained Programs

Dated back to Charnes et al. (1958), Charnes and Cooper (1963), RCCPs have been extensively studied and applied in various domains to address uncertain constraints in decision-making problems. For example, RCCPs can be used to ensure a reliable and high-quality communication network by restricting the probability of network congestion or data transmission failures (see, e.g., Wang et al. 2014) to be small. RCCPs can also be employed to ensure environmental regulations and safety constraints are met with high probability, e.g., in the area of emissions or pollutant concentrations when uncertainties are prevalent (see, e.g., Simic 2016). In the context of power systems, RCCPs address the challenges posed by uncertainties related to renewable energy sources, demand fluctuations, transmission line failures, and other stochastic factors that can impact system performance (see, e.g., Wu et al. 2014, Lubin et al. 2015, Cho and Papavasiliou 2023). For a comprehensive review of applications in RCCPs, interested readers are referred to Ahmed and Shapiro (2008), Ahmed and Xie (2018), Küçükyavuz and Jiang (2022).

Despite their significance and wide applicability, RCCPs face primary challenges arising from their nonconvex feasible regions. Under finite support, RCCPs can be reformulated as mixed-integer programs (see, e.g., Ruszczyński 2002, Luedtke and Ahmed 2008). To improve the efficiency of solving RCCPs, researchers have focused on reducing the values of big-M coefficients and strengthening the corresponding formulations (see, e.g., Qiu et al. 2014, Song et al. 2014, Song and Shen 2016, Deng and Shen 2016, Zhang et al. 2020). For example, Song et al. (2014) developed a computationally efficient big-M coefficient strengthening procedure for regular chance constrained binary packing problems. Meanwhile, existing methods focus on either inner or outer approximations of RCCPs (see, e.g., Calafiore and Campi 2006, Nemirovski and Shapiro 2006, 2007, Ahmed et al. 2017, Jiang and Xie 2022, 2023). For example, Nemirovski and Shapiro (2007) proposed convex inner approximations of the chance constraint by using CVaR approximation, which has been recently improved by ALSO-X and ALSO-X# (Jiang and Xie 2022, 2023). Ahmed et al. (2017) introduced several outer approximations for RCCPs based on nonanticipative Lagrangian dual. It has been reported that effective upper and lower bounds based on inner and outer approximations can improve the running time of the mixed-integer solvers (see, e.g., Fattahi et al. 2018). However, to the best of our knowledge, there is currently a lack of a systematic approach to effectively integrating inner and outer approximations. We fill this gap by developing a variable fixing procedure that can harvest the past efforts of inner and outer approximations and significantly enhance the solution performances of solving an RCCP.

## 1.2. Overview and Relevant Literature of Distributionally Robust Chance Constrained Programs under Wasserstein Ambiguity Set

A DRCCP has been recently widely studied in the literature (see, e.g., Küçükyavuz and Jiang 2022 and the references therein). Particularly, we study DRCCP under type $q-$Wasserstein ambiguity set. We believe the developed methodology in this work can also be directly applied to DRCCPs under other statistical distance-based ambiguity sets. The type $q-$Wasserstein ambiguity set that the distance between two distributions via the minimum transportation plan is no larger than a radius $\theta$ can be formally defined as

$$\mathcal{P}_q = \left\{\mathbb{P} \colon \mathbb{P}\left\{\tilde{\boldsymbol{\xi}} \in \Xi\right\} = 1, W_q(\mathbb{P}, \widehat{\mathbb{P}}) \le \theta\right\},$$

where for any $q \in [1, \infty]$, the $q-$Wasserstein distance is defined as

$$W_q(\mathbb{P}_1, \mathbb{P}_2) = \inf\left\{\left[\int_{\Xi\times\Xi} \|\boldsymbol{\xi}_1 - \boldsymbol{\xi}_2\|_p^q \, \mathbb{Q}(d\boldsymbol{\xi}_1, d\boldsymbol{\xi}_2)\right]^{\frac{1}{q}} : \begin{array}{c}\mathbb{Q} \text{ is a joint distribution of } \tilde{\boldsymbol{\xi}}_1 \text{ and } \tilde{\boldsymbol{\xi}}_2 \\ \text{with marginals } \mathbb{P}_1 \text{ and } \mathbb{P}_2, \text{ respectively}\end{array}\right\},$$

$\theta \geq 0$ is the Wasserstein radius, and $\widehat{\mathbb{P}}$ denotes the reference distribution induced by random parameters $\tilde{\boldsymbol{\xi}}$. Note that if $q=\infty$, the $\infty-$Wasserstein distance is reduced to

$$W_\infty(\mathbb{P}_1,\mathbb{P}_2)=\inf\left\{\text{ess.sup}\,\|\boldsymbol{\xi}_1-\boldsymbol{\xi}_2\|_p\,\mathbb{Q}(d\boldsymbol{\xi}_1,d\boldsymbol{\xi}_2):\begin{array}{c}\mathbb{Q}\text{ is a joint distribution of }\tilde{\boldsymbol{\xi}}_1\text{ and }\tilde{\boldsymbol{\xi}}_2\\ \text{with marginals }\mathbb{P}_1\text{ and }\mathbb{P}_2,\text{ respectively}\end{array}\right\}.$$

We thus focus on the following DRCCP:

$$v_q^*=\min_{\boldsymbol{x}\in\mathcal{X}}\left\{\boldsymbol{c}^\top\boldsymbol{x}:\ \inf_{\mathbb{P}\in\mathcal{P}_q}\mathbb{P}\left\{\tilde{\boldsymbol{\xi}}:\ \boldsymbol{a}_i(\boldsymbol{x})^\top\tilde{\boldsymbol{\xi}}\leq b_i(\boldsymbol{x}),\forall i\in[I]\right\}\geq 1-\varepsilon\right\}. \tag{1}$$

We use the corresponding equivalent reformulations of DRCCP (1) under type $\infty-$Wasserstein ambiguity set and type $q-$Wasserstein ambiguity set with $q\in[1,\infty)$, which are discussed in detail in the following sections.

In literature, DRCCPs have emerged as a popular approach to addressing decision-making problems under uncertainty when the distributional information is not fully known (see, e.g., Zymler et al. 2013, Hanasusanto et al. 2015, 2017, Xie and Ahmed 2018, Xie 2021, Ji and Lejeune 2021, Shen and Jiang 2021, Chen et al. 2022, Ho-Nguyen et al. 2022, Küçükyavuz and Jiang 2022, Jiang and Xie 2022, Chen et al. 2023, Shen and Jiang 2023, Ho-Nguyen et al. 2023, Jiang and Xie 2023), such as energy (see, e.g., Xie and Ahmed 2017, Zhou et al. 2021), transportation (see, e.g., Ghosal and Wiesemann 2020, Zhao and Zhang 2020), and telecommunications (see, e.g., Li et al. 2016, Zhai et al. 2022, Li et al. 2022a,b, 2023). For instance, in the updated lecture notes of Shapiro et al. (2021), the authors used DRCCPs to optimize investment portfolios while ensuring the probability of incurring unacceptable losses stays within a predetermined limit. When designing and optimizing the wireless communication networks (see, e.g., Li et al. 2022a,b, 2023), DRCCPs have been employed to optimize the reliability and capacity constraints with a high level of confidence. For a comprehensive review of DRCCPs, interested readers are referred to a recent survey from Küçükyavuz and Jiang (2022).

However, similar to RCCPs, solving DRCCPs to optimality can still be challenging due to their non-convex feasible regions. When the reference distribution is finitely supported, DRCCPs under type-1 Wasserstein ambiguity set can be reformulated as mixed-integer programs using the big-M method (see, e.g., Xie 2021, Chen et al. 2022). Since then, researchers have dedicated their efforts to enhancing the efficiency of solving DRCCPs by focusing on two aspects: reducing the values of big-M coefficients and strengthening the formulations (see, e.g., Wang et al. 2021, Ho-Nguyen et al. 2022, 2023, Porras et al. 2023). For example, Ho-Nguyen et al. (2022, 2023) applied quantile information to improve the big-M coefficients in the DRCCPs. In this work, we extend the approach developed by Song et al. (2014) to DRCCPs and present efficient methods for finding better big-M coefficients. Prior literature primarily focused on either inner or outer approximations of DRCCPs (see, e.g., Xie 2021, Jiang and Xie 2022, 2023, Chen et al. 2023). For example, Xie (2021) introduced VaR outer approximation for DRCCPs, while Chen et al. (2023) discussed CVaR inner approximation for DRCCPs. In this work, we unify inner and outer approximations of DRCCPs through variable fixing, which will significantly benefit the exact methods.

### 1.3. Relevant Literature of Variable Fixing

Fixing a certain number of variables in the optimization problems is an effective approach to accelerate the solution process (see, e.g., Savelsbergh 1994, Anstreicher et al. 1996, 1999, Fischetti and Lodi 2010, Posta et al. 2012, Wu et al. 2018, Atamturk and Gómez 2020, Li et al. 2022c). For example, Li et al. (2022c) used variable fixing techniques to derive effective optimality cuts on the binary variables in the D-optimal data fusion problem, resulting in more efficient branch and cut algorithms. Wu et al. (2018) developed algorithms in a min-max regret generalized assignment problem based on variable fixing. As far as we know, our work is the first one that exploits the variable fixing for the RCCPs or DRCCPs to enhance the optimization process.

### 1.4. Summary of Contributions

This paper studies variable fixing techniques for solving RCCPs and DRCCPs. Our main contributions are summarized below:

(i) We provide a generic framework for variable fixing in RCCPs and DRCCPs, harvesting the known upper and lower bounds to derive efficient optimality cuts;

(ii) We extend the big-M coefficient strengthening approach developed by Song et al. (2014) to DRCCPs and present efficient algorithms for improving big-M coefficients. Combining with variable fixing, we provide numerical evidence to solve the standard instances of DRCCPs under type $q-$Wasserstein ambiguity set to optimality with a much shorter time, where the parameter $q \in \{1, \infty\}$;

(iii) Under type $q-$Wasserstein ambiguity set with $q \in (1, \infty)$, since DRCCPs, in general, may not admit a mixed-integer convex programming formulation, we introduce a new conservative approximation to improve the known conservative conditional value-at-risk (CVaR) approximation while still allowing for a MIP reformulation and maintaining a reasonable solution time;

(iv) We provide theoretical asymptotic analysis of variable fixing. This analysis shows that the number of scenarios can be fixed under some particular DRCCPs or RCCPs; and

(v) We provide extensive numerical studies to demonstrate the effectiveness of our proposed methods. Our variable fixing methods close the gap for all reported instances.

The roadmap of contributions in our paper is shown in Figure 1. It is worth mentioning that all the instances in this paper are standard and can be found in Song et al. (2014).

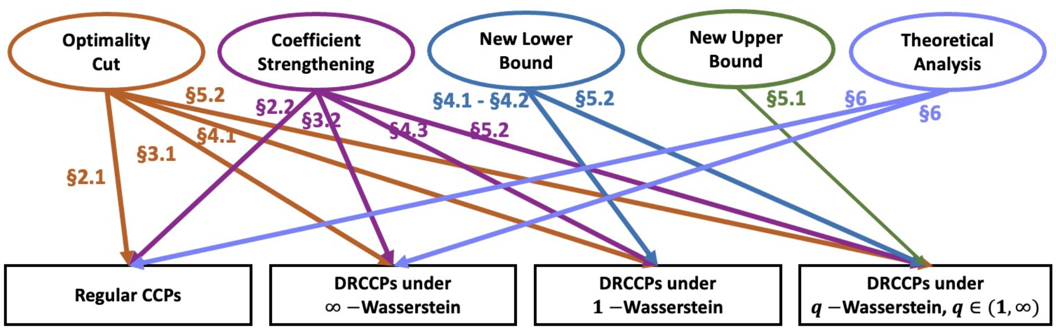



**Figure 1** A Roadmap of the Main Results in This Paper.

**Organization.** The remainder of the paper is organized as follows. Section 2 presents an overview of variable fixing and reviews how to strengthen big-M coefficients in RCCPs. Section 3, Section 4, and Section 5 study variable fixing optimality cuts and big-M coefficients improvement for DRCCPs under different types of Wasserstein ambiguity set. Section 6 provides a theoretical understanding of variable fixing. Section 7 numerically illustrates the proposed methods. Section 8 concludes the paper.

**Notation.** The following notation is used throughout the paper. We use bold letters (e.g., $\boldsymbol{x}, \boldsymbol{A}$) to denote vectors and matrices and use corresponding non-bold letters to denote their components. We let $\|\cdot\|_*$ denote the dual norm of a general norm $\|\cdot\|$. We let $\mathbf{e}$ be the vector or matrix of all ones, and let $\mathbf{e}_i$ be the $i$th standard basis vector. Given an integer $n$, we let $[n] := \{1, 2, \ldots, n\}$, and use $\mathbb{R}^n_+ := \{\boldsymbol{x} \in \mathbb{R}^n : x_i \geq 0, \forall i \in [n]\}$. Given a real number $\tau$, we let $(\tau)_+ := \max\{\tau, 0\}$. Given a finite set $I$, we let $|I|$ denote its cardinality. We let $\tilde{\boldsymbol{\xi}}$ denote a random vector and denote its realizations by $\boldsymbol{\xi}$. Given a vector $\boldsymbol{x} \in \mathbb{R}^n$, let $\text{supp}(\boldsymbol{x})$ be its support, i.e., $\text{supp}(\boldsymbol{x}) := \{i \in [n] : x_i \neq 0\}$. Given a probability distribution $\mathbb{P}$ on $\Xi$, we use $\mathbb{P}\{A\}$ to denote $\mathbb{P}\{\tilde{\boldsymbol{\xi}} : \text{condition } A(\tilde{\boldsymbol{\xi}}) \text{ holds}\}$ when $A(\boldsymbol{\xi})$ is a condition on $\boldsymbol{\xi}$, and to denote $\mathbb{P}\{\tilde{\boldsymbol{\xi}} : \tilde{\boldsymbol{\xi}} \in A\}$ when $A \subseteq \Xi$ is $\mathbb{P}-$measurable. We use $\lfloor x \rfloor$ to denote the largest integer $y$ satisfying $y \leq x$, for any $x \in \mathbb{R}$. For a given set $R$, the indicator function $\mathbb{I}(\boldsymbol{x} \in R) = 1$ if $\boldsymbol{x} \in R$, and 0, otherwise. Additional notations are introduced as needed.

## 2. Variable Fixing in RCCPs

Exploiting the discrete distribution $\widehat{\mathbb{P}}$, we can equivalently rewrite RCCP as

$$v^* = \min_{\boldsymbol{x}\in\mathcal{X}} \left\{ \boldsymbol{c}^\top \boldsymbol{x} \colon \sum_{j\in[N]} \mathbb{I}\left\{ \boldsymbol{a}_i(\boldsymbol{x})^\top \widehat{\boldsymbol{\xi}}^j \le b_i(\boldsymbol{x}), \forall i \in [I] \right\} \ge N - \lfloor N\varepsilon \rfloor \right\}, \tag{2}$$

where $\lfloor\cdot\rfloor$ denotes the floor function and $\mathbb{I}(\cdot)$ is the zero-one indicator function. Introducing binary variables $\boldsymbol{z}$ to replace the indicator functions and choosing appropriate big-M coefficients $M_{i,j}$ for each $i\in[I]$, $j\in[N]$ (e.g., $M_{i,j} \ge \max_{\boldsymbol{x}\in\mathcal{X}}\{\boldsymbol{a}_i(\boldsymbol{x})^\top \widehat{\boldsymbol{\xi}}^j - b_i(\boldsymbol{x})\}$), RCCP (2) can be written as the following mixed-integer linear program:

$$v^* = \min_{\substack{\boldsymbol{x}\in\mathcal{X},\\ \boldsymbol{z}\in\{0,1\}^N}} \left\{ \boldsymbol{c}^\top \boldsymbol{x} \colon \begin{array}{l} \boldsymbol{a}_i(\boldsymbol{x})^\top \widehat{\boldsymbol{\xi}}^j \le b_i(\boldsymbol{x}) + M_{i,j}(1-z_j), \forall i\in[I], j\in[N], \\ \sum_{j\in[N]} z_j \ge N - \lfloor N\varepsilon \rfloor \end{array} \right\}. \tag{3}$$

This section aims to select one or multiple binary variables $\boldsymbol{z}$ of RCCP (3) to restrict them to be either one or zero and then solve the restricted RCCP to obtain effective optimality cuts on these binary variables if the restricted problem has a higher objective value than $v^*$ or the upper bound of a heuristic approach. This procedure has the potential to significantly decrease the number of binary variables in RCCP while still attaining the same optimal objective value.

### 2.1. Variable Fixing for the RCCP

Suppose that $\mathcal{S}_0$ and $\mathcal{S}_1$ are sets representing the index sets of scenarios to be removed (i.e., $z_j = 0$ for each $j\in\mathcal{S}_0$) and selected (i.e., $z_j = 1$ for each $j\in\mathcal{S}_1$), respectively. If we introduce the additional constraints $\sum_{j\in\mathcal{S}_0} z_j = 0$ and $\sum_{j\in\mathcal{S}_1} z_j = |\mathcal{S}_1|$ to RCCP (3), the corresponding Restricted RCCP can be formulated as follows:

$$\bar{v}(\mathcal{S}_0,\mathcal{S}_1) = \min_{\substack{\boldsymbol{x}\in\mathcal{X},\\ \boldsymbol{z}\in\{0,1\}^N}} \left\{ \boldsymbol{c}^\top \boldsymbol{x} \colon \begin{array}{l} \boldsymbol{a}_i(\boldsymbol{x})^\top \widehat{\boldsymbol{\xi}}^j \le b_i(\boldsymbol{x}) + M_{i,j}(1-z_j), \forall i\in[I], j\in[N], \\ \sum_{j\in[N]} z_j \ge N - \lfloor N\varepsilon \rfloor, \sum_{j\in\mathcal{S}_0} z_j = 0, \sum_{j\in\mathcal{S}_1} z_j = |\mathcal{S}_1| \end{array} \right\}. \tag{Restricted RCCP}$$

Hence, we can deduce that $\bar{v}(\mathcal{S}_0,\mathcal{S}_1) \ge v^*$. For instance, when $\mathcal{S}_0 = \mathcal{S}_1 = \emptyset$, both the Restricted RCCP and the original RCCP (3) coincide, resulting in $\bar{v}(\mathcal{S}_0,\mathcal{S}_1) = v^*$. However, in cases where $\mathcal{S}_0$ and $\mathcal{S}_1$ are not empty, we have $\bar{v}(\mathcal{S}_0,\mathcal{S}_1) > v^*$, indicating that at least one constraint in the sets $\mathcal{S}_0$ and $\mathcal{S}_1$ is violated by any optimal solution of the RCCP. Consequently, we can derive an effective optimality cut.

THEOREM 1. *For any two disjoint sets $\mathcal{S}_0,\mathcal{S}_1 \subseteq [N]$ with $|\mathcal{S}_0| \le \lfloor N\varepsilon \rfloor$, if $\bar{v}(\mathcal{S}_0,\mathcal{S}_1) > v^*$, then the following inequality is valid for any optimal solution of RCCP (3):*

$$\sum_{j\in\mathcal{S}_0} z_j + \sum_{j\in\mathcal{S}_1} (1-z_j) \ge 1,$$

*with $\sum_{j\in[N]} z_j \ge N - \lfloor N\varepsilon \rfloor, \boldsymbol{z}\in\{0,1\}^N$.*

*Proof.* Let $\boldsymbol{z}^*$ be an optimal solution of RCCP (3) with $\sum_{j\in\mathcal{S}_0} z_j^* = 0, \sum_{j\in\mathcal{S}_1} z_j^* = |\mathcal{S}_1|$, then $\boldsymbol{z}^*$ is feasible to the Restricted RCCP and achieves the same objective value $v^*$, which violates the assumption that $\bar{v}(\mathcal{S}_0,\mathcal{S}_1) > v^*$. This indicates that one of the equality constraints added is violated. Since $\boldsymbol{z}^*$ is binary, we know one of the following inequalities is valid:

$$\sum_{j\in\mathcal{S}_0} z_j \ge 1, \sum_{j\in\mathcal{S}_1} z_j \le |\mathcal{S}_1| - 1.$$

Since sets $\mathcal{S}_0,\mathcal{S}_1$ are disjoint, we obtain a desired optimality cut. □

Notably, in alignment with the convention commonly employed in the variable fixing literature (as seen in, for example, section 1.2 of Savelsbergh 1994), we designate the process of deriving the optimality cut in Theorem 1 as the "variable fixing" technique applied to the RCCP (3).

However, performing variable fixing can be computationally expensive, since it necessitates solving Restricted RCCP to optimality, and also requires knowledge of the optimal objective value $v^*$ of RCCP (3). In order to address this issue, we resort to using computationally tractable yet slightly conservative inner approximations of RCCP (3), and outer approximations of Restricted RCCP. Let $v^U$ be the upper bound of RCCP (3), and $\bar{v}^L(\mathcal{S}_0,\mathcal{S}_1)$ be the lower bound of Restricted RCCP. If $\bar{v}^L(\mathcal{S}_0,\mathcal{S}_1) > v^U$, then the result in Theorem 1 still holds. To find $v^U$, we can use convex inner approximations such as the CVaR approximation (as described in, for example, Nemirovski and Shapiro 2007), ALSO-X approximation (explored in Ahmed et al. 2017, Jiang and Xie 2022), or the more recent ALSO-X# approximation presented in Jiang and Xie (2023). On the other hand, to compute $\bar{v}^L(\mathcal{S}_0,\mathcal{S}_1)$, we can employ dual formulations and the primal counterparts from Ahmed et al. (2017), such as their formulation (23) or formulation (24). Additionally, Ahmed et al. (2017), Ahmed and Xie (2018) offer further alternative choices of valid upper bound $v^U$ and easily computable $\bar{v}^L(\mathcal{S}_0,\mathcal{S}_1)$.

Specifically, to derive the optimality cuts, we start by introducing the constraints $\sum_{j\in\mathcal{S}_0} z_j = 0$ and $\sum_{j\in\mathcal{S}_1} z_j = |\mathcal{S}_1|$ into the dual formulation (as presented in, for example, formulation (23) or formulation (24) in Ahmed et al. 2017). Following the conventions in Ahmed et al. (2017), let us suppose that $\mathcal{X} \subseteq \{\boldsymbol{x} : \boldsymbol{D}^\top \boldsymbol{x} \le \boldsymbol{d}\}$. Then, we can solve the following restricted dual formulation (4) to find the optimality cuts.

Corollary 1. *Suppose that set $\mathcal{X} \subseteq \{\boldsymbol{x} : \boldsymbol{D}^\top \boldsymbol{x} \le \boldsymbol{d}\}$, for any two disjoint sets $\mathcal{S}_0,\mathcal{S}_1 \subseteq [N]$ with $|\mathcal{S}_0| \le \lfloor N\varepsilon \rfloor$. Then one can solve the following problem:*

$$\widehat{v}(\mathcal{S}_0,\mathcal{S}_1) = \min_{\substack{\boldsymbol{x},\boldsymbol{u},\\ \boldsymbol{w},\boldsymbol{z}\in[0,1]^N,\widehat{y}}} \left\{ \widehat{y} : \begin{array}{l} \widehat{y} \ge \boldsymbol{c}^\top \boldsymbol{u}^j + v^U(1-z_j), \widehat{y} \ge \boldsymbol{c}^\top \boldsymbol{w}^j + v^U z_j, \forall j \in [N]\setminus(\mathcal{S}_1\cup\mathcal{S}_0),\\ z_j \boldsymbol{a}_i \left(\boldsymbol{u}^j/z_j\right)^\top \widehat{\boldsymbol{\xi}}^j \le z_j b_i(\boldsymbol{u}^j/z_j), \forall i\in[I], j\in[N],\\ \boldsymbol{D}^\top \boldsymbol{u}^j \le \boldsymbol{d} z_j, \forall j\in[N], \boldsymbol{D}^\top \boldsymbol{w}^j \le \boldsymbol{d}(1-z_j), \forall j \in [N],\\ \boldsymbol{x} = \boldsymbol{u}^j + \boldsymbol{w}^j, \forall j\in[N],\\ \sum_{j\in[N]} z_j \ge N - \lfloor N\varepsilon\rfloor, \sum_{j\in\mathcal{S}_0} z_j = 0, \sum_{j\in\mathcal{S}_1} z_j = |\mathcal{S}_1| \end{array} \right\}. \tag{4}$$

*If $\widehat{v}(\mathcal{S}_0,\mathcal{S}_1) \ge v^U$, then the following inequality is valid for any optimal solution of RCCP (3):*

$$\sum_{j\in\mathcal{S}_0} z_j + \sum_{j\in\mathcal{S}_1}(1-z_j) \ge 1,$$

*with $\sum_{j\in[N]} z_j \ge N - \lfloor N\varepsilon\rfloor, \boldsymbol{z}\in\{0,1\}^N$.*

We make the following remarks on Corollary 1:

(i) The result in Corollary 1 can be further extended to DRCCPs under the type $q-$Wasserstein ambiguity set. The detailed extension are presented in the following sections; and

(ii) Alternatively, instead of solving the restricted dual problem (4), one can verify the feasibility of the restricted dual bound. If the following problem is infeasible, then the same result as in Corollary 1 holds:

$$\min_{\substack{\boldsymbol{x},\boldsymbol{u},\\ \boldsymbol{w},\boldsymbol{z}\in[0,1]^N}} \left\{ 0 : \begin{array}{l} \boldsymbol{c}^\top \boldsymbol{u}^j \le v^U z_j, \forall j\in[N], \boldsymbol{c}^\top \boldsymbol{w}^j \le v^U(1-z_j), \forall j \in [N],\\ z_j \boldsymbol{a}_i \left(\boldsymbol{u}^j/z_j\right)^\top \widehat{\boldsymbol{\xi}}^j \le z_j b_i(\boldsymbol{u}^j/z_j), \forall i\in[I], j\in[N],\\ \boldsymbol{D}^\top \boldsymbol{u}^j \le \boldsymbol{d} z_j, \forall j\in[N], \boldsymbol{D}^\top \boldsymbol{w}^j \le \boldsymbol{d}(1-z_j), \forall j \in [N],\\ \boldsymbol{x} = \boldsymbol{u}^j + \boldsymbol{w}^j, \forall j\in[N],\\ \sum_{j\in[N]} z_j \ge N - \lfloor N\varepsilon\rfloor, \sum_{j\in\mathcal{S}_0} z_j = 0, \sum_{j\in\mathcal{S}_1} z_j = |\mathcal{S}_1| \end{array} \right\}.$$

It is essential to highlight that, besides the variable fixing procedure described in Corollary 1, there exists a straightforward and effective alternative method. This approach involves fixing the variables individually for each scenario using the upper bound $v^U$, as summarized below.

COROLLARY 2. *Let* $\widehat{\eta}_j = \min_{\boldsymbol{x}\in\mathcal{X}}\{\boldsymbol{c}^\top\boldsymbol{x} : \boldsymbol{a}_i(\boldsymbol{x})^\top\widehat{\boldsymbol{\xi}}^j \le b_i(\boldsymbol{x}), \forall i\in[I]\}$ *for each* $j\in[N]$. *Suppose that for some* $j\in[N]$, *we have* $\widehat{\eta}_j > v^U$. *Then* $z_j = 0$.

*Proof.* The inequality $\widehat{\eta}_j > v^U$ implies that

$$\widehat{\eta}_j = \min_{\boldsymbol{x}\in\mathcal{X}}\left\{\boldsymbol{c}^\top\boldsymbol{x} : \boldsymbol{a}_i(\boldsymbol{x})^\top\widehat{\boldsymbol{\xi}}^j \le b_i(\boldsymbol{x}), \forall i\in[I]\right\} > v^U \ge v^*.$$

If at optimality, scenario $j$ were satisfied, we must have $v^* \ge \widehat{\eta}_j$, a contradiction. Therefore, we have $z_j = 0$. □

Notice that the optimality cuts found in Corollary 1 and Corollary 2 can reduce the big-M coefficients in RCCP (3), which is detailed in the next subsection.

### 2.2. Big-M Coefficients Strengthening for RCCP (3)

By employing variable fixing techniques, we can further enhance the big-M coefficients in the RCCP (3). Following the variable fixing procedure outlined in Theorem 1, we can reduce these big-M coefficients using the approach described in section 3.1 of Song et al. (2014). Specifically, after identifying a subset $\widehat{\mathcal{S}}_0$ of scenarios that must be violated via variable fixing (i.e., $z_j = 0$ for each $j\in\widehat{\mathcal{S}}_0$), or each given scenario $j\in[N]\setminus\widehat{\mathcal{S}}_0$, we calculate

$$\eta_{i,j}(j') := \max_{\boldsymbol{x}}\left\{\boldsymbol{a}_i(\boldsymbol{x})^\top\widehat{\boldsymbol{\xi}}^j - b_i(\boldsymbol{x}) \colon \boldsymbol{a}_i(\boldsymbol{x})^\top\widehat{\boldsymbol{\xi}}^{j'} \le b_i(\boldsymbol{x}), \boldsymbol{x}\in\mathcal{X}\right\},$$

then we sort $\{\eta_{i,j}(j')\}_{j'\in[N]\setminus\widehat{\mathcal{S}}_0}$ in the nondecreasing order and select the $\lfloor\varepsilon N\rfloor - |\widehat{\mathcal{S}}_0| + 1$ one as the corresponding big-M coefficient. In our numerical study, we see that using the simple variable fixing in Corollary 2, we can reduce the strengthened big-M coefficients significantly. We will discuss how to strengthen the big-M coefficients in DRCCPs, elaborated in the following sections.

### 2.3. Implementation of the Variable Fixing Procedure

The implementation of the variable fixing procedure involves the following steps:

1. Pre-compute Upper Bound: Using a reliable approximation algorithm, such as ALSO-X# as proposed in Jiang and Xie (2023), calculate an upper bound denoted as $v^U$ for RCCP (3).
2. Evaluate Restricted Lower Bound: Next, evaluate the restricted lower bound of RCCP (3) by fixing the values of certain binary variables.
3. Compare Bounds and Derive Optimality Cuts: Compare the restricted lower bound with the pre-computed upper bound $v^U$. Based on the comparison results, we can derive effective optimality cuts according to Corollary 1 and Corollary 2.
4. Strengthen Big-M Coefficients: Reduce the value of big-M coefficients in RCCP (3) using the approach proposed by Song et al. (2014).

The detailed procedure of the variable fixing process is provided in Algorithm 1, where the algorithm outlines the steps for efficiently handling the binary variables $\boldsymbol{z}$ in RCCP (3). By applying the variable fixing procedure, the algorithm helps refine the solution space and potentially leads to better performances when solving RCCP (3).

We use the following example to illustrate how Algorithm 1 works.

EXAMPLE 1. Consider an RCCP with 5 equiprobable scenarios (i.e., $\widehat{\mathbb{P}}\{\tilde{\boldsymbol{\xi}} = \widehat{\boldsymbol{\xi}}^j\} = 1/5$ for each $j\in[5]$), risk level $\varepsilon = 1/2$, set $\mathcal{X} = [0,1]^2$, function $\boldsymbol{a}_1(\boldsymbol{x})^\top\widehat{\boldsymbol{\xi}} - b_1(\boldsymbol{x}) = \widehat{\xi}_1 - \widehat{\boldsymbol{\xi}}_2^\top\boldsymbol{x}$, and $\widehat{\xi}_1^1 = 1$, $\widehat{\xi}_1^2 = 2$,

**Algorithm 1** Variable Fixing Procedure for RCCP (3)

1: Pre-compute an upper bound $v^U$ for RCCP (3)
2: Fix the variables individually for each scenario based on Corollary 2
3: Apply the optimality cuts from Step 2 and find the optimality cuts based on Corollary 1 by comparing the restricted lower bound $\widehat{v}(\mathcal{S}_0,\mathcal{S}_1)$ of the RCCP (3) and upper bound $v^U$
4: Strengthen big-M coefficients in RCCP (3) using method in section 3.1 of Song et al. (2014)

$\widehat{\xi}_1^3=3/2$, $\widehat{\xi}_1^4=2$, $\widehat{\xi}_1^5=3/2$, $\widehat{\boldsymbol{\xi}}_2^1=(2/3,4)^\top$, $\widehat{\boldsymbol{\xi}}_2^2=(5/2,2)^\top$, $\widehat{\boldsymbol{\xi}}_2^3=(5,2)^\top$, $\widehat{\boldsymbol{\xi}}_2^4=(2,3)^\top$, $\widehat{\boldsymbol{\xi}}_2^5=(1,8/3)^\top$. In this case, RCCP (3) reduces to the following mixed-integer linear program:

$$v^*=\min_{\substack{\boldsymbol{x}\in[0,1]^2,\\ \boldsymbol{z}\in\{0,1\}^5}}\left\{-x_1-x_2\colon \begin{array}{c} \frac{2}{3}x_1+4x_2\le z_1+M_{1,1}(1-z_1), \frac{5}{2}x_1+2x_2\le 2z_2+M_{1,2}(1-z_2),\\ 5x_1+2x_2\le \frac{3}{2}z_3+M_{1,3}(1-z_3), 2x_1+3x_2\le 2z_4+M_{1,4}(1-z_4),\\ x_1+\frac{8}{3}x_2\le \frac{3}{2}z_5+M_{1,5}(1-z_5), \sum_{j\in[5]} z_j\ge 3 \end{array}\right\}.$$

In this particular example, the upper bound is obtained using ALSO-X# from Jiang and Xie (2023) is $v^U=-0.8571$ with an error bound of $[-10^{-4},10^{-4}]$. For each scenario, we compute $\widehat{\boldsymbol{\eta}}=[-13/12,-1.0,-3/4,-1,-19/16]$. Based on Corollary 2, we derive an optimality cut $z_3^*=0$.

Next, let us consider the situation where we add $z_5=0$ along with the optimality cut $z_3^*=0$. Due to the chance constraint, this implies that $z_1=z_2=z_4=1$. With these fixed values for the binary variables, we can determine the objective function value, which is found to be $v=-0.8269$ with an error bound of $[-10^{-4},10^{-4}]$. Importantly, we observe that $v>v^U$, which means that after applying the variable fixing procedure, we find another optimality cut $z_5^*=1$.

Notice that big-M coefficients can be further reduced after fixing $z_3^*=0$. Without fixing, we use the approach described in section 3.1 of Song et al. (2014) and we have

$$M_{1,1}=2, M_{1,2}=\frac{2}{3}, M_{1,3}=\frac{11}{3}, M_{1,4}=\frac{9}{16}, M_{1,5}=\frac{1}{2}.$$

After fixing $z_3^*=0$, we do not need to compute $M_{1,3}$, and for other four big-M coefficients, we have

$$M_{1,1}=\frac{5}{3}, M_{1,2}=\frac{2}{3}, M_{1,4}=\frac{9}{16}, M_{1,5}=\frac{5}{18}.$$

Two of these coefficients, $M_{1,1}$ and $M_{1,5}$, have been improved. ⋄

In summary, the variable fixing procedure is instrumental in refining the solution space and can expedite the solution procedure of exact big-M methods, which is further illustrated in the numerical study Section 7.

## 3. Variable Fixing for DRCCPs under Type $\infty-$Wasserstein Ambiguity Set

In this section, we consider the variable fixing in DRCCP (1) under type $\infty-$Wasserstein ambiguity set with the discrete reference distribution $\widehat{\mathbb{P}}$.

### 3.1. Variable Fixing Procedure for DRCCPs under Type $\infty-$Wasserstein Ambiguity Set

According to the equivalent reformulation in proposition 8 of Jiang and Xie (2022), DRCCP (1) under type $\infty-$Wasserstein ambiguity set can be written as

$$v_\infty^*=\min_{\boldsymbol{x}\in\mathcal{X}}\left\{\boldsymbol{c}^\top\boldsymbol{x}\colon \sum_{j\in[N]}\mathbb{I}\left\{\theta\left\|\boldsymbol{a}_i(\boldsymbol{x})\right\|_*+\boldsymbol{a}_i(\boldsymbol{x})^\top\widehat{\boldsymbol{\xi}}^j\le b_i(\boldsymbol{x}),\forall i\in[I]\right\}\ge N-\lfloor N\varepsilon\rfloor\right\}. \quad (5)$$

Introducing binary variables $\boldsymbol{z}$ to replace the indicator functions and choosing an appropriate big-M coefficient $M_{i,j}$ for each $i \in [I]$, $j \in [N]$ (e.g., $M_{i,j} \geq \max_{\boldsymbol{x}\in\mathcal{X}}\{\theta\|\boldsymbol{a}_i(\boldsymbol{x})\|_* + \boldsymbol{a}_i(\boldsymbol{x})^\top\widehat{\boldsymbol{\xi}}^j - b_i(\boldsymbol{x})\}$), DRCCP (5) can be written as the following mixed-integer convex program:

$$v_\infty^* = \min_{\substack{\boldsymbol{x}\in\mathcal{X},\\ \boldsymbol{z}\in\{0,1\}^N}} \left\{ \boldsymbol{c}^\top\boldsymbol{x} \colon \begin{array}{l} \theta\left\|\boldsymbol{a}_i(\boldsymbol{x})\right\|_* + \boldsymbol{a}_i(\boldsymbol{x})^\top\widehat{\boldsymbol{\xi}}^j \leq b_i(\boldsymbol{x}) + M_{i,j}(1-z_j), \forall i\in[I], j\in[N], \\ \sum_{j\in[N]} z_j \geq N - \lfloor N\varepsilon \rfloor \end{array} \right\}. \tag{6}$$

Notice that RCCP (2) and DRCCP (6) differ only by the presence of the term $\theta\|\boldsymbol{a}_i(\boldsymbol{x})\|_*$ for each $i \in [I]$ in DRCCP (6). Hence, all the results regarding finding optimality cuts through variable fixing in Section 2.1 can be readily extended to DRCCP (6). For instance, the optimality cuts can be derived by comparing the restricted lower bound $\widehat{v}_\infty(\mathcal{S}_0,\mathcal{S}_1)$ of DRCCP (6) with the upper bound $v_\infty^U$ of DRCCP (6), as detailed below.

COROLLARY 3. *Suppose that $\mathcal{X} \subseteq \{\boldsymbol{x} : \boldsymbol{D}^\top\boldsymbol{x} \leq \boldsymbol{d}\}$, for any two disjoint sets $\mathcal{S}_0, \mathcal{S}_1 \subseteq [N]$ with $|\mathcal{S}_0| \leq \lfloor N\varepsilon \rfloor$. Then one can solve the following problem:*

$$\widehat{v}_\infty(\mathcal{S}_0,\mathcal{S}_1) = \min_{\substack{\boldsymbol{x},\boldsymbol{u},\\ \boldsymbol{w},\boldsymbol{z}\in[0,1]^N,\widehat{y}}} \left\{ y \colon \begin{array}{l} \widehat{y} \geq \boldsymbol{c}^\top\boldsymbol{u}^j + v_\infty^U(1-z_j), \widehat{y} \geq \boldsymbol{c}^\top\boldsymbol{w}^j + v_\infty^U z_j, \forall j \in [N]\setminus(\mathcal{S}_1\cup\mathcal{S}_0), \\ \theta\left\|\boldsymbol{a}_i\left(\boldsymbol{u}^j\right)\right\|_* + z_j\boldsymbol{a}_i\left(\boldsymbol{u}^j/z_j\right)^\top\widehat{\boldsymbol{\xi}}^j \leq z_j b_i(\boldsymbol{u}^j/z_j), \forall i\in[I], j\in[N], \\ \boldsymbol{D}^\top\boldsymbol{u}^j \leq \boldsymbol{d}z_j, \forall j\in[N], \boldsymbol{D}^\top\boldsymbol{w}^j \leq \boldsymbol{d}(1-z_j), \forall j\in[N], \\ \boldsymbol{x} = \boldsymbol{u}^j + \boldsymbol{w}^j, \forall j\in[N], \\ \sum_{j\in[N]} z_j \geq N - \lfloor N\varepsilon\rfloor, \sum_{j\in\mathcal{S}_0} z_j = 0, \sum_{j\in\mathcal{S}_1} z_j = |\mathcal{S}_1| \end{array} \right\}. \tag{7}$$

*If $\widehat{v}_\infty(\mathcal{S}_0,\mathcal{S}_1) \geq v_\infty^U$, where $v_\infty^U$ is an upper bound of DRCCP (6), then the following inequality is valid for any optimal solution of DRCCP (6):*

$$\sum_{j\in\mathcal{S}_0} z_j + \sum_{j\in\mathcal{S}_1}(1-z_j) \geq 1,$$

*with $\sum_{j\in[N]} z_j \geq N - \lfloor N\varepsilon \rfloor, \boldsymbol{z} \in \{0,1\}^N$.*

It is worth noting that when the dual norm is $\|\cdot\|_* = \|\cdot\|_p$ with $p \in (1,\infty)$, the restricted dual bound problem (7) comprises a minimum of $NI$ conic constraints, posing a considerable challenge in numerical computations. As a result, it may be more practical to use this procedure solely when $p$ takes on values from the set $\{1,\infty\}$.

Moreover, the variable fixing technique described in Corollary 2 can be straightforwardly extended to address DRCCP (6) for any $p \in [1,\infty]$.

COROLLARY 4. *Let $\widehat{\eta}_{\infty,j} = \min_{\boldsymbol{x}\in\mathcal{X}}\{\boldsymbol{c}^\top\boldsymbol{x} : \theta\|\boldsymbol{a}_i(\boldsymbol{x})\|_* + \boldsymbol{a}_i(\boldsymbol{x})^\top\widehat{\boldsymbol{\xi}}^j \leq b_i(\boldsymbol{x}), \forall i \in [I]\}$ for each $j \in [N]$. Suppose that for some $j \in [N]$, we have $\widehat{\eta}_{\infty,j} > v_\infty^U$. Then $z_j = 0$.*

In summary, the solution approach for DRCCP (6) involves the following steps. First, we calculate the upper bound for the problem. Then, we apply the variable fixing technique outlined in Corollary 4 to fix the variables individually for each scenario. If the value of $p$ falls within the set $\{1,\infty\}$, we proceed to check the restricted lower bound of DRCCP (6) and compare it with the upper bound using Corollary 3. Our numerical study confirms that this systematic approach is highly effective in deriving optimality cuts and significantly expediting the solution process of DRCCP (6).

### 3.2. Big-M Coefficients Strengthening for DRCCPs under Type $\infty-$Wasserstein Ambiguity Set

In this subsection, we focus on strengthening the big-M coefficients by building on the insights and discussions presented in Section 2.2. Specifically, for a given scenario $j \in [N]$, we compute

$$\eta_{i,j}(j'|\theta) = \max_{\boldsymbol{x}\in\mathcal{X}} \left\{ \theta\left\|\boldsymbol{a}_i(\boldsymbol{x})\right\|_* + \boldsymbol{a}_i(\boldsymbol{x})^\top\widehat{\boldsymbol{\xi}}^j - b_i(\boldsymbol{x}) \colon \theta\left\|\boldsymbol{a}_i(\boldsymbol{x})\right\|_* + \boldsymbol{a}_i(\boldsymbol{x})^\top\widehat{\boldsymbol{\xi}}^{j'} \leq b_i(\boldsymbol{x}) \right\}, \tag{8}$$

which is, in general, NP-hard to solve for any convex $L_p$ norm with $\|\cdot\|_* = \|\cdot\|_p$ and $p \in [1,\infty)$ and $\theta > 0$.

THEOREM 2. *Suppose $\theta > 0$. For any dual norm $\|\cdot\|_* = \|\cdot\|_p$ and $p \in [1,\infty)$, solving Problem (8), in general, is NP-hard.*

*Proof.* See Appendix A.1. □

Despite the difficulty of computing $\eta_{i,j}(j'|\theta)$ for general norms, it turns out that for the inf-norm (i.e., when $p = \infty$), Problem (8) can be tractable.

PROPOSITION 1. *Suppose that in Problem (8), set $\mathcal{X}$ is compact and convex and the dual norm $\|\cdot\|_* = \|\cdot\|_\infty$. Then Problem (8) is equivalent to solving $2n$ tractable convex programs, i.e., $\eta_{i,j}(j'|\theta) = \max_{\tau\in[n]} \max_{k\in[n]} \max_{\ell\in[2]} \eta_{i,j}(j',\tau,\ell|\theta)$, where for each $\tau \in [n]$, we have*

$$\eta_{i,j}(j',\tau,1|\theta) = \max_{\boldsymbol{x}\in\mathcal{X}} \left\{\theta a_{i\tau}(\boldsymbol{x}) + \boldsymbol{a}_i(\boldsymbol{x})^\top \widehat{\boldsymbol{\xi}}^j - b_i(\boldsymbol{x}) \colon \theta \|\boldsymbol{a}_i(\boldsymbol{x})\|_\infty + \boldsymbol{a}_i(\boldsymbol{x})^\top \widehat{\boldsymbol{\xi}}^{j'} \le b_i(\boldsymbol{x})\right\},$$

$$\eta_{i,j}(j',\tau,2|\theta) = \max_{\boldsymbol{x}\in\mathcal{X}} \left\{-\theta a_{i\tau}(\boldsymbol{x}) + \boldsymbol{a}_i(\boldsymbol{x})^\top \widehat{\boldsymbol{\xi}}^j - b_i(\boldsymbol{x}) \colon \theta \|\boldsymbol{a}_i(\boldsymbol{x})\|_\infty + \boldsymbol{a}_i(\boldsymbol{x})^\top \widehat{\boldsymbol{\xi}}^{j'} \le b_i(\boldsymbol{x})\right\}.$$

*Proof.* See Appendix A.2. □

In the context of the general dual norm $\|\cdot\|_* = \|\cdot\|_p$ with $p \in [1,\infty)$, we aim to identify nontrivial conditions under which Problem (8) becomes more manageable to solve. To accomplish this, we first employ strong duality and present an equivalent reformulation of Problem (8).

PROPOSITION 2. *Suppose that in Problem (8), set $\mathcal{X}$ is compact and convex. Let*

$$\bar{v}^P_{\infty,i,j,1}(j'|\theta) = \max_{\boldsymbol{x}\in\mathcal{X}} \left\{\boldsymbol{a}_i(\boldsymbol{x})^\top \left[\widehat{\boldsymbol{\xi}}^j - \widehat{\boldsymbol{\xi}}^{j'}\right] \colon \theta \|\boldsymbol{a}_i(\boldsymbol{x})\|_* + \boldsymbol{a}_i(\boldsymbol{x})^\top \widehat{\boldsymbol{\xi}}^{j'} \le b_i(\boldsymbol{x})\right\}, \tag{9a}$$

$$\bar{v}^P_{\infty,i,j,2}(j'|\theta) = \min_{0\le\alpha\le1} \max_{\boldsymbol{x}\in\mathcal{X}} \left\{(\alpha-1)b_i(\boldsymbol{x}) + (1-\alpha)\theta\|\boldsymbol{a}_i(\boldsymbol{x})\|_* + \sum_{k\in[n]} \left(\widehat{\xi}^j_k - \alpha\widehat{\xi}^{j'}_k\right) a_{ik}(\boldsymbol{x})\right\}. \tag{9b}$$

*Let $\bar{\eta}_{i,j}(j'|\theta)$ be the minimum between $\bar{v}^P_{\infty,i,j,1}(j'|\theta)$ and $\bar{v}^P_{\infty,i,j,2}(j'|\theta)$, i.e., $\bar{\eta}_{i,j}(j'|\theta) = \min\{\bar{v}^P_{\infty,i,j,1}(j'|\theta), \bar{v}^P_{\infty,i,j,2}(j'|\theta)\}$. Then*

*(i) The optimal value of Problem (8) is upper bounded by $\bar{\eta}_{i,j}(j'|\theta)$, i.e., $\eta_{i,j}(j'|\theta) \le \bar{\eta}_{i,j}(j'|\theta)$; and*

*(ii) When the inner maximization in Problem (9b) admits a unique solution with $\alpha = 0$, the optimal value of Problem (8) is equal to $\bar{\eta}_{i,j}(j'|\theta)$, i.e., $\eta_{i,j}(j'|\theta) = \bar{\eta}_{i,j}(j'|\theta)$.*

*Proof.* See Appendix A.3. □

We remark that Problem (9a) can be viewed as a convex optimization problem, which is easy to solve. However, Problem (9b) can still be challenging to solve. For example, suppose $\widehat{\boldsymbol{\xi}}^j = \widehat{\boldsymbol{\xi}}^{j'} = \boldsymbol{0}, \boldsymbol{a}_i(\boldsymbol{x}) = \boldsymbol{x}, b_i(\boldsymbol{x}) = b_i > \max_{\boldsymbol{x}\in\mathcal{X}} \theta\|\boldsymbol{x}\|_*$, then optimal $\alpha^*$ in Problem (9b) is $\alpha^* = 0$ and the remaining problem (9b) is $\bar{v}^P_{\infty,i,j,2}(j'|\theta) = \max_{\boldsymbol{x}\in\mathcal{X}} \theta\|\boldsymbol{x}\|_*$, which, in general, is NP-hard, according to Theorem 2.

Consequently, to address this complexity, we further explore the structural properties of set $\mathcal{X}$ such that the upper bound of Problem (8) can be efficiently computable using the golden section search method. One such case is when set $\mathcal{X} = [0,1]^n$, the dual norm $\|\cdot\|_* = \|\cdot\|_p$ with $p \in (1,\infty)$, and $\boldsymbol{a}_i(\boldsymbol{x}) = \boldsymbol{x}$, $b_i(\boldsymbol{x}) = b_i$ for each $i \in [I]$. Particularly, if we impose the additional assumption that the empirical samples are nonnegative, i.e., $\widehat{\boldsymbol{\xi}}^j \ge \boldsymbol{0}$ for each $j \in [N]$, in fact, the easily computable upper bound $\bar{\eta}_{i,j}(j'|\theta)$ is exact.

PROPOSITION 3. *Suppose $\boldsymbol{a}_i(\boldsymbol{x}) = \boldsymbol{x}$, $b_i(\boldsymbol{x}) = b_i$ for each $i \in [I]$, the dual norm $\|\cdot\|_* = \|\cdot\|_p$ with $p \in (1,\infty)$, and set $\mathcal{X} = [0,1]^n$. Then*

*(i) The upper bound of Problem (8) can be efficiently computable;*

*(ii) When the empirical samples are nonnegative $\widehat{\boldsymbol{\xi}}^j \ge \boldsymbol{0}$ for all $j \in [N]$, the upper bound $\bar{\eta}_{i,j}(j'|\theta)$ is exact.*

*Proof.* See Appendix A.4. □

For simplicity, we consider the set $\mathcal{X}=[0,1]^n$ in Proposition 3. It is important to note that there are other possible choices for the set $\mathcal{X}$, and these alternatives can be investigated through analogous proofs, similar to those outlined in the proofs of Step 1 and Step 2 in Proposition 3. The summarized outcomes are as follows.

COROLLARY 5. *Suppose $\boldsymbol{a}_i(\boldsymbol{x})=\boldsymbol{x}$, $b_i(\boldsymbol{x})=b_i$ for each $i\in[I]$, the dual norm $\|\cdot\|_*=\|\cdot\|_p$ with $p\in(1,\infty)$. The upper bound of Problem (8) can be efficiently computable if*
*(i) $\mathcal{X}=\{\boldsymbol{x}:\sum_{k\in[n]}x_k\le\tau, 0\le x_k\le 1,\forall k\in[n]\}$; or*
*(ii) $\mathcal{X}=\{\boldsymbol{x}:\sum_{k\in[n]}2^{k-1}x_k\le\tau, 0\le x_k\le 1,\forall k\in[n]\}$.*

When the dual norm $\|\cdot\|_*=\|\cdot\|_1$, Problem (8) can be solved more efficiently using sorting without the need to resort to the dual formulation or employ the golden search method in Proposition 3 (see the details in Appendix A.4).

PROPOSITION 4. *Suppose $\boldsymbol{a}_i(\boldsymbol{x})=\boldsymbol{x}$, $b_i(\boldsymbol{x})=b_i$ for each $i\in[I]$, the dual norm $\|\cdot\|_*=\|\cdot\|_1$, and set $\mathcal{X}=[0,1]^n$. Problem (8) can be efficiently computable.*

*Proof.* See Appendix A.5. □

### 3.3. Implementation of the Variable Fixing Procedure

Similar to Algorithm 1, a step-by-step procedure for the variable fixing process aimed at solving the DRCCP (6) is shown in Algorithm 2. It is important to note that the findings discussed in this section have been put into the numerical implementation in Section 7.

**Algorithm 2** Variable Fixing Procedure for DRCCP (6)

1: Pre-compute an upper bound $v_\infty^U$ for DRCCP (6)
2: Fix the variables individually for each scenario based on Corollary 4
3: When $p\in\{1,\infty\}$, apply the optimality cuts from Step 2 and find the optimality cuts based on Corollary 3 by comparing the restricted lower bound $\widehat{v}_\infty(\mathcal{S}_0,\mathcal{S}_1)$ of the DRCCP (6) and upper bound $v_\infty^U$
4: Strengthen big-M coefficients in DRCCP (6)

## 4. Variable Fixing in DRCCPs under Type 1−Wasserstein Ambiguity Set

In this section, we discuss the variable fixing under type 1−Wasserstein ambiguity set. To begin with, we first introduce the notions of value-at-risk (VaR) and conditional value-at-risk (CVaR). Given a random variable $\tilde{X}$, let $\widehat{\mathbb{P}}$ and $F_{\widehat{\mathbb{P}}}(\cdot)$ be its probability distribution and cumulative distribution function, respectively. For a given risk level $\varepsilon\in(0,1)$, $(1-\varepsilon)$ VaR of $\tilde{X}$ is

$$\widehat{\mathbb{P}}\text{-VaR}_{1-\varepsilon}(\tilde{X}):=\min_s\left\{s:F_{\widehat{\mathbb{P}}}(s)\ge 1-\varepsilon\right\},$$

and the corresponding CVaR is defined as

$$\widehat{\mathbb{P}}\text{-CVaR}_{1-\varepsilon}(\tilde{X}):=\min_\beta\left\{\beta+\frac{1}{\varepsilon}\mathbb{E}_{\widehat{\mathbb{P}}}[\tilde{X}-\beta]_+\right\}.$$

In this section, we align with the assumptions in recent DRCCP literature such as Xie (2021), Chen et al. (2022). Specifically, we adopt the assumption that $\|\boldsymbol{a}_i(\boldsymbol{x})\|_*=\|\boldsymbol{a}_1(\boldsymbol{x})\|_*$ holds for all $i\in[I]$. According to the equivalent reformulations in Xie (2021), Chen et al. (2022), the DRCCP (1) under type 1−Wasserstein ambiguity set can be equivalently reformulated as:

$$v_1^*=\min_{\substack{\boldsymbol{x}\in\mathcal{X},\lambda,\gamma,\\ \boldsymbol{s},\boldsymbol{y},\boldsymbol{z}}}\quad \boldsymbol{c}^\top\boldsymbol{x}, \tag{10a}$$

$$\text{s.t.} \quad \theta\lambda - \varepsilon\gamma \le \frac{1}{N}\sum_{j\in[N]} y_j, \tag{10b}$$

$$y_j + \gamma \le s_j, \forall j \in [N], \tag{10c}$$

$$s_j \le b_i(\boldsymbol{x}) - \boldsymbol{a}_i(\boldsymbol{x})^\top \widehat{\boldsymbol{\xi}}^j + M_{i,j,1}(1-z_j), \forall i \in [I], j \in [N], \tag{10d}$$

$$s_j \le M_{i,j,2} z_j, \forall i \in [I], j \in [N], \tag{10e}$$

$$\|\boldsymbol{a}_1(\boldsymbol{x})\|_* \le \lambda, \forall i \in [I], \tag{10f}$$

$$\sum_{j\in[N]} z_j \ge N - \lfloor N\varepsilon \rfloor + \mathbb{I}\{\theta > 0\}, \tag{10g}$$

$$\lambda > 0, \gamma \ge 0, s_j \ge 0, \forall j \in [N], y_j \le 0, \forall j \in [N], \boldsymbol{z} \in \{0,1\}^N, \tag{10h}$$

where for each $i \in [I], j \in [N]$, we define

$$M_{i,j,1} \ge \max_{\boldsymbol{x}\in\mathcal{X}} \left\{ \boldsymbol{a}_i(\boldsymbol{x})^\top \widehat{\boldsymbol{\xi}}^j - b_i(\boldsymbol{x}) \right\}, M_{i,j,2} \ge \max_{\boldsymbol{x}\in\mathcal{X}} \left\{ b_i(\boldsymbol{x}) - \boldsymbol{a}_i(\boldsymbol{x})^\top \widehat{\boldsymbol{\xi}}^j \right\}.$$

### 4.1. A Big-M Free Formulation and Its Corresponding Variable Fixing

Inspired by the "big-M free" formulations introduced in Ahmed et al. (2017) using disjunctive programming, we present an equivalent reformulation for the DRCCP (10) under the conditions that the compact set $\mathcal{X} \subseteq \{\boldsymbol{x} : \boldsymbol{D}^\top \boldsymbol{x} \le \boldsymbol{d}\}$ and $\|\boldsymbol{a}_i(\boldsymbol{x})\|_* = \|\boldsymbol{a}_1(\boldsymbol{x})\|_*$ for each $i \in [I]$. This reformulation takes the form:

$$v_1^* = \min_{\substack{\boldsymbol{x},\boldsymbol{u},\boldsymbol{w},\boldsymbol{z},\gamma,\\ \boldsymbol{s},\boldsymbol{y},\gamma,\widehat{y}}} \quad \widehat{y}, \tag{11a}$$

$$\text{s.t.} \quad \widehat{y} \ge \boldsymbol{c}^\top \boldsymbol{u}^j + \ell_1(1-z_j), \widehat{y} \ge \boldsymbol{c}^\top \boldsymbol{w}^j + \ell_1 z_j, \forall j \in [N], \tag{11b}$$

$$\theta \|\boldsymbol{a}_1(\boldsymbol{x})\|_* - \varepsilon\gamma \le \frac{1}{N}\sum_{j\in[N]} y_j, \tag{11c}$$

$$\boldsymbol{D}^\top \boldsymbol{u}^j \le \boldsymbol{d} z_j, \forall j \in [N], \boldsymbol{D}^\top \boldsymbol{w}^j \le \boldsymbol{d}(1-z_j), \forall j \in [N], \tag{11d}$$

$$\boldsymbol{x} = \boldsymbol{u}^j + \boldsymbol{w}^j, \forall j \in [N], \tag{11e}$$

$$y_j + \gamma \le s_j, \forall j \in [N], z_j \boldsymbol{a}_i\left(\boldsymbol{u}^j/z_j\right)^\top \widehat{\boldsymbol{\xi}}^j + s_j \le z_j b_i(\boldsymbol{u}^j/z_j), \forall i \in [I], j \in [N], \tag{11f}$$

$$\sum_{j\in[N]} z_j \ge N - \lfloor N\varepsilon \rfloor + \mathbb{I}\{\theta > 0\}, \tag{11g}$$

$$\boldsymbol{z} \in \{0,1\}^N, \gamma \ge 0, \boldsymbol{s} \ge \boldsymbol{0}, \boldsymbol{y} \le \boldsymbol{0}, \tag{11h}$$

where $\ell_1$ is a lower bound of DRCCP (10), e.g., one can use the quantile bound (see, e.g., Song et al. 2014) as a valid lower bound. The validity of formulation (11) is provided through the following theorem.

THEOREM 3. *Suppose that set $\mathcal{X} \subseteq \{\boldsymbol{x} : \boldsymbol{D}^\top \boldsymbol{x} \le \boldsymbol{d}\}$ and $\|\boldsymbol{a}_i(\boldsymbol{x})\|_* = \|\boldsymbol{a}_1(\boldsymbol{x})\|_*$ for all $i \in [I]$. Then the DRCCP (10) and DRCCP (11) are equivalent.*

*Proof.* According to the reformulation in corollary 1 of Xie (2021), we can rewrite DRCCP (1) under type 1−Wasserstein ambiguity set as:

$$v_1^* = \min_{\boldsymbol{x}\in\mathcal{X}} \left\{ \boldsymbol{c}^\top \boldsymbol{x} : \begin{array}{l} \theta\varepsilon^{-1} \|\boldsymbol{a}_1(\boldsymbol{x})\|_* + \widehat{\mathbb{P}}\text{-CVaR}_{1-\varepsilon}\left[ -\min_{i\in[I]} \left( b_i(\boldsymbol{x}) - \boldsymbol{a}_i(\boldsymbol{x})^\top \widehat{\boldsymbol{\xi}} \right)_+ \right] \le 0, \\ \widehat{\mathbb{P}}\left\{ \boldsymbol{a}_i(\boldsymbol{x})^\top \widehat{\boldsymbol{\xi}} - b_i(\boldsymbol{x}) \le 0, \forall i \in [I] \right\} \ge 1-\varepsilon \end{array} \right\}.$$

By the dual representation of $\text{CVaR}_{1-\varepsilon}(\cdot)$ (see, e.g., Shapiro and Ahmed 2004), we have

$$v_1^* = \min_{\boldsymbol{x}\in\mathcal{X}} \left\{ \boldsymbol{c}^\top \boldsymbol{x} : \begin{array}{l} \theta\varepsilon^{-1} \|\boldsymbol{a}_1(\boldsymbol{x})\|_* + \max\limits_{\substack{0\le p_j\le \frac{1}{N\varepsilon}, \forall j\in[N], \\ \sum_{j\in[N]} p_j = 1}} \left[ \sum_{j\in[N]} p_j \left[ -\min_{i\in[I]} \left( b_i(\boldsymbol{x}) - \boldsymbol{a}_i(\boldsymbol{x})^\top \widehat{\boldsymbol{\xi}}^j \right)_+ \right] \right] \le 0, \\ \widehat{\mathbb{P}} \left\{ \boldsymbol{a}_i(\boldsymbol{x})^\top \widehat{\boldsymbol{\xi}} - b_i(\boldsymbol{x}) \le 0, \forall i \in [I] \right\} \ge 1-\varepsilon \end{array} \right\}.$$

Introducing a binary variable $z_j$ for each $j \in [N]$ to denote the $(\cdot)_+$ function, we have

$$v_1^* = \min_{\boldsymbol{x}\in\mathcal{X}, \boldsymbol{z}\in\{0,1\}^N} \left\{ \boldsymbol{c}^\top \boldsymbol{x} : \begin{array}{l} \theta\varepsilon^{-1} \|\boldsymbol{a}_1(\boldsymbol{x})\|_* + \max\limits_{\substack{0\le p_j\le \frac{1}{N\varepsilon}, \forall j\in[N], \\ \sum_{j\in[N]} p_j = 1}} \left[ \sum_{j\in[N]} p_j \left[ -\min_{i\in[I]} \left( b_i(\boldsymbol{x}) - \boldsymbol{a}_i(\boldsymbol{x})^\top \widehat{\boldsymbol{\xi}}^j \right) z_j \right] \right] \le 0, \\ \widehat{\mathbb{P}} \left\{ \boldsymbol{a}_i(\boldsymbol{x})^\top \widehat{\boldsymbol{\xi}} - b_i(\boldsymbol{x}) \le 0, \forall i \in [I] \right\} \ge 1-\varepsilon \end{array} \right\}.$$

In DRCCP (10), we know that $\sum_{j\in[N]} z_j$ must be at least $N - \lfloor N\varepsilon \rfloor + 1$ when $\theta > 0$. On the other hand, in its RCCP counterpart (i.e., $\theta = 0$ in DRCCP (10)), $\sum_{j\in[N]} z_j$ must be at least $N - \lfloor N\varepsilon \rfloor$. Therefore, we can incorporate these conditions:

$$v_1^* = \min_{\boldsymbol{x}\in\mathcal{X}, \boldsymbol{z}} \left\{ \boldsymbol{c}^\top \boldsymbol{x} : \begin{array}{l} \theta\varepsilon^{-1} \|\boldsymbol{a}_1(\boldsymbol{x})\|_* + \max\limits_{\substack{0\le p_j\le \frac{1}{N\varepsilon}, \forall j\in[N], \\ \sum_{j\in[N]} p_j = 1}} \left[ \sum_{j\in[N]} p_j \left[ \max_{i\in[I]} \left( \boldsymbol{a}_i(\boldsymbol{x})^\top \widehat{\boldsymbol{\xi}}^j z_j - b_i(\boldsymbol{x}) z_j \right) \right] \right] \le 0, \\ \sum_{j\in[N]} z_j \ge N - \lfloor N\varepsilon \rfloor + \mathbb{I}\{\theta > 0\}, \boldsymbol{z} \in \{0,1\}^N \end{array} \right\}.$$

Dualizing the inner maximum of the first constraint, we have

$$v_1^* = \min_{\boldsymbol{x}\in\mathcal{X}, \boldsymbol{z}} \left\{ \boldsymbol{c}^\top \boldsymbol{x} : \begin{array}{l} \theta\varepsilon^{-1} \|\boldsymbol{a}_1(\boldsymbol{x})\|_* + \min\limits_{\beta\le 0} \left[ \beta + \dfrac{1}{N\varepsilon} \sum_{j\in[N]} \left[ \max_{i\in[I]} \left( \boldsymbol{a}_i(\boldsymbol{x})^\top \widehat{\boldsymbol{\xi}}^j z_j - b_i(\boldsymbol{x}) z_j \right) - \beta \right]_+ \right] \le 0, \\ \sum_{j\in[N]} z_j \ge N - \lfloor N\varepsilon \rfloor + \mathbb{I}\{\theta > 0\}, \boldsymbol{z} \in \{0,1\}^N \end{array} \right\}.$$

Replacing the minimum operator over $\beta \le 0$ with the existence of $\beta$, we have

$$v_1^* = \min_{\boldsymbol{x}\in\mathcal{X}, \beta\le 0, \boldsymbol{z}} \left\{ \boldsymbol{c}^\top \boldsymbol{x} : \begin{array}{l} \theta \|\boldsymbol{a}_1(\boldsymbol{x})\|_* + \varepsilon\beta + \dfrac{1}{N} \sum_{j\in[N]} \left[ \max_{i\in[I]} \left( \boldsymbol{a}_i(\boldsymbol{x})^\top \widehat{\boldsymbol{\xi}}^j z_j - b_i(\boldsymbol{x}) z_j \right) - \beta \right]_+ \le 0, \\ \sum_{j\in[N]} z_j \ge N - \lfloor N\varepsilon \rfloor + \mathbb{I}\{\theta > 0\}, \boldsymbol{z} \in \{0,1\}^N \end{array} \right\}.$$

Introducing slack variables $\boldsymbol{\phi}$ and $\boldsymbol{\psi}$, we have

$$v_1^* = \min_{\boldsymbol{x}\in\mathcal{X}, \beta\le 0, \boldsymbol{z}, \boldsymbol{\phi}\ge \boldsymbol{0}, \boldsymbol{\psi}\le \boldsymbol{0}} \left\{ \boldsymbol{c}^\top \boldsymbol{x} : \begin{array}{l} \theta \|\boldsymbol{a}_1(\boldsymbol{x})\|_* + \varepsilon\beta + \dfrac{1}{N} \sum_{j\in[N]} \phi_j \le 0, \\ \phi_j \ge \psi_j - \beta, \forall j \in [N], \\ \psi_j \ge \boldsymbol{a}_i(\boldsymbol{x})^\top \widehat{\boldsymbol{\xi}}^j z_j - b_i(\boldsymbol{x}) z_j, \forall i \in [I], j \in [N], \\ \sum_{j\in[N]} z_j \ge N - \lfloor N\varepsilon \rfloor + \mathbb{I}\{\theta > 0\}, \boldsymbol{z} \in \{0,1\}^N \end{array} \right\}.$$

However, the reformulation above is bilinear due to terms $\{\boldsymbol{a}_i(\boldsymbol{x})^\top \widehat{\boldsymbol{\xi}}^j z_j\}_{i\in[I],j\in[N]}$, $\{b_i(\boldsymbol{x})z_j\}_{i\in[I],j\in[N]}$. To address this, we use the extended formulation from proposition 9 in Ahmed et al. (2017). With the presumption that set $\mathcal{X}\subseteq\{\boldsymbol{x}:\boldsymbol{D}^\top\boldsymbol{x}\le\boldsymbol{d}\}$, we obtain the following equivalent reformulation:

$$v_1^* = \min_{\substack{\boldsymbol{x},\boldsymbol{u},\boldsymbol{w},\boldsymbol{z},\\ \beta\le 0,\boldsymbol{\phi}\ge\boldsymbol{0},\boldsymbol{\psi}\le\boldsymbol{0},\widehat{y}}} \left\{ \widehat{y}: \begin{array}{l} \widehat{y}\ge \boldsymbol{c}^\top\boldsymbol{u}^j+\ell_1(1-z_j), \widehat{y}\ge \boldsymbol{c}^\top\boldsymbol{w}^j+\ell_1 z_j, \forall j\in[N],\\ \theta\left\|\boldsymbol{a}_1(\boldsymbol{x})\right\|_* +\varepsilon\beta+\dfrac{1}{N}\displaystyle\sum_{j\in[N]}\phi_j\le 0,\\ \phi_j\ge\psi_j-\beta,\forall j\in[N], \psi_j\ge z_j\boldsymbol{a}_i\left(\boldsymbol{u}^j/z_j\right)^\top\widehat{\boldsymbol{\xi}}^j - z_j b_i(\boldsymbol{u}^j/z_j), \forall i\in[I], j\in[N],\\ \boldsymbol{D}^\top\boldsymbol{u}^j\le\boldsymbol{d}z_j,\forall j\in[N], \boldsymbol{D}^\top\boldsymbol{w}^j\le\boldsymbol{d}(1-z_j),\forall j\in[N],\\ \boldsymbol{x}=\boldsymbol{u}^j+\boldsymbol{w}^j,\forall j\in[N],\\ \displaystyle\sum_{j\in[N]} z_j\ge N-\lfloor N\varepsilon\rfloor+\mathbb{I}\left\{\theta>0\right\}, \boldsymbol{z}\in\{0,1\}^N \end{array} \right\},$$

where $\ell_1$ is a lower bound of DRCCP (10). Let $\gamma=-\beta$, $\boldsymbol{y}=-\boldsymbol{\phi}$, and $\boldsymbol{s}=-\boldsymbol{\psi}$, we have the desired result in (11). This completes the proof. ☐

To solve DRCCP (11) to optimality, we employ variable fixing techniques to generate optimality cuts. To facilitate this process, we first propose a computationally efficient lower bound for DRCCP (11). In literature, Ahmed et al. (2017) introduced a nonlinear programming formulation to calculate a lower bound for the RCCP. Further details can be found in section 5.2 of Ahmed et al. (2017). We extend this approach to suit DRCCP (11):

$$\ell_1^{NLP} = \min_{\substack{\boldsymbol{x},\boldsymbol{u},\boldsymbol{w},\boldsymbol{z},\\ \gamma\ge 0,\boldsymbol{s}\ge\boldsymbol{0},\boldsymbol{y}\le\boldsymbol{0},\widehat{y}}} \left\{\widehat{y}: \begin{array}{l} \boldsymbol{c}^\top\boldsymbol{u}^j\le\widehat{y}z_j, \boldsymbol{c}^\top\boldsymbol{w}^j\le\widehat{y}(1-z_j),\forall j\in[N],\\ \text{(11c)-(11g)}, \boldsymbol{z}\in[0,1]^N \end{array}\right\}. \tag{12}$$

According to proposition 7 and proposition 8 in Ahmed et al. (2017), one can use a linear programming based approach for solving $\ell_1^{NLP}$ (12) as summarized below.

Corollary 6. *Let $\ell_1^{NLP}=\ell_1^*$, then $v_1^L(\ell_1^*)=\ell_1^*$, where*

$$v_1^L(\ell_1) = \min_{\substack{\boldsymbol{x},\boldsymbol{u},\boldsymbol{w},\boldsymbol{z},\\ \gamma\ge 0,\boldsymbol{s}\ge\boldsymbol{0},\boldsymbol{y}\le\boldsymbol{0},\widehat{y}}} \left\{y: \begin{array}{l} \widehat{y}\ge\boldsymbol{c}^\top\boldsymbol{u}^j+\ell_1(1-z_j), \widehat{y}\ge\boldsymbol{c}^\top\boldsymbol{w}^j+\ell_1 z_j,\forall j\in[N],\\ \text{(11c)-(11g)}, \boldsymbol{z}\in[0,1]^N \end{array}\right\}. \tag{13}$$

It is important to recognize that an iterative approach can be used to solve (13) until the convergence. Following the convention in Ahmed et al. (2017), we denote the optimal value of Problem (12) as the dual bound for DRCCP (10).

We consider the restricted Problem (13) to generate optimality cuts, where we replace $\ell_1$ with $v_1^U$ (i.e., the upper bound of DRCCP (10)) and add the constraints $\sum_{j\in\mathcal{S}_1} z_j=|\mathcal{S}_1|$ and $\sum_{j\in\mathcal{S}_0} z_j=0$. For the restricted problem, if the optimal value is no less than $v_1^U$, we can successfully identify optimality cuts. The result is shown below.

Corollary 7. *Suppose that set $\mathcal{X}\subseteq\{\boldsymbol{x}:\boldsymbol{D}^\top\boldsymbol{x}\le\boldsymbol{d}\}$, $\|\boldsymbol{a}_i(\boldsymbol{x})\|_*=\|\boldsymbol{a}_1(\boldsymbol{x})\|_*$ for all $i\in[I]$, and $v_1^U$ denotes an upper bound of DRCCP (11). For any two disjoint sets $\mathcal{S}_0,\mathcal{S}_1\subseteq[N]$ with $|\mathcal{S}_0|\le\lfloor N\varepsilon\rfloor$, one can solve the following problem:*

$$\widehat{v}_1(\mathcal{S}_0,\mathcal{S}_1) = \min_{\substack{\boldsymbol{x},\boldsymbol{u},\boldsymbol{w},\boldsymbol{z}\in[0,1]^N,\\ \gamma\ge 0,\boldsymbol{s}\ge\boldsymbol{0},\boldsymbol{y}\le\boldsymbol{0},\widehat{y}}} \left\{\widehat{y}: \begin{array}{l} \widehat{y}\ge\boldsymbol{c}^\top\boldsymbol{u}^j+v_1^U(1-z_j), \widehat{y}\ge\boldsymbol{c}^\top\boldsymbol{w}^j+v_1^U z_j,\forall j\in[N]\setminus(\mathcal{S}_1\cup\mathcal{S}_0),\\ \text{(11c)-(11g)}, \displaystyle\sum_{j\in\mathcal{S}_0} z_j=0, \sum_{j\in\mathcal{S}_1} z_j=|\mathcal{S}_1| \end{array}\right\}. \tag{14}$$

*If $\widehat{v}_1(\mathcal{S}_0,\mathcal{S}_1)\ge v_1^U$, then the following inequality is valid for any optimal solution of DRCCP (11):*

$$\sum_{j\in\mathcal{S}_0} z_j+\sum_{j\in\mathcal{S}_1}(1-z_j)\ge 1,$$

*with $\sum_{j\in[N]} z_j\ge N-\lfloor N\varepsilon\rfloor+\mathbb{I}\{\theta>0\}, \boldsymbol{z}\in\{0,1\}^N$.*

*Proof.* Note that if we let $\ell_1 = v_1^U$ in $v^L(\ell_1)$ (13), we have $v_1^L(v_1^U) < v_1^U$. Therefore, if we have $\widehat{v}_1(\mathcal{S}_0, \mathcal{S}_1) \geq v_1^U$, following the similar proof as Theorem 1, we can obtain an optimality cut. $\square$

Similarly, the variable fixing technique outlined in Corollary 2 can be easily extended to DRCCP (10).

Corollary 8. *Let* $\widehat{\eta}_j = \min_{\boldsymbol{x}\in\mathcal{X}}\{\boldsymbol{c}^\top\boldsymbol{x} : \boldsymbol{a}_i(\boldsymbol{x})^\top\widehat{\boldsymbol{\xi}}^j \leq b_i(\boldsymbol{x}), \forall i \in [I]\}$ *for each* $j \in [N]$*. If for some* $j \in [N]$*, the value* $\widehat{\eta}_j > v_1^U$*, then we must have* $z_j = 0$*.*

### 4.2. VaR Outer Approximation

The VaR outer approximation (see, e.g., theorem 3 in Xie 2021) can provide a different lower bound for DRCCP (10), which admits the following formulation:

$$v_1^{\mathrm{VaR}} = \min_{\boldsymbol{x}\in\mathcal{X}, \widehat{\boldsymbol{z}}} \quad \boldsymbol{c}^\top\boldsymbol{x}, \tag{15a}$$

$$\text{s.t.} \quad \theta\varepsilon^{-1}\left\|\boldsymbol{a}_i(\boldsymbol{x})\right\|_* + \boldsymbol{a}_i(\boldsymbol{x})^\top\widehat{\boldsymbol{\xi}}^j \leq b_i(\boldsymbol{x}) + M_{i,j}^{\mathrm{VaR},1}(1-\widehat{z}_j), \forall i \in [I], j \in [N], \tag{15b}$$

$$\sum_{j\in[N]} \widehat{z}_j \geq N - \lfloor N\varepsilon \rfloor, \widehat{\boldsymbol{z}} \in \{0,1\}^N, \tag{15c}$$

where for each $i \in [I], j \in [N]$, we have

$$M_{i,j}^{\mathrm{VaR},1} \geq \max_{\boldsymbol{x}\in\mathcal{X}} \left\{\theta\varepsilon^{-1}\left\|\boldsymbol{a}_i(\boldsymbol{x})\right\|_* + \boldsymbol{a}_i(\boldsymbol{x})^\top\widehat{\boldsymbol{\xi}}^j - b_i(\boldsymbol{x})\right\}.$$

Note that the big-M coefficients $\{M_{i,j}^{\mathrm{VaR},1}\}_{i\in[I],j\in[N]}$ can be strengthened efficiently following the discussions in Section 3.

However, it might be difficult to find the optimal objective value of VaR outer approximation (15). Nonetheless, we can derive the valid inequalities by observing that the feasible region of DRCCP (10) is included in that of the VaR outer approximation (15).

Proposition 5. *Suppose the Wasserstein radius* $\theta > 0$*. For a given DRCCP feasible solution* $\widehat{\boldsymbol{x}} \in \mathcal{X}$*, let* $\widehat{\boldsymbol{z}}$ *be a solution of* VaR *outer approximation* (15) *induced by* $\widehat{\boldsymbol{x}}$ *and* $\bar{\boldsymbol{z}}$ *be a solution of DRCCP* (10) *induced by* $\widehat{\boldsymbol{x}}$*, respectively. Then, we have* $\widehat{z}^j \leq \bar{z}^j$ *for each* $j \in [N]$*.*

*Proof.* For a DRCCP feasible solution $\widehat{\boldsymbol{x}} \in \mathcal{X}$, since $\widehat{\boldsymbol{z}}$ is a solution of VaR outer approximation (15) induced by $\widehat{\boldsymbol{x}}$. Suppose $\widehat{z}^j = 1$ for a given $j \in [N]$, then we have

$$\theta\varepsilon^{-1}\left\|\boldsymbol{a}_i(\widehat{\boldsymbol{x}})\right\|_* + \boldsymbol{a}_i(\widehat{\boldsymbol{x}})^\top\widehat{\boldsymbol{\xi}}^j \leq b_i(\widehat{\boldsymbol{x}}), \forall i \in [I],$$

which implies that we cannot have $\boldsymbol{a}_i(\widehat{\boldsymbol{x}})^\top\widehat{\boldsymbol{\xi}}^j > b_i(\bar{\boldsymbol{x}})$ for each $i \in [I]$. Corresponding, since $\bar{\boldsymbol{z}}$ is a solution of DRCCP (10) induced by $\widehat{\boldsymbol{x}}$, the value of $\bar{z}^j$ must be set to 1 and we have $\widehat{z}^j \leq \bar{z}^j$.

On the other hand, suppose $\bar{z}^j = 1$, the inequality always holds $\widehat{z}^j \leq \bar{z}^j$. This completes the proof. $\square$

Proposition 5 can help enhance the VaR lower bound (15). Specifically, we can combine VaR outer approximation (15) with the relaxed DRCCP (10) to get an improved VaR lower bound, that is,

$$\bar{v}_1^L = \min_{\substack{\boldsymbol{x}\in\mathcal{X},\lambda,\gamma,\\ \boldsymbol{s},\boldsymbol{y},\boldsymbol{z}\in[0,1]^N, \widehat{\boldsymbol{z}}\in\{0,1\}^N}} \left\{\boldsymbol{c}^\top\boldsymbol{x} : \begin{array}{l} (10b)\text{-}(10c), (10f)\text{-}(10h), (10d)\text{-}(10e), (15b)\text{-}(15c), \\ \widehat{z}_j \leq z_j, \forall j \in [N] \end{array}\right\}. \tag{16}$$

In the numerical implementation, we choose the maximum between this improved VaR lower bound (16) and the dual bound (13) as the best lower bound of DRCCP (10).

### 4.3. Coefficients Strengthening for DRCCP (10)

Similarly, we can strengthen big-M coefficients in DRCCP (10). Different from the previous method described in Section 2.2, we strengthen the big-M coefficients using the relaxed VaR outer approximation, that is,

$$M_{i,j,1} \geq \max_{\boldsymbol{x}\in\mathcal{X},\widehat{\boldsymbol{z}}} \left\{ \boldsymbol{a}_i(\boldsymbol{x})^\top \widehat{\boldsymbol{\xi}}^j - b_i(\boldsymbol{x}) \colon (15b), \sum_{j\in[N]} \widehat{z}_j \geq N - \lfloor N\varepsilon \rfloor, \widehat{\boldsymbol{z}} \in [0,1]^N \right\}, \tag{17a}$$

$$M_{i,j,2} \geq \max_{\boldsymbol{x}\in\mathcal{X},\widehat{\boldsymbol{z}}} \left\{ b_i(\boldsymbol{x}) - \boldsymbol{a}_i(\boldsymbol{x})^\top \widehat{\boldsymbol{\xi}}^j \colon (15b), \sum_{j\in[N]} \widehat{z}_j \geq N - \lfloor N\varepsilon \rfloor, \widehat{\boldsymbol{z}} \in [0,1]^N \right\}. \tag{17b}$$

We remark that Ho-Nguyen et al. (2023) introduced a big-M coefficient strengthening procedure based on mixing inequalities, our proposed strengthening procedure (18) differs from theirs in the following two aspects: (i) While the strengthening procedure in section 4.2 of Ho-Nguyen et al. (2023) relies on the sorting, we obtain the values of $M_{i,j,1}$, $M_{i,j,2}$ by incorporating the relaxed VaR approximation; (ii) Ho-Nguyen et al. (2023) focused solely on strengthening $M_{i,j,1}$ and did not strengthen $M_{i,j,2}$. In contrast, our approach involves strengthening both $M_{i,j,1}$ and $M_{i,j,2}$. Through our numerical study, we observe that strengthening $M_{i,j,2}$ improves efficiency. We provide detailed comparisons in our numerical study in Section 7 and Appendix C.

It is worth mentioning that the parameter $\gamma$ plays a vital role in solving DRCCP (10), i.e., the high-quality lower and upper bounds of $\gamma$ can speed up the solution process. Hence, to enhance the effectiveness, we employ the following lower and upper bounds for $\gamma$ to strengthen DRCCP (10):

$$\gamma^L = \min_{\boldsymbol{x}\in\mathcal{X},\lambda,\gamma,\boldsymbol{s},\boldsymbol{y},\boldsymbol{z}\in\{0,1\}^N} \left\{ \gamma \colon \boldsymbol{c}^\top \boldsymbol{x} \leq v_1^U, (10b)\text{-}(10c), (10f)\text{-}(10h), (10d)\text{-}(10e) \right\}, \tag{18a}$$

$$\gamma^U = \max_{\boldsymbol{x}\in\mathcal{X},\lambda,\gamma,\boldsymbol{s},\boldsymbol{y},\boldsymbol{z}\in\{0,1\}^N} \left\{ \gamma \colon \boldsymbol{c}^\top \boldsymbol{x} \leq v_1^U, (10b)\text{-}(10c), (10f)\text{-}(10h), (10d)\text{-}(10e) \right\}. \tag{18b}$$

We demonstrate the effectiveness of strengthening $\gamma$ for solving DRCCP (10) in Section 7.

### 4.4. The Implementation of the Variable Fixing Procedure

In summary, the variable fixing procedure to address DRCCP (10) is outlined in Algorithm Algorithm 3. It is important to note that the results presented in this section have been implemented in the numerical study Section 7.

**Algorithm 3** Variable Fixing Procedure for DRCCP (10)

1: Pre-compute lower bound $v_1^L$ and upper bound $v_1^U$ for DRCCP (10)
2: Fix the variables individually for each scenario based on Corollary 8
3: Apply the optimality cuts from Step 2 and find the optimality cuts based on Corollary 7 by comparing the restricted lower bound $\widehat{v}_1(\mathcal{S}_0, \mathcal{S}_1)$ of the DRCCP (10) and upper bound $v_1^U$
4: Solve the improved VaR lower bound (16)
5: Strengthen big-M coefficients (17) and the coefficient $\gamma$ (18)

## 5. Variable Fixing for DRCCPs under Type $q-$Wasserstein Ambiguity Set with $q \in (1, \infty)$

In this section, we discuss the variable fixing under type $q-$Wasserstein ambiguity set with $q \in (1, \infty)$. According to reformulation in Jiang and Xie (2023), under type $q-$Wasserstein ambiguity set with $q \in (1, \infty)$, DRCCP (1) can be simplified to

$$v_q^* = \min_{\boldsymbol{x} \in \mathcal{X}} \left\{ \boldsymbol{c}^\top \boldsymbol{x} : \begin{array}{l} \theta^q \varepsilon^{-1} \left\| \boldsymbol{a}_1(\boldsymbol{x}) \right\|_*^q + \widehat{\mathbb{P}}\text{-CVaR}_{1-\varepsilon} \left[ -\min_{i \in [I]} \left( b_i(\boldsymbol{x}) - \boldsymbol{a}_i(\boldsymbol{x})^\top \widehat{\boldsymbol{\xi}} \right)_+^q \right] \leq 0, \\ \widehat{\mathbb{P}} \left\{ \boldsymbol{a}_i(\boldsymbol{x})^\top \widehat{\boldsymbol{\xi}} - b_i(\boldsymbol{x}) \leq 0, \forall i \in [I] \right\} \geq 1 - \varepsilon \end{array} \right\}. \tag{19}$$

where the corresponding CVaR approximation with $\theta > 0$ is

$$v_q^{\text{CVaR}} = \min_{\boldsymbol{x} \in \mathcal{X}} \left\{ \boldsymbol{c}^\top \boldsymbol{x} : \theta \varepsilon^{-\frac{1}{q}} \left\| \boldsymbol{a}_1(\boldsymbol{x}) \right\|_* + \widehat{\mathbb{P}}\text{-CVaR}_{1-\varepsilon} \left[ \max_{i \in [I]} \left\{ \boldsymbol{a}_i(\boldsymbol{x})^\top \widehat{\boldsymbol{\xi}} - b_i(\boldsymbol{x}) \right\} \right] \leq 0 \right\}. \tag{20}$$

Regarding DRCCP (19) and CVaR approximation (20), there are two primary limitations: (i) DRCCP (19) may not admit a mixed-integer convex programming reformulation (see, e.g., proposition 3 in appendix A.3 of Jiang and Xie 2023); and (ii) CVaR approximation can be quite conservative (see the discussions in Nemirovski and Shapiro 2007, Chen et al. 2023).

### 5.1. A New Inner Approximation of DRCCP (19)

To address the limitations of DRCCP (19) and CVaR approximation (20), we introduce a new conservative approximation that enhances CVaR approximation while still allowing for a MIP reformulation with $\theta > 0$, that is,

$$\widehat{v}_q = \min_{\boldsymbol{x} \in \mathcal{X}} \left\{ \boldsymbol{c}^\top \boldsymbol{x} : \theta \varepsilon^{-\frac{1}{q}} \left\| \boldsymbol{a}_1(\boldsymbol{x}) \right\|_* + \widehat{\mathbb{P}}\text{-CVaR}_{1-\varepsilon} \left[ -\min_{i \in [I]} \left( b_i(\boldsymbol{x}) - \boldsymbol{a}_i(\boldsymbol{x})^\top \widehat{\boldsymbol{\xi}} \right)_+ \right] \leq 0 \right\}. \tag{21}$$

For comparison purposes, we use VaR outer approximation for DRCCP (19). According to theorem 3 in Xie (2021), VaR outer approximation of DRCCP (19) can be written as

$$v_q^{\text{VaR}} = \min_{\boldsymbol{x} \in \mathcal{X}} \left\{ \boldsymbol{c}^\top \boldsymbol{x} : \widehat{\mathbb{P}} \left\{ \tilde{\boldsymbol{\xi}} : \theta \varepsilon^{-\frac{1}{q}} \left\| \boldsymbol{a}_1(\boldsymbol{x}) \right\|_* + \boldsymbol{a}_i(\boldsymbol{x})^\top \widehat{\boldsymbol{\xi}} \leq b_i(\boldsymbol{x}), \forall i \in [I] \right\} \geq 1 - \varepsilon \right\}. \tag{22}$$

THEOREM 4. *When $\left\| \boldsymbol{a}_i(\boldsymbol{x}) \right\|_* = \left\| \boldsymbol{a}_1(\boldsymbol{x}) \right\|_*$ for all $i \in [I]$ and $\theta > 0$, under type $q-$Wasserstein ambiguity set with $q \in (1, \infty)$, the following inequalities hold:*

$$v_q^{\text{VaR}} \leq v_q^* \leq \widehat{v}_q \leq v_q^{\text{CVaR}}.$$

*Proof.* Based on theorem 3 in Xie (2021), we have $v_q^{\text{VaR}} \leq v_q^*$. By the definition of $(\cdot)_+$ function,we conclude that $\widehat{v}_q \leq v_q^{\text{CVaR}}$. Then, it remains to prove $v_q^* \leq \widehat{v}_q$. Since $\text{CVaR}_{1-\varepsilon}[\cdot]$ is a coherent convex measure (see the details in Rockafellar et al. 2000) and according to Jensen inequality (see, e.g., theorem 3.3 in Rudin 1987), for all $\boldsymbol{x} \in \mathcal{X}$, we have

$$-\widehat{\mathbb{P}}\text{-CVaR}_{1-\varepsilon} \left[ -\min_{i \in [I]} \left( b_i(\boldsymbol{x}) - \boldsymbol{a}_i(\boldsymbol{x})^\top \widehat{\boldsymbol{\xi}} \right)_+ \right] \leq \left[ -\widehat{\mathbb{P}}\text{-CVaR}_{1-\varepsilon} \left[ -\min_{i \in [I]} \left( b_i(\boldsymbol{x}) - \boldsymbol{a}_i(\boldsymbol{x})^\top \widehat{\boldsymbol{\xi}} \right)_+ \right]^q \right]^{\frac{1}{q}},$$

which implies that

$$\begin{aligned} & \left\{ \boldsymbol{x} : \theta \varepsilon^{-\frac{1}{q}} \left\| \boldsymbol{a}_1(\boldsymbol{x}) \right\|_* \leq -\widehat{\mathbb{P}}\text{-CVaR}_{1-\varepsilon} \left[ -\min_{i \in [I]} \left( b_i(\boldsymbol{x}) - \boldsymbol{a}_i(\boldsymbol{x})^\top \widehat{\boldsymbol{\xi}} \right)_+ \right] \right\} \\ \subseteq & \left\{ \boldsymbol{x} : \theta \varepsilon^{-\frac{1}{q}} \left\| \boldsymbol{a}_1(\boldsymbol{x}) \right\|_* \leq \left[ -\widehat{\mathbb{P}}\text{-CVaR}_{1-\varepsilon} \left[ -\min_{i \in [I]} \left( b_i(\boldsymbol{x}) - \boldsymbol{a}_i(\boldsymbol{x})^\top \widehat{\boldsymbol{\xi}} \right)_+ \right]^q \right]^{\frac{1}{q}} \right\}. \end{aligned}$$

Thus, we have $v_q^* \leq \widehat{v}_q$. This completes the proof. □

It is important to note that the inner approximation (21) and the recent convex approximation–ALSO-X# from Jiang and Xie (2023) may not be directly comparable. Below is an example.

EXAMPLE 2. Consider a single DRCCP under type 2−Wasserstein ambiguity set with $\theta = 0.8$ and dual norm $\|\cdot\|_* = \|\cdot\|_2$. Assume that the empirical distribution has 4 equiprobable scenarios (i.e., $N = 4$, $\widehat{\mathbb{P}}\{\tilde{\boldsymbol{\xi}} = \widehat{\boldsymbol{\xi}}^i\} = 1/N$), risk parameter $\varepsilon = 1/2$, set $\mathcal{X} = [0,1]^3$, $\boldsymbol{c} = (-2,-3,-2)^\top$, function $\boldsymbol{a}_1(\boldsymbol{x})^\top \tilde{\boldsymbol{\xi}} - b_1(\boldsymbol{x}) = \boldsymbol{x}^\top \tilde{\boldsymbol{\xi}} - 3$, $\widehat{\boldsymbol{\xi}}^1 = (2,2,6)^\top$, $\widehat{\boldsymbol{\xi}}^2 = (3,3,2)^\top$, $\widehat{\boldsymbol{\xi}}^3 = (6,4,8)^\top$, and $\widehat{\boldsymbol{\xi}}^4 = (7,2,2)^\top$. In this example, numerically, we can solve ALSO-X# $v_2^{A\#}$ and the inner approximation $\widehat{v}_2$, where the approximated objective values are $v_2^{A\#} = -1.9531$ and $\widehat{v}_2 = -1.9433$ with error bound $[-10^{-4}, 10^{-4}]$.

◇

In the numerical implementation, we choose the minimum between the inner approximation (21) and ALSO-X# as the upper bound of DRCCP (19).

### 5.2. Variable Fixing for Inner Approximation (21)

According to the discussions from the previous subsection, with the aim of efficiently solving the inner approximation (21), we proceed by introducing its equivalent reformulation. Particularly, when set $\mathcal{X} \subseteq \{\boldsymbol{x} : \boldsymbol{D}^\top \boldsymbol{x} \leq \boldsymbol{d}\}$ and $\|\boldsymbol{a}_i(\boldsymbol{x})\|_* = \|\boldsymbol{a}_1(\boldsymbol{x})\|_*$ for all $i \in [I]$, according to Theorem 3, inner approximation (21) can be written as

$$
\begin{aligned}
\widehat{v}_q = \min_{\boldsymbol{x},\boldsymbol{u},\boldsymbol{w},\boldsymbol{z},\gamma,\boldsymbol{s},\boldsymbol{y},\widehat{y}} \quad & \widehat{y}, && \text{(23a)}\\
\text{s.t.} \quad & \widehat{y} \geq \boldsymbol{c}^\top \boldsymbol{u}^j + \ell_q(1 - z_j), \widehat{y} \geq \boldsymbol{c}^\top \boldsymbol{w}^j + \ell_q z_j, \forall j \in [N], && \text{(23b)}\\
& \theta \varepsilon^{1-\frac{1}{q}} \|\boldsymbol{a}_1(\boldsymbol{x})\|_* - \varepsilon\gamma \leq \frac{1}{N} \sum_{j\in[N]} y_j, && \text{(23c)}\\
& \boldsymbol{D}^\top \boldsymbol{u}^j \leq \boldsymbol{d} z_j, \forall j \in [N], \boldsymbol{D}^\top \boldsymbol{w}^j \leq \boldsymbol{d}(1 - z_j), \forall j \in [N], && \text{(23d)}\\
& \boldsymbol{x} = \boldsymbol{u}^j + \boldsymbol{w}^j, \forall j \in [N], && \text{(23e)}\\
& y_j + \gamma \leq s_j, \forall j \in [N], z_j \boldsymbol{a}_i\left(\boldsymbol{u}^j / z_j\right)^\top \widehat{\boldsymbol{\xi}}^j + s_j \leq z_j b_i(\boldsymbol{u}^j / z_j), \forall i \in [I], j \in [N], && \text{(23f)}\\
& \sum_{j\in[N]} z_j \geq N - \lfloor N\varepsilon \rfloor + \mathbb{I}\{\theta > 0\}, && \text{(23g)}\\
& \boldsymbol{z} \in \{0,1\}^N, \gamma \geq 0, \boldsymbol{s} \geq \boldsymbol{0}, \boldsymbol{y} \leq \boldsymbol{0}, && \text{(23h)}
\end{aligned}
$$

where $\ell_q$ is a lower bound of DRCCP (19), e.g., one can use VaR outer approximation (22) as a lower bound. Building upon Corollary 7, we can leverage the inner approximation formulation (23) to derive optimality cuts.

COROLLARY 9. *Suppose that set $\mathcal{X} \subseteq \{\boldsymbol{x} : \boldsymbol{D}^\top \boldsymbol{x} \leq \boldsymbol{d}\}$ and $\|\boldsymbol{a}_i(\boldsymbol{x})\|_* = \|\boldsymbol{a}_1(\boldsymbol{x})\|_*$ for all $i \in [I]$, for any two disjoint sets $\mathcal{S}_0, \mathcal{S}_1 \subseteq [N]$ with $|\mathcal{S}_0| \leq \lfloor N\varepsilon \rfloor$. Then one can solve the following problem:*

$$
\widehat{v}_q(\mathcal{S}_0, \mathcal{S}_1) = \min_{\substack{\boldsymbol{x},\boldsymbol{u},\boldsymbol{w},\boldsymbol{z}\in[0,1]^N,\\ \gamma\geq 0,\boldsymbol{s}\geq\boldsymbol{0},\boldsymbol{y}\leq\boldsymbol{0},\widehat{y}}} \left\{ y : \begin{array}{l} \widehat{y} \geq \boldsymbol{c}^\top \boldsymbol{u}^j + v_q^U(1 - z_j), \widehat{y} \geq \boldsymbol{c}^\top \boldsymbol{w}^j + v_q^U z_j, \forall j \in [N] \setminus (\mathcal{S}_1 \cup \mathcal{S}_0), \\ \text{(23c)-(23g)}, \\ \sum_{j\in\mathcal{S}_0} z_j = 0, \sum_{j\in\mathcal{S}_1} z_j = |\mathcal{S}_1| \end{array} \right\}. \quad (24)
$$

*If $\widehat{v}_q(\mathcal{S}_0, \mathcal{S}_1) \geq v_q^U$, where $v_q^U$ is an upper bound of DRCCP (19), then the following inequality is valid for any optimal solution of inner approximation (21):*

$$
\sum_{j\in\mathcal{S}_0} z_j + \sum_{j\in\mathcal{S}_1} (1 - z_j) \geq 1,
$$

*with $\sum_{j\in[N]} z_j \geq N - \lfloor N\varepsilon \rfloor + \mathbb{I}\{\theta > 0\}, \boldsymbol{z} \in \{0,1\}^N$.*

It is noteworthy that the only difference between the inner approximation (21) and DRCCP (11) pertains to the coefficient $q$ in the constraint (23c). As such, the discussions closely mirror those presented in Section 4 and are therefore omitted for brevity.

## 6. Theoretical Analysis of Variable Fixing

In this section, we study the theoretical perspective on variable fixing. We examine several specific DRCCPs (5) under the type $\infty-$Wasserstein ambiguity set. Notably, since these DRCCPs degenerate to RCCPs when the Wasserstein radius $\theta$ equals 0, the results in this section also hold for the corresponding RCCPs. More specifically, for every case presented, we conduct an asymptotic analysis to determine the proportion of scenarios whose corresponding binary variables $\boldsymbol{z}$ can be fixed to zero.

### 6.1. Joint DRCCPs with a Continuous Reference Distribution

In this subsection, we consider a joint DRCCP with right-hand uncertainty and a continuous reference distribution. In particular, we assume that set $\mathcal{X} = \mathbb{R}^n$, $I = n$, the uncertainty constraint is $\boldsymbol{a}_i(\boldsymbol{x})^\top \boldsymbol{\xi} - b_i(\boldsymbol{x}) = \xi_i - x_i$, and the random parameter $\tilde{\xi}_i$ is continuous for each $i \in [n]$. That is, we consider the following DRCCP:

$$v_\infty^* = \min_{\boldsymbol{x}\in\mathbb{R}^n} \left\{ \boldsymbol{c}^\top \boldsymbol{x} \colon \inf_{\mathbb{P}\in\mathcal{P}_\infty} \mathbb{P}\left\{ \tilde{\boldsymbol{\xi}} \colon \tilde{\xi}_i \le x_i, \forall i \in [n] \right\} \ge 1-\varepsilon \right\},$$

that is,

$$v_\infty^* = \min_{\boldsymbol{x}\in\mathbb{R}^n} \left\{ \sum_{i\in[n]} x_i : \widehat{\mathbb{P}}\left\{ \widehat{\xi}_i + \theta \le x_i, \forall i \in [n] \right\} \ge 1-\varepsilon \right\}, \tag{25a}$$

Suppose the reference distribution $\widehat{\mathbb{P}}$ is log-concave independent and identically distributed (see, e.g., Prékopa 1973, 1980). Then, an optimal solution of DRCCP (25a) is $x_1^* = x_2^* = \cdots = x_n^*$ and the optimal objective value of DRCCP (25a) is

$$v_\infty^* = \sum_{i\in[n]} x_i^* = n\theta + nF_{\widehat{\xi}}^{-1}\left( [1-\varepsilon]^{\frac{1}{n}} \right),$$

where $F_{\widehat{\xi}}^{-1}(\cdot)$ denotes the inverse cumulative distribution function of $\widehat{\xi}$.

Then, we provide the theoretical analysis of Corollary 4 to demonstrate its efficacy in variable fixing.

PROPOSITION 6. *Suppose that in a joint DRCCP (25a) and the reference distribution $\widehat{\mathbb{P}}$ is independent and identically distributed. Let $\widehat{F}(\cdot)$ denote the cumulative distribution function of $\mathbf{e}^\top \widehat{\boldsymbol{\xi}}$. Then, using the variable fixing technique described in Corollary 4, asymptotically $(1-\widehat{F}(nF_{\widehat{\xi}}^{-1}([1-\varepsilon]^{\frac{1}{n}}))) \times 100\%$ of scenarios can be successfully identified, and their corresponding binary variables can be fixed to zero.*

*Proof.* Let us consider the variable fixing based on Corollary 4. For one particular scenario $\widehat{\boldsymbol{\xi}}$, we fix the variables if

$$\widehat{\eta}_\infty = \min_{\boldsymbol{x}\in\mathbb{R}^n} \left\{ \sum_{i\in[n]} x_i : \widehat{\xi}_i + \theta \le x_i, \forall i \in [n] \right\} > v_\infty^*,$$

which implies that scenario $\widehat{\boldsymbol{\xi}}$ has to be violated if $\mathbf{e}^\top \widehat{\boldsymbol{\xi}} > v_\infty^* - n\theta$. Hence, asymptotically, we fix $\widehat{\mathbb{P}}\{\mathbf{e}^\top \widehat{\boldsymbol{\xi}} > v_\infty^* - n\theta\}$ portion of scenarios to be zero, that is,

$$\widehat{\mathbb{P}}\left\{ \mathbf{e}^\top \widehat{\boldsymbol{\xi}} > v_\infty^* - n\theta \right\} = 1 - \widehat{\mathbb{P}}\left\{ \mathbf{e}^\top \widehat{\boldsymbol{\xi}} \le nF_{\widehat{\xi}}^{-1}\left( [1-\varepsilon]^{\frac{1}{n}} \right) \right\} = 1 - \widehat{F}\left( nF_{\widehat{\xi}}^{-1}\left( [1-\varepsilon]^{\frac{1}{n}} \right) \right).$$

□

We remark that the result in Proposition 6 depends on the dimension of $\boldsymbol{x}$, denoted as $n$, and risk parameter $\varepsilon$. With additional assumptions on the distribution of $\widehat{\xi}$, where the cumulative distribution of $\mathbf{e}^\top\widehat{\boldsymbol{\xi}}$ can be easy to evaluate, we are able to simplify the result in Proposition 6. For example, $\xi_i$ follows a normal distribution or an exponential distribution for each $i\in[n]$.

COROLLARY 10. *Under the same assumptions in Proposition 6, the followings must hold:*

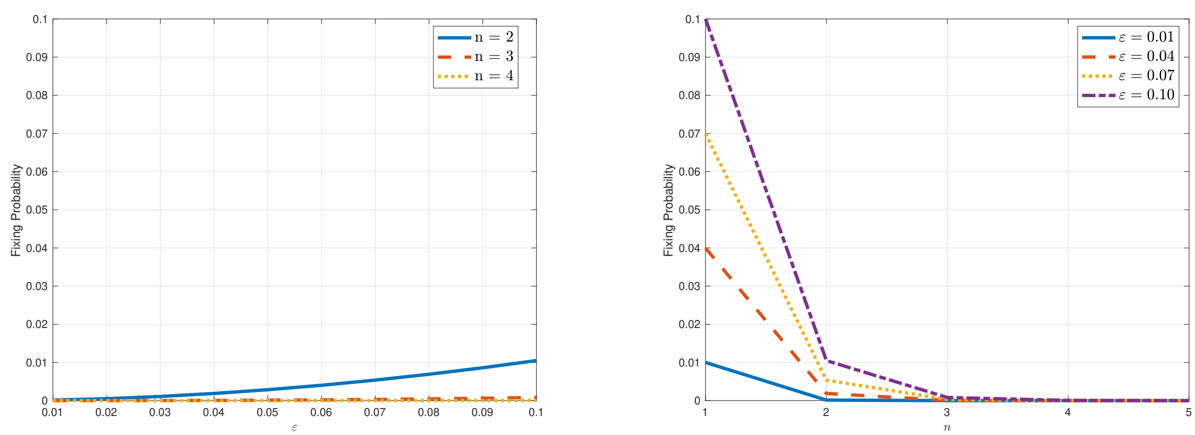
*(i) suppose each $\widehat{\xi}_i$ follows a normal distribution, i.e., $\widehat{\xi}_i\sim\mathcal{N}(\mu,\sigma)$ for all $i\in[n]$. Then, asymptotically $(1-\Phi(\sqrt{n}\Phi^{-1}[(1-\varepsilon)^{\frac{1}{n}}]))\times 100\%$ of scenarios can be successfully identified, and their corresponding binary variables can be fixed to zero; and*

*(ii) suppose each $\widehat{\xi}_i$ follows an exponential distribution, i.e., for all $i\in[n]$, $\widehat{\xi}_i\sim\mathrm{Exp}(\widehat{\lambda})$ with cumulative distribution function $F(\widehat{\xi},\widehat{\lambda})=1-e^{-\widehat{\lambda}\widehat{\xi}}$ when $\widehat{\xi}\geq 0$, 0 otherwise. Then, asymptotically*

$$\left[1-\left[1-(1-\varepsilon)^{\frac{1}{n}}\right]^n\left(\sum_{i=0}^{n-1}\frac{1}{i!}\left(-n\log\left[1-(1-\varepsilon)^{\frac{1}{n}}\right]\right)^i\right)\right]\times 100\%$$

*of scenarios can be successfully identified, and their corresponding binary variables can be fixed to zero.*

To illustrate the result, we plot the curves with some particular choices of $n$ and $\varepsilon$ in case (i) of Corollary 10. When the dimension of $\boldsymbol{x}$ is small, or the risk parameter $\varepsilon$ is not very small, simply applying the variable fixing technique described in Corollary 4 can be quite effective. We observe that when $n$ increases or $\varepsilon$ decreases, the portion of scenarios we can fix decreases.




(a) Comparisons among Different $n$ (b) Comparisons among Different $\varepsilon$

**Figure 2** Portion of Violated Scenarios Identified by Part (i) in Corollary 10.

## 6.2. Binary Single DRCCPs with i.i.d. Bernoulli Random Parameters

Let us consider a single distributional robust chance constrained set covering problem (5), where the set $\mathcal{X}\subseteq\{0,1\}^n$, the affine mappings are $\boldsymbol{a}_1(\boldsymbol{x})=-\boldsymbol{x}$, $b_1(\boldsymbol{x})=b_1\geq 0$, and the binary support $\Xi\subseteq\{0,1\}^n$. That is, we consider the following single DRCCP:

$$v^*_\infty=\min_{\boldsymbol{x}\in\{0,1\}^n}\left\{\boldsymbol{c}^\top\boldsymbol{x}\colon\inf_{\mathbb{P}\in\mathcal{P}_\infty}\mathbb{P}\left\{\boldsymbol{x}^\top\tilde{\boldsymbol{\xi}}\geq b_1\right\}\geq 1-\varepsilon\right\}. \tag{26}$$

We first present an equivalent reformulation of DRCCP (26).

LEMMA 1. *Suppose the binary support $\Xi \subseteq \{0,1\}^n$. Then DRCCP (26) is equivalent to*

$$v_\infty^* = \min_{\boldsymbol{x}\in\{0,1\}^n} \left\{ \boldsymbol{c}^\top \boldsymbol{x} \colon \widehat{\mathbb{P}}\left\{ \boldsymbol{x}^\top \widehat{\boldsymbol{\xi}} \geq b_1 + \lfloor \theta^p \rfloor \right\} \geq 1-\varepsilon \right\}. \tag{27}$$

*Proof.* By the definition of type $\infty-$Wasserstein ambiguity set, we write DRCCP (26) as

$$v_\infty^* = \min_{\boldsymbol{x}\in\{0,1\}^n} \left\{ \boldsymbol{c}^\top \boldsymbol{x} \colon \widehat{\mathbb{P}}\left\{ \min_{\boldsymbol{\xi}\in\{0,1\}^n} \left\{ \boldsymbol{x}^\top \boldsymbol{\xi} \colon \|\boldsymbol{\xi} - \widehat{\boldsymbol{\xi}}\|_p \leq \theta \right\} \geq b_1 \right\} \geq 1-\varepsilon \right\}.$$

Since the support is binary, we can write the norm constraint as $\|\boldsymbol{\xi} - \widehat{\boldsymbol{\xi}}\|_p = (\sum_{i\in[n]} |\xi_i - \widehat{\xi}_i|^p)^{1/p} = (\sum_{i\in[n]} |\xi_i - \widehat{\xi}_i|)^{1/p} \leq \theta$, which can be simplified as $\sum_{i\in[n]} |\xi_i - \widehat{\xi}_i| \leq \theta^p$. For a given binary decision $\boldsymbol{x}$, let us consider the minimization problem: $\min_{\boldsymbol{\xi}\in\{0,1\}^n} \{\sum_{i\in[n]} \xi_i x_i : \sum_{i\in[n]} |\xi_i - \widehat{\xi}_i| \leq \theta^p\}$. Denote $\mathcal{T} = \{k \in [n] : x_k = 1\}$, we have

$$\min_{\boldsymbol{\xi}\in\{0,1\}^n} \left\{ \sum_{i\in\mathcal{T}} \xi_i x_i + \sum_{i\in[n]\setminus\mathcal{T}} \xi_i x_i : \sum_{i\in\mathcal{T}} |\xi_i - \widehat{\xi}_i| \leq \theta^p \right\} = \max\left\{ \boldsymbol{x}^\top \widehat{\boldsymbol{\xi}} - \lfloor \theta^p \rfloor, 0 \right\}.$$

Therefore, we arrive at DRCCP (27). □

Lemma 1 motivates us to study a special DRCCP in favor of variable fixing analysis.

PROPOSITION 7. *Suppose that in a single DRCCP (26) and the reference distribution $\widehat{\mathbb{P}}$ is independent and identically distributed with $\widehat{\xi}_i \sim \text{Binomial}(1, \widehat{p})$ for each $i \in [n]$. Assume $\boldsymbol{c} = \mathbf{e}$. Then, by the variable fixing technique described in Corollary 4, asymptotically $\sum_{i=0}^{b_1+\lfloor\theta^p\rfloor-1} \binom{n}{i} (1-\widehat{p})^{n-i} \widehat{p}^i \times 100\%$ of scenarios can be successfully identified, and their corresponding binary variables can be fixed to zero.*

*Proof.* Since $\widehat{\boldsymbol{\xi}}$ is independent and identically distributed for each $i \in [n]$ and $\boldsymbol{c} = \mathbf{e}$, we have

$$v_\infty^* = \min_{\boldsymbol{x}\in\{0,1\}^n} \left\{ \sum_{i\in[n]} x_i \colon \widehat{\mathbb{P}}\left\{ \sum_{i\in[n]} \widehat{\xi}_i x_i \geq b_1 + \lfloor \theta^p \rfloor \right\} \geq 1-\varepsilon \right\}.$$

For one particular scenario $\widehat{\boldsymbol{\xi}}$, we fix the variables if $\widehat{\eta} = \min_{\boldsymbol{x}} \left\{ \sum_{i\in[n]} x_i \colon \sum_{i\in[n]} \widehat{\xi}_i x_i \geq b_1 + \lfloor \theta^p \rfloor \right\} > v_\infty^*$. Notice that the value of $\widehat{\eta}$ is either $b_1 + \lfloor \theta^p \rfloor$ or $\infty$. Then, if $\sum_{i\in[n]} \widehat{\xi}_i < b_1 + \lfloor \theta^p \rfloor$, we can fix the corresponding scenario, that is, asymptotically

$$\mathbb{P}\left\{ \sum_{i\in[n]} \widehat{\xi}_i < b_1 + \lfloor \theta^p \rfloor \right\} \times 100\% = \sum_{i=0}^{b_1+\lfloor\theta^p\rfloor-1} \binom{n}{i} (1-\widehat{p})^{n-i} \widehat{p}^i \times 100\%$$

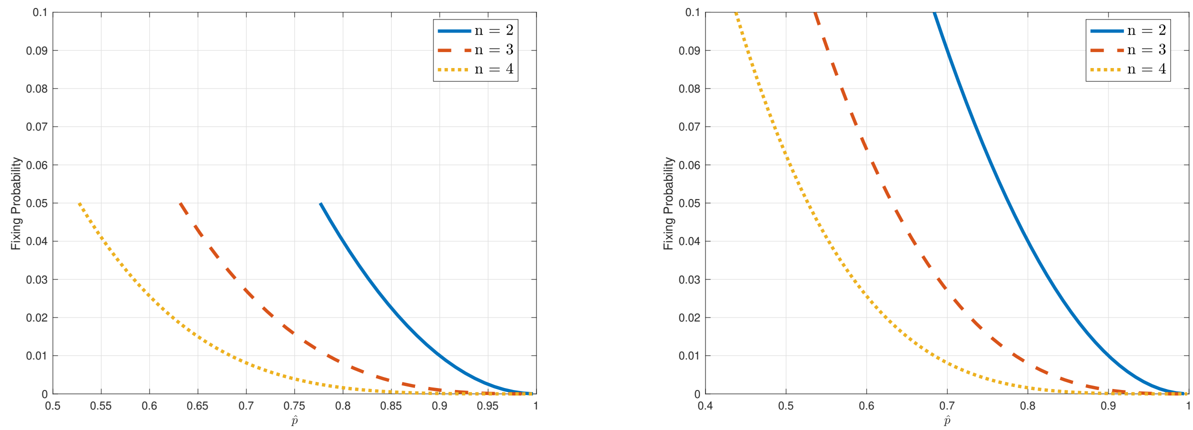
of scenarios can be identified, and their corresponding binary variables can be fixed to zero. □

Surprisingly, Proposition 7 shows that the variable fixing technique presented in Corollary 4 is more or less independent of the risk parameter $\varepsilon$. To illustrate the result, we let $b_1 = 1, \theta \in (0,1)$ in DRCCP (27), and we present the corresponding curves with varying $n$ and $\widehat{p}$ in Figure 3. We observe that when $\widehat{p}$ increases or $n$ increases, the portion of scenarios we can fix decreases.

### 6.3. Single DRCCPs with Elliptical Reference Distributions

In this subsection, we consider a single DRCCP (5) with the elliptical reference distribution. An elliptical distribution $\mathbb{P}_{\mathrm{E}}(\boldsymbol{\mu}, \boldsymbol{\Sigma}, \widehat{g})$ is characterized by three parameters, a location parameter $\boldsymbol{\mu}$, a positive semi-definite matrix $\boldsymbol{\Sigma}$, and a generating function $\widehat{g}$, and its probability density function $\widehat{f}$ can be expressed as:

$$\widehat{f}(\boldsymbol{x}) = \bar{k} \cdot \widehat{g}\left( \frac{1}{2} (\boldsymbol{x} - \boldsymbol{\mu})^\top \boldsymbol{\Sigma}^{-1} (\boldsymbol{x} - \boldsymbol{\mu}) \right),$$




(a) Comparisons among Different $\widehat{p}$ with $\varepsilon = 0.05$ (b) Comparisons among Different $\widehat{p}$ with $\varepsilon = 0.10$

**Figure 3** Portion of Violated Scenarios Identified by Corollary 4 in Proposition 7 with $b_1 = 1, \theta \in (0,1)$.

where $\bar{k}$ is a positive normalization scalar. Specifically, for standard univariate elliptical distribution $\mathbb{P}_{\mathrm{E}}(0, 1, \widehat{g})$, its probability density function is $\varphi(z) = \bar{k}\widehat{g}(z^2/2)$ and the corresponding cumulative distribution function is denoted as $\Phi(\tau) = \int_{-\infty}^{\tau} \bar{k}\widehat{g}(z^2/2)dz$. In DRCCP (5), suppose that the affine mappings are $\boldsymbol{a}_1(\boldsymbol{x}) = \boldsymbol{x}$, $b_1(\boldsymbol{x}) = b_1$, the random parameters $\widehat{\boldsymbol{\xi}}$ follow a joint elliptical distribution with $\widehat{\boldsymbol{\xi}} \sim \mathbb{P}_{\mathrm{E}}(\boldsymbol{\mu}, \boldsymbol{\Sigma}, \widehat{g})$, and the norm defining the Wasserstein distance is the generalized Mahalanobis norm associated with the matrix $\boldsymbol{\Sigma}$, i.e., $\|\boldsymbol{y}\| = \sqrt{\boldsymbol{y}^\top \boldsymbol{\Sigma}^\dagger \boldsymbol{y}}$, for some $\boldsymbol{y} \in \mathbb{R}^n$, where $\boldsymbol{\Sigma}^\dagger$ is the pseudo-inverse. According to the reformulation in proposition 10 of Jiang and Xie (2022), DRCCP (5) can be simplified as

$$v_\infty^* = \min_{\boldsymbol{x} \in \mathcal{X}} \left\{ \boldsymbol{c}^\top \boldsymbol{x} \colon \boldsymbol{\mu}^\top \boldsymbol{x} + \left(\Phi^{-1}(1-\varepsilon) + \theta\right) \sqrt{\boldsymbol{x}^\top \boldsymbol{\Sigma} \boldsymbol{x}} - b_1 \le 0 \right\},$$

which allows us to conduct a theoretical analysis of Corollary 4 to demonstrate its effectiveness in variable fixing.

Proposition 8. *Suppose that in a single DRCCP (5), the reference distribution $\widehat{\mathbb{P}}$ is elliptical with affine mappings $\boldsymbol{a}_1(\boldsymbol{x}) = \boldsymbol{x}$, $b_1(\boldsymbol{x}) = b_1 > 0$, $\boldsymbol{\Sigma} = \boldsymbol{I}$, $\boldsymbol{\mu} = \bar{\mu}\mathbf{e}$, $\boldsymbol{c} = -\mathbf{e}$, set $\mathcal{X} = \mathbb{R}^n$. Then, by the variable fixing technique described in Corollary 4, asymptotically*

$$\widehat{\mathbb{P}} \left\{ \sum_{i \in [n]} \widehat{\xi}_i + \sqrt{\widehat{\gamma}} > \left[ n\bar{\mu} + \sqrt{n} \left[ \theta + \Phi^{-1}(1-\varepsilon) \right] \right] \middle| \widehat{\gamma} \ge 0 \right\} \times 100\%$$

*of scenarios can be successfully identified, and their corresponding binary variables can be fixed to zero, where $\widehat{\gamma} = (\sum_{i \in [n]} \widehat{\xi}_i)^2 - n \sum_{i \in [n]} \widehat{\xi}_i^2 + n\theta^2$.*

*Proof.* From the variable fixing technique described in Corollary 4, for one particular $\widehat{\boldsymbol{\xi}}$, we fix the variables if

$$\widehat{\eta}_\infty = \min_{\boldsymbol{x}} \left\{ -\mathbf{e}^\top \boldsymbol{x} \colon \widehat{\boldsymbol{\xi}}^\top \boldsymbol{x} + \theta \|\boldsymbol{x}\|_2 - b_1 \le 0 \right\} > v_\infty^*. \tag{28a}$$

Let $\alpha$ be its dual variable of the constraint in (28a). Then its dual problem is

$$\max_{\alpha \ge 0} \left\{ -b_1/\alpha \colon \left\| -\alpha \mathbf{e} + \widehat{\boldsymbol{\xi}} \right\|_2 \le \theta \right\}. \tag{28b}$$

Conditioning on $\widehat{\gamma}\geq 0$, i.e., $(\sum_{i\in[n]}\widehat{\xi}_i)^2\geq n[\sum_{i\in[n]}\widehat{\xi}_i^2-\theta^2]$, an optimal solution of (28b) is

$$\alpha_{\max}=\frac{1}{n}\left[\sum_{i\in[n]}\widehat{\xi}_i+\sqrt{\left(\sum_{i\in[n]}\widehat{\xi}_i\right)^2-n\left[\sum_{i\in[n]}\widehat{\xi}_i^2-\theta^2\right]}\right],$$

with the optimal objective value $\widehat{\eta}_\infty=-b_1/\alpha_{\max}$.

Following similar derivations, we have

$$v_\infty^*=\max_{\beta\geq 0}\left\{-\frac{b_1}{\beta}:\ \|-\beta\mathbf{e}+\boldsymbol{\mu}\|_2\leq\theta+\Phi^{-1}(1-\varepsilon)\right\},$$

and its optimal solution is $\beta_{\max}=\bar{\mu}+[\theta+\Phi^{-1}(1-\varepsilon)]/\sqrt{n}$ with the optimal value $v_\infty^*=-b_1/\beta_{\max}$. For $v_\infty^*$, we construct its primal solution as $x_i=b_1/[n\bar{\mu}+\sqrt{n}[\theta+\Phi^{-1}(1-\varepsilon)]]$, which yields the same objective value.

Then $\widehat{\eta}>v_\infty^*$ is equivalent to $\alpha_{\max}>\beta_{\max}$, which implies that

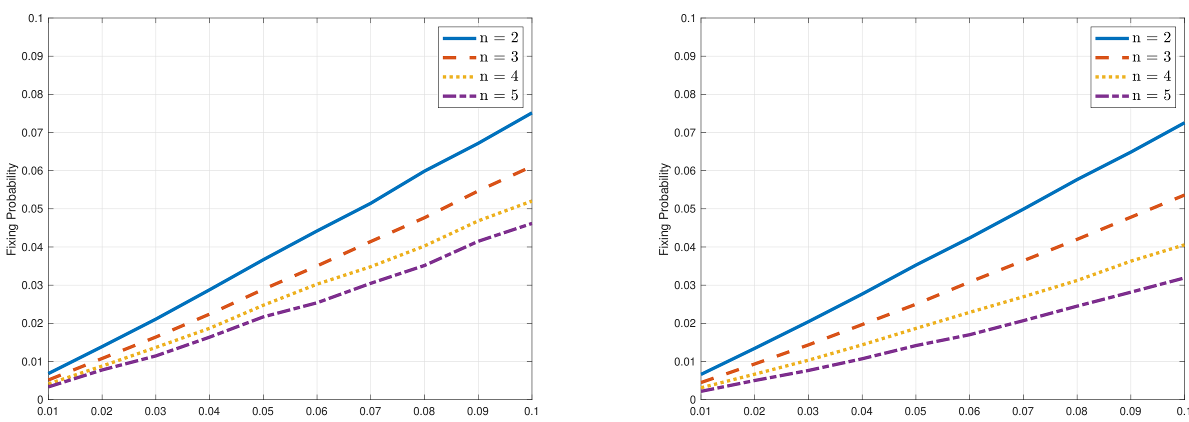
$$\sum_{i\in[n]}\widehat{\xi}_i+\sqrt{\left(\sum_{i\in[n]}\widehat{\xi}_i\right)^2-n\left[\sum_{i\in[n]}\widehat{\xi}_i^2-\theta^2\right]}>n\bar{\mu}+\sqrt{n}\left[\theta+\Phi^{-1}(1-\varepsilon)\right].$$

Asymptotically, we fix

$$\widehat{\mathbb{P}}\left\{\sum_{i\in[n]}\widehat{\xi}_i+\sqrt{\widehat{\gamma}}>\left[n\bar{\mu}+\sqrt{n}\left[\theta+\Phi^{-1}(1-\varepsilon)\right]\right]\,\middle|\,\widehat{\gamma}\geq 0\right\}\times 100\%$$

of the scenario to be zero. This completes the proof. □

We then generate $10^6$ samples to evaluate the fixing probability in Proposition 8. For each instance, we set $\bar{\mu}=-10$, and the outcomes are presented in Figure 4. We observe that when $n$ increases or $\varepsilon$ decreases, the portion of scenarios we can fix decreases.




(a) Comparisons among Different $\varepsilon$ with $\theta=1$ (b) Comparisons among Different $\varepsilon$ with $\theta=2$

**Figure 4** Portion of Violated Scenarios Identified by Corollary 4

## 7. Numerical Study

In this section, we numerically demonstrate the effectiveness of the proposed methods. All the instances in this section are executed in Python 3.9 with calls to solver Gurobi (version 9.5.2 with default settings) on a personal PC with an Apple M1 Pro processor and 16G of memory. We use "UB" and "LB" to denote the best upper bound and the best lower bound found by the big-M model, and use "GAP" to denote its optimality gap as GAP (%) $= (|\text{UB} - \text{LB}|)/(|\text{LB}|) \times 100$. In all our experiments, we set the time limit of each instance to $14,400$ seconds (i.e., 4 hours) with the default optimality gap tolerance of 0.01%. We evaluate proposed methods on two sets of instances, *1-7-1* and *1-7-5* from Song et al. (2014) with set $\mathcal{X} = [0,1]^n$, $n = 50$, risk parameter $\varepsilon \in \{0.10, 0.20\}$, and Wasserstein radius $\theta \in \{0, 0.10, 0.20\}$. For each instance, we solve each method 5 times and report the average performance. We separate our discussions into four cases: an RCCP, a DRCCP under type $\infty-$Wasserstein ambiguity set, a DRCCP under type $1-$Wasserstein ambiguity set, and a DRCCP under type $2-$Wasserstein ambiguity set. Codes of the numerical experiments are available at https://github.com/jnan97/Variable_Fixing.

**Case I. Testing an RCCP.** The RCCP that we test admits the following form:

$$v^* = \min_{\boldsymbol{x} \in [0,1]^n} \left\{ \boldsymbol{c}^\top \boldsymbol{x} \colon \frac{1}{N} \sum_{j \in [N]} \mathbb{I}\left[ \sum_{i \in [n]} \xi_i^j x_i \le b_j \right] \ge 1 - \varepsilon \right\}.$$

To solve each testing instance, our approach comprises the following three steps:

**Step 1.** We use the dual bound (see, e.g., Ahmed et al. 2017) as the outer approximation $v^L$ and ALSO-X# (see, e.g., Jiang and Xie 2023) as the inner approximation $v^U$, respectively. Then, we identify the cuts based on Corollary 2. We use "P" to denote the running time of this step.

**Step 2.** We then strengthen the big-M coefficients according to the discussions in Section 2.2. We use "S" to denote the running time of this step.

**Step 3.** Finally, we execute the big-M method.

We compare the numerical results for the following three methods: (i) Big-M method with fixing & strengthening; (ii) Big-M method with strengthening; (iii) Vanilla big-M method. In the first method, we initialize the solver with the solution ALSO-X#. Additionally, we incorporate the inequalities $\boldsymbol{c}^\top \boldsymbol{x} \ge v^L$ and $\boldsymbol{c}^\top \boldsymbol{x} \le v^U$ into the solver. From Step 1, we identify type $z_j = 0$ cut by Corollary 2. We then record the number of type $z_j = 0$ cuts as "# Cuts (A)" in our numerical results. For the second method, we implement the above Step 2 and Step 3. For the third method, we implement Step 3. The result is displayed in Table 1. Our methods consistently exhibit superior performance compared to other approaches, consistently achieving faster and more stable solutions across all instances. There are total 20 instances reported in Table 1, we also provide its performance profile for all instances in Figure 5. That is, we use the horizontal axis to represent the logarithmic scale of running time and use the vertical axis to represent the number of instances solved to optimality up to that point in time. Our method demonstrates the ability to close the gap with a faster total running time, while other methods perform worse and are unable to close the gap even with a longer total running time. If we only compare the solver's running time, our method performs much better than other methods.

**Case II. Testing a DRCCP under type $\infty-$Wasserstein ambiguity set.** Let us consider the following DRCCP:

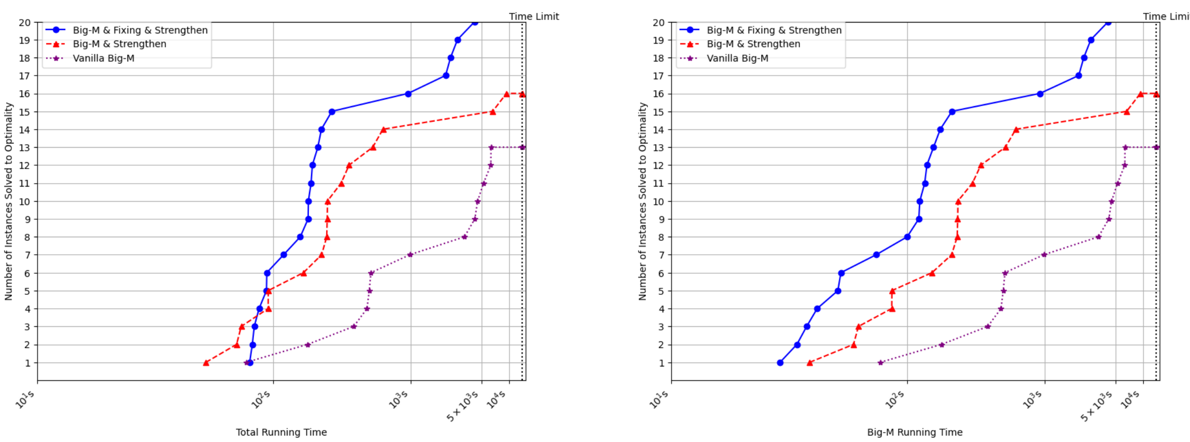
$$v_\infty^* = \min_{\boldsymbol{x} \in [0,1]^n} \left\{ \boldsymbol{c}^\top \boldsymbol{x} \colon \frac{1}{N} \sum_{j \in [N]} \mathbb{I}\left[ \theta \|\boldsymbol{x}\|_2 + \sum_{i \in [n]} \xi_i^j x_i \le b_j \right] \ge 1 - \varepsilon \right\}.$$

For each instance, the following three steps are employed:

**Step 1.** We solve the quantile bound (see, e.g., Song et al. 2014, Ahmed et al. 2017) as the outer approximation $v_\infty^L$ and ALSO-X# (see, e.g., Jiang and Xie 2023) as the inner approximation $v_\infty^U$, respectively. Then, we identify the cuts based on Corollary 4. We use "P" to denote the running time of this step.

**Table 1** Numerical Results of an RCCP with Instances *1-7-1-1000* and *1-7-5-1000* from Song et al. (2014) and $N = 1000$

| Dataset | $\varepsilon$ | Case | Big-M & Fixing & Strengthening | | | | | | Big-M & Strengthening | | | | Vanilla big-M | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | | | GAP | # of Cuts (A) | Time (s) P | Time (s) S | Time (s) Solver | Time (s) Total | GAP | Time (s) S | Time (s) Solver | Time (s) Total | GAP | Time (s) |
| 1-7-1-1000 | 0.1 | 1 | 0.00% | 44 | 31.68 | 11.13 | 30.83 | 73.64 | 0.00% | 11.81 | 6517.51 | 6529.32 | 0.05% | 14400.00 |
| | | 2 | 0.00% | 41 | 35.03 | 11.00 | 145.27 | 191.31 | 0.00% | 12.00 | 468.20 | 480.20 | 0.12% | 14400.00 |
| | | 3 | 0.00% | 47 | 36.45 | 10.79 | 44.33 | 91.57 | 0.00% | 11.68 | 567.21 | 578.89 | 0.00% | 977.33 |
| | | 4 | 0.00% | 10 | 35.77 | 11.59 | 117.57 | 164.93 | 0.00% | 12.00 | 207.80 | 219.80 | 0.00% | 464.58 |
| | | 5 | 0.00% | 40 | 33.73 | 10.87 | 127.89 | 172.48 | 0.00% | 12.05 | 210.45 | 222.51 | 0.00% | 6248.18 |
| | 0.2 | 1 | 0.00% | 100 | 31.82 | 9.53 | 2305.09 | 2346.44 | 0.24% | 12.19 | 14387.81 | 14400.00 | 0.36% | 14400.00 |
| | | 2 | 0.00% | 96 | 29.59 | 9.57 | 2056.57 | 2095.73 | 0.02% | 11.87 | 14388.13 | 14400.00 | 0.25% | 14400.00 |
| | | 3 | 0.00% | 91 | 28.72 | 9.73 | 2713.46 | 2751.91 | 0.05% | 12.01 | 14387.99 | 14400.00 | 0.13% | 14400.00 |
| | | 4 | 0.00% | 95 | 29.35 | 9.88 | 905.32 | 944.54 | 0.09% | 11.80 | 14388.20 | 14400.00 | 0.24% | 14400.00 |
| | | 5 | 0.00% | 88 | 30.10 | 9.84 | 4102.81 | 4142.74 | 0.00% | 12.49 | 9356.12 | 9368.61 | 0.05% | 14400.00 |
| 1-7-5-1000 | 0.1 | 1 | 0.00% | 45 | 39.56 | 10.54 | 27.99 | 78.09 | 0.00% | 11.47 | 51.17 | 62.64 | 0.00% | 163.70 |
| | | 2 | 0.00% | 42 | 41.22 | 11.16 | 23.82 | 76.20 | 0.00% | 12.05 | 31.81 | 43.86 | 0.00% | 432.92 |
| | | 3 | 0.00% | 34 | 37.89 | 11.21 | 42.87 | 91.96 | 0.00% | 11.97 | 81.67 | 93.64 | 0.00% | 450.59 |
| | | 4 | 0.00% | 12 | 37.16 | 11.62 | 34.35 | 83.12 | 0.00% | 12.12 | 54.17 | 66.29 | 0.00% | 70.68 |
| | | 5 | 0.00% | 15 | 35.43 | 11.57 | 191.56 | 238.56 | 0.00% | 11.90 | 81.87 | 93.77 | 0.00% | 4464.96 |
| | 0.2 | 1 | 0.00% | 101 | 38.97 | 9.68 | 67.07 | 115.72 | 0.00% | 12.12 | 190.62 | 202.74 | 0.00% | 5175.37 |
| | | 2 | 0.00% | 104 | 37.06 | 9.55 | 100.23 | 146.84 | 0.00% | 11.79 | 305.37 | 317.17 | 0.00% | 4161.60 |
| | | 3 | 0.00% | 89 | 36.49 | 9.81 | 119.32 | 165.62 | 0.00% | 11.91 | 209.46 | 221.36 | 0.00% | 3246.94 |
| | | 4 | 0.00% | 96 | 34.15 | 9.84 | 131.87 | 175.86 | 0.00% | 11.92 | 141.97 | 153.89 | 0.00% | 343.38 |
| | | 5 | 0.00% | 90 | 31.74 | 9.91 | 160.46 | 202.12 | 0.00% | 11.99 | 265.62 | 277.61 | 0.00% | 6225.72 |




(a) Total Running Time Comparisons of Table 1

(b) big-M Running Time Comparisons of Table 1

**Figure 5** Comparisons among different methods to solve an RCCP. The horizontal axis represents the logarithmic scale of running and the vertical axis represents the number of instances solved to optimality.

**Step 2.** We implement the algorithms in Proposition 3 to strengthen the big-M coefficients and use "S" to denote the running time of this step.

**Step 3.** Finally, we execute the big-M method.

We compare the numerical results for the following three methods: (i) Big-M method with fixing & strengthening; (ii) Big-M method with strengthening; and (iii) Vanilla big-M method. For the first method, we initialize the solver with the ALSO-X# solution. Besides, we add the inequalities $\boldsymbol{c}^\top \boldsymbol{x} \geq v_\infty^L$ and $\boldsymbol{c}^\top \boldsymbol{x} \leq v_\infty^U$ to the solver. From Step 1, we identify type $z_j = 0$ cut by Corollary 4. We then record the number of type $z_j = 0$ cuts as "# Cuts (A)" in our numerical results. The results are displayed in Table 2 and Table 3. It is seen that in terms of the solver's running time, the "Big-M & Fixing & Strengthening" approach is significantly faster than the other two methods. For most instances, the "Big-M & Fixing & Strengthening" approach consistently outperforms the others in terms of the total running time. This suggests that incorporating the variable fixing step greatly reduces computational time. We observe that for some instances, the Vanilla big-M method outperforms the other two methods in terms of the total running time, since strengthening the big-M coefficients takes longer. However, after applying fixing and strengthening, all the instances can be consistently easier to solve. In Figure 6, we provide the performance profile for the instances in Table 2 and Table 3, where the horizontal axis represents the logarithmic scale of running and the vertical axis represents the number of instances solved to optimality up to the time. It is seen that our methods consistently outperform other approaches by solving all instances significantly faster.

**Table 2** Numerical Results of a DRCCP under Type $\infty-$Wasserstein Ambiguity Set with Instances *1-7-1-1000* from Song et al. (2014) and $N = 1000$

| $\varepsilon$ | $\theta$ | Case | Big-M & Fixing & Strengthening | | | | | | Big-M & Strengthening | | | | Vanilla big-M | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | | | GAP | # of Cuts (A) | Time (s) | | | | GAP | Time (s) | | | GAP | Time (s) |
| | | | | | P | S | Solver | Total | | S | Solver | Total | | |
| 0.1 | 0.1 | 1 | 0.00% | 44 | 2.74 | 239.24 | 26.52 | 268.49 | 0.00% | 261.84 | 63.20 | 325.033 | 0.10% | 14400.00 |
| | | 2 | 0.00% | 41 | 2.86 | 237.00 | 40.01 | 279.87 | 0.00% | 257.21 | 73.13 | 330.342 | 0.00% | 4212.78 |
| | | 3 | 0.00% | 47 | 2.68 | 234.02 | 36.81 | 273.50 | 0.00% | 256.90 | 56.77 | 313.667 | 0.00% | 278.92 |
| | | 4 | 0.00% | 10 | 2.69 | 250.78 | 55.20 | 308.67 | 0.00% | 261.72 | 110.74 | 372.453 | 0.00% | 358.35 |
| | | 5 | 0.00% | 40 | 2.53 | 236.95 | 31.77 | 271.26 | 0.00% | 261.75 | 202.66 | 464.407 | 0.19% | 14400.00 |
| 0.1 | 0.2 | 1 | 0.00% | 44 | 2.50 | 238.73 | 25.97 | 267.19 | 0.00% | 260.30 | 89.92 | 350.22 | 0.00% | 271.27 |
| | | 2 | 0.00% | 41 | 2.74 | 242.71 | 192.00 | 437.44 | 0.00% | 263.32 | 65.60 | 328.92 | 0.00% | 222.88 |
| | | 3 | 0.00% | 47 | 2.84 | 235.15 | 33.57 | 271.56 | 0.00% | 257.72 | 68.55 | 326.27 | 0.00% | 504.59 |
| | | 4 | 0.00% | 10 | 2.53 | 254.07 | 75.24 | 331.84 | 0.00% | 260.69 | 74.70 | 335.39 | 0.00% | 371.44 |
| | | 5 | 0.00% | 40 | 2.39 | 271.12 | 39.32 | 312.82 | 0.00% | 292.17 | 9192.53 | 9486.70 | 0.21% | 14400.00 |
| 0.2 | 0.1 | 1 | 0.00% | 100 | 2.62 | 210.77 | 151.44 | 364.83 | 0.00% | 260.61 | 455.25 | 715.86 | 0.49% | 14400.00 |
| | | 2 | 0.00% | 96 | 2.59 | 211.94 | 484.01 | 698.54 | 0.00% | 260.62 | 502.13 | 762.75 | 0.00% | 5982.19 |
| | | 3 | 0.00% | 91 | 2.42 | 212.90 | 422.22 | 637.55 | 0.00% | 258.47 | 451.12 | 709.59 | 0.00% | 7459.28 |
| | | 4 | 0.00% | 95 | 2.55 | 214.82 | 89.69 | 307.06 | 0.00% | 262.07 | 233.01 | 495.08 | 0.00% | 392.11 |
| | | 5 | 0.00% | 88 | 2.33 | 219.20 | 220.01 | 441.54 | 0.00% | 264.26 | 530.54 | 794.80 | 0.24% | 14400.00 |
| 0.2 | 0.2 | 1 | 0.00% | 100 | 2.88 | 208.39 | 103.26 | 314.53 | 0.00% | 260.99 | 435.21 | 696.21 | 0.71% | 14400.00 |
| | | 2 | 0.00% | 96 | 2.34 | 212.54 | 269.56 | 484.44 | 0.00% | 263.49 | 3972.70 | 4236.19 | 0.00% | 6940.54 |
| | | 3 | 0.00% | 91 | 2.41 | 213.92 | 361.15 | 577.48 | 0.00% | 263.30 | 7284.95 | 7548.25 | 0.00% | 9282.48 |
| | | 4 | 0.00% | 95 | 2.75 | 212.74 | 142.21 | 357.70 | 0.00% | 261.96 | 419.94 | 681.90 | 0.29% | 14400.00 |
| | | 5 | 0.00% | 88 | 2.43 | 240.98 | 181.49 | 424.90 | 0.00% | 294.98 | 7274.92 | 7569.90 | 0.26% | 14400.00 |

**Table 3** Numerical Results of a DRCCP under Type $\infty-$Wasserstein Ambiguity Set with Instances *1-7-5-1000* from Song et al. (2014) and $N = 1000$

| $\varepsilon$ | $\theta$ | Case | Big-M & Fixing & Strengthening | | | | | | Big-M & Strengthening | | | | Vanilla big-M | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | | | GAP | # of Cuts (A) | Time (s) | | | | GAP | Time (s) | | | GAP | Time (s) |
| | | | | | P | S | Solver | Total | | S | Solver | Total | | |
| 0.1 | 0.1 | 1 | 0.00% | 45 | 2.62 | 243.35 | 28.96 | 274.93 | 0.00% | 267.71 | 38.09 | 305.80 | 0.00% | 251.34 |
| | | 2 | 0.00% | 42 | 2.69 | 241.76 | 22.53 | 266.98 | 0.00% | 267.38 | 38.78 | 306.16 | 0.00% | 144.18 |
| | | 3 | 0.00% | 34 | 2.63 | 246.36 | 20.42 | 269.42 | 0.00% | 268.88 | 61.48 | 330.36 | 0.00% | 181.80 |
| | | 4 | 0.00% | 12 | 2.67 | 259.55 | 17.20 | 279.42 | 0.00% | 264.62 | 36.64 | 301.26 | 0.00% | 5154.11 |
| | | 5 | 0.00% | 15 | 2.93 | 257.68 | 36.02 | 296.63 | 0.00% | 267.17 | 38.60 | 305.77 | 0.00% | 106.51 |
| 0.1 | 0.2 | 1 | 0.00% | 45 | 2.60 | 239.54 | 28.93 | 271.07 | 0.00% | 269.44 | 55.83 | 325.27 | 0.00% | 58.34 |
| | | 2 | 0.00% | 42 | 2.95 | 246.17 | 25.96 | 275.08 | 0.00% | 269.18 | 26.86 | 296.04 | 0.00% | 241.52 |
| | | 3 | 0.00% | 34 | 2.67 | 250.12 | 26.22 | 279.01 | 0.00% | 269.78 | 41.56 | 311.34 | 0.00% | 5176.38 |
| | | 4 | 0.00% | 12 | 2.58 | 260.35 | 36.96 | 299.89 | 0.00% | 266.61 | 48.48 | 315.09 | 0.00% | 158.37 |
| | | 5 | 0.00% | 15 | 2.39 | 256.24 | 43.78 | 302.41 | 0.00% | 263.31 | 54.89 | 318.20 | 0.00% | 256.57 |
| 0.2 | 0.1 | 1 | 0.00% | 99 | 2.58 | 215.57 | 49.21 | 267.36 | 0.00% | 269.53 | 74.96 | 344.49 | 0.00% | 1836.19 |
| | | 2 | 0.00% | 105 | 2.34 | 213.36 | 51.84 | 267.54 | 0.00% | 270.77 | 81.40 | 352.16 | 0.00% | 2357.91 |
| | | 3 | 0.00% | 89 | 2.63 | 222.09 | 80.94 | 305.66 | 0.00% | 268.43 | 99.39 | 367.82 | 0.00% | 310.70 |
| | | 4 | 0.00% | 96 | 2.51 | 219.01 | 122.26 | 343.78 | 0.00% | 266.99 | 145.37 | 412.36 | 0.00% | 1260.25 |
| | | 5 | 0.00% | 90 | 2.42 | 217.30 | 97.60 | 317.32 | 0.00% | 265.28 | 171.24 | 436.52 | 0.00% | 3081.89 |
| 0.2 | 0.2 | 1 | 0.00% | 99 | 2.70 | 217.55 | 44.16 | 264.41 | 0.00% | 271.13 | 153.52 | 424.65 | 0.00% | 6335.77 |
| | | 2 | 0.00% | 105 | 2.53 | 216.60 | 41.19 | 260.32 | 0.00% | 267.03 | 266.67 | 533.70 | 0.34% | 14400.00 |
| | | 3 | 0.00% | 89 | 2.77 | 218.84 | 116.33 | 337.93 | 0.00% | 265.86 | 126.36 | 392.23 | 0.00% | 8313.88 |
| | | 4 | 0.00% | 96 | 2.40 | 221.32 | 97.16 | 320.88 | 0.00% | 273.84 | 222.65 | 496.50 | 0.00% | 3867.32 |
| | | 5 | 0.00% | 90 | 3.71 | 225.89 | 68.13 | 297.72 | 0.00% | 275.18 | 264.43 | 539.62 | 0.00% | 6240.27 |

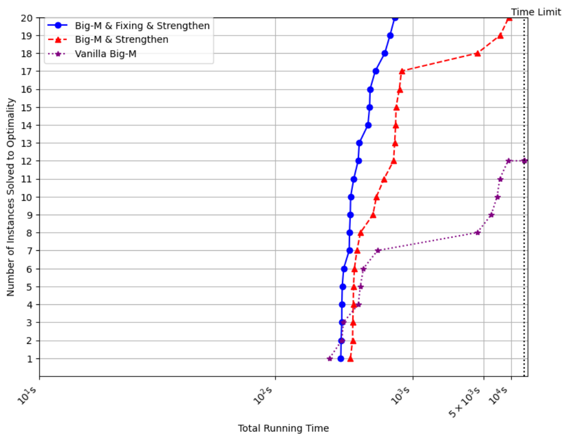
**Case III. Testing a DRCCP under type $1-$Wasserstein ambiguity set.** Let us consider the following DRCCP with the dual norm $\|\cdot\|_* = \|\cdot\|_\infty$:

$$v_1^* = \min_{\boldsymbol{x}\in[0,1]^n} \left\{ \boldsymbol{c}^\top \boldsymbol{x} \colon \inf_{\mathbb{P}\in\mathcal{P}_1} \mathbb{P}\left\{ \tilde{\boldsymbol{\xi}} \colon \boldsymbol{x}^\top \tilde{\boldsymbol{\xi}} \le \tilde{b} \right\} \ge 1-\varepsilon \right\}.$$

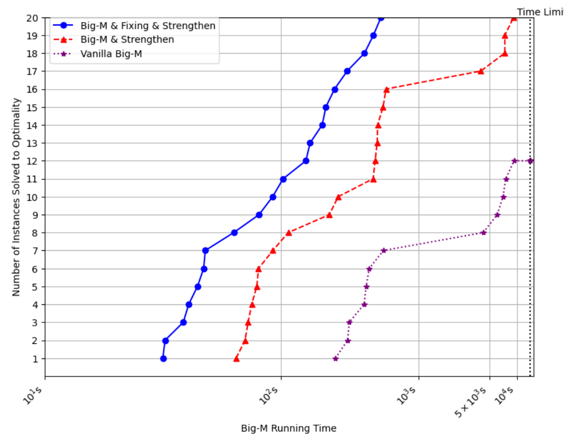
Again, we use the instances *1-7-1* and *1-7-5* from Song et al. (2014). For each instance, we follow the five steps below:

**Step 1.** We solve the quantile bound and use ALSO-X# (see, e.g., Jiang and Xie 2023) as the inner approximation. Then, we identify the cuts based on Corollary 8. We use "P" to denote the running time of this step.

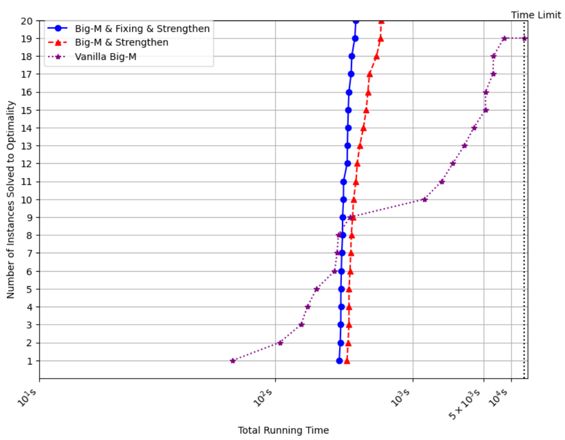
**Step 2.** We solve the dual bound (13) as the outer approximation. We use "DB" to denote the running time of this step.



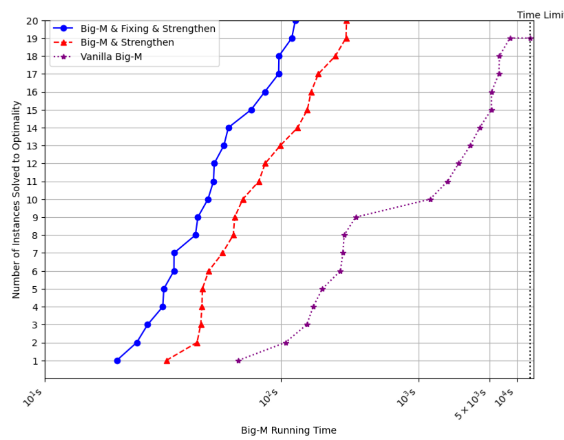

(a) Total Running Time Comparisons of Table 2




(b) big-M Running Time Comparisons of Table 2




(c) Total Running Time Comparisons of Table 3




(d) big-M Running Time Comparisons of Table 3

**Figure 6** Comparisons among different methods to solve a DRCCP under type $\infty-$Wasserstein ambiguity set. The horizontal axis represents the logarithmic scale of running and the vertical axis represents the number of instances solved to optimality.

**Step 3.** We fix the variables based on Corollary 7. We first sort the values of $\widehat{\eta}_j$ with $\widehat{\eta}_j \leq v_1^U$ in the descending order. For the initial $0.1 \times \lfloor N\varepsilon \rfloor$ scenarios, we verify whether $z_j = 0$. For the subsequent $0.2 \times \lfloor N\varepsilon \rfloor$ scenarios, we check if $z_j + z_{j+1} \leq 1$. We use "F" to denote the running time of this step.
**Step 4.** We solve the improved VaR lower bound (16) and an improved upper bound presented in Appendix B, respectively. Then, we strengthen the coefficient $\gamma$ from (18) and big-M coefficients from (17). We use "S" to denote the running time of this step.
**Step 5.** Finally, we execute the big-M method.

We compare the total running time for the following three methods: (i) Big-M method with fixing & strengthening; (ii) Big-M method with strengthening; and (iii) Vanilla big-M method. For the first method, we initialize the solver with the solution of the improved upper bound from Appendix B. Besides, we add the inequalities $\boldsymbol{c}^\top \boldsymbol{x} \geq v_1^L$ and $\boldsymbol{c}^\top \boldsymbol{x} \leq v_1^U$. We set the time limits for each optimization problem in Step 4 as 1200 seconds. We referred to Steps 1 to 4 collectively as the "Pre-compute" process. We record the number of cuts into two categories: category (A) represents the number of type $z_j = 0$ cuts, and category (B) represents the number of type $z_j + z_{j+1} \leq 1$ cuts. Note that the cuts identified from category (A) are based on Corollary 7 and Corollary 8. It is important to note that other types of cuts can be derived from Corollary 7. However, since category (A) and category (B) are sufficient to close the gap, for the sake of time, we do not explore these alternative cuts in our experimentation. In the second method, we use (17) to strengthen big-M coefficients to optimality. We also compare this big-M coefficient strengthening technique with the approach outlined in section

4.2 of Ho-Nguyen et al. (2023). The results can be founded in Appendix C. Based on our numerical study findings, it is evident that both the big-M method with strengthening and the Vanilla big-M method are unable to effectively close the gap in nearly all instances when $N=500$. To ensure a fair comparison, we use $N=500$ for this numerical case. Table 4 and Table 5, we report the detailed numerical results. In all instances, we show that our "Big-M & Fixing & Strengthening" approach consistently outperforms other approaches dramatically. In Figure 7, we provide the performance profile for the instances in Table 4 and Table 5, where the horizontal axis represents the logarithmic scale of running and the vertical axis represents the number of instances solved to optimality up to the time. Our approach succeeds in closing the gap for all the instances within a reasonable time. In contrast, other methods can only achieve optimality for a limited number of instances and may fail to close the gap within the time limit.

**Table 4** Numerical Results of a DRCCP under Type 1−Wasserstein Ambiguity Set with Instances *1-7-1-500* from Song et al. (2014) and $N=500$

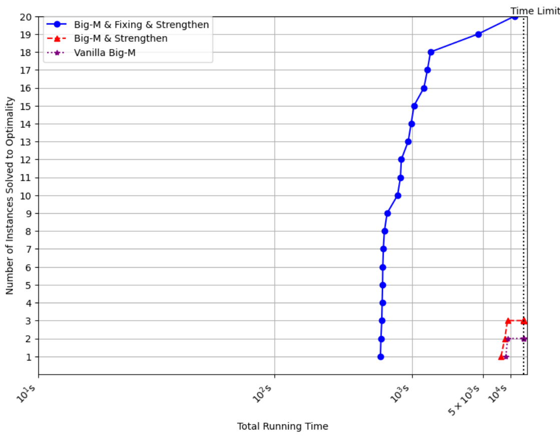
| $\varepsilon$ | $\theta$ | Case | Big-M & Fixing & Strengthening | | | | | | | | | Big-M & Strengthening | | | | Vanilla big-M | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | | | GAP | Pre-compute Time (s) | | | | # of Cuts | | Time (s) | | GAP | Time (s) | | | GAP | Time (s) |
| | | | | P | DB | F | S | A | B | Solver | Total | | S | Solver | Total | | |
| 0.1 | 0.1 | 1 | 0.00% | 1.07 | 131.70 | 131.30 | 282.14 | 15 | 8 | 6.31 | 552.51 | 0.00% | 259.68 | 9007.25 | 9266.92 | 0.00% | 9312.72 |
| | | 2 | 0.00% | 1.08 | 119.27 | 115.28 | 288.78 | 13 | 5 | 32.23 | 556.64 | 0.08% | 261.48 | 14138.52 | 14400.00 | 0.11% | 14400.00 |
| | | 3 | 0.00% | 1.09 | 111.99 | 149.58 | 287.28 | 17 | 4 | 18.91 | 568.84 | 0.00% | 265.26 | 8364.19 | 8629.45 | 0.00% | 8866.28 |
| | | 4 | 0.00% | 1.07 | 129.41 | 127.80 | 280.15 | 18 | 2 | 5.42 | 543.85 | 0.00% | 262.52 | 7550.28 | 7812.80 | 0.04% | 14400.00 |
| | | 5 | 0.00% | 1.04 | 143.04 | 105.86 | 273.19 | 15 | 7 | 16.35 | 539.47 | 0.21% | 259.86 | 14140.14 | 14400.00 | 0.35% | 14400.00 |
| 0.1 | 0.2 | 1 | 0.00% | 1.06 | 96.40 | 116.41 | 279.63 | 13 | 5 | 69.84 | 563.34 | 0.23% | 274.32 | 14125.68 | 14400.00 | 0.31% | 14400.00 |
| | | 2 | 0.00% | 1.08 | 99.35 | 118.37 | 466.57 | 15 | 1 | 290.73 | 976.10 | 0.39% | 258.33 | 14141.67 | 14400.00 | 0.45% | 14400.00 |
| | | 3 | 0.00% | 1.03 | 90.97 | 120.51 | 312.16 | 14 | 1 | 55.43 | 580.09 | 0.27% | 261.94 | 14138.06 | 14400.00 | 0.29% | 14400.00 |
| | | 4 | 0.00% | 1.10 | 115.32 | 133.41 | 293.75 | 16 | 1 | 17.61 | 561.19 | 0.17% | 266.36 | 14133.64 | 14400.00 | 0.20% | 14400.00 |
| | | 5 | 0.00% | 1.07 | 110.84 | 112.57 | 329.76 | 11 | 6 | 56.86 | 611.11 | 0.33% | 275.13 | 14124.87 | 14400.00 | 0.58% | 14400.00 |
| 0.2 | 0.1 | 1 | 0.00% | 1.09 | 186.54 | 248.18 | 312.26 | 53 | 6 | 41.32 | 789.39 | 0.35% | 262.58 | 14137.42 | 14400.00 | 0.38% | 14400.00 |
| | | 2 | 0.00% | 1.09 | 161.48 | 231.14 | 346.23 | 42 | 21 | 628.72 | 1368.66 | 0.42% | 261.92 | 14138.08 | 14400.00 | 0.55% | 14400.00 |
| | | 3 | 0.00% | 1.04 | 189.11 | 264.79 | 347.50 | 43 | 4 | 231.93 | 1034.36 | 0.42% | 265.94 | 14134.06 | 14400.00 | 0.49% | 14400.00 |
| | | 4 | 0.00% | 1.03 | 140.10 | 223.83 | 289.08 | 50 | 6 | 148.65 | 802.69 | 0.48% | 262.21 | 14137.79 | 14400.00 | 0.54% | 14400.00 |
| | | 5 | 0.00% | 1.01 | 161.13 | 245.58 | 282.36 | 51 | 4 | 55.78 | 745.86 | 0.19% | 266.94 | 14133.07 | 14400.00 | 0.28% | 14400.00 |
| 0.2 | 0.2 | 1 | 0.00% | 1.06 | 153.79 | 259.54 | 359.54 | 50 | 6 | 496.22 | 1270.14 | 0.59% | 258.22 | 14141.78 | 14400.00 | 0.62% | 14400.00 |
| | | 2 | 0.00% | 1.07 | 126.78 | 223.90 | 661.17 | 41 | 14 | 3329.93 | 4342.85 | 0.81% | 261.01 | 14139.00 | 14400.00 | 0.83% | 14400.00 |
| | | 3 | 0.00% | 1.07 | 152.61 | 247.71 | 1705.61 | 39 | 6 | 9132.78 | 11239.78 | 1.19% | 257.82 | 14142.18 | 14400.00 | 1.27% | 14400.00 |
| | | 4 | 0.00% | 1.03 | 135.50 | 231.21 | 513.65 | 49 | 3 | 578.11 | 1459.50 | 0.72% | 258.79 | 14141.21 | 14400.00 | 0.84% | 14400.00 |
| | | 5 | 0.00% | 1.04 | 132.26 | 244.91 | 314.52 | 48 | 6 | 226.93 | 919.65 | 0.45% | 259.01 | 14140.99 | 14400.00 | 0.63% | 14400.00 |

**Table 5** Numerical Results of a DRCCP under Type 1−Wasserstein Ambiguity Set with Instances *1-7-5-500* from Song et al. (2014) and $N=500$

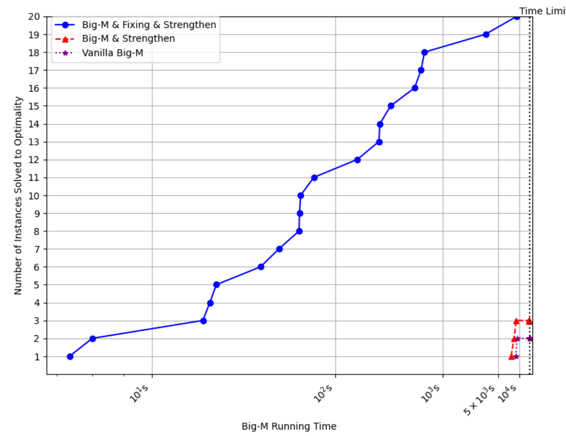
| $\varepsilon$ | $\theta$ | Case | Big-M & Fixing & Strengthening | | | | | | | | | Big-M & Strengthening | | | | Vanilla big-M | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | | | GAP | Pre-compute Time (s) | | | | # of Cuts | | Time (s) | | GAP | Time (s) | | | GAP | Time (s) |
| | | | | P | DB | F | S | A | B | Solver | Total | | S | Solver | Total | | |
| 0.1 | 0.1 | 1 | 0.00% | 1.02 | 104.00 | 103.44 | 273.98 | 14 | 4 | 1.88 | 484.32 | 0.18% | 258.23 | 14141.77 | 14400.00 | 0.19% | 14400.00 |
| | | 2 | 0.00% | 1.06 | 102.47 | 97.27 | 267.75 | 20 | 4 | 5.39 | 473.94 | 0.16% | 258.27 | 14141.73 | 14400.00 | 0.24% | 14400.00 |
| | | 3 | 0.00% | 1.01 | 113.95 | 97.72 | 272.34 | 17 | 2 | 4.75 | 489.78 | 0.08% | 261.58 | 14138.42 | 14400.00 | 0.15% | 14400.00 |
| | | 4 | 0.00% | 1.03 | 113.83 | 123.10 | 270.59 | 15 | 0 | 17.48 | 526.03 | 0.00% | 264.27 | 9837.24 | 10101.51 | 0.00% | 13829.29 |
| | | 5 | 0.00% | 0.99 | 108.99 | 105.60 | 277.11 | 17 | 7 | 3.24 | 495.94 | 0.02% | 264.54 | 14135.46 | 14400.00 | 0.04% | 14400.00 |
| 0.1 | 0.2 | 1 | 0.00% | 1.02 | 91.64 | 115.85 | 282.33 | 11 | 1 | 7.77 | 498.61 | 0.23% | 265.56 | 14134.44 | 14400.00 | 0.28% | 14400.00 |
| | | 2 | 0.00% | 1.13 | 101.97 | 100.21 | 283.16 | 18 | 4 | 19.47 | 505.93 | 0.19% | 271.49 | 14128.51 | 14400.00 | 0.44% | 14400.00 |
| | | 3 | 0.00% | 1.01 | 93.21 | 96.29 | 282.30 | 15 | 2 | 13.21 | 486.02 | 0.12% | 261.27 | 14138.73 | 14400.00 | 0.19% | 14400.00 |
| | | 4 | 0.00% | 1.07 | 108.98 | 107.96 | 291.32 | 15 | 0 | 38.24 | 547.57 | 0.22% | 270.09 | 14129.91 | 14400.00 | 0.27% | 14400.00 |
| | | 5 | 0.00% | 1.01 | 97.35 | 113.14 | 302.47 | 12 | 3 | 70.04 | 584.00 | 0.08% | 265.96 | 14134.04 | 14400.00 | 0.12% | 14400.00 |
| 0.2 | 0.1 | 1 | 0.00% | 1.06 | 180.16 | 220.73 | 300.57 | 46 | 11 | 59.24 | 761.75 | 0.24% | 257.38 | 14142.62 | 14400.00 | 0.31% | 14400.00 |
| | | 2 | 0.00% | 1.13 | 218.80 | 215.88 | 261.07 | 53 | 13 | 13.97 | 710.85 | 0.19% | 264.99 | 14135.01 | 14400.00 | 0.33% | 14400.00 |
| | | 3 | 0.00% | 1.05 | 186.46 | 191.07 | 255.93 | 59 | 11 | 25.02 | 659.51 | 0.18% | 265.56 | 14134.45 | 14400.00 | 0.20% | 14400.00 |
| | | 4 | 0.00% | 1.12 | 227.73 | 235.35 | 271.96 | 52 | 13 | 21.65 | 757.81 | 0.19% | 268.32 | 14131.68 | 14400.00 | 0.24% | 14400.00 |
| | | 5 | 0.00% | 1.04 | 148.83 | 214.04 | 271.05 | 50 | 11 | 9.77 | 644.73 | 0.34% | 259.83 | 14140.17 | 14400.00 | 0.38% | 14400.00 |
| 0.2 | 0.2 | 1 | 0.00% | 1.04 | 142.22 | 226.21 | 511.78 | 40 | 10 | 2564.28 | 3445.53 | 0.28% | 264.13 | 14135.87 | 14400.00 | 0.32% | 14400.00 |
| | | 2 | 0.00% | 1.03 | 159.95 | 224.37 | 296.90 | 50 | 10 | 53.51 | 735.76 | 0.41% | 273.49 | 14126.51 | 14400.00 | 0.41% | 14400.00 |
| | | 3 | 0.00% | 1.05 | 152.54 | 210.72 | 276.82 | 54 | 10 | 210.72 | 851.85 | 0.26% | 279.46 | 14120.54 | 14400.00 | 0.28% | 14400.00 |
| | | 4 | 0.00% | 1.06 | 172.02 | 217.10 | 500.34 | 46 | 10 | 221.59 | 1112.11 | 0.31% | 281.83 | 14118.17 | 14400.00 | 0.32% | 14400.00 |
| | | 5 | 0.00% | 1.03 | 137.40 | 224.02 | 312.50 | 46 | 7 | 165.99 | 840.94 | 0.44% | 269.85 | 14130.15 | 14400.00 | 0.45% | 14400.00 |

**Case IV. Testing a DRCCP under type 2−Wasserstein ambiguity set.** Let us consider the following DRCCP with the dual norm $\|\cdot\|_* = \|\cdot\|_\infty$:

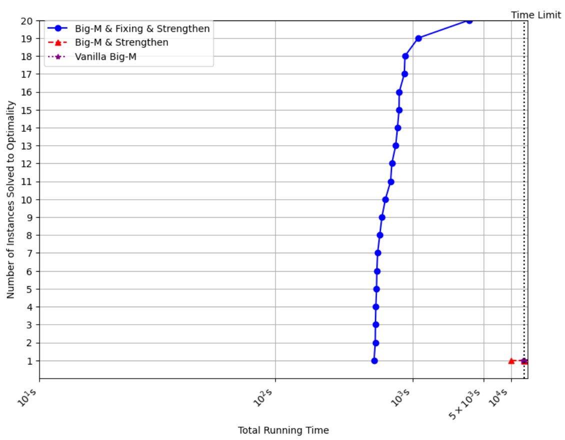
$$v_2^* = \min_{\boldsymbol{x}\in[0,1]^n} \left\{ \boldsymbol{c}^\top \boldsymbol{x} \colon \inf_{\mathbb{P}\in\mathcal{P}_2} \mathbb{P}\left\{ \tilde{\boldsymbol{\xi}} \colon \boldsymbol{x}^\top \tilde{\boldsymbol{\xi}} \le \tilde{b} \right\} \ge 1-\varepsilon \right\}.$$



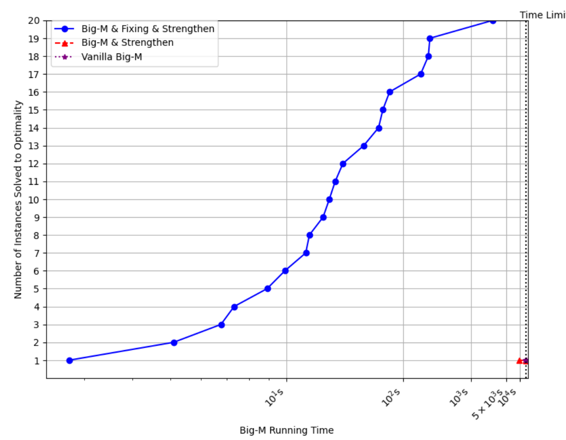

(a) Total Running Time Comparisons of Table 4




(b) big-M Running Time Comparisons of Table 4




(c) Total Running Time Comparisons of Table 5




(d) big-M Running Time Comparisons of Table 5

**Figure 7** Comparisons among different methods to solve a DRCCP under type 1−Wasserstein ambiguity set. The horizontal axis represents the logarithmic scale of running and the vertical axis represents the number of instances solved to optimality.

Since this DRCCP may not admit a mixed-integer convex programming reformulation (see the proof in Appendix A.3 of Jiang and Xie 2023), we report the gap between the best lower bound and best upper bound for this DRCCP from numerical experiments, where the best upper bound is the minimum of inner approximation (21) and ALSO-X#, and the best lower bound is the improved VaR lower bound (see the formulation in Appendix D). We employ the same instances as in Case III and follow similar steps accordingly. Since the big-M method with strengthening and Vanilla big-M method may not be able to solve the inner approximation (21) to optimality, we only provide the numerical results of the best "Big-M & Fixing & Strengthening" approach. We denote the best upper bound and best lower bound as "Best UB" and "Best LB," respectively. The term "Bound GAP" represents the gap between the best upper bound and best lower bound, i.e., $\text{Bound GAP}\ (\%) = (|\text{Best UB} - \text{Best LB}|)/(|\text{Best LB}|) \times 100$. The detailed numerical results are displayed in Table 6 and Table 7, which highlight the effectiveness of our "Big-M & Fixing & Strengthening" approach in providing upper and lower bounds for this DRCCP. We observe that the gap between the best upper bound and best lower bound achieved by our approach is very small, typically from 0.20% to 0.50%. This indicates that our method can accurately approximate the optimal solution of thw DRCCP, ensuring a high level of precision in the obtained bounds.

**Table 6** Numerical Results of a DRCCP under Type 2−Wasserstein Ambiguity Set with Instances *1-7-1-500* from Song et al. (2014) and $N = 500$

| $\varepsilon$ | $\theta$ | Case | Big-M & Fixing & Strengthening for Solving (21) | | | | | | | | | Best UB | Best LB | Bound GAP |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | | | GAP | Pre-compute Time (s) | | | | # Cuts | | Time (s) | | | | |
| | | | | P | DB | F | S | A | B | Solver | Total | | | |
| 0.1 | 0.1 | 1 | 0.00% | 1.00 | 151.43 | 120.86 | 272.96 | 21 | 7 | 4.95 | 551.20 | -16668.21 | -16722.67 | 0.33% |
| | | 2 | 0.00% | 1.02 | 111.84 | 125.19 | 285.44 | 17 | 5 | 7.70 | 531.19 | -16747.49 | -16796.13 | 0.29% |
| | | 3 | 0.00% | 0.99 | 106.68 | 126.00 | 269.67 | 20 | 3 | 4.08 | 507.42 | -16731.69 | -16785.97 | 0.32% |
| | | 4 | 0.00% | 1.01 | 137.94 | 120.94 | 267.21 | 21 | 9 | 4.97 | 532.07 | -16624.30 | -16680.61 | 0.34% |
| | | 5 | 0.00% | 1.00 | 140.13 | 114.63 | 271.22 | 22 | 2 | 5.84 | 532.81 | -16683.48 | -16743.52 | 0.36% |
| 0.1 | 0.2 | 1 | 0.00% | 1.02 | 134.89 | 110.37 | 279.67 | 17 | 7 | 4.65 | 530.60 | -16646.54 | -16717.80 | 0.43% |
| | | 2 | 0.00% | 1.01 | 112.28 | 111.09 | 276.58 | 15 | 3 | 10.27 | 511.23 | -16723.72 | -16791.25 | 0.40% |
| | | 3 | 0.00% | 1.00 | 102.80 | 122.81 | 289.94 | 17 | 6 | 6.04 | 522.59 | -16702.37 | -16782.72 | 0.48% |
| | | 4 | 0.00% | 1.03 | 140.01 | 113.99 | 276.68 | 20 | 5 | 4.85 | 536.55 | -16607.48 | -16675.05 | 0.41% |
| | | 5 | 0.00% | 1.07 | 129.90 | 139.29 | 285.56 | 19 | 6 | 8.61 | 564.43 | -16661.95 | -16738.96 | 0.46% |
| 0.2 | 0.1 | 1 | 0.00% | 1.06 | 209.38 | 228.97 | 274.72 | 57 | 4 | 19.75 | 733.88 | -16901.22 | -16954.33 | 0.31% |
| | | 2 | 0.00% | 1.02 | 178.04 | 224.29 | 285.65 | 50 | 13 | 54.59 | 743.57 | -16953.85 | -17014.68 | 0.36% |
| | | 3 | 0.00% | 1.02 | 175.48 | 234.49 | 287.16 | 44 | 8 | 42.19 | 740.33 | -16904.58 | -16953.15 | 0.29% |
| | | 4 | 0.00% | 1.02 | 171.44 | 230.95 | 267.87 | 54 | 5 | 18.27 | 689.54 | -16886.35 | -16950.49 | 0.38% |
| | | 5 | 0.00% | 1.05 | 185.68 | 242.01 | 260.23 | 53 | 8 | 19.18 | 708.14 | -16945.52 | -17010.20 | 0.38% |
| 0.2 | 0.2 | 1 | 0.00% | 1.04 | 180.43 | 238.24 | 305.59 | 53 | 6 | 116.68 | 841.97 | -16880.38 | -16951.33 | 0.42% |
| | | 2 | 0.00% | 1.06 | 151.58 | 232.85 | 328.36 | 43 | 19 | 1040.28 | 1754.13 | -16930.07 | -17012.18 | 0.48% |
| | | 3 | 0.00% | 1.05 | 189.25 | 244.83 | 308.81 | 43 | 5 | 133.22 | 877.15 | -16879.91 | -16950.54 | 0.42% |
| | | 4 | 0.00% | 1.06 | 148.14 | 233.32 | 299.55 | 50 | 6 | 151.33 | 833.39 | -16862.51 | -16947.86 | 0.50% |
| | | 5 | 0.00% | 1.06 | 156.36 | 250.08 | 276.14 | 51 | 5 | 40.85 | 724.48 | -16922.36 | -17007.05 | 0.50% |

**Table 7** Numerical Results of a DRCCP under Type 2−Wasserstein Ambiguity Set with Instances *1-7-5-500* from Song et al. (2014) and $N = 500$

| $\varepsilon$ | $\theta$ | Case | Big-M & Fixing & Strengthening for Solving (21) | | | | | | | | | Best UB | Best LB | Bound GAP |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | | | GAP | Pre-compute Time (s) | | | | # Cuts | | Time (s) | | | | |
| | | | | P | DB | F | S | A | B | Solver | Total | | | |
| 0.1 | 0.1 | 1 | 0.00% | 1.03 | 145.43 | 106.21 | 277.18 | 17 | 3 | 2.26 | 532.11 | -17720.04 | -17750.66 | 0.17% |
| | | 2 | 0.00% | 0.99 | 127.75 | 86.94 | 273.61 | 23 | 3 | 4.53 | 493.82 | -17730.25 | -17769.61 | 0.22% |
| | | 3 | 0.00% | 0.99 | 153.60 | 96.65 | 274.31 | 20 | 2 | 3.37 | 528.92 | -17648.42 | -17681.15 | 0.19% |
| | | 4 | 0.00% | 1.02 | 106.77 | 100.03 | 264.54 | 17 | 1 | 2.59 | 474.95 | -17727.31 | -17773.25 | 0.26% |
| | | 5 | 0.00% | 0.99 | 140.94 | 104.81 | 275.91 | 22 | 2 | 3.49 | 526.14 | -17685.80 | -17722.59 | 0.21% |
| 0.1 | 0.2 | 1 | 0.00% | 1.02 | 127.02 | 121.02 | 280.43 | 15 | 3 | 1.86 | 531.34 | -17705.08 | -17747.31 | 0.24% |
| | | 2 | 0.00% | 1.02 | 124.69 | 96.94 | 274.68 | 21 | 5 | 4.43 | 501.75 | -17713.00 | -17761.34 | 0.27% |
| | | 3 | 0.00% | 1.02 | 129.21 | 97.06 | 280.77 | 18 | 4 | 4.61 | 512.68 | -17630.60 | -17675.75 | 0.26% |
| | | 4 | 0.00% | 1.05 | 126.53 | 103.50 | 268.79 | 15 | 0 | 4.69 | 504.56 | -17710.79 | -17768.22 | 0.32% |
| | | 5 | 0.00% | 1.01 | 152.76 | 95.96 | 280.55 | 20 | 4 | 4.17 | 534.45 | -17671.31 | -17719.75 | 0.27% |
| 0.2 | 0.1 | 1 | 0.00% | 1.03 | 198.18 | 209.06 | 262.69 | 54 | 12 | 12.63 | 683.59 | -17863.89 | -17910.53 | 0.26% |
| | | 2 | 0.00% | 1.03 | 217.31 | 197.77 | 264.09 | 61 | 9 | 16.47 | 696.67 | -17890.76 | -17933.41 | 0.24% |
| | | 3 | 0.00% | 1.08 | 196.37 | 171.72 | 252.30 | 64 | 5 | 7.54 | 629.00 | -17841.16 | -17876.78 | 0.20% |
| | | 4 | 0.00% | 1.07 | 228.12 | 215.07 | 254.33 | 57 | 13 | 8.13 | 706.71 | -17889.62 | -17946.26 | 0.32% |
| | | 5 | 0.00% | 1.02 | 180.28 | 199.29 | 262.13 | 54 | 8 | 18.44 | 661.16 | -17849.98 | -17890.32 | 0.23% |
| 0.2 | 0.2 | 1 | 0.00% | 1.00 | 179.21 | 221.24 | 278.02 | 46 | 12 | 32.82 | 712.29 | -17845.53 | -17907.38 | 0.35% |
| | | 2 | 0.00% | 1.03 | 212.11 | 222.44 | 276.53 | 55 | 11 | 10.85 | 722.96 | -17875.92 | -17929.19 | 0.30% |
| | | 3 | 0.00% | 1.07 | 183.77 | 203.32 | 267.90 | 60 | 10 | 15.15 | 671.21 | -17826.09 | -17875.66 | 0.28% |
| | | 4 | 0.00% | 1.00 | 220.57 | 210.83 | 262.55 | 53 | 12 | 38.28 | 733.23 | -17870.43 | -17945.19 | 0.42% |
| | | 5 | 0.00% | 1.01 | 148.74 | 183.22 | 268.41 | 51 | 10 | 11.85 | 613.22 | -17834.13 | -17887.57 | 0.30% |

## 8. Conclusion

This study introduced a systematic framework for implementing variable fixing techniques within the context of Robust Chance-Constrained Programs (RCCPs) and Distributionally Robust Chance-Constrained Programs (DRCCPs) by integrating inner or outer approximations. We derived optimality cuts by probing the restricted outer approximations and comparing them with the inner ones for RCCPs or DRCCPs under type $q-$Wasserstein ambiguity set with $q \in \{1, \infty\}$. We provided a new conservative approximation for DRCCP under type $q-$Wasserstein ambiguity set with $q \in (1, \infty)$. We conducted a theoretical analysis of variable fixing techniques to evaluate the proportion of scenarios that should be fixed to be violated. We showcased the effectiveness of our proposed methods by reducing the running time and closing the gap for all reported instances.

## Acknowledgment

This research has been supported by the National Science Foundation grant 2246414.

## References

Ahmed S, Luedtke J, Song Y, Xie W (2017) Nonanticipative duality, relaxations, and formulations for chance-constrained stochastic programs. *Mathematical Programming* 162(1):51–81.

Ahmed S, Shapiro A (2008) Solving chance-constrained stochastic programs via sampling and integer programming. *State-of-the-art decision-making tools in the information-intensive age*, 261–269 (Informs).

Ahmed S, Xie W (2018) Relaxations and approximations of chance constraints under finite distributions. *Mathematical Programming* 170:43–65.

Anstreicher KM, Fampa M, Lee J, Williams J (1996) Continuous relaxations for constrained maximum-entropy sampling. *Integer Programming and Combinatorial Optimization: 5th International IPCO Conference Vancouver, British Columbia, Canada, June 3–5, 1996 Proceedings 5*, 234–248 (Springer).

Anstreicher KM, Fampa M, Lee J, Williams J, et al. (1999) Using continuous nonlinear relaxations to solve. constrained maximum-entropy sampling problems. *Mathematical programming* 85(2):221–240.

Atamturk A, Gómez A (2020) Safe screening rules for l0-regression from perspective relaxations. *International conference on machine learning*, 421–430 (PMLR).

Calafiore GC, Campi MC (2006) The scenario approach to robust control design. *IEEE Transactions on automatic control* 51(5):742–753.

Charnes A, Cooper WW (1963) Deterministic equivalents for optimizing and satisficing under chance constraints. *Operations research* 11(1):18–39.

Charnes A, Cooper WW, Symonds GH (1958) Cost horizons and certainty equivalents: an approach to stochastic programming of heating oil. *Management science* 4(3):235–263.

Chen Z, Kuhn D, Wiesemann W (2022) Data-driven chance constrained programs over wasserstein balls. *Operations Research* .

Chen Z, Kuhn D, Wiesemann W (2023) On approximations of data-driven chance constrained programs over wasserstein balls. *Operations Research Letters* 51(3):226–233.

Cho J, Papavasiliou A (2023) Exact mixed-integer programming approach for chance-constrained multi-area reserve sizing. *IEEE Transactions on Power Systems* .

Deng Y, Shen S (2016) Decomposition algorithms for optimizing multi-server appointment scheduling with chance constraints. *Mathematical Programming* 157:245–276.

Fattahi S, Lavaei J, Atamtürk A (2018) A bound strengthening method for optimal transmission switching in power systems. *IEEE Transactions on Power Systems* 34(1):280–291.

Fischetti M, Lodi A (2010) Heuristics in mixed integer programming. *Wiley Encyclopedia of Operations Research and Management Science* .

Ghosal S, Wiesemann W (2020) The distributionally robust chance-constrained vehicle routing problem. *Operations Research* 68(3):716–732.

Hanasusanto GA, Roitch V, Kuhn D, Wiesemann W (2015) A Distributionally Robust Perspective on Uncertainty Quantification and Chance Constrained Programming. *Mathematical Programming* 151(1):35–62.

Hanasusanto GA, Roitch V, Kuhn D, Wiesemann W (2017) Ambiguous joint chance constraints under mean and dispersion information. *Operations Research* 65(3):751–767.

Ho-Nguyen N, Kılınç-Karzan F, Küçükyavuz S, Lee D (2022) Distributionally robust chance-constrained programs with right-hand side uncertainty under wasserstein ambiguity. *Mathematical Programming* 196(1):641–672.

Ho-Nguyen N, Kilinç-Karzan F, Küçükyavuz S, Lee D (2023) Strong formulations for distributionally robust chance-constrained programs with left-hand side uncertainty under wasserstein ambiguity. *INFORMS Journal on Optimization* 5(2):211–232.

Ji R, Lejeune MA (2021) Data-driven distributionally robust chance-constrained optimization with wasserstein metric. *Journal of Global Optimization* 79(4):779–811.

Jiang N, Xie W (2022) ALSO-X and ALSO-X+: Better Convex Approximations for Chance Constrained Programs. *Operations Research* 70(6):3581–3600.

Jiang N, Xie W (2023) ALSO-X#: Better Convex Approximations for Distributionally Robust Chance Constrained Programs. *arXiv preprint arXiv:2302.01737* .

Küçükyavuz S, Jiang R (2022) Chance-constrained optimization under limited distributional information: a review of reformulations based on sampling and distributional robustness. *EURO Journal on Computational Optimization* 100030.

Li B, Rong Y, Sun J, Teo KL (2016) A distributionally robust linear receiver design for multi-access space-time block coded mimo systems. *IEEE Transactions on Wireless Communications* 16(1):464–474.

Li S, Hou YT, Lou W, Jalaian BA, Russell S (2022a) Maximizing energy efficiency with channel uncertainty under mutual interference. *IEEE Transactions on Wireless Communications* 21(10):8476–8488.

Li S, Jiang N, Chen Y, Hou YT, Lou W, Xie W (2022b) $D^2$BF—Data-Driven Beamforming in MU-MIMO with Channel Estimation Uncertainty. *IEEE INFOCOM 2022-IEEE Conference on Computer Communications*, 120–129 (IEEE).

Li S, Li C, Huang Y, Jalaian BA, Hou YT, Lou W (2023) Enhancing resilience in mobile edge computing under processing uncertainty. *IEEE Journal on Selected Areas in Communications* 41(3):659–674.

Li Y, Fampa M, Lee J, Qiu F, Xie W, Yao R (2022c) D-optimal data fusion: Exact and approximation algorithms. *arXiv preprint arXiv:2208.03589* .

Lubin M, Dvorkin Y, Backhaus S (2015) A robust approach to chance constrained optimal power flow with renewable generation. *IEEE Transactions on Power Systems* 31(5):3840–3849.

Luedtke J, Ahmed S (2008) A sample approximation approach for optimization with probabilistic constraints. *SIAM Journal on Optimization* 19(2):674–699.

Nemirovski A, Shapiro A (2006) Scenario approximations of chance constraints. *Probabilistic and randomized methods for design under uncertainty* 3–47.

Nemirovski A, Shapiro A (2007) Convex approximations of chance constrained programs. *SIAM Journal on Optimization* 17(4):969–996.

Porras Á, Domínguez C, Morales JM, Pineda S (2023) Tight and compact sample average approximation for joint chance-constrained problems with applications to optimal power flow. *INFORMS Journal on Computing* .

Posta M, Ferland JA, Michelon P (2012) An exact method with variable fixing for solving the generalized assignment problem. *Computational Optimization and Applications* 52:629–644.

Prékopa A (1973) On logarithmic concave measures and functions. *Acta Scientiarum Mathematicarum* 34:335–343.

Prékopa A (1980) Logarithmic concave measurement and related topics. *Stochastic programming* 63–82.

Qiu F, Ahmed S, Dey SS, Wolsey LA (2014) Covering linear programming with violations. *INFORMS Journal on Computing* 26(3):531–546.

Rockafellar RT, Uryasev S, et al. (2000) Optimization of conditional value-at-risk. *Journal of risk* 2:21–42.

Rudin W (1987) *Real and Complex Analysis, 3rd Ed.* (USA: McGraw-Hill, Inc.), ISBN 0070542341.

Ruszczyński A (2002) Probabilistic programming with discrete distributions and precedence constrained knapsack polyhedra. *Mathematical Programming* 93:195–215.

Savelsbergh MW (1994) Preprocessing and probing techniques for mixed integer programming problems. *ORSA Journal on Computing* 6(4):445–454.

Shapiro A, Ahmed S (2004) On a class of minimax stochastic programs. *SIAM Journal on Optimization* 14(4):1237–1249.

Shapiro A, Dentcheva D, Ruszczynski A (2021) *Lectures on stochastic programming: modeling and theory* (SIAM).

Shen H, Jiang R (2021) Convex chance-constrained programs with wasserstein ambiguity. *arXiv preprint arXiv:2111.02486* .

Shen H, Jiang R (2023) Chance-constrained set covering with wasserstein ambiguity. *Mathematical Programming* 198(1):621–674.

Simic V (2016) Interval-parameter chance-constraint programming model for end-of-life vehicles management under rigorous environmental regulations. *Waste Management* 52:180–192.

Sion M (1958) On general minimax theorems. *Pacific Journal of mathematics* 8(1):171–176.

Song Y, Luedtke JR, Küçükyavuz S (2014) Chance-constrained binary packing problems. *INFORMS Journal on Computing* 26(4):735–747.

Song Y, Shen S (2016) Risk-averse shortest path interdiction. *INFORMS Journal on Computing* 28(3):527–539.

Wang KY, So AMC, Chang TH, Ma WK, Chi CY (2014) Outage constrained robust transmit optimization for multiuser miso downlinks: Tractable approximations by conic optimization. *IEEE Transactions on Signal Processing* 62(21):5690–5705.

Wang S, Li J, Mehrotra S (2021) Chance-constrained multiple bin packing problem with an application to operating room planning. *INFORMS Journal on Computing* 33(4):1661–1677.

Wu H, Shahidehpour M, Li Z, Tian W (2014) Chance-constrained day-ahead scheduling in stochastic power system operation. *IEEE Transactions on Power Systems* 29(4):1583–1591.

Wu W, Iori M, Martello S, Yagiura M (2018) Exact and heuristic algorithms for the interval min-max regret generalized assignment problem. *Computers & Industrial Engineering* 125:98–110.

Xie W (2021) On distributionally robust chance constrained programs with wasserstein distance. *Mathematical Programming* 186(1):115–155.

Xie W, Ahmed S (2017) Distributionally robust chance constrained optimal power flow with renewables: A conic reformulation. *IEEE Transactions on Power Systems* 33(2):1860–1867.

Xie W, Ahmed S (2018) On deterministic reformulations of distributionally robust joint chance constrained optimization problems. *SIAM Journal on Optimization* 28(2):1151–1182.

Zhai J, Jiang Y, Shi Y, Jones CN, Zhang XP (2022) Distributionally robust joint chance-constrained dispatch for integrated transmission-distribution systems via distributed optimization. *IEEE Transactions on Smart Grid* 13(3):2132–2147.

Zhang Z, Denton BT, Xie X (2020) Branch and price for chance-constrained bin packing. *INFORMS Journal on Computing* 32(3):547–564.

Zhao S, Zhang K (2020) A distributionally robust stochastic optimization-based model predictive control with distributionally robust chance constraints for cooperative adaptive cruise control under uncertain traffic conditions. *Transportation Research Part B: Methodological* 138:144–178.

Zhou Y, Yu W, Zhu S, Yang B, He J (2021) Distributionally robust chance-constrained energy management of an integrated retailer in the multi-energy market. *Applied Energy* 286:116516.

Zymler S, Kuhn D, Rustem B (2013) Distributionally robust joint chance constraints with second-order moment information. *Mathematical Programming* 137(1):167–198.

# Appendix A Proofs

# Proofs in Section 3

## A.1 Proof of Theorem 2

THEOREM 2. *Suppose $\theta > 0$. For any dual norm $\|\cdot\|_* = \|\cdot\|_p$ and $p \in [1, \infty)$, solving Problem (8), in general, is NP-hard.*

*Proof.* This poof reduces Problem (8) to the feasibility of a generic binary integer program. Consider the following NP-complete problem — feasibility problem of a binary integer program:

Does there exist a feasible solution to the binary program $\mathcal{X} = \{\boldsymbol{x} \in \{-1,1\}^n : \boldsymbol{A}\boldsymbol{x} \geq \boldsymbol{b}\}$?

In Problem (8), suppose $\widehat{\boldsymbol{\xi}}^j = \widehat{\boldsymbol{\xi}}^{j'} = \boldsymbol{0}, \boldsymbol{a}_i(\boldsymbol{x}) = \boldsymbol{x}, b_i(\boldsymbol{x}) = b_i > \max_{\boldsymbol{x}\in\mathcal{X}} \theta\|\boldsymbol{x}\|_*$. In this case, the constraint of Problem (8) is redundant and Problem (8) reduces to

$$\eta_{i,j}(j'|\theta) = \max_{\boldsymbol{x}\in\mathcal{X}} \theta\|\boldsymbol{x}\|_*,$$

which has the optimal value $\theta\sqrt[p]{n}$ if and only if there exists a feasible binary solution to the set $\mathcal{X}$. Hence, solving Problem (8) is NP-hard. □

## A.2 Proof of Proposition 1

PROPOSITION 1. *Suppose that in Problem (8), set $\mathcal{X}$ is compact and convex and the dual norm $\|\cdot\|_* = \|\cdot\|_\infty$. Then Problem (8) is equivalent to solving $2n$ tractable convex programs, i.e., $\eta_{i,j}(j'|\theta) = \max_{\tau\in[n]} \max_{k\in[n]} \max_{\ell\in[2]} \eta_{i,j}(j',\tau,\ell|\theta)$, where for each $\tau \in [n]$, we have*

$$\eta_{i,j}(j',\tau,1|\theta) = \max_{\boldsymbol{x}\in\mathcal{X}} \left\{ \theta a_{i\tau}(\boldsymbol{x}) + \boldsymbol{a}_i(\boldsymbol{x})^\top \widehat{\boldsymbol{\xi}}^j - b_i(\boldsymbol{x}) \colon \theta \left\|\boldsymbol{a}_i(\boldsymbol{x})\right\|_\infty + \boldsymbol{a}_i(\boldsymbol{x})^\top \widehat{\boldsymbol{\xi}}^{j'} \leq b_i(\boldsymbol{x}) \right\},$$
$$\eta_{i,j}(j',\tau,2|\theta) = \max_{\boldsymbol{x}\in\mathcal{X}} \left\{ -\theta a_{i\tau}(\boldsymbol{x}) + \boldsymbol{a}_i(\boldsymbol{x})^\top \widehat{\boldsymbol{\xi}}^j - b_i(\boldsymbol{x}) \colon \theta \left\|\boldsymbol{a}_i(\boldsymbol{x})\right\|_\infty + \boldsymbol{a}_i(\boldsymbol{x})^\top \widehat{\boldsymbol{\xi}}^{j'} \leq b_i(\boldsymbol{x}) \right\}.$$

*Proof.* When the dual norm $\|\cdot\|_* = \|\cdot\|_\infty$, we have $\|\boldsymbol{a}_i(\boldsymbol{x})\|_\infty = \theta \max\{a_{i\tau}(\boldsymbol{x}), -a_{i\tau}(\boldsymbol{x})\}_{\tau\in[n]}$. In this case, we write Problem (8) as

$$\eta_{i,j}(j'|\theta) = \max_{\boldsymbol{x}\in\mathcal{X}} \theta \max\{a_{i\tau}(\boldsymbol{x}), -a_{i\tau}(\boldsymbol{x})\}_{\tau\in[n]} + \boldsymbol{a}_i(\boldsymbol{x})^\top \widehat{\boldsymbol{\xi}}^j - b_i(\boldsymbol{x}),$$
$$\text{s.t.}\, \theta \left\|\boldsymbol{a}_i(\boldsymbol{x})\right\|_\infty + \boldsymbol{a}_i(\boldsymbol{x})^\top \widehat{\boldsymbol{\xi}}^{j'} \leq b_i(\boldsymbol{x}).$$

Then we can simplify it as

$$\eta_{i,j}(j'|\theta) = \max_{\tau\in[n]} \left\{ \begin{array}{l} \max_{\boldsymbol{x}\in\mathcal{X}} \left\{ \theta a_{i\tau}(\boldsymbol{x}) + \boldsymbol{a}_i(\boldsymbol{x})^\top \widehat{\boldsymbol{\xi}}^j - b_i(\boldsymbol{x}) \colon \theta \left\|\boldsymbol{a}_i(\boldsymbol{x})\right\|_\infty + \boldsymbol{a}_i(\boldsymbol{x})^\top \widehat{\boldsymbol{\xi}}^{j'} \leq b_i(\boldsymbol{x}) \right\}, \\ \max_{\boldsymbol{x}\in\mathcal{X}} \left\{ -\theta a_{i\tau}(\boldsymbol{x}) + \boldsymbol{a}_i(\boldsymbol{x})^\top \widehat{\boldsymbol{\xi}}^j - b_i(\boldsymbol{x}) \colon \theta \left\|\boldsymbol{a}_i(\boldsymbol{x})\right\|_\infty + \boldsymbol{a}_i(\boldsymbol{x})^\top \widehat{\boldsymbol{\xi}}^{j'} \leq b_i(\boldsymbol{x}) \right\} \end{array} \right\},$$

which is equivalent to solving $2n$ tractable convex programs and selecting the best one with the largest optimal value. □

## A.3 Proof of Proposition 2

PROPOSITION 2. *Suppose that in Problem (8), set $\mathcal{X}$ is compact and convex. Let*

$$\bar{v}^P_{\infty,i,j,1}(j'|\theta) = \max_{\boldsymbol{x}\in\mathcal{X}} \left\{ \boldsymbol{a}_i(\boldsymbol{x})^\top \left[ \widehat{\boldsymbol{\xi}}^j - \widehat{\boldsymbol{\xi}}^{j'} \right] : \theta \left\|\boldsymbol{a}_i(\boldsymbol{x})\right\|_* + \boldsymbol{a}_i(\boldsymbol{x})^\top \widehat{\boldsymbol{\xi}}^{j'} \leq b_i(\boldsymbol{x}) \right\}, \tag{9a}$$

$$\bar{v}^P_{\infty,i,j,2}(j'|\theta) = \min_{0\leq\alpha\leq1} \max_{\boldsymbol{x}\in\mathcal{X}} \left\{ (\alpha-1)b_i(\boldsymbol{x}) + (1-\alpha)\theta \left\|\boldsymbol{a}_i(\boldsymbol{x})\right\|_* + \sum_{k\in[n]} \left( \widehat{\xi}^j_k - \alpha\widehat{\xi}^{j'}_k \right) a_{ik}(\boldsymbol{x}) \right\}. \tag{9b}$$

*Let $\bar{\eta}_{i,j}(j'|\theta)$ be the minimum between $\bar{v}^P_{\infty,i,j,1}(j'|\theta)$ and $\bar{v}^P_{\infty,i,j,2}(j'|\theta)$, i.e., $\bar{\eta}_{i,j}(j'|\theta) = \min\{\bar{v}^P_{\infty,i,j,1}(j'|\theta), \bar{v}^P_{\infty,i,j,2}(j'|\theta)\}$. Then*

*(i) The optimal value of Problem (8) is upper bounded by $\bar{\eta}_{i,j}(j'|\theta)$, i.e., $\eta_{i,j}(j'|\theta) \le \bar{\eta}_{i,j}(j'|\theta)$; and*
*(ii) When the inner maximization in Problem (9b) admits a unique solution with $\alpha = 0$, the optimal value of Problem (8) is equal to $\bar{\eta}_{i,j}(j'|\theta)$, i.e., $\eta_{i,j}(j'|\theta) = \bar{\eta}_{i,j}(j'|\theta)$.*

*Proof.* Let $\alpha$ be the dual variable of constraint in Problem (8). Then the Lagrangian function for Problem (8) is

$$\mathcal{L}(\boldsymbol{x},\alpha) = (\alpha-1)b_i(\boldsymbol{x}) + (1-\alpha)\theta\,\|\boldsymbol{a}_i(\boldsymbol{x})\|_* + \sum_{k\in[n]} \left(\widehat{\xi}_k^j - \alpha\widehat{\xi}_k^{j'}\right) a_{ik}(\boldsymbol{x}),$$

and its dual problem can be written as

$$\min_{\alpha\ge 0}\max_{\boldsymbol{x}\in\mathcal{X}} \mathcal{L}(\boldsymbol{x},\alpha) = \min\left\{\min_{0\le\alpha\le 1}\max_{\boldsymbol{x}\in\mathcal{X}} \mathcal{L}(\boldsymbol{x},\alpha), \min_{\alpha\ge 1}\max_{\boldsymbol{x}\in\mathcal{X}} \mathcal{L}(\boldsymbol{x},\alpha)\right\}.$$

By weak duality, we have

$$\eta_{i,j}(j'|\theta) \le \min_{\alpha\ge 0}\max_{\boldsymbol{x}\in\mathcal{X}} \mathcal{L}(\boldsymbol{x},\alpha).$$

We split the proof into three steps.
**Step 1.** For $\alpha \ge 1$, we provide an equivalent reformulation of the dual problem we consider, that is, the dual problem we consider is

$$\min_{\alpha\ge 1}\max_{\boldsymbol{x}\in\mathcal{X}} \mathcal{L}(\boldsymbol{x},\alpha) = \min_{\alpha\ge 1}\max_{\boldsymbol{x}\in\mathcal{X}} \left\{(\alpha-1)b_i(\boldsymbol{x}) + (1-\alpha)\theta\,\|\boldsymbol{a}_i(\boldsymbol{x})\|_* + \sum_{k\in[n]} \left(\widehat{\xi}_k^j - \alpha\widehat{\xi}_k^{j'}\right) a_{ik}(\boldsymbol{x})\right\},$$

which is convex in $\alpha$ for a given $\boldsymbol{x} \in \mathcal{X}$ and is concave in $\boldsymbol{x}$ for a given $\alpha \ge 1$. Since set $\mathcal{X}$ is convex and compact, we apply Sion's minimax theorem (see, e.g., Sion 1958) and interchange the min and max operators, that is,

$$\max_{\boldsymbol{x}\in\mathcal{X}}\min_{\alpha\ge 1} \left\{(\alpha-1)b_i(\boldsymbol{x}) + (1-\alpha)\theta\,\|\boldsymbol{a}_i(\boldsymbol{x})\|_* + \sum_{k\in[n]} \left(\widehat{\xi}_k^j - \alpha\widehat{\xi}_k^{j'}\right) a_{ik}(\boldsymbol{x})\right\}.$$

Assuming $\bar{\alpha} = \alpha - 1 \ge 0$, we have

$$\max_{\boldsymbol{x}\in\mathcal{X}}\min_{\bar{\alpha}\ge 0} \left\{\boldsymbol{a}_i(\boldsymbol{x})^\top \left[\widehat{\boldsymbol{\xi}}^j - \widehat{\boldsymbol{\xi}}^{j'}\right] - \bar{\alpha}\left[\theta\,\|\boldsymbol{a}_i(\boldsymbol{x})\|_* + \boldsymbol{a}_i(\boldsymbol{x})^\top \widehat{\boldsymbol{\xi}}^{j'} - b_i(\boldsymbol{x})\right]\right\}.$$

Optimizing over $\bar{\alpha}$, we have

$$\max_{\boldsymbol{x}\in\mathcal{X}} \left\{\boldsymbol{a}_i(\boldsymbol{x})^\top \left[\widehat{\boldsymbol{\xi}}^j - \widehat{\boldsymbol{\xi}}^{j'}\right] : \theta\,\|\boldsymbol{a}_i(\boldsymbol{x})\|_* + \boldsymbol{a}_i(\boldsymbol{x})^\top \widehat{\boldsymbol{\xi}}^{j'} \le b_i(\boldsymbol{x})\right\} := \bar{v}^P_{\infty,i,j,1}(j'|\theta).$$

**Step 2.** For $0 \le \alpha \le 1$, the dual problem we consider is

$$\begin{aligned}\min_{0\le\alpha\le 1}\max_{\boldsymbol{x}\in\mathcal{X}} \mathcal{L}(\boldsymbol{x},\alpha) = &\min_{0\le\alpha\le 1}\max_{\boldsymbol{x}\in\mathcal{X}} \left\{(\alpha-1)b_i(\boldsymbol{x}) + (1-\alpha)\theta\,\|\boldsymbol{a}_i(\boldsymbol{x})\|_* + \sum_{k\in[n]} \left(\widehat{\xi}_k^j - \alpha\widehat{\xi}_k^{j'}\right) a_{ik}(\boldsymbol{x})\right\}\\ &:= \bar{v}^P_{\infty,i,j,2}(j'|\theta).\end{aligned}$$

Thus, Part (i) follows directly by the weak duality.
**Step 3.** It remains to prove that when the inner maximization in Problem (9b) admits a unique solution with $\alpha = 0$, we have $\eta_{i,j}(j'|\theta) = \bar{\eta}_{i,j}(j'|\theta)$. By weak duality, we have $\bar{\eta}_{i,j}(j'|\theta) \ge \eta_{i,j}(j'|\theta)$. Thus, we only need to show that $\bar{\eta}_{i,j}(j'|\theta) \le \eta_{i,j}(j'|\theta)$. These are two cases to discuss.

- **Case 1.** Suppose that there exists one optimal solution $\boldsymbol{x}^*$ of Problem (8) such that the constraint in Problem (8) is not binding, i.e., $\theta\|\boldsymbol{a}_i(\boldsymbol{x}^*)\|_* + \boldsymbol{a}_i(\boldsymbol{x}^*)^\top \widehat{\boldsymbol{\xi}}^{j'} < b_i(\boldsymbol{x}^*)$, then Problem (8) has the same optimal value as

$$\eta_{i,j}(j'|\theta) = \max_{\boldsymbol{x}\in\mathcal{X}} \left\{ \theta \left\| \boldsymbol{a}_i(\boldsymbol{x}) \right\|_* + \boldsymbol{a}_i(\boldsymbol{x})^\top \widehat{\boldsymbol{\xi}}^j - b_i(\boldsymbol{x}) \right\},$$

which is the exactly same problem by setting $\alpha = 0$ in $\bar{v}^P_{\infty,i,j,2}(j'|\theta)$ (9b). Thus, $\eta_{i,j}(j'|\theta) \geq \bar{v}^P_{\infty,i,j,2}(j'|\theta) \geq \bar{\eta}_{i,j}(j'|\theta)$.
- **Case 2.** For any optimal solution of Problem (8), the constraint in Problem (8) is binding, i.e., let $\boldsymbol{x}^*$ be an arbitrary optimal solution to Problem (8), we have $\theta\|\boldsymbol{a}_i(\boldsymbol{x}^*)\|_* + \boldsymbol{a}_i(\boldsymbol{x}^*)^\top \widehat{\boldsymbol{\xi}}^{j'} = b_i(\boldsymbol{x}^*)$. We split the following discussions into two cases.

  **Case 2.1.** Suppose that there exists an optimal solution of Problem (9a) such that the constraint in Problem (9a) is binding. In this subcase, Problem (9a) and Problem (8) coincide. Then, we must have $\eta_{i,j}(j'|\theta) = \bar{v}^P_{\infty,i,j,1}(j'|\theta) \geq \bar{\eta}_{i,j}(j'|\theta)$.

  **Case 2.2.** Suppose that for any optimal solution of Problem (9a), the constraint in Problem (9a) is not binding. In this case, we suppose that $\bar{v}^P_{\infty,i,j,1}(j'|\theta) > \eta_{i,j}(j'|\theta)$ (otherwise, the proof is done). It remains to show that $\bar{v}^P_{\infty,i,j,2}(j'|\theta) \leq \eta_{i,j}(j'|\theta)$. Let us define

$$\bar{v}^P_{\infty,i,j,2}(j'|\theta,\alpha) = \max_{\boldsymbol{x}\in\mathcal{X}} \left\{ (\alpha-1)b_i(\boldsymbol{x}) + (1-\alpha)\theta \left\| \boldsymbol{a}_i(\boldsymbol{x}) \right\|_* + \sum_{k\in[n]} \left( \widehat{\xi}^j_k - \alpha \widehat{\xi}^{j'}_k \right) a_{ik}(\boldsymbol{x}) \right\},$$

and $\bar{v}^P_{\infty,i,j,2}(j'|\theta,\widehat{\alpha}^*) = \bar{v}^P_{\infty,i,j,2}(j'|\theta)$. Obviously, $\bar{v}^P_{\infty,i,j,1}(j'|\theta) = \bar{v}^P_{\infty,i,j,2}(j'|\theta,\alpha=1) > \eta_{i,j}(j'|\theta)$ and $\eta_{i,j}(j'|\theta) \leq \bar{v}^P_{\infty,i,j,2}(j'|\theta,\alpha=0)$. If $\eta_{i,j}(j'|\theta) = \bar{v}^P_{\infty,i,j,2}(j'|\theta,\alpha=0)$, then the proof is done. Hence, suppose that $\eta_{i,j}(j'|\theta) < \bar{v}^P_{\infty,i,j,2}(j'|\theta,\alpha=0)$. The subdifferential of $\bar{v}^P_{\infty,i,j,2}(j'|\theta,\alpha)$ with respect to $\alpha$ is

$$\partial_\alpha \bar{v}^P_{\infty,i,j,2}(j'|\theta,\alpha) = \operatorname{conv}\left\{ b_i(\bar{\boldsymbol{x}}^*) - \theta \left\| \boldsymbol{a}_i(\bar{\boldsymbol{x}}^*) \right\|_* - \boldsymbol{a}_i(\bar{\boldsymbol{x}}^*)^\top \widehat{\boldsymbol{\xi}}^{j'} : \bar{\boldsymbol{x}}^* \in \arg\max_{\boldsymbol{x}\in\mathcal{X}} \mathcal{L}(\boldsymbol{x},\alpha) \right\}.$$

According to our presumption, for any $\bar{\boldsymbol{x}}^* \in \arg\max_{\boldsymbol{x}\in\mathcal{X}} \mathcal{L}(\boldsymbol{x},1)$, we have $b_i(\bar{\boldsymbol{x}}^*) - \theta \left\| \boldsymbol{a}_i(\bar{\boldsymbol{x}}^*) \right\|_* - \boldsymbol{a}_i(\bar{\boldsymbol{x}}^*)^\top \widehat{\boldsymbol{\xi}}^{j'} > 0$. Hence, $\partial_\alpha \bar{v}^P_{\infty,i,j,2}(j'|\theta,1) \subseteq \mathbb{R}_{++}$. On the other hand, when the inner maximization in Problem (9b) admits a unique solution with $\alpha = 0$, we have $\bar{\boldsymbol{x}}^*_0 \in \arg\max_{\boldsymbol{x}\in\mathcal{X}} \bar{v}^P_{\infty,i,j,2}(j'|\theta,0)$, that is,

$$\bar{\boldsymbol{x}}^*_0 \in \arg\max_{\boldsymbol{x}\in\mathcal{X}} \left\{ -b_i(\boldsymbol{x}) + \theta \left\| \boldsymbol{a}_i(\boldsymbol{x}) \right\|_* + \sum_{k\in[n]} \widehat{\xi}^j_k a_{ik}(\boldsymbol{x}) \right\}.$$

The subdifferential of $\bar{v}^P_{\infty,i,j,2}(j'|\theta,\alpha)$ with respect to $\alpha$ at $\alpha=0$ is

$$\partial_\alpha \bar{v}^P_{\infty,i,j,2}(j'|\theta,0) = \left\{ b_i(\bar{\boldsymbol{x}}^*_0) - \theta \left\| \boldsymbol{a}_i(\bar{\boldsymbol{x}}^*_0) \right\|_* - \boldsymbol{a}_i(\bar{\boldsymbol{x}}^*_0)^\top \widehat{\boldsymbol{\xi}}^{j'} \right\}.$$

Given that optimizing over $\alpha$ in $\bar{v}^P_{\infty,i,j,2}(j'|\theta,\alpha)$ is a one-dimension convex optimization problem and the assumption that $\eta_{i,j}(j'|\theta) < \bar{v}^P_{\infty,i,j,2}(j'|\theta,\alpha=0)$, we have

$$\partial_\alpha \bar{v}^P_{\infty,i,j,2}(j'|\theta,0) \subseteq \mathbb{R}_{--}.$$

Therefore, the optimal $\widehat{\alpha}^*$ of $\bar{v}^P_{\infty,i,j,2}(j'|\theta,\alpha)$ must be in the interior of $[0,1]$, i.e., $\widehat{\alpha}^* \in (0,1)$. Then, we have the following necessary and sufficient KKT conditions:

$$0 \in \partial_\alpha \bar{v}^P_{\infty,i,j,2}(j'|\theta,\widehat{\alpha}^*),$$

which implies that there exists an optimal solution $\bar{\boldsymbol{x}}^* \in \arg\max_{\boldsymbol{x}\in\mathcal{X}} \mathcal{L}(\boldsymbol{x},\widehat{\alpha}^*)$ such that $b_i(\bar{\boldsymbol{x}}^*) - \theta \left\| \boldsymbol{a}_i(\bar{\boldsymbol{x}}^*) \right\|_* - \boldsymbol{a}_i(\bar{\boldsymbol{x}}^*)^\top \widehat{\boldsymbol{\xi}}^{j'} = 0$. Thus, we have $\bar{v}^P_{\infty,i,j,2}(j'|\theta,\widehat{\alpha}^*) = \eta_{i,j}(j'|\theta)$.

This completes the proof. □

## A.4 Proof of Proposition 3

PROPOSITION 3. *Suppose $\boldsymbol{a}_i(\boldsymbol{x})=\boldsymbol{x}$, $b_i(\boldsymbol{x})=b_i$ for each $i\in[I]$, the dual norm $\|\cdot\|_*=\|\cdot\|_p$ with $p\in(1,\infty)$, and set $\mathcal{X}=[0,1]^n$. Then*

*(i) The upper bound of Problem (8) can be efficiently computable;*

*(ii) When the empirical samples are nonnegative $\widehat{\boldsymbol{\xi}}^j\geq\mathbf{0}$ for all $j\in[N]$, the upper bound $\bar{\eta}_{i,j}(j'|\theta)$ is exact.*

*Proof.* With the presumptions, Problem (8) reduces to

$$\eta_{i,j}(j'|\theta)=\max_{\boldsymbol{x}\in[0,1]^n}\left\{\theta\left\|\boldsymbol{x}\right\|_p+\boldsymbol{x}^\top\widehat{\boldsymbol{\xi}}^j-b_i:\,\theta\left\|\boldsymbol{x}\right\|_p+\boldsymbol{x}^\top\widehat{\boldsymbol{\xi}}^{j'}\leq b_i\right\}. \tag{30a}$$

According to Proposition 2, we are going to check that $\bar{v}^P_{\infty,i,j,1}(j'|\theta)$ and $\bar{v}^P_{\infty,i,j,2}(j'|\theta)$ can be efficiently computable. We split the proof into three steps.

**Step 1.** For $0\leq\alpha\leq1$, we know

$$\bar{v}^P_{\infty,i,j,2}(j'|\theta)=\min_{0\leq\alpha\leq1}\max_{\boldsymbol{x}\in[0,1]^n}\left\{(\alpha-1)b_i+(1-\alpha)\theta\|\boldsymbol{x}\|_p+\sum_{k\in[n]}\left(\widehat{\xi}^j_k-\alpha\widehat{\xi}^{j'}_k\right)x_k\right\}. \tag{30b}$$

Given that the inner maximization problem is a convex maximization problem, it follows that its solution must lie in the set of extreme points, which can be represented as $\boldsymbol{x}\in\{0,1\}^n$. Therefore, for a given $0\leq\alpha\leq1$, we can recast the inner maximization problem as:

$$\widehat{f}(\alpha)=\max_{\boldsymbol{x}\in\{0,1\}^n}\left\{(\alpha-1)b_i+(1-\alpha)\theta\|\boldsymbol{x}\|_p+\sum_{k\in[n]}\left(\widehat{\xi}^j_k-\alpha\widehat{\xi}^{j'}_k\right)x_k\right\}. \tag{30c}$$

Notice that in this case, when $\alpha=0$, Problem (30c) reduces to

$$\max_{\boldsymbol{x}\in\{0,1\}^n}\left\{-b_i+\theta\|\boldsymbol{x}\|_p+\boldsymbol{x}^\top\widehat{\boldsymbol{\xi}}^j\right\}.$$

Then, we sort $\{\widehat{\xi}^j_k-\alpha\widehat{\xi}^{j'}_k\}_{k\in[n]}$ in the nonincreasing order, i.e., let $\bar{a}_k=\widehat{\xi}^j_k-\alpha\widehat{\xi}^{j'}_k$ for all $k\in[n]$, for a permutation $\sigma$ of $[n]$ such that $\bar{a}_{\sigma_1}\geq\bar{a}_{\sigma_2}\geq\cdots\geq\bar{a}_{\sigma_{\bar{\tau}}}\geq\bar{a}_{\sigma_{\bar{\tau}+1}}\geq\cdots\geq\bar{a}_{\sigma_n}$, where

$$\bar{\tau}=\min_{j\in[n]}\left\{j:\,(1-\alpha)\theta\left[j^{1/p}-(j-1)^{1/p}\right]+\bar{a}_{\sigma_j}<0\right\}.$$

Thus, an optimal solution of Problem (30c) is

$$x^*_{\sigma_k}=\begin{cases}1, & \forall k\in[\bar{\tau}],\\ 0 & \forall k\in[\bar{\tau}+1,n],\end{cases}$$

and the objective value of Problem (30c) is

$$\widehat{f}(\alpha)=(\alpha-1)b_i+(1-\alpha)\theta\bar{\tau}^{1/p}+\sum_{j\in[\bar{\tau}]}\bar{a}_{\sigma_j}. \tag{30d}$$

For the outer minimization problem over $\alpha$, we apply the golden section search to find the optimal $\alpha$ efficiently, which is detailed in Algorithm 4.

**Step 2.** Given that optimizing over $\alpha$ in the outer minimization of Problem (30a) is a one-dimension convex optimization problem, we can apply the golden section search Algorithm 4 to find the optimal $\alpha$ efficiently. Thus, it remains to show that inner maximization can be solved efficiently. For any given

$\alpha > 1$, by introducing Lagrangian multipliers for the constraints $\theta|\boldsymbol{x}|_p + \boldsymbol{x}^\top \widehat{\boldsymbol{\xi}}^{j'} \leq b_i$ and $\boldsymbol{x} \in [0,1]^n$, the Lagrangian function for Problem (30a) can be formulated as:

$$\mathcal{L}(\boldsymbol{x},\alpha,\boldsymbol{\gamma},\boldsymbol{\mu}) = (\alpha-1)b_i + (1-\alpha)\theta\|\boldsymbol{x}\|_p + \mathbf{e}^\top\boldsymbol{\gamma} + \sum_{k\in[n]}\left(\widehat{\xi}_k^j - \alpha\widehat{\xi}_k^{j'} - \gamma_k + \mu_k\right)x_k, \tag{30e}$$

and its corresponding dual problem can be written as

$$\min_{\substack{\alpha>1,\boldsymbol{\gamma}\geq\mathbf{0},\\ \boldsymbol{\mu}\geq\mathbf{0}}} \max_{\boldsymbol{x}} \left\{(\alpha-1)b_i + (1-\alpha)\theta\|\boldsymbol{x}\|_p + \mathbf{e}^\top\boldsymbol{\gamma} + \sum_{k\in[n]}\left(\widehat{\xi}_k^j - \alpha\widehat{\xi}_k^{j'} - \gamma_k + \mu_k\right)x_k\right\}. \tag{30f}$$

Letting $1/p + 1/\bar{q} = 1$, according to Hölder's inequality, the dual problem (30f) is equivalent to

$$\widehat{f}(\alpha) = \min_{\boldsymbol{\gamma}\geq\mathbf{0},\boldsymbol{\mu}\geq\mathbf{0}} \left\{(\alpha-1)b_i + \mathbf{e}^\top\boldsymbol{\gamma} \colon \left[\sum_{k\in[n]}\left|\widehat{\xi}_k^j - \alpha\widehat{\xi}_k^{j'} - \gamma_k + \mu_k\right|^{\bar{q}}\right]^{\frac{1}{\bar{q}}} \leq (\alpha-1)\theta\right\}. \tag{30g}$$

Here, we define $\bar{a}_k = \widehat{\xi}_k^j - \alpha\widehat{\xi}_k^{j'}$ for each $k\in[n]$ and $\mathcal{S}_2 = \{k\colon \bar{a}_k > 0, \forall k\in[n]\}$. We sort the elements in $\bar{\boldsymbol{a}}$ with a nonincreasing order, i.e., $\bar{a}_{\sigma_1} \geq \bar{a}_{\sigma_2} \geq \cdots \geq \bar{a}_{\sigma_{|\mathcal{S}_2|-1}} \geq \bar{a}_{\sigma_{|\mathcal{S}_2|}}$ and let $\sigma_{t_1}$ be the largest scenario with $t_1 \leq |S_2|$ such that

$$\bar{a}_{\sigma_{t_1}} - \frac{1}{t_1^{\frac{1}{\bar{q}}}}\left[(\alpha-1)^{\bar{q}}\theta^{\bar{q}} - \sum_{j=\sigma_{t_1+1}}^{\sigma_{|\mathcal{S}_2|}} \bar{a}_j^{\bar{q}}\right]^{\frac{1}{\bar{q}}} > 0.$$

We further define $\mathcal{S}_3 \subseteq \mathcal{S}_2$ with $\mathcal{S}_3 = \{\sigma_1,\sigma_2,\ldots,\sigma_{t_1}\}$ and $|\mathcal{S}_3| = t_1$. Then, we construct a primal feasible solution $\bar{\boldsymbol{x}}$ as

$$\bar{x}_i = 1, \forall i\in\mathcal{S}_3, \bar{x}_i = \frac{\bar{a}_i^{\frac{\bar{q}}{p}} t_1^{\frac{1}{p}}}{[(\alpha-1)\theta]^{\frac{\bar{q}}{p}}\left[1 - \sum_{j=\sigma_{t_1+1}}^{\sigma_{|\mathcal{S}_2|}} \frac{\bar{a}_j^{\bar{q}}}{(\alpha-1)^{\bar{q}}\theta^{\bar{q}}}\right]^{\frac{1}{p}}}, \forall i\in\mathcal{S}_2\setminus\mathcal{S}_3, \bar{x}_i = 0, \forall i\in[n]\setminus\mathcal{S}_2,$$

while a dual feasible solution $(\bar{\boldsymbol{\gamma}},\bar{\boldsymbol{\mu}})$ is

$$\bar{\gamma}_i = 0, \forall i\in[n]\setminus\mathcal{S}_3, \bar{\gamma}_i = \bar{a}_i - \frac{1}{t_1^{\frac{1}{\bar{q}}}}\left[(\alpha-1)^{\bar{q}}\theta^{\bar{q}} - \sum_{j=\sigma_{t_1+1}}^{\sigma_{|\mathcal{S}_2|}} \bar{a}_j^{\bar{q}}\right]^{\frac{1}{\bar{q}}}, \forall i\in\mathcal{S}_3,$$
$$\bar{\mu}_i = 0, \forall i\in\mathcal{S}_2, \bar{\mu}_i = -\bar{a}_i, \forall i\in[n]\setminus\mathcal{S}_2.$$

In this way, both solutions yield the same objective value as

$$(\alpha-1)b_i + \sum_{i\in\mathcal{S}_3}\left[\bar{a}_i - |\mathcal{S}_3|^{1-\frac{1}{\bar{q}}}\left[(\alpha-1)^{\bar{q}}\theta^{\bar{q}} - \sum_{j\in\mathcal{S}_2\setminus\mathcal{S}_3}\bar{a}_j^{\bar{q}}\right]^{\frac{1}{\bar{q}}}\right].$$

Therefore, according to the weak duality, both primal and dual solutions are optimal. We can conclude that for a given $\alpha > 1$, the optimal objective value of Problem (30g) is

$$\widehat{f}(\alpha) = (\alpha-1)b_i + \sum_{i\in\mathcal{S}_3}\left[\bar{a}_i - |\mathcal{S}_3|^{1-\frac{1}{\bar{q}}}\left[(\alpha-1)^{\bar{q}}\theta^{\bar{q}} - \sum_{j\in\mathcal{S}_2\setminus\mathcal{S}_3}\bar{a}_j^{\bar{q}}\right]^{\frac{1}{\bar{q}}}\right]. \tag{30h}$$

**Step 3.** For Part(ii), with the nonnegative empirical samples $\widehat{\boldsymbol{\xi}}^j \geq \mathbf{0}$ for all $j\in[N]$, $\boldsymbol{x} = \mathbf{e}$ is the unique solution of Problem (30c) when $\alpha = 0$, which satisfies the uniqueness assumption in Part (ii) of Proposition 2. Therefore, by calculating the values of $\bar{v}^P_{\infty,i,j,1}(j'|\theta)$ and $\bar{v}^P_{\infty,i,j,2}(j'|\theta)$, we can determine the exact optimal value of $\eta_{i,j}(j'|\theta)$. This completes the proof. □

**Algorithm 4** Golden Section Search Method

**Input:** Let $\underline{\alpha}$ and $\overline{\alpha}$ denote the lower and upper bounds of the optimal value of $\alpha$, respectively, and let $\delta_1$ denote the stopping tolerance parameter
**while** $\overline{\alpha}-\underline{\alpha}>\delta_1$ **do**
  $\alpha_1=((\sqrt{5}-1)\underline{\alpha}+(3-\sqrt{5})\overline{\alpha})/2$, $\alpha_2=((3-\sqrt{5})\underline{\alpha}+(\sqrt{5}-1)\overline{\alpha})/2$
  Calculate $\widehat{f}(\alpha_1)$ and $\widehat{f}(\alpha_2)$ using (30d) for the cases where $0\le\alpha_1\le 1$ or $0\le\alpha_2\le 1$, and using (30h) for the cases where $\alpha_1\ge 1$ or $\alpha_2\ge 1$
  **if** $\widehat{f}(\alpha_1)\ge\widehat{f}(\alpha_2)$ set $\underline{\alpha}=\alpha_1$; **else** set $\overline{\alpha}=\alpha_2$
**end while**
**Output:** $\alpha^*=(\overline{\alpha}+\underline{\alpha})/2$

### A.5 Proof of Proposition 4

PROPOSITION 4. *Suppose $\boldsymbol{a}_i(\boldsymbol{x})=\boldsymbol{x}$, $b_i(\boldsymbol{x})=b_i$ for each $i\in[I]$, the dual norm $\|\cdot\|_*=\|\cdot\|_1$, and set $\mathcal{X}=[0,1]^n$. Problem (8) can be efficiently computable.*

*Proof.* With the presumptions, Problem (8) can be written as

$$\eta_{i,j}(j'|\theta)=\max_{\boldsymbol{x}\in[0,1]^n}\left\{\theta\sum_{k\in[n]}x_k+\boldsymbol{x}^\top\widehat{\boldsymbol{\xi}}^j-b_i:\theta\sum_{k\in[n]}x_k+\boldsymbol{x}^\top\widehat{\boldsymbol{\xi}}^{j'}\le b_i\right\}. \tag{31}$$

We first define the following four sets:

$$\mathcal{S}_{+,+}=\left\{k\in[n]:\widehat{\xi}_k^j+\theta\ge 0,\widehat{\xi}_k^{j'}+\theta\ge 0\right\},\mathcal{S}_{-,-}=\left\{k\in[n]:\widehat{\xi}_k^j+\theta<0,\widehat{\xi}_k^{j'}+\theta<0\right\},$$
$$\mathcal{S}_{+,-}=\left\{k\in[n]:\widehat{\xi}_k^j+\theta\ge 0,\widehat{\xi}_k^{j'}+\theta<0\right\},\mathcal{S}_{-,-}=\left\{k\in[n]:\widehat{\xi}_k^j+\theta<0,\widehat{\xi}_k^{j'}+\theta\ge 0\right\}.$$

Based on the principle of monotonicity, there exists an optimal solution $\boldsymbol{x}^*$ for Problem (31) that must possess the following property:

$$x_k^*=1,\forall k\in\mathcal{S}_{+,-},\text{ and } x_k^*=0,\forall k\in\mathcal{S}_{-,-}.$$

Then, Problem (31) can be reduced to

$$\begin{aligned}\max_{\boldsymbol{x}\in[0,1]^{|\mathcal{S}_{+,+}\cup\mathcal{S}_{-,-}|}}\quad&\sum_{k\in\mathcal{S}_{+,+}\cup\mathcal{S}_{-,-}}\left(\widehat{\xi}_k^j+\theta\right)x_k+\sum_{k\in\mathcal{S}_{+,-}}\left(\widehat{\xi}_k^j+\theta\right)-b_i,\\ \text{s.t.}\quad&\sum_{k\in\mathcal{S}_{+,+}\cup\mathcal{S}_{-,-}}\left(\widehat{\xi}_k^{j'}+\theta\right)x_k\le b_i-\sum_{k\in\mathcal{S}_{+,-}}\left(\widehat{\xi}_k^{j'}+\theta\right).\end{aligned}$$

Since the coefficients in set $\mathcal{S}_{-,-}$ are negative, we change the variables as $x_i=1-x_i$ for each $i\in\mathcal{S}_{-,-}$. In this way, Problem (31) is equivalent to

$$\begin{aligned}\max_{\boldsymbol{x}\in[0,1]^{|\mathcal{S}_{+,+}\cup\mathcal{S}_{-,-}|}}\quad&\sum_{k\in\mathcal{S}_{+,+}\cup\mathcal{S}_{-,-}}\left(\widehat{\xi}_k^j+\theta\right)x_k+\sum_{k\in\mathcal{S}_{+,-}}\left(\widehat{\xi}_k^j+\theta\right)-\sum_{k\in\mathcal{S}_{-,-}}\left|\widehat{\xi}_k^j+\theta\right|-b_i,\\ \text{s.t.}\quad&\sum_{k\in\mathcal{S}_{+,+}\cup\mathcal{S}_{-,-}}\left|\widehat{\xi}_k^{j'}+\theta\right|x_k+\le b_i+\sum_{k\in\mathcal{S}_{-,-}}\left|\widehat{\xi}_k^{j'}+\theta\right|+\sum_{k\in\mathcal{S}_{+,-}}\left|\widehat{\xi}_k^{j'}+\theta\right|.\end{aligned}$$

We then compute the ratio $(|\widehat{\xi}_k^j+\theta|)/(|\widehat{\xi}_k^{j'}+\theta|)$ for each $k\in\mathcal{S}_{+,+}\cup\mathcal{S}_{-,-}$ and then sort these values in nonincreasing order, i.e., for a permutation $\sigma$ of set $\mathcal{S}_{+,+}\cup\mathcal{S}_{-,-}$, we have

$$\frac{|\widehat{\xi}_{\sigma_1}^j+\theta|}{|\widehat{\xi}_{\sigma_1}^{j'}+\theta|}\ge\frac{|\widehat{\xi}_{\sigma_2}^j+\theta|}{|\widehat{\xi}_{\sigma_2}^{j'}+\theta|}\ge\cdots\ge\frac{|\widehat{\xi}_{\sigma_{|\mathcal{S}_{+,+}\cup\mathcal{S}_{-,-}|}}^j+\theta|}{|\widehat{\xi}_{\sigma_{|\mathcal{S}_{+,+}\cup\mathcal{S}_{-,-}|}}^{j'}+\theta|}.$$

Therefore, with $\boldsymbol{x} \in [0,1]^{|\mathcal{S}_{+,+}\cup\mathcal{S}_{-,-}|}$, an optimal solution of $\boldsymbol{x}^*$ is

$$x^*_{\sigma_k} = \begin{cases} 1, & \forall k \in [\ell], \\ \frac{1}{\left|\theta+\widehat{\xi}^{j'}_{\sigma_{\ell+1}}\right|}\left[b_i + \sum_{k\in\mathcal{S}_{-,-}\cup\mathcal{S}_{+,-}}\left|\widehat{\xi}^j_k+\theta\right| - \sum_{k\in[\ell]}\left|\theta+\widehat{\xi}^{j'}_{\sigma_k}\right|\right], & k=\ell+1, \\ 0, & \forall k\in[\ell+2, |\mathcal{S}_{+,+}\cup\mathcal{S}_{-,-}|], \end{cases}$$

where $\sum_{k\in[\ell]}|\theta+\widehat{\xi}^{j'}_{\sigma_k}| \le b_i + \sum_{k\in\mathcal{S}_{-,-}\cup\mathcal{S}_{+,-}}|\widehat{\xi}^j_k+\theta| < \sum_{k\in[\ell+1]}|\theta+\widehat{\xi}^{j'}_{\sigma_k}|$.

This completes the proof. □

## Appendix B An Improved Upper Bound of DRCCP (10) under Type 1−Wasserstein Ambiguity Set

Let $(\boldsymbol{x}^L, \lambda^L, \gamma^L, \boldsymbol{s}^L, \boldsymbol{y}^L, \boldsymbol{z}^L, \widehat{\boldsymbol{z}}^L)$ be an optimal solution of the improved VaR lower bound (16), then we use $\widehat{\boldsymbol{z}}^L$ information to DRCCP (10). This integration allows us to obtain an improved upper bound, that is,

$$\bar{v}^U_1 = \min_{\substack{\boldsymbol{x}\in\mathcal{X},\lambda,\gamma,\\ \boldsymbol{s},\boldsymbol{y},\boldsymbol{z}\in\{0,1\}^N}} \left\{ \boldsymbol{c}^\top\boldsymbol{x}: \begin{array}{l} \text{(10b)-(10c), (10f)-(10h), (10d)-(10e),} \\ \widehat{z}^L_j \le z_j, \forall j\in[N] \end{array} \right\}. \tag{32}$$

In the numerical implementation, we choose the minimum of improved upper bound (32) and ALSO-X# as the best upper bound of DRCCP (10). And we use the solution from the improved upper bound (32) to warm-start the big-M method.

## Appendix C Numerical Comparisons of Different Strengthen Techniques with Ho-Nguyen et al. (2023) under Type 1−Wasserstein Ambiguity Set

To compare the effectiveness of the coefficient strengthening, we implement the big-M coefficient strengthening procedure in section 4.2 of Ho-Nguyen et al. (2023). We consider the same setting in Case III of Section 7 and report the detailed numerical comparisons in Table 8 and Table 9. For all the instances, when evaluating the gap or the running time, we find that the big-M coefficient strengthening method in Section 4.3 can outperform that of the one in section 4.2 of Ho-Nguyen et al. (2023). However, we remark that both strengthening methods may not be able to close the gap in most reported instances, while our "Big-M & Fixing & Strengthening" approach can solve all instances and significantly reduce the total running time.

## Appendix D An Improved VaR Lower Bound under Type $q$−Wasserstein Ambiguity Set with $q\in(1,\infty)$

Similar to the discussions in the previous sections, inner approximation (21) can be written as

$$\widehat{v}_q = \min_{\boldsymbol{x}\in\mathcal{X},\lambda,\gamma,\boldsymbol{s},\boldsymbol{y},\boldsymbol{z}\in\{0,1\}^N} \quad \boldsymbol{c}^\top\boldsymbol{x}, \tag{33a}$$

$$\text{s.t.} \quad \theta\varepsilon^{1-\frac{1}{q}}\lambda - \varepsilon\gamma \le \frac{1}{N}\sum_{j\in[N]} y_j, \tag{33b}$$

$$y_j + \gamma \le s_j, \forall j\in[N], \tag{33c}$$

$$s_j \le b_i(\boldsymbol{x}) - \boldsymbol{a}_i(\boldsymbol{x})^\top\widehat{\boldsymbol{\xi}}^j + M^1_{i,j}(1-z_j), \forall i\in[I], j\in[N], \tag{33d}$$

$$s_j \le M^2_{i,j} z_j, \forall i\in[I], j\in[N], \tag{33e}$$

$$\|\boldsymbol{a}_1(\boldsymbol{x})\|_* \le \lambda, \forall i\in[I], \tag{33f}$$

**Table 8** Numerical Comparisons of Big-M Coefficient Strengthening Methods for a DRCCP under Type 1−Wasserstein Ambiguity Set with Instances *1-7-1-500* from Song et al. (2014) and $N = 500$

| $\varepsilon$ | $\theta$ | Case | Big-M & Fixing & Strengthen in Case III of Section 7 | | Big-M & Strengthening in Case III of Section 7 | | | | Big-M & Strengthening from Section 4.2 of Ho-Nguyen et al. (2023) | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | | | GAP | Total Time (s) | GAP | Time (s) | | | GAP | Time (s) | | |
| | | | | | | S | Solver | Total | | S | Solver | Total |
| 0.1 | 0.1 | 1 | 0.00% | 552.51 | 0.00% | 259.68 | 9007.25 | 9266.92 | 0.00% | 2.82 | 9760.62 | 9763.44 |
| | | 2 | 0.00% | 556.64 | 0.08% | 261.48 | 14138.52 | 14400.00 | 0.10% | 2.87 | 14397.13 | 14400.00 |
| | | 3 | 0.00% | 568.84 | 0.00% | 265.26 | 8364.19 | 8629.45 | 0.00% | 2.85 | 8673.92 | 8676.77 |
| | | 4 | 0.00% | 543.85 | 0.00% | 262.52 | 7550.28 | 7812.80 | 0.03% | 2.88 | 14397.12 | 14400.00 |
| | | 5 | 0.00% | 539.47 | 0.21% | 259.86 | 14140.14 | 14400.00 | 0.24% | 2.93 | 14397.07 | 14400.00 |
| 0.1 | 0.2 | 1 | 0.00% | 563.34 | 0.23% | 274.32 | 14125.68 | 14400.00 | 0.26% | 2.87 | 14397.13 | 14400.00 |
| | | 2 | 0.00% | 976.10 | 0.39% | 258.33 | 14141.67 | 14400.00 | 0.39% | 2.95 | 14397.05 | 14400.00 |
| | | 3 | 0.00% | 580.09 | 0.27% | 261.94 | 14138.06 | 14400.00 | 0.27% | 2.90 | 14397.10 | 14400.00 |
| | | 4 | 0.00% | 561.19 | 0.17% | 266.36 | 14133.64 | 14400.00 | 0.18% | 2.91 | 14397.09 | 14400.00 |
| | | 5 | 0.00% | 611.11 | 0.33% | 275.13 | 14124.87 | 14400.00 | 0.41% | 2.87 | 14397.13 | 14400.00 |
| 0.2 | 0.1 | 1 | 0.00% | 789.39 | 0.35% | 262.58 | 14137.42 | 14400.00 | 0.36% | 2.99 | 14397.01 | 14400.00 |
| | | 2 | 0.00% | 1368.66 | 0.42% | 261.92 | 14138.08 | 14400.00 | 0.49% | 2.94 | 14397.06 | 14400.00 |
| | | 3 | 0.00% | 1034.36 | 0.42% | 265.94 | 14134.06 | 14400.00 | 0.44% | 2.85 | 14397.15 | 14400.00 |
| | | 4 | 0.00% | 802.69 | 0.48% | 262.21 | 14137.79 | 14400.00 | 0.52% | 2.88 | 14397.12 | 14400.00 |
| | | 5 | 0.00% | 745.86 | 0.19% | 266.94 | 14133.07 | 14400.00 | 0.20% | 2.92 | 14397.08 | 14400.00 |
| 0.2 | 0.2 | 1 | 0.00% | 1270.14 | 0.59% | 258.22 | 14141.78 | 14400.00 | 0.59% | 2.91 | 14397.09 | 14400.00 |
| | | 2 | 0.00% | 4342.85 | 0.81% | 261.01 | 14139.00 | 14400.00 | 0.81% | 2.90 | 14397.10 | 14400.00 |
| | | 3 | 0.00% | 11239.78 | 1.19% | 257.82 | 14142.18 | 14400.00 | 1.20% | 2.88 | 14397.12 | 14400.00 |
| | | 4 | 0.00% | 1459.50 | 0.72% | 258.79 | 14141.21 | 14400.00 | 0.80% | 2.84 | 14397.16 | 14400.00 |
| | | 5 | 0.00% | 919.65 | 0.45% | 259.01 | 14140.99 | 14400.00 | 0.52% | 2.88 | 14397.12 | 14400.00 |

**Table 9** Numerical Comparisons of Big-M Coefficient Strengthening Methods for a DRCCP under Type 1−Wasserstein Ambiguity Set with Instances *1-7-5-500* from Song et al. (2014) and $N = 500$

| $\varepsilon$ | $\theta$ | Case | Big-M & Fixing & Strengthen in Case III of Section 7 | | Big-M & Strengthening in Case III of Section 7 | | | | Big-M & Strengthening from Section 4.2 of Ho-Nguyen et al. (2023) | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | | | GAP | Total Time (s) | GAP | Time (s) | | | GAP | Time (s) | | |
| | | | | | | S | Solver | Total | | S | Solver | Total |
| 0.1 | 0.1 | 1 | 0.00% | 484.32 | 0.18% | 258.23 | 14141.77 | 14400.00 | 0.19% | 2.84 | 14397.16 | 14400.00 |
| | | 2 | 0.00% | 473.94 | 0.16% | 258.27 | 14141.73 | 14400.00 | 0.19% | 2.84 | 14397.16 | 14400.00 |
| | | 3 | 0.00% | 489.78 | 0.08% | 261.58 | 14138.42 | 14400.00 | 0.09% | 2.86 | 14397.14 | 14400.00 |
| | | 4 | 0.00% | 526.03 | 0.00% | 264.27 | 9837.24 | 10101.51 | 0.00% | 2.83 | 11384.92 | 11387.75 |
| | | 5 | 0.00% | 495.94 | 0.02% | 264.54 | 14135.46 | 14400.00 | 0.02% | 2.85 | 14397.15 | 14400.00 |
| 0.1 | 0.2 | 1 | 0.00% | 498.61 | 0.23% | 265.56 | 14134.44 | 14400.00 | 0.24% | 2.84 | 14397.16 | 14400.00 |
| | | 2 | 0.00% | 505.93 | 0.19% | 271.49 | 14128.51 | 14400.00 | 0.36% | 2.86 | 14397.14 | 14400.00 |
| | | 3 | 0.00% | 486.02 | 0.12% | 261.27 | 14138.73 | 14400.00 | 0.15% | 2.80 | 14397.20 | 14400.00 |
| | | 4 | 0.00% | 547.57 | 0.22% | 270.09 | 14129.91 | 14400.00 | 0.22% | 2.81 | 14397.19 | 14400.00 |
| | | 5 | 0.00% | 584.00 | 0.08% | 265.96 | 14134.04 | 14400.00 | 0.09% | 2.84 | 14397.16 | 14400.00 |
| 0.2 | 0.1 | 1 | 0.00% | 761.75 | 0.24% | 257.38 | 14142.62 | 14400.00 | 0.25% | 2.85 | 14397.15 | 14400.00 |
| | | 2 | 0.00% | 710.85 | 0.19% | 264.99 | 14135.01 | 14400.00 | 0.29% | 2.82 | 14397.18 | 14400.00 |
| | | 3 | 0.00% | 659.51 | 0.18% | 265.56 | 14134.45 | 14400.00 | 0.18% | 2.84 | 14397.16 | 14400.00 |
| | | 4 | 0.00% | 757.81 | 0.19% | 268.32 | 14131.68 | 14400.00 | 0.22% | 2.84 | 14397.16 | 14400.00 |
| | | 5 | 0.00% | 644.73 | 0.34% | 259.83 | 14140.17 | 14400.00 | 0.35% | 2.83 | 14397.17 | 14400.00 |
| 0.2 | 0.2 | 1 | 0.00% | 3445.53 | 0.28% | 264.13 | 14135.87 | 14400.00 | 0.28% | 2.81 | 14397.19 | 14400.00 |
| | | 2 | 0.00% | 735.76 | 0.41% | 273.49 | 14126.51 | 14400.00 | 0.41% | 2.83 | 14397.17 | 14400.00 |
| | | 3 | 0.00% | 851.85 | 0.26% | 279.46 | 14120.54 | 14400.00 | 0.26% | 2.80 | 14397.20 | 14400.00 |
| | | 4 | 0.00% | 1112.11 | 0.31% | 281.83 | 14118.17 | 14400.00 | 0.32% | 2.86 | 14397.14 | 14400.00 |
| | | 5 | 0.00% | 840.94 | 0.44% | 269.85 | 14130.15 | 14400.00 | 0.45% | 2.85 | 14397.15 | 14400.00 |

$$\sum_{j\in[N]} z_j \geq N - \lfloor N\varepsilon \rfloor + \mathbb{I}\{\theta > 0\}, \tag{33g}$$

$$\lambda > 0, \gamma \geq 0, s_j \geq 0, \forall j \in [N], y_j \leq 0, \forall j \in [N]. \tag{33h}$$

It is worth noting that the lower bound of DRCCP (19) can be found by using the improved VaR lower bound, that is,

$$\bar{v}_q^L = \min_{\substack{\boldsymbol{x}\in\mathcal{X},\lambda,\gamma,\\ \boldsymbol{s},\boldsymbol{y},\boldsymbol{z},\widehat{\boldsymbol{z}}}} \left\{ \boldsymbol{c}^\top \boldsymbol{x} : \begin{array}{l} \theta\varepsilon^{-\frac{1}{q}} \left\|\boldsymbol{a}_i(\boldsymbol{x})\right\|_* + \boldsymbol{a}_i(\boldsymbol{x})^\top \widehat{\boldsymbol{\xi}}^j \leq b_i(\boldsymbol{x}) + M_{i,j}^{\mathrm{VaR},q}(1-\widehat{z}_j), \forall i \in [I], j \in [N], \\ \sum_{j\in[N]} \widehat{z}_j \geq N - \lfloor N\varepsilon \rfloor, \widehat{\boldsymbol{z}} \in \{0,1\}^N, \boldsymbol{z} \in [0,1]^N, \widehat{z}_j \leq z_j, \forall j \in [N], \\ (33b)\text{-}(33h) \end{array} \right\},$$

where for each $i \in [I], j \in [N]$, we have

$$M_{i,j}^{\mathrm{VaR},q} \geq \max_{\boldsymbol{x}\in\mathcal{X}} \left\{ \theta\varepsilon^{-\frac{1}{q}} \left\|\boldsymbol{a}_i(\boldsymbol{x})\right\|_* + \boldsymbol{a}_i(\boldsymbol{x})^\top \widehat{\boldsymbol{\xi}}^j - b_i(\boldsymbol{x}) \right\}.$$