On deterministic reformulations of distributionally robust joint chance constrained optimization problems

Weijun Xie* and Shabbir Ahmed†

School of Industrial & Systems Engineering
Georgia Institute of Technology, Atlanta, GA 30332

May 8, 2017

Abstract

A joint chance constrained optimization problem involves multiple uncertain constraints, i.e., constraints with stochastic parameters, that are jointly required to be satisfied with probability exceeding a prespecified threshold. In a distributionally robust joint chance constrained optimization problem (DRCCP), the joint chance constraint is required to hold for all probability distributions of the stochastic parameters from a given ambiguity set. In this work, we consider DRCCP involving convex nonlinear uncertain constraints and an ambiguity set specified by convex moment constraints. We investigate deterministic reformulations of such problems and conditions under which such deterministic reformulations are convex. In particular we show that a DRCCP can be reformulated as a convex program if one the following conditions hold: (i) there is a single uncertain constraint, (ii) the ambiguity set is defined by a single moment constraint, (iii) the ambiguity set is defined by linear moment constraints, and (iv) the uncertain and moment constraints are positively homogeneous with respect to uncertain parameters. We further show that if the decision variables are binary and the uncertain constraints are linear then a DRCCP can be reformulated as a deterministic mixed integer convex program. Finally, we present a numerical study to illustrate that the proposed mixed integer convex reformulation can be solved efficiently by existing solvers.

1 Introduction

1.1 Problem Setting

We consider a distributionally robust chance constrained program (DRCCP) of the form (c.f. [4, 10, 12, 24]):

$$
\begin{aligned}
v^* = \min\ & c^\top x, && \text{(1a)}\\
\text{s.t. }\ & x \in S, && \text{(1b)}\\
& \inf_{\mathbb{P}\in\mathcal{P}} \mathbb{P}[\boldsymbol{\xi} : F(x,\boldsymbol{\xi}) \geq 0] \geq 1-\epsilon. && \text{(1c)}
\end{aligned}
$$

*wxie33@gatech.edu
†sahmed@isye.gatech.edu

where $x \in \mathbb{R}^n$ is a decision vector; the vector $c \in \mathbb{R}^n$ denotes the objective coefficients; the set $S \subseteq \mathbb{R}^n$ denotes deterministic constraints on x; the random vector $\boldsymbol{\xi}$ supported on $\Xi \subset \mathbb{R}^m$ denotes uncertain constraint coefficients with a realization denoted by ξ; the mapping $F(x, \xi) := (f_1(x, \xi), \ldots, f_I(x, \xi))^\top$ with $f_i(x, \xi) : \mathbb{R}^n \times \Xi \to \mathbb{R}$ for all $i \in [I] := \{1, \ldots, I\}$ defines a set of uncertain constraints on x; the ambiguity set $\mathcal{P}$ denotes a set of probability measures $\mathbb{P}$ on the space Ξ with a sigma algebra $\mathcal{F}$; and $\epsilon \in (0, 1)$ denotes a risk tolerance. In (1) we seek a decision vector x to minimize a linear objective $c^\top x$ subject to a set of deterministic constraints defined by S, and a chance constraint (1c) that is required to hold for any probability distribution from the ambiguity set $\mathcal{P}$ with a probability of $1 - \epsilon$. Note that when $|I| = 1$ the constraint (1c) involves a *single* chance constraint and if $|I| \geq 2$ it involves a *joint* chance constraint.

The primary difficulty of (1) is due to the distributionally robust chance constraint (1c). Let us denote the feasible region induced by (1c) as

$$Z_{\mathrm{D}} := \left\{ x \in \mathbb{R}^n : \inf_{\mathbb{P} \in \mathcal{P}} \mathbb{P}[\boldsymbol{\xi} : F(x, \boldsymbol{\xi}) \geq 0] \geq 1 - \epsilon \right\}. \tag{2}$$

In this paper we study deterministic reformulations of the set Z_{D} and its convexity properties. Our study is restricted to the convex, moment constrained setting (cf. [19, 20]), i.e. under the following assumptions.

(A1) Each function $f_i(x, \xi)$ in the mapping $F(x, \xi) := (f_1(x, \xi), \ldots, f_I(x, \xi))^\top$ is concave in x for any fixed ξ, and is convex in ξ for any fixed x.

(A2) The random vector $\boldsymbol{\xi}$ is supported on a nonempty closed convex set $\Xi \subseteq \mathbb{R}^m$.

(A3) The ambiguity set $\mathcal{P}$ is nonempty and is defined by moment constraints:

$$\mathcal{P} = \{\mathbb{P} \in \mathcal{P}_0\,(\Xi) : \mathbb{E}_{\mathbb{P}}[\phi_t(\boldsymbol{\xi})] = g_t, t \in \mathcal{T}_1, \mathbb{E}_{\mathbb{P}}[\phi_t(\boldsymbol{\xi})] \geq g_t, t \in \mathcal{T}_2, \} \tag{3}$$

where $\mathcal{P}_0\,(\Xi)$ denotes the set of all of probability measures on Ξ with a sigma algebra $\mathcal{F}$, and for each $t \in \mathcal{T}_1 \cup \mathcal{T}_2$, the moment function $\phi_t : \Xi \to \mathbb{R}$ is a real valued continuous function and g_t is a scalar. Furthermore, for each $t \in \mathcal{T}_1$, the function $\phi_t(\xi)$ is linear, and for each $t \in \mathcal{T}_2$, the function $\phi_t(\xi)$ is concave.

1.2 Contributions

Even under the above convexity assumptions the set Z_{D} is nonconvex in general, making (1) a difficult optimization problem. Moreover it is not described by explicit functions, and so is not suitable for direct optimization as a mathematical program. In this paper we first provide a deterministic approximation of Z_{D} that is nearly tight and then identify a variety of settings under which Z_{D} is convex. The main results of this paper are summarized next.

1. We propose a deterministic conservative approximation of Z_{D}, which is in general nonconvex and can be formulated as an optimization problem involving biconvex constraints.

2. If there is a single uncertain constraint, i.e. $|I| = 1$, we prove that the proposed deterministic approximation is exact and reduces to a tractable convex program. This result is a generalization of existing works (e.g., [4, 24, 26]) to arbitrary convex ambiguity sets rather than those involving only first and second moment constraints.

3. We prove that if the ambiguity set $\mathcal{P}$ contains only one moment inequality, i.e. $|\mathcal{T}_1| = 0$ and $|\mathcal{T}_2| = 1$, then Z_{D} is a tractable convex program; and if the ambiguity set contains only one

moment linear equality, i.e. $|\mathcal{T}_1| = 1$ and $|\mathcal{T}_2| = 0$, then Z_{D} is equivalent to the disjunction of two tractable convex programs.

4. We prove that if $\Xi = \mathbb{R}^m$ and the moment functions $\{\phi_t(\xi)\}_{t\in\mathcal{T}_1\cup\mathcal{T}_2}$ are linear, then Z_{D} is equivalent to the feasible region of a robust convex program.

5. We prove that if Ξ is a closed convex cone, the function $f_i(x,\xi)$ for any $i \in [I]$ is of the (separable) form $f_i(x,\xi) = w_i(x) - h_i(\xi)$ where $h_i(\xi)$ is positively homogeneous on Ξ, and the moment functions $\{\phi_t(\xi)\}_{t\in\mathcal{T}_2}$ are positively homogeneous on Ξ, then set Z_{D} is convex. This result is a generalization of [11], where the authors assumed that $w_i(\cdot)$ and $h_i(\cdot)$ are affine functions for each $i \in [I]$.

6. When the decision variables are pure binary (i.e. $S \subseteq \{0,1\}^n$) and uncertain constraints are linear, we show that the proposed deterministic approximation can be reformulated as a mixed integer convex program. We also present a numerical study to demonstrate that the proposed reformulation can be effectively solved using a standard solver.

1.3 Connection to existing works

Nonlinear chance constraints have been studied in the context of a variety of applications, e.g., wireless communication [14], transportation [24], facility location [15], and power systems [25]. A number of works have proposed convex reformulations of distributionally robust chance constraints with different types of nonlinear moment ambiguity sets. For example, [11] studied mean dispersion ambiguity set, [5, 7, 12, 26] incorporated second moment into ambiguity set, and [22] considered constraints on the coefficient of variation. Next, we will review existing works on single and joint chance constraints separately.

In the case of a single uncertain constraint, i.e., $|I| = 1$, there has been significant efforts in identifying settings where Z_{D} can be reformulated by deterministic convex constraints. For example, with known mean and covariance of ξ, the authors in [4] showed that the set Z_{D} can be formulated as a second order cone program (SOCP). Recently, more efforts have been made to derive tractable reformulation of the set Z_{D}. For instance, in [26], the authors showed that with given range of first- and second- order moments, the set Z_{D} can be reformulated as a semidefinite program (SDP). These tractability results have been generalized to nonlinear uncertain constraints in [24]. In [10], the authors demonstrated that the set Z_{D} is convex when $\mathcal{P}$ involves conic moment constraints or unimodality of $\mathbb{P}$. Generalizing the above mentioned earlier works, this paper demonstrates that for any ambiguity set with convex moment constraints, when there is a single chance constraint, the set Z_{D} can be reformulated as a convex program.

Tractability results for a joint DRCCP (i.e., $|I| > 1$) are very rare. It has been shown in [10] that optimization over the set Z_{D} is in general NP-hard. Therefore, much of the earlier works built approximation of the set Z_{D} instead of deriving its exact reformulation. For example, in [17], the authors suggested that using Bonferroni's inequality to decompose a joint chance constraint into $|I|$ different single chance constraints whose sum of risk parameters is no larger than ϵ. With such decomposition, any approximation scheme proposed for a single chance constraint could be directly applied. However, Bonferroni's inequality is not tight in general (c.f. [5, 26]). Thus, in [5], the authors proposed to improve Bonferroni's inequality by scaling each uncertain constraint with a positive number and converting them into a single constraint. For any fixed scaler, they were able to provide a conservative SOCP approximation. Later, it was shown in [26] that by optimizing over the scaling parameters, the feasible region of the proposed scaling method is nearly exact to set $Z_{\rm D}$ when $\mathcal{P}$ is described by first- and second- order moments. This result was established using strong duality of SDP. However, in this case, the corresponding deterministic reformulation of (1) turns out to be a bilinear optimization problem, which is naturally hard to solve (c.f. [2]). Recently, [11] derived a tractable reformulation under the restricted assumption that the stochastic mapping $F(x,\xi)$ is separate and affine in (x,ξ), Ξ is a closed convex solid cone and the ambiguity set is defined by mean and dispersion constraints, where the dispersion function is positively homogeneous on the cone Ξ. We extend the results of [26] to any ambiguity set with convex moment constraints and show that the approximation yields a mixed integer convex program when the decision vector x is binary. Unlike [11], we show that a DRCCP with single moment constraint is tractable by relaxing their assumptions on the set Ξ and mapping $F(x,\xi)$, and we also provide new sufficient conditions under which joint DRCCP is tractable.

The remainder of the paper is organized as follows. Section 2 presents some preliminary results to be used subsequently. Section 3 proposes an equivalent deterministic reformulation of the set $Z_{\rm D}$ and develops a tight approximation of the set $Z_{\rm D}$ via a system of biconvex constraints. Section 4 provides various sufficient conditions for the convexity of the set $Z_{\rm D}$. Section 5 demonstrates that the proposed tight approximation of $Z_{\rm D}$ yields a mixed integer convex program when the decision variables are binary and the uncertain constraints are linear. A numerical study is presented to test the proposed formulation. Finally, Section 6 concludes the paper.

2 Preliminaries

In this section we present some standard results and then define a special function associated with the set $Z_{\rm D}$ that will be used in our analysis. We begin by defining some of the notation that will be used throughout.

2.1 Notation

Recall that we use $\boldsymbol{\xi}$ to denote a random vector and ξ to denote a realization of $\boldsymbol{\xi}$. We let $\mathbf{e}$ be the all-ones vector. For a positive integer m, let $[m] := \{1,\ldots,m\}$ and $\mathbb{R}^m_+ = \{x \in \mathbb{R}^m : x \geq 0\}, \mathbb{R}^m_{++} = \{x \in \mathbb{R}^m : x_i > 0, \forall i \in [m]\}$. We let $\mathcal{M}_+(\Xi)$ denote the cone of all nonnegative measures on Ξ. Given a vector $\widehat{F} = (\widehat{f}_1,\ldots,\widehat{f}_I)^\top$, the indicator function $\mathbb{I}_+(\widehat{F})$ is equal to 1 if $\widehat{f}_i \geq 0$ for all $i \in [I]$,0 otherwise. For a statement χ, we let $\mathbb{I}(\chi)$ be 1 if χ is true, 0, otherwise. Given a function $\widehat{f}(\cdot)$, we use $\operatorname{dom} \widehat{f}$ to denote its domain.

2.2 Some standard results

As is common in the distributionally robust optimization literature (e.g., [19]), our deterministic reformulation of $Z_{\rm D}$ relies on dualizing the optimization problem appearing in the left-hand-side of the chance constraint defining (2). Towards this, in addition to Assumptions (A1) - (A3), we will make the following constraint qualification assumption on $\mathcal{P}$ throughout the rest of this paper.

(A4) (Slater's condition) there exists a probability measure $\mathbb{P}$ satisfying the constraints defining $\mathcal{P}$ for any sufficiently small perturbation of $\{g_t\}_{t\in\mathcal{T}_1\cup\mathcal{T}_2}$.

We will use the following strong duality result.

Lemma 1. *Let $\mathcal{P}$ be defined as in* (3) *and suppose Assumptions (A1) - (A4) hold. Let $\psi(\xi)$ be a Lebesgue measurable function on $(\Xi, \mathcal{F})$ such that $|\mathbb{E}_{\mathbb{P}}[\psi(\boldsymbol{\xi})]| < \infty$ for any $\mathbb{P} \in \mathcal{P}$, then*

$$\inf_{\mathbb{P}\in\mathcal{P}} \mathbb{E}_{\mathbb{P}}[\psi(\boldsymbol{\xi})], \tag{4}$$

is equivalent to the following mathematical program:

$$\max \quad \lambda + \sum_{t\in\mathcal{T}_1\cup\mathcal{T}_2} g_t\gamma_t, \tag{5a}$$

$$\text{s.t.} \quad \lambda + \sum_{t\in\mathcal{T}_1\cup\mathcal{T}_2} \phi_t(\xi)\gamma_t \le \psi(\xi), \forall \xi \in \Xi, \tag{5b}$$

$$\gamma_t \ge 0, t \in \mathcal{T}_2. \tag{5c}$$

Proof. Note that $\mathcal{M}_+(\Xi)$ denote the cone of all nonnegative measures on Ξ. Then (4) can be formulated as

$$\inf_{\mu\in\mathcal{M}_+(\Xi)} \quad \int_\Xi \psi(\xi)d\mu(\xi)$$
$$\text{s.t.} \int_\Xi \phi_t(\xi)d\mu(\xi) = g_t, \forall t \in \mathcal{T}_1, \int_\Xi \phi_t(\xi)d\mu(\xi) \ge g_t, \forall t \in \mathcal{T}_2, \int_\Xi d\mu(\xi) = 1.$$

The dual of the above semi-infinite linear program is (5). Due to Assumption (A4), Theorem 5.99 in [3] implies that strong duality holds and the set of optimal solutions are bounded. □

Next, we note that Z_{D} is closed.

Lemma 2. *Under Assumptions (A1) - (A3), Z_{D} is closed.*

Proof. For any given $\mathbb{P} \in \mathcal{P}$, let

$$Z_{\mathbb{P}} := \{x \in \mathbb{R}^n : \mathbb{P}[\boldsymbol{\xi} : F(x, \boldsymbol{\xi}) \ge 0] \ge 1 - \epsilon\}.$$

From Proposition 1.7 in [13], since $F(x, \xi)$ is continuous in x, $Z_{\mathbb{P}}$ is a closed set. By definition, $Z_{\mathrm{D}} = \bigcap_{\mathbb{P}\in\mathcal{P}} Z_{\mathbb{P}}$ and it is well known that any intersection of closed set is also closed. Thus, Z_{D} is closed. □

Finally, we mention a result from convex programming that will be useful.

Lemma 3. *(Convex Theorem of Alternatives, [1]) Consider the convex inequality system*

$$(S_1) \qquad \begin{array}{l} f(x) < c, \\ g_i(x) \le 0, i \in [m], \\ x \in X, \end{array}$$

where $c \in \mathbb{R}$ is a constant, $f(x), \{g_i(x)\}_{i\in[m]}$ are convex functions defined on $\mathbb{R}^n$ and $X \subseteq \mathbb{R}^n$ is a nonempty convex set. Assume that there exist $\bar{x}$ such that $g_i(\bar{x}) < 0$ for each $i \in [m]$. Then system (S_1) is unsolvable if and only if system (S_2) is solvable, where (S_2) is defined as

$$(S_2) \qquad \begin{array}{l} \inf_{x\in X}[f(x) + \sum_{i\in[m]} \lambda_i g_i(x)] \ge c, \\ \lambda_i \ge 0, i \in [m]. \end{array}$$

2.3 ϕ-conjugate functions

Our subsequent constructions will make use of the following function associated with the ingredients defining Z_{D}.

Definition 1. Let $f(x,\xi)$ be a function which is convex in $\xi \in \Xi$ for all x. Given functions $\{\phi_t(\xi)\}_{t\in\mathcal{T}_1\cup\mathcal{T}_2}$ as defined in (3) the *ϕ-conjugate* of f corresponding to weights $(\gamma,\alpha) \in (\mathbb{R}^{|\mathcal{T}_1|} \times \mathbb{R}_+^{|\mathcal{T}_2|}) \times \mathbb{R}_+$ is

$$\psi_f(\gamma,\alpha,x) := \sup_{\xi\in\Xi} \left(\sum_{t\in\mathcal{T}_1\cup\mathcal{T}_2} \phi_t(\xi)\gamma_t - \alpha f(x,\xi) \right)$$

For notational simplicity, when $f(x,\xi) = 0$ for all (x,ξ), we denote

$$\psi_0(\gamma) := \psi_0(\gamma,\alpha,x).$$

Note that evaluating $\psi_f(\gamma,\alpha,x)$ amounts to solving a concave maximization problem. Often, such a problem is tractable. Next we establish some properties of ψ_f. Recall, that given a closed convex cone $\mathcal{C}$, a function $f : \mathcal{C} \to \mathbb{R}$ is *positively homogeneous* on $\mathcal{C}$ if $f(\lambda x) = \lambda f(x)$ for any $x \in \mathcal{C}$ and $\lambda \geq 0$.

Lemma 4. *Let $f(x,\xi)$ be concave in $x \in \mathbb{R}^n$ and convex in $\xi \in \Xi$. Then its ϕ-conjugate function $\psi_f(\gamma,\alpha,x)$ has the following properties:*

(i) $\psi_f(\cdot,\cdot,x)$ is jointly convex in $(\gamma,\alpha) \in (\mathbb{R}^{|\mathcal{T}_1|} \times \mathbb{R}_+^{|\mathcal{T}_2|}) \times \mathbb{R}_+$ for any given x;

(ii) $\psi_f(\cdot,\alpha,\cdot)$ is jointly convex in $(\gamma,x) \in (\mathbb{R}^{|\mathcal{T}_1|} \times \mathbb{R}_+^{|\mathcal{T}_2|}) \times \mathbb{R}^n$ for any given α;

(iii) if Ξ is a closed convex cone, $f(x,\cdot)$ and $\{\phi_t(\cdot)\}_{t\in\mathcal{T}_1\cup\mathcal{T}_2}$ are positively homogeneous on Ξ, then

$$\psi_f(\gamma,\alpha,x) = \begin{cases} 0, & \text{if } (\gamma,\alpha,x) \in \operatorname{dom} \psi_f, \\ +\infty, & \text{otherwise.} \end{cases}$$

Proof. Parts (i) and (ii) follow from the fact that the supremum of a set of convex functions is convex. We only show part (iii). If Ξ is a closed convex cone, $f(x,\cdot)$ and $\{\phi_t(\cdot)\}_{t\in\mathcal{T}_1\cup\mathcal{T}_2}$ are positively homogeneous on Ξ, then we must have $\psi_f(\gamma,\alpha,x) \leq 0$ for any $(\gamma,\alpha,x) \in \operatorname{dom} \psi_f$. Otherwise, there would exist $\bar{\xi} \in \Xi$ such that

$$\sum_{t\in\mathcal{T}_1\cup\mathcal{T}_2} \phi_t(\bar{\xi})\gamma_t - \alpha f(x,\bar{\xi}) > 0$$

and so

$$\sup_{\xi\in\Xi} \left(\sum_{t\in\mathcal{T}_1\cup\mathcal{T}_2} \phi_t(\xi)\gamma_t - \alpha f(x,\xi) \right) \geq \lim_{\lambda\to\infty} \sum_{t\in\mathcal{T}_1\cup\mathcal{T}_2} \phi_t(\lambda\bar{\xi})\gamma_t - \alpha f(x,\lambda\bar{\xi}) = \infty$$

where the first inequality comes from $\lambda\bar{\xi} \in \Xi$ for any $\lambda \geq 0$ and the first equality due to the positive homogeneity of $f(x,\cdot)$ and $\{\phi_t(\cdot)\}_{t\in\mathcal{T}_1\cup\mathcal{T}_2}$. Also from positive homogeneity it follows that $\sum_{t\in\mathcal{T}_1\cup\mathcal{T}_2} \phi_t(0)\gamma_t - \alpha f(x,0) = 0$. Thus $\psi_f(\gamma,\alpha,x) = 0$ for all $(\gamma,\alpha,x) \in \operatorname{dom} \psi_f$. ☐

Note that in part (iii) of Lemma 4, if Ξ is a polyhedral cone, $f(x,\xi)$ is linear in ξ for any given x and $\{\phi_t(\xi)\}_{t\in\mathcal{T}_1\cup\mathcal{T}_2}$ are all linear functions, then by strong duality of linear programming, $\psi_f(\gamma,\alpha,x)$ is equal to the characteristic function of a set defined by linear inequalities on (γ,α) for any given x.

Next we present a few examples of $\mathcal{P}$ whose associated ϕ-conjugate function and its domain can be explicitly computed. We omit the calculations for brevity.

Proposition 1. *Let*

$$\mathcal{P} = \left\{\mathbb{P} \in \mathcal{P}_0\left(\mathbb{R}^n\right) : \mathbb{E}_{\mathbb{P}}[\|\boldsymbol{\xi}\|_q] \leq g_1\right\},$$

with $q \geq 1, \Xi = \mathbb{R}^m$, *and* $f(x,\xi) = b - \mu^\top x - \xi^\top x$. *Here* $\sum_{t\in\mathcal{T}_1\cup\mathcal{T}_2} \phi_t(\xi)\gamma_t = -\gamma\|\xi\|_q$. *Then*

$$\psi_f(\gamma,\alpha,x) = \begin{cases} -\alpha(b-\mu^\top x), & \text{if } \|\alpha x\|_{\frac{q}{q-1}} \leq \gamma,\ \gamma \geq 0 \\ +\infty, & \text{otherwise.} \end{cases}$$

Proposition 2. *Let*

$$\mathcal{P} = \left\{\mathbb{P} \in \mathcal{P}_0\left(\mathbb{R}^m\right) : \mathbb{E}_{\mathbb{P}}[w_t^\top \boldsymbol{\xi}] = g_t, t \in \mathcal{T}_1, \mathbb{E}_{\mathbb{P}}[w_t^\top \boldsymbol{\xi}] \geq g_t, t \in \mathcal{T}_2\right\},$$

with $\Xi = \mathbb{R}^m$, $f(x,\xi) = (Ax+a)^\top\xi + Bx + b$, *and* $\sum_{t\in\mathcal{T}_1\cup\mathcal{T}_2} \phi_t(\xi)\gamma_t = \sum_{t\in\mathcal{T}_1\cup\mathcal{T}_2} \gamma_t w_t^\top \xi$. *Then*

$$\psi_f(\gamma,\alpha,x) = \begin{cases} -\alpha(Bx+b), & \text{if } \alpha(Ax+a) = \sum_{t\in\mathcal{T}_1\cup\mathcal{T}_2} \gamma_t w_t,\ \gamma_t \geq 0\ \forall\, t \in \mathcal{T}_2 \\ +\infty, & \text{otherwise.} \end{cases}$$

Proposition 3. *Let*

$$\mathcal{P} = \left\{\mathbb{P} \in \mathcal{P}_0\left(\mathbb{R}^m\right) : \mathbb{E}_{\mathbb{P}}[\boldsymbol{\xi}] = 0, \mathbb{E}_{\mathbb{P}}[\|\boldsymbol{\xi}\|_q] \leq g\right\},$$

with $q \geq 1$, $\Xi = \mathbb{R}^m$, *and* $f(x,\xi) = h(x) + B\xi$ *where* $h(x)$ *is concave in* x. *Here* $\sum_{t\in\mathcal{T}_1\cup\mathcal{T}_2} \phi_t(\xi)\gamma_t = \gamma_1^\top\xi - \gamma_2\|\xi\|_q$ *Then*

$$\psi_f(\gamma,\alpha,x) = \begin{cases} -\alpha h(x), & \text{if } \|\gamma_1 - \alpha B^\top\|_{\frac{q}{q-1}} \leq \gamma_2,\ \gamma_2 \geq 0 \\ +\infty, & \text{otherwise.} \end{cases}$$

Proposition 4. *Let*

$$\mathcal{P} = \left\{\mathbb{P} \in \mathcal{P}_0\left(\mathbb{R}^m\right) : \mathbb{E}_{\mathbb{P}}[\boldsymbol{\xi}] = 0, \mathbb{E}_{\mathbb{P}}[\boldsymbol{\xi}\boldsymbol{\xi}^\top] \preceq \Sigma\right\},$$

with $\Xi = \mathbb{R}^m$, *and* $f(x,\xi) = (Ax+a)^\top\xi + Bx + b$. *Here* $\sum_{t\in\mathcal{T}_1\cup\mathcal{T}_2} \phi_t(\xi)\gamma_t = \gamma_1^\top\xi - \langle\gamma_2, \xi\xi^\top\rangle$. *Then*

$$\psi_f(\gamma,\alpha,x) = \begin{cases} -\alpha(Bx+b) + \min\left\{t : T(t,\gamma,\alpha,x) \succeq 0\right\} & \text{if } \gamma_2 \succeq 0 \\ +\infty & \text{otherwise,} \end{cases}$$

with $T(t,\gamma,\alpha,x) = \begin{bmatrix} t & -\frac{1}{2}(\gamma_1 - \alpha(Ax+a))^\top \\ -\frac{1}{2}(\gamma_1 - \alpha(Ax+a)) & \gamma_2 \end{bmatrix}$.

Proposition 5. *Let*

$$\mathcal{P} = \left\{\mathbb{P} \in \mathcal{P}_0\left(\mathbb{R}^m\right) : \mathbb{E}_{\mathbb{P}}[\boldsymbol{\xi}] = 0, \mathbb{E}_{\mathbb{P}}[(h^\top\boldsymbol{\xi})^2] \leq g(h^\top\mu)^2\right\},$$

with $\Xi = \mathbb{R}^m$, *and* $f(x,\xi) = (Ax+a)^\top\xi + Bx + b$. *Here* $\sum_{t\in\mathcal{T}_1\cup\mathcal{T}_2} \phi_t(\xi)\gamma_t = \gamma_1^\top\xi - \gamma_2 g(h^\top\xi)^2$. *Then*

$$\psi_f(\gamma,\alpha,x) = \begin{cases} -\alpha(Bx+b) + \min\left\{t : T(t,\gamma,\alpha,x) \succeq 0\right\}, & \text{if } \gamma_2 \geq 0 \\ +\infty & \text{otherwise,} \end{cases}$$

with $T(t,\gamma,\alpha,x) = \begin{bmatrix} t & -\frac{1}{2}(\gamma_1 - \alpha(Ax+a))^\top \\ -\frac{1}{2}(\gamma_1 - \alpha(Ax+a)) & \gamma_2 g h h^\top \end{bmatrix}$.

The mean deviation ambiguity set in Proposition 3 has been studied in [11]. The first and second- moment ambiguity set in Proposition 4 has been studied in [5, 7, 12, 26], while the mean and coefficient of variation ambiguity set in Proposition 5 has been studied in [22].

3 Deterministic formulations

In this section we present two deterministic formulations associated with Z_{D}. The first is a direct reformulation using the strong duality result in Lemma 1. The second is an approximate formulation of Z_{D} via a biconvex program which is shown to be nearly tight.

3.1 Direct reformulation

Lemma 1 is sufficient to derive the following deterministic reformulation of Z_{D}. We will investigate the convexity of this reformulation in Section 4.

Theorem 1. *Suppose Assumptions (A1) - (A4) hold, then*

$$Z_{\mathrm{D}} = \left\{ \begin{array}{ll} & \lambda + \sum_{t \in \mathcal{T}_1 \cup \mathcal{T}_2} g_t \gamma_t \geq 1 - \epsilon, \quad (6a) \\ x : & \lambda + \psi_0 (\gamma) \leq 1, \quad (6b) \\ & \lambda + \psi_{f_i} (\gamma, \alpha_i, x) \leq 0, \forall i \in I_1(x), \quad (6c) \\ & \gamma_t \geq 0, \forall t \in \mathcal{T}_2, \alpha_i \geq 0, \forall i \in I_1(x), \quad (6d) \end{array} \right\}$$

where the functions $\psi_0 (\cdot), \psi_\cdot (\cdot, \cdot, \cdot)$ are ϕ-conjugate functions as defined in Definition 1, and

$$I_1(x) := \{i \in [I] : \exists \xi \in \Xi, f_i(x, \xi) < 0\}.$$

Proof. Note that

$$Z_D = \left\{ x \in \mathbb{R}^n : \inf_{\mathbb{P} \in \mathcal{P}} \mathbb{E}[\mathbb{I}_+(F(x, \boldsymbol{\xi}))] \geq 1 - \epsilon \right\}.$$

By Lemma 1, the left-hand side of the constraint defining Z_{D} is equivalent to

$$\max \quad \lambda + \sum_{t \in \mathcal{T}_1 \cup \mathcal{T}_2} g_t \gamma_t \tag{7a}$$

$$\text{s.t.} \quad \lambda + \sum_{t \in \mathcal{T}_1 \cup \mathcal{T}_2} \phi_t(\xi) \gamma_t \leq \mathbb{I}_+(F(x, \xi)), \ \forall \xi \in \Xi, \tag{7b}$$

$$\gamma_t \geq 0, t \in \mathcal{T}_2. \tag{7c}$$

From the definition of $\mathbb{I}_+(F(x, \xi))$, constraints (7b) are equivalent to

$$\lambda + \sum_{t \in \mathcal{T}_1 \cup \mathcal{T}_2} \phi_t(\xi) \gamma_t \leq 1, \forall \xi \in \Xi, \tag{8a}$$

$$\lambda + \sum_{t \in \mathcal{T}_1 \cup \mathcal{T}_2} \phi_t(\xi) \gamma_t \leq 0, \forall \xi \in \bigcup_{i \in [I]} \{\xi \in \Xi : f_i(x, \xi) < 0\}. \tag{8b}$$

Constraint (8a) is equivalent to

$$\lambda + \sup_{\xi \in \Xi} \left(\sum_{t \in \mathcal{T}_1 \cup \mathcal{T}_2} \phi_t(\xi) \gamma_t \right) \leq 1,$$

which is equivalent to (6b) by the definition of $\psi_0 (\gamma)$. Next we focus on reformulating (8b).

Given x, let $\Xi_i(x) := \{\xi \in \Xi : \ f_i(x, \xi) < 0\}$ for all $i \in [I]$. Note that by definition of $I_1(x)$ we have that $\Xi_i(x) \neq \emptyset$ for all $i \in I_1(x)$. Thus (8b) is equivalent to

$$\lambda + \sup_{\xi \in \Xi_i(x)} \left(\sum_{t \in \mathcal{T}_1 \cup \mathcal{T}_2} \phi_t(\xi)\gamma_t \right) \leq 0, \forall i \in I_1(x). \tag{9}$$

For any given x, since $f_i(x, \xi)$ and $\sum_{t \in \mathcal{T}_1 \cup \mathcal{T}_2} \phi_t(\xi)\gamma_t$ are continuous in ξ, Ξ is a closed set and by definition $\Xi_i(x) \neq \emptyset$ for each $i \in I_1(x)$, so we can replace the strict inequalities in $\Xi_i(x)$ by non-strict ones. Thus, (9) is equivalent to

$$\lambda + \sup_{\substack{\xi \in \Xi \\ f_i(x,\xi) \leq 0}} \left(\sum_{t \in \mathcal{T}_1 \cup \mathcal{T}_2} \phi_t(\xi)\gamma_t \right) \leq 0, \forall i \in I_1(x). \tag{10}$$

For any given $\gamma \in \mathbb{R}^{|\mathcal{T}_1|} \times \mathbb{R}_+^{|\mathcal{T}_2|}$, $\lambda \in \mathbb{R}$ and $x \in S$, (10) implies that the following constraint system on ξ is insolvable

$$\lambda + \sum_{t \in \mathcal{T}_1 \cup \mathcal{T}_2} \phi_t(\xi)\gamma_t > 0, \xi \in \Xi, f_i(x, \xi) \leq 0, \tag{11}$$

for each $i \in I_1(x)$.

By definition of the set $I_1(x)$, we have that there exists $\bar{\xi} \in \Xi$ such that $f_i(x, \bar{\xi}) < 0$. Thus by Lemma 3, (11) is equivalent to that there exists an $\alpha_i \geq 0$,

$$\lambda + \sup_{\xi \in \Xi} \left(\sum_{t \in \mathcal{T}_1 \cup \mathcal{T}_2} \phi_t(\xi)\gamma_t - \alpha_i f_i(x, \xi) \right) \leq 0,$$

for each $i \in I_1(x)$. By definition of ψ_{f_i} the above system is equivalent to (6c). □

We remark that reformulation (6) of Z_{D} is not convex since each function $\psi_{f_i}(\cdot)$ is not in general convex for $i \in [I]$, and also because the index set $I_1(x)$ depends on x. In the subsequent sections, we will explore the tractability of the set Z_{D} by establishing conditions under which $\{\psi_{f_i}(\gamma, \alpha_i, x)\}_{i \in [I]}$ are convex and $I_1(x)$ can be replaced by $[I]$.

We also remark that by Lemma 3.1 in [20], for any distribution $\mathbb{P} \in \mathcal{P}$, there exists a discrete distribution $\widehat{\mathbb{P}} \in \mathcal{P}$ with a finite support of at most $1 + |\mathcal{T}_1| + |\mathcal{T}_2|$ points. Therefore, there exists a worst-case distribution $\mathbb{P}^*$ which achieves the infimum in (1c) has a finite support with points $\{\xi_j^*\}_{j \in [J]}$ with $J \leq 1 + |\mathcal{T}_1| + |\mathcal{T}_2|$, i.e., $\mathbb{P}^*$ satisfies

$$\mathbb{P}^*(\boldsymbol{\xi}) = \sum_{j \in [J]} p_j^* \mathbb{I}(\boldsymbol{\xi} = \xi_j^*).$$

We first observe that $\lambda + \psi_0(\gamma)$ can be lower bounded by $1 - \epsilon$.

Corollary 1. *For any $(\lambda, \gamma, \alpha, x)$ satisfying (6), we must have $\lambda + \psi_0(\gamma) \geq 1 - \epsilon$.*

Proof. For any given $\mathbb{P} \in \mathcal{P}$, (3) yields

$$\int_{\Xi} \phi_t(\xi)\mathbb{P}(d\xi) = g_t, \forall t \in \mathcal{T}_1, \int_{\Xi} \phi_t(\xi)\mathbb{P}(d\xi) \geq g_t, \forall t \in \mathcal{T}_2.$$

Since $\gamma_t \geq 0$ for each $t \in \mathcal{T}_2$ by aggregating the above inequalities with multipliers $\{\gamma_t\}_{t \in \mathcal{T}_1 \cup \mathcal{T}_2}$, we obtain

$$\sum_{t \in \mathcal{T}_1 \cup \mathcal{T}_2} \int_{\Xi} \phi_t(\xi)\gamma_t \mathbb{P}(d\xi) \geq \sum_{t \in \mathcal{T}_1 \cup \mathcal{T}_2} g_t\gamma_t. \tag{12}$$

Since

$$\psi_0(\gamma) := \sup_{\xi\in\Xi}\left(\sum_{t\in\mathcal{T}_1\cup\mathcal{T}_2}\phi_t(\xi)\gamma_t\right) \geq \sum_{t\in\mathcal{T}_1\cup\mathcal{T}_2}\int_{\Xi}\phi_t(\xi)\gamma_t\mathbb{P}(d\xi),$$

therefore (12) and $\lambda + \sum_{t\in\mathcal{T}_1\cup\mathcal{T}_2} g_t\gamma_t \geq 1-\epsilon$ imply

$$\lambda + \psi_0(\gamma) \geq 1-\epsilon.$$

□

We also observe that for each $i \in I_1(x)$, α_i must be strictly positive. This observation is key to the proofs of several main results in subsequent sections, for instance, it allows us to prove the convexity of the set $Z_{\rm D}$ when $\Xi = \mathbb{R}^m$ and $\{\phi_t(\xi)\}_{t\in\mathcal{T}_1\cup\mathcal{T}_2}$ are all linear functions.

Corollary 2. *For any x satisfying* (6)*, we must have $\alpha_i > 0$ for all $i \in I_1(x)$.*

Proof. If $I_1(x) = \emptyset$, then we are done. Now let us assume that $I_1(x) \neq \emptyset$. Suppose $\alpha_{i_0} = 0$ for some $i_0 \in I_1(x)$. Then from (6c), we have $\lambda + \psi_{f_{i_0}}(\gamma, 0, x) \leq 0$, which is equivalent to $\lambda + \psi_0(\gamma) \leq 0$. This yields a contradiction to Corollary 1 that $\lambda + \psi_0(\gamma) \geq 1-\epsilon$. □

3.2 Biconvex approximation

Recently, in [5], [12] and [26], the authors derived **CVaR** approximation of a joint DRCCP with linear uncertain constraints, which yields an almost exact feasible region of a DRCCP. The construction of such an approximation scheme is outlined below. First, for any given positive vector $\alpha \in \mathbb{R}^I_{++}$, we can convert the joint chance constraints into a single one as

$$\begin{aligned} Z_{\rm D} &= \left\{x \in \mathbb{R}^n : \alpha \in \mathbb{R}^I_{++}, \inf_{\mathbb{P}\in\mathcal{P}} \mathbb{P}[\boldsymbol{\xi} : \max_{i\in[I]}\{\alpha_i(-f_i(x,\xi))\} \leq 0] \geq 1-\epsilon\right\} \\ &= \left\{x \in \mathbb{R}^n : \alpha \in \mathbb{R}^I_{++}, \sup_{\mathbb{P}\in\mathcal{P}} \mathbb{P}[\boldsymbol{\xi} : \max_{i\in[I]}\{\alpha_i(-f_i(x,\xi))\} > 0] \leq \epsilon\right\} \end{aligned}$$

where the second equality is due to $\mathbb{P}[\boldsymbol{\xi} : \max_{i\in[I]}\{\alpha_i(-f_i(x,\xi))\} > 0] + \mathbb{P}[\boldsymbol{\xi} : \max_{i\in[I]}\{\alpha_i(-f_i(x,\xi))\} \leq 0] = 1$. For any given probability measure $\mathbb{P} \in \mathcal{P}$, apply the **CVaR** approximation of [17] to the above chance constraint, which yields a conservative approximation of $Z_{\rm D}$ as

$$Z_{\rm D} \supseteq \left\{x : \alpha \in \mathbb{R}^I_{++}, \sup_{\mathbb{P}\in\mathcal{P}}\left\{\inf_{\beta\in\mathbb{R}}\left\{\beta + \frac{1}{\epsilon}\mathbb{E}_{\mathbb{P}}\left[\left(\max_{i\in[I]}\{\alpha_i(-f_i(x,\boldsymbol{\xi}))\} - \beta\right)_+\right]\right\}\right\} \leq 0\right\}.$$

We further note that if we interchange the infimum with the supremum, then by standard minimax argument, set $Z_{\rm D}$ is further approximated by

$$\begin{aligned} Z_{\rm D} &\supseteq \left\{x : \alpha \in \mathbb{R}^I_{++}, \inf_{\beta\in\mathbb{R}}\left\{\beta + \frac{1}{\epsilon}\sup_{\mathbb{P}\in\mathcal{P}}\mathbb{E}_{\mathbb{P}}\left[\left(\max_{i\in[I]}\{\alpha_i(-f_i(x,\boldsymbol{\xi}))\} - \beta\right)_+\right]\right\} \leq 0\right\} \\ &\supseteq Z_C = \left\{x : \alpha \in \mathbb{R}^I_{++}, \beta \in \mathbb{R}, \beta + \frac{1}{\epsilon}\sup_{\mathbb{P}\in\mathcal{P}}\mathbb{E}_{\mathbb{P}}\left[\left(\max_{i\in[I]}\{\alpha_i(-f_i(x,\boldsymbol{\xi}))\} - \beta\right)_+\right] \leq 0\right\} \end{aligned} \quad (13)$$

where the second inclusion is because infimum might not be achieved by any β.

The relation (13) leads us to reformulate Z_C as a disjunction of two sets by distinguishing whether $\beta = 0$ or not.

Theorem 2. *$Z_C = X_C \cup Y_C$, where*

$$X_C = \{x : F(x,\xi) \geq 0, \forall \xi \in \Xi\} \tag{14}$$

and

$$Y_C = \left\{ x : \begin{aligned} & \lambda + \sum_{t \in \mathcal{T}_1 \cup \mathcal{T}_2} g_t \gamma_t \geq 1 - \epsilon, && (15a) \\ & \lambda + \psi_0(\gamma) \leq 1, && (15b) \\ & \lambda + \psi_{f_i}(\gamma, \alpha_i, x) \leq 0, \forall i \in [I], && (15c) \\ & \gamma_t \geq 0, \forall t \in \mathcal{T}_2, \alpha_i \geq 0, \forall i \in [I]. && (15d) \end{aligned} \right\}$$

Proof. We separate the proof into three parts.

(i) Note that in (13), we must have $\beta \leq 0$; otherwise, as $(\max_{i\in[I]}\{\alpha_i(-f_i(x,\xi))\} - \beta)_+ \geq 0$ for all $\xi \in \Xi$, thus the expectation in (13) is always nonnegative, which implies that the left-hand side of (13) is strictly positive, a contradiction.

(ii) For any $x \in Z_C$, there exists $(\alpha, \beta) \in \mathbb{R}^I_{++} \times \mathbb{R}_-$ such that

$$\beta\epsilon + \sup_{\mathbb{P}\in\mathcal{P}} \mathbb{E}_{\mathbb{P}}\left[\left(\max_{i\in[I]}\{\alpha_i(-f_i(x,\boldsymbol{\xi}))\} - \beta\right)_+\right] \leq 0.$$

Now we distinguish whether $\beta = 0$ or $\beta < 0$.

(a) If $\beta = 0$, then (13) yields

$$\sup_{\mathbb{P}\in\mathcal{P}} \mathbb{E}_{\mathbb{P}}\left[\left(\max_{i\in[I]}\{\alpha_i(-f_i(x,\boldsymbol{\xi}))\}\right)_+\right] \leq 0 \tag{16}$$

which is equivalent to

$$\inf_{\mathbb{P}\in\mathcal{P}} \mathbb{P}[F(x,\xi) \geq 0] = 1 \geq 1 - \epsilon.$$

By continuity of $F(x,\xi)$, $\inf_{\mathbb{P}\in\mathcal{P}} \mathbb{P}[F(x,\xi) \geq 0] = 1$ implies that $F(x,\xi) \geq 0$ for all $\xi \in \Xi$, which implies that the feasible solution x must be in X_C.

(b) Suppose $\beta < 0$. Divide (13) by $-\beta$ and add ϵ on both sides, then we have

$$\frac{1}{-\beta} \sup_{\mathbb{P}\in\mathcal{P}} \mathbb{E}_{\mathbb{P}}\left[\left(\max_{i\in[I]}\{\alpha_i(-f_i(x,\boldsymbol{\xi}))\} - \beta\right)_+\right] \leq \epsilon. \tag{17}$$

Since $\beta < 0$, we can redefine α_i as $\alpha_i/(-\beta)$ for each $i \in [I]$. Thus, (17) yields

$$\sup_{\mathbb{P}\in\mathcal{P}} \mathbb{E}_{\mathbb{P}}\left[\left(\max_{i\in[I]}\{\alpha_i(-f_i(x,\boldsymbol{\xi}))\} + 1\right)_+\right] \leq \epsilon.$$

Subtracting one on both sides and flipping the sign of inequality yields

$$\inf_{\mathbb{P}\in\mathcal{P}} \mathbb{E}_{\mathbb{P}}\left[1 - \left(\max_{i\in[I]}\{\alpha_i(-f_i(x,\boldsymbol{\xi}))\} + 1\right)_+\right] \geq 1 - \epsilon. \tag{18}$$

By Lemma 1, for any given $\alpha \in \mathbb{R}^I_{++}$, the infimum in the left-hand side of (18) is equivalent to

$$\max\ \lambda + \sum_{t\in\mathcal{T}_1\cup\mathcal{T}_2} g_t\gamma_t, \tag{19a}$$

$$\text{s.t. } \lambda + \sum_{t\in\mathcal{T}_1\cup\mathcal{T}_2} \phi_t(\xi)\gamma_t \le 1 - \left(\max_{i\in[I]} \{\alpha_i(-f_i(x,\xi))\} + 1\right)_+, \forall\xi\in\Xi, \tag{19b}$$

$$\gamma_t \ge 0, t\in\mathcal{T}_2. \tag{19c}$$

Since the maximum value of (19) should be no smaller than $1-\epsilon$, therefore, we can replace the maximization over λ, γ by existing λ, γ. In addition, breaking down the maximum in (19b), (18) is equivalent to

$$\lambda + \sum_{t\in\mathcal{T}_1\cup\mathcal{T}_2} g_t\gamma_t \ge 1-\epsilon, \tag{20a}$$

$$\lambda + \sum_{t\in\mathcal{T}_1\cup\mathcal{T}_2} \phi_t(\xi)\gamma_t \le 1, \forall\xi\in\Xi, \tag{20b}$$

$$\lambda + \left[\sum_{t\in\mathcal{T}_1\cup\mathcal{T}_2} \phi_t(\xi)\gamma_t - \alpha_i f_i(x,\xi)\right] \le 0, \forall\xi\in\Xi, i\in[I], \tag{20c}$$

$$\gamma_t \ge 0, t\in\mathcal{T}_2, \tag{20d}$$

for some $\alpha\in\mathbb{R}^I_{++}$. By definition of ψ_f, (20a)-(20c) are equivalent to (15a)-(15c). Therefore, $(\lambda,\gamma,\alpha,x)$ satisfies (15). Thus, $x\in Y_C$.

This implies that $Z_C \subseteq X_C \cup Y_C$.

(iii) Now let $x\in X_C\cup Y_C$. If $x\in X_C$, then choose $\beta=0, \alpha=\mathbf{e}$, thus $x\in Z_C$. If $x\in Y_C$, there exists $(\lambda',\gamma',\alpha')$ such that $(\lambda',\gamma',\alpha',x)$ satisfies (15) and we must have $\alpha'>0$ from Corollary 2. In (13), let $\beta=1, \alpha=\alpha'$. Then by Lemma 1, the dual reformulation (13) is equivalent to the set Z_C. Thus, $x\in Z_C$.

□

Remark 1. To solve (1) over set $S\cap Z_C$, one can optimize $c^\top x$ over $S\cap X_C$ and $S\cap Y_C$ separately, then choose the minimum value.

Remark 2. We note that the left-hand sides of the constraint system (15) are biconvex in α and (λ,γ,x), i.e., for any given $\alpha\in\mathbb{R}^I_+$, they are convex in (λ,γ,x), and also convex in α for any given (λ,γ,x).

We observe that Y_C is quite similar to (6) except that index set $I_1(x)$ is equal to $[I]$ in (15). Indeed, we next show that Z_{D} is equivalent to Y_C under a certain condition.

Theorem 3. *Let*

$$\mathcal{C} := \left\{x\in Z_{\mathrm{D}} : \exists i\in[I], \inf_{\xi\in\Xi} f_i(x,\xi) = 0\right\}.$$

If $\mathcal{C}=\emptyset$*, then*

$$Y_C = Z_C = Z_{\mathrm{D}}.$$

Proof. Since Theorem 2 implies that $Y_C\subseteq Z_C\subseteq Z_{\mathrm{D}}$, we only need to show that $Z_{\mathrm{D}}\subseteq Y_C$. We first rewrite set Y_C as

$$Y_C = \left\{ x : \begin{array}{ll} \lambda + \sum_{t \in \mathcal{T}_1 \cup \mathcal{T}_2} g_t \gamma_t \geq 1 - \epsilon, & (21a) \\ \lambda + \psi_0(\gamma) \leq 1, & (21b) \\ \lambda + \psi_{f_i}(\gamma, \alpha_i, x) \leq 0, \forall i \in I_1(x), & (21c) \\ \lambda + \psi_{f_i}(\gamma, \alpha_i, x) \leq 0, \forall i \in [I] \setminus I_1(x), & (21d) \\ \gamma_t \geq 0, \forall t \in \mathcal{T}_2, \alpha_i \geq 0, \forall i \in [I]. & (21e) \end{array} \right\}$$

By definition we have $[I] \setminus I_1(x) = \{i \in [I] : f_i(x, \xi) \geq 0, \forall \xi \in \Xi\}$. For each $i \in [I] \setminus I_1(x)$, (21d) is equivalent to

$$\lambda + \sum_{t \in \mathcal{T}_1 \cup \mathcal{T}_2} \phi_t(\xi) \gamma_t \leq \alpha_i f_i(x, \xi), \forall \xi \in \Xi.$$

In the above reformulation, by taking supremum over the left-hand side and using the fact that $\lambda + \psi_0(\gamma) \leq 1$, we observe that

$$Y_C \supseteq \left\{ x : \begin{array}{l} 1 \leq \alpha_i \inf_{\xi \in \Xi} f_i(x, \xi), \forall i \in [I] \setminus I_1(x), \\ (21a), (21b), (21c), \\ \gamma_t \geq 0, \forall t \in \mathcal{T}_2, \alpha_i \geq 0, \forall i \in [I]. \end{array} \right\} \quad (22)$$

Using the fact that $\alpha_i > 0$ from the proof of Theorem 2, the right-hand side in (22) is equivalent to

$$Y_C \supseteq \widehat{Y}_C := \left\{ x : \begin{array}{l} 0 < \inf_{\xi \in \Xi} f_i(x, \xi), \forall i \in [I] \setminus I_1(x), \\ (21a), (21b), (21c), \\ \gamma_t \geq 0, \forall t \in \mathcal{T}_2, \alpha_i \geq 0, \forall i \in I_1(x). \end{array} \right\} \quad (23)$$

From definition of $\widehat{Y}_C$, we have

$$Z_{\mathrm{D}} \setminus Y_C \subseteq Z_{\mathrm{D}} \setminus \widehat{Y}_C \subseteq \mathcal{C} := \{x \in Z_{\mathrm{D}} : \exists i \in [I], \inf_{\xi \in \Xi} f_i(x, \xi) = 0\}$$

where the first inclusion is due to $\widehat{Y}_C \subseteq Y_C$, and the second inclusion is because for any $x \in Z_{\mathrm{D}}$ but $x \notin \mathcal{C}$, we must have $x \in \widehat{Y}_C$. □

A direct observation from the proof of Theorem 3 is the sufficient conditions when α could be bounded. This observation is useful for binary DRCCP which will be discussed in subsequent sections.

Corollary 3. *If S is compact and $I_1(x) = [I]$ for all $x \in Z_{\mathrm{D}}$ (i.e., $Z_{\mathrm{D}} = Y_C$), then there exists an $M \in \mathbb{R}^I_{++}$ such that in $S \cap Y_C$, $\alpha_i \leq M_i$ for each $i \in [I]$.*

Proof. By Theorem 3, we have

$$Z_{\mathrm{D}} \setminus Z_C \subseteq Z_{\mathrm{D}} \setminus Y_C \subseteq \mathcal{C} := \{x \in Z_{\mathrm{D}} : \exists i \in [I], \inf_{\xi \in \Xi} f_i(x, \xi) = 0\}.$$

If $I = 1$, then we have

$$\mathcal{C} \subseteq \{x \in \mathbb{R}^n : f_1(x, \xi) \geq 0, \forall \xi \in \Xi\} \subseteq X_C.$$

Thus,

$$Z_{\mathrm{D}} \subseteq Z_C \cup X_C = Z_C,$$

i.e. $Z_{\mathrm{D}} = Z_C$. □

Next we observe that $Z_C = Z_{\mathrm{D}}$ when there is a single uncertain constraint.

Corollary 4. *When $I = 1$, we have $Z_C = Z_{\mathrm{D}}$.*

Proof. By Theorem 3, we have

$$Y_C \subseteq Z_C \subseteq Z_{\mathrm{D}} = \mathrm{cl}(Y_C).$$

Thus,

$$Z_{\mathrm{D}} \setminus Z_C \subseteq Z_{\mathrm{D}} \setminus Y_C \subseteq \mathcal{C} := \{x \in \mathbb{R}^n : \exists i \in [I] \setminus I_1(x), \inf_{\xi \in \Xi} f_i(x, \xi) = 0\}$$

where the first inclusion is due to $Y_C \subseteq Z_C$, and the second inclusion is because for any $x \in Z_{\mathrm{D}}$ but $x \notin \mathcal{C}$, we must have $x \in Y_C$.

If $I = 1$, then we have

$$\mathcal{C} \subseteq \{x \in \mathbb{R}^n : f_1(x, \xi) \geq 0, \forall \xi \in \Xi\} \subseteq X_C.$$

Thus,

$$Z_{\mathrm{D}} \subseteq Z_C \cup X_C = Z_C,$$

i.e. $Z_{\mathrm{D}} = Z_C$. □

Remark 3. A special case of Corollary 4 has been observed by [24, 26] for a DRCCP with first- and second- moment constraints. Here, we provide a different proof and our results apply to a DRCCP with more general convex moment constraints.

Despite the tightness of Z_C, due to the biconvex terms in set Y_C, it is nonconvex in general. However, as shown in Sections 4 and 5, in some cases, it is possible that these biconvex terms can be convexified.

4 Convexity conditions for Z_{D}

In this section, we will explore some settings under which the set Z_{D} is convex. The results in the first two subsections are derived by constructing a new formulation which projects out dual variables λ, α in (6) and proving that the new formulation is convex and equivalent to the set Z_{D}. The subsequent two results exploit positive homogeneity of mappings $\{\phi_t(\xi)\}_{t\imath\mathcal{T}_1\cup\mathcal{T}_2}$ or $F(x, \xi)$.

4.1 Single uncertain constraint

We show that if there is a single uncertain constraint (i.e., $I = 1$), then the set Z_{D} is convex.

Theorem 4. *When $I = 1$, then*

$$Z_{\mathrm{D}} = Z_C = \left\{ x : \begin{array}{ll} -\sum_{t \in \mathcal{T}_1 \cup \mathcal{T}_2} g_t \gamma_t + (1-\epsilon)\psi_0(\gamma) + \epsilon \psi_{f_1}(\gamma, 1, x) \leq 0, & (24a) \\ \gamma_t \geq 0, \forall t \in \mathcal{T}_2, & (24b) \end{array} \right\}$$

which is a convex set.

Proof. Note that $Z_C = Z_{\mathrm{D}}$ from Corollary 4. By definition of the set Z_C in (13), we can obtain

$$Z_C = \left\{ x : \alpha_1 > 0, \beta \in \mathbb{R}, \beta + \frac{1}{\epsilon} \sup_{\mathbb{P} \in \mathcal{P}} \mathbb{E}_{\mathbb{P}} \left[\left(\alpha_1 \left(-f_1(x, \boldsymbol{\xi}) \right) - \beta \right)_+ \right] \leq 0 \right\} \tag{25}$$

By replacing β with $-\beta$ and flipping the sign of inequality, we have that (25) is equivalent to

$$Z_C = \left\{ x : \alpha_1 > 0, \beta \in \mathbb{R}, \inf_{\mathbb{P}\in\mathcal{P}} \mathbb{E}_{\mathbb{P}} \left[\min\left(\alpha_1 f_1(x, \boldsymbol{\xi}), \beta\right)\right] \geq (1-\epsilon)\beta \right\} \tag{26}$$

Then by Lemma 1 and the proof of Theorem 2, the infimum in the left-hand side of (26) is equivalent to

$$Z_C = \left\{ x : \begin{array}{ll} \lambda + \sum_{t\in\mathcal{T}_1\cup\mathcal{T}_2} g_t\gamma_t \geq (1-\epsilon)\beta & \text{(27a)} \\ \lambda + \psi_0(\gamma) \leq \beta & \text{(27b)} \\ \lambda + \psi_{f_1}(\gamma, \alpha_1, x) \leq 0, & \text{(27c)} \\ \gamma_t \geq 0, \forall t \in \mathcal{T}_2, \alpha_1 > 0. & \text{(27d)} \end{array} \right\}$$

Next, we project out variables λ, β from set Z_C in (27) by Fourier-Motzkin elimination procedure and obtain

$$Z_C = \left\{ x : \begin{array}{ll} -\sum_{t\in\mathcal{T}_1\cup\mathcal{T}_2} g_t\gamma_t + (1-\epsilon)\psi_0(\gamma) + \epsilon\psi_{f_1}(\gamma, \alpha_1, x) \leq 0, & \text{(28a)} \\ \gamma_t \geq 0, \forall t \in \mathcal{T}_2, \alpha_1 > 0. & \text{(28b)} \end{array} \right\}$$

We note that in (28a), α_1 must be positive and finite. Thus, by scaling it to be 1, we obtain (24).

Since $\phi_0(\cdot)$ and $\phi_{f_1}(\cdot, 1, \cdot)$ are convex functions, Z_{D} is a convex set. $\square$

Theorem 4 implies that for a single DRCCP, Z_{D} can be always reformulated as a convex set. This has been observed by [12, 24, 26] where $\mathcal{P}$ is only constrained by first- and second- moments. Here, we extend this result to more general moment constraints.

4.2 Single moment constraint

Here we consider the case that $\mathcal{P}$ is described by a single moment constraint. First, we show that for one inequality constraint (i.e., $|\mathcal{T}_1| = 0, |\mathcal{T}_2| = 1$), the set Z_{D} is convex. The main idea behind the proof is to project out the α variables from characterization (6) in Theorem 1.

Theorem 5. *If $|\mathcal{T}_1| = 0$ and $|\mathcal{T}_2| = 1$, then*

$$Z_{\mathrm{D}} = \left\{ x : \begin{array}{ll} -g_1\widehat{\gamma}_i + (1-\epsilon)\psi_0(\widehat{\gamma}_i) + \epsilon\psi_{f_i}(\widehat{\gamma}_i, 1, x) \leq 0, \forall i \in [I], & \text{(29a)} \\ \widehat{\gamma}_i \geq 0, \forall i \in [I], & \text{(29b)} \end{array} \right\}$$

which is a convex set.

Proof. Let Z_{D}^* be the set defined on the right-hand side of (29), which is clearly convex. We will first show that set Z_{D} is equivalent to $\tilde{Z}_{\mathrm{D}}$ by projecting out dual multiplier λ and aggregating two types of constraints into one. Next, we show that the consolidated set $\tilde{Z}_{\mathrm{D}}$ is equivalent to the convex set Z_{D}^*. The proof proceeds in three steps.

(a) First, using Fourier-Motzkin method to project out variable λ in (6), we can reformulate Z_{D} as

$$Z_{\mathrm{D}} = \left\{ x : \begin{array}{ll} -g_1\gamma_1 + \psi_0(\gamma_1) \leq \epsilon, & \text{(30a)} \\ -g_1\gamma_1 + \psi_{f_i}(\gamma_1, \alpha_i, x) \leq \epsilon - 1, \forall i \in I_1(x), & \text{(30b)} \\ \gamma_1 \geq 0, \alpha_i > 0, \forall i \in [I], & \text{(30c)} \end{array} \right\}$$

where in (30c), we let $\alpha_i > 0$ for all $i \in [I]$ due to Corollary 2.

Next, we can get a relaxation of the Z_{D} by aggregating the two constraints (30a), (30b) above together as $(1-\epsilon) \times$ (30a) $+ \epsilon \times$ (30b):

$$\tilde{Z_{\mathrm{D}}} = \left\{ x : \begin{array}{l} -g_1\gamma_1 + (1-\epsilon)\psi_0(\gamma_1) + \epsilon\psi_{f_i}(\gamma_1, \alpha_i, x) \le 0, \forall i \in I_1(x), \\ \gamma_1 \ge 0, \alpha_i > 0, \forall i \in [I]. \end{array} \right\} \begin{array}{r} (31a) \\ (31b) \end{array}$$

Set $\tilde{Z_{\mathrm{D}}}$ is a relaxation of Z_{D}, i.e. $Z_{\mathrm{D}} \subseteq \tilde{Z_{\mathrm{D}}}$.

Next we show that $\tilde{Z_{\mathrm{D}}} \subseteq Z_{\mathrm{D}}$. Recall that $I_1(x) := \{i \in [I] : \exists \xi \in \Xi, f_i(x, \xi) < 0\}$. Given a point $x \in \tilde{Z_{\mathrm{D}}}$, we consider two cases $I_1(x) = \emptyset$ and $I_1(x) \neq \emptyset$. If $I_1(x) = \emptyset$, then let $\gamma_1 = 0$ and $\alpha = \mathbf{e}$ ($\mathbf{e}$ is the all-one vector). Clearly (γ_1, α, x) satisfies constraints in (30). Hence $x \in Z_{\mathrm{D}}$.

Now suppose that $I_1(x) \neq \emptyset$. As $x \in \tilde{Z_{\mathrm{D}}}$, there exists (γ_1, α) such that (γ_1, α, x) satisfies constraints in (31). First of all, we claim that $\gamma_1 > 0$; otherwise, suppose that $\gamma_1 = 0$, then by (31a), Definition 1 and the fact that $\alpha \in \mathbb{R}^I_{++}$, we have

$$\sup_{\xi \in \Xi} -\alpha_i f_i(x, \xi) \le 0,$$

which implies that $f_i(x, \xi) \ge 0$ for all $\xi \in \Xi$ and $i \in I_1(x)$, contradicting $I_1(x) \neq \emptyset$. Thus, we must have $\gamma_1 > 0$.

Next, we claim that $-g_1\gamma_1 + \psi_0(\gamma_1) > 0$. Indeed, by Assumption (A4), there exists a probability measure $\mathbb{P} \in \mathcal{P}_0(\Xi)$ such that $\mathbb{E}_{\mathbb{P}}[\phi_1(\boldsymbol{\xi})] > g_1$. Thus

$$g_1 < \mathbb{E}_{\mathbb{P}}[\phi_1(\boldsymbol{\xi})] \le \sup_{\xi \in \Xi} \phi_1(\xi).$$

Since $\gamma_1 > 0$, we must have

$$-g_1\gamma_1 + \psi_0(\gamma_1) = \gamma_1[-g_1 + \sup_{\xi \in \Xi} \phi_1(\xi)] > 0.$$

Define $\bar{\gamma}_1 = \epsilon\gamma_1 / [-g_1\gamma_1 + \psi_0(\gamma_1)]$. Thus,

$$\begin{aligned} & -g_1\bar{\gamma}_1 + \psi_0(\bar{\gamma}_1) = -g_1\epsilon\gamma_1 / [-g_1\gamma_1 + \psi_0(\gamma_1)] + \sup_{\xi \in \Xi} (\epsilon\gamma_1\phi_1(\xi) / [-g_1\gamma_1 + \psi_0(\gamma_1)]) \\ & = \frac{\epsilon}{-g_1\gamma_1 + \psi_0(\gamma_1)} [-g_1\gamma_1 + \psi_0(\gamma_1)] = \epsilon, \end{aligned}$$

where the first and second equalities are from the definition of $\psi_0(\cdot)$ and construction of $\bar{\gamma}_1$; and similarly for each $i \in I_1(x)$,

$$\begin{aligned} & -g_1\bar{\gamma}_1 + \psi_{f_1}(\bar{\gamma}_1, \alpha_i, x) = -g_1\epsilon\gamma_1 / [-g_1\gamma_1 + \psi_0(\gamma_1)] \\ & + \sup_{\xi \in \Xi} (\epsilon\gamma_1\phi_1(\xi) / [-g_1\gamma_1 + \psi_0(\gamma_1)] - \epsilon\alpha_i f_i(x, \xi) / [-g_1\gamma_1 + \psi_0(\gamma_1)]) \\ & = \frac{\epsilon}{-g_1\gamma_1 + \psi_0(\gamma_1)} [-g_1\gamma_1 + \psi_{f_i}(\gamma, \alpha_i, x)] \\ & \le \frac{\epsilon}{-g_1\gamma_1 + \psi_0(\gamma_1)} \left[-g_1\gamma_1 + \frac{g_1\gamma_1 - (1-\epsilon)\psi_0(\gamma)}{\epsilon} \right] = \epsilon - 1, \end{aligned}$$

where the first and second equalities are from the definition of $\psi_0(\cdot)$ and construction of $\bar{\gamma}_1$, and the last inequality is due to (31a).

Hence, $(\bar{\gamma}_1, \alpha, x)$ satisfy the constraints in (30); i.e., $x \in Z_{\mathrm{D}}$. Thus, $\tilde{Z_{\mathrm{D}}} = Z_{\mathrm{D}}$.

(b) Now we show that $Z^*_{\mathrm{D}} \subseteq \tilde{Z_{\mathrm{D}}}$. Given a point $x \in Z^*_{\mathrm{D}}$, if $I_1(x) = \emptyset$, then clearly $x \in \tilde{Z_{\mathrm{D}}}$.

Suppose $I_1(x) \neq \emptyset$. Since $x \in Z_{\mathrm{D}}^*$, there must exist a vector $\widehat{\gamma}$ such that $(\widehat{\gamma}, x)$ satisfies (29). For any $i \in I_1(x)$, we must have $\widehat{\gamma}_i > 0$; otherwise, (31a) yields $\psi_{f_i}(0, 1, x) \leq 0$; i.e., $f_i(x, \xi) \geq 0$ for all $\xi \in \Xi$, contradicting $i \in I_1(x)$.

Let $\gamma_1 = \max_{i \in I_1(x)} \widehat{\gamma}_i$ and set $\alpha_i = \gamma_1 / \widehat{\gamma}_i$ for each $i \in I_1(x)$ and $\alpha_i = 1$, otherwise. Then for each $i \in I_1(x)$,

$$
\begin{aligned}
& -g_1\gamma_1 + (1-\epsilon)\psi_0(\gamma_1) + \epsilon\psi_{f_i}(\gamma_1, \alpha_i, x) \\
&= -g_1\gamma_1 + (1-\epsilon)\sup_{\xi \in \Xi}(\gamma_1\phi_1(\xi)) + \epsilon \sup_{\xi \in \Xi}(\gamma_1\phi_1(\xi) - \gamma_1/\widehat{\gamma}_i f_i(x, \xi)) \\
&= \frac{\gamma_1}{\widehat{\gamma}_i}\left[-g_1\widehat{\gamma}_i + (1-\epsilon)\sup_{\xi \in \Xi}(\widehat{\gamma}_i\phi_1(\xi)) + \epsilon\sup_{\xi \in \Xi}(\widehat{\gamma}_i\phi_1(\xi) - f_i(x, \xi))\right] \\
&= \frac{\gamma_1}{\widehat{\gamma}_i}\left[-g_1\widehat{\gamma}_i + (1-\epsilon)\psi_0(\widehat{\gamma}_i) + \epsilon\psi_{f_i}(\widehat{\gamma}_i, 1, x)\right] \leq 0,
\end{aligned}
$$

where the first three equalities are from the definition of $\psi_\cdot(\cdot)$ and construction of γ_1, α, and the last inequality due to (29a) and the fact that $\gamma_1 = \max_{j \in I_1(x)} \widehat{\gamma}_j \geq \widehat{\gamma}_i > 0$.

Thus, (γ_1, α, x) satisfies constraints (31), i.e. $x \in \tilde{Z}_{\mathrm{D}}$.

(c) Next we show that $\tilde{Z}_{\mathrm{D}} \subseteq Z_{\mathrm{D}}^*$. Given $x \in \tilde{Z}_{\mathrm{D}}$, there exists (γ_1, α) such that (γ_1, α, x) satisfies constraints in (31). For each $i \in [I] \setminus I_1(x)$, let $\widehat{\gamma}_i = 0$; otherwise, let $\widehat{\gamma}_i = \gamma_1/\alpha_i$.

Then for each $i \in [I] \setminus I_1(x)$, we have

$$
-g_1\widehat{\gamma}_i + (1-\epsilon)\psi_0(\widehat{\gamma}_i) + \epsilon\psi_{f_i}(\widehat{\gamma}_i, 1, x) = \epsilon\sup_{\xi \in \Xi}(-f_i(x, \xi)) \leq 0,
$$

where the first equality is from the definition of $\psi_\cdot(\cdot, \cdot, \cdot)$ and $\widehat{\gamma}_i = 0$, and the first inequality is due to $i \in [I] \setminus I_1(x)$, thus $f_i(x, \xi) \geq 0$ for all $\xi \in \Xi$. On the other hand, for for each $i \in I_1(x)$, we have

$$
\begin{aligned}
& -g_1\widehat{\gamma}_i + (1-\epsilon)\psi_0(\widehat{\gamma}_i) + \epsilon\psi_{f_i}(\widehat{\gamma}_i, 1, x) \\
&= -g_1\gamma_1/\alpha_i + (1-\epsilon)\sup_{\xi \in \Xi}(\phi_1(\xi)\gamma_1/\alpha_i) + \epsilon\sup_{\xi \in \Xi}(\phi_1(\xi)\gamma_1/\alpha_i - f_i(x, \xi)) \\
&= \frac{1}{\alpha_i}\left[-g_1\gamma_1 + (1-\epsilon)\sup_{\xi \in \Xi}(\gamma_1\phi_1(\xi)) + \epsilon\sup_{\xi \in \Xi}(\gamma_1\phi_1(\xi) - \alpha_i f_i(x, \xi))\right] \leq 0,
\end{aligned}
$$

where the first two equalities are due to the definition of $\psi_\cdot(\cdot, \cdot, \cdot)$ and $\widehat{\gamma}_i = \gamma_1/\alpha_i$, and the last inequality is due to (31a) and $\alpha_i > 0$.

Thus, $(\widehat{\gamma}, x)$ satisfies the constraints in (29); i.e., $x \in \tilde{Z}_{\mathrm{D}}$.

Thus, $Z_{\mathrm{D}} = \tilde{Z}_{\mathrm{D}} = Z_{\mathrm{D}}^*$. Since $\phi_0(\cdot)$ and $\{\phi_{f_i}(\cdot, 1, \cdot)\}_{i \in [I]}$ are convex functions, Z_{D} is a convex set. □

We remark that the proof of Theorem 5 only holds for the case of one moment inequality. If there is more than one moment inequality, it is difficult to project out the dual multipliers $\{\alpha_i\}_{i \in [I]}$.

Another observation is that in the reformulation (29), the constraints (29a) are I replications of (24a). Indeed, let us consider the following set

$$
Z_O = \left\{x : \inf_{\mathbb{P} \in \mathcal{P}} \mathbb{P}[\boldsymbol{\xi} : f_i(x, \boldsymbol{\xi}) \geq 0] \geq 1 - \epsilon, \forall i \in [I]\right\} \tag{32}
$$

which relaxes the requirement to satisfy all uncertain constraints, and is an outer approximation of Z_{D}. Next, we show that the relaxed set Z_O is equivalent to Z_{D} if there is only one moment inequality in $\mathcal{P}$.

Proposition 6. *If $|\mathcal{T}_1| = 0$ and $|\mathcal{T}_2| = 1$, then $Z_O = Z_{\mathrm{D}}$.*

Proof. Since set Z_O consists of I single DRCCP, therefore Theorem 4, Z_O is equivalent to

$$Z_O = \left\{ x : \begin{array}{l} -g_1\gamma_{1i} + (1-\epsilon)\psi_0\left(\gamma_{1i}\right) + \epsilon\psi_{f_i}\left(\gamma_{1i}, 1, x\right) \le 0, \forall i \in [I], \\ \gamma_{1i} \ge 0, \forall i \in [I], \end{array} \right\}$$

which clearly equals to (29) defining set Z_{D}. □

The result in Proposition 6 does not hold for a general ambiguity set $\mathcal{P}$ as illustrated by the following example:

Example 1. Let $n = 1, I = 2$ and $f_1(x,\xi) = \boldsymbol{\xi}x + T, f_2(x,\xi) = -\boldsymbol{\xi}x + T$ and

$$\mathcal{P} = \left\{\mathbb{P} \in \mathcal{P}_0\left(\mathbb{R}\right) : \mathbb{E}_{\mathbb{P}}[\boldsymbol{\xi}] = 0, \mathbb{E}_{\mathbb{P}}[\boldsymbol{\xi}^2] \le 1\right\},$$

then according to Theorem 3.1 [4] and Theorem 2 [23], we can reformulate sets Z_O and Z_{D} as follows:

$$Z_O = \left\{ x : |x| \le \sqrt{\frac{\epsilon}{1-\epsilon}T} \right\},$$

$$Z = \left\{ x : |x| \le \sqrt{\epsilon}T \right\}.$$

Clearly, when $\epsilon \to 1$, $Z_O \to \mathbb{R}$ but Z is always bounded. Hence the distance between Z and Z_O can be arbitrarily large.

Note that if there is one linear moment equality, we can reformulate Z_{D} as a disjunction of two sets by treating an equality constraint as two inequalities, then applying the same technique of Theorem 5.

Theorem 6. *If $|\mathcal{T}_1| = 1$ and $|\mathcal{T}_2| = 0$, then*

$$Z_{\mathrm{D}} = \bar{Z_{\mathrm{D}}}^1 \cup \bar{Z_{\mathrm{D}}}^2,$$

where

$$\bar{Z_{\mathrm{D}}}^1 = \left\{ x : \begin{array}{l} -g_1\widehat{\gamma}_i + (1-\epsilon)\psi_0\left(\widehat{\gamma}_i\right) + \epsilon\psi_{f_i}\left(\widehat{\gamma}_i, 1, x\right) \le 0, \forall i \in [I], \quad (33a) \\ \widehat{\gamma}_i \ge 0, \forall i \in [I], \quad (33b) \end{array} \right\}$$

$$\bar{Z_{\mathrm{D}}}^2 = \left\{ x : \begin{array}{l} -g_1\widehat{\gamma}_i + (1-\epsilon)\psi_0\left(\widehat{\gamma}_i\right) + \epsilon\psi_{f_i}\left(\widehat{\gamma}_i, 1, x\right) \le 0, \forall i \in [I], \quad (34a) \\ \widehat{\gamma}_i \le 0, \forall i \in [I], \quad (34b) \end{array} \right\}$$

and $\bar{Z_{\mathrm{D}}}^1, \bar{Z_{\mathrm{D}}}^2$ are convex sets.

Proof. From the proof of Theorem 5, we know

$$Z_{\mathrm{D}} = \left\{ x : \begin{array}{l} -g_1\gamma_1 + (1-\epsilon)\psi_0\left(\gamma_1\right) + \epsilon\psi_{f_i}\left(\gamma_1, \alpha_i, x\right) \le 0, \forall i \in I_1(x), \quad (35a) \\ \alpha_i \ge 0, \forall i \in [I], \quad (35b) \end{array} \right\}$$

and if $\gamma_1 \ge 0$, then $Z_{\mathrm{D}} = Z_{\mathrm{D}}^1$, while if $\gamma_1 \le 0$, then $Z_{\mathrm{D}} = Z_{\mathrm{D}}^2$. The conclusion follows by combining these two sets together with disjunction. □

Different from a result in [11], where the authors only showed tractability of a linear DRCCP with right-hand-side uncertainty (i.e., $f_i(x,\xi) = a_i^\top x - \xi^\top b^i$ with constants $a_i \in \mathbb{R}^n, b^i \in \mathbb{R}^m$), our results in Theorems 5 and 6 apply to both right-hand-side and left-hand-side uncertainty.

The following example shows an application of Theorem 5.

Example 2. Consider a stochastic multi-dimensional continuous knapsack problem. There are n items and I knapsacks, where for each item j, c_j represents its value, and $\boldsymbol{\xi}_i$ represents ith knapsack's random weight vector, and let b_i be the total capacity of ith knapsack. The variable x_j denotes the portion of jth item being picked. Suppose that we know the total absolute deviation of weight, thus $\mathcal{P}$ is defined as

$$\mathcal{P} = \left\{\mathbb{P} \in \mathcal{P}_0\left(\Xi\right) : \mathbb{E}_{\mathbb{P}}[\|\tilde{\boldsymbol{\xi}}\|_1] \leq g\right\},$$

where $\Xi = \mathbb{R}^{n\times I}$ and $\boldsymbol{\xi}_i = \tilde{\boldsymbol{\xi}}_i + \mu_i$ for all $i \in [I]$.

Now the distributionally robust multi-dimensional continuous knapsack problem is formulated as

$$\begin{aligned} v^* = \max_{x\in[0,1]^n} \quad & c^\top x, \\ \text{s.t.} \quad & \inf_{\mathbb{P}\in\mathcal{P}} \mathbb{P}[x^\top(\mu_i + \tilde{\xi}_i) \leq b_i, \forall i \in [I]] \geq 1-\epsilon. \end{aligned} \tag{36}$$

Note that for each $i \in [I]$, the ith uncertain constraint is $f_i(x,\xi) = b_i - x^\top \mu_i - x^\top\tilde{\xi}_i$. By Proposition 1 with $q = 1$, we have

$$\psi_{f_i}(\gamma, \alpha_i, x) = \begin{cases} -\alpha_i(b_i - \mu_i^\top x), & \text{if } \|\alpha_i x\|_\infty \leq \gamma, \\ +\infty, & \text{otherwise}, \end{cases}$$

and $\psi_0(\gamma) = 0$.

Thus, according to Theorem 5, Problem (36) is equivalent to the following linear program

$$v^* = \max \quad c^\top x, \tag{37a}$$

$$\text{s.t.} \quad x \in [0,1]^n, \tag{37b}$$

$$g\gamma_i \leq \epsilon(b_i - \mu_i^\top x), \forall i \in [I], \tag{37c}$$

$$\|x\|_\infty \leq \gamma_i, \forall i \in [I], \tag{37d}$$

$$\gamma_i \geq 0, \forall i \in [I]. \tag{37e}$$

4.3 Linear moment constraints

Here, we show that if $\Xi = \mathbb{R}^m$ and the ambiguity set is defined only by linear moment constraints, then the set Z_{D} is equal to X_C as defined in (14). Hence, DRCCP is equivalent to a robust convex program.

Theorem 7. *Suppose $\Xi = \mathbb{R}^m$ and $\{\phi_t(\xi)\}_{t\in\mathcal{T}_1\cup\mathcal{T}_2}$ are all linear functions, then*

$$Z_{\mathrm{D}} = X_C = \left\{x \in \mathbb{R}^n : f_i(x,\xi) \geq 0, \forall \xi \in \mathbb{R}^m, i \in [I]\right\}. \tag{38}$$

Proof. Since $\{\phi_t(\xi)\}_{t\in\mathcal{T}_1\cup\mathcal{T}_2}$ are all linear functions, let $\phi_t(\xi) = w_t^\top \xi$ for all $t \in \mathcal{T}_1\cup\mathcal{T}_2$. We only need to show that $I_1(x) = \emptyset$ for any given $x \in Z_{\mathrm{D}}$, where $I_1(x) := \{i \in [I] : \exists \xi \in \Xi, f_i(x,\xi) < 0\}$.

Suppose that $I_1(x) \neq \emptyset$ for some $x \in Z_{\mathrm{D}}$. Since $\phi_t(\xi) = w_t^\top \xi$ for all $t \in \mathcal{T}_1\cup\mathcal{T}_2$ and $\Xi = \mathbb{R}^m$, (6b) in Theorem 1 yields have

$$\lambda + \sup_{\xi\in\mathbb{R}^m} \sum_{t\in\mathcal{T}_1\cup\mathcal{T}_2} w_t^\top \xi \gamma_t \leq 1,$$

which implies that

$$\sum_{t\in\mathcal{T}_1\cup\mathcal{T}_2} w_t\gamma_t = 0.$$

Meanwhile, in (3), for any $\mathbb{P}\in\mathcal{P}_0(\Xi)$, we have $\mathbb{E}_{\mathbb{P}}[w_t^\top\boldsymbol{\xi}] = g_t$ for all $t\in\mathcal{T}_1$ and $\mathbb{E}_{\mathbb{P}}[w_t^\top\boldsymbol{\xi}]\geq g_t$ for all $t\in\mathcal{T}_2$. Multiplying these equalities and inequalities with $\{\gamma_t\}_{t\in\mathcal{T}_1\cup\mathcal{T}_2}$ (note that $\gamma_t\geq 0$ for each $t\in\mathcal{T}_2$), we have

$$\mathbb{E}_{\mathbb{P}}\left[\sum_{t\in\mathcal{T}_1\cup\mathcal{T}_2}\gamma_t w_t^\top\boldsymbol{\xi}\right]\geq\sum_{t\in\mathcal{T}_1\cup\mathcal{T}_2} g_t\gamma_t,$$

i.e., $\sum_{t\in\mathcal{T}_1\cup\mathcal{T}_2} g_t\gamma_t\leq 0$ as $\sum_{t\in\mathcal{T}_1\cup\mathcal{T}_2} w_t\gamma_t = 0$.

Now from (6c) in Theorem 1, for each $i\in I_1(x)$, we have

$$\lambda+\psi_{f_i}(\gamma,\alpha_i,x) = \lambda+\sup_{\xi\in\Xi}\left(\sum_{t\in\mathcal{T}_1\cup\mathcal{T}_2}\gamma_t w_t^\top\xi-\alpha_i f_i(x,\xi)\right) = \lambda+\sup_{\xi\in\Xi}\left(-\alpha_i f_i(x,\xi)\right),$$

where the first equality is due to the definition of $\phi_\cdot(\cdot,\cdot,\cdot)$ and the second equality is because of $\sum_{t\in\mathcal{T}_1\cup\mathcal{T}_2} w_t\gamma_t = 0$. As we know $i\in I_1(x)$ and $\alpha_i>0$ from Corollary 2, we must have $\sup_{\xi\in\Xi}(-\alpha_i f_i(x,\xi))>0$. Hence, (6c) (i.e., $\lambda+\psi_{f_i}(\gamma,\alpha_i,x)\leq 0$) implies that $\lambda<0$.

On the other hand, (6a) (i.e., $\lambda+\sum_{t\in\mathcal{T}_1\cup\mathcal{T}_2} g_t\gamma_t\geq 1-\epsilon$) implies that $\lambda\geq 1-\epsilon$ since $\sum_{t\in\mathcal{T}_1\cup\mathcal{T}_2} g_t\gamma_t\leq 0$. Thus, we have a contradiction. □

This proposition suggests us that only considering first-moment information might not provide us a sufficient characterization of the ambiguous set and hence more nonlinear moment constraints are needed for a more realistic reformulation.

4.4 Nonlinear positively homogeneous moment constraints

Now we consider the case of multiple (possibly, nonlinear) moment constraints. Let us begin with the following technical lemma.

Lemma 5. *Suppose that Ξ is a closed convex cone, $\phi_t(\xi)$ is positively homogeneous on Ξ for each $t\in\mathcal{T}_2$, then Z_{D} is equivalent to*

$$Z_{\mathrm{D}} = \left\{ x\in\mathbb{R}^n : \begin{array}{ll} \sum_{t\in\mathcal{T}_1\cup\mathcal{T}_2} g_t\gamma_t\geq-\epsilon, & \text{(39a)}\\ \bar{\psi}_0(\gamma)\leq 0, & \text{(39b)}\\ 1+\psi_{f_i}(\gamma,\alpha_i,x)\leq 0,\forall i\in I_1(x), & \text{(39c)}\\ \gamma_t\geq 0,\forall t\in\mathcal{T}_2,\alpha_i\geq 0,\forall i\in I_1(x), & \text{(39d)} \end{array}\right\}$$

where convex mapping $\bar{\psi}_0$ describes the domain of ϕ_0, i.e.,

$$\psi_0(\gamma) = \begin{cases} 0, & \textit{if } \bar{\psi}_0(\gamma)\leq 0,\\ +\infty, & \textit{otherwise,}\end{cases}$$

Proof. Since $\phi_t(\xi)$ is linear for each $t\in\mathcal{T}_1$ and is positively homogeneous on Ξ for each $t\in\mathcal{T}_2$, by part (iii) of Lemma 4, (6) reduces to

$$Z_{\mathrm{D}} = \left\{ x : \begin{array}{ll} \lambda + \sum_{t \in \mathcal{T}_1 \cup \mathcal{T}_2} g_t \gamma_t \geq 1 - \epsilon, & \text{(40a)} \\ \lambda \leq 1, & \text{(40b)} \\ \lambda + \psi_{f_i}(\gamma, \alpha_i, x) \leq 0, \forall i \in I_1(x), & \text{(40c)} \\ (39b), (39d). & \end{array} \right\}$$

It remains to show that for any x in Z_{D}, we always have $\lambda = 1$.

Consider an $x \in Z_{\mathrm{D}}$ with $(\lambda, \gamma, \alpha)$ such that $(\lambda, \gamma, \alpha, x)$ satisfies (40) and $\lambda < 1$. By Corollary 1 and $\psi_0(\gamma) = 0$ in its domain, we must have $\lambda \geq 1 - \epsilon$. Now construct a new solution $(\bar{\lambda}, \bar{\gamma}, \bar{\alpha}, x)$ as $\bar{\lambda} = 1, \bar{\gamma} = \frac{\gamma}{\lambda}, \bar{\alpha} = \frac{\alpha}{\lambda}$. Clearly, $(\bar{\lambda}, \bar{\gamma}, \bar{\alpha}, x)$ also satisfies (40). Thus, we can always set $\lambda = 1$ in (40), which yields (39).

By Lemma 4, $\psi_0(\gamma)$ is convex in λ, therefore $\bar{\psi}_0$ is a convex mapping. □

Next, we identify sufficient conditions for the set Z_{D} to be convex if the moment constraints are defined by positively homogeneous functions $\{\phi_t(\xi)\}_{t \in \mathcal{T}_2}$.

Theorem 8. *Suppose that Ξ is a closed convex cone, $\phi_t(\xi)$ is positively homogeneous on Ξ for each $t \in \mathcal{T}_2$, and for each $i \in [I]$, $f_i(x, \xi) = w_i(x) - h_i(\xi)$, where $w_i(x) : \mathbb{R}^n \to$ is a concave function and $h_i(\xi) : \Xi \to \mathbb{R}$ is a concave function and positively homogeneous on Ξ. Then Z_{D} is formulated as the following convex set*

$$Z_{\mathrm{D}} = \left\{ x : \begin{array}{ll} \sum_{t \in \mathcal{T}_1 \cup \mathcal{T}_2} g_t \gamma_t \geq -\epsilon, & \text{(41a)} \\ 1 \leq \alpha_i w_i(x), \forall i \in [I] \setminus I_2, & \text{(41b)} \\ 0 \leq w_i(x), \forall i \in I_2, & \text{(41c)} \\ \bar{\psi}_0(\gamma) \leq 0, & \text{(41d)} \\ \bar{\psi}_{-h_i}(\gamma, -\alpha_i) \leq 0, \forall i \in [I] \setminus I_2, & \text{(41e)} \\ \gamma_t \geq 0, \forall t \in \mathcal{T}_2, \alpha_i \geq 0, \forall i \in [I] \setminus I_2, & \text{(41f)} \end{array} \right\}$$

where $I_2 := \{i \in [I] : h_i(\xi) \leq 0, \forall \xi \in \Xi\}$, and for each $i \in [I]$,

$$\psi_{f_i}(\gamma, \alpha_i, x) = -\alpha_i w_i(x) + \psi_{-h_i}(\gamma, \alpha_i, x) = \begin{cases} -\alpha_i w_i(x), & \text{if } \bar{\psi}_{-h_i}(\gamma, \alpha_i) \leq 0, \\ +\infty, & \text{otherwise}, \end{cases}$$

with the convex mapping $\bar{\psi}_{-h_i}(\gamma, \alpha_i)$.

Proof. We first note that $0 \leq w_i(x)$ for all $i \in I_2$ for each $x \in Z_{\mathrm{D}}$. We prove it by contradiction. Suppose that $I_1(x) \cap I_2 \neq \emptyset$. Thus, in (39c), for each $i \in I_1(x) \cap I_2$, we have

$$\begin{aligned} 0 \geq 1 + \psi_{f_i}(\gamma, \alpha_i, x) = \psi_{f_i}(\gamma, \alpha_i, x) &= 1 - \alpha_i w_i(x) + \psi_{-h_i}(\gamma, \alpha_i, x) \\ &= \begin{cases} 1 - \alpha_i w_i(x), & \text{if } \bar{\psi}_{-h_i}(\gamma, \alpha_i) \leq 0, \\ +\infty, & \text{otherwise}, \end{cases} \end{aligned}$$

where the last inequality due to positive homogeneity of $\{\phi_t(\xi)\}_{t \in \mathcal{T}_1 \cup \mathcal{T}_2}$ and $h_i(\xi)$. Thus, we must have $\bar{\psi}_{-h_i}(\gamma, \alpha_i) \leq 0$ and $\alpha_i w_i(x) \geq 1$. This implies that $w_i(x) \geq \sup_{\xi} h_i(\xi)$, i.e. $f_i(x, \xi) = w_i(x) - h_i(\xi) \geq 0$ for all $\xi \in \Xi$. Therefore, $i \notin I_1(x)$, contradiction.

Thus, set Z_{D} is now equivalent to $\{0 \leq w_i(x), \forall i \in I_2\} \cap \tilde{Z_{\mathrm{D}}}$, where

$$\tilde{Z_{\mathrm{D}}} = \left\{ x \in \mathbb{R}^n : \inf_{\mathbb{P} \in \mathcal{P}} \mathbb{P}[\boldsymbol{\xi} : w_i(x) \geq h_i(\boldsymbol{\xi}), \forall i \in [I] \setminus I_2] \geq 1 - \epsilon \right\}$$

Now from the proof in Theorem 2, we know

$$\tilde{Z_{\mathrm{D}}} \setminus \tilde{Y}_C \subseteq \mathcal{C} := \{x \in \tilde{Z_{\mathrm{D}}} : \exists i \in [I] \setminus I_2, w_i(x) = \sup_{\xi \in \Xi} h_i(\xi)\},$$

where

$$\tilde{Y}_C = \left\{ x \in \mathbb{R}^n : \begin{array}{l} \sum_{t \in \mathcal{T}_1 \cup \mathcal{T}_2} g_t \gamma_t \geq -\epsilon, \\ \lambda + \psi_0(\gamma) \leq 1, \\ \lambda + \psi_{f_i}(\gamma, \alpha_i, x) \leq 0, \forall i \in [I] \setminus I_2, \\ \gamma_t \geq 0, \forall t \in \mathcal{T}_2, \alpha_i \geq 0, \forall i \in [I] \setminus I_2. \end{array} \right\} \tag{42}$$

Since for each $i \in [I] \setminus I_2$, we have $h_i(\xi) > 0$ for some $\xi \in \Xi$, thus by positive homogeneity, we must have $\sup_{\xi \in \Xi} h_i(\xi) = \infty$. Hence, $\mathcal{C} = \emptyset$ and $\tilde{Y}_C = \tilde{Z_{\mathrm{D}}}$. By by Lemma 5, we can $\lambda = 1$ in set $\tilde{Y}_C$. Thus,

$$Z_{\mathrm{D}} = \{0 \leq w_i(x), \forall i \in I_2\} \cap \tilde{Z_{\mathrm{D}}} = \{0 \leq w_i(x), \forall i \in I_2\} \cap \tilde{Y}_C$$

leads to

$$Z_{\mathrm{D}} = \left\{ x \in \mathbb{R}^n : \begin{array}{l} \sum_{t \in \mathcal{T}_1 \cup \mathcal{T}_2} g_t \gamma_t \geq -\epsilon, \\ \bar{\psi}_0(\gamma) \leq 0, \\ 0 \leq w_i(x), \forall i \in I_2, \\ 1 + \psi_{f_i}(\gamma, \alpha_i, x) \leq 0, \forall i \in [I] \setminus I_2, \\ \gamma_t \geq 0, \forall t \in \mathcal{T}_2, \alpha_i \geq 0, \forall i \in [I] \setminus I_2. \end{array} \right\}$$

Note that for each $i \in I$, $\psi_{f_i}(\gamma, \alpha_i, x) = -\alpha_i w_i(x) + \psi_{-h_i}(\gamma, \alpha_i, x)$. Since function h_i is positive homogeneous in ξ and irrelevant with x, by Lemma 4, $\psi_{-h_i}(\gamma, \alpha_i, x)$ is convex in (γ, α_i) and hence its domain is. Hence, we arrive at (41). □

Note that (41b) is a convex constraint for each $i \in [I]$ and is second order cone representable by introducing a new variable $0 \leq q_i \leq w_i(x)$. Then (41b) is equivalent to

$$2^2 + (\alpha_i - q_i)^2 \leq (\alpha_i + q_i)^2, 0 \leq q_i \leq w_i(x), \forall i \in [I].$$

Note that Theorem 2 in [11] is a special case of Theorem 8, where in [11], it is assume that $w_i(x)$ is an affine function and $h_i(\xi)$ is a linear function for each $i \in [I]$.

The following example demonstrates an application of Theorem 8.

Example 3. Consider a stochastic lot-sizing problem. There are I time periods and at each time period $i \in [I]$, ξ_i represents the random demand and c_i, f_i are the production cost and fixed cost, respectively. The production for each time period cannot exceed M. There are two types of decision variables, x_i represents production level and y_i represents production set up at time i, i.e., $y_i = 1$ if $x_i > 0$; 0, otherwise. Suppose that we know the mean of the demand at each period and the total deviation of demand, thus $\mathcal{P}$ is defined as

$$\mathcal{P} = \left\{ \mathbb{P} \in \mathcal{P}_0(\Xi) : \mathbb{E}_{\mathbb{P}}[\tilde{\boldsymbol{\xi}}] = 0, \mathbb{E}_{\mathbb{P}}[\|\tilde{\boldsymbol{\xi}}\|_q] \leq g \right\},$$

where $\Xi = \mathbb{R}^I$ and $\boldsymbol{\xi} = \tilde{\boldsymbol{\xi}} + \mu$.

Now the entire distributionally robust lot-sizing problem is formulated as

$$\begin{aligned} v^* = \min \quad & c^\top x + f^\top y, \\ \text{s.t.} \quad & x_i \leq M y_i, \forall i \in [I], \end{aligned} \tag{43a}$$

$$\inf_{\mathbb{P}\in\mathcal{P}} \mathbb{P}\left[\sum_{j=1}^{i} x_j \geq \sum_{j=1}^{i}(\tilde{\xi}_j+\mu_j), \forall i \in [I]\right] \geq 1-\epsilon, \tag{43b}$$

$$y \in \{0,1\}^I, x \geq 0. \tag{43c}$$

Let us define matrix $A \in \mathbb{R}^{I\times I}$ as $A(i,j) = 1$ if $j \leq i$; 0, otherwise; and $a_{i\cdot}$ denote ith row of A. From Proposition 3, we have $w_i(x) = a_{i\cdot}(x-\mu), h_i(\xi) = a_{i\cdot}\tilde{\xi}$, thus $\bar{\psi}_0(\gamma) = \|\gamma_1\|_{\frac{q}{q-1}} - \gamma_2$, $\bar{\psi}_{-h_i}(\gamma, -\alpha_i) = \|\gamma_1 + \alpha_i a_{i\cdot}\|_{\frac{q}{q-1}} - \gamma_2$ in Theorem 8. It is easy to see that

$$I_2 = \left\{ i \in [I] : \sup_{\tilde{\xi}\in\mathbb{R}^I} \left[\sum_{j=1}^{i} \tilde{\xi}_j\right] \leq 0 \right\} = \emptyset.$$

Thus, set Z_{D} is reformulated as

$$Z_C = \left\{ x : \begin{array}{ll} g\gamma_2 \leq \epsilon, & \text{(44a)} \\ 1 \leq \alpha_i a_{i\cdot}(x-\mu), \forall i \in [I], & \text{(44b)} \\ \|\gamma_1\|_{\frac{q}{q-1}} \leq \gamma_2, \|\gamma_1 + \alpha_i a_{i\cdot}\|_{\frac{q}{q-1}} \leq \gamma_2, \forall i \in [I], & \text{(44c)} \\ \gamma_2 \geq 0, \alpha_i \geq 0, \forall i \in [I]. & \text{(44d)} \end{array} \right\}$$

In the above formulation, note that a larger γ_2 value implies a larger feasible region. Thus, at optimality, we must have $\gamma_2 = \frac{\epsilon}{g}$. Therefore, Problem (43) is equivalent to the following mixed integer convex program:

$$v^* = \min \quad c^\top x + f^\top y, \tag{45a}$$

$$\text{s.t.} \quad x_i \leq M y_i, \forall i \in [I], \tag{45b}$$

$$1 \leq \alpha_i a_{i\cdot}(x-\mu), \forall i \in [I], \tag{45c}$$

$$\|\gamma_1\|_{\frac{q}{q-1}} \leq \frac{\epsilon}{g}, \|\gamma_1 + \alpha_i a_{i\cdot}\|_{\frac{q}{q-1}} \leq \frac{\epsilon}{g}, \forall i \in [I], \tag{45d}$$

$$\alpha_i \geq 0, \forall i \in [I], y \in \{0,1\}^I, x \in \mathbb{R}^I_+. \tag{45e}$$

Note that variables α in (45) can be interpreted as a safety buffer, which guarantee that the inequalities (45c) are robust.

5 Binary DRCCP

In this section, we consider the case of binary decision variables and general moment ambiguity set $\mathcal{P}$ defined in (3), i.e., $S \subseteq \{0,1\}^n$ and linear chance constraints. We will first derive a mixed integer convex reformulation and then present a numerical study. Please note that the convexity results in Section 4 also apply to binary DRCCP.

5.1 Mixed integer convex formulation

Proposition 7. *Suppose $S \subseteq \{0,1\}^n$, $f_i(x,\xi) = (A^i x + a^i)^\top \xi + B^i x + b^i$ for each $i \in [I]$ and α in* (15) *can be upper bounded by a vector M for any $x \in S$. Consider a convex set*

$$\bar{Y}_C = \left\{ x : \begin{array}{ll} \lambda + \sum_{t \in \mathcal{T}_1 \cup \mathcal{T}_2} g_t \gamma_t \geq 1 - \epsilon, & \text{(46a)} \\ \lambda + \psi_0(\gamma) \leq 1, & \text{(46b)} \\ \lambda + \psi_{\bar{f}_i}(\gamma, 1, (\alpha_i, y_i)) \leq 0, \forall i \in [I], & \text{(46c)} \\ 0 \leq y_j^i \leq M_i x_j, \alpha_i - M_i(1 - x_j) \leq y_j^i \leq \alpha_i, \forall i \in [I], j \in [n], & \text{(46d)} \\ \gamma_t \geq 0, \forall t \in \mathcal{T}_2, \alpha_i \geq 0, \forall i \in [I], & \text{(46e)} \end{array} \right\}$$

where for each $i \in [I]$, $\bar{f}_i(\alpha_i, y_i) = (A^i y_i + a^i \alpha_i)^\top \xi + B^i y_i + b^i \alpha_i$, then

$$S \cap \bar{Y}_C = S \cap Y_C \subseteq S \cap Z_{\mathrm{D}}.$$

Proof. When $f_i(x,\xi) = (A^i x + a^i)^\top \xi + B^i x + b^i$ for each $i \in [I]$, by Definition 1, we have

$$\psi_{f_i}(\gamma, \alpha_i, x) = \sup_{\xi \in \Xi} \left(\sum_{t \in \mathcal{T}_1 \cup \mathcal{T}_2} \phi_t(\xi)\gamma_t - \alpha_i((A^i x + a^i)^\top \xi + B^i x + b^i) \right).$$

Since $\alpha_i \leq M_i$ for each $i \in [I]$, let us define new variables y such that $y^i = \alpha_i x$, which can be linearized via McCormick inequalities [16] as $0 \leq y_j^i \leq M_i x_j, \alpha_i - M_i(1 - x_j) \leq y_j^i \leq \alpha_i$, i.e., (46d). This linearization is exact for any $x \in \{0,1\}^n$. Thus, $S \cap \bar{Y}_C = S \cap Y_C$. □

Proposition 7 tells that to optimize over $S \cap Y_C$ is equivalent to optimize over $S \cap \bar{Y}_C$, which is a mixed integer convex set instead of a mixed integer nonconvex set. Above we have assumed the existence of vector M and one sufficient condition is Corollary 3. Moreover, the next theorem shows that under some other cases, the variables α in (46) have closed-form bounds.

Theorem 9. *Suppose the ambiguity set be defined as*

$$\mathcal{P} = \left\{ \mathbb{P} \in \mathcal{P}_0(\Xi) : \mathbb{E}_{\mathbb{P}}[\boldsymbol{\xi}_i] = \mu_i, \mathbb{E}_{\mathbb{P}}[(\boldsymbol{\xi}_i - \mu_i)(\boldsymbol{\xi}_i - \mu_i)^\top] \preceq \Sigma_i, \forall i \in [I] \right\},$$

where $\Xi = \mathbb{R}^{n \times I}$, and $\Sigma_i \succ 0$ for each $i \in [I]$. Suppose $S \subseteq \{0,1\}^n$, $f_i(x,\xi) = (Ax + a)^\top \xi_i + B^i x + b^i$ for each $i \in [I]$, then

$$S \cap (X_C \cup \bar{Y}_C) = S \cap Z_{\mathrm{D}}.$$

Sets X_C and $\bar{Y}_C$ in (46) *are defined as*

$$X_C = \left\{ x : Ax + a = 0, B^i x + b^i \geq 0, \forall i \in [I] \right\}, \tag{47}$$

and

$$\bar{Y}_C = \left\{ x : \begin{array}{l} \lambda - \sum_{j\in[I]} \left[\left\langle A^\top \Sigma_j A, w_{\cdot\cdot j} \right\rangle + 2a^\top \Sigma_j A z_{\cdot j} + a^\top \Sigma_j a \gamma_{2j} \right] \geq 1 - \epsilon, \\ \lambda + \sum_{j\in[I]} t_{0j} \leq 1, \\ \lambda + \sum_{j\in[I]} t_{ij} \leq \alpha_i \left((Ax+a)^\top \mu_i + b^i \right) + B^i y^i, \forall i \in [I], \\ \gamma_{1j}^2 \leq 4 t_{ij} \gamma_{2j}, \forall i \in [I], j \in [I] \setminus \{i\}, \\ (\gamma_{1i} - \alpha_i)^2 \leq 4 t_{ii} \gamma_{2i}, \gamma_{1i}^2 \leq 4 t_{0i} \gamma_{2i}, \forall i \in [I], \\ 0 \leq y_j^i \leq M_i x_j, \alpha_i - M_i(1 - x_j) \leq y_j^i \leq \alpha_i, \forall i \in [I], j \in [n], \\ 0 \leq z_{ij} \leq \frac{\epsilon}{\underline{\delta}\eta} x_i, \gamma_{2j} - \frac{\epsilon}{\underline{\delta}\eta}(1 - x_i) \leq z_{ij} \leq \gamma_{2j}, \forall i, j \in [I], \\ 0 \leq w_{ikj} \leq \frac{\epsilon}{\underline{\delta}\eta} x_i, 0 \leq w_{ikj} \leq \frac{\epsilon}{\underline{\delta}\eta} x_k, \\ \gamma_{2j} - \frac{\epsilon}{\underline{\delta}\eta}(2 - x_i - x_j) \leq w_{ikj} \leq \gamma_{2j}, \forall i, j, k \in [I], \\ \gamma_{2i} \geq 0, \alpha_i \geq 0, \forall i \in [I]. \end{array} \right\} \tag{48}$$

with

$$M_i = \frac{4\epsilon}{\underline{\delta}\eta} \left[(b^i + \mu_i^\top a^i + \|B^i + \mu_i^\top A^i\|_1) + \sqrt{(b^i + \mu_i^\top a^i + \|B^i + \mu_i^\top A^i\|_1)^2 + \underline{\delta}\eta - \frac{\underline{\delta}\eta}{2\epsilon}} \right]$$

for each $i \in [I]$, where $\eta = \min_{x\in\{0,1\}^n : Ax+a\neq 0} \|Ax + a\|_2^2$, $\underline{\delta}$ is the smallest eigenvalue of matrices $\{\Sigma_j\}_{j\in[I]}$, and $w_{\cdot\cdot j}$ denotes the matrix (w_{ikj}) for each j and $z_{\cdot j}$ denotes the vector (z_{ij}) for each j.

Proof. We will separate the proof into four parts.

(i) Suppose for any $x \in \{0, 1\}^n$ such that $Ax + a = 0$, then by (1c), we must have $B^i x + b^i \geq 0$ for all $i \in [I]$. Hence, this implies that $x \in X_C$. Also note that for any $x \in S \cap Z_{\mathrm{D}} \setminus X_C$, we must have $I_1(x) = [I]$, i.e., $x \in S \cap Y_C$. Therefore, $S \cap (X_C \cup Y_C) = S \cap Z_{\mathrm{D}}$.

It remains to show the existence of vector M such that $\alpha_i \leq M_i$ for each $i \in [I]$ and $x \in S \cap Z_{\mathrm{D}} \setminus X_C$.

(ii) From now on, we assume that $\|Ax + a\|_2 \neq 0$ for all $x \in S \cap Y_C$. Define $\xi_i = \tilde{\xi}_i + \mu_i$ for each $i \in [I]$. Then the ambiguity set is equivalent to

$$\mathcal{P} = \left\{ \mathbb{P} \in \mathcal{P}_0(\Xi) : \mathbb{E}_{\mathbb{P}}[\tilde{\boldsymbol{\xi}}_i] = 0, \mathbb{E}_{\mathbb{P}}[\tilde{\boldsymbol{\xi}}_i \tilde{\xi}_i^\top] \preceq \Sigma_i, \forall i \in [I] \right\},$$

where $\Xi = \mathbb{R}^{n\times I}$. Also, the uncertain constraint is $f_i(x, \tilde{\boldsymbol{\xi}}_i) = (Ax + a)^\top \tilde{\boldsymbol{\xi}}_i + (Ax + a)^\top \mu_i + B^i x + b^i$ for each $i \in [I]$.

For any given $x \in Y_C$, replacing $\tilde{\boldsymbol{\xi}}_i$ by $\boldsymbol{\zeta}_i = (Ax + a)^\top \tilde{\boldsymbol{\xi}}_i$ and by the standard random variable changing (c.f. [8]) and Theorem 1 in [18], the ambiguity set can be further replaced as

$$\mathcal{P} = \left\{ \mathbb{P} \in \mathcal{P}_0\left(\mathbb{R}^I\right) : \mathbb{E}_{\mathbb{P}}[\boldsymbol{\zeta}_i] = 0, \mathbb{E}_{\mathbb{P}}[\boldsymbol{\zeta}_i^2] \leq (Ax + a)^\top \Sigma_i (Ax + a), \forall i \in [I] \right\},$$

with the uncertain constraint $f_i(x, \boldsymbol{\zeta}_i) = \boldsymbol{\zeta}_i + (Ax + a)^\top \mu_i + B^i x + b^i$ for each $i \in [I]$.

(iii) Next by Definition 1, we have

$$\psi_{f_i}(\gamma, \alpha_i, x) = \sum_{j \in [I], j \neq i} \sup_{\zeta_j \in \mathbb{R}} \left[\gamma_{1j}\zeta_j - \gamma_{2j}\zeta_j^2 \right]$$
$$+ \sup_{\zeta_i \in \mathbb{R}} \left[\gamma_{1i}\zeta_i - \gamma_{2i}\zeta_i^2 - \alpha_i \left(\zeta_i + (Ax+a)^\top \mu_i + B^i x + b^i \right) \right]$$
$$= \min \left\{ \sum_{j \in [I]} t_{ij} - \alpha_i \left((Ax+a)^\top \mu_i + B^i x + b^i \right) : \gamma_{1j}^2 \leq 4 t_{ij} \gamma_{2j}, \forall j \in [I] \setminus \{i\}, \right.$$
$$\left. (\gamma_{1i} - \alpha_i)^2 \leq 4 t_{ii} \gamma_{2i} \right\}$$

where the second equality is due to 1-dimensional S-Lemma in [1]. Similarly,

$$\psi_0(\gamma) = \min \left\{ \sum_{j \in [I]} t_{0j} : \gamma_{1j}^2 \leq 4 t_{0j} \gamma_{2j}, \forall j \in [I] \right\}.$$

Then by replacing minimum operator with its equivalent "existence" argument, set Y_C can be formulated as

$$Y_C = \left\{ x : \begin{array}{ll} \lambda - \sum_{j \in [I]} \gamma_{2j} (Ax+a)^\top \Sigma_j (Ax+a) \geq 1 - \epsilon, & \text{(49a)} \\ \lambda + \sum_{j \in [I]} t_{0j} \leq 1, & \text{(49b)} \\ \lambda + \sum_{j \in [I]} t_{ij} \leq \alpha_i \left((Ax+a)^\top \mu_i + B^i x + b^i \right), \forall i \in [I], & \text{(49c)} \\ \gamma_{1j}^2 \leq 4 t_{ij} \gamma_{2j}, \forall i \in [I], j \in [I] \setminus \{i\}, & \text{(49d)} \\ (\gamma_{1i} - \alpha_i)^2 \leq 4 t_{ii} \gamma_{2i}, \forall i \in [I], & \text{(49e)} \\ \gamma_{1j}^2 \leq 4 t_{0j} \gamma_{2j}, \forall j \in [I], & \text{(49f)} \\ \gamma_{2i} \geq 0, \alpha_i \geq 0, \forall i \in [I]. & \text{(49g)} \end{array} \right\}$$

(iv) Now we show the existence of upper bounds on $\{\alpha_i\}_{i \in [I]}$ and $\{\gamma_{2i}\}_{i \in [I]}$ with following four steps.

(a) First of all, we observe that $t_{0j} \geq 0$ for each $j \in [I]$; otherwise, it contradicts that (49f). Thus, (49b) implies that $\lambda \leq 1$ and hence $\sum_{j \in [I]} \gamma_{2j} (Ax+a)^\top \Sigma_j (Ax+a) \leq \epsilon$.
Let $\underline{\delta}$ be the smallest eigenvalue of matrices $\{\Sigma_j\}_{j \in [I]}$. Hence, $\Sigma_j \succeq \underline{\delta} I_e$ for each $j \in [I]$, where I_e is the identity matrix.
Clearly, $\underline{\delta} > 0$ since $\Sigma_j \succ 0$ for each $j \in [I]$. Thus, $\sum_{j \in [I]} \gamma_{2j} (Ax+a)^\top \Sigma_j (Ax+a) \leq \epsilon$ and $\Sigma_j \succeq \underline{\delta} I$ for each $j \in [I]$ imply that

$$\underline{\delta} \|Ax+a\|_2^2 \sum_{j \in [I]} \gamma_{2j} \leq \epsilon.$$

Let $\eta = \min_{x \in \{0,1\}^n : Ax+a \neq 0} \|Ax+a\|_2^2$, thus γ_{2j} is bounded by $\frac{\epsilon}{\underline{\delta}\eta}$.

(b) Therefore, (49d), (49e) and (49f) can be relaxed by replacing $\{\gamma_{2j}\}_{j \in [I]}$ with their lower bound $\frac{\epsilon}{\underline{\delta}\eta}$ as below

$$t_{ij} \geq \frac{\delta\eta}{4\epsilon}\gamma_{1j}^2, \forall i \in [I], j \in [I] \setminus \{i\}, \tag{50}$$

$$t_{ii} \geq \frac{\delta\eta}{4\epsilon}(\gamma_{1i} - \alpha_i)^2, \forall i \in [I], \tag{51}$$

$$t_{0j} \geq \frac{\delta\eta}{4\epsilon}\gamma_{1j}^2, \forall j \in [I]. \tag{52}$$

(c) Since $t_{ij}, t_{0j} \geq 0$ for all $i \in [I], j \in [I] \setminus \{i\}$ and $\Sigma_j \succ 0$ for all $j \in [I]$, thus (49a), (49b) and (49c) imply that

$$\begin{aligned}
&\lambda \geq 1 - \epsilon,\\
&\lambda + t_{ii} \leq \alpha_i[b^i + \mu_i^\top a + (B^i + \mu_i^\top A)x], \forall i \in [I],\\
&\lambda + t_{0i} \leq 1, \forall i \in [I].
\end{aligned}$$

Together with (51) and (52), these above inequalities are further reduced to

$$\begin{aligned}
&-2\lambda \leq -2(1-\epsilon),\\
&\lambda + \frac{\delta\eta}{4\epsilon}(\gamma_{1i} - \alpha_i)^2 - \alpha_i[b^i + \mu_i^\top a + (B^i + \mu_i^\top A)x] \leq 0, \forall i \in [I],\\
&\lambda + \frac{\delta\eta}{4\epsilon}\gamma_{1i}^2 \leq 1, \forall i \in [I].
\end{aligned}$$

Summing these inequalities up for each $i \in [I]$ yields

$$\frac{\delta\eta}{4\epsilon}\left[(\gamma_{1i} - \alpha_i)^2 + \gamma_{1i}^2\right] - \alpha_i[b^i + \mu_i^\top a + (B^i + \mu_i^\top A)x] \leq 2\epsilon - 1.$$

(d) Using the fact that $(r+s)^2 + s^2 \geq \frac{r^2}{2}$, we have

$$\frac{\delta\eta}{8\epsilon}\alpha_i^2 - \alpha_i[b^i + \mu_i^\top a + (B^i + \mu_i^\top A)x] \leq 2\epsilon - 1, \forall j \in [I].$$

As $x \in \{0,1\}^n$, we have $(B^i + \mu_i^\top A)x \leq \|B^i + \mu_i^\top A^i\|_1$. Thus we arrive at the following inequality

$$\frac{\delta\eta}{8\epsilon}\alpha_i^2 - \alpha_i(b^i + \mu_i^\top a + \|B^i + \mu_i^\top A\|_1) + 1 - 2\epsilon \leq 0, \forall i \in [I].$$

Hence, α can be upper bounded by

$$M_i = \frac{4\epsilon}{\underline{\delta}\eta}\left[(b^i + \mu_i^\top a^i + \|B^i + \mu_i^\top A^i\|_1) + \sqrt{(b^i + \mu_i^\top a^i + \|B^i + \mu_i^\top A^i\|_1)^2 + \underline{\delta}\eta - \frac{\delta\eta}{2\epsilon}}\right]$$

for each $i \in [I]$.

(v) As $\alpha_i \leq M_i$ and $x \in \{0,1\}^n$ for each $i \in [I]$, let us define new variables y such that $y^i = \alpha_i x$, which can be linearized via McCormick inequalities as

$$0 \leq y_j^i \leq M_i x_j, \alpha_i - M_i(1 - x_j) \leq y_j^i \leq \alpha_i.$$

Also since $\gamma_{2j} \leq \frac{\epsilon}{\underline{\delta}\eta}$, thus, let $z_{ij} = \gamma_{2j}x_i, w_{ikj} = \gamma_{2j}x_i x_k$ for all $i, k, j \in [I]$, which also can be linearized via McCormick inequalities as

$$\begin{aligned}
&0 \leq z_{ij} \leq \frac{\epsilon}{\underline{\delta}\eta}x_i, \gamma_{2j} - \frac{\epsilon}{\underline{\delta}\eta}(1 - x_i) \leq z_{ij} \leq \gamma_{2j},\\
&0 \leq w_{ikj} \leq \frac{\epsilon}{\underline{\delta}\eta}x_i, 0 \leq w_{ikj} \leq \frac{\epsilon}{\underline{\delta}\eta}x_k, \gamma_{2j} - \frac{\epsilon}{\underline{\delta}\eta}(2 - x_i - x_j) \leq w_{ikj} \leq \gamma_{2j}.
\end{aligned}$$

Thus, we arrive at (48).

□

The following example illustrates an application of Theorem 9. This example has been studied in [6], where the authors presented several different heuristic (approximate) algorithms. Instead, we show that the feasible region of this problem can be approximated almost exactly as a mixed integer second order conic program (SOCP). Thus, any mixed integer SOCP approach could be used to solve it.

Example 4. (multi-dimensional binary knapsack problem) Consider a variant of Example 2 where $x \in \{0,1\}^n$, i.e. $x_j = 1$ if jth item being picked, 0 otherwise. Suppose that we know the mean of weight vector of each knapsack and its second moment, thus $\mathcal{P}$ is defined as

$$\mathcal{P} = \left\{\mathbb{P} \in \mathcal{P}_0\left(\Xi\right) : \mathbb{E}_{\mathbb{P}}[\tilde{\boldsymbol{\xi}}_i] = 0, \mathbb{E}_{\mathbb{P}}[\tilde{\boldsymbol{\xi}}_i \tilde{\boldsymbol{\xi}}_i^\top] \preceq \Sigma_i, \forall i \in [I]\right\},$$

where $\Xi = \mathbb{R}^{n\times I}$ and $\boldsymbol{\xi}_i = \tilde{\boldsymbol{\xi}}_i + \mu_i$. Without loss of generality, we assume that $\mu_i \geq 0, \Sigma_i \succ 0$ for each $i \in [I]$.

Now the entire distributionally robust multi-dimensional knapsack problem is formulated as

$$\begin{aligned} v^* = \max \quad & c^\top x, \\ \text{s.t.} \quad & x \in \{0,1\}^n, \\ & \inf_{\mathbb{P}\in\mathcal{P}} \mathbb{P}[F(x,\tilde{\boldsymbol{\xi}}) \geq 0] \geq 1-\epsilon, \end{aligned}$$

where $f_i(x,\tilde{\xi}) = b_i - x^\top \mu_i - x^\top \tilde{\xi}_i$ for each $i \in [I]$.

By Theorem 9, we must have $S \cap Z_{\mathrm{D}} = S \cap (X_C \cup \bar{Y}_C)$, where $X_C = \{0\}$ and $\bar{Y}_C$ is equivalent to

$$\bar{Y}_C = \left\{ x : \begin{array}{l} \lambda - \sum_{j\in[I]} \langle \Sigma_j, w_{\cdot\cdot j}\rangle \geq 1-\epsilon, \\ \lambda + \sum_{j\in[I]} t_{0j} \leq 1, \\ \lambda + \sum_{j\in[I]} t_{ij} \leq \alpha_i b_i - \mu_i^\top y^i, \forall i \in [I], \\ \gamma_{1j}^2 \leq 4t_{ij}\gamma_{2j}, \forall i \in [I], j \in [I]\setminus\{i\}, \\ (\gamma_{1i}-\alpha_i)^2 \leq 4t_{ii}\gamma_{2i}, \forall i \in [I], \\ \gamma_{1j}^2 \leq 4t_{0j}\gamma_{2j}, \forall j \in [I], \\ 0 \leq y_j^i \leq M_i x_j, \alpha_i - M_i(1-x_j) \leq y_j^i \leq \alpha_i, \forall i \in [I], j \in [n], \\ 0 \leq w_{ikj} \leq \frac{\epsilon}{\underline{\delta}\eta} x_i, 0 \leq w_{ikj} \leq \frac{\epsilon}{\underline{\delta}\eta} x_k, \\ \gamma_{2j} - \frac{\epsilon}{\underline{\delta}\eta}(2-x_i-x_j) \leq w_{ikj} \leq \gamma_{2j}, \forall i,j,k \in [I] \\ \gamma_{2i} \geq 0, \alpha_i \geq 0, \forall i \in [I]. \end{array} \right\} \tag{53}$$

where

$$M_i = \frac{4\epsilon}{\underline{\delta}\eta}\left[(b^i + \|\mu_i\|_1) + \sqrt{(b^i + \|\mu_i\|_1)^2 + \underline{\delta}\eta - \frac{\underline{\delta}\eta}{2\epsilon}}\right]$$

for each $i \in [I]$ with $\eta = \min_{x\in\{0,1\}^n:x\neq 0} \|x\|_2^2 = 1$ and $\underline{\delta}$ the smallest eigenvalue of matrices $\{\Sigma_j\}_{j\in[I]}$.

5.2 Numerical illustration

In this section, we present a numerical study to illustrate the strength of proposed formulation (53) corresponding to the multidimensional knapsack problem in Example 4. The instances are constructed from the problem set *mk-20-10* in [21]. The instances in this set are named *1-4-multi-N-i*, where $N \in \{100, 500, 1000, 3000\}$ denotes sample size of the weight vector and there are 5 different instances for each sample size (i.e., $i \in \{1, 2, 3, 4, 5\}$). Each instance has 20 decision variables and 10 knapsack constraints, i.e, $n = 20, I = 10$. We compute μ, σ of the weight vector for each knapsack as the sample mean and covariance from the provided data.

Our first approach is to solve the mixed integer SOCP (53) exactly. We notice that the explicit upper bounds of α, γ_2 could be quite loose. Hence, instead, we enhance these bounds by maximizing these variables over the continuous relaxation of (53).

We compare our approach with the heuristic one proposed in [6]. The authors formulate to maximize $c^\top x$ over set $Z_C \cap \{0,1\}^n$ in (13) as the following mixed integer nonconvex program

$$\begin{aligned}
\max_{x\in\{0,1\}^n} \quad & c^\top x, \\
\text{s.t. } & \lambda - \sum_{j\in[I]} \langle \Sigma_j, \gamma_{2j}\rangle \geq (1-\epsilon)\beta, \\
& \lambda + \sum_{j\in[I]} t_{0j} \leq \beta, \\
& \lambda + \sum_{j\in[I]} t_{ij} \leq \alpha_i (b_i - \mu_i^\top x), \forall i \in [I], \\
& \begin{bmatrix} t_{ij} & -\frac{1}{2}\gamma_{1j}^\top \\ -\frac{1}{2}\gamma_{1j} & \gamma_{2j} \end{bmatrix} \succeq 0, \forall i \in [I], j \in [I]\setminus\{i\}, \\
& \begin{bmatrix} t_{ii} & -\frac{1}{2}(\gamma_{1i} + \alpha_i x)^\top \\ -\frac{1}{2}(\gamma_{1i} + \alpha_i x) & \gamma_{2i} \end{bmatrix} \succeq 0, \forall i \in [I], \\
& \begin{bmatrix} t_{0j} & -\frac{1}{2}\gamma_{1j}^\top \\ -\frac{1}{2}\gamma_{1j} & \gamma_{2j} \end{bmatrix} \succeq 0, \forall j \in [I], \\
& \gamma_{2i} \succeq 0, \alpha_i > 0, \forall i \in [I].
\end{aligned} \tag{54}$$

The solution approach is summarized below. First of all, given an α, solve the continuous relaxation of (54) with α fixed, which is an SDP. Let $\widehat{x}$ be the corresponding optimal solution. Then, fix $x = \widehat{x}$, change the objective function to $\max \lambda - \sum_{j\in[I]} \langle \Sigma_j, \gamma_{2j}\rangle$ (i.e., maximize the largest probability) and solve the corresponding continuous relaxation problem with optimal solution $\widehat{\alpha}$. In the next step, let $\alpha = \widehat{\alpha}$, and iterate. This procedure terminates whenever the values of α and x no longer change. Suppose, at the end of this procedure, $\alpha = \widehat{\alpha}, x = \widehat{x}$. In general, $\widehat{x} \in [0,1]^n$ is not binary. So the final step is randomized rounding, i.e. treat $x_i \in \{0,1\}$ as a Bernoulli random variable with probability $\widehat{x}_i$ for each $i \in [n]$, generate a sample $\tilde{x}$, then check the feasibility of $\tilde{x}$ to (54). This step could repeat multiple times until finding several candidate solutions (e.g., 5 solutions) and of course, choosing the best one as the output.

We use commercial solver CPLEX for the first approach, while CVX for the second approach.

The results are listed in Table 1. We use $v_{\rm SOCP}, t_{\rm SOCP}$ to denote the objective value and the total running time of first approach (including big-M strengthening time), while use $v_{\rm H}, t_{\rm H}, {\rm gap}$ for the objective value, the total running time and the optimality gap of the second approach (the heuristic one). All instances were executed on a laptop with a 2.67 GHz processor and 4GB RAM, while CPLEX 12.5.1 and CVX 2.1 were used with their default settings.

Table 1: Performance comparison of exact and the heuristic methods

ϵ	N	Index	Exact Approach (53)		Heuristic Approach by [6]		
			$v_{\rm SOCP}$	$t_{\rm SOCP}$	$v_{\rm H}$	$t_{\rm H}$	gap(%)
0.05	100	1	2925	2291	2860	3069	2.2%
		2	3950	8412	3750	2852	5.1%
		3	2910	3740	2580	3111	11.3%
		4	3430	2061	1900	3298	44.6%
		5	3860	3756	3510	12294	9.1%
0.1	100	1	3140	8945	2860	4972	8.9%
		2	4460	5546	4330	3850	2.9%
		3	4140	3914	2610	4537	37.0%
		4	4100	5597	4010	3462	2.2%
		5	4380	7429	3935	4698	10.2%
0.05	500	1	3480	3286	1710	4846	50.9%
		2	3400	3471	2000	4564	41.2%
		3	3800	2080	3680	3917	3.2%
		4	3050	4300	1690	3525	44.6%
		5	4010	3897	3980	3198	0.7%
0.1	500	1	4100	5304	2580	3207	37.1%
		2	4270	2188	2030	3194	52.5%
		3	4480	4269	4380	3506	2.2%
		4	4050	5197	3760	3376	7.2%
		5	4530	5870	4240	3350	6.4%
0.05	1000	1	4150	1430	2490	3431	40.0%
		2	3280	789	2580	5451	21.3%
		3	4060	1369	3920	3428	3.4%
		4	2640	1125	2170	3489	17.8%
		5	2560	1300	2090	4096	18.4%
0.1	1000	1	4560	2527	4550	4100	0.2%
		2	4280	2655	2610	4543	39.0%
		3	4560	4304	4360	3441	4.4%
		4	3500	1517	2490	3877	28.9%
		5	3150	3424	2490	3720	21.0%
0.05	3000	1	2905	2053	2530	5455	12.9%
		2	2770	1701	2330	4482	15.9%
		3	2840	2546	2610	3844	8.1%
		4	2860	2361	2240	3830	21.7%
		5	2770	3087	2290	7839	17.3%
0.1	3000	1	3680	2026	2170	3904	41.0%
		2	3460	1765	2170	3837	37.3%
		3	3610	4282	2580	7537	28.5%
		4	3560	4609	2490	4069	30.1%
		5	3470	2097	2170	3512	37.5%

In Table 1, we observe that the solution time of both methods are in general quite similar, while on average, the exact approach (3463s) takes less time than the heuristic one (4268s). If we compare the solution quality, it can be seen that the solution of the heuristic method is quite unpredictable, i.e., for some instances, it finds a very good solution but for others, it does not. The average gap of the heuristic solutions is around 20%. On the other hand, the exact approach can find the optimal solution within an hour and a half for majority of instances (34 over 40 instances). These results demonstrate the effectiveness of the exact approach proposed in this paper.

6 Conclusion

In this paper, we studied a distributionally robust chance constrained problem (DRCCP) with joint nonlinear uncertain constraints under convex moment ambiguity sets. We identified a number of general sufficient conditions where such a DRCCP can be reformulated as a (mixed integer) convex program. In this work, we assume the uncertain mapping $F(x, \xi)$ is concave in x and convex in ξ. A future direction is to generalize these results to a broader family of uncertain mappings, for example, similar to [24], when $F(x, \xi)$ is concave in x and quasi-convex in ξ. In addition, a convex moment ambiguity set in general will not converge to the true distribution if more empirical data is available to estimate moments. Therefore, extension to other type of ambiguity sets (e.g., KL divergence [12] or Wasserstein metric [9] based ambiguity set) would be very valuable. Finally, for binary DRCCP, the big-M formulation in Proposition 7 could be very weak in general, thus it is of interest to develop sophisticated solution approaches for solving the corresponding mixed integer convex programs efficiently.

Acknowledgment

This research has been supported in part by the Georgia Tech Algorithms and Randomness Center (ARC) and the National Science Foundation award 1633196. The authors thank the referees and editors for constructive comments that helped improve this paper.

References

[1] Ben-Tal, A. and Nemirovski, A. (2001). *Lectures on modern convex optimization: analysis, algorithms, and engineering applications*, volume 2. Siam.

[2] Bennett, K. P. and Mangasarian, O. L. (1993). Bilinear separation of two sets inn-space. *Computational Optimization and Applications*, 2(3):207–227.

[3] Bonnans, J. F. and Shapiro, A. (2000). *Perturbation analysis of optimization problems*. Springer Science & Business Media.

[4] Calafiore, G. C. and El Ghaoui, L. (2006). On distributionally robust chance-constrained linear programs. *Journal of Optimization Theory and Applications*, 130(1):1–22.

[5] Chen, W., Sim, M., Sun, J., and Teo, C.-P. (2010). From cvar to uncertainty set: Implications in joint chance-constrained optimization. *Operations research*, 58(2):470–485.

[6] Cheng, J., Delage, E., and Lisser, A. (2014). Distributionally robust stochastic knapsack problem. *SIAM Journal on Optimization*, 24(3):1485–1506.

[7] Delage, E. and Ye, Y. (2010). Distributionally robust optimization under moment uncertainty with application to data-driven problems. *Operations research*, 58(3):595–612.

[8] Durrett, R. (1996). *Probability: theory and examples*. Cambridge university press.

[9] Esfahani, P. M. and Kuhn, D. (2015). Data-driven distributionally robust optimization using the wasserstein metric: performance guarantees and tractable reformulations. Available at http://www.optimization-online.org/DB_FILE/2015/05/4899.pdf.

[10] Hanasusanto, G. A., Roitch, V., Kuhn, D., and Wiesemann, W. (2015). A distributionally robust perspective on uncertainty quantification and chance constrained programming. *Mathematical Programming*, 151(1):35–62.

[11] Hanasusanto, G. A., Roitch, V., Kuhn, D., and Wiesemann, W. (2017). Ambiguous joint chance constraints under mean and dispersion information. *Operations Research*.

[12] Jiang, R. and Guan, Y. (2015). Data-driven chance constrained stochastic program. *Mathematical Programming*, pages 1–37.

[13] Kall, P. and Wallace, S. W. (1994). *Stochastic programming*. John Wiley & Sons, Chichester.

[14] Kim, K.-L. H. S.-J. and Boyd, S. (2008). Power allocation with outage probability specifications in wireless shadowed fading channels via geometric programming. *RR, Information Systems Laboratory, Stanford University*.

[15] Lejeune, M. A. and Margot, F. (2016). Solving chance-constrained optimization problems with stochastic quadratic inequalities. *Operations Research*, 64(4):939–957.

[16] McCormick, G. P. (1976). Computability of global solutions to factorable nonconvex programs: Part I- convex underestimating problems. *Mathematical programming*, 10(1):147–175.

[17] Nemirovski, A. and Shapiro, A. (2006). Convex approximations of chance constrained programs. *SIAM Journal on Optimization*, 17(4):969–996.

[18] Popescu, I. (2007). Robust mean-covariance solutions for stochastic optimization. *Operations Research*, 55(1):98–112.

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

[20] Shapiro, A. and Kleywegt, A. (2002). Minimax analysis of stochastic problems. *Optimization Methods and Software*, 17(3):523–542.

[21] Song, Y., Luedtke, J. R., and Küçükyavuz, S. (2014). Chance-constrained binary packing problems. *INFORMS Journal on Computing*.

[22] Wiesemann, W., Kuhn, D., and Sim, M. (2014). Distributionally robust convex optimization. *Operations Research*, 62(6):1358–1376.

[23] Xie, W. and Ahmed, S. (2016). Distributionally robust chance constrained optimal power flow with renewables: A conic reformulation. Available at http://www.optimization-online.org/DB_FILE/2016/12/5759.pdf.

[24] Yang, W. and Xu, H. (2016). Distributionally robust chance constraints for non-linear uncertainties. *Mathematical Programming*, 155(1-2):231–265.

[25] Zhang, H. and Li, P. (2010). Probabilistic analysis for optimal power flow under uncertainty. *IET Generation, Transmission & Distribution*, 4(5):553–561.

[26] Zymler, S., Kuhn, D., and Rustem, B. (2013). Distributionally robust joint chance constraints with second-order moment information. *Mathematical Programming*, 137(1-2):167–198.