On Solving Two-Stage Distributionally Robust Disjunctive Programs with a General Ambiguity Set*

Manish Bansal[1,*]

Department of Industrial and Systems Engineering, Virginia Tech, Blacksburg, VA, USA

Sanjay Mehrotra

Department of Industrial Engineering and Management Sciences, Northwestern University, Evanston, IL, USA

Abstract

We introduce two-stage distributionally robust disjunctive programs (TSDR-DPs) with disjunctive constraints in both stages and a general ambiguity set for the probability distributions. The TSDR-DPs subsume various classes of two-stage distributionally robust programs where the second stage problems are non-convex programs (such as mixed binary programs, semi-continuous program, nonconvex quadratic programs, separable non-linear programs, etc.). TSDR-DP is an optimization model in which the degree of risk aversion can be chosen by decision makers. It generalizes two-stage stochastic disjunctive program (risk-neutral) and two-stage robust disjunctive program (most-conservative). To our knowledge, the foregoing special cases of TSDR-DPs have not been studied until now. In this paper, we develop decomposition algorithms, which utilize Balas' linear programming equivalent for deterministic disjunctive programs or his sequential convexification approach within L-shaped method, to solve TSDR-DPs. We present sufficient conditions under which our algorithms are finitely convergent. These algorithms generalize the distributionally robust integer L-shaped algorithm of Bansal et al. (SIAM J. on Optimization 28: 2360-2388, 2018) for TSDR mixed binary programs, a subclass of TSDR-DPs. Furthermore, we formulate a semi-continuous program (SCP) as a disjunctive program and use our results for TSDR-DPs to solve general two-stage distributionally robust SCPs (TSDR-SCPs) and TSDR-SCP having semi-continuous inflow set in the second stage.

Keywords: stochastic programming, distributionally robust disjunctive programs, decomposition algorithms, semi-continuous program, reverse convex program

---

*Corresponding Author

Email addresses: `bansal@vt.edu` (Manish Bansal), `mehrotra@northwestern.edu` (Sanjay Mehrotra)

[1]Funded by National Science Foundation Grant CMMI 1824897, which is gratefully acknowledged

1. Introduction

Disjunctive programming (DP) is a well-known area in optimization where a linear programming problem has disjunctive constraints, i.e. linear constraints with "or" ($\vee$, disjunctive) operations. More specifically, DP optimizes over a union of polyhedra $\mathcal{R}_i = \{z \in \mathbb{R}^n_+ : E^i z \geq f^i\}$, denoted by $\mathcal{R} := \cup_i \mathcal{R}_i = \{z \in \mathbb{R}^n_+ : \vee_i (E^i z \geq f^i)\}$. In this paper, we introduce two-stage distributionally robust disjunctive programs (TSDR-DPs) where both the first and second stages have disjunctive constraints, and the random parameters in the second stage follow the worst-case distribution belonging to an ambiguity set of probability distributions. We write a TSDR-DP as follows:

$$\min \left\{ \bar{c}^T \bar{x} + \max_{P \in \mathfrak{P}} \mathbb{E}_P[\mathcal{Q}_\omega(\bar{x})] : \bigvee_{s \in S} (A_s \bar{x} \geq b_s), \bar{x} \in \{0,1\}^{\bar{p}} \right\},$$

where $|S|$ is finite, a random vector associated to uncertain data parameters is defined by probability distribution $P \in \mathfrak{P}$ (a set of distributions) with support Ω, and for any scenario $\omega \in \Omega$ and a finite set H,

$$\mathcal{Q}_\omega(\bar{x}) := \min \left\{ g_\omega^T y_\omega : W_\omega y_\omega \geq r_\omega - \overline{T}_\omega \bar{x}, \bigvee_{h \in H} \left(D^h_{\omega,1} y_\omega \geq d^h_{\omega,0} - \overline{D}^h_{\omega,2} \bar{x} \right), y_\omega \in \mathbb{R}^q_+ \right\}.$$

Here, the parameters $\bar{c} \in \mathbb{R}^{\bar{p}}$, $A_s \in \mathbb{R}^{\overline{m}_1 \times \bar{p}}$, $b_s \in \mathbb{R}^{\overline{m}_1}$ for $s \in S$, and for each $\omega \in \Omega$, $g_\omega \in \mathbb{R}^q$, $W_\omega \in \mathbb{R}^{m_2 \times q}$, $\overline{T}_\omega \in \mathbb{R}^{m_2 \times \bar{p}}$, and $r_\omega \in \mathbb{R}^{m_2}$. Likewise, $D^h_{\omega,1}$, $\overline{D}^h_{\omega,2}$, and $d^h_{\omega,0}$ are real matrices/vectors of appropriate dimensions. Note that the parameters of the first stage disjunctive constraints are deterministic, and therefore, we reformulate the first stage feasible region, i.e. $\{\bar{x} \in \{0,1\}^{\bar{p}} : \bar{x} \in \bigvee_{s \in S} (A_s \bar{x} \geq b_s)\}$ using binary variables with linear constraints (Nemhauser and Wolsey, 1988):

$$A_s \bar{x} \geq b_s - M(1 - \chi_s)\mathbf{1}, \;\; s \in S, \tag{1a}$$

$$\sum_{s \in S} \chi_s = 1, \tag{1b}$$

$$\bar{x} \in \{0,1\}^{\bar{p}}, \chi_s \in \{0,1\}, \;\; s \in S, \tag{1c}$$

where M is a constant and $\mathbf{1}$ is a vector of all ones. The constant M is selected such that $A_s \bar{x} \geq b_s - M\mathbf{1}$ for all $\bar{x} \in \{0,1\}^{\bar{p}}$ and $s \in S$. This formulation, defined by (1a)-(1c), has only binary variables which in a compact form can be written as $\{x \in \{0,1\}^p : Ax \geq b\}$ where $x = (\bar{x}, \{\chi_s\}_{s \in S})$, $A \in \mathbb{R}^{m_1 \times p}$, and $b \in \mathbb{R}^{m_1}$. Note that this reformulation has $m_1 = \overline{m}_1 \times |S| + 1$ linear constraints and $p = \bar{p} + |S|$ binary variables, in comparison to $\bar{p}$ binary variables and $|S|$ disjunctive constraints, each defined by $\overline{m}_1$ linear constraints, in the original formulation.

In light of the above reformulation, in the rest of the paper, we utilize the following definition of the TSDR-DP (without loss of generality):

$$\min \left\{ c^T x + \max_{P \in \mathfrak{P}} \mathbb{E}_P[\mathcal{Q}_\omega(x)] : Ax \geq b, x \in \{0,1\}^p \right\}, \tag{2}$$

where

$$\mathcal{Q}_\omega(x) = \min \; g_\omega^T y_\omega \tag{3a}$$
$$s.t. \; W_\omega y_\omega \geq r_\omega - T_\omega x \tag{3b}$$
$$\bigvee_{h \in H} \left(D_{\omega,1}^h y_\omega \geq d_{\omega,0}^h - D_{\omega,2}^h x \right) \tag{3c}$$
$$y_\omega \in \mathbb{R}_+^q, \tag{3d}$$

for $\omega \in \Omega$, $c \in \mathbb{R}^p$, and $T_\omega \in \mathbb{R}^{m_2 \times p}$, $\omega \in \Omega$. The formulation defined by (3), constraint (3c), the function $\mathcal{Q}_\omega(x)$, and the set of distributions $\mathfrak{P}$ are referred to as the second-stage subproblem, the disjunctive constraint in the *disjunctive normal form*, the recourse function, and the ambiguity set, respectively. In this paper, we also consider TSDR-DPs where the disjunctive constraint in the second stage are defined in *conjunctive normal form*, i.e., for $\omega \in \Omega$

$$\mathcal{Q}_\omega(x) = \min \; g_\omega^T y_\omega \tag{4a}$$
$$s.t. \; W_\omega y_\omega \geq r_\omega - T_\omega x \tag{4b}$$
$$\bigwedge_{j=1}^{\bar{m}_2} \left(\bigvee_{i \in H_j} \eta_{\omega,1}^i y_\omega \geq \eta_{\omega,0}^i - \eta_{\omega,2}^i x \right) \tag{4c}$$
$$y_\omega \in \mathbb{R}_+^q, \tag{4d}$$

where $\wedge$ denotes "and" or conjunction operation, $\bar{m}_2$ and $|H_j|$ for all j are finite, and $\eta_{\omega,1}^i \in \mathbb{R}^q$, $\eta_{\omega,0}^i \in \mathbb{R}$, and $\eta_{\omega,2}^i \in \mathbb{R}^p$. Observe that, since the logical operations $\vee$ (disjunction) and $\wedge$ (conjunction) obey the distributive law, i.e., $(a_1 \wedge a_2) \vee (b_1 \wedge b_2) = (a_1 \vee b_1) \wedge (a_1 \vee b_2) \wedge (a_2 \vee b_1) \wedge (a_2 \vee b_2)$, the disjunctive constraint (3c) can also be written in the conjunctive normal form, i.e., (4c), where each disjunction j contains exactly one inequality from each system of inequalities in the corresponding disjunctive constraint, and $|H_j| = |H|$. Conversely, since $(a_1 \vee b_1) \wedge (a_2 \vee b_2) = (a_1 \wedge a_2) \vee (a_1 \wedge b_2) \vee (b_1 \vee a_2) \vee (b_1 \vee b_2)$, the disjunctive constraint (4c) can also be written in the disjunctive normal form, i.e., (3c). For example, $z \in \{0,1\}^2 = \{(z_1, z_2) : \wedge_{i=1}^2 ((z_i = 0) \vee (z_i = 1))\} = \{(z_1, z_2) : (z_1 = 0, z_2 = 0) \vee (z_1 = 0, z_2 = 1) \vee (z_1 = 1, z_2 = 0) \vee (z_1 = 1, z_2 = 1)\}$.

To study TSDR-DPs, we assume that

1. $X := \{x \in \{0,1\}^p : Ax \geq b\}$ is non-empty.
2. $\mathcal{K}_\omega(x) := \{y_\omega : \text{(3b)-(3d) hold}\}$ is non-empty and $\mathcal{Q}_\omega(x) > -\infty$ for all $x \in X$ and $\omega \in \Omega$ (relatively complete recourse).

3. Each probability distribution $P \in \mathfrak{P}$ has finite support Ω, i.e. $|\Omega|$ is finite.

4. There exists an algorithm that provides a probability distribution $P \in \mathfrak{P}$, i.e., $\{p_\omega\}_{\omega \in \Omega}$ where p_ω is the probability of occurrence of scenario $\omega \in \Omega$, by solving the optimization problem:

$$\mathcal{Q}(x) := \max_{P \in \mathfrak{P}} \mathbb{E}_P[\mathcal{Q}_\omega(x)] \tag{5}$$

   for a given $x \in X$.

We refer to the optimization problem (5) as the distribution separation problem corresponding to an ambiguity set, and the algorithm to solve this problem is referred to as the *distribution separation algorithm*.

In the literature, the ambiguity set has been defined in various different ways such as using linear constraints on the first two moments of the distribution (Bertsimas and Popescu, 2005; Dupacová, 1987; Prékopa, 1995; Scarf, 1958), conic constraints to describe the set of distributions with moments (Bertsimas et al., 2010; Delage and Ye, 2010), measure bounds and general moment constraints (Shapiro and Ahmed, 2004), Kantorovich distance or Wasserstein metric (Pflug et al., 2012; Pflug and Wozabal, 2007; Wozabal, 2012), ζ-structure metrics (Zhao and Guan, 2015), ϕ-divergences such as χ^2 distance and Kullback-Leibler divergence (Ben-Tal et al., 2012; Calafiore, 2007; Jiang and Guan, 2015; Love and Bayraksan, 2016; Wang et al., 2016; Yankolu and den Hertog, 2012), and Prokhorov metrics (Erdogan and Iyengar, 2006). Readers can refer to Hanasusanto et al. (2015); Shapiro (2018) and references therein for literature review on the origin of the distributionally robust optimization framework (Scarf, 1958). In this paper, we consider a general class of ambiguity sets in our decomposition algorithms for solving TSDR-DPs. We show that the developed decomposition algorithms identify an optimal solution of TSDR-DPs. Additionally, we demonstrate that these algorithms are finitely convergent if the distribution seperation problem (5) can be solved in finite iterations. Few examples of such ambiguity sets are: moment matching set (defined using bounds on moments), Kantorovich set (defined using Wasserstein metric), and total variational set (Rahimian et al., 2018; Sun and Xu, 2015), defined for a finite sample space Ω. The distribution separation problem associated with these ambiguity sets are linear programs, which can be solved in finite iterations. It is important to note that in several cases when the ambiguity set has special structures, the distribution separation problem (5) has solutions with closed formulation; see Shapiro (2017, 2018) and references therein. Therefore, selecting such a structured ambiguity set will further benefit our algorithms.

It is important to note that the class of TSDR-DPs subsumes classes of two-stage distributionally robust programs where the second stage is a non-convex program such as mixed binary programs, semi-continuous program, nonconvex quadratic programs, separable non-linear programs, reverse convex programs (DPs with infinitely many terms), etc. Refer to Balas (1974, 1998) for details. To our knowledge, the TSDR-DP, in its general form, has not been studied before. In this paper, we develop decomposition algorithms, which utilize linear programming equivalent for deterministic disjunctive programs or sequential convexification approach of Balas (1979, 1998) and distribution separation algorithms within L-shaped method, to solve TSDR-DPs. We present sufficient conditions under which our algorithms are finitely convergent. Furthermore, we showcase the significance of studying TSDR-DPs by introducing two-stage distributionally robust semi-continuous programs (TSDR-SCPs). We write a DP equivalent of a semi-continuous program (SCP) which is a linear program with semi-continuity restrictions on some continuous variables, i.e. a variable belongs to a set of the form $[0, \underline{l}] \cup [\bar{l}, \bar{u}]$ where $0 \leq \underline{l} \leq \bar{l} \leq \bar{u}$. We use our results for TSDR-DPs to solve general TSDR-SCPs and TSDR-SCP having semi-continuous inflow set (Angulo et al., 2013) in the second stage. Note that by setting $\underline{l} = \bar{l}$, the semi-continuous variable becomes continuous and by setting $\underline{l} = 0$ and $\bar{l} = \bar{u} = 1$, the semi-continuous variable becomes binary. Therefore, two-stage distributionally robust mixed binary programs (TSDR-MBPs) with binary variables in the first stage and mixed binary programs in the second stage are special cases of TSDR-SCPs or TSDR-DPs.

1.1 Motivation for studying TSDR-DPs

The main motivation to study TSDR-DPs are as follows: (1) their application in formulating various applied optimization problems using disjunctive constraints with uncertain data-parameters, and (2) they generalize various well-known optimization modeling frameworks. More specifically, deterministic DPs have been utilized to solve problems arising in wide range of applications. This includes power systems (Wood et al., 2013), transportation (Gunluk et al., 2007), synthesis of chemical process systems (Trkay and Grossmann, 1996; Grossmann et al., 1999), network design problems (Bertsekas and Gallager, 1992), and many more. Therefore, the results presented for TSDR-DPs can also solve two-stage stochastic (risk-neutral), robust, or distributionally robust variants of these applied optimization problems. Furthermore, in addition to two-stage distributionally robust programs where the second stage is a non-convex program such as mixed binary programs, semi-continuous program, nonconvex quadratic programs, separable non-linear programs, reverse convex programs, etc., various special cases of TSDR-DPs include:

1) TSDR-MBP: Recently, Bansal et al. (2018a) introduced TSDR-MBP, a special case of TSDR-DP, and presented a decomposition algorithm which utilizes distribution separation procedure and parametric cuts within Benders' algorithm (Benders, 1962) to solve TSDR-MBPs. The authors referred to this algorithm as distributionally robust integer (DRI) L-shaped algorithm because it generalizes the integer L-shaped algorithm (Laporte and Louveaux, 1993) developed for a special case of TSDR-MBP where $\mathfrak{P}$ is singleton, i.e. two-stage stochastic mixed binary program. The algorithms and their finite convergence results presented in this paper for TSDR-DPs generalize the results of Bansal et al. (2018a) for TSDR-MBPs.

2) TSDR Linear Programs: An extensively studied special case of TSDR-DP is the class of TSDR linear programs (TSDR-LPs), i.e. TSDR-DP where S and H are empty sets and the first stage has no binary restrictions (Bertsimas et al., 2010; Hanasusanto and Kuhn, 2016; Love and Bayraksan, 2016). More specifically, Bertsimas et al. (2010) considered TSDR-LP where the ambiguity set is

defined using multivariate distributions with known first and second moments and risk is incorporated in the model using a convex nondecreasing piecewise linear function on the second stage costs. They showed that the corresponding problem has semidefinite programming reformulations. Jiang and Guan (2015) presented a sample average approximation algorithm to solve a special case of TSDR-MBP with binary variables only in the first stage where the ambiguity set is defined using the l_1-norm on the space of all (continuous and discrete) probability distributions. Recently, Love and Bayraksan (2016) developed a decomposition algorithm for solving TSDR-LP where the ambiguity set is defined using ϕ-divergence. Hanasusanto and Kuhn (2016) provided a conic programming reformulations for TSDR-LP where the ambiguity set comprises of a Wasserstein ball centered at a discrete distribution. Bansal et al. (2018a) also developed a finitely convergent decomposition algorithm for TSDR-LP where the ambiguity set $\mathfrak{P}$ is defined by a polytope.

3) Two-Stage Stochastic Mixed Binary Programs: Another well-studied special case of TSDR-DP is the class of two-stage stochastic mixed binary programs (TSS-MBPs) (Carøe and Tind, 1997), i.e. TSDR-DP with mixed binary programs in the second stage, only binary variables in the first stage, and $|\mathfrak{P}| = 1$. Various studies (Bansal et al., 2018b; Gade et al., 2014; Laporte and Louveaux, 1993; Ntaimo, 2009; Sen and Higle, 2005; Sherali and Fraticelli, 2002) have utilized globally valid parametric cuts in (x, y_ω) space to solve the subproblems. Readers can refer to Kücükyavuz and Sen (2017) for a comprehensive survey on algorithms for TSS-MBPs.

TSDR-DP is an optimization model in which the degree of risk aversion can be chosen by decision makers. It generalizes: (a) Two-stage stochastic disjunctive program (TSS-DP), which is TSDR-DP with a singleton $\mathfrak{P} = \{P_0\}$; and (b) Two-stage robust disjunctive program (TSR-DP), i.e. TSDR-DP with a set $\mathfrak{P}$ that consists of all probability distributions supported on Ω. In the literature, the TSS-DP and TSR-DP, in their general forms, have not been studied.

Remark 1. *Note that the TSDR-DP is at least as hard as the TSS-MBP (a special case of TSDR-MBP and TSDR-DP) which is an #P-hard problem (Dyer and Stougie, 2006).*

Remark 2. *In another direction, Bansal and Zhang (2018) introduced two-stage distributionally robust p-order conic integer programs (TSDR-CMIPs) in which the first stage has only integer variables and the second-stage problems have p-order conic constraints along with integer variables. They introduced structured CMIPs in the second stage of TSDR-CMIPs and provided convex programming equivalent for them using parametric inequalities.*

1.2 An illustrative example of TSDR-DP

In this section, we provide an example of TSDR-DP, defined by (2)-(3), which we will use to illustrate reformulations and algorithm presented in this paper for TSDR-DP in the disjunctive normal form.

Example 1. *Consider the following TSDR-DP defined over sample space* $\Omega := \{\omega_1, \omega_2, \omega_3\}$*:*

$$\text{Minimize} \quad \overline{x}_1 - \overline{x}_2 - \overline{x}_3 + \max_{\{p_\omega\}_{\omega\in\Omega}\in\mathfrak{P}_E} \sum_{\omega\in\Omega} p_\omega \mathcal{Q}_\omega(\overline{x}) \tag{6a}$$

$$\text{subject to} \quad (\overline{x}_1 + \overline{x}_2 \geq 2) \vee (-\overline{x}_2 - \overline{x}_3 \geq 0) \tag{6b}$$

$$\overline{x}_1, \overline{x}_2, \overline{x}_3 \in \{0, 1\}, \tag{6c}$$

where $\mathfrak{P}_E = \left\{(p_{\omega_1}, p_{\omega_2}, p_{\omega_3}) \in \mathbb{R}^3_+ : \sum_{\omega\in\Omega} p_\omega = 1, p_{\omega_1} - p_{\omega_2} + p_{\omega_3} \geq 0.1\right\},$

$$\mathcal{Q}_\omega(\overline{x}) = \quad \text{Minimize} \quad 2y^1_\omega - y^2_\omega \tag{7a}$$

$$\text{subject to} \quad -y^1_\omega - y^2_\omega \geq d_\omega + \overline{x}_1 + \overline{x}_2, \tag{7b}$$

$$\left(y^1_\omega \geq 1 + \overline{x}_2\right) \vee \left(y^2_\omega \geq 2 - \overline{x}_1 - \overline{x}_3\right), \tag{7c}$$

$$y^1_\omega, y^2_\omega \in \mathbb{R}_+, \tag{7d}$$

for $\omega \in \Omega$, $d_{\omega_1} = -5$, $d_{\omega_2} = -6$, *and* $d_{\omega_3} = -4$. *To demonstrate the nonconvexity of the feasible region of the second stage problem, i.e.,* $\{y_\omega \in \mathbb{R}^2_+ : (7\text{b}) - (7\text{d})\}$, *we sketch it for three pairs of* $(\overline{x}, \omega)$*; see Figure 1.*

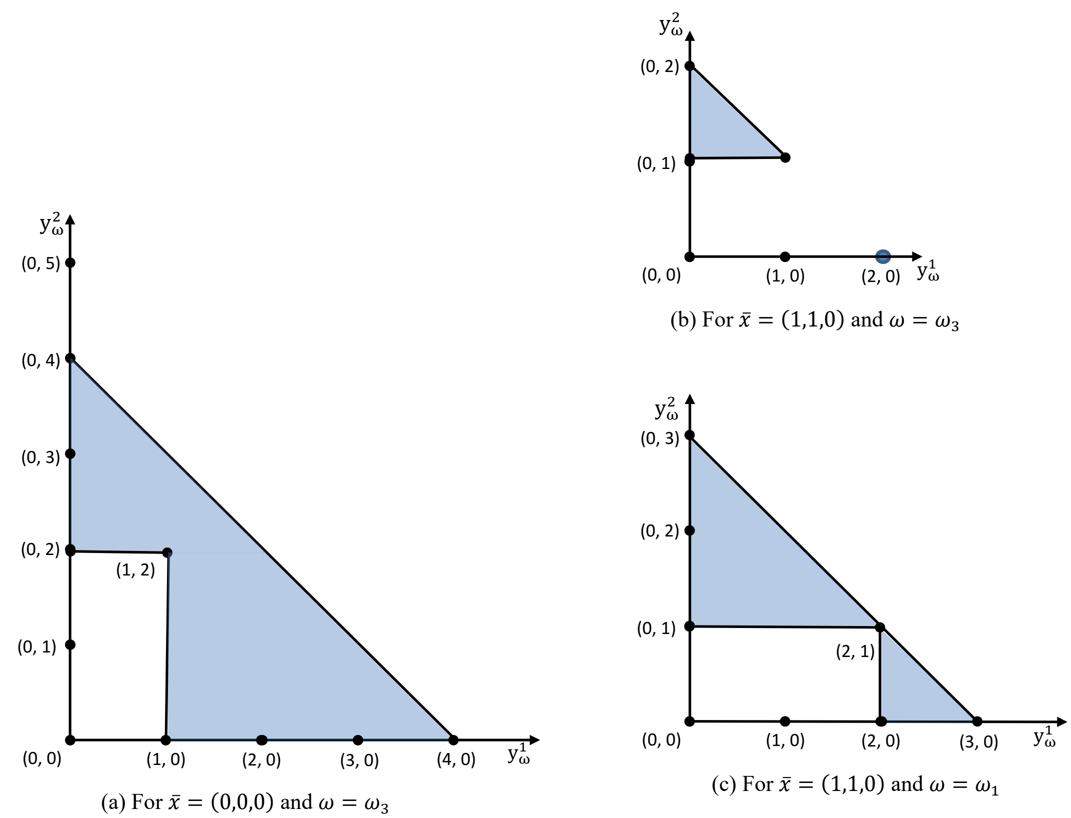

Figure 1: Feasible region of the second stage problem for given $(\overline{x}, \omega)$, denoted by shaded region. Note that in Fig. 1(b), $y_\omega = (2, 0)$ is also a feasible solution.

Similar to (1)*, the first stage feasible region is reformulated using two additional binary variables and big-M coefficient, i.e.. $M = 2$, as follows:*

$$\textit{Minimize} \quad \overline{x}_1 - \overline{x}_2 - \overline{x}_3 + \max_{\{p_\omega\}_{\omega\in\Omega}\in\mathfrak{P}_E} \sum_{\omega\in\Omega} p_\omega \mathcal{Q}_\omega(x) \tag{8a}$$

$$\textit{subject to} \quad \overline{x}_1 + \overline{x}_2 \geq 2 - M\,(1-\chi_{s_1}) = 2\chi_{s_1}, \tag{8b}$$

$$-\overline{x}_2 - \overline{x}_3 \geq -M\,(1-\chi_{s_2}) = -2 + 2\chi_{s_2}, \tag{8c}$$

$$x := (\overline{x}_1, \overline{x}_2, \overline{x}_3, \chi_{s_1}, \chi_{s_2}) \in \{0,1\}^5. \tag{8d}$$

Since the coefficients associated with variables χ_{s_1} and χ_{s_2} in the second stage subproblem (7) *are zero, $\mathcal{Q}_\omega(x)$ is same as $\mathcal{Q}_\omega(\overline{x})$. Additionally, this example satisfies all assumptions made for TSDR-DPs, i.e., (a) the set of first stage feasible solutions is non-empty; (b) relatively complete recourse; (c) $|\Omega|$ is finite; and (d) there exists an algorithm to solve the distribution separation problem as it is a linear problem.* □

Remark 3. *For numerical examples of TSDR-MBP (a special case of TSDR-DP in the conjunctive normal form), please refer to Bansal et al. (2018a).*

1.3 Organization of this paper

In Section 2, we briefly review the results developed in Balas (1979, 1998) and Balas et al. (1993) for the disjunctive programming problems. In Section 3, we provide a decomposition algorithm to solve general TSDR-DPs using a linear programming equivalent of Balas (1979, 1998) for DPs. We present an alternative decomposition algorithm, which utilizes the sequential convexification approach within L-shaped method, to solve TSDR-DPs with facial DPs and sequentially convexifiable programs in the second stage. We prove that the foregoing algorithms are finitely convergent. In Section 4, we study TSDR-SCPs and present linear programming equivalent for the second stage of TSDR-SCPs and a relaxation of a two-stage distributionally robust semi-continuous network flow problem. Finally, we provide concluding remarks in Section 5.

2. Necessary Background on Disjunctive Programming

In this section, we briefly review the results developed in Balas (1979, 1998) and Balas et al. (1993) for disjunctive programming problems to provide the necessary background for the results in the following sections (readers can also refer to Chapter 10 of Jnger et al. (2009) for details). As mentioned before, a disjunctive program is a linear program with disjunctive constraints, i.e. linear inequalities connected by $\vee$ ("or", disjunction) logical operations. Given non-empty polyhedra $\mathcal{R}_i := \{z \in \mathbb{R}^n_+ : E^i z \geq f^i\}$, $i \in L$, the *disjunctive normal form* representation for the set

$\mathcal{R} = \cup_{i\in L}\mathcal{R}_i = \left\{z \in \mathbb{R}^n_+ : \vee_{i\in L}\left(E^i z \geq f^i\right)\right\}$, where

$$E^i = \begin{pmatrix} E_1 \\ E_2^i \end{pmatrix} \text{ and } f^i = \begin{pmatrix} f_1 \\ f_2^i \end{pmatrix}, \quad i \in L.$$

Another way to represent a disjunctive set is using *conjunctive normal form*, where a set of points satisfies multiple disjunctive constraints and each disjunction contains exactly one inequality, i.e.,

$$\mathcal{R}^{CN} := \left\{z \in \mathbb{R}^n_+ : E_1 z \geq f_1, \bigwedge_{j=1}^{m}\left(\bigvee_{i\in L_j} \tilde{\phi}^i z \geq \tilde{\phi}^i_0\right)\right\}.$$

Here, $\tilde{\phi}^i \in \mathbb{R}^n$, $\tilde{\phi}^i_0 \in \mathbb{R}$ for all i, and $|L_j|$ is finite for all j. Note that as mentioned in the previous section, since the logical operations $\vee$ (disjunction) and $\wedge$ ("and" or conjunction) obey the distributive law, the set $\mathcal{R}$ can also be written in the conjunctive normal form where each disjunction j contains exactly one inequality from the system $E_2^i z \geq f_2^i$, $i \in L$, and $|L_j| = |L|$ for $j = 1, \ldots, m$.

Definition 1. *A linear programming relaxation of $\mathcal{R}$ and $\mathcal{R}^{CN}$ has no disjunctive constraint and is given by $\mathcal{R}_0 := \{z \in \mathbb{R}^n_+ : E_1 z \geq f_1\}$.*

Definition 2. *The disjunctive set $\mathcal{R}^{CN}$ (or conjunctive normal form of $\mathcal{R}$) is called* facial *if each inequality $\tilde{\phi}^i z \geq \tilde{\phi}^i_0$, $i \in L_j$, $j = 1, \ldots, m$, defines a face of $\mathcal{R}_0$.*

Definition 3. *An* extended formulation *of $\mathcal{R}$, denoted by $\mathcal{R}_{EF} := \{(z, \upsilon) \in \mathbb{R}^n_+ \times \mathbb{R}^{n_1}_+ : B_z z + B_\upsilon \upsilon \geq \beta_0\}$, is referred to as a* tight extended formulation *of $\mathcal{R}$ if and only if $conv(\mathcal{R}) = conv\left(Proj_z(\mathcal{R}_{EF})\right)$.*

2.1 Convex hull description of disjunctive programs

Balas (1979, 1998) provided a tight extended formulation for $\mathcal{R}$ and the convex hull description of $\mathcal{R}$ in the original space. Theorem 1 provides a tight extended formulation for the convex hull of the points satisfying disjunctive constraints. Theorem 2 provides the convex hull description of the union of the polyhedra, $\cup_{i\in L}\mathcal{R}_i$, in the original z-space.

Theorem 1 (Balas (1979, 1998))**.** *The closure of convex hull of $\cup_{i\in L}\mathcal{R}_i$ is the projection of the following extended formulation* (9) *onto the z-space:*

$$z = \sum_{i\in L} \zeta^i, \tag{9a}$$

$$E^i\zeta^i \geq f^i\zeta^i_0, \quad i \in L, \tag{9b}$$

$$\sum_{i\in L} \zeta^i_0 = 1, \tag{9c}$$

$$(\zeta^i, \zeta^i_0) \geq 0, \quad i \in L. \tag{9d}$$

Theorem 2 (Balas (1979, 1998)). *Let $\mathcal{R}_{TEF}$ be defined by $\{(z, \{\zeta^i, \zeta_0^i\}_{i\in L}) : (9) \text{ hold}\}$ and $\mathcal{R}_{hull} = conv(\cup_{i\in L}\mathcal{R}_i)$ be full dimensional (or has non-empty interior). Then, the projection of $\mathcal{R}_{TEF}$ onto the z-space is given by:*

$$Proj_z(\mathcal{R}_{TEF}) = \{z \in \mathbb{R}^n_+ : \alpha z \geq \beta \text{ for all } (\alpha, \beta) \in \mathcal{C}_0\},$$

where $\mathcal{C}_0 := \{(\alpha, \beta) \in \mathbb{R}^{n+1} : \alpha = \sigma^i E^i, \beta = \sigma^i f^i \text{ for some } \sigma^i \geq 0, i \in L\}$. The cone of all valid inequalities for $\mathcal{R}_{hull}$, denoted by $\mathcal{R}^*_{hull}$, is same as the polyhedral cone $\mathcal{C}_0$, i.e.

$$\begin{aligned}\mathcal{R}^*_{hull} &:= \{(\alpha, \beta) \in \mathbb{R}^{n+1} : \alpha z \geq \beta \text{ for all } z \in \mathcal{R}_{hull}\} \\ &= \{(\alpha, \beta) \in \mathbb{R}^{n+1} : \alpha = \sigma^i E^i, \beta = \sigma^i f^i \text{ for some } \sigma^i \geq 0, i \in L\}.\end{aligned}$$

Moreover, the inequality $\alpha z \geq \beta$ defines a facet of $\mathcal{R}_{hull}$ if and only if (α, β) is an extreme ray of the cone $\mathcal{R}^*_{hull}$.

Letting

$$W^h_\omega := \begin{pmatrix} W_\omega \\ D^h_{\omega,1} \end{pmatrix}, \; T^h_\omega := \begin{pmatrix} T_\omega \\ D^h_{\omega,2} \end{pmatrix}, \; r^h_\omega := \begin{pmatrix} r_\omega \\ d^h_{\omega,0} \end{pmatrix},$$

and $\mathcal{K}^h_\omega(x) := \{y_\omega \in \mathbb{R}^q_+ : W^h_\omega y_\omega \geq r^h_\omega - T^h_\omega x\} \neq \emptyset$ for $(\omega, x, h) \in (\Omega, X, H)$, we get

$$\mathcal{K}_\omega(x) = \bigcup_{h\in H} \mathcal{K}^h_\omega(x) = \left\{ y_\omega \in \mathbb{R}^q_+ : \bigvee_{h\in H} \left(W^h_\omega y_\omega \geq r^h_\omega - T^h_\omega x \right) \right\},$$

where $\mathcal{K}^h_\omega(x)$, for each (ω, x, h), is a polyhedral set. Also, let $X_{LP} := \{x \in \mathbb{R}^p : Ax \geq b, 0 \leq x_i \leq 1, i = 1, \ldots, p\}$ and $\mathcal{K}^\omega_{LP}(x) := \{y_\omega \in \mathbb{R}^q_+ : (3b) \text{ hold}\}$ for $(\omega, x) \in (\Omega, X)$ be the set of feasible solutions for the linear programming relaxation of the first and second stage problems, respectively. We also define a deterministic equivalent, also referred to as extensive formulation, of TSDR-DP as follows:

$$\min \quad c^T x + \max_{P\in\mathfrak{P}} \left\{ \mathbb{E}_P \left[g^T_\omega y_\omega \right] \right\} \tag{10a}$$

$$\text{s.t.} \quad Ax \geq b \tag{10b}$$

$$\bigvee_{h\in H} \left(T^h_\omega x + W^h_\omega y_\omega \geq r^h_\omega \right), \omega \in \Omega \tag{10c}$$

$$x \in \{0, 1\}^p \tag{10d}$$

$$y_\omega \in \mathbb{R}^q_+, \omega \in \Omega. \tag{10e}$$

Let $\mathcal{F} := \{(x, \{y_\omega\}_{\omega\in\Omega}) : (10b) - (10e) \text{ hold}\}$ be the feasible region of the extensive formulation.

We define a substructure of $\mathcal{F}$ for each $\omega \in \Omega$:

$$\mathcal{F}_\omega = \left\{ (x, y_\omega) \in X \times \mathbb{R}^q_+ : \bigvee_{h \in H} \left(T^h_\omega x + W^h_\omega y_\omega \geq r^h_\omega \right) \right\}.$$

Notice that $\mathcal{F} = \cap_{\omega \in \Omega} \mathcal{F}_\omega$.

2.2 Sequential convexification

Balas (1979, 1998) also exhibited a property of disjunctive programs according to which the convex hull of a set of points satisfying multiple disjunctive constraints, where each disjunction contains exactly one inequality, can be derived by sequentially generating the convex hull of points satisfying only one disjunctive constraint. This property is referred to as the *sequential convexification.* A subclass of DPs for which the sequential convexification property holds is called sequentially convexifiable DPs. Balas (1979, 1998) showed that the so-called *facial* DPs are sequentially convexifiable (Theorem 3). Interestingly, all pure binary and mixed 0-1 programming problems are facial DPs, while general mixed (or pure) integer programs are not facial DPs (Balas, 1998; Balas et al., 1993; Balas and Perregaard, 2002).

Theorem 3 (Balas (1979, 1998))**.** *If $\mathcal{R}^{CN}$ is facial then $\Pi_m = conv(\mathcal{R}^{CN})$, where $\Pi_0 := \mathcal{R}_0$ and*

$$\Pi_j = conv \left(\Pi_{j-1} \cap \left\{ z : \bigvee_{i \in L_j} \tilde{\phi}^i z \geq \tilde{\phi}^i_0 \right\} \right),$$

for $j = 1, \ldots, m$.

According to Theorem 3, the convex hull of the facial disjunctive set $\mathcal{R}^{CN}$ can be obtained in a sequence of m steps, where at each step the convex hull of points satisfying only one disjunctive constraints is generated. Later, Balas et al. (1989) extended the sequential convexification property for a general non-convex set with multiple constraints. They provided the necessary and sufficient conditions under which reverse convex programs (DPs with infinitely many terms) are sequentially convexifiable, and present classes of problems, in addition to facial DPs, which always satisfy the sequential convexification property.

3. Decomposition Algorithms for TSDR-DPs

We present two decomposition algorithms similar to the Benders' decomposition and L-shaped method (Van Slyke and Wets, 1969) to solve: (i) General TSDR-DPs (2) by extending the result of Balas (1979, 1998) for DP (Theorem 1) to get a linear programming equivalent for the second stage of TSDR-DPs, i.e. $\mathcal{K}_\omega(x)$; and (ii) TSDR-DPs (2) where disjunctive constraints (4c) in the second stage are sequentially convexifiable. We also provide conditions under which these algorithms are finitely convergent and illustrate our results for general TSDR-DPs using Example 1.

3.1 Decomposition algorithm for general TSDR-DPs

In the following theorem, we first extend the result of Theorem 1 (Balas, 1979, 1998) by providing conditions to get a linear programming equivalent for the second stage disjunctive programs of the TSDR-DPs, i.e. $\mathcal{K}_\omega(x)$. Thereafter, we utilize this result to develop a decomposition algorithm for TSDR-DPs.

Theorem 4. *For all $(x, \omega) \in (X, \Omega)$, the convex hull of $\mathcal{K}_\omega(x)$ is the projection of $\mathcal{K}^\omega_{tight}(x)$ onto y_ω space where*

$$\begin{aligned}\mathcal{K}^\omega_{tight}(x) := \Bigg\{ &\sum_{h\in H} \xi^h_{\omega,1} - y_\omega = 0, \sum_{h\in H} \xi^h_{\omega,2} = x,\\ &W^h_\omega \xi^h_{\omega,1} + T^h_\omega \xi^h_{\omega,2} \geq r^h_\omega \xi^h_{\omega,0},\ h \in H,\\ &\sum_{h\in H} \xi^h_{\omega,0} = 1,\\ &y_\omega \in \mathbb{R}^q_+, \xi^h_{\omega,1} \in \mathbb{R}^q_+, \xi^h_{\omega,2} \in \mathbb{R}^p_+, \xi^h_{\omega,0} \in \mathbb{R}_+, h \in H \Bigg\}.\end{aligned} \tag{11}$$

Proof. Using Theorem 1, we derive a tight extended formulation for $\mathcal{F}_\omega$, $\omega \in \Omega$, which is given by

$$\begin{aligned}\mathcal{F}^\omega_{tight} := \Bigg\{ &y_\omega = \sum_{h\in H} \xi^h_{\omega,1}, x = \sum_{h\in H} \xi^h_{\omega,2},\\ &W^h_\omega \xi^h_{\omega,1} + T^h_\omega \xi^h_{\omega,2} \geq r^h_\omega \xi^h_{\omega,0},\ h \in H,\\ &\sum_{h\in H} \xi^h_{\omega,0} = 1,\\ &x \in X, y_\omega \in \mathbb{R}^q_+, \xi^h_{\omega,1} \in \mathbb{R}^q_+, \xi^h_{\omega,2} \in \mathbb{R}^p_+, \xi^h_{\omega,0} \in \mathbb{R}_+, h \in H \Bigg\}.\end{aligned} \tag{12}$$

Therefore, $conv(\mathcal{F}_\omega) = conv(\text{Proj}_{x,y_\omega}(\mathcal{F}^\omega_{tight}))$ for $\omega \in \Omega$. Let $\hat{x} \in X$ and $\bar{y}_\omega \in \mathcal{K}_\omega(\hat{x})$. It is easy to see that $(\hat{x}, \bar{y}_\omega) \in \mathcal{F}_\omega$. Likewise for $\hat{y}_\omega \in conv(\mathcal{K}_\omega(\hat{x}))$, $(\hat{x}, \hat{y}_\omega) \in conv(\mathcal{F}_\omega)$ which implies that there exists $\hat{\xi}_\omega \in \mathbb{R}^{(q+p+1)|H|}_+$ such that $(\hat{x}, \hat{y}_\omega, \hat{\xi}_\omega) \in \mathcal{F}^\omega_{tight}$. Since $\mathcal{K}^\omega_{tight}(\hat{x}) = \text{Proj}_{x=\hat{x},y_\omega}\left(\mathcal{F}^\omega_{tight}\right)$, $\left(\hat{y}_\omega, \hat{\xi}_\omega\right) \in \mathcal{K}^\omega_{tight}(\hat{x})$. It implies that $conv(\mathcal{K}_\omega(x)) \subseteq \text{Proj}_{y_\omega}(\mathcal{K}^\omega_{tight}(x))$ for all $x \in X$.

Next, we consider the case when $(\hat{y}_\omega, \hat{\xi}_\omega) \in \mathcal{K}^\omega_{tight}(\hat{x})$ for $\hat{x} \in X$, which also means that $(\hat{x}, \hat{y}_\omega, \hat{\xi}_\omega) \in \mathcal{F}^\omega_{tight}$. Since $conv(\mathcal{F}_\omega) = \text{Proj}_{x,y_\omega}(\mathcal{F}^\omega_{tight})$, $(\hat{x}, \hat{y}_\omega)$ can be written as a convex combination of extreme points of $conv(\mathcal{F}_\omega)$. More specifically, $\hat{y}_\omega = \sum_i \lambda_i \tilde{y}_{\omega,i}$ such that $0 \leq \lambda_i \leq 1$, $\sum_i \lambda_i = 1$, and $(\hat{x}, \tilde{y}_{\omega,i})$ is an extreme point of $conv(\mathcal{F}_\omega)$. Now, for each $\hat{x} \in X$, i.e. $\hat{x} \in \{0,1\}^p$, $x = \hat{x}$ defines a face of $conv(\mathcal{F}_\omega)$ and because of the relatively complete recourse assumption, each extreme point of $conv(\mathcal{F}_\omega) \cap \{x = \hat{x}\}$ has y_ω component belonging to $\mathcal{K}_\omega(\hat{x})$. Therefore, $\tilde{y}_{\omega,i} \in \mathcal{K}_\omega(\hat{x})$ or $\hat{y}_\omega \in conv(\mathcal{K}_\omega(\hat{x}))$. Hence, $conv(\mathcal{K}_\omega(x)) \supseteq \text{Proj}_{y_\omega}(\mathcal{K}^\omega_{tight}(x))$ for all $x \in X$. □

Example 1 (continued). *Let* $X = \{x : \text{(8b)} - \text{(8d)} \text{ hold}\}$, $\mathcal{K}_\omega(x) = \{y_\omega : \text{(7b)} - \text{(7d)} \text{ hold}\}$, *and*

$$\mathcal{F}_\omega = \Big\{(x, y_\omega) \in X \times \mathbb{R}^2_+ : \; -\overline{x}_1 - \overline{x}_2 - y^1_\omega - y^2_\omega \geq d_\omega, \\ \big(-\overline{x}_2 + y^1_\omega \geq 1\big) \vee \big(\overline{x}_1 + \overline{x}_3 + y^2_\omega \geq 2\big) \Big\}$$

for $\omega \in \Omega$. *Using Theorems 1 and 4, the convex hull of* $\mathcal{K}_\omega(x)$ *for all* $(x, \Omega) \in (X, \Omega)$ *is given by the projection of* $\mathcal{K}^\omega_{tight}(x)$ *onto* $\big(y^1_\omega, y^2_\omega\big)$ *space, where*

$$\mathcal{K}^\omega_{tight}(x) = \Big\{(y^1_\omega, y^2_\omega, \xi^1_{\omega,1}, \xi^2_{\omega,1}, \xi^1_{\omega,2}, \xi^2_{\omega,2}, \xi^1_{\omega,0}, \xi^2_{\omega,0}) \in \mathbb{R}_+ \times \mathbb{R}_+ \times \mathbb{R}^2_+ \times \mathbb{R}^2_+ \times \mathbb{R}^3_+ \times \mathbb{R}^3_+ \times \mathbb{R}_+ \times \mathbb{R}_+ : \\ \xi^1_{\omega,1,i} + \xi^2_{\omega,1,i} - y^i_\omega = 0, \; i = 1, 2, \quad \xi^1_{\omega,2,j} + \xi^2_{\omega,2,j} = \overline{x}_j, \; j = 1, 2, 3, \\ \xi^1_{\omega,1,1} - \xi^1_{\omega,2,2} \geq \xi^1_{\omega,0}, \quad \xi^2_{\omega,1,2} + \xi^2_{\omega,2,1} + \xi^2_{\omega,2,3} \geq 2\xi^2_{\omega,0}, \quad \xi^1_{\omega,0} + \xi^2_{\omega,0} = 1, \\ -\xi^h_{\omega,1,1} - \xi^h_{\omega,1,2} - \xi^h_{\omega,2,1} - \xi^h_{\omega,2,2} \geq d_\omega \xi^h_{\omega,0}, \; h = 1, 2\Big\}. \qquad \square$$

Decomposition algorithm for TSDR-DPs. The pseudocode of our algorithm which utilizes Theorem 4 to solve general TSDR-DPs (2) is given by Algorithm 1. Let LB and UB be the lower and upper bound, respectively, on the optimal solution value of a given TSDR-DP. We denote the following linear programming equivalent of the second stage disjunctive program (3a)-(3d) by SLP(ω, x) for $(\omega, x) \in (\Omega, X)$:

$$\mathcal{Q}^\omega_{tight}(x) := \min\big\{g^T_\omega y_\omega : y_\omega \in \operatorname{Proj}_{y_\omega}(\mathcal{K}^\omega_{tight}(x))\big\}, \tag{13}$$

where $\mathcal{K}^\omega_{tight}(x)$ is defined by (11). Also, let $\pi^*_\omega(x)$ be the optimal dual multipliers obtained by solving the linear program SLP(ω, x) for a given $(\omega, x) \in (\Omega, X)$, and $\mathcal{K}^\omega_{tight}(x)$, defined by (11), be written in a compact form as

$$\Big\{N_{\omega,2}y_\omega + \sum_{h \in H}\Big(N^h_{\omega,3}\xi^h_{\omega,1} + N^h_{\omega,4}\xi^h_{\omega,2} + N^h_{\omega,5}\xi^h_{\omega,0}\Big) \geq \Delta_\omega - N_{\omega,1}x \\ y_\omega \in \mathbb{R}^q_+, \xi^h_{\omega,1} \in \mathbb{R}^q_+, \xi^h_{\omega,2} \in \mathbb{R}^p_+, \xi^h_{\omega,0} \in \mathbb{R}_+, h \in H\Big\},$$

where $N_{\omega,1}$, $N_{\omega,2}$, $N^h_{\omega,3}$, $N^h_{\omega,4}$, and $N^h_{\omega,5}$ are matrices/vectors associated with x, y_ω, $\xi^h_{\omega,1}$, $\xi^h_{\omega,2}$, and $\xi^h_{\omega,0}$, respectively, in the system of (in)equalities (11), and $\Delta_\omega = [0, 0, \ldots, 0, 1]^T$. Then, the corresponding *optimality cut*, OCS$(\pi^*_{\omega,0}(x), \{p_\omega\}_{\omega\in\Omega})$, is

$$\sum_{\omega \in \Omega} p_\omega \Big\{\pi^*_{\omega,0}(x)^T\left(\Delta_\omega - N_{\omega,1}x\right)\Big\} \leq \theta, \tag{14}$$

where $\{p_\omega\}_{\omega\in\Omega}$ is obtained by solving the distribution separation problem associated to the ambiguity set $\mathfrak{P}$, i.e. $\max_{P \in \mathfrak{P}}\big\{\mathbb{E}_P[\mathcal{Q}_\omega(x)] = \max_{P\in\mathfrak{P}} \mathbb{E}_P[\mathcal{Q}^\omega_{tight}(x)]\big\}$. These cuts help in deriving a lower bounding approximation of the first stage problem (2), i.e.

$$
\begin{aligned}
\min \ & c^T x + \theta \\
s.t. \ & x \in X \\
& \sum_{\omega \in \Omega} p_\omega^k \left(\pi_{\omega,0}^*(x^k)^T (\Delta_\omega - N_{\omega,1} x) \right) \leq \theta, \quad \text{for } k = 1, \ldots, l,
\end{aligned}
\tag{15}
$$

where $x^k \in X$ for $k = 1 \ldots, l$ and $\{p_\omega^k\}_{\omega \in \Omega} = \arg\max_{P \in \mathfrak{P}} \mathbb{E}_P[\mathcal{Q}_{tight}^\omega(x^k)]$. We denote problem (15) by $\mathcal{M}_l$ for $l \in \mathbb{Z}_+$ and refer to it as the master problem at iteration l. Note that $\mathcal{M}_0$ is the master problem without any optimality cut.

Algorithm 1 Decomposition Algorithm for TSDR-DPs in the Disjunctive Normal Form (2)

1: Initialization: $l \leftarrow 1$, $LB \leftarrow -\infty$, $UB \leftarrow \infty$. Assume $x^1 \in X$.
2: **while** $UB - LB > \epsilon$ **do** ▷ ϵ is a pre-specified tolerance
3:     **for** $\omega \in \Omega$ **do**
4:         Solve linear program SLP(ω, x^l) and store the following:
5:         $y_\omega^*(x^l) \leftarrow$ optimal solution; $\mathcal{Q}_{tight}^\omega(x^l) \leftarrow$ optimal solution value;
6:         $\pi_\omega^*(x^l) \leftarrow$ optimal dual multipliers;
7:     **end for**
8:     Solve distribution separation problem using $\mathcal{Q}_{tight}^\omega(x^l)$, $\omega \in \Omega$, to get $\{p_\omega^l\}_{\omega \in \Omega}$;
9:     **if** $UB > c^T x^l + \sum_{\omega \in \Omega} p_\omega^l \mathcal{Q}_{tight}^\omega(x^l)$ **then**
10:         $UB \leftarrow c^T x^l + \sum_{\omega \in \Omega} p_\omega^l \mathcal{Q}_{tight}^\omega(x^l)$;
11:         **if** $UB \leq LB + \epsilon$ **then**
12:             Go to Line 21;
13:         **end if**
14:     **end if**
15:     Derive optimality cut OCS$(\pi_{\omega,0}^*(x^l), \{p_\omega^l\}_{\omega \in \Omega})$ using (14);
16:     Add OCS$(\pi_{\omega,0}^*(x^l), \{p_\omega^l\}_{\omega \in \Omega})$ to $\mathcal{M}_{l-1}$ to get $\mathcal{M}_l$;
17:     Solve master problem $\mathcal{M}_l$ (a mixed binary program) using specialized lift-and-project algorithm of Balas et al. (1993);
18:     $(x^{l+1}, \theta^{l+1}) \leftarrow$ optimal solution of $\mathcal{M}_l$; $LB \leftarrow$ optimal solution value of $\mathcal{M}_l$;
19:     $l \leftarrow l + 1$;
20: **end while**
21: **return** $(x^l, \{y_\omega^*(x^l)\}_{\omega \in \Omega})$, UB

We initialize Algorithm 1 by setting LB to negative infinity, UB to positive infinity, iteration counter l to 1, and selecting a first stage feasible solution $x^1 \in X$ (Line 1). For each x^l and $\omega \in \Omega$, we solve linear program SLP(ω, x^l) to get optimal solution $y_\omega^*(x^l)$, the optimal objective value $\mathcal{Q}_{tight}^\omega(x^l) := g_\omega^T y_\omega^*(x^l)$, and the optimal dual multipliers $\pi_\omega^*(x^l)$ (Lines 3-7). In Line 8, we solve distribution separation problem for $x = x^l$ to get $\{p_\omega^l\}_{\omega \in \Omega}$. This provides us a feasible solution $(x^l, y_{\omega_1}^*(x^l), \ldots, y_{\omega_{|\Omega|}}^*(x^l))$ for the original problem as $y_\omega^*(x^l) \in \mathcal{K}_\omega(x^l)$ for all $\omega \in \Omega$, and a better upper bound UB if the solution value corresponding to this solution is smaller than the existing best known upper bound (Lines 9-13). In Lines 15-16, we then augment the master problem $\mathcal{M}_{l-1}$ to get $\mathcal{M}_l$ by adding optimality cut OCS$(\pi_{\omega,0}^*(x^l), \{p_\omega^l\}_{\omega \in \Omega})$, i.e. (14), to $\mathcal{M}_{l-1}$.

Notice that $\mathcal{M}_l$ is a mixed binary program where $\theta \in \mathbb{R}$ is the continuous variable. We utilize specialized lift-and-project algorithm of Balas et al. (1993) to solve $\mathcal{M}_l$ which terminates after a finite number of iterations (see Page 227 of Conforti et al. (2014)). Additionally, since it is a lower bounding approximation of (2), we use the optimal solution value associated with $\mathcal{M}_l$ to update LB. Observe that $\mathcal{M}_{l-1}$ is a relaxation of $\mathcal{M}_l$ for each $l \geq 1$ which means that after every iteration $UB - LB$ either decreases or remains same as in previous iteration. This is because UB and LB are non-increasing and non-decreasing, respectively, with respect to the iterations. Hence, whenever this difference becomes zero or reaches a pre-specified tolerance ϵ (Line 2 or Lines 10-12), we terminate the algorithm after returning the optimal solution $(x^l, \{y^*_\omega(x^l)\}_{\omega\in\Omega})$ and the optimal objective value UB.

Theorem 5 (Optimality Result)**.** *Algorithm 1 provides an optimal solution for the TSDR-DP* (2)*.*

Proof. Let $(x^*, \{y^*_\omega(x^*)\}_{\omega\in\Omega})$ be an optimal solution of a given TSDR-DP (2) instance, $(x^l,$ $\{y^*_\omega(x^l)\}_{\omega\in\Omega}) \in \mathcal{F}$ be a feasible solution utilized in iteration l of Algorithm 1, and (x^{l+1}, θ^{l+1}) be the solution obtained after solving master problem $\mathcal{M}_l$ in the iteration l. Observe that the upper bound $UB \leq c^T x^l + \sum_{\omega\in\Omega} p^l_\omega g^T_\omega y^*_\omega(x^l)$ as $(x^l, \{y^*_\omega(x^l)\}_{\omega\in\Omega}) \in \mathcal{F}$. Now, there are two possibilities: (i) $x^{l+1} = x^l$ or (ii) $x^{l+1} \neq x^l$. In the first case, i.e., $x^{l+1} = x^l$, we obtain $\theta^{l+1} \geq \sum_{\omega\in\Omega} p^l_\omega g^T_\omega y^*_\omega(x^l)$ by substituting $x = x^l$ in the optimality cut OCS$(\pi^*_{\omega,0}(x^l), \{p^l_\omega\}_{\omega\in\Omega})$, i.e. Inequality (14). Since θ^{l+1} is a lower bound approximation of the recourse function,

$$\theta^{l+1} \leq \max_{P\in\mathfrak{P}} \left\{\mathbb{E}_P\left[\mathcal{Q}_\omega(x^l)\right]\right\} = p^l_\omega g^T_\omega y^*_\omega(x^l)$$

as $\{p^l_\omega\}_{\omega\in\Omega}$ is the optimal solution of the distribution separation problem associated with $\mathfrak{P}$ and first stage feasible solution x^l. This implies $\theta^{l+1} = p^l_\omega g^T_\omega y^*_\omega(x^l)$ and $LB = c^T x^l + \theta^{l+1}$ is equal to UB. Since the termination condition, $UB - LB = 0$, has been satisfied, the algorithm will terminate at this iteration after returning the optimal solution $(x^l, \{y^*_\omega(x^l)\}_{\omega\in\Omega})$. In the second case, i.e. $x^{l+1} \neq x^l$, another iteration of "while" loop is performed and these iterations are repeated at most $|X|$ times (which is finite), until $(x^*, \{y^*_\omega(x^*)\}_{\omega\in\Omega}) = (x^k, \{y^*_\omega(x^k)\}_{\omega\in\Omega})$ for some $k \geq l+1$. □

Theorem 6 (Convergence Result)**.** *Algorithm 1 converges in finitely many iterations if the distribution separation algorithm associated to the ambiguity set $\mathfrak{P}$ for solving* (5) *is finitely convergent.*

Proof. We first prove that the number of iterations, l, is bounded from above by a finite number. In Algorithm 1, at the end of each iteration l, either $x^{l+1} \neq x^l$, or $x^{l+1} = x^l$. Since $|X|$ is finite because all the variables in the first stage of TSDR-DP are binary, the case $x^{l+1} \neq x^l$ can happen only finite number of times. On the other hand, in case $x^{l+1} = x^l$, we obtain $\theta^{l+1} \geq \sum_{\omega\in\Omega} p^l_\omega g^T_\omega y^*_\omega(x^l)$ by substituting $x = x^l$ in the optimality cut OCS$(\pi^*_{\omega,0}(x^l), \{p^l_\omega\}_{\omega\in\Omega})$, i.e. (14). Since θ^{l+1} is a lower bound approximation of the recourse function, $\theta^{l+1} \leq \max_{P\in\mathfrak{P}} \left\{\mathbb{E}_P\left[\mathcal{Q}_\omega(x^l)\right]\right\} = p^l_\omega g^T_\omega y^*_\omega(x^l)$ as $\{p^l_\omega\}_{\omega\in\Omega}$ is the optimal solution of the distribution separation problem associated with $\mathfrak{P}$ and first stage feasible solution x^l. This implies $\theta^{l+1} = p^l_\omega g^T_\omega y^*_\omega(x^l)$ and $LB = c^T x^l + \theta^{l+1}$ is equal to UB. Since termination condition, $UB - LB = 0$, has been satisfied, the algorithm will terminate at this iteration after returning the optimal solution $(x^l, \{y^*_\omega(x^l)\}_{\omega\in\Omega})$. In a nutshell, we get $l \leq |X| + 1$ which means the "while" loop, i.e. Lines 2-20, is repeated finite number of times.

Next, we ensure that Lines 3-19 in Algorithm 1 are performed in finite iterations. In these lines, for a given $x^l \in X$, we solve $|\Omega|$ number of linear programs, i.e. SLP(ω, x^l) for all $\omega \in \Omega$, distribution separation problem associated with the ambiguity set $\mathfrak{P}$, and a master problem $\mathcal{M}_l$. Since $|X|$ is finite, $\mathcal{M}_l$ is a mixed binary program (which can be solved in finite iterations using specialized lift-and-project algorithm of Balas et al. (1993)), and the distribution separation algorithm is finitely convergent (because of our assumption), it is clear that Lines 3-19 are performed in finite iterations. This implies that the overall algorithm also terminates in finitely many iterations. This completes the proof. □

Example 1 (continued). *We illustrate an iteration of Algorithm 1 to solve this example. We initialize the algorithm by solving a binary program,*

$$\begin{aligned}
\min\ & \overline{x}_1 - \overline{x}_2 - \overline{x}_3 \\
s.t.\ & \overline{x}_1 + \overline{x}_2 - 2\chi_{s_1} \geq 0, \\
& -\overline{x}_2 - \overline{x}_3 - 2\chi_{s_2} \geq -2, \\
& \overline{x}_1, \overline{x}_2, \overline{x}_3, \chi_{s_1}, \chi_{s_2} \in \{0, 1\},
\end{aligned}$$

to get $x^1 = (\overline{x}_1, \overline{x}_2, \overline{x}_3, \chi_{s_1}, \chi_{s_2}) = (1, 1, 1, 1, 0)$, and by setting $UB = \infty$ and $LB = -\infty$. Then, we solve the second stage problems for $x = x^1$ and $\omega \in \Omega$, and store

$$\begin{aligned}
&y^*_{\omega_1}(x^1) = (0, 3),\ \xi^2_{\omega_1,1,2} = 3,\ \xi^2_{\omega_1,2,j} = 1, j = 1, 2, 3,\ \xi^2_{\omega_1,0} = 1,\ \mathcal{Q}_{\omega_1}(x^1) = \mathcal{Q}^{\omega_1}_{tight}(x^1) = -3, \\
&y^*_{\omega_2}(x^1) = (0, 4),\ \xi^2_{\omega_2,1,2} = 4,\ \xi^2_{\omega_2,2,j} = 1, j = 1, 2, 3,\ \xi^2_{\omega_2,0} = 1,\ \mathcal{Q}_{\omega_2}(x^1) = \mathcal{Q}^{\omega_2}_{tight}(x^1) = -4, \\
&y^*_{\omega_3}(x^1) = (0, 2),\ \xi^2_{\omega_3,1,2} = 2,\ \xi^2_{\omega_3,2,j} = 1, j = 1, 2, 3,\ \xi^2_{\omega_3,0} = 1,\ \mathcal{Q}_{\omega_3}(x^1) = \mathcal{Q}^{\omega_3}_{tight}(x^1) = -2,
\end{aligned}$$

and $\xi^1_{\omega,1,i} = \xi^2_{\omega,1,1} = \xi^1_{\omega,2,j} = \xi^1_{\omega,0} = 0$ for $i \in \{1, 2\}$, $j \in \{1, 2, 3\}$, and $\omega \in \Omega$. Using $\mathcal{Q}_\omega(x^1)$, $\omega \in \Omega$, we also solve distribution separation problem,

$$\max\Big\{ -3p_{\omega_1} - 4p_{\omega_2} - 2p_{\omega_3} :\ (p_{\omega_1}, p_{\omega_2}, p_{\omega_3}) \in \mathfrak{P}_E \Big\},$$

and get optimal solution, i.e., $p^1_{\omega_1} = p^1_{\omega_2} = 0$ and $p^1_{\omega_3} = 1$. This also leads to an improved upper bound $UB = (1 - 1 - 1 + (-2)) = -3$.

Next, we derive an optimality cut OCS$(\pi^*_{\omega,0}(x^1), \{p^1_\omega\}_{\omega\in\Omega})$ as follows. After solving the second stage problems for $x = x^1$, we store optimal dual multipliers $\pi^*_{\omega,0}(x^1)$, $\omega \in \Omega$ and utilize them to *derive the optimality cut* (14),

$$\overline{x}_1 + \overline{x}_2 - 4 \leq \theta.$$

This cut is then added to $\mathcal{M}_0$ *to get* $\mathcal{M}_1$*, i.e.,*

$$\min \big\{\overline{x}_1 - \overline{x}_2 - \overline{x}_3 + \theta : \overline{x}_1 + \overline{x}_2 - 2\chi_{s_1} \geq 0,\ -\overline{x}_2 - \overline{x}_3 - 2\chi_{s_2} \geq -2,$$
$$\overline{x}_1 + \overline{x}_2 - 4 \leq \theta,\ \overline{x}_1, \overline{x}_2, \overline{x}_3, \chi_{s_1}, \chi_{s_2} \in \{0,1\}\big\}.$$

We solve $\mathcal{M}_1$ *and update LB to -4. We repeat these iterations until the termination condition (LB = UB) is satisfied. Note that the optimal solution of this example is* $(\overline{x}_1, \overline{x}_2, \overline{x}_3) = (0, 0, 0)$ *and optimal solution value is -4.* □

3.2 Decomposition algorithm for TSDR-DP with sequentially convexifiable DPs in the second stage

We present a decomposition algorithm to solve TSDR-DPs with sequentially convexifiable DPs in the second stage, which we denote by TSDR-SC-DPs, by harnessing the benefits of sequential convexification property within L-shaped method. As mentioned before, Balas (1979, 1998) introduced this property for a subclass of DPs, referred to as sequentially convexifiable DPs, according to which the convex hull of a set of points satisfying multiple disjunctive constraints, where each disjunction contains exactly one inequality, can be derived by sequentially generating the convex hull of points satisfying only one disjunctive constraint. Balas (1979, 1998) showed that the facial DPs are sequentially convexifiable and later, Balas et al. (1989) extended the sequential convexification property for a general non-convex set with multiple constraints. They provide the necessary and sufficient conditions under which reverse convex programs (DPs with infinitely many terms) are sequentially convexifiable and present classes of problems, in addition to the facial DPs, which always satisfy the sequential convexification property. In light of this discussion, it is clear that our algorithm for TSDR-SC-DPs will also solve various subclasses of TSDR-DPs.

The pseudocode of our decomposition algorithm for TSDR-SC-DP is presented in Algorithm 2. It is important to note that Algorithm 2 is similar in structure to the distributionally robust integer L-shaped algorithm of Bansal et al. (2018a) for TSDR-MBP (a special case of TSDR-SC-DP), except how parametric cuts are developed to solve the subproblems to optimality (discussed in Section 3.2.1). However for the sake of completeness of this paper, we explain all steps of Algorithm 2: Let the lower bound and upper bound on the optimal objective value of a given TSDR-SC-DP instances be denoted by LB and UB, respectively. We define subproblem $\mathcal{S}_\omega(x)$ for $(\omega, x) \in (\Omega, X)$

as follows:

$$
\begin{aligned}
\mathcal{Q}^s_\omega(x) := \min \ & g_\omega^T y_\omega && \text{(16a)}\\
s.t. \ & W_\omega y_\omega \geq r_\omega - T_\omega x && \text{(16b)}\\
& \alpha^t_\omega y_\omega \geq \beta^t_\omega - \psi^t_\omega x, \quad t = 1, \ldots, \tau_\omega && \text{(16c)}\\
& y_\omega \in \mathbb{R}^q_+, && \text{(16d)}
\end{aligned}
$$

where $\alpha^t_\omega \in \mathbb{Q}^q$, $\psi^t_\omega \in \mathbb{Q}^p$, and $\beta^t_\omega \in \mathbb{Q}$ are the coefficients associated with the parametric inequalities. We will discuss how these parametric inequalities, referred to as the parametric lift-and-project cuts, are developed in succession using sequential convexification approach of Balas (1979, 1998) in Section 3.2.1. Let $\pi^*_\omega(x) = (\pi^*_{\omega,0}(x), \pi^*_{\omega,1}(x), \ldots, \pi^*_{\omega,\tau_\omega}(x))^T$ be the optimal dual multipliers obtained by solving $\mathcal{S}_\omega(x)$ for a given $x \in X$ and $\omega \in \Omega$. Here, $\pi^*_{\omega,0}(x) \in \mathbb{R}^{m_2}$ corresponds to constraints (16b) and $\pi^*_{\omega,t}(x) \in \mathbb{R}$ corresponds to constraint (16c) for $t = 1, \ldots, \tau(\omega)$.

Algorithm 2 Decomposition Algorithm for TSDR-SC-DPs using Parametric Lift-and-Project Cuts

1: Initialization: $l \leftarrow 1$, $LB \leftarrow -\infty$, $UB \leftarrow \infty$, $\tau_\omega \leftarrow 0$ for all $\omega \in \Omega$. Assume $x^1 \in X$.
2: **while** $UB - LB > \epsilon$ **do** ▷ `ϵ is a pre-specified tolerance`
3:     **for** $\omega \in \Omega$ **do**
4:         Solve linear program $\mathcal{S}_\omega(x^l)$;
5:         $y^*_\omega(x^l) \leftarrow$ optimal solution; $\mathcal{Q}^s_\omega(x^l) \leftarrow$ optimal solution value;
6:     **end for**
7:     **if** $y^*_\omega(x^l) \notin \mathcal{K}_\omega(x^l)$ for some $\omega \in \Omega$ **then**
8:         **for** $\omega \in \Omega$ where $y^*_\omega(x^l) \notin \mathcal{K}_\omega(x^l)$ **do** ▷ `Add parametric inequalities`
9:             Add a parametric cut to $\mathcal{S}_\omega(x)$ as explained in Section 3.2.1;
10:             Set $\tau_\omega \leftarrow \tau_\omega + 1$ and solve linear program $\mathcal{S}_\omega(x^l)$;
11:             $y^*_\omega(x^l) \leftarrow$ optimal solution; $\mathcal{Q}^s_\omega(x^l) \leftarrow$ optimal solution value;
12:         **end for**
13:     **end if**
14:     Solve distribution separation problem using $\mathcal{Q}^s_\omega(x^l)$, $\omega \in \Omega$, to get $\{p^l_\omega\}_{\omega\in\Omega}$;
15:     **if** $y^*_\omega(x^l) \in \mathcal{K}_\omega(x^l)$ for all $\omega \in \Omega$ and $UB > c^T x^l + \sum_{\omega\in\Omega} p^l_\omega \mathcal{Q}^s_\omega(x^l)$ **then**
16:         $UB \leftarrow c^T x^l + \sum_{\omega\in\Omega} p^l_\omega \mathcal{Q}^s_\omega(x^l)$;
17:         **if** $UB \leq LB + \epsilon$ **then**
18:             Go to Line 28;
19:         **end if**
20:     **end if**
21:     $\pi^*_\omega(x^l) \leftarrow$ optimal dual multipliers obtained by solving $\mathcal{S}_\omega(x^l)$ for all $\omega \in \Omega$;
22:     Derive optimality cut $\mathrm{OC}(\pi^*_\omega(x^l), \{p^l_\omega\}_{\omega\in\Omega})$ using (17);
23:     Add $\mathrm{OC}(\pi^*_\omega(x^l), \{p^l_\omega\}_{\omega\in\Omega})$ to $\mathcal{M}_{l-1}$ to get $\mathcal{M}_l$;
24:     Solve master problem $\mathcal{M}_l$ using specialized lift-and-project algorithm of Balas et al. (1993);
25:     $(x^{l+1}, \theta^{l+1}) \leftarrow$ optimal solution of $\mathcal{M}_l$; $LB \leftarrow$ optimal solution value of $\mathcal{M}_l$;
26:     Set $l \leftarrow l + 1$;
27: **end while**
28: **return** $(x^l, \{y^*_\omega(x^l)\}_{\omega\in\Omega})$, UB

We derive a lower bounding approximation of the first stage problem (2) using the following *optimality cut*, $\mathrm{OC}(\pi^*_\omega(x), \{p_\omega\}_{\omega\in\Omega})$:

$$\sum_{\omega\in\Omega} p_\omega \Big\{ \pi^*_{\omega,0}(x)^T \left(r_\omega - T_\omega x\right) + \sum_{t=1}^{\tau_\omega} \pi^*_{\omega,t}(x) \left(\beta^t_\omega - \psi^t_\omega x\right) \Big\} \leq \theta, \tag{17}$$

where $\{p_\omega\}_{\omega\in\Omega}$ is obtained by solving the distribution separation problem associated to the ambiguity set $\mathfrak{P}$. More specifically, the lower bound approximation of the first stage problem (2), referred as the master problem $\mathcal{M}_l$ for $l \in \mathbb{Z}_+$, is given by:

$$\min\left\{c^T x + \theta : x \in X \text{ and } \mathrm{OC}(\pi^*_\omega(x^k), \{p^k_\omega\}_{\omega\in\Omega}) \text{ holds, for } k = 1, \dots, l\right\} \tag{18}$$

where $x^k \in X$ for $k = 1\dots, l$ and $\{p^k_\omega\}_{\omega\in\Omega} = \arg\max_{P\in\mathfrak{P}} \mathbb{E}_P[\mathcal{Q}^s_\omega(x^k)]$. Note that $\mathcal{M}_0$ is the master problem without any optimality cut and the optimality cut $\mathrm{OC}(\pi^*_\omega(x), \{p_\omega\}_{\omega\in\Omega})$ is valid because $\max_{P\in\mathfrak{P}} \mathbb{E}_P[\mathcal{Q}^s_\omega(x)] \leq \max_{P\in\mathfrak{P}} \mathbb{E}_P[\mathcal{Q}_\omega(x)] \leq \theta$ as $\mathcal{Q}^s_\omega(x) \leq \mathcal{Q}_\omega(x)$ for all $\omega \in \Omega$ and $x \in X$.

Now, we initialize Algorithm 2 by setting lower bound LB to negative infinity, upper bound UB to positive infinity, iteration counter l to 1, number of parametric inequalities τ_ω for all $\omega \in \Omega$ to zero, and by selecting a first stage feasible solution $x^1 \in X$ (Line 1). For each x^l and $\omega \in \Omega$, we solve linear programs $\mathcal{S}_\omega(x^l)$ to get optimal solution $y^*_\omega(x^l)$ and the optimal solution value $\mathcal{Q}^s_\omega(x^l) := g^T_\omega y^*_\omega(x^l)$ (Lines 3-6). In case $y^*_\omega(x^l) \notin \mathcal{K}_\omega(x^l)$ for any $\omega \in \Omega$, in Lines 8-12, we develop parametric lift-and-project cut for sequentially convexifiable DPs (explained in Section 3.2.1), add it to $\mathcal{S}_\omega(x)$, resolve the updated subproblem $\mathcal{S}_\omega(x)$ by fixing $x = x^l$, and obtain its optimal solution $y^*_\omega(x^l)$ along with the optimal solution value. Then, in Line 14, we solve the distribution separation problem associated to the ambiguity set $\mathfrak{P}$ using $\mathcal{Q}^s_\omega(x^l)$ and obtain the optimal solution, i.e. $\{p^l_\omega\}_{\omega\in\Omega}$. Whereas, in case $y^*_\omega(x^l) \in \mathcal{K}_\omega(x^l)$ for all $\omega \in \Omega$, then $(x^l, y^*_{\omega_1}(x^l), \dots, y^*_{\omega_{|\Omega|}}(x^l))$ is a feasible solution for the original problem. Moreover, if the solution value corresponding to thus obtained feasible solution is smaller than the existing upper bound, we update UB in Lines 15-16. In Lines 22-23, we augment the master problem $\mathcal{M}_{l-1}$ to get $\mathcal{M}_l$ by adding optimality cut $\mathrm{OC}(\pi^*_\omega(x^l), \{p^l_\omega\}_{\omega\in\Omega})$, i.e. (17), to $\mathcal{M}_{l-1}$.

Notice that $\mathcal{M}_l$ is a mixed binary program where $\theta \in \mathbb{R}$ is the continuous variable. We utilize specialized lift-and-project algorithm of Balas et al. (1993) to solve $\mathcal{M}_l$ which terminates after a finite number of iterations (see Page 227 of Conforti et al. (2014)). Additionally, since it is a lower bounding approximation of (2), we use the optimal value associated with $\mathcal{M}_l$ to update LB. Observe that $\mathcal{M}_{l-1}$ is a relaxation of $\mathcal{M}_l$ for each $l \geq 1$. It means that after every iteration $UB - LB$ either decreases or remains same as in previous iteration. This is because UB and LB are non-increasing and non-decreasing, respectively, with respect to the iterations. Hence, whenever this difference becomes zero or reaches a pre-specified tolerance ϵ (Line 2 or Lines 17-19), we terminate the algorithm after returning the optimal solution $(x^l, \{y_\omega(x^l)\}_{\omega\in\Omega})$ and the optimal objective value UB (Line 28).

In the following section, we discuss how to solve subproblems for a given $x \in X$, i.e. $\mathcal{S}_\omega(x)$, for all $\omega \in \Omega$. Also in Section 3.2.2, we investigate the conditions under which Algorithm 2 solves TSDR-SC-DPs in finitely many iterations.

3.2.1 Solving subproblems using parametric cuts

Here we present how a parametric lift-and-project cut of the form $\alpha^t_\omega y_\omega \geq \beta^t_\omega - \psi^t_\omega x$ where $t = \tau_\omega + 1$, is generated in Algorithm 2 (Line 9). Given a first stage feasible solution x^l at iteration l, assume that there exists an $\bar{\omega} \in \Omega$ such that the optimal solution of $\mathcal{S}_{\bar{\omega}}(x^l)$, i.e. $y^*_{\bar{\omega}}(x^l)$, does not belong to $\mathcal{K}_{\bar{\omega}}(x^l)$. This implies that there exists a disjunctive constraint, $\bigvee_{i \in H_j} \left(\eta^i_{\bar{\omega},1} y_{\bar{\omega}} \geq \eta^i_{\bar{\omega},0} - \eta^i_{\bar{\omega},2} x\right)$, $j \in \{1, \ldots, \bar{m}_2\}$, which is not satisfied by the point $(x^l, \{y^*_\omega(x^l)\}_{\omega \in \Omega})$. In order to generate an inequality which cuts this point, we first use Theorem 1 to get a tight extended formulation for the closed convex hull of

$$W_{\bar{\omega}} y_{\bar{\omega}} + T_{\bar{\omega}} x \geq r_{\bar{\omega}} \tag{19}$$

$$\alpha^t_{\bar{\omega}} y_{\bar{\omega}} + \psi^t_{\bar{\omega}} x \geq \beta^t_{\bar{\omega}},\ t = 1, \ldots, \tau_{\bar{\omega}} \tag{20}$$

$$\bigvee_{i \in H_j} \left(\eta^i_{\bar{\omega},1} y_{\bar{\omega}} + \eta^i_{\bar{\omega},2} x \geq \eta^i_{\bar{\omega},0}\right) \tag{21}$$

$$x \in X_{LP}, y_{\bar{\omega}} \in \mathbb{R}^q_+ \tag{22}$$

where $|H_j| = |H|$. Then, we project this tight extended formulation in the lifted space to the $(x, y_{\bar{\omega}})$ space using Theorem 2. Let $\mathcal{F}^j_{\bar{\omega}} = \{(x, y_{\bar{\omega}}) \in X_{LP} \times \mathbb{R}^q_+ : (19) - (21) \text{ hold}\}$ and its linear programming equivalent in the lifted space be given by

$$\begin{aligned}
\mathcal{F}^{\bar{\omega},j}_{tight} := \Bigg\{ & \sum_{i \in H_j} \xi^i_{\bar{\omega},1} - y_{\bar{\omega}} = 0, \sum_{i \in H_j} \xi^i_{\bar{\omega},2} - x = 0 \\
& W_{\bar{\omega}} \xi^i_{\bar{\omega},1} + T_{\bar{\omega}} \xi^i_{\bar{\omega},2} \geq r_{\bar{\omega}} \xi^i_{\bar{\omega},0}, \quad i \in H_j \\
& \alpha^t_{\bar{\omega}} \xi^i_{\bar{\omega},1} + \psi^t_{\bar{\omega}} \xi^i_{\bar{\omega},2} \geq \beta^t_{\bar{\omega}} \xi^i_{\bar{\omega},0},\ i \in H_j, t = 1, \ldots, \tau_{\bar{\omega}} \\
& \eta^i_{\bar{\omega},1} \xi^i_{\bar{\omega},1} + \eta^i_{\bar{\omega},2} \xi^i_{\bar{\omega},2} \geq \eta^i_{\bar{\omega},0} \xi^i_{\bar{\omega},0},\ i \in H_j \\
& \sum_{i \in H_j} \xi^i_{\bar{\omega},0} = 1 \\
& x \in X_{LP}, y_{\bar{\omega}} \in \mathbb{R}^q_+, \xi^i_{\bar{\omega},1} \in \mathbb{R}^q_+, \xi^i_{\bar{\omega},2} \in \mathbb{R}^p_+, \xi^i_{\bar{\omega},0} \in \mathbb{R}_+, i \in H_j \Bigg\}.
\end{aligned}$$

Let $\widehat{W}_{\bar{\omega}}$, $\widehat{T}_{\bar{\omega}}$, and $\widehat{r}_{\bar{\omega}}$ denote the constraint matrix associated with $y_{\bar{\omega}}$ variables, constraint matrix associated with x variables, and right-hand side vector, respectively, in the following system of inequalities:

$$
\begin{aligned}
W_{\bar{\omega}} y_{\bar{\omega}} + T_{\bar{\omega}} x &\geq r_{\bar{\omega}}, \\
Ax &\geq b, \\
x_i &\geq 0, i = 1, \ldots, p, \\
-x_i &\geq -1, i = 1, \ldots, p, \\
I y_{\bar{\omega}} \quad &\geq 0.
\end{aligned}
$$

Using Theorem 2, we derive the projection of $\mathcal{F}^{\bar{\omega},j}_{tight}$ onto the $(x, y_{\bar{\omega}})$ space, i.e. $\text{Proj}_{x,y_{\bar{\omega}}}(\mathcal{F}^{\bar{\omega},j}_{tight})$. This projection is given by

$$
\{(x, y_{\bar{\omega}}) \in \mathbb{R}^p \times \mathbb{R}^q : \alpha y_{\bar{\omega}} + \psi x \geq \beta \text{ for all } (\alpha, \psi, \beta) \in \mathcal{C}^j_{\omega}\} \tag{23}
$$

where

$$
\begin{aligned}
\mathcal{C}^j_{\omega} := \Bigg\{ &(\alpha, \psi, \beta) \in \mathbb{R}^p \times \mathbb{R}^q \times \mathbb{R} : \\
&\alpha = \sigma^i \begin{pmatrix} \widehat{W}_{\bar{\omega}} \\ \eta^i_{\bar{\omega},1} \\ \alpha^1_{\bar{\omega}} \\ \vdots \\ \alpha^{\tau_{\bar{\omega}}}_{\bar{\omega}} \end{pmatrix}, \ \psi = \sigma^i \begin{pmatrix} \widehat{T}_{\bar{\omega}} \\ \eta^i_{\bar{\omega},2} \\ \psi^1_{\bar{\omega}} \\ \vdots \\ \psi^{\tau_{\bar{\omega}}}_{\bar{\omega}} \end{pmatrix}, \ \beta = \sigma^i \begin{pmatrix} \widehat{r}_{\bar{\omega}} \\ \eta^i_{\bar{\omega},0} \\ \beta^1_{\bar{\omega}} \\ \vdots \\ \beta^{\tau_{\bar{\omega}}}_{\bar{\omega}}, \end{pmatrix} \\
&\text{for some } \sigma^i = \left(\sigma^{i,0} \sigma^{i,1}_c, \ldots, \sigma^{i,\tau_{\bar{\omega}}}_c\right) \in \mathbb{R}^{m_1+m_2+2p+1}_+ \times \mathbb{R}^{\tau_{\bar{\omega}}}_+, \ i \in H_j \Bigg\}.
\end{aligned}
$$

Next, we solve the following cut-generating linear program (CGLP) to find the most violated parametric lift-and-project cut among the defining inequalities of (23) for $(x^l, y^*_{\omega}(x^l))$:

$$
\max\{\beta - \alpha y^*_{\omega}(x^l) - \psi x^l : (\alpha, \psi, \beta) \in \mathcal{C}^j_{\bar{\omega}} \cap \mathcal{N}_{\bar{\omega}}\}, \tag{24}
$$

where $\mathcal{N}_{\bar{\omega}}$ is a normalization set (defined by one or more constraints) which truncates the cone $\mathcal{C}^j_{\bar{\omega}}$. Let $(\alpha^*, \psi^*, \beta^*)$ be the optimal solution for (24). Then, for $t = \tau_{\bar{\omega}} + 1$, we set $\alpha^t_{\bar{\omega}} = \alpha^*$, $\psi^t_{\bar{\omega}} = \psi^*$, and $\beta^t_{\bar{\omega}} = \beta^*$ to get the required parametric lift-and-project cut in Line 9 of Algorithm 2.

3.2.2 Optimality and finite convergence

We prove that Algorithm 2 solves TSDR-SC-DP to optimality and also present conditions under which it converges in finitely many iterations.

Theorem 7 (Optimality Result)**.** *Algorithm 2 provides an optimal solution for TSDR-SC-DPs.*

Proof. Let x^l and x^{l+1} be the first stage feasible solution obtained in iteration $l-1$ and l, respectively, of Algorithm 2, and $y^*_{\omega}(x^l)$ be the solution obtained by solving subproblem $\mathcal{S}_{\omega}(x^l)$ for

$\omega \in \Omega$. At the end of iteration l, there are three possibilities: (i) $x^{l+1} = x^l$ and $y^*_\omega(x^l) \in \mathcal{K}_\omega(x^l)$ for all $\omega \in \Omega$, (ii) $x^{l+1} = x^l$ and $y^*_\omega(x^l) \notin \mathcal{K}_\omega(x^l)$ for some $\omega \in \Omega$, or (iii) $x^{l+1} \neq x^l$. In Case (i), $(x^l, \{y^*_\omega(x^l)\}_{\omega\in\Omega}) \in \mathcal{F}$ and therefore, the upper bound $UB \leq c^T x^l + \sum_{\omega\in\Omega} p^l_\omega g^T_\omega y^*_\omega(x^l)$. Moreover, by substituting $x = x^l$ in the optimality cut $\text{OCS}(\pi^*_{\omega,0}(x^l), \{p^l_\omega\}_{\omega\in\Omega})$, i.e. Inequality (17), we get $\theta^{l+1} \geq \sum_{\omega\in\Omega} p^l_\omega g^T_\omega y^*_\omega(x^l)$. Since θ^{l+1} is a lower bound approximation of the recourse function, $\theta^{l+1} \leq \max_{P\in\mathfrak{P}} \left\{\mathbb{E}_{\xi_P}\left[\mathcal{Q}_\omega(x^l)\right]\right\} = p^l_\omega g^T_\omega y^*_\omega(x^l)$ as $\{p^l_\omega\}_{\omega\in\Omega}$ is the optimal solution of the distribution separation problem associated with $\mathfrak{P}$ and first stage feasible solution x^l. This implies $\theta^{l+1} = p^l_\omega g^T_\omega y^*_\omega(x^l)$ and $LB = c^T x^l + \theta^l$, which is equal to UB. Since termination condition, $UB - LB = 0$, has been satisfied, the algorithm will terminate at this iteration after returning the optimal solution $(x^l, \{y^*_\omega(x^l)\}_{\omega\in\Omega})$.

In Case (ii), we add lift-and-project cuts (Line 9) to the subproblems and resolve them in the subsequent iterations. Clearly, $x = x^l$ defines a face of $conv(\mathcal{F}_\omega)$ as $x^l \in \{0,1\}^p$ and since we assume relatively complete recourse, each extreme point of $conv(\mathcal{F}_\omega) \cap \{x = x^l\}$ has $y_\omega \in \mathcal{K}_\omega(x^l)$ for all $\omega \in \Omega$. Therefore, by repeating the "while" loop, i.e. adding globally valid parametric lift-and-projects cuts to the subproblems (with sequentially convexifiable DPs), we reach an iteration $k \geq l+1$ such that $y^*_\omega(x^k) \in \mathcal{K}_\omega(x^k)$ for all $\omega \in \Omega$ where $x^k = x^l$. This happens because the class of parametric lift-and-project is sufficient to describe the convex hull of $\mathcal{F}_\omega$ for TSDR-SC-DPs (Balas, 1998; Balas et al., 1989). Finally in Case (iii), i.e. $x^{l+1} \neq x^l$, another iteration of "while" loop is performed and these iterations are repeated until $x^k = x^{k+1}$ for some $k \geq l+1$. $\square$

Theorem 8 (Convergence Result)**.** *Algorithm 2 solves TSDR-SC-DP with facial DPs in the second stage in finitely many iterations if the distribution separation algorithm associated to the ambiguity set $\mathfrak{P}$ for solving* (5) *is finitely convergent.*

Proof. The arguments for this proof are similar to the finite convergence proof for distributionally robust integer L-shaped algorithm of Bansal et al. (2018a) for TSDR-MBP (a special case of TSDR-SC-DP), except the last paragraph of this proof. However for the sake of completeness, we provide all arguments. Since all the variables in the first stage of TSDR-SC-DP are binary, the number of first stage feasible solutions $|X|$ is finite. In Algorithm 2, for a given $x^l \in X$, we solve $|\Omega|$ number of linear programs, i.e. $\mathcal{S}_\omega(x^l)$ for all $\omega \in \Omega$, the distribution separation problem associated with the ambiguity set $\mathfrak{P}$, and a master problem $\mathcal{M}_l$ (after adding an optimality cut which requires a linear program to be solved). Notice that the master problem is a mixed binary program and can be solved using a finite number of cutting planes by specialized lift-and-project algorithm of Balas et al. (1993). Therefore, Lines 3-26 in Algorithm 2 are performed in finite iterations because we assume that the distribution separation algorithm is finitely convergent.

Now we have to ensure that the "while" loop in Line 2 terminates after finite iterations and provides the optimal solution. Notice that at the end of iteration l, either of the following two cases can happen: (i) $x^{l+1} \neq x^l$, or (ii) $x^{l+1} = x^l$. In the first case where $(x^{l+1}, \theta^{l+1}) \neq (x^l, \theta^l)$, (x^l, θ^l) will not be visited again in future iterations because the optimality cut generated in Line 22 cuts-off the point (x^l, θ^l) and this case can happen only a finite number of times because $|X|$ is finite. The second case can further be divided into two subcases: In the first subcase, let $y^*_\omega(x^l) \in \mathcal{K}_\omega(x^l)$ for all $\omega \in \Omega$. From the extensive formulation of TSDR-SC-DP, i.e. (10a)-(10e) where constraint (10c) is written in conjunctive normal form using (4b) and (4c), it is clear that $(x^l, \{y^*_\omega(x^l)\}_{w\in\Omega}) \in \mathcal{F}$ and hence, $UB = c^T x^l + \sum_{\omega\in\Omega} p^l_\omega g^T_\omega y^*_\omega(x^l)$ where $\{p^l_\omega\}_{\omega\in\Omega}$ is an optimal solution of the distribution separation problem associated with $\mathfrak{P}$ and first stage feasible solution x^l. Since (x^{l+1}, θ^{l+1}) is an optimal solution of $\mathcal{M}_l$, $\theta^{l+1} \geq \max_{P\in\mathfrak{P}} \left\{\mathbb{E}_P\left[\mathcal{Q}_\omega(x^l)\right]\right\} = \sum_{\omega\in\Omega} p^l_\omega g^T_\omega y^*_\omega(x^l)$ and $LB = c^T x^{l+1} + \theta^{l+1} \geq c^T x^l + \sum_{\omega\in\Omega} p^l_\omega g^T_\omega y^*_\omega(x^l) = UB$. This implies that $UB = LB$ and $(x^l, \{y^*_\omega(x^l)\}_{\omega\in\Omega})$ is the optimal solution, and we get $(x^{l+1}, \theta^{l+1}) = (x^l, \theta^l)$. Hence, in this subcase the algorithm terminates after returning the optimal solution and optimal objective value UB.

In the second subcase, let $y^*_{\bar\omega}(x^l) \notin \mathcal{K}_{\bar\omega}(x^l)$ for some $\bar\omega \in \Omega$. For this subcase, we derive a lift-and-project cut (Line 9) in $(x, y_{\bar\omega})$ subspace to cut-off the point $(x^l, y^*_{\bar\omega}(x^l))$, project this cutting plane to $y_{\bar\omega}$ space, add this (globally valid) parametric cut to $\mathcal{S}_\omega(x)$, and resolve the linear program for $x = x^l$. Since we assume relatively complete recourse, for each $\omega \in \Omega$, $\mathcal{K}_\omega(x^l)$ and its relaxations are nonempty. Notice that $x^l \in ver(X)$ where $ver(X)$ is the set of vertices of $conv(X)$ because X is defined by binary variables only. Hence, $x = x^l$ defines a face of $conv(\mathcal{F}_{\bar\omega})$ and since we assume relatively complete recourse, each extreme point of $conv(\mathcal{F}_{\bar\omega}) \cap \{x = x^l\}$ has $y_{\bar\omega} \in \mathcal{K}_{\bar\omega}(x^l)$. Therefore, according to the results of Jeroslow (1980), we can obtain $y^*_{\bar\omega}(x^l) \in \mathcal{K}_{\bar\omega}(x^l)$, by adding a finite number of parametric lift-and-projects cuts to $\mathcal{S}_\omega(x)$. This step can be repeated until $y^*_\omega(x^l) \in \mathcal{K}_\omega(x^l)$ for all $\omega \in \Omega$. As explained above, under such conditions, our algorithm terminates and returns the optimal solution after a finite number of iterations as $|\Omega|$ is finite. This completes the proof. $\square$

4. Two-stage distributionally robust semi-continuous programs

In this section, we showcase the significance of studying TSDR-DPs by introducing two-stage distributionally robust semi-continuous programs (TSDR-SCPs) where the first stage has binary variables and the second stage problems have semi-continuous variables. Specifically, TSDR-SCP is defined by (2), where

$$
\begin{aligned}
\mathcal{Q}_\omega(x) := \min\ & g^T_\omega y_\omega && (25a)\\
\text{s.t. }\ & W_\omega y_\omega \geq r_\omega - T_\omega x, && (25b)\\
& y^i_\omega \in [0, \underline{l}^i_\omega] \cup [\bar{l}^i_\omega, u^i_\omega], \quad i = 1, \ldots, q_1, && (25c)\\
& y^i_\omega \geq 0, \qquad\qquad\qquad i = q_1 + 1, \ldots, q, && (25d)
\end{aligned}
$$

such that $0 \leq \underline{l}^i_\omega < \bar{l}^i_\omega \leq u^i_\omega$ for $i = 1, \ldots, q_1$ and $\omega \in \Omega$. Note that by setting $\underline{l}^i_\omega = \bar{l}^i_\omega$ for all i and ω, the semi-continuous variables become continuous, and by setting $\underline{l}^i_\omega = 0$ and $\bar{l}^i_\omega = u^i_\omega = 1$, the semi-continuous variables become binary. Therefore, the two-stage distributionally robust problems with mixed 0-1 programs in the second stage, i.e., TSDR-MBPs, (Bansal et al., 2018a) are special cases of the TSDR-SCPs. In the following sections, we provide linear programming equivalent for the second stage: (i) TSDR-SCPs using DPs (Section 4.1), and (ii) TSDR semi-continuous network flow problem (Section 4.2).

4.1 Linear programming equivalent for the second stage of TSDR-SCPs

We re-write the semi-continuity constraints (25c) as disjunctive constraints,

$$\bigwedge_{i=1}^{q_1} \left\{ \left(0 \le y_\omega^i \le \underline{l}_\omega^i\right) \vee \left(\bar{l}_\omega^i \le y_\omega^i \le \bar{u}_\omega^i\right) \right\}, \tag{26}$$

thereby, showing that the class of TSDR-DPs subsumes the TSDR-SCPs. Next, in order to convexify $\mathcal{K}_\omega(x)$, $(\omega, x) \in \Omega \times X$, for the TSDR-SCP, we assume that $q_1 = 2$ (for the sake of convenience) and hence, the constraints (25c) or (26) in the disjunctive normal form is given by:

$$\begin{pmatrix} 0 \le y_\omega^1 \le \underline{l}_\omega^1 \\ 0 \le y_\omega^2 \le \underline{l}_\omega^2 \end{pmatrix} \bigvee \begin{pmatrix} 0 \le y_\omega^1 \le \underline{l}_\omega^1 \\ \bar{l}_\omega^2 \le y_\omega^2 \le \bar{u}_\omega^2 \end{pmatrix} \bigvee \begin{pmatrix} \bar{l}_\omega^1 \le y_\omega^1 \le \bar{u}_\omega^1 \\ 0 \le y_\omega^2 \le \underline{l}_\omega^2 \end{pmatrix} \bigvee \begin{pmatrix} \bar{l}_\omega^1 \le y_\omega^1 \le \bar{u}_\omega^1 \\ \bar{l}_\omega^2 \le y_\omega^2 \le \bar{u}_\omega^2 \end{pmatrix}. \tag{27}$$

Corollary 1. *Assuming $q = q_1 = 2$, a tight extended formulation for $\mathcal{K}_\omega(x) := \{y_\omega : (25\text{b})-(25\text{d})\}$, $(\omega, x) \in (\Omega, X)$, is given by*

$$\begin{aligned}
\mathcal{K}_{tight}^\omega(x) := \Bigg\{ & \sum_{h\in H} \xi_{1,1,\omega}^h - y_\omega^1 = 0, \sum_{h\in H} \xi_{1,2,\omega}^h - y_\omega^2 = 0, \\
& \sum_{h\in H} \xi_{\omega,2}^h = x, \\
& W_\omega(\xi_{1,1,\omega}^h, \xi_{1,2,\omega}^h)^T \ge r_\omega \xi_{\omega,0}^h - T_\omega \xi_{\omega,2}^h, \quad h \in H, \\
& D_{1,1}^h \xi_{1,1,\omega}^h + D_{1,2}^h \xi_{1,2,\omega}^h \le d_{\omega,0}^h \xi_{\omega,0}^h, \qquad h \in H, \\
& \sum_{h\in H} \xi_{\omega,0}^h = 1, y_\omega \in \mathbb{R}_+^2, \xi_{1,1,\omega}^h \in \mathbb{R}_+, \xi_{1,2,\omega}^h \in \mathbb{R}_+, \xi_{\omega,2}^h \in \mathbb{R}, \xi_{\omega,0}^h \in \mathbb{R}_+, \ h \in H \Bigg\}
\end{aligned}$$

where $H := \{1, 2, 3, 4\}$, $D_{1,1}^h = [-1 \quad 1 \quad 0 \quad 0]^T$, $D_{1,2}^h = [0 \quad 0 \quad -1 \quad 1]^T$ for all $h \in H$, and

$$d_{\omega,0}^1 = \begin{bmatrix} 0 \\ \underline{l}_\omega^1 \\ 0 \\ \underline{l}_\omega^2 \end{bmatrix}, d_{\omega,0}^2 = \begin{bmatrix} 0 \\ \underline{l}_\omega^1 \\ -\bar{l}_\omega^2 \\ \bar{u}_\omega^2 \end{bmatrix}, d_{\omega,0}^3 = \begin{bmatrix} -\bar{l}_\omega^1 \\ \bar{u}_\omega^1 \\ 0 \\ \underline{l}_\omega^2 \end{bmatrix}, d_{\omega,0}^4 = \begin{bmatrix} -\bar{l}_\omega^1 \\ \bar{u}_\omega^1 \\ -\bar{l}_\omega^2 \\ \bar{u}_\omega^2 \end{bmatrix}.$$

4.2 Two-stage distributionally robust semi-continuous network flow problem

Angulo et al. (2013) studied the so-called semi-continuous inflow set, defined by

$$\begin{aligned}\mathcal{S}(t,h) := \big\{(\theta,\eta) \in \mathbb{R}^n \times \mathbb{R}^n :\\ &\sum_{i\in N}\theta_i \geq \beta,\\ &t_i + \eta_i \geq \theta_i, && i \in N,\\ &\theta_i \in \{0\} \cup [h_i, \infty), && i \in N,\\ &\eta_i \in \{0\} \cup [l_i, \infty), && i \in N\big\},\end{aligned}$$

where $N := \{1, \ldots, n\}$ and provided a tight and compact extended formulation for $\mathcal{S}(0,h)$ which is given as follows:

$$\begin{aligned}\mathcal{S}_{tight} := \big\{(\theta,\eta,\gamma) \in \mathbb{R}^n \times \mathbb{R}^n \times \mathbb{R}^{|L|} :\\ &\sum_{i\in N\setminus L}\frac{\theta_i}{\max\{\beta,h_i\}} + \sum_{i\in L}\gamma_i \geq 1,\\ &\frac{\eta_i}{l_i} \geq \gamma_i, && i \in L,\\ &\frac{\theta_i}{\max\{\beta,h_i\}} \geq \gamma_i, && i \in L,\\ &\theta_i \geq 0, && i \in N,\\ &\eta_i \geq \theta_i, && i \in N\big\},\end{aligned}$$

where $L := \{i \in N : \max\{\beta, h_i\} < l_i\}$. They showed that $\mathcal{S}(t,h)$ arises as substructure in general semi-continuous network flow problem and semi-continuous transportation problem. Here, we consider the following TSDR-SCP with semi-continuous inflow set in the second stage, defined by (2) where

$$\mathcal{Q}_\omega(x) := \min \; g_{\omega,1}y_\omega + g_{\omega,2}z_\omega \tag{28}$$

$$s.t. \; \sum_{i\in N} y^i_\omega \geq d_\omega, \tag{29}$$

$$y^i_\omega - z^i_\omega \leq 0, \qquad i \in N, \tag{30}$$

$$x_i \leq d_\omega, \qquad i \in N, \tag{31}$$

$$y^i_\omega \in \{0\} \cup [x_i, \infty), \qquad i \in N, \tag{32}$$

$$z^i_\omega \in \{0\} \cup [l^i_\omega, \infty), \qquad i \in N. \tag{33}$$

Here, $g_{\omega,1} \in \mathbb{R}^n$, $g_{\omega,2} \in \mathbb{R}^n$, and $l_\omega \in \mathbb{R}^n_+$. Let $\mathcal{K}_\omega(x) := \{(y_\omega, z_\omega) : (29) - (33)\}$.

Proposition 1. *For each $(x,\omega) \in (X,\Omega)$,*

$$\begin{aligned}
\mathcal{K}^{\omega}_{tight}(x) := \big\{(y_\omega, z_\omega, u_\omega) \in \mathbb{R}^n \times \mathbb{R}^n \times \mathbb{R}^{|L_\omega|} :\\
&\frac{z^i_\omega}{l^i_\omega} \geq u^i_\omega, && i \in L_\omega,\\
&\frac{y^i_\omega}{d_\omega} \geq u^i_\omega, && i \in L_\omega,\\
&x_i \leq d_\omega, && i \in N,\\
&y^i_\omega \geq 0, && i \in N,\\
&z^i_\omega \geq y^i_\omega, && i \in N,\\
&\sum_{i \in N \setminus L_\omega} \frac{y^i_\omega}{d_\omega} + \sum_{i \in L_\omega} u^i_\omega \geq 1\big\},
\end{aligned}$$

where $L_\omega := \{i \in N : d_\omega < l^i_\omega\}$, is a tight extended formulation of $\mathcal{K}_\omega(x)$, i.e. $conv(\mathcal{K}_\omega(x)) = Proj_{y_\omega, z_\omega}(\mathcal{K}^{\omega}_{tight}(x))$.

Proof. For each $\omega \in \Omega$, we substitute y^i_ω, z^i_ω, x_i, and d_ω in $\mathcal{K}_\omega(x)$ by θ_i, η_i, h_i, and β, respectively, to get $\mathcal{S}(0,h)$ where $\beta \geq h_i$ for all $i \in N$. Now, by applying the compact extended formulation for $\mathcal{S}(0,h)$ we get a tight extended formulation for $\mathcal{K}_\omega(x)$, i.e. $\mathcal{K}^{\omega}_{tight}(x)$ for all $x \in X$. This completes the proof. □

5. Conclusion and Future Work

We introduced two-stage distributionally robust disjunctive programs (TSDR-DPs) with general ambiguity set. We extended the results of Balas (1979, 1998) developed for deterministic disjunctive programs to TSDR-DPs. More specifically, we provided linear programming equivalent for the second stage of TSDR-DPs and utilized it within a decomposition algorithm to solve general TSDR-DPs. Additionally, by utilizing the sequential convexification approach of Balas (1979), we developed another decomposition algorithm to solve TSDR-DPs where second stage programs are facial DPs (in finite iterations), and sequentially convexifiable DPs (which include some nonconvex programs such as nonconvex quadratic programs, separable non-linear programs, reverse convex programs, etc.). Furthermore, we showcased the significance of studying TSDR-DPs by reformulating TSDR semi-continuous programs (TSDR-SCPs) as TSDR-DPs and then deriving linear programming equivalent for the second stage SCPs. We also provide linear programming equivalent for TSDR-SCP with semi-continuous inflow set in the second stage.

A potential future extension can be to perform computational study on the performance of the algorithms presented in this paper for solving TSDR-DPs and various applied optimization problems which can be formulated as TSDR-DPs. It would involve consideration of various strategies to generate parametric lift-and-project inequalities for different classes of TSDR-SC-DPs and in addition, there is a scope to further strengthen these inequalities using various cut-generating procedures such as mixed integer rounding (MIR) (Nemhauser and Wolsey, 1990), mixing (Günlük and Pochet, 2001), continuous multi-mixing (Bansal and Kianfar, 2015), mingled n-step cycling (Bansal and Kianfar, 2017), and many more.

References

Angulo G, Ahmed S, Dey SS. Semi-continuous network flow problems. Mathematical Programming 2013;145(1-2):565–99.

Balas E. Intersection cuts from disjunctive constraints. MSRR No 330, Carnegie-Mellon University 1974;.

Balas E. Disjunctive Programming. In: Hammer P, Johnson EL, Korte BH, editors. Annals of Discrete Mathematics. Elsevier; volume 5 of *Discrete Optimization II Proceedings of the Advanced Research Institute on Discrete Optimization and Systems Applications of the Systems Science Panel of NATO and of the Discrete Optimization Symposium*; 1979. p. 3–51.

Balas E. Disjunctive programming: Properties of the convex hull of feasible points. Discrete Applied Mathematics 1998;89(13):3–44.

Balas E, Ceria S, Cornujols G. A lift-and-project cutting plane algorithm for mixed 01 programs. Mathematical Programming 1993;58(1-3):295–324.

Balas E, Perregaard M. Lift-and-project for mixed 01 programming: recent progress. Discrete Applied Mathematics 2002;123(13):129–54.

Balas E, Tama JM, Tind J. Sequential convexification in reverse convex and disjunctive programming. Mathematical Programming 1989;44(1-3):337–50.

Bansal M, Huang KL, Mehrotra S. Decomposition algorithms for two-stage distributionally robust mixed binary programs. SIAM Journal on Optimization 2018a;28(3):2360–83.

Bansal M, Huang KL, Mehrotra S. Tight second-stage formulations in two-stage stochastic mixed integer programs. SIAM Journal on Optimization 2018b;28(1):788–819.

Bansal M, Kianfar K. n-step cycle inequalities: facets for continuous multi-mixing set and strong cuts for multi-module capacitated lot-sizing problem. Mathematical Programming 2015;154(1):113–44.

Bansal M, Kianfar K. Facets for continuous multi-mixing set with general coefficients and bounded integer variables. Discrete Optimization 2017;16:1–25.

Bansal M, Zhang Y. Two-stage stochastic and distributionally robust p-order conic mixed integer programs. Technical Report 2018;Available at `http://www.optimization-online.org/DB_HTML/2018/05/6630.html`.

Ben-Tal A, den Hertog D, De Waegenaere A, Melenberg B, Rennen G. Robust Solutions of Optimization Problems Affected by Uncertain Probabilities. Management Science 2012;59(2):341–57.

Benders JF. Partitioning procedures for solving mixed-variables programming problems. Numerische Mathematik 1962;4(1):238–52.

Bertsekas DP, Gallager . Data Networks. 2nd ed. Englewood Cliffs, N.J: Pearson, 1992.

Bertsimas D, Doan XV, Natarajan K, Teo CP. Models for minimax stochastic linear optimization problems with risk aversion. Mathematics of Operations Research 2010;35(3):580–602.

Bertsimas D, Popescu I. Optimal inequalities in probability theory: A convex optimization approach. SIAM Journal on Optimization 2005;15(3):780–804.

Calafiore G. Ambiguous Risk Measures and Optimal Robust Portfolios. SIAM Journal on Optimization 2007;18(3):853–77.

Carøe CC, Tind J. A cutting-plane approach to mixed 0–1 stochastic integer programs. European Journal of Operational Research 1997;101(2):306–16.

Conforti M, Cornujols G, Zambelli G. Integer Programming. volume 271 of *Graduate Texts in Mathematics*. Cham: Springer International Publishing, 2014.

Delage E, Ye Y. Distributionally robust optimization under moment uncertainty with application to data-driven problems. Operations Research 2010;58:595–612.

Dupacová J. The minimax approach to stochastic programming and an illustrative application. Stochastics 1987;20:73–88.

Dyer M, Stougie L. Computational complexity of stochastic programming problems. Mathematical Programming 2006;106(3):423–32.

Erdogan E, Iyengar G. Ambiguous chance constrained problems and robust optimization. Mathematical Programming 2006;107(1-2):37–61.

Gade D, Kücükyavuz S, Sen S. Decomposition algorithms with parametric gomory cuts for two-stage stochastic integer programs. Mathematical Programming 2014;144(1-2):39–64.

Grossmann IE, Caballero JA, Yeomans H. Mathematical programming approaches to the synthesis of chemical process systems. Korean Journal of Chemical Engineering 1999;16(4):407–26.

Gunluk O, Lee J, Weismantel R. MINLP Strengthening for Separable Convex Quadratic Transportation-Cost UFL. IBM Technical Report 2007;:18.

Günlük O, Pochet Y. Mixing mixed-integer inequalities. Mathematical Programming 2001;90(3):429–57.

Hanasusanto GA, Kuhn D. Conic programming reformulations of two-stage distributionally robust linear programs over wasserstein balls 2016;under review. Available at https://arxiv.org/abs/1609.07505.

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

Jeroslow R. A cutting-plane game for facial disjunctive programs. SIAM Journal on Control and Optimization 1980;18(3):264–81.

Jiang R, Guan Y. Risk-averse two-stage stochastic program with distributional ambiguity 2015;Available at http://www.optimization-online.org/DB_FILE/2015/05/4908.pdf.

Jnger M, Liebling TM, Naddef D, Nemhauser GL, Pulleyblank WR, Reinelt G, Rinaldi G, Wolsey LA. 50 Years of Integer Programming 1958-2008: From the Early Years to the State-of-the-Art. Springer Science & Business Media, 2009.

Kücükyavuz S, Sen S. An introduction to two-stage stochastic mixed-integer programming. In: The Operations Research Revolution (INFORMS Tutorials in Operations Research. 2017. p. 1–27. doi:doi: 10.1287/educ.2017.0171.

Laporte G, Louveaux FV. The integer L-shaped method for stochastic integer programs with complete recourse. Operations Research Letters 1993;13(3):133–42.

Love DK, Bayraksan G. Phi-divergence constrained ambiguous stochastic programs for data-driven optimization 2016;under review. Available at http://www.optimization-online.org/DB_HTML/2016/03/5350.html.

Nemhauser GL, Wolsey LA. Integer and combinatorial optimization. New York, NY, USA: Wiley-Interscience, 1988.

Nemhauser GL, Wolsey LA. A recursive procedure to generate all cuts for 01 mixed integer programs. Mathematical Programming 1990;46(1-3):379–90.

Ntaimo L. Disjunctive decomposition for two-stage stochastic mixed-binary programs with random recourse. Operations Research 2009;58(1):229–43.

Pflug G, Pichlera A, Wozabalb D. The 1/n investment strategy is optimal under high model ambiguity. Journal of Banking and Finance 2012;36(2):410–7.

Pflug G, Wozabal D. Ambiguity in portfolio selection. Quantitative Finance 2007;7(4):435–442,.

Prékopa A. Stochastic Programming. Kluwer Academic, 1995.

Rahimian H, Bayraksan G, Homem-de Mello T. Identifying effective scenarios in distributionally robust stochastic programs with total variation distance. Mathematical Programming 2018;:1–38.

Scarf H. A min-max solution of an inventory problem. In: Studies in the Mathematical Theory of Inventory and Production. 1958. p. 201–9.

Sen S, Higle JL. The C^3 theorem and a D^2 algorithm for large scale stochastic mixed-integer programming: set convexification. Mathematical Programming 2005;104(1):1–20.

Shapiro A. Distributionally robust stochastic programming. SIAM Journal on Optimization 2017;27(4):2258–75.

Shapiro A. Tutorial on risk neutral, distributionally robust and risk averse multistage stochastic programming. 2018. Available at http://www.optimization-online.org/DB_HTML/2018/02/6455.html.

Shapiro A, Ahmed S. On a class of minimax stochastic programs. SIAM Journal on Optimization 2004;14(4):1237–49.

Sherali HD, Fraticelli BM. A modification of benders' decomposition algorithm for discrete subproblems: An approach for stochastic programs with integer recourse. Journal of Global Optimization 2002;22(1-4):319–42.

Sun H, Xu H. Convergence Analysis for Distributionally Robust Optimization and Equilibrium Problems. Mathematics of Operations Research 2015;41(2):377–401.

Trkay M, Grossmann IE. Logic-based MINLP algorithms for the optimal synthesis of process networks. Computers & Chemical Engineering 1996;20(8):959–78.

Van Slyke R, Wets R. L-shaped linear programs with applications to optimal control and stochastic programming. SIAM Journal on Applied Mathematics 1969;17(4):638–63.

Wang Z, Glynn PW, Ye Y. Likelihood robust optimization for data-driven problems. Computational Management Science 2016;13(2):241–61.

Wood AJ, Wollenberg BF, Shebl GB. Power Generation, Operation, and Control. 3rd ed. Hoboken, New Jersey: Wiley-Interscience, 2013.

Wozabal D. A framework for optimization under ambiguity. Annals of Operations Research 2012;193(1):21–47.

Yankolu h, den Hertog D. Safe Approximations of Ambiguous Chance Constraints Using Historical Data. INFORMS Journal on Computing 2012;25(4):666–81.

Zhao C, Guan Y. Data-driven risk-averse two-stage stochastic program with ζ-structure probability metrics 2015;under review. Available at http://www.optimization-online.org/DB_HTML/2015/07/5014.html.