Semi-Infinite Generalized Disjunctive and Mixed Integer Convex Programs with(out) Uncertainty[1]

Manish Bansal
Grado Department of Industrial and Systems Engineering
Virginia Tech, Blacksburg, VA 24060 {bansal@vt.edu}

Abstract. In this paper, we introduce semi-infinite generalized disjunctive programs that are defined by logical propositions along with disjunctions of sets of logical equations and infinite number of algebraic inequalities. We denote these programs by SIGDPs. For SIGDPs with linear and convex inequalities, we present new reformulations: semi-infinite mixed-binary/disjunctive linear programs and semi-infinite mixed-binary/disjunctive convex programs, respectively. These results are also applicable for solving SIGDPs with nonconvex functions using their convex underestimators. Even for finite GDPs, this leads to introduction of new reformulations that have no big-M parameters and have lesser number of variables, in comparison to reformulations known in the literature for finite GDPs. We also present a tight extended formulation for semi-infinite disjunctive convex programs. Additionally, we study semi-infinite mixed integer convex program after binarizing integer variables (a special case of SIGDP) and present: (a) semi-infinite convex programming equivalent in higher dimensional space, (b) hierarchy of relaxations between the continuous and convex hull of its feasible region, and (c) sequential convexification approach. We showcase the applications of these results for solving robust 0-1 convex programs and distributionally robust convex programs as well. Furthermore, we provide Lagrangian relaxation based approaches embedded with aforementioned results for deriving lower bounds to solve two-stage stochastic and distributionally robust SIGDPs with general ambiguity set.

Keywords. Semi-infinite programs; 0-1 convex programs; generalized disjunctive programs; distributionally robust optimization; hierarcy of relaxation; extended formulation.

1. Introduction

In Leibniz's algebra of concepts and Boolean algebra, disjunction ("or") and conjunction ("and") are logical operators that are used in constructing logical statements and are denoted by $\vee$ and $\wedge$, respectively. A logical disjunctive statement $\psi_1 \vee \psi_2$ is true when at least one of the input statements (disjuncts, i.e., ψ_1 and ψ_2) is true. In the field of mathemat-

[1]First Draft: December 25, 2022

ical optimization, disjunctive programming is a well-known algebraic modeling framework that allows disjunctions of polyhedral sets as feasible region [3, 4, 5, 7, 15, 30, 31], thereby subsuming 0-1 programs, mixed 0-1 programs, and programs with discrete variables [32]. A disjunctive linear program (DLP) is defined as

$$\min_{x \in \mathbb{R}^n_+} \left\{ \sum_{i=1}^{n} c_i x_i : \bigvee_{h \in H} \left(A^h x \geq b^h \right) \right\}, \tag{1}$$

where A^h and b^h are finite dimensional vectors, and the disjunctive constraint imply that the feasible region $\mathcal{D}$ is the union of polyhedral sets $\mathcal{D}^h := \{x \in \mathbb{R}^n_+ : A^h x \geq b^h\}$ for $h \in H$. In simple words, $\mathcal{D} = \cup_{h \in H} \mathcal{D}^h$. Raman and Grossman [40] introduced a new generalized modeling framework, referred to as generalized disjunctive programs (GDPs), by incorporating logical and algebraic equations within disjunctions, and conjunction with logical propositions. A motivation behind introduction of GDPs was to address optimization problems arising in complicated chemical engineering processes, such as finding optimal design of reactive distillation columns [29] and optimal synthesis of process networks [48]. Lately, even a truck loading problem has been modelled as GDP [38]. A GDP in its general form is defined as follows [40]:

$$\min \Bigg\{ f(x) : q(x) \leq 0 \tag{2a}$$

$$\bigvee_{h \in H_k} \begin{bmatrix} \Gamma_{hk} \\ g^j_{hk}(x) \leq 0, \ j \in J \end{bmatrix} \quad \text{for } k \in K, \tag{2b}$$

$$\Omega(\Gamma) = \text{True}, \tag{2c}$$

$$x \in \mathbb{R}^n_+, \Gamma_{hk} \in \{\text{True, False}\}, h \in H_k, k \in K \Bigg\}, \tag{2d}$$

where f, q, and g^j_{hk} are functions of x, index sets K, H_k, and J are finite, and Γ_{hk} is a Boolean variable associated to disjunct h of kth disjunctive constraint (2b). There are $|K|$ disjunctive constraints (2b) comprising of algebraic inequalities and Boolean variables (logical equations). The algebraic inequalities, $g^j_{hk}(x) \leq 0$ for $j \in J$, are imposed whenever the Boolean variable Γ_{hk} is True, and are ignored otherwise. Constraints (2c) represent logical propositions connecting the Boolean variables using conjunctive and disjunctive operators; for an example, $(\Gamma_{11} \vee \Gamma_{12} \vee \ldots \vee \Gamma_{1k}) \wedge (\neg \Gamma_{21} \vee \Gamma_{32} \vee \ldots \vee \Gamma_{k,k-1}) = \text{True}$ is a logical proposition where $\neg$ denotes negation of the Boolean variable. Depending on definition of functions f, q, and g^j_{hk}, the GDPs (2) have been classified into three categories: (i) Linear GDPs, denoted by GDLPs [41], (ii) Convex GDPs, denoted by GDCPs [36], and (iii) Nonconvex GDPs [26].

Notice that DLP (1) is a special case of GDLP. Refer to Section 2 for a review of literature on DLPs and GDPs.

Since the development of simplex method, duality, and sensitivity analysis for finite linear programs, i.e., DLP (1) with $|H| = 1$, researchers have been generalizing these results (starting in 1960s) for semi-infinite linear programs (SILPs), i.e., linear programs with infinite number of constraints but finite number of variables, or vice-versa [10, 19, 20, 25, 27, 28][2]. In parallel, Charnes et al. [21] and Ben-Tal et al. [13] derived optimality conditions for semi-infinite convex programs (SICPs), and Kortanek and No [34] extended a so-called central cutting plane algorithm of Gribik [25] for SILPs to SICPs. Recently, in [8], the author introduced semi-infinite disjunctive linear programs (SIDLPs) where each disjunct has infinite number of linear constraints, i.e.,

$$\min_{x \in \mathbb{R}^n_+} \left\{ \sum_{i=1}^{n} c_i x_i : \bigvee_{h \in H} \left(\sum_{i=1}^{n} a^h_{ij} x \geq b^h_j, \quad j \in J \right) \right\}, \tag{3}$$

where J is infinite. Note that SIDLPs subsume semi-infinite mixed binary/discrete linear programs [8]. Refer to Section 2 for a review of literature on semi-infinite programs as well.

Contributions and Organization of this Paper. We introduce semi-infinite GDP (SIGDP) which is a GDP (2) with infinite number of algebraic inequalities in each disjunct of disjunctive inequalities (2b), i.e., $|J|$ is infinite. For SIGDPs with linear and convex functions, we present new reformulations, i.e., semi-infinite mixed binary linear and convex programs without big-M parameters, and equivalent semi-infinite disjunctive linear and convex programs, respectively (Section 3). These results are also applicable for solving SIGDPs with nonconvex functions using their convex underestimators [47]. The SIGDP with linear and convex functions are denoted by SIGDLP and SIGDCP, respectively. Interestingly, even for finite GDP, these reformulations without big-M parameters are new and they require lesser number of variables in comparison to known big-M free reformulations in [36, 40].

In Section 4, we reformulate semi-infinite mixed integer convex programs as semi-infinite mixed binary convex programs (SIMBCPs) by "binarizing" the integer variables [23], and then study SIMBCPs that are defined by

$$\min_{\mathbf{x} \in [0,1]^n} \left\{ \sum_{i=1}^{n} c_i x_i : g_j(x) \leq 0, j \in J; \quad (x_i = 0 \vee x_i = 1 \text{ for } i = 1, \ldots, p) \right\}, \tag{4}$$

[2]The traces of genesis of semi-infinite programs can be found in seminal work of mathematician Haar (back in 1924) [1], published in German in an obscure journal. In the early literature, SILPs were also referred to as "Haar Programs".

where g_j is a convex function and $|J|$ is infinite. We present the following results for SIMBCPs: (a) semi-infinite convex programming equivalent in higher dimensional space, (b) hierarchy of relaxations between the continuous and convex hull of the feasible region of (4), and (c) sequential convexification approach for them. We showcase that semi-infinite mixed-discrete convex programs, robust 0-1 convex problem, and distributionally robust 0-1 convex programs can be reformulated as SIMBCPs, and as a consequence, the foregoing results are applicable for solving these extensions/variants of MBCP/SIMBCPs as well.

In Section 5, we study SIGDPs with uncertain data parameters. Specifically, we introduce two-stage stochastic and distributionally robust SIGDPs where both stages have SIGDPs, i.e., disjunctions of sets of logical equations and infinite number of algebraic inequalities along with logical propositions. A two-stage distributionally robust SIGDP (DR-SIGDP) is defined as follows. Let (x, Γ) and (z, Γ^2) be the first-stage (here-and-now) and second-stage (wait-and-see) decisions, respectively. The second-stage decisions are made after realization of uncertain data and they depend on the first-stage decisions. To address uncertainty, we consider a random variable defined over a finite sample space $\{\omega_1, \ldots, \omega_m\}$ with probability distribution $\{p_\omega\}_\omega$ belonging to a set of distributions $\mathfrak{P}$, referred to as ambiguity set. Akin to DR programs [42], a decision maker minimizes sum of the first-stage cost function $f_1(x)$ and expected second-stage cost for worst-case probability distribution from the ambiguity set $\mathfrak{P}$. Note that this ambiguity set can be defined in different ways such as using bounds on moments or an ϵ-ball around a reference distribution. In this paper, we consider DR-SIGDP with a general ambiguity set that is defined by (2) where

$$f(x) = f_1(x) + \max_{P \in \mathfrak{P}} \left(\sum_{\omega=\omega_1}^{\omega_m} p_\omega \mathcal{Q}_\omega(x) \right)$$

and

$$\mathcal{Q}_\omega(x) := \min \Bigg\{ f^2_\omega(z_\omega) : q_2(z_\omega, x) \leq 0, \quad \Omega_\omega(\Gamma^{2,\omega}) = \text{True}, \tag{5a}$$

$$\bigvee_{h \in H^\omega_k} \begin{bmatrix} \Gamma^{2,\omega}_{hk} \\ \tilde{g}^{\omega,j}_{hk}(z_\omega, x) \leq 0, \ j \in J \end{bmatrix} \quad \text{for } k \in K, \tag{5b}$$

$$z_\omega \in \mathbb{R}^n_+, \ \Gamma^{2,\omega}_{hk} \in \{\text{True, False}\}, \ h \in H^\omega_k, k \in K \Bigg\}. \tag{5c}$$

In case $\mathfrak{P}$ is singleton, i.e., the probability distribution is known, the DR-SIGDP reduces to stochastic (risk-neutral) SIGDP. Other special cases include DR-SILP, DR-SIDLP, DR-SIGDLP, DR-SIMICP, DR-SIGDCP, and many more. As per our knowledge, even stochastic

and DR-SILP have not been studied in the literature. We provide Lagrangean relaxation based approaches for deriving lower bounds to solve stochastic and DR-SIGDP, which are also applicable for solving DR-SILP, DR-SIDLP, DR-SIGDLP, DR-SIGDCP, and DR-SIMBCP. Table 1 lists modeling frameworks discussed in this paper and "*" denotes a new framework introduced and studied in this paper for the first time in the literature.

Modeling Framework	Abbr.	Semi-Infinite (SI)	Distributionally Robust (DR)
Linear Program	LP	SILP [19, 25]	DR-SILP*
Disjunctive Linear Program	DLP	SIDLP [8]	DR-SIDLP*
Mixed Binary Linear Program	MBLP	SIMBLP [8, 45]	DR-SIMBLP*
Mixed Integer Convex Program	MICP	SIMICP*	DR-SIMICP*
Mixed Binary Convex Program	MBCP	SIMBCP*	DR-SIMBCP*
Generalized Disjunctive Program	GDP	SIGDP*	DR-SIGDP*
Generalized DLP	GDLP	SIGDLP*	DR-SIGDLP*
GD Convex Program	GDCP	SIGDCP*	DR-SIGDCP*

Table 1: Abbreviation of Algebraic Modeling Frameworks

2. Literature Review

In this section, we review literature related to disjunctive programs, generalized disjunctive programs, semi-infinite linear/convex/mixed-binary/disjunctive programs, and two-stage distributionally robust programs.

Finite DPs and GDPs. The DPs and GDPs provide alternate frameworks to algebraically model problems with discrete and logical variables. Balas [4, 5] studied convex hull of feasible region of DLPs, $conv(\mathcal{D})$, by introducing a tight extended linear programming formulation in higher dimensional space and cutting planes in the original space. Blair [15] and Jeroslow [30, 31] also developed cutting planes for DLPs with a facial property. In contrast, Beaumont [12] presented a branch-and-bound based exact algorithm for directly solving DLPs. Since DLPs subsume mixed binary programs, researchers thereafter shifted their focus toward utilizing the foregoing cutting planes to derive disjunctive cuts for mixed binary linear and convex programs [6, 16, 33, 46]. In another direction, Raman and Grossman [40] presented a big-M reformulation and a solution approach for GDLPs. Thereafter, Lee and Grossman [36] introduced a tight extended formulation that is big-M free but has additional variables and constraints. They also presented a branch-and-bound algorithm (an extension of algorithm in [12] for DLPs) for GDCPs. Extensions of approaches known for mixed integer nonlinear programs such as outer-approximation have also been developed for solving GDLPs and GDCPs [48]; refer to [26] for a survey on GDPs.

Semi-Infinite Programs. In early 1960s, Charnes et al. [19, 20] introduced duality theory for SILPs, and in 1970s, Gustafson and Kortanek [27] and Gribik [25] developed numerical methods and a central cutting-plane approach, respectively, for solving SILPs. Thereafter, in late 1980s, Anderson and Lewis [2] presented an extension of the simplex algorithm for SILPs. Less than a decade ago, Basu et al. [10] extended a well-known projection algorithm, Fourier-Motzkin elimination, for solving LPs to solve SILPs. In a separate paper, they performed sensitivity analysis for SILPs as well [11]. As mentioned before, Charnes et al. [21] and Ben-Tal et al. [13] derived optimality conditions for SICPs, and Kortanek and No [34] extended the central cutting plane algorithm for SILPs to SICPs. Readers can refer to survey by Hettich and Kortanek [28] and Shapiro [43] for details and other results for SILPs and SICPs.

Recently, Bansal [8] amalgamated the concepts of "semi-infiniteness" with DLPs by introducing SIDLPs (3). The author presented a tight extended formulation (a SILP equivalent) and valid inequalities for SIDPs. Using these results, he developed partial convex hulls, a sequential convexification approach, and a family of cutting planes for SIMBLPs. Additionally, a branch-and-cut based exact algorithm for solving semi-infinite binary programs (SIBLPs) is presented. He also showcased how the approaches for SIBLPs, SIMBLPs, and SIDPs can be utilized for solving disjunctive convex programs, SIMBLPs with discrete variables, robust knapsack problem, distributionally robust chance-constrained programs, and various combinatorial optimization problems, in particular implicit hitting set problem [37]. As per our knowledge, no other study in the literature considers semi-infinite programs with discrete variables, except a Reformulation-Linearization Technique based SILP equivalent for SIMBLP by Sherali and Adams [45]. Since SIMBLPs belong to the family of SIDLPs, the foregoing reformulation is also derived in [8] as a special case of results for SIDLP.

Two-Stage Distributionally Robust Programs. Dantzig [22] presented the concept of two-stage decision-making in the presence of uncertainty, thereby leading to two-stage stochastic linear programs. In late 1950s, Scarf [42] introduced the concept of min-max stochastic (now popularly known as distributionally robust) optimization where the ambiguity set is defined using linear constraints on the first two moments of the distribution [14, 24, 39, 42]. Since then, both frameworks have evolved significantly with the development of decomposition methods, cutting-plane algorithms, reformulation techniques, and Lagrangean relaxation based approaches for two-stage stochastic and distributionally robust binary/integer/disjunctive/linear programs (refer to [9, 35] and references therein). As mentioned before, in the context of semi-infiniteness, even two-stage stochastic semi-infinite linear programs have not been studied in the literature.

3. Semi-Infinite Generalized Disjunctive Programs

In this section, we present reformulations for SIGDPs (2) with linear and convex functions.

3.1 SIGDP with Linear Inequalities

We define SIGDLP as follows:

$$\min \left\{ c^\top x : Dx \geq f,\ (2c) - (2d) \text{ hold}, \bigvee_{h \in H_k} \begin{bmatrix} \Gamma_{hk} \\ (a^j_{hk})^\top x \geq b^j_{hk},\ j \in J \end{bmatrix} \text{ for } k \in K \right\}, \tag{6}$$

where $a^j_{hk} \in \mathbb{R}^n$ and $b^j_{hk} \in \mathbb{R}$, and $|J|$ is infinite. For finite J, Raman and Grossmann [40] introduced a big-M reformulation using additional binary variables, $\lambda_{hk} \in \{0, 1\}$ for $h \in H_k$ and $k \in K$. Lee and Grossmann [36] presented another reformulation without big-M parameters for GDLPs, but using more variables, i.e., $(\mu_{hk}, \lambda_{hk}) \in \mathbb{R}_+ \times \{0, 1\}$ for $h \in H_k$ and $k \in K$, and constraints. In this paper, we introduce a new reformulation for SIGDLP without big-M parameters and with only $y_{hk} \in \{0, 1\}$ for $h \in H_k$ and $k \in K$ additional variables, which is applicable for finite index set J as well.

<u>New Reformulation for GDLPs and SIGDLPs.</u> Assuming that $(a^j_{hk})^\top x \geq 0$ for all x, the new reformulation for (6) is given by

$$\min \Big\{ c^\top x : Dx \geq f, \tag{7a}$$

$$(a^j_{hk})^\top x \geq b^j_{hk} y_{hk},\ j \in J, h \in H_k, k \in K, \tag{7b}$$

$$\sum_{h \in H_k} y_{hk} = 1,\ k \in K, \tag{7c}$$

$$Ey \geq \tau, \tag{7d}$$

$$x \in \mathbb{R}^n_+, y_{hk} \in \{0, 1\},\ h \in H_k, k \in K \Big\}, \tag{7e}$$

where $y_{hk} = 1$ implies that Boolean variable $\Gamma_{hk} =$ True and constraints (7d) represent logic propositions (2c). Note that the assumption $(a^j_{hk})^\top x \geq 0$ is automatically satisfied when $a^j_{hk} \geq 0$, which occurs in many applications. Moreover, reformulation (7) is a semi-infinite mixed binary linear program (SIMBLP). Therefore, exact algorithm, cutting planes, partial convex hulls, and extended formulations presented in [8] for SIMBLPs are applicable for solving this new reformulation as well. We present Lagrangian relaxations of (7) such that this relaxation can be decomposed into subproblems with only one binary variable y_{hk} in each subproblem.

We first introduce copies of x variables by adding constraints $x = x_{hk}$ for $h \in H_k$ and $k \in K$. Thereafter, we relax these constraints along with constraints (7c) and (7d) using dual/Lagrangian multipliers $\overline{\eta}$, $\hat{\eta}$, and $\tilde{\eta}$.

$$\min \sum_{k \in K} \overline{\eta}_k \left(\sum_{h \in H_k} y_{hk} - 1 \right) + \hat{\eta}^T (Ey - \tau) + \sum_{k \in K} \sum_{h \in H_k} \left(\bar{c}^\top x_{hk} + \tilde{\eta}_{hk} (x_{hk} - x) \right) \tag{8a}$$

$$\text{s.t.} \quad Dx_{hk} \geq f, \quad h \in H_k, k \in K, \tag{8b}$$

$$(a^j_{hk})^\top x_{hk} \geq b^j_{hk} y_{hk}, \quad j \in J, h \in H_k, k \in K, \tag{8c}$$

$$x_{hk} \in \mathbb{R}^n_+, y_{hk} \in \{0, 1\}, \; h \in H_k, k \in K. \tag{8d}$$

Since x is unbounded, we pick $\tilde{\eta}_{hk}$ such that $\sum_{k \in K} \sum_{h \in H_k} \tilde{\eta}_{hk} = 0$. Let the objective function (8a) in simplified form be written as

$$\mathcal{L}(\alpha) = \min \sum_{k \in K} \sum_{h \in H_k} \left(\alpha^1_{hk} y_{hk} + \alpha^2_{hk} x_{hk} + \alpha^0 \right).$$

Then the relaxation (8) is separable. In other words, $\mathcal{L}(\alpha) = \sum_{k \in K} \sum_{h \in H_k} \mathcal{L}_{hk}(\alpha)$ where

$$\begin{aligned} \mathcal{L}_{hk}(\alpha) = \min \; & \alpha^1_{hk} y_{hk} + \alpha^2_{hk} x_{hk} + \alpha^0 \\ \text{s.t.} \quad & Dx_{hk} \geq f, \qquad (a^j_{hk})^\top x_{hk} \geq b^j_{hk} y_{hk}, \quad j \in J, \\ & x_{hk} \in \mathbb{R}^n_+, y_{hk} \in \{0, 1\} \end{aligned}$$

is a subproblem. There are two ways to solve this subproblem. First method is a branch-and-bound approach where we solve SILP relaxation with $y_{hk} \in [0, 1]$ at the root node, and create branches (if needed) by setting $y_{hk} = 0$ for left node and $y_{hk} = 1$ for right node. At each node, we have an SILP that can be solved using algorithms of [10, 25, 34]. Another way to solve the subproblems is deriving its SIDLP/SILP equivalent. Note that $y_{hk} \in \{0, 1\}$ is equivalent to $(y_{hk} = 0) \vee (y_{hk} = 1)$. Therefore, the feasible region of each subproblem is equivalent to an SIDLP.

Remark 1. *The best Lagrangian relaxation is derived by computing* $\alpha^* = \arg\max \mathcal{L}(\alpha)$.

Next, we present an SIDLP equivalent of SIGDLP, so that results in [8] for the former program are applicable to solve the latter program as well. Specifically, an equivalent SIDLP formulation for SIGDLP is given by:

$$\min \Bigg\{ c^\top x : Dx \geq f, \quad \sum_{h \in H_k} y_{hk} = 1, \quad Ey \geq \tau, \tag{9a}$$

$$\bigvee_{h \in H_k} \begin{bmatrix} y_{hk} = 1 \\ (a^j_{hk})^\top x \geq b^j_{hk}, \ j \in J \end{bmatrix} \text{ for } k \in K, \tag{9b}$$

$$x \in \mathbb{R}^n_+, y_{hk} \in [0,1] \text{ for } h \in H_k, k \in K \Bigg\}. \tag{9c}$$

Proposition 1. *Formulations* (6) *and* (9) *are equivalent.*

Proof. The arguments are same as arguments of Proposition 2.1.1 in [41] for finite J. □

We introduce the following SILP relaxation for the feasible region of (9) by convexifying each of the $|K|$ disjunctive constraints in (9b) using semi-infinite linear programming equivalent for SIDLPs (Theorem 1 in [8]).

$$Dx \geq f, \quad \sum_{h \in H_k} y_{hk} = 1, \quad Ey \geq \tau, \tag{10a}$$

$$(a^j_{hk})^\top \xi^1_{hk} \geq b^j_{hk} \xi^{0,1}_{hk}, \quad j \in J, h \in H_k, k \in K, \tag{10b}$$

$$\sum_{h \in H_k} \xi^1_{hk} = x, \quad \sum_{h \in H_k} \xi^{0,1}_{hk} = 1, \quad k \in K, \tag{10c}$$

$$e^\top_{hk} \xi^2_{hk} = \xi^{0,2}_{hk}, \quad h \in H_k, k \in K, \tag{10d}$$

$$\sum_{h \in H_k} \xi^2_{hk} = y_k, \quad \sum_{h \in H_k} \xi^{0,2}_{hk} = 1, \quad k \in K, \tag{10e}$$

$$x \in \mathbb{R}^n_+, y_{hk} \in [0,1] \text{ for } h \in H_k, k \in K, \tag{10f}$$

$$\xi^1_{hk} \in \mathbb{R}^n_+, \xi^2_{hk} \in \mathbb{R}^{|H_k|}_+, \xi^{0,1}_{hk} \in \mathbb{R}_+, \xi^{0,2}_{hk} \in \mathbb{R}_+, \ h \in H_k, k \in K, \tag{10g}$$

where $y_k = [y_{1k}, y_{2k}, \ldots, y_{|H_k|k}]^\top$ and e_{hk} is an $|H_k|$ dimensional unit vector with h^{th} component as 1. This relaxation can be solved using algorithms of [10, 25, 34] for SILPs.

3.2 SIGDP with Convex Functions

We first introduce semi-infinite disjunctive convex programs (SIDCP) and introduce a semi-infinite convex programming equivalent for it. This generalizes Theorem 1 of [8] for SIDLPs and Theorem 1 of Ceria and Soares [18] for finite disjunctive convex programs. Thereafter, we introduce SIDCP equivalent and a semi-infinite mixed binary convex programming reformulation without big-M parameters for SIGDCPs.

Theorem 1. *Let*

$$\mathcal{D}_C := \left\{ \mathbf{x} \in \mathbb{R}^n : \bigvee_{h \in H} \mathcal{D}_C^h \right\} \text{ where } \mathcal{D}_C^h := \left\{ \mathbf{x} \in \mathbb{R}_+^n : g_j^h(x) \leq 0 \text{ for all } j \in J \right\}$$

for $h \in H$, is a nonempty and bounded semi-infinite convex program defined using infinite number of differentiable and continuous convex functions $g_j^h(x)$. The convex hull of $\mathcal{D}_C$ is equivalent to the projection of the following extended formulation onto the x-space:

$$x = \sum_{h \in H} \zeta^h, \tag{11a}$$

$$\overline{g}_j^h(\zeta^h, \zeta_0^h) \leq 0, \quad h \in H \text{ and } j \in J, \tag{11b}$$

$$\sum_{h \in H} \zeta_0^h = 1, \tag{11c}$$

$$(\zeta^h, \zeta_0^h) \geq 0, \quad h \in H, \tag{11d}$$

where convex function $\overline{g}_j^h(\zeta^h, \zeta_0^h) = \zeta_0^h g_j^h(\zeta^h/\zeta_0^h)$ for $\zeta_0^h > 0$ and 0, otherwise.

Proof. Let $\left(\widehat{x}, \widehat{\zeta}^1, \ldots, \widehat{\zeta}^{|H|}, \widehat{\zeta}_0^1, \ldots, \widehat{\zeta}_0^{|H|}\right)$ be a point that satisfies constraints of (11). Define $\widehat{H} := \{h \in H : \widehat{\zeta}_0^h = 0\}$. Because of constraint (11c), we know that $\widehat{H} \subset H$. As a result, we get

$$\widehat{x} = \sum_{h \in H \setminus \widehat{H}} \widehat{\zeta}^h, \qquad g_j^h(\widehat{\zeta}^h/\widehat{\zeta}_0^h) \leq 0,\ h \in H \setminus \widehat{H}, j \in J, \qquad \sum_{h \in H \setminus \widehat{H}} \widehat{\zeta}_0^h = 1, \quad (\widehat{\zeta}^h, \widehat{\zeta}_0^h) \geq 0,\ h \in H.$$

For $h \in H \setminus \widehat{H}$, $\widehat{\nu}^h = \widehat{\zeta}^h/\widehat{\zeta}_0^h \in \mathcal{D}_C^h$ as $g_j^h(\widehat{\nu}) \leq 0$ for all $j \in J$. Substituting $\widehat{\zeta}^h = \widehat{\nu}^h \widehat{\zeta}_0^h$ in the above constraints gives

$$\widehat{x} = \sum_{h \in H \setminus \widehat{H}} \widehat{\nu}^h \widehat{\zeta}_0^h, \qquad \sum_{h \in H \setminus \widehat{H}} \widehat{\zeta}_0^h = 1, \quad \widehat{\zeta}_0^h \geq 0,\ h \in H \setminus \widehat{H}. \tag{12}$$

Constraints (12) shows that $\widehat{x}$ is equal to convex combination of $\{\widehat{\nu}^h\}_{h \in H \setminus \widehat{H}}$ points. Since $\widehat{\nu}^h \in \mathcal{D}_C^h$ for $h \in H \setminus \widehat{H}$, point $\widehat{x} \in conv(\cup_{h \in H \setminus \widehat{H}} \mathcal{D}_C^h) \subseteq conv(\mathcal{D}_C)$. This implies that $conv(\mathcal{D}_C)$ subsumes projection of (11) onto the x-space.

Now, consider a point $\widehat{x} \in conv(\mathcal{D}_C) = conv(\cup_h \mathcal{D}_C^h)$. We write $\widehat{x}$ as a convex combination of points $\overline{x}_h \in \mathcal{D}_C^h$, $h \in H$:

$$\widehat{x} = \sum_{h \in H} \lambda_h \overline{x}_h, \quad \sum_{h \in H} \lambda_h = 1, \quad \lambda_h \geq 0, h \in H.$$

Let $\widehat{\zeta}^h = \lambda_h \overline{x}_h$ and $\widehat{\zeta}_0^h = \lambda_h$. Clearly, point $\left(\widehat{x}, \{\widehat{\zeta}^h, \widehat{\zeta}_0^h\}_{h \in H}\right)$ satisfies constraints of (11). This implies that for every point in $conv(\mathcal{D}_C)$, there exists a point in the extended space defined by (11). In other words, projection of (11) onto the x-space subsumes $conv(\mathcal{D}_C)$. $\square$

SIDCP Equivalent for SIGDCP. The GDP (2) with infinite index set J and convex (differentiable and continuous) functions $f(x)$, $q(x)$, and $g_{hk}^j(x)$ is referred to as SIGDCP. Its SIDCP equivalent is given by

$$\min \Bigg\{ f(x) : q(x) \leq 0, \quad \sum_{h \in H_k} y_{hk} = 1, \quad Ey \geq \tau, \tag{13a}$$

$$\bigvee_{h \in H_k} \begin{bmatrix} y_{hk} = 1 \\ g_{hk}^j(x) \leq 0, \ j \in J \end{bmatrix} \text{ for } k \in K, \tag{13b}$$

$$x \in \mathbb{R}_+^n, y_{hk} \in [0,1] \text{ for } h \in H_k, k \in K \Bigg\}. \tag{13c}$$

Using Theorem 1, we can easily derive SICP relaxation of the feasible region of (13) which can then be solved using algorithm of [34].

New Reformulation Without Big-M Parameters for SIGDCP. We re-define SIGDCP as follows:

$$\min \left\{ f(x) : \ q(x) \leq 0, (2c) - (2d) \text{ hold}, \bigvee_{h \in H_k} \begin{bmatrix} \Gamma_{hk} \\ \widehat{g}_{hk}^j(x) + b_{hk}^j \leq 0, \ j \in J \end{bmatrix} \text{ for } k \in K \right\}, \tag{14}$$

where $\widehat{g}_{hk}^j(x)$ is a convex function. Assuming that $\widehat{g}_{hk}^j(x) \leq 0$ for all x, the SIGDCP (14) can be reformulated as the following SIMBCP:

$$\min \Bigg\{ f(x) : q(x) \leq 0, \quad \sum_{h \in H_k} y_{hk} = 1, \ k \in K, \quad Ey \geq \tau, \tag{15a}$$

$$\widehat{g}_{hk}^j(x) + b_{hk}^j y_{hk} \leq 0, \ j \in J, h \in H_k, k \in K, \tag{15b}$$

$$x \in \mathbb{R}_+^n, y_{hk} \in \{0,1\}, \ h \in H_k, k \in K \Bigg\}. \tag{15c}$$

Observe that this reformulation is a special case of SIMBCP (4) that has not been studied in the literature, except when $g_j(x)$ is an affine function [8, 45]. In Section 4, we present results for SIMBCP (4) that are applicable for solving (15) as well.

Remark 2. *Problem* (14) *can also be defined using constraints* $\widehat{g}_{hk}^j(x) \geq b_{hk}^j$ *where* $\widehat{g}_{hk}^j(x)$ *is a concave function.*

Remark 3 (Extensions of SIGDPs with Nonconvex Functions)**.** *For SIGDPs where functions $f(x)$, $q(x)$, and/or $g^j_{hk}(x)$ are nonconvex, we can utilize convex underestimators of these functions (refer to Chapter 2 of [47]) to get SIGDCP relaxations.*

4. Semi-Infinite Mixed Binary Convex Programs

Recall that we define SIMBCPs as follows: $\min\{c^\top x : x \in \mathcal{C}_0,\ x_i \in \{0,1\} \text{ for } i = 1,\dots,p\}$ where

$$\mathcal{C}_0 := \{x \in \mathbb{R}^n_+ : g_j(x) \le 0 \text{ for } j \in J\},$$

is a convex set. Without any loss of generality, we assume that $\{g_j(x) \le 0 : j \in J\}$ subsumes constraints $0 \le x_i \le 1$ for all i, and convex objective function $f(x)$ expressed as a constraints $f(x) \le \theta$. Let $\mathcal{C} := \mathcal{C}_0 \cap \{x_i \in \{0,1\} : i = 1,\dots,p\}$ be a semi-infinite mixed binary convex set where $|J|$ is infinite.

Theorem 2. *The convex hull of $\mathcal{C}$ is equivalent to the projection of the following convex set (in the extended space) onto the original x-space:*

$$\overline{g}_j(\xi^h, \xi^h_0) \le 0, \quad \textit{for all } j \in J \textit{ and } h \in \{1,\dots,|\mathbb{K}|\}, \tag{16a}$$

$$\xi^h_k = 0, k \in K_1, \quad \xi^h_k = \xi^h_{k,0}, k \in K_2, \quad \textit{for all } h \in \{1,\dots,|\mathbb{K}|\}, (K_1,K_2) \in \mathbb{K}, \tag{16b}$$

$$x_i = \sum_{h=1}^{|\mathbb{K}|} \xi^h_i, \quad i \in \{1,\dots,n\}, \tag{16c}$$

$$\sum_{h=1}^{|\mathbb{K}|} \xi^h_{k,0} = 1,\ k \in K_2 \textit{ where } (K_1,K_2) \in \mathbb{K}, \qquad (\xi^h, \xi^h_0) \ge 0 \textit{ for all } h \in \{1,\dots,|\mathbb{K}|\}. \tag{16d}$$

where $\mathbb{K} := \{(K_1,K_2) : |K_1 \cup K_2| = p, K_1 \cap K_2 = \emptyset, K_1, K_2 \subseteq \{1,\dots,p\}\}$ and for $j \in J$, convex function $\overline{g}_j(\xi^h, \xi^h_0) = \xi^h_0 g_j(\xi^h/\xi^h_0)$ for $\xi^h_0 > 0$ and 0, otherwise.

Proof. The disjunctive normal form of set $\mathcal{C}$ is given by

$$\bigvee_{(K_1,K_2)\in\mathbb{K}} \begin{pmatrix} g_j(x) \le 0, j \in J \\ x_k = 0\ \forall k \in K_1, \quad x_k = 1\ \forall k \in K_2 \end{pmatrix}.$$

Using Theorem 1 with $H \equiv \mathbb{K}$, we derive a convex programming equivalent of $\mathcal{C}$ in an extended space which is same as (16). □

Next, we present a hierarchy of relaxations between the continuous and convex hull representation of the set $\mathcal{C}$, i.e., $\mathcal{C}_0$ and $conv(\mathcal{C})$, respectively. A relaxation of $\mathcal{C}$ can be defined

as

$$\mathcal{C}_t^R := \mathcal{C}_0 \cap \{x_i = 0 \vee x_i = 1, \text{ for } i = 1, \ldots, t\} \tag{17}$$

where $t \in \{1, \ldots, p\}$ and binary restrictions on $\{x_i\}_{i=t+1}^p$ variables have been relaxed. Clearly,

$$\mathcal{C} = \mathcal{C}_p^R \subseteq \mathcal{C}_{p-1}^R \subseteq \ldots \subseteq \mathcal{C}_1^R \subseteq \mathcal{C}_0^R = \mathcal{C}_0,$$

and therefore, $conv(\mathcal{C}) \subseteq conv(\mathcal{C}_{t+1}^R) \subseteq conv(\mathcal{C}_t^R) \subseteq \mathcal{C}_0$ for any $t \in \{1, \ldots, p-1\}$. Now, set $\mathcal{C}_t^R$ in the disjunctive normal form is written as

$$\bigvee_{(K_1,K_2)\in\mathbb{K}_t} \begin{pmatrix} g_j(x) \leq 0, j \in J \\ x_k = 0 \ \forall k \in K_1, \quad x_k = 1 \ \forall k \in K_2 \end{pmatrix},$$

where $\mathbb{K}_t := \{(K_1, K_2) : |K_1 \cup K_2| = t, K_1 \cap K_2 = \emptyset, K_1, K_2 \subseteq \{1, \ldots, t\}\}$. Note that $\mathbb{K}_p = \mathbb{K}$.

Corollary 1 (Hierarchical Relaxations)**.** *For any $t \in \{1, \ldots, p\}$, the convex hull of $\mathcal{C}_t^R$ is equivalent to the projection of convex set* (16) *where $\mathbb{K}$ is replaced by $\mathbb{K}_t$, onto the original x-space.*

Proposition 2. *Let $\mathcal{C}_{1,m}^R := \mathcal{C}_0 \cap \{x_m = 0 \vee x_m = 1\}$ for a given $m \in \{1, \ldots, p\}$. Again, $conv(\mathcal{C}) \subseteq conv(\mathcal{C}_{1,m}^R)$, and the semi-infinite convex programming equivalent of $conv(\mathcal{C}_{1,m}^R)$ is given by* (16) *where $\mathbb{K}$ is replaced by $\mathbb{K}_{1,m} := \{(K_1, K_2) : |K_1 \cup K_2| = 1, K_1 \cap K_2 = \emptyset, K_1, K_2 \subseteq \{m\}\}$, i.e., $\mathbb{K}_{1,m} = \{(\emptyset, \{m\}), (\{m\}, \emptyset)\}$.*

Theorem 3 (Sequential Convexification)**.** *For $t \in \{1, \ldots, p-1\}$,*

$$conv(\mathcal{C}_{t+1}^R) = Proj_x \left(conv \left(\mathcal{M}_t \cap \{x_{t+1} = 0 \vee x_{t+1} = 1\}\right)\right),$$

where convex set $\mathcal{M}_t$ is given by (16) *where $\mathbb{K}$ is replaced by $\mathbb{K}_t$.*

Proof. It is easy to observe that $|\mathbb{K}_{t+1}| = 2|\mathbb{K}_t|$ as

$$\mathbb{K}_{t+1} = \big\{(K_1 \cup \{t+1\}, K_2), (K_1, K_2 \cup \{t+1\}) \text{ such that } (K_1, K_2) \in \mathbb{K}_t\big\}, \tag{18}$$

Also, $\mathcal{M}_t \cap \{(x_{t+1} = 0) \vee (x_{t+1} = 1)\} = (\mathcal{M}_t \cap \{x_{t+1} = 0\}) \vee (\mathcal{M}_t \cap \{x_{t+1} = 1\})$, and we obtain its following convex programming equivalent in the extended space using Theorem 1.

$$\hat{g}_j(\sigma^{h,1}, \sigma_0^{h,1}, \bar{\sigma}_0^1) \leq 0, \quad j \in J, h \in \{1, \ldots, |\mathbb{K}_t|\} \tag{19a}$$

$$\hat{g}_j(\sigma^{h,2}, \sigma_0^{h,2}, \bar{\sigma}_0^2) \leq 0, \quad j \in J, h \in \{1, \ldots, |\mathbb{K}_t|\} \tag{19b}$$

$$\sigma_k^{h,1} = 0, k \in K_1, \quad \sigma_k^{h,1} - \sigma_{k,0}^{h,1} = 0, k \in K_2, \quad h \in \{1, \ldots, |\mathbb{K}_t|\}, (K_1, K_2) \in \mathbb{K}_t, \tag{19c}$$

$$\sigma_k^{h,2} = 0, k \in K_1, \quad \sigma_k^{h,2} - \sigma_{k,0}^{h,2} = 0, k \in K_2, \quad h \in \{1, \ldots, |\mathbb{K}_t|\}, (K_1, K_2) \in \mathbb{K}_t, \tag{19d}$$

$$\sigma^{x,1} - \sum_{h=1}^{|\mathbb{K}_t|} \sigma^{h,1} = 0, \quad \sigma^{x,2} - \sum_{h=1}^{|\mathbb{K}_t|} \sigma^{h,2} = 0, \quad x = \sigma^{x,1} + \sigma^{x,2}, \tag{19e}$$

$$\xi^h = \sigma^{h,1} + \sigma^{h,2}, \quad \xi_0^h = \sigma_0^{h,1} + \sigma_0^{h,2}, \quad h \in \{1, \ldots, |\mathbb{K}_t|\}, \tag{19f}$$

$$\sum_{h=1}^{|\mathbb{K}_t|} \sigma_{k,0}^{h,1} = \bar{\sigma}_{k,0}^1, \qquad \sum_{h=1}^{|\mathbb{K}_t|} \sigma_{k,0}^{h,2} = \bar{\sigma}_{k,0}^2, \qquad \bar{\sigma}_{k,0}^1 + \bar{\sigma}_{k,0}^2 = 1, k \in K_2, \text{ where } (K_1, K_2) \in \mathbb{K}_t, \tag{19g}$$

$$\sigma_{t+1}^{x,1} = 0\bar{\sigma}_0^{x,1}, \quad \sigma_{t+1}^{x,2} = \bar{\sigma}_0^{x,2}, \quad \bar{\sigma}_0^{x,1} + \bar{\sigma}_0^{x,2} = 1, \qquad (x, \sigma, \xi, \bar{\sigma}_0) \geq 0, \tag{19h}$$

where for $e \in \{1, 2\}$, function $\hat{g}_j(\sigma^{h,e}, \sigma_0^{h,e}, \bar{\sigma}_0^e) = \bar{\sigma}_0^e \overline{g}_j \left(\sigma^{h,e}/\bar{\sigma}_0^e, \sigma_0^{h,e}/\bar{\sigma}_0^e\right) = \sigma_0^{h,e} g_j \left(\sigma^{h,e}/\sigma_0^{h,e}\right)$ when $\sigma_0^{h,e}, \bar{\sigma}_0^e > 0$, and 0 otherwise.

We simplify constraints (19e) and (19g) to get

$$x = \sum_{h=1}^{|\mathbb{K}_t|} \left(\sigma^{h,1} + \sigma^{h,2}\right), \qquad \sum_{h=1}^{|\mathbb{K}_t|} \left(\sigma_{k,0}^{h,1} + \sigma_{k,0}^{h,2}\right) = 1, \quad k \in K_2, \text{ where } (K_1, K_2) \in \mathbb{K}_t, \tag{20}$$

which along with $\bar{\sigma}_0^{x,1} + \bar{\sigma}_0^{x,2} = 1$ are equivalent to (16c) and (16d) where $\mathbb{K}$ is replaced with $\mathbb{K}_{t+1}$. Likewise, constraints (19c), (19d), and first two equations of (19h) together can be written as (16b) where $\mathbb{K}$ is replaced by $\mathbb{K}_{t+1}$. Finally, (19a) and (19b) lead to (16a) with $\mathbb{K}_{t+1}$. As a consequence, (19a)-(19h) reduces to $\mathcal{M}_{t+1}$ whose projection onto x-space gives $conv(\mathcal{C}_{t+1}^R)$ because of Corollary 1. This completes the proof. □

Special Cases of SIMBCPs – SIMBLPs and MBCPs. Sherali and Adams presented hierarchy of relaxations between continuous and convex hull representation of $\mathcal{C}$ with affine $g_j(x)$ functions having finite [44] and infinite [45] index set J; Refer to Theorem 1 of [44] and Remark 2 of [45]. They also provide convex hull representations in the extended space for these special cases of SIMBCPs. We observe that by setting $\xi_i^h = x_i \xi_0^h$ for $i = 1, \ldots, n$, and

$$\xi_0^h = \left(\prod_{k \in K_2} x_k\right) \left(\prod_{k \in K_1} (1 - x_k)\right),$$

in (16), Theorem 2 and Corollary 1 of our paper reduce to Theorem 1 of [45] and Theorem 1 of [44], respectively. In [46], sequential convexification approach of [6] for mixed binary linear programs is extended to mixed binary convex programs (SIMBCP with finite J). Theorem 3 of our paper further generalizes these results for infinite J.

4.1 Extensions of SIMBCPs

We also consider the following extensions of SIMBCPs.

(*i*) *Semi-Infinite Mixed-Discrete Convex Programs.* This class of problem is defined over set:

$$\left\{\mathbf{x} \in \mathbb{R}_+^n : \left(g_j(x) \leq 0, j \in J\right) \wedge \left(\bigvee_{l=1}^{m} (x_i = \alpha_l), i = 1, \ldots, p\right)\right\}, \tag{21}$$

where $\alpha_i \in \mathbb{R}_+$ and m is finite. Moreover, the defining constraints include $x_i \leq 1$ for $i = p+1, \ldots, n$. Using auxiliary binary variables $\nu_{il} = 1$, if $x_i = \alpha_l$ and 0 otherwise, the disjunctive constraints in set (21) reduces to:

$$\left(x_i = \sum_{l=1}^{m} \alpha_l \nu_{il} \text{ for } i = 1, \ldots, p; \quad \sum_{l=1}^{m} \nu_{il} = 1 \text{ for } i = 1, \ldots, p\right), \tag{22}$$

and hence an optimization problem defined over set (21) with the objective of minimizing a convex function belongs to the class of SIMBCPs.

(*ii*) *Robust 0-1 Convex Problem.* A robust 0-1 convex problem is defined as follows:

$$\min_{\mathbf{x}} \left\{ \max_{u \in \mathcal{U}} f(x, u) : x \in \mathcal{X} \cap \{0, 1\}^n \right\},$$

where $f(x, u)$ is a convex function for any fixed $u \in \mathcal{U}$ and $\mathcal{U}$ is an uncertainty set defined using polytope or an ellipsoidal. This problem is equivalent to

$$\min_{\mathbf{x} \in \mathcal{X} \cap \{0,1\}^n} \left\{ \theta : f(x, u) \leq \theta \text{ for } u \in \mathcal{U} \right\},$$

which is SIMBCP when $\mathcal{X}$ is a polyhedron.

(*iii*) *Distributionally Robust Convex Programs* is defined as

$$\min_{\mathbf{x} \in \mathcal{X}} \left\{ \Theta : \sum_{\omega = \omega_1}^{\omega_m} p_\omega H(x, \omega) \leq \Theta, \quad \{p_\omega\}_{\omega \in \Omega} \in \mathfrak{P} \right\},$$

which is a semi-infinite program. Assuming that $H(x,.)$ is a convex function, it is equivalent to SIDCP, semi-infinite mixed-discrete convex program, or SIMBCP when $\mathcal{X}$ is a disjunctive set, mixed-discrete set, or mixed binary set, respectively.

5. Stochastic and Distributionally Robust SIGDPs

In this section, we study stochastic and distributionally robust SIGDPs. Specifically, we amalgamate reformulations presented in the preceding sections with a Lagrangian relaxation based approach (also referred to as dual decomposition method) to derive lower bounds for these programs. Recall that stochastic SIGDP is defined as

$$\min \left\{ f_1(x) + \sum_{\omega \in \Xi} p_\omega \mathcal{Q}_\omega(x) : x \in \mathcal{X} \right\}, \tag{23}$$

where $\mathcal{Q}_\omega(x)$ is given by (5), $\Xi := \{\omega_1, \ldots, \omega_m\}$, and $\mathcal{X} := \{x : q(x) \leq 0, (2b) - (2d) \text{ hold}\}$. The deterministic equivalent formulation of (23) is a large-scale SIGDP:

$$\min\ f_1(x) + \sum_{\omega \in \Xi} p_\omega f^2_\omega(z_\omega) \tag{24a}$$

$$\text{s.t. } q_2(z_\omega, x) \leq 0, \quad \Omega_\omega(\Gamma^{2,\omega}) = \text{True}, \quad \text{for all } \omega \in \Xi, \tag{24b}$$

$$\bigvee_{h \in H^\omega_k} \begin{bmatrix} \Gamma^{2,\omega}_{hk} \\ \tilde{g}^{\omega,j}_{hk}(z_\omega, x) \leq 0,\ j \in J \end{bmatrix}, \quad \text{for all } k \in K, \omega \in \Xi, \tag{24c}$$

$$x \in \mathcal{X}, z_\omega \in \mathbb{R}^n_+,\ \Gamma^{2,\omega}_{hk} \in \{\text{True, False}\},\ h \in H^\omega_k, k \in K, \omega \in \Xi. \tag{24d}$$

Akin to usual dual decomposition method [17], we create replicas of x variables for each scenario using constraints $x = x_\omega$ for $\omega \in \Xi$, and then we relax these constraints using dual/Lagrangian multipliers $\mu = \{\mu_\omega\}_{\omega \in \Omega}$. This results in the following Lagrangian relaxation of (23): $\mathcal{L}(\mu) = \sum_{\omega \in \Xi} p_\omega \mathcal{L}_\omega(\mu)$ where

$$\mathcal{L}_\omega(\mu) = \min\ f_1(x_\omega) + f^2_\omega(z_\omega) + \mu^\top_\omega x_\omega \tag{25a}$$

$$\text{s.t. } q_2(z_\omega, x_\omega) \leq 0, \quad \Omega_\omega(\Gamma^{2,\omega}) = \text{True}, \tag{25b}$$

$$\bigvee_{h \in H^\omega_k} \begin{bmatrix} \Gamma^{2,\omega}_{hk} \\ \tilde{g}^{\omega,j}_{hk}(z_\omega, x_\omega) \leq 0,\ j \in J \end{bmatrix}, \quad \text{for all } k \in K, \tag{25c}$$

$$x_\omega \in \mathcal{X}, z_\omega \in \mathbb{R}^n_+,\ \Gamma^{2,\omega}_{hk} \in \{\text{True, False}\},\ h \in H^\omega_k, k \in K, \tag{25d}$$

for $\omega \in \Xi$ and $\sum_{\omega \in \Xi} \mu_\omega = 0$ to avoid unbounded solution value because of unbounded x.

Observe that for any given μ and $\omega \in \Xi$, formulation (25) is an SIGDP in (x_ω, z_ω) space. Below we present three different reformulations of (25) where for the sake of simplicity, we drop subscript/superscript ω.

(a) *Semi-Infinite Disjunctive Programming Reformulation.*

$$\min\ f_1(x) + f^2(z) + \mu^\top x \tag{26a}$$

$$\text{s.t. } q(x) \le 0, \quad \sum_{h \in H_k} y^1_{hk} = 1, \quad E^1 y^1 \ge \tau^1, \tag{26b}$$

$$q_2(z, x) \le 0, \quad \sum_{h \in H_k} y^2_{hk} = 1, \quad E^2 y^2 \ge \tau^2, \tag{26c}$$

$$\bigvee_{h \in H_k} \begin{bmatrix} y^1_{hk} = 1 \\ g^j_{hk}(x) \le 0,\ j \in J \end{bmatrix} \quad \text{for all } k \in K, \tag{26d}$$

$$\bigvee_{h \in H_k} \begin{bmatrix} y^2_{hk} = 1 \\ \tilde{g}^j_{hk}(z, x) \le 0,\ j \in J \end{bmatrix} \quad \text{for all } k \in K, \tag{26e}$$

$$x \in \mathbb{R}^n_+, y^1_{hk} \in [0, 1], z \in \mathbb{R}^n_+,\ y^2_{hk} \in [0, 1],\ h \in H_k, k \in K, \tag{26f}$$

where $E^1 y^1 \ge \tau^1$ and $E^2 y^2 \ge \tau^2$ correspond to $\Omega(\Gamma) = \text{True}$ and $\Omega_\omega(\Gamma^{2,\omega}) = \text{True}$, respectively.

(b) *Semi-Infinite Mixed Binary Reformulation with Big-M Parameters.*

$$\begin{aligned}
&\min\ f_1(x) + f^2(z) + \mu^\top x \\
&\text{s.t. } q(x) \le 0, \quad \sum_{h \in H_k} y^1_{hk} = 1, \quad E^1 y^1 \ge \tau^1, \\
&\qquad q_2(z, x) \le 0, \quad \sum_{h \in H_k} y^2_{hk} = 1, \quad E^2 y^2 \ge \tau^2, \\
&\qquad g^j_{hk}(x) \le M(1 - y^1_{hk}),\ j \in J, h \in H_k, k \in K, \\
&\qquad \tilde{g}^j_{hk}(z, x) \le M(1 - y^2_{hk}),\ j \in J, h \in H_k, k \in K, \\
&\qquad x \in \mathbb{R}^n_+, y^1_{hk} \in \{0, 1\}, z \in \mathbb{R}^n_+,\ y^2_{hk} \in \{0, 1\},\ h \in H_k, k \in K,
\end{aligned}$$

where M is a large number.

(c) *Semi-Infinite Mixed Binary Reformulation without Big-M Parameters.* Assuming $g^j_{hk}(x) = \widehat{g}^{j,1}_{hk}(x) + b^{j,1}_{hk}$ and $\tilde{g}^j_{hk}(x) = \widehat{g}^{j,2}_{hk}(z, x) + b^{j,2}_{hk}$ such that $\widehat{g}^{j,1}_{hk}(x) \le 0$ for all x and $\widehat{g}^{j,2}_{hk}(z, x) \le 0$

for all (z, x), we get the following reformulation without any big-M parameter:

$$\min\ f_1(x) + f^2(z) + \mu^\top x \tag{27a}$$

$$\text{s.t. } q(x) \le 0, \quad \sum_{h \in H_k} y^1_{hk} = 1, \quad E^1 y^1 \ge e^1, \tag{27b}$$

$$q_2(z, x) \le 0, \quad \sum_{h \in H_k} y^2_{hk} = 1, \quad E^2 y^2 \ge e^2, \tag{27c}$$

$$\widehat{g}^{j,1}_{hk}(x) + b^{j,1}_{hk} y^1_{hk} \le 0,\ j \in J, h \in H_k, k \in K, \tag{27d}$$

$$\widehat{g}^{j,2}_{hk}(z, x) + b^{j,2}_{hk} y^2_{hk} \le 0,\ j \in J, h \in H_k, k \in K, \tag{27e}$$

$$x \in \mathbb{R}^n_+, y^1_{hk} \in \{0, 1\}, z \in \mathbb{R}^n_+,\ y^2_{hk} \in \{0, 1\},\ h \in H_k, k \in K. \tag{27f}$$

Remark 4. *The best Lagrangian relaxation lower bound is computed by solving optimization problem* $\max\{\mathcal{L}(\mu) : \sum_{\omega \in \Xi} \mu_\omega = 0\}$.

Remark 5. *As we know that SIGDPs subsume SIGDLPs, SIGDCPs, SIDLPs, SIDCPs, SIMBLPs, and SIMBCPs. The aforementioned reformulations and the dual decomposition approach are applicable for solving stochastic variants of all these special cases of SIGDPs.*

Distributionally Robust SIGDPs. Next, we study DR-SIGDPs defined for a general ambiguity set $\mathfrak{P}$:

$$\kappa = \min_{x \in \mathcal{X}} \left\{ f_1(x) + \max_{\{p_\omega\} \in \mathfrak{P}} \sum_{\omega \in \Xi} p_\omega \mathcal{Q}_\omega(x) \right\} = \min_{x \in \mathcal{X}} \max_{\{p_\omega\} \in \mathfrak{P}} \sum_{\omega \in \Xi} p_\omega \left(f_1(x) + \mathcal{Q}_\omega(x) \right) \tag{28}$$

$$\ge \max_{\{p_\omega\} \in \mathfrak{P}} \min_{x \in \mathcal{X}} \sum_{\omega \in \Xi} p_\omega \left(f_1(x) + \mathcal{Q}_\omega(x) \right), \tag{29}$$

where the last inequality is because of the well-known max-min inequality. Observe that for any given $\{p_\omega\}_\omega \in \mathfrak{P}$, the inner minimization problem of (29) is equivalent to (23). Therefore, solving

$$\max_{P \in \mathfrak{P}} \sum_{\omega \in \Xi} p_\omega \mathcal{L}_\omega(\mu)$$

for appropriate μ, returns a lower bound for the optimal objective value of DR-SIGDP, i.e., κ.

6. Conclusion

We introduced semi-infinite generalized disjunctive programs (SIGDPs) with linear and/or convex functions, and reformulated them as semi-infinite mixed-binary/disjunctive linear or convex, respectively, programs. This led to new reformulations even for finite GDPs that

does not require big-M parameters and have lesser number of variables, in comparison to other reformulations known in the literature. We presented tight extended formulation for semi-infinite disjunctive convex programs. For semi-infinite mixed binary convex programs (SIMBCPs), we derived semi-infinite convex programming equivalent in higher dimensional space, hierarchy of relaxations, and a sequential convexification approach. We demonstrated that these results are also applicable for solving semi-infinite mixed discrete convex programs, robust 0-1 convex programs and distributionally robust convex programs. Lastly, we presented Lagrangian relaxation based approaches for deriving lower bounds to solve two-stage stochastic and distributionally robust SIGDPs with general ambiguity set. Since SIGDPs subsume various modeling frameworks such as SIMBCP, semi-infinite disjunctive convex program, and many more, the foregoing approaches have a wide range of applications.

Acknowledgment. This work is supported by National Science Foundation Grant CMMI-1824897, which is gratefully acknowledged.

References

[1] Haar A. *Uber lineare ungleichungen.* A M. Kir. Ferencz Jozsef Tudomanyegyetem Baratai Egyesuletenek, 1924.

[2] E. J. Anderson and A. S. Lewis. An extension of the simplex algorithm for semi-infinite linear programming. *Mathematical Programming*, 44(1):247–269, May 1989.

[3] Egon Balas. Intersection cuts from disjunctive constraints. *MSRR No. 330, Carnegie-Mellon University*, 1974.

[4] Egon Balas. Disjunctive Programming. In P.L. Hammer, E. L. Johnson, and B. H. Korte, editors, *Annals of Discrete Mathematics*, 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*, pages 3–51. Elsevier, 1979.

[5] Egon Balas. Disjunctive programming: Properties of the convex hull of feasible points. *Discrete Applied Mathematics*, 89(1–3):3–44, December 1998.

[6] Egon Balas, Sebastián Ceria, and Gérard Cornuéjols. A lift-and-project cutting plane algorithm for mixed 0–1 programs. *Mathematical Programming*, 58(1-3):295–324, January 1993.

[7] Egon Balas, Joseph M. Tama, and Jørgen Tind. Sequential convexification in reverse convex and disjunctive programming. *Mathematical Programming*, 44(1-3):337–350, May 1989.

[8] Manish Bansal. Semi-infinite mixed binary and disjunctive programs: Applications to implicit hitting set and set-covering with infinite demand points. *Technical Report (under review)*, pages 1–27, Oct. 2022.

[9] Manish Bansal and Yingqiu Zhang. Scenario-based cuts for structured two-stage stochastic and distributionally robust p-order conic mixed integer programs. *Journal of Global Optimization*, January 2021.

[10] Amitabh Basu, Kipp Martin, and Christopher Thomas Ryan. Projection: A Unified Approach to Semi-Infinite Linear Programs and Duality in Convex Programming. *Mathematics of Operations Research*, 40(1):146–170, February 2015.

[11] Amitabh Basu, Kipp Martin, and Christopher Thomas Ryan. Strong duality and sensitivity analysis in semi-infinite linear programming. *Mathematical Programming*, 161(1):451–485, January 2017.

[12] Nicholas Beaumont. An algorithm for disjunctive programs. *European Journal of Operational Research*, 48(3):362–371, October 1990.

[13] A. Ben-Tal, L. Kerzner, and S. Zlobec. Optimality conditions for convex semi-infinite programming problems. *Naval Research Logistics Quarterly*, 27(3):413–435, 1980.

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

[15] Charles E. Blair. Facial disjunctive programs and sequence of cutting-planes. *Discrete Applied Mathematics*, 2(3):173–179, September 1980.

[16] Pierre Bonami. Lift-and-Project Cuts for Mixed Integer Convex Programs. In Oktay Günlük and Gerhard J. Woeginger, editors, *Integer Programming and Combinatoral Optimization*, number 6655 in Lecture Notes in Computer Science, pages 52–64. Springer Berlin Heidelberg, 2011.

[17] Claus C. Caroe and Rudiger Schultz. Dual decomposition in stochastic integer programming. *Operations Research Letters*, 24(1):37–45, February 1999.

[18] Sebastián Ceria and João Soares. Convex programming for disjunctive convex optimization. *Mathematical Programming*, 86(3):595–614, December 1999.

[19] A. Charnes, W. W. Cooper, and K. Kortanek. Duality in Semi-Infinite Programs and Some Works of Haar and Carathéodory. *Management Science*, 9(2):209–228, January 1963.

[20] A. Charnes, W. W. Cooper, and K. Kortanek. On Representations of Semi-Infinite Programs which Have No Duality Gaps. *Management Science*, 12(1):113–121, September 1965.

[21] A. Charnes, W. W. Cooper, and K. O. Kortanek. On the theory of semi-infinite programming and a generalization of the kuhn-tucker saddle point theorem for arbitrary convex functions. *Naval Research Logistics Quarterly*, 16(1):41–52, 1969.

[22] George B. Dantzig. Linear Programming under Uncertainty. *Management Science*, 1(3-4):197–206, April 1955.

[23] Sanjeeb Dash, Oktay Günlük, and Robert Hildebrand. Binary extended formulations of polyhedral mixed-integer sets. *Mathematical Programming*, 170(1):207–236, July 2018.

[24] J. Dupacová. The minimax approach to stochastic programming and an illustrative application. *Stochastics*, 20:73–88, 1987.

[25] P. R. Gribik. A central-cutting-plane algorithm for semi-infinite programming problems. In R. Hettich, editor, *Semi-Infinite Programming*, Lecture Notes in Control and Information Sciences, pages 66–82, Berlin, Heidelberg, 1979. Springer.

[26] Ignacio E. Grossmann and Juan P. Ruiz. Generalized Disjunctive Programming: A Framework for Formulation and Alternative Algorithms for MINLP Optimization. In Jon Lee and Sven Leyffer, editors, *Mixed Integer Nonlinear Programming*, The IMA Volumes in Mathematics and its Applications, pages 93–115, New York, NY, 2012. Springer.

[27] S. A. Gustafson and K. O. Kortanek. Numerical treatment of a class of semi-infinite programming problems. *Naval Research Logistics Quarterly*, 20(3):477–504, 1973.

[28] R. Hettich and K. O. Kortanek. Semi-Infinite Programming: Theory, Methods, and Applications. *SIAM Review*, 35(3):380–429, 1993.

[29] Jennifer R. Jackson and Ignacio E. Grossmann. A disjunctive programming approach for the optimal design of reactive distillation columns. *Computers & Chemical Engineering*, 25(11):1661–1673, November 2001.

[30] R. Jeroslow. A cutting-plane game for facial disjunctive programs. *SIAM Journal on Control and Optimization*, 18(3):264–281, May 1980.

[31] R. G. Jeroslow. Cutting-Plane Theory: Disjunctive Methods. In E. L. Johnson P.L. Hammer, B. H. Korte and G. L. Nemhauser, editor, *Annals of Discrete Mathematics*, volume 1 of *Studies in Integer Programming*, pages 293–330. Elsevier, 1977.

[32] Michael Jünger, Thomas M. Liebling, Denis Naddef, George L. Nemhauser, William R. Pulleyblank, Gerhard Reinelt, Giovanni Rinaldi, and Laurence A. Wolsey. *50 Years of Integer Programming 1958-2008: From the Early Years to the State-of-the-Art*. Springer Science & Business Media, November 2009.

[33] Mustafa R. Kilinc, Jeff Linderoth, and James Luedtke. Lift-and-project cuts for convex mixed integer nonlinear programs. *Mathematical Programming Computation*, 9(4):499–526, December 2017.

[34] K. Kortanek and H. No. A Central Cutting Plane Algorithm for Convex Semi-Infinite Programming Problems. *SIAM Journal on Optimization*, 3(4):901–918, November 1993.

[35] Gilbert Laporte and Francois V. Louveaux. The integer L-shaped method for stochastic integer programs with complete recourse. *Operations Research Letters*, 13(3):133–142, April 1993.

[36] Sangbum Lee and Ignacio E. Grossmann. New algorithms for nonlinear generalized disjunctive programming. *Computers & Chemical Engineering*, 24(9):2125–2141, October 2000.

[37] Erick Moreno-Centeno and Richard M. Karp. The Implicit Hitting Set Approach to Solve Combinatorial Optimization Problems with an Application to Multigenome Alignment. *Operations Research*, 61(2):453–468, April 2013.

[38] Juan M. Novas, Juan Ignacio Ramello, and María Analía Rodríguez. Generalized disjunctive programming models for the truck loading problem: A case study from the non-alcoholic beverages industry. *Transportation Research Part E: Logistics and Transportation Review*, 140:1366–5545, August 2020.

[39] A. Prékopa. *Stochastic Programming*. Kluwer Academic, 1995.

[40] R. Raman and I. E. Grossmann. Modelling and computational techniques for logic based integer programming. *Computers & Chemical Engineering*, 18(7):563–578, July 1994.

[41] Nicolas Sawaya and Ignacio Grossmann. A hierarchy of relaxations for linear generalized disjunctive programming. *European Journal of Operational Research*, 216(1):70–82, January 2012.

[42] H. Scarf. A min-max solution of an inventory problem. In *Studies in the Mathematical Theory of Inventory and Production*, chapter 12, pages 201–209. 1958.

[43] Alexander Shapiro. Semi-infinite programming, duality, discretization and optimality conditions. *Optimization*, 58(2):133–161, February 2009.

[44] H. Sherali and W. Adams. A hierarchy of relaxations between the continuous and convex hull representations for zero-one programming problems. *SIAM Journal on Discrete Mathematics*, 3(3):411–430, August 1990.

[45] Hanif D. Sherali and Warren P. Adams. A Reformulation-Linearization Technique (RLT) for semi-infinite and convex programs under mixed 0-1 and general discrete restrictions. *Discrete Applied Mathematics*, 157(6):1319–1333, March 2009.

[46] Robert A. Stubbs and Sanjay Mehrotra. A branch-and-cut method for 0-1 mixed convex programming. *Mathematical Programming*, 86(3):515–532, December 1999.

[47] Mohit Tawarmalani and Nikolaos V. Sahinidis. *Convexification and Global Optimization in Continuous and Mixed-Integer Nonlinear Programming*, volume 65 of *Nonconvex Optimization and Its Applications*. Springer US, Boston, MA, 2002.

[48] Metin Türkay and Ignacio E. Grossmann. Logic-based MINLP algorithms for the optimal synthesis of process networks. *Computers & Chemical Engineering*, 20(8):959–978, August 1996.