Small independent branching formulations for unions of $V$-polyhedra

Joey Huchette
Operations Research Center, M.I.T., huchette@mit.edu, http://www.mit.edu/~huchette

Juan Pablo Vielma
Sloan School of Management, M.I.T., jvielma@mit.edu, http://web.mit.edu/jvielma/www

We present a framework for constructing strong mixed-integer formulations for logical disjunctive constraints. Our approach is a generalization of the logarithmically-sized formulations of Vielma and Nemhauser for SOS2 constraints [50], and we offer a complete characterization of its expressive power. We apply the framework to a variety of disjunctive constraints, producing novel small and strong formulations for outer approximations of multilinear terms, generalizations of special ordered sets, piecewise linear functions over a variety of domains, and obstacle avoidance constraints.

90C11
Primary: Integer Programming; secondary: Piecewise Linear, Polyhedra

History:

1. Introduction

A central modeling primitive in mathematical optimization is the disjunctive constraint: any feasible solution must satisfy at least one of some fixed, finite collection of alternatives. This type of constraint is general enough to capture structures as diverse as boolean satisfiability, complementarity constraints, special-ordered sets, and (bounded) integrality. The special case of polyhedral disjunctive constraints corresponds to the form

$$x \in \bigcup_{i=1}^{d} P^{i},$$

where we have that each $P^{i} \subseteq \mathbb{R}^{n}$ is a polyhedron (assumed for the moment to be bounded). In this work, we will focus on $V$-polyhedra; that is, we have the description of the $P^{i}$ in terms of their extreme points $\text{ext}(P^{i})$.

We are particularly interested in the case where constraint (1) is primitive, and so we are interested in modeling it in a generic, composable way. In particular, if (1) is embedded in a larger, more complex optimization problem

$$\min_{(x,y) \in Q \cap (1)} f(x,y),$$

we hope for a mathematical description sufficiently structured such that we may use more advanced algorithmic approaches, beyond naïve enumeration, to solve (2). In particular, $Q$ could be described by any number of different types of constraints: linear inequalities, conic constraints, integrality conditions, or additional disjunctive constraints. In this context, it is well known that merely constructing the convex hull $\text{Conv}(\bigcup_{i=1}^{d} P^{i})$ is not sufficient for solving (2); we will need a formulation for $\bigcup_{i=1}^{d} P^{i}$ directly.

Mixed-integer programming (MIP) has emerged as an incredibly expressive modeling methodology, with advanced computational methods capable of solving many problems of practical interest, often at very large scale [12, 26]. Constraint (1) with polyhedral sets $P^{i}$ is particularly well-suited for a mixed-integer programming approach. Indeed, standard formulations for (1) were presented in [25], and are ideal, or as strong as possible with respect to their continuous linear programming

\[^{1}\text{See [46, Example 2] for a simple example.}\]
relaxations (see Section 4.1 for a formal definition). However, this formulation requires introducing $d$ auxiliary binary variables, which may be impractically large, especially in the context of the larger problem (2).

However, it is sometimes possible to construct ideal formulations with considerably fewer auxiliary variables. In particular, a string of recent work [3, 34, 46, 50] has presented ideal formulations for certain highly structured constraints such as SOS2 [8] with only $O(\log(d))$ auxiliary binary variables and additional constraints (excluding variable bounds). Moreover, these formulations have proven practically useful, and indeed the most performant by a significant margin for a large swath of instances of the problem classes for which they have been applied [48]. However, the construction procedure for these formulations is complex and ad-hoc, hindering the construction, analysis, and implementation of such formulations for new constraints.

Of particular interest are the formulations of Vielma and Nemhauser for SOS2 [50], which fit into the independent branching (IB) scheme framework. That is, they find some (particularly structured) polyhedra $Q_{1,j}$ and $Q_{2,j}$ such that (1) can be rewritten as

$$\bigcup_{i=1}^{d} P_i = \bigcap_{j=1}^{t} (Q_{1,j} \cup Q_{2,j}). \quad (3)$$

This representation represents the disjunctive constraint in term of a series of simple choices between two alternatives. Given such a representation, it is straightforward to construct a simple, small, and ideal formulation for (1) by formulating each of the $t$ alternatives separately, and then combining them. Furthermore, as the polyhedra $P_i$ are $V$-polyhedra, the construction of the independent branching scheme-based formulation is purely combinatorial, based on the vertices that are shared between the different polyhedra $P_i$. As we will see, we can therefore approach formulating (1) combinatorially, by studying the shared structure amongst the vertices.

In this work we generalize and provide a systematic study of the applicability and limitations of the independent branching approach. The contributions of this work can be categorized in the following way.

1. We generalize the notion of independent branching schemes to allow for multiple alternatives, and provide an exact characterization of when there exists any independent branching representation for (1), in terms of the graphical representation of the shared vertices between the polyhedra $P_i$. In particular:
   (a) We demonstrate that the widely-used cardinality constraints cannot be expressed by any independent branching schemes with few alternatives. We argue that this negative result provides theoretical justification for the practical observation that both MIP formulations and simple constraint branching schemes seems incapable of modeling cardinality constraints effectively.
   (b) We show that arbitrary piecewise linear functions in the plane can be modeled with at most three alternatives, and provide a polynomial-time verifiable condition for representability with two alternatives.
   (c) We argue that nonconvex polygonal set avoidance constraints are always representable with two alternatives.

2. We provide an exact characterization of when there exists a two-alternative independent branching representation for (1) of size $t$, in terms of the classical biclique covering problem. This relation allows the algorithmic construction of small independent branching formulations for (1).

3. We apply our framework to a variety of constraints of the form (1) to give an indication of the expressive power of the IB scheme approach. In particular:
   (a) We construct explicit, small (i.e. logarithmic in $d$), ideal formulations for generalizations of the special ordered sets of Beale and Tomlin [8], piecewise linear functions over arbitrary 2-dimensional grid triangulations, and outer-approximating discretizations of multilinear terms.
   (b) We provide matching lower bounds for these constructions, showing that they are asymptotically optimal with respect to the size of any possible MIP formulation.
(c) We present our novel ideal formulations for discretizations of multilinear terms as a generalization of the popular logarithmically-sized formulation of Misener et al. [32, 33] for nonconvex quadratic optimization, which we show is not ideal in general.

2. Preliminaries: Definitions, notation, and nomenclature

A (bounded) \( V \)-polyhedra (or polyhedra in \( V \)-form) is a set \( P \subseteq \mathbb{R}^n \) that can be expressed as

\[
P = \text{Conv}(V) \overset{\text{def}}{=} \left\{ \sum_{v \in V} \lambda_v v : \lambda \in \Delta^V \right\}
\]

for some finite set of vectors \( V \subseteq \mathbb{R}^n \), where \( \Delta^V \overset{\text{def}}{=} \{ \lambda \in \mathbb{R}^V_+ : \sum_{v \in V} \lambda_v = 1 \} \) is the standard simplex. According to the celebrated Minkowski-Weyl Theorem (e.g. [14, Corollary 3.14]), any polyhedral disjunctive constraint (1) can be expressed as the union of \( V \)-polyhedra in terms of their extreme points \( \text{ext}(P^i) \).

When the disjunctive constraint (1) is a union of \( V \)-polyhedra, it suffices to consider only the combinatorial structure of the extreme points of the polyhedra \( P^i \). To see why, consider \( J = \bigcup_{i=1}^d \text{ext}(P^i) \) as the ground set and take \( S = \{ \text{ext}(P^i) \}_{i=1}^d \subseteq 2^J \) as the collection of extreme points for each of the polyhedra. We can then define a corresponding disjunctive constraint that is purely combinatorial on the sets \( S \).

**Definition 1.** A combinatorial disjunctive constraint (CDC) induced by sets \( S \) is

\[
\lambda \in \text{CDC}(S) \overset{\text{def}}{=} \bigcup_{S \in S} Q(S),
\]

where \( Q(S) \overset{\text{def}}{=} \{ \lambda \in \Delta^J : \lambda_{J,S} \leq 0 \} \) is the face that \( S \subseteq J \) induces on the standard simplex.

Our goal for this work is to construct formulations for these combinatorial disjunctive constraints directly, since it is straightforward to construct a corresponding formulation for (1) as

\[
\left\{ \sum_{v \in J} \lambda_v v : \lambda \in \text{CDC}(S) \right\}.
\]

We emphasize here that formulation (4) allows us to divorce the problem-specific data (i.e. the values \( v \in J \)) from the underlying combinatorial structure encapsulated in \( \text{CDC}(S) \). As such, we can construct a single, strong formulation for a given structure \( \text{CDC}(S) \) and change the problem data at-will, while still using the same underlying formulation.

We also note that if we wish to model the case where the polyhedra \( P^i \) are unbounded, a result of Jeroslows and Lowe [25] [45, Proposition 11.2] tells us that we may only construct a (binary) MIP formulation for (1) if the recession cones \( \text{rec}(P^i) \) coincide for each \( P^i \). In the case this condition is met, we may formulate (1) with

\[
\left\{ \sum_{v \in J} \lambda_v v + \sum_{r \in R} \mu_r r : \lambda \in \text{CDC}(S), \mu \in \mathbb{R}_+^R \right\},
\]

where \( R \) is the shared set of extreme rays for each of the \( P^i \). Therefore, we will restrict our attention to the case where each of the \( P^i \) are bounded, as formulating the unbounded case is a straightforward extension.

In the remainder of the paper, we will make the following assumptions on \( S \) that are without loss of generality.

\footnote{For the moment we are assuming that the \( P^i \) are bounded; the unbounded case is more delicate, as we will discuss shortly.}
Assumption 1. We assume the following about $\mathcal{S}$:

- $\mathcal{S}$ is irredundant: there do not exist distinct $S,T \in \mathcal{S}$ such that $S \subseteq T$.
- $\mathcal{S}$ covers the ground set: $\bigcup_{S \in \mathcal{S}} S = J$.

We will say that a set $S \subseteq J$ is a feasible set with respect to $\text{CDC}(\mathcal{S})$ if $Q(S) \subseteq \text{CDC}(\mathcal{S})$ (equivalently, if $S \subseteq T$ for some $T \in \mathcal{S}$) and that it is an infeasible set otherwise.

3. Motivating examples As mentioned, we can represent any polyhedral disjunctive constraint (1) as the union of $\mathcal{V}$-polyhedra. However, there are many disjunctive constraints for which the $\mathcal{V}$-form of (1) is especially natural; we now present some as running examples that we will return to throughout.

One motif that will appear repeatedly will be the graph of a continuous piecewise linear function. That is, given some bounded domain $\Omega \subset \mathbb{R}^n$ and some polyhedral partition $\bigcup_{i=1}^d P^i = \Omega$ (i.e. the relative interiors do not overlap), we are interested in modeling a continuous function $f : \Omega \rightarrow \mathbb{R}$ such that $x \in P^i \Rightarrow f(x) = a_i^T x + b_i$ for some appropriate $a_i \in \mathbb{R}^n$ and $b_i \in \mathbb{R}$. In order to model the graph $\text{gr}(f;\Omega) \equiv \{(x,f(x)) : x \in \Omega\}$, we can construct a formulation for $\text{CDC}(\mathcal{S})$ and express

$$\text{gr}(f) = \left\{ \left( \sum_{v \in J} \lambda_v v, \sum_{v \in J} \lambda_v f(v) \right) : \lambda \in \text{CDC}(\mathcal{S}) \right\}, \quad (6)$$

where we will use the notation $\text{gr}(f) \equiv \text{gr}(f;\Omega)$ when $\Omega$ is clear from context.\footnote{Formally, $\text{relint}(P^i) \cap \text{relint}(P^j) = \emptyset$ for each $i \neq j$.}

3.1. SOS2 Consider a univariate (nonconvex) piecewise linear function $f$ characterized by $N$ breakpoints $x^1 < x^2 < \ldots < x^N$. We may model the graph of this function via (6), where $\mathcal{S} = \{\{\tau, \tau + 1\} : \tau \in [N-1]\}$. This is the SOS2 constraint of Beale and Tomlin [8], and can be equivalent to requiring that, given an ordering of the ground set $J = [N] \equiv \{1, \ldots, N\}$, at most two components of $\lambda$ may be nonzero, and they must be consecutive in this ordering.

3.2. SOS$k$ A generalization of the special ordered sets considers the case where at most $k$ consecutive components of $\lambda$ may be nonzero at once. In particular, if $J = [N]$, we have $\mathcal{S} = \{\{\tau, \tau + 1, \ldots, \tau + k - 1\} : \tau \in [N-k+1]\}$. This constraint may arise, for example, in chemical process scheduling problems, where an activated machine may only be on for $k$ consecutive time units and must produce a fixed quantity during that period [18, 27].

3.3. Cardinality constraints An extremely common constraint is optimization is the cardinality constraint of degree $\ell$, where at most $\ell$ components of $\lambda$ may be nonzero. This corresponds to $\mathcal{S} = \{I \subseteq J : |I| = \ell\}$. A particularly compelling application of the cardinality constraint is in portfolio optimization [10, 11, 13, 47], where it is often advantageous to limit the number of investments to some fixed number $\ell$ to minimize transaction costs, or to allow differentiation from the performance of the market as a whole.

3.4. Grid triangulations of the plane Consider a rectangular region in the plane $\Omega = [0,M] \times [0,N]$, and the regular grid points $J = \{0, \ldots, M\} \times \{0, \ldots, N\}$. A grid triangulation $\mathcal{S}$ of $\Omega$ is then a set $\mathcal{S}$ where:

- Each $S \in \mathcal{S}$ is a triangle: $|S| = 3$
- $\mathcal{S}$ partitions $\Omega$; $\bigcup_{S \in \mathcal{S}} \text{Conv}(S) = \Omega$ and $\text{relint}(\text{Conv}(S)) \cap \text{relint}(\text{Conv}(S')) = \emptyset$ for each $S, S' \in \mathcal{S}$

We note that the results to follow can potentially be extended to certain discontinuous piecewise linear functions by working instead with the epigraph of $f$; we point the interested reader to [48, 49] for further discussion.
S is on a regular grid: $S \subset J$ for each $S \in \mathcal{S}$, and $\|v - w\|_2 \leq 1$ for each $v, w \in S$.

Grid triangulations are often used to model bivariate, non-separable, piecewise linear functions [48, 50] using (6). As concrete examples, consider Figure 1, where we depict three different triangulations with $M = N = 2$.

![Figure 1](image)

**Figure 1.** Three grid triangulations of $\Omega = [0, 2] \times [0, 2]$: the Union Jack (J1) [44] (Left), the K1 [28] (Center), and a more idiosyncratic construction (Right).

### 3.5. Obstacle avoidance

Consider an unmanned aerial vehicle (UAV) which you would like to navigate through an area with fixed obstacles. At any given time, you wish to impose the constraint that the location of the vehicle $x \in \mathbb{R}^2$ must lie in some (nonconvex) region $\Omega \subset \mathbb{R}^2$, which is the plane, less any obstacles in the area. MIP formulations of this constraint has received renewed interest as a useful primitive for path planning [9, 16, 38, 31].

We may model $x \in \Omega$ by partitioning the region $\Omega$ with polyhedra such that $\Omega = \bigcup_{i=1}^d P_i$. Traditional approaches to modeling constraints (1) of this form use a linear inequality description for each of the polyhedra $P_i$ and construct a corresponding big-M formulation [37, 38], which will not be ideal in general. In the $\mathcal{V}$-polyhedra framework, we will instead take the $P_i$ as $\mathcal{V}$-polyhedra that partition $\Omega$, and be able to construct small, ideal formulations.

### 3.6. General piecewise linear functions in the plane

Consider again a (potentially nonconvex) region $\Omega \subset \mathbb{R}^2$. We would like to model a (also potentially nonconvex) piecewise linear function $f$ with domain over $\Omega$. Take $\{P_i\}_{i=1}^d$ as the set of pieces of the domain, and the corresponding ground set as $J = \bigcup_{i=1}^d \text{ext}(P_i)$ and sets as $S = \{\text{ext}(P_i)\}_{i=1}^d$. We may then model the piecewise linear function via the graph representation (6). Note that this representation is a generalization of both the SOS2 and grid triangulation construction mentioned above.

### 3.7. Discretizations of multilinear terms

Consider a multilinear function $f(x_1, \ldots, x_\eta) = \prod_{i=1}^\eta x_i$, defined over some box domain $\Omega = \{l, u\} \subset \mathbb{R}^\eta$. This function appears often in optimization models [19], but is nonconvex, and often leads to problems which are difficult to solve to global optimality in practice [4, 39, 51]. As a result, computational techniques will often “relax” the graph of the function $\text{gr}(f)$ with a convex outer approximation, which is easier to optimize over [42].

For the bilinear case ($\eta = 2$), the well-known McCormick envelope [30] describes the convex hull of $\text{gr}(f)$. Although traditionally stated in an inequality description, we may equivalently describe the convex hull via its four extreme points, which are readily available in closed form. For higher-dimensional multilinear terms, the convex hull has $2^\eta$ extreme points, and can be constructed in a similar manner.

---

5 Note that these relaxations are still useful in the context of global optimization, coupled with algorithmic techniques such as spatial branch-and-bound.
Misener et al. [32, 33] propose a computational technique for optimizing problems with bilinear terms where, instead of modeling the graph over a single region $\Omega = [l, u] \subset \mathbb{R}^2$, they discretize the region in a regular fashion and apply the McCormick envelope to each subregion. They model this constraint as a union of polyhedra, where each subregion enjoys a tighter relaxation of the bilinear term. Additionally, they propose a logarithmically-sized formulation for the union. However, it is not ideal (see Appendix A), it only applies for bilinear terms ($\eta = 2$), and it is specialized for a particular type of discretization (namely, only discretizing along one component $x_1$).

For a more general setting, we have that the extreme points of the convex hull of the graph $\text{Conv}(\text{gr}(f))$ are given by $\{(x, f(x)) : x \in \text{ext}(\Omega)\}$ [40], where it is easy to see that $\text{ext}(\Omega) = \prod_{i=1}^{\eta} [l_i, u_i]$. Consider a grid imposed on $[l, u] \subset \mathbb{R}^\eta$; that is, along each component $i \in [\eta]$, we partition $[l_i, u_i]$ along the points $l_i = h_0^i < h_1^i < \cdots < h_{d_i}^i = u_i$. This yields $\prod_{i=1}^{\eta} d_i$ subregions; denote them by $R \overset{\text{def}}{=} \{\prod_{i=1}^{\eta} [h_{l_i}^i, h_{u_i}^i] : \ell \in \prod_{i=1}^{\eta} [d_i]\}$.

We will take the polyhedra partition $\Omega$ as $P^R \overset{\text{def}}{=} \text{Conv}(\text{gr}(f; R))$, the sets as $S = \{\text{ext}(P^R)\}_{R \in \mathbb{R}^\eta}$, and the ground set as $J = \bigcup\{S \in S\}$. In particular, we have that $J = \prod_{i=1}^{\eta} [h_0^i, h_1^i, \ldots, h_{d_i}^i]$. By taking the $\mathbb{V}$-polyhedron perspective, we can apply our framework to construct small, ideal formulations for tight outer approximations of higher-order multilinear functions over general grid discretizations.

4. MIP formulations for combinatorial disjunctive constraints Using standard MIP formulation techniques, we now present formulations for CDC($S$) for comparison with our approach. In particular, we will argue that the framework we will present later can lead to formulations that are smaller (in terms of the number of auxiliary variables), and enjoy other favorable properties we enumerate in Section 5.

4.1. MIP formulations: Definitions, size, and strength Formally, we say that a (binary) MIP formulation $F$ for a constraint $x \in Q \subseteq \mathbb{R}^{n_1}$ is the composition of linear inequalities (the linear programming (LP) relaxation, or just relaxation in the context of this work)

$$R \overset{\text{def}}{=} \{(x, y, z) \in [l^x, u^x] \times [l^y, u^y] \times [0, 1]^{n_3} : Ax + By + Cz \leq d\}$$

with (binary) integrality conditions

$$F \overset{\text{def}}{=} R \cap (\mathbb{R}^{n_1} \times \mathbb{R}^{n_2} \times \{0, 1\}^{n_3})$$

(7)

such that $\text{Proj}_x(F) = Q$.

Throughout, we will be interested in ways of understanding both the strength of a given formulation, as well as in quantifying the size or complexity of a formulation. We say that a formulation is ideal if each extreme point of the relaxation naturally satisfies the integrality conditions, i.e. $\text{Proj}_x(\text{ext}(R)) \subseteq \{0, 1\}^{n_3}$. The choice of name is apt, as this is the strongest possible MIP formulation we can expect.

As a measure of the complexity of the formulation, we count the number of auxiliary continuous variables $y$ and continuous binary variables $z$ used by the formulation, as well as the number of inequalities in our description of $R$. We ignore the size of $x$, since this is intrinsic to the constraint we wish to model. We will say that a formulation is extended if there are auxiliary continuous variables $y$ in the representation $F$ (that is, $n_2 > 0$) and non-extended otherwise ($n_2 = 0$). Furthermore, as suggested by the definition of $R$, we distinguish between variable bounds (e.g. $l^y \leq y \leq u^y$) and general inequalities ($Ax + By + Cz \leq d$), as modern MIP solvers are able to incorporate variable bounds with minimal extra computational cost.

6 To handle the unbounded case, we allow the variable bounds $l^x \leq x \leq u^x$ and $l^y \leq y \leq u^y$ to take infinite values.

7 An ideal formulation is also sharp, i.e. its relaxation projects down to the convex hull of the set we are formulating ($\text{Proj}_x(R) = \text{Conv}(Q)$).
4.2. Existing formulations for combinatorial disjunctive constraints  

A standard formulation for CDC(S) adapted from Jeroslow and Lowe [25] is

\[ \lambda_v = \sum_{S \subseteq v \in S} \gamma_v^S \quad \forall v \in J \]  
\[ z_S = \sum_{v \in S} \gamma_v^S \quad \forall S \in S \]  
\[ \sum_{S \subseteq S} z_S = 1 \]  
\[ \gamma_v^S \in \Delta_v \quad \forall S \in S \]  
\[ (\lambda, z) \in \Delta^J \times \{0, 1\}^S. \]  

This formulation has \( d = |S| \) auxiliary binary variables, \( \sum_{S \subseteq S} |S| \) auxiliary continuous variables, and no general inequalities. Additionally, it is ideal.

Using Proposition 9.3 from [45], we can construct an ideal MIP formulation with fewer auxiliary binary variables:

\[ \lambda_v = \sum_{S \subseteq v \in S} \gamma_v^S \quad \forall v \in J \]  
\[ \sum_{S \subseteq v \in S} \lambda_v^S = 1 \]  
\[ \sum_{S \subseteq v \in S} h_v^S \gamma_v^S = z \]  
\[ \gamma_v^S \geq 0 \quad \forall S \in S \]  
\[ z \in \{0, 1\}^r, \]  

where \( \{h_v^S\}_{S \subseteq S} \subseteq \{0, 1\}^r \) is some set of distinct binary vectors. This formulation is actually a generalization of (8), which we recover if we take \( h^S = e_S \in \mathbb{R}^S \) as the canonical unit vectors. If instead we take \( r \) to be as small as possible while remaining distinct, we recover \( r = \lfloor \log_2(d) \rfloor \). Therefore, formulation (9) yields an ideal extended formulation for (1) with \( \lfloor \log_2(d) \rfloor \) auxiliary binary variables, \( \sum_{S \subseteq S} |S| \) auxiliary continuous variables, and no general inequalities. The following corollary shows that this is the smallest number of auxiliary binary variables we may hope for.

**Proposition 1.** If the sets \( S \) are irreduntant, then any binary MIP formulation for CDC(S) must have at least \( \lfloor \log_2(d) \rfloor \) auxiliary binary variables.

**Proof** See Appendix B. \( \square \)

The formulations thus far have been extended formulations, as they are constructed by formulating each polyhedra \( P^i \) separately and then aggregating them, rather than working with the combinatorial structure underlying the shared extreme points. Therefore, each of these formulations requires a copy of the multiplier \( \gamma_v^S \) for each set \( S \in S \) for which \( v \in S \), and so \( \sum_{i=1}^d |S| \) auxiliary continuous variables total.

In contrast, we can construct non-extended formulations for CDC(S) that work directly on the \( \lambda \) variables and the underlying combinatorial structure of \( S \). An example of a non-extended formulation for CDC is the widely used ad-hoc formulation (see [45, Section 6] and the references therein) given by

\[ \lambda_v \leq \sum_{S \subseteq v \in S} z_S \quad \forall v \in J \]  
\[ \sum_{S \subseteq v \in S} z_v = 1 \]  
\[ (\lambda, z) \in \Delta^J \times \{0, 1\}^S. \]
This formulation is not necessarily ideal, and it requires no auxiliary continuous variables, \( d \) auxiliary binary variables, and \(|J|\) general inequalities.

In summary, we have seen an ideal extended formulation (9) for CDC(\( S \)) with relatively few auxiliary binary variables, but potentially many auxiliary continuous variables. On the other end of the spectrum, we have a non-extended formulation (10) with no auxiliary continuous variables, but which requires relatively many auxiliary binary variables and which may fail to be ideal. However, we know that in special cases we can construct ideal, non-extended formulations with only \( O(\log(d)) \) auxiliary variables and constraints (e.g. SOS1, SOS2, and particular 2-dimensional grid triangulations [46, 50]). This work provides a framework for constructing such small, strong, non-extended MIP formulations for CDC(\( S \)), which are automatically ideal, and in the best case will have \( O(\log(d)) \) auxiliary binary variables and general inequality constraints.

5. Independent branching schemes Vielma and Nemhauser [50] introduced the notion of an independent branching scheme as a natural framework for constructing formulations for combinatorial disjunctive constraints. The independent branching scheme is a logically equivalent way of expressing a CDC in terms of a conjunction of dichotomies: that is, as a series of choices

**Definition 2.** A \( k \)-way independent branching (IB) scheme of depth \( t \) for CDC(\( S \)) is given by a family of sets \( (L_1^j, \ldots, L_k^j) \) (where each \( L_i^j \subseteq 2^J \)) for \( j \in [t] \), where

\[
\text{CDC}(S) = \bigcap_{j=1}^{t} \left( \bigcup_{i=1}^{k} Q(L_i^j) \right). \tag{11}
\]

We say that such an IB scheme has depth \( t \), and that each \( j \in [t] \) yields a corresponding level of the IB scheme \( \bigcup_{i=1}^{k} Q(L_i^j) \), given by the \( k \) alternatives \( Q(L_i^j) \).

An equivalent way of understanding these representations, which we will be using for the remainder of this work, is by eschewing the polyhedra \( Q(L_i^j) \) and working directly on the underlying set \( L_i^j \). That is, a valid \( k \)-way IB scheme satisfies the condition that

\( T \subseteq 2^J \) is a feasible set \( \Longleftrightarrow \forall j \in [t], \exists i \in [k] \) s.t. \( T \subseteq L_i^j \).

First, we observe that, due to our assumption that \( S \) covers the ground set, we have that for each element \( v \in J \), there will be at least one alternative \( i \in [k] \) such that \( v \in L_i^j \) for each level \( j \in [t] \). We will use this extensively in the analysis to come, as it simplifies some otherwise tedious case analyses. Second, we see that this definition can capture potential schemes with a variable number of alternatives in each level by adding empty alternatives \( L_i^j = \emptyset \), provided we take \( k \) as the maximum number of alternatives for all levels. For notational simplicity, we say that a 2-way IB scheme is a pairwise IB scheme, and in this case we write the sets as \( \{(L_i^j, R_i^j)\} \) as in [50]. In contrast, we will call the case with \( k > 2 \) a multi-way IB scheme.

In this form, we have replaced the monolithic constraint CDC(\( S \)) by \( t \) constraints, each of which require the selection between \( k \) alternatives. Given this form, we may use standard techniques to construct a corresponding mixed-integer formulation.

**Proposition 2.** Given an independent branching scheme \( \{(L_1^j, \ldots, L_k^j)\} \) for CDC(\( S \)), the following is a valid formulation for CDC(\( S \)):

\[
\sum_{v \notin L_i^j} \lambda_v \leq 1 - z_i^j \quad \forall j \in [t], \forall i \in [k] \tag{12a}
\]
\[ \sum_{i=1}^{k} z_i^j = 1 \quad \forall j \in [t] \quad (12b) \]
\[ \lambda \in \Delta^j \quad (12c) \]
\[ z^j \in \{0, 1\}^k \quad \forall j \in [t]. \quad (12d) \]

The formulation is known to be ideal for \( k = 2 \) [48, 50]. It has no auxiliary continuous variables, \( kt \) auxiliary binary variables, and \( kt \) general inequalities.

5.1. Constraint branching via independent branching-based formulations

The canonical algorithmic technique for mixed-integer programming is some variation of branch-and-bound [29], which implicitly enumerates all possible values for the binary variables. In its simplest form, a sequence of problems are solved, starting with the relaxation of the MIP formulation, after which a binary variable \( z_i \) is chosen for branching. That is, the current problem is branched into two subproblems: one with the additional constraint \( z_i \leq 0 \), another with \( z_i \geq 1 \). Repeating this procedure, the subproblems form a (binary) tree whose leaves correspond to all \( 2^n \) possible values for the \( n \) binary variables in formulation (7). At any given subproblem, the augmented relaxation to be solved is described by the set of binary variables fixed to zero, and the set of those fixed to one.

The spirit of constraint branching is to allow richer branching decisions. For example, a branching decision might be between \( k \) alternatives of the form \( \{Q^j\}_{i=1}^k \), where each \( Q^j \) is the intersection of the relaxation \( R \) with some new set of general inequality constraints. The concept has significant overlap with the broader field of constraint programming [6, 24], which has been recognized and exploited in the mixed-integer programming literature [1, 5, 22, 36, 41]. More complex constraint branching can often lead to a more balanced branch-and-bound tree, in the sense that the number of sets \( i \in [d] \) such that \( P^i \subseteq Q^j \) is roughly the same for each alternative \( j \in [k] \) (see, for example, [45, Section 8] and [52] for more discussion). This is often a crucial computational feature, as unbalanced trees typically produce weak dual bounds, which leads to long solution times. However, as only specialized versions (i.e. SOS1 or SOS2) are natively supported by MIP solvers such as Gurobi and CPLEX, constraint branching is considerably more difficult to implement efficiently than binary variable branching, which is the default.

For the pairwise \((k = 2)\) case, the formulation (12) induces a constraint branching in the sense that, when traditional binary branching is applied \((z_i \leq 0 \text{ or } z_i^j \geq 1, \text{ respectively})\), the constraints (12a,12b) impose a constraint branching \((\sum_{v \in \mathcal{S}}^j \lambda_v \leq 0 \text{ or } \sum_{v \in \mathcal{S}}^j \lambda_v \leq 0, \text{ respectively})\). In this way, (12) naturally enjoys the computational benefits of constraint branching, all within the binary variable branching framework for which modern MIP solvers are optimized.

Furthermore, we can now motivate in what sense an independent branching scheme is “independent,” per the discussion in [50, Section 3]. In the binary variable branching applied to a formulation of the form (9) or (10), the set of \( \lambda \) variables fixed to zero when \( z_i^j \leq 0 \text{ or } z_i^j \geq 1 \) is not independent of the sets of \( z \) variables previously branched on. For example, consider SOS2(5) (i.e. \( S = \{\{1, 2\}, \{2, 3\}, \{3, 4\}, \{4, 5\}\} \)). For this particular instance, formulation (9) for an appropriate choice of \( \{h^3\}_{v \in \mathcal{S}} \) is

\[
\begin{align*}
\lambda_1 &= \gamma_1^{(1, 2)}, \\
\lambda_2 &= \gamma_2^{(1, 2)} + \gamma_2^{(2, 3)}, \\
\lambda_3 &= \gamma_3^{(2, 3)} + \gamma_3^{(3, 4)}, \\
\lambda_4 &= \gamma_4^{(3, 4)} + \gamma_4^{(4, 5)}, \\
\lambda_5 &= \gamma_5^{(4, 5)}, \\
\gamma_1^{(1, 2)} + &\gamma_2^{(1, 2)} + \gamma_2^{(2, 3)} + \gamma_3^{(2, 3)} + \gamma_3^{(3, 4)} + \gamma_4^{(3, 4)} + \gamma_4^{(4, 5)} + \gamma_5^{(4, 5)} = 1, \\
\gamma_1^{(1, 2)} + &\gamma_2^{(1, 2)} + \gamma_2^{(2, 3)} + \gamma_3^{(2, 3)} = z_1, \\
\gamma_1^{(1, 2)} + &\gamma_2^{(1, 2)} + \gamma_3^{(3, 4)} + \gamma_4^{(3, 4)} = z_2, \\
\gamma_{\gamma}^{S} &\geq 0 \quad \forall \gamma, \mathcal{S}, \\
(\lambda, z) &\in \Delta^5 \times \{0, 1\}^2.
\end{align*}
\]
and a pairwise IB-based formulation (simplified slightly from (12)) is

\begin{align}
\lambda_1 + \lambda_2 & \leq z_1 \tag{14a} \\
\lambda_4 + \lambda_5 & \leq 1 - z_1 \tag{14b} \\
\lambda_3 & \leq z_2 \tag{14c} \\
\lambda_1 + \lambda_5 & \leq 1 - z_2 \tag{14d} \\
\sum_{j=1}^i z_j &= 1 \tag{14e} \\
(\lambda, z) &\in \Delta^5 \times \{0,1\}^2. \tag{14f}
\end{align}

In Figure 2, we see the first two levels of the branch-and-bound trees for both formulations, where we first branch on \(z_1\), and then on \(z_2\). At the second level (branching on \(z_2\)), the formulation is always able to prove that either \(\lambda_1 = 0\) or \(\lambda_3 = 0\), depending on whether \(z_2 = 0\) or \(z_2 = 1\). The branch is also able to prove that \(\lambda_3 = 0\) in one subproblem. However, the subproblem under which the formulation can prove this is conditional on the value to which \(z_1\) is fixed.

In contrast, in independent branching schemes, the branching decisions fix components of the \(\lambda\) variables to zero independently of the values to which the other binary variables have been previously fixed in the tree. For example, in Figure 2, we see that regardless of the value to which \(z_1\) is fixed, branching on \(z_2\) is always able to prove that either \(\lambda_1 = 0\) (if \(z_2 = 0\)) or \(\lambda_3 = \lambda_5 = 0\) (if \(z_2 = 1\)). This has the potential to simplify branching rules (i.e. choosing which variable \(z_i\) to branch on), a notoriously difficult and computationally important part of the algorithmic performance of a MIP solver (see, for example, [2]).

Finally, we make the tautological observation that, if there is any constraint branching of the form (11), this immediately implies a corresponding independent branching scheme, and vice versa. From the definition of \(Q(\cdot)\), each alternative in (11) is defined by the set of components of \(\lambda\) which it fixes to zero. From this, we can see that there is an equivalency between independent branching schemes, and “simple” constraint branching representations whose alternatives do not use general inequalities, and so the nonexistence of IB schemes implies the nonexistence of these “simple” constraint branchings.

\begin{figure}
\centering
\includegraphics[width=\textwidth]{fig2}
\caption{The branch-and-bound trees for (13) (Left) and (14) (Right), when \(z_1\) is first to branch on, and then \(z_2\). The sets inside the nodes is \(I \subset [5]\), the set of components of \(\lambda\) where the subproblem can prove that \(\lambda_v = 0\) for each \(v \in I\). The text on the lines show the branching decision (e.g. \(z_2 = 1\)), and the variables \(\lambda_v\) for which the subproblem is able to prove that \(\lambda_v = 0\) independently, along with the new variables it is able to prove are zero given the previous branching decisions. This figure is adapted from [50, Figure 2].}
\end{figure}
6. Independent branching scheme representability To start, we observe that the independent branching approach is not sufficiently general to capture every possible formulation for CDC(S). In particular, there is the restriction that each alternative \( Q(L^i_j) \) restricts the variables to lie on a single face of the standard simplex. A natural first question is then: given a family of sets \( S \), do any \( k \)-way IB schemes exist for CDC(S)? To begin, we provide both a positive and negative example; we will then introduce a general characterization.

6.1. SOS2 Vielma and Nemhauser [50] construct a pairwise IB scheme for SOS2(N) constraints of depth logarithmic in \( N \). Recall that \( J = \{ N \} \) and \( S = \{ \{ \tau, \tau + 1 \} : \tau \in [N - 1] \} \) for SOS2. The construction is built around a Gray code [43], or sequence of distinct binary vectors \( \{ v^i \}_{i=1}^{N-1} \subseteq \{0,1\}^{\log_2(N-1)} \) where each adjacent pair \( (v^i, v^{i+1}) \) differs in exactly one component. Notationally, here and throughout, take \( v^0 \triangleq v^1 \) and \( v^N \triangleq v^{N-1} \). The pairwise IB scheme is then given by

\[
L^t = \{ \tau \in [N] : v^T_j = 1 \text{ or } v^T_j = 1 \} \quad \forall j \in \{ \log_2(N - 1) \}
\]

\[
R^t = \{ \tau \in [N] : v^T_j = 0 \text{ or } v^T_j = 0 \} \quad \forall j \in \{ \log_2(N - 1) \}
\]

We observe that the resulting formulation matches the lower bound from Proposition 1 with respect to the number of auxiliary binary variables.

6.2. Cardinality constraints However, the multi-way IB scheme approach is not capable of capturing every possible CDC; we now see that there do not exist multi-way IB schemes for the cardinality constraint with few alternatives.

**Proposition 3.** Any \( k \)-way independent branching scheme for the cardinality constraint of degree \( \ell \) must have \( k > \ell \).

**Proof** Consider some infeasible point \( \lambda \in \Delta^J \setminus \text{CDC}(S) \); that is, \( T \triangleq \{ v \in J : \lambda_v > 0 \} \) is such that \( |T| > \ell \). In order for an IB scheme to be valid for CDC(S), there must exist some level \( j \in [n] \) that renders \( \lambda \) infeasible. That is, \( T \not\subseteq L^j_i \) for each \( i \in [n] \), so for each \( i \) there exists some component \( v(i) \in T \setminus L^j_i \). However, the set \( V \triangleq \{ v(i) \}_{i=1}^{N-1} \) is a feasible set if \( k \leq \ell \). Therefore, in order for the proposed IB scheme to be valid, there must be some set \( L^j_i \) with \( V \subseteq L^j_i \) for every \( j \in [n] \), which contradicts our construction that \( v(i) \notin L^j_i \), and so \( v(i) \not\in V \subseteq L^j_i \) for each \( i \). Therefore, we may not have \( k \leq \ell \).

Furthermore, it is straightforward to construct an independent branching scheme for cardinality constraints of degree \( \ell \) from the “conjunctive normal form” [7] with \( k = \ell + 1 \) (see the proof of Theorem 1 for the construction). In other words, the lower bound in Proposition 3 is tight.

Both specialized constraint branching schemes for cardinality constraints [23] and the binary variable branching induced by standard formulations for cardinality constraints are quite imbalanced. The existence of a pairwise independent branching scheme for cardinality constraints would likely have finally produced the sought-after balanced constraint branching. However, Proposition 3 implies that such a balanced constraint branching cannot be produced via IB schemes, or equivalently by constraint branchings that do not use general inequalities.

6.3. Characterization of multi-way branching scheme representability To generalize the results in the previous section, we will characterize the existence of \( k \)-way IB scheme exists for a generic combinatorial disjunctive constraint and any choice of \( k \). In particular, the simple observation underlying the proof of Proposition 3 generalizes to the notion of a “blocker” set that independent branching schemes are incapable of expressing. These blocker sets turn out to be exactly the minimal infeasible subsets of \( J \), and the cardinality of these sets gives a lower bound on the number of dichotomies any IB scheme must have.
Theorem 1. A k-way IB scheme exists if and only if each minimal infeasible set \( T \subseteq J \) has \( |T| \leq k \).

Proof To show the condition is necessary, we use the same approach as in the proof of Proposition 3. If a minimal infeasible subset \( T \subseteq J \) with \( |T| > k \) exists, then for any potential IB scheme, there must exist some level \( j \in [\ell] \) with \( T \supseteq L_j^i \) for all \( i \in [k] \); that is, for each \( i \), there exists some \( v(i) \in T \setminus L_j^i \). However, the set \( \{v(i)\}_{i=1}^k \subseteq T \) must be feasible from minimality of \( T \), a contradiction.

To show the condition is sufficient, we offer an explicit construction. Take \( \mathcal{T} \) as the set of all minimal infeasible sets; from assumption, \( |T| \leq k \) for each \( T \in \mathcal{T} \). Our construction is then:

\[
L_i^T = \begin{cases} 
J \setminus \{T_i\} & i \leq |T| \\
\emptyset & \text{o.w.}
\end{cases} \quad \forall i \in [k], \forall T \in \mathcal{T},
\]

where \( T_i \) is the \( i \)-th element of \( T \) (for a given arbitrary ordering). Every infeasible set \( T \in \mathcal{T} \) has some level that renders it infeasible: namely, level \( T \). Since all minimal infeasible sets are contained in \( \mathcal{T} \), this implies that every infeasible set for the CDC (even those with cardinality strictly greater than \( k \)) is infeasible for this IB scheme. It only remains to show that each feasible set for CDC(S) is feasible for this IB scheme as well.

Choose some \( S \in \mathcal{S} \). For each \( T \in \mathcal{T} \) that is infeasible for the IB scheme, \( T \supseteq S \), and so there exists some \( v \in T \setminus S \). If \( v = T_i \), then we have that \( S \in L_i^T \), rendering \( S \) feasible for this level \( T \). \( \Box \)

Throughout, we will say that CDC(S) is \( k \)-way IB-representable (or pairwise IB-representable for \( k = 2 \)) if the condition in Theorem 1 holds.

It is doubtful that the IB scheme presented in the proof for Theorem 1 will prove practically useful, since in the worst case, the family of infeasible sets of cardinality at most \( k \) will have size on the order of \(|J|^k\). However, we may turn our attention to finding smaller IB schemes later, once we have established that the remainder of our motivating problems are representable.

SOSk Consider some infeasible set \( T \subseteq J \). It is infeasible since the nonzero elements are not all within a \( k \)-consecutive sequence, so we may reduce \( T \) to the infeasible set \( \{\min_{r \in T} \tau, \max_{r \in T} \tau\} \). Therefore, SOSk is pairwise IB-representable.

Cardinality constraints To slightly reframe the proof of Proposition 3, we see that for any set \( T \subseteq J \) with \( |T| > k \), it is impossible to reduce \( T \) to some smaller set \( T' \subsetneq T \) with \( |T'| \leq k \) such that \( T' \) is also infeasible.

Grid triangulations Consider some infeasible set \( T \subseteq J \). If there are some \( v, w \in T \) such that \( \|v - w\|_\infty > 1 \), then there does not exist any triangle on the grid that contains both, so \( \{v, w\} \) is also an infeasible set. Otherwise, we have that \( T \subset \{r, r + 1\} \times \{s, s + 1\} \) for some \( r, s \), and that \( T \) contains elements in both of the triangles in this square. For each of the two triangles, we can select an element of \( T \) that is not contained in the other triangle, which yields an infeasible pair contained in \( T \). Therefore, any grid triangulation is pairwise IB-representable.

Obstacle avoidance and general piecewise linear functions in the plane In Section 7.3, we derive more general conditions for the representability of obstacle avoidance constraint and piecewise linear functions in the plane. In particular, we show that they are always 3-way IB-representable, and that if a (polynomial-time verifiable) condition holds, then the constraint is pairwise IB-representable as well.

Discretization of multilinear terms Similarly as for the grid triangulation case, each infeasible set \( T \subseteq J \) must necessarily contain two elements \( v, w \in T \) with \( \|v - w\|_\infty > 1 \), and so we have that \( T \) can be reduced to the infeasible pair \( \{v, w\} \). Therefore, any discretization of this form is pairwise IB-representable.
7. Pairwise independent branching schemes

The pairwise independent branching scheme framework was initially introduced by Vielma and Nemhauser [50], where it was used to model particularly structured piecewise linear functions. In the remainder of this work, we will offer a more complete picture of the expressive powers of pairwise IB schemes in particular, and provide an algorithmic framework for constructing them.

7.1. Graphical representation of CDCs

A natural object through which to study the combinatorial structure of a pairwise IB-representable CDC is the incidence graph induced by the sets \( S \). This graph is given by \( H_S \triangleq (J, E) \), where the nodes are the ground set \( J \), and the edges are the feasible pairs \( E \triangleq \{ (u, v) \in J \times J : u \neq v, \{u, v\} \text{ is a feasible set} \} \). We will primarily work with its complement, the conflict graph \( H_S^c \triangleq (J, \bar{E}) \), where \( \bar{E} \triangleq \{ (u, v) \in J \times J : u \neq v, \{u, v\} \text{ is an infeasible set} \} \) is the set of all infeasible pairs of elements of \( J \). The crucial observation that allows us to reduce questions of pairwise representability to this graph is that of Theorem 1, as the edges of the conflict graph of a pairwise IB-representable CDC are exactly the minimal infeasible sets.

Given the graph \( H_S^c \), we will be able to provide both a more refined answer to the question of general pairwise IB representability addressed in Section 6.3 (can I construct any pairwise IB scheme for \( S \)?) and a characterization of the existence of small IB schemes (can I construct a pairwise IB scheme for \( S \) of a given depth \( t \)?) In particular, we will see that general representation hinges on an identification between the sets \( S \) and the maximal independent sets of \( H_S^c \), and that the existence of pairwise IB schemes of a given depth \( t \) will be equivalent to the existence of a particular graphical decomposition of \( H_S^c \) of size \( t \).

7.2. Existence of pairwise IB schemes

In Section 6.3, we saw that we can characterize pairwise IB representability in terms of the existence of minimal infeasible sets. However, we may also do this question graphically, and provide a characterization of general pairwise representability in terms of the maximal independent sets of the conflict graph \( H_S^c \).

**Theorem 2.** CDC(\( S \)) is pairwise IB-representable if and only if the sets \( S \) are exactly the maximal independent sets of \( H_S^c \).

**Proof** First we see that each set \( S \in \mathcal{S} \) induces an independent set in \( H_S^c \), as \( (u, v) \in H_S \) for each pair \( u, v \in S \). For the “if” direction, suppose that the one-to-one correspondence exists. We would like to show that each minimal infeasible set \( T \subseteq J \) must have \( |T| = 2 \). From the covering assumption that \( J = \bigcup \{ S \in \mathcal{S} \} \), we must have that \( |T| \geq 2 \). Suppose for contradiction that there exists a minimal infeasible set \( T \) with \( |T| \geq 3 \). From minimality, each set \( T' \subseteq T \) is a feasible set, i.e. \( T' \subseteq S \) for some \( S \in \mathcal{S} \). In particular, each pair \( \{r, s\} \nsubseteq T \) is feasible, and so does not induce an edge in \( H_S^c \). Therefore, \( T \) forms an independent set in \( H_S^c \), and so is contained in a maximal independent set. However, this implies that \( T \subseteq S \) for some \( S \in \mathcal{S} \) from the one-to-one correspondence, a contradiction of its infeasibility.

For the “only if” direction, suppose that \( S \) is pairwise IB-representable, and so all minimal infeasible sets have cardinality 2. First, we show that each set \( S \in \mathcal{S} \) is a maximal independent set of \( H_S^c \). It is clear that each feasible set \( S \in \mathcal{S} \) induces an independent set of \( H_S^c \) \( ((u, v) \in E \) for each \( u, v \in S) \). To show that \( S \) is maximal, it is, for contradiction, that there is some independent set \( S' \) that strictly contains \( S \) with \( S' = S \cup \{v\} \) for some \( v \in J \setminus S \). From the irredundancy assumption, since \( S \subseteq S' \) and \( S \in \mathcal{S} \), it must be that \( S' \) is an infeasible set. Furthermore, since \( S \) is pairwise IB-representable, this means that there is some \( u \in S \) such that \( (u, v) \) is an infeasible pair. However, this implies that \( (u, v) \in \bar{E} \), which means that \( S' \) cannot be an independent set of \( H_S^c \), a contradiction.

Second, to show that each maximal independent set corresponds with a set in \( \mathcal{S} \), consider a maximal independent set \( C \). If \( C \) is a feasible set for CDC(\( S \)), it must be that it must be an element of \( \mathcal{S} \). To see this, note that if \( C \) were strictly contained in some \( S \in \mathcal{S} \) with \( C \nsubseteq S \), this implies
that $S$ induces a strictly larger independent set in $H_\mathcal{S}$ containing $C$, contradicting its maximality. Therefore, we just need to show that $C$ is a feasible set. If not, then as $\mathcal{S}$ is pairwise IB-representable, there exists some pair $u,v \in C$ such that $\{u,v\}$ is an infeasible set. However, $(u,v)$ does not induce an edge in $H_\mathcal{S}$ as $C$ is an independent set in the graph, a contradiction. \hfill \Box

Therefore, verifying general pairwise representability reduces to enumerating the maximal independent sets of $H_\mathcal{S}$ and identifying them to exactly the sets $S$. As an example, we can see that, for cardinality constraint of degree $\ell$ with $2 \leq \ell < |J|$, the only maximal independent set of $H_\mathcal{S}$ is the entire ground set $J$, which certainly cannot be identified with $S = \{S \subset J : |S| = \ell\}$.

7.3. Application: Polygonal partitions of the plane Consider a (not necessarily convex) bounded domain in the plane $\Omega \subset \mathbb{R}^2$ and a partition of $\Omega$ given by polyhedra $\{P_i\}_{i=1}^d$. For a partition, we need that $\bigcup_{i=1}^d P_i = \Omega$ and $\text{relint}(P_i) \cap \text{relint}(P_j) = \emptyset$ for each $i,j \in [d]$. We will describe these polyhedra in V-form, and so the corresponding combinatorial disjunctive constraint is given by $S = \{\text{ext}(P_i)\}_{i=1}^d$ and $J = \bigcup \{S \in \mathcal{S}\}$. We additionally will require that the polyhedra do not have “internal vertices:” if $v \in J \cap P_i$, then $v \text{\ in } \text{ext}(P_i)$.\footnote{We note in passing that under the internal vertex condition, the sets $S$ correspond to the maximal elements of a polyhedral complex \cite[Section 5.1]{53}.}

In this setting, minimal infeasible sets have a natural interpretation.

**Theorem 3.** Take bounded $\Omega \subset \mathbb{R}^2$ and a polyhedral partition with no internal vertices, given by the polyhedra $\{P_i\}_{i=1}^d$ and the corresponding sets $\mathcal{S} = \{\text{ext}(P_i)\}_{i=1}^d$. The minimal sets $T \subseteq J$ infeasible for CDC($\mathcal{S}$) are exactly the independent sets $T$ in $H_\mathcal{S}$ that are infeasible for CDC($\mathcal{S}$) with $|T| \leq 3$.

**Proof** Take some independent set $T$ in $H_\mathcal{S}$, where $T$ is infeasible and $|T| \leq 3$. It is easy to see that $T$ must be minimal, as every strictly contained subset of $T$ has cardinality at most 2, and every pair $u,v \in T$ must be feasible in order for $(u,v) \notin E$. This immediately gives one direction of the result.

To show the other direction, consider a minimal infeasible set $T$. First, we can see that $T$ is an independent set in $H_\mathcal{S}$. If this is not the case, i.e. there is some edge $(u,v) \in E$ with $u,v \in T$, then $\{u,v\}$ is an infeasible set strictly contained within $T$, contradicting minimality. Therefore, it just remains to show that $|T| \leq 3$. Assume for contradiction that $r \overset{\text{def}}{=} |T| > 3$, and label the points $T = \{v_i\}_{i=1}^r$. First, we show that the points may not be in general position, i.e. that w.l.o.g. $v^r \in \text{Conv}(\{v_i\}_{i=1}^{r-1})$. Then, we argue that the points not being in general position implies that $\{v_j\}_{j=1}^{r-1}$ is also an infeasible set, violating the minimality condition.

Assume for contradiction that the points are in general position; that is, that none can be written as a convex combination of the others. This implies that $\text{ext}(\text{Conv}(T)) = T$. Assume that the ordering $\{v^1,\ldots,v^r\}$ forms a path around the edges of $\text{Conv}(T)$; that is, $v^i$ and $v^j$ both lie on an edge of $\text{Conv}(T)$ if and only if $|i - j| = 1$ or $\{i,j\} = \{1,r\}$.

Choose some set $S^1 \subseteq \mathcal{S}$ and some $2 < j < r$ such that $v^1,v^j \in S^1$ and $v^2 \notin S^1$; the associated polyhedra is $P^1$. Such a set exists, else $T$ is not a minimal infeasible set (choose instead $T \leftarrow T \setminus \{v^2\}$). Now choose $S^2 \subseteq \mathcal{S}$ such that $v^2,v^r \in S^2$; the associated polyhedra is $P^2$. Such set exists, as $(v^2,v^r) \in E$. As the nodes $v^1,v^2,v^j,v^r$ are interlaced along the boundary of $\text{Conv}(T)$, we have that $\text{Conv}(\{v^1,v^2\}) \cap \text{Conv}(\{v^2,v^j\}) \subseteq \text{Conv}(T)$ is nonempty. As each of the four points is on the boundary of $\text{Conv}(T)$, and the points are in general position, it follows that $\text{Conv}(\{v^1,v^2\}) \cap \text{Conv}(\{v^2,v^r\}) = \text{relint}(\text{Conv}(\{v^1,v^2\})) \cap \text{relint}(\text{Conv}(\{v^2,v^r\}))$. Therefore, there must exist some point $y$ with $y \in \text{relint}(\text{Conv}(\{v^1,v^2\})) \subseteq \text{relint}(P^1)$ and $y \in \text{relint}(\text{Conv}(\{v^2,v^r\})) \subseteq \text{relint}(P^2)$. However, this implies that $\text{relint}(P^1) \cap \text{relint}(P^2) \neq \emptyset$, which contradicts the assumption that our sets partition the region $\Omega$. 

Finally, it just remains to show that \( \{v^i\}_{i=1}^{r-1} \) is also an infeasible set, and therefore \( \{v^i\}_{i=1}^{r} \) cannot be a minimal infeasible set. Assume for contradiction that it is not: i.e. that there exists some \( j \) such that \( \{v^i\}_{i=1}^{r-1} \subseteq \text{ext}(P^j) \). But this implies that \( v^r \in J \) and \( v^r \in \text{Conv}(T) \subseteq P^j \), yet \( v^r \notin \text{ext}(P^j) \), a contradiction of the internal vertices assumption.

In other words, every polyhedral partition of the plane is 3-way independent branching-representable, and pairwise IB representability can be checked in time polynomial in \(|J|\) (for example, by enumerating the subsets of \( J \) of cardinality 3). To illustrate, in Figure 3 we depict the three possible cases for a partition with respect to Theorem 3: it 1) does not satisfy the internal vertices condition, 2) admits a pairwise IB scheme, 3) does not admit a pairwise IB scheme, but does admit a 3-way IB scheme.

![Figure 3. Partitions of a nonconvex region in the plane that: do not satisfy the internal vertices condition (Left), admit a pairwise IB scheme (Center), and admit a 3-way IB scheme but not a pairwise one (Right).](image)

Furthermore, we can argue that we can always represent a obstacle avoidance constraint in such a way that it admits a pairwise IB scheme. Inspecting Figure 3, we see that the region \( \Omega \) is the same in each, and it is only the partition of \( \Omega \) that can potentially lead to constraints that are not pairwise IB-representable. Therefore, the obstacle avoidance constraint is invariant to the specification of the partition, and it is always possible to construct the partition in such a way that satisfies the conditions of Theorem 3.

7.4. Representation at a given depth

Once a CDC has been shown to be pairwise IB-representable in general, a natural next question is: what is the smallest possible depth at which we may construct an IB scheme? That is, we ask if there exists a pairwise IB scheme for CDC(\( S \)) of some given depth \( t \). The answer to this question reduces to the existence of a graphical decomposition of the conflict graph \( H_{S}^c \).

**Definition 3**. A biclique cover of graph \( G = (J,E) \) is a collection of complete bipartite graphs given by \( \{ (A^j,B^j) \}_{j=1}^{t} \) (with disjoint sets \( A^j, B^j \subseteq J, A^j \cap B^j = \emptyset \)) such that the edgesets of the bipartite graphs \( E^j = A^j \times B^j \) cover \( E \) (i.e. \( \bigcup_{j=1}^{t} E^j = E \)). We can now show that, given a biclique cover for \( H_{S}^c \) of size \( t \), we can construct a corresponding IB scheme for CDC(\( S \)) of depth \( t \).

---

9 Note that this is not the case for piecewise linear functions over \( \Omega \), as the choice of the partition is intimately connected with the values the function may take.
THEOREM 4. Given a biclique cover $\{(A^j,B^j)\}_{j=1}^t$ of the conflict graph $H_S^c$ for pairwise IB-representable CDC($S$), the following is a pairwise IB scheme for CDC($S$) of depth $t$:

$$L^j = \bigcup \left\{ S \in S : S \cap A^j \neq \emptyset \text{ or } S \cap B^j = \emptyset \right\} \quad \forall j \in [t] \quad (15a)$$

$$R^j = \bigcup \left\{ S \in S : S \cap A^j = \emptyset \text{ or } S \cap B^j \neq \emptyset \right\} \quad \forall j \in [t]. \quad (15b)$$

**Proof** From the definition of $L^j$ and $R^j$, each set $S \in S$ is feasible for this IB scheme, in that sense that, for each $j \in [t]$, either $S \subseteq L^j$ or $S \subseteq R^j$. This follows, as either $S \cap A^j = \emptyset$ or $S \cap B^j \neq \emptyset$. As CDC($S$) is pairwise IB-representable, each minimal infeasible set has cardinality 2. Therefore, if we show each pair $(a,b) \in \tilde{E}$ is infeasible for the proposed IB scheme, this verifies that each infeasible set for CDC($S$) is also infeasible for the proposed IB scheme. This then implies that $\{(L^j,R^j)\}_{j=1}^t$ is a valid pairwise IB scheme for CDC($S$).

As $\{(A^j,B^j)\}_{j=1}^t$ is a biclique cover for $H_S^c$, there is some level $j \in [t]$ with $(a,b) \in \tilde{E}^j$. W.l.o.g., take $a \in A^j$, $b \in B^j$. From the covering assumption on $S$, we have that $a \in L^j$ and $b \in R^j$, as there must exist some $S,S' \in S$ with $a \in S$ and $b \in S'$, and therefore $S \cap A^j \neq \emptyset$ and $S' \cap B^j \neq \emptyset$. Then we have that some level $j$ of the proposed IB scheme renders $(a,b)$ infeasible ($(a,b) \notin L^j$ and $(a,b) \notin R^j$) if $a \in L^j \setminus R^j$ and $b \in R^j \setminus L^j$.

To show that $b \notin L^j$, it suffices from the definition of $L^j$ to show that, for each $S \in S$ with $b \in S$, we have that $S \cap A^j = \emptyset$ and $S \cap B^j \neq \emptyset$. The second follows immediately as $b \in B^j$, so we focus on the first. If both $S \cap A^j \neq \emptyset$ and $S \cap B^j = \emptyset$ (which, again, is true if $b \in B^j$ and $b \in S$), this implies some $r,s \in S$ with $r \in A^j$ and $s \in B^j$, which in turn implies that both $(r,s) \in \tilde{E}$ and $(r,s) \in \tilde{E}$, a contradiction. Therefore, we cannot have some $S$ with $S \cap A^j \neq \emptyset$ and with $b \in S$, and so $b \notin L^j$. A similar argument holds to show that $a \notin R^j$. □

Furthermore, we can show the reverse implication as well: for any pairwise independent branching scheme (that satisfies a minimality condition), there exists a corresponding biclique cover for $H_S^c$ that can be used to reconstruct the IB scheme via (15). That is, a pairwise IB scheme for CDC($S$) of depth $t$ exists if and only if there exists a biclique cover for $H_S^c$ of depth $t$.

**Theorem 5.** If a valid pairwise IB scheme $\{(L^j,R^j)\}_{j=1}^t$ for CDC($S$) is minimal with respect to set inclusion (i.e. $\{(\tilde{L}^j,\tilde{R}^j)\}_{j=1}^t$ is not a valid pairwise IB scheme for CDC($S$) for sets $\tilde{L}^j \subseteq L^j$ and $\tilde{R}^j \subseteq R^j$ for each $j \in [t]$, with at least one containment strict), then there exists a biclique cover $\{(A^j,B^j)\}_{j=1}^t$ such that the construction in (15) holds.

**Proof** Presume there is a a valid IB scheme for CDC($S$), $\{(L^j,R^j)\}_{j=1}^t$, that satisfies the stated minimality conditions. We posit that the associated biclique cover is given by $A^j \overset{\Delta}{=} L^j \setminus R^j$ and $B^j \overset{\Delta}{=} R^j \setminus L^j$ for each $j \in [t]$. Clearly $A^j \cap B^j = \emptyset$. Therefore, to show that this yields a valid biclique cover for $H_S^c$, it suffices to show that $\tilde{E} = \bigcup_{j=1}^t \tilde{E}^j$ (where, again, $\tilde{E}^j = A^j \times B^j$).

First, see that $\tilde{E} \subseteq \bigcup_{j=1}^t \tilde{E}^j$, as each minimal infeasible set $\{a,b\}$ will have some level $j \in [t]$ with (w.l.o.g.) $a \in L^j \setminus R^j = A^j$ and $b \in R^j \setminus L^j = B^j$, and so $(a,b) \in \tilde{E}^j$. Second, to show that $\bigcup_{j=1}^t \tilde{E}^j \subseteq \tilde{E}$, take some arbitrary $j \in [t]$ and some edge $(a,b) \in \tilde{E}^j$. From the definition of our biclique cover, we have that w.l.o.g. $a \in A^j \overset{\Delta}{=} L^j \setminus R^j$ and $b \in B^j \overset{\Delta}{=} R^j \setminus L^j$. Therefore, $(a,b)$ is an infeasible set for the IB scheme, and thus for CDC($S$) as well, and so $(a,b) \in \tilde{E}$.

Finally, we must show that (15) holds for our construction of $\{(A^j,B^j)\}_{j=1}^t$; we only show (15a), as (15b) follows analogously. For the remainder, note that we may rewrite (15a) as $\bigcup \{S \in S : S \cap (L^j \setminus R^j) \neq \emptyset \text{ or } S \cap (R^j \setminus L^j) = \emptyset \}$. To see that $L^j \subseteq \bigcup \{S \in S : S \cap (L^j \setminus R^j) \neq \emptyset \text{ or } S \cap (R^j \setminus L^j) = \emptyset \}$, see if this were not true, then there is some element $v \in L^j$ such that, for all $S \in S$ with $v \in S$, we have that $S \cap (L^j \setminus R^j) = \emptyset$ and $S \cap (R^j \setminus L^j) = \emptyset$. This necessarily implies that $S \subseteq R^j$ (as $L^j \cup R^j = J$ by the covering assumption), which in turn implies that $v \in L^j \cap R^j$. However, shrinking the set $L^j \leftarrow L^j \setminus \{v\}$ yields a valid IB scheme as well. This is because each feasible set $S \in S$ containing $v \in S$ is still feasible w.r.t. the IB scheme, as
we have established that such a \( S \subseteq R_j \) necessarily. Moreover, as \( v \in L_j \cap R_j \), there are no infeasible edges \((u, v) \in \bar{E}\) which are separated by level \( j \). Therefore, we have a strictly smaller valid IB scheme, which contradicts the minimality of the IB scheme.

To see that \( L_j \supseteq \bigcup \{ S \in \mathcal{S} : S \cap A_j \neq \emptyset \text{ or } S \cap B_j = \emptyset \} \), it suffices to show that for all \( S \in \mathcal{S} \) such that either \( S \cap (L_j \setminus R_j) \neq \emptyset \) or \( S \cap (R_j \setminus L_j) = \emptyset \), \( S \subseteq L_j \) necessarily. Consider the first case: \( S \cap (L_j \setminus R_j) \neq \emptyset \). If this is not so (that is, \( S \subseteq L_j \) ), then there are \( u, v \in S \) with \( u \in L_j \setminus R_j \) and \( v \notin L_j \) (and so \( v \in R_j \setminus L_j \) by the covering assumption). However, this implies that \((u, v) \in \bar{E}\) (as \( u, v \in S \in \mathcal{S} \)), and the IB scheme renders \((u, v) \) infeasible with its \( j \)-th level, a contradiction of its validity. Now consider the second case: \( S \cap (R_j \setminus L_j) = \emptyset \). By the covering assumption \((L_j \cup R_j = J)\), this immediately implies that \( S \subseteq L_j \).

Given Theorems 4 and 5, we can now naturally frame the problem of finding a minimum depth pairwise IB scheme as the minimum biclique cover problem \([17, 20]\). Unfortunately, the decision version of this problem is known to be NP-complete \([35]\) and inapproximable within a factor of \(|J|^{1/3-\epsilon}\) if \( P \neq NP\) \([21]\), even for bipartite graphs. However, we note that it is simple to construct a MIP feasibility problem for finding a pairwise IB scheme of a given depth \( t \), which gives us a way to algorithmically find the smallest pairwise IB scheme for a specific (fixed) CDC.

**Proposition 4.** A pairwise IB scheme of depth \( t \) exists for pairwise independent branching-representable CDC(\( S \)) if and only if the following is admits a feasible solution:

\[
\begin{align*}
\sum_{j=1}^{\bar{E}} z_{j}^{r,s} & \geq 1 \quad \forall (r, s) \in \bar{E} \\
\sum_{j=1}^{E} z_{j}^{r,s} & = 0 \quad \forall (r, s) \in E \\
x^{r} & \in \{0,1\}^{d} \quad \forall r \in J \\
z^{r,s} & \in \{0,1\}^{d} \quad \forall (r, s) \in E \cup \bar{E}.
\end{align*}
\]

Moreover, for any feasible solution \((x, z)\), a pairwise IB scheme is given by \( L_j = \{ r \in J : x^{r} = 0 \} \) and \( R_j = \{ r \in J : x^{r} = 1 \} \) for each \( j \in [\bar{t}] \).

**Proof** See Appendix C.

Furthermore, we can restate the MIP formulation from \([50]\) (which is a special case of \((12)\) with \( k = 2 \)) in terms of biclique covers of \( H_{S}^{z} \).

**Proposition 5 (Theorem 5, \([50]\); Theorem 1, \([48]\)).** If CDC(\( S \)) is pairwise independent branching-representable and \( \{(A_j, B_j)\}_{j=1}^{\bar{t}} \) is a biclique cover for \( H_{S}^{z} \), then the following is an ideal formulation for CDC(\( S \)):

\[
\begin{align*}
\sum_{v \in A_j} \lambda_v & \leq z_{j} \quad \forall j \in [\bar{t}] \\
\sum_{v \in B_j} \lambda_v & \leq 1 - z_{j} \quad \forall j \in [\bar{t}] \\
(\lambda, z) & \in \Delta^{J} \times \{0,1\}^{\bar{t}}.
\end{align*}
\]

We end the section by noting that the relation between biclique covers and independent sets has also been exploited in the study of boolean functions, particularly in the equivalence between posiforms and maximum weighted stable sets (e.g. \([15, \text{Theorem 13.16}]\)). In fact, formulation \((17)\) is
reminiscent of formulation (13.45–13.50) in [15, Theorem 13.13]. The main difference between these formulations is that in the context of [15] the $\lambda$ variables will be binary variables not constrained to lie in the unit simplex. For this reason inequalities (17a–17b) appear disaggregated in [15, Theorem 13.13] in the form $\lambda_v \leq z_j$ for all $v \in A^j$, $j \in [l]$. However, the resulting formulation is not ideal (See [50, Section 5] for more details). Still, the combinatorial aspects of this connection could prove useful for constructing small IB schemes.

In the next section, we will explore instances where we can, in closed form, construct small (asymptotically optimal) IB schemes for families of particularly structured CDCs.

8. Illustrative examples

With a framework to construct pairwise independent branching schemes for arbitrary (pairwise IB-representable) CDCs, we now return to some of our motivating examples. We will apply our methodology to these specific structures, and produce small, closed-form IB schemes. In particular, this allows us to construct novel, small MIP formulations for these constraints.

8.1. A simple IB scheme of size $|J|$ 

To start, we show that any pairwise IB-representable CDC admits an IB scheme of depth $|J|$. If $|J|$ is smaller than $|S|$, this already offers a drop in size from (10).

**Proposition 6.** For pairwise IB-representable CDC($S$), a biclique cover for $H^S_v$ is given by:

$$A^v = \{v\}, \quad B^v = \{u \in J : (u, v) \in E\} \quad \forall v \in J.$$ 

**Proof** By construction of the sets, we see that each $(r, s) \in \bar{E}^v \equiv A^v \times B^v$ corresponds to an infeasible edge: that is, $\bar{E}^v \subseteq \bar{E}$ for each $v$, and so $\bigcup_{v \in J} \bar{E}^v \subseteq \bar{E}$. Furthermore, each infeasible edge $(r, s) \in \bar{E}$ is infeasible for levels $r$ and $s$, and so $\bar{E} \subseteq \bigcup_{v \in J} \bar{E}^v$. Therefore, this construction forms a valid biclique cover of the conflict graph.

This gives us an upper bound of $|J|$ on the minimum depth for any pairwise IB-representable CDC.

8.2. SOS3

As a concrete example, we will now study two instances of the SOS3($N$) constraint for small values of $N$. First, consider the instance with $N = 6$, where $S = \{\{1, 2, 3\}, \{2, 3, 4\}, \{3, 4, 5\}, \{4, 5, 6\}\}$. Therefore, $|S| = 4$, yielding a lower bound of depth $\log_2(4) = 2$ from Proposition 1. However, there does not exist a biclique cover of depth 2 (which can be verified via Proposition 4), though one of depth 3 does exist:

$$A^1 = \{1\}, \quad B^1 = \{4, 5, 6\}$$
$$A^2 = \{1, 2\}, \quad B^2 = \{5, 6\}$$
$$A^3 = \{1, 2, 3\}, \quad B^3 = \{6\}.$$ 

We can see the proposed IB scheme on the left side of Figure 4. For clarity, the associated MIP formulation for the CDC from Proposition 5 is

$$\begin{align*}
\lambda_1 & \leq z_1 \\
\lambda_1 + \lambda_2 & \leq z_2 \\
\lambda_1 + \lambda_2 + \lambda_3 & \leq z_3 \\
(\lambda, z) & \in \Delta^6 \times \{0, 1\}^3.
\end{align*}$$

Next, we consider $N = 10$, where we also cannot attain the $\log_2(8) = 3$ lower bound. However, a biclique for this the conflict graph of this constraint is

$$A^1 = \{1, 8, 9, 10\}, \quad B^1 = \{4, 5\}$$
$$A^2 = \{1, 2, 10\}, \quad B^2 = \{5, 6, 7\}$$
$$A^3 = \{1, 2, 3, 9, 10\}, \quad B^3 = \{6\}$$
$$A^4 = \{1, 2, 3, 4\}, \quad B^4 = \{7, 8, 9, 10\},$$
as seen on the right side of Figure 4. The corresponding MIP formulation is

\[
\begin{align*}
\lambda_1 + \lambda_8 + \lambda_9 + \lambda_{10} \leq z_1 & \quad \lambda_4 + \lambda_5 \leq 1 - z_1 \\
\lambda_1 + \lambda_2 + \lambda_{10} \leq z_2 & \quad \lambda_5 + \lambda_6 + \lambda_7 \leq 1 - z_2 \\
\lambda_1 + \lambda_2 + \lambda_3 + \lambda_9 + \lambda_{10} \leq z_3 & \quad \lambda_6 \leq 1 - z_3 \\
\lambda_1 + \lambda_2 + \lambda_3 + \lambda_4 \leq z_4 & \quad \lambda_7 + \lambda_8 + \lambda_9 + \lambda_{10} \leq 1 - z_4
\end{align*}
\]

\[(\lambda, z) \in \Delta^{10} \times \{0, 1\}^4.\]

![Figure 4](image.png)

**Figure 4.** Visualizations of the biclique covers presented in the text for SOS3 (Left) and SOS3(10) (Right). Each row corresponds to some level \(j\), and the elements of \(A^j\) and \(B^j\) are the green and blue nodes, respectively.

### 8.3. SOS\(k\)

The construction for SOS3(6) above suggests a more general construction for SOS\(k(N)\) when \(k \leq N/2\) (assume for convenience that \(N\) is even). Consider the sets given by

\[
A^j = \{1, \ldots, j\} \cup \{j + N/2 + k, \ldots, N\}, \quad B^j = \{j + k, \ldots, j + N/2\}
\]

for each \(j \in \lfloor N/2 \rfloor\). It is straightforward to see that this yields a biclique cover of the conflict graph for SOS\(k(N)\) of depth \(N/2\). Therefore, with this simple operation, we have constructed an ideal formulation for SOS\(k(N)\) with size strictly smaller than \(N\), the size of the naïve nonextended formulation (10).

Based on the second example in Section 8.2, we know that this construction is, in general, not the smallest possible. Indeed, we can construct an IB scheme of depth \(\log_2(N/k) + O(k)\). See Figure 5 for an example of the construction.

**Theorem 6.** Take some fixed \(N \in \mathbb{Z}_+, k \in [N]\), and \(\{v^i\}_{i=1}^{[N/k]-1} \subseteq \{0, 1\}^{[\log_2([N/k]-1)]}\) as any Gray code where \(v^0 \equiv v^1\) and \(v^{[N/k]} \equiv v^{[N/k]-1}\). Then the following is a biclique cover for the conflict graph of SOS\(k(N)\) of depth \([\log_2([N/k]-1)] + 3k\):

\[
\begin{align*}
A^{1,j} & = \left\{ \tau \in J : v_j^{[\tau/k]-1} = 0 \right\} \\
B^{1,j} & = \left\{ \tau \in J : v_j^{[\tau/k]-1} = 1 \right\} \\
A^{2,j'} & = \bigcup_{i=0}^{\lfloor N/3 \rfloor} \{ \tau \in J : \tau = j' + (3i - 3)k \} \\
B^{2,j'} & = \bigcup_{i=0}^{\lfloor N/3 \rfloor} \{ \tau \in J : j' + (3i - 2)k \leq \tau \leq j' + (3i - 1)k \}
\end{align*}
\]
for all \( j \in [\lfloor \log_2(\lfloor N/k \rfloor - 1) \rfloor] \) and \( j' \in [3k] \).

**Proof** Assume for simplicity of notation that \( N/k = \lfloor N/k \rfloor \) (i.e. \( N/k \in \mathbb{Z}_+ \)). If this is not the case, augment \( J \) with \( [N/k] - N/k \) dummy nodes, apply the construction that follows, and then remove the dummy nodes from the resulting sets at the end.

Consider a partition the ground set \( J \) with sets of the form

\[
S^l = \{ k(l - 1) + 1, \ldots, k l \} \quad \forall l \in [N/k].
\]

The idea is to impose SOS2 on the sets \( \{S^l\}_{l=1}^{N/k} \) to separate all infeasible edges that are “far apart,” and then use another construction to separate the remaining infeasible edges that are “closer together.” That is, we construct a “two-stage” biclique cover, where the first stage is given by \( \{(A^1_j, B^1_j)\}_{j=1}^{\lceil \log_2(N/k - 1) \rceil} \) and the second by \( \{(A^2_j', B^2_j')\}_{j'=1}^{3k} \). This yields the desired biclique cover. For illustration, see Figure 5, where we present the construction for SOS3(26). In this example, the first stage has \( \lceil \log_2(\lfloor 26/3 \rfloor - 1) \rceil = 3 \) levels, and the second stage has \( 3k = 9 \) levels.

Formally, consider some infeasible edge \((r, s) \in \bar{E}\) (w.l.o.g. \( r < s \)). From definition of the SOSk constraint, this implies that \( s - r \geq k \). We consider two cases. If \( s - r \geq 2k \) (i.e. “far away”), then \( r \in S^l \) and \( s \in S^{l'} \) uniquely for \( \ell' - \ell \geq 2 \). Therefore, the first stage IB scheme separates this edge, since from construction it induces an aggregated SOS2 constraint on the sets \( \{S^l\}_{l=1}^{N/k} \). If \( k \leq s - r < 2k \) (i.e. “close together”), note that \( \bigcup_{j=1}^{3k} A^{2,j} = J \), and so the \( j' \) such that \( r \in A^{2,j'} \).

Then, from definition, \( s \in B^{2,j'} \), and so this level will separate the infeasible edge \((r, s) \in \bar{E}\).

Finally, if \( s - r < k \), then this is a feasible edge \((r, s) \in E\), and we must show that neither first nor second stage separates the nodes. To see this, note that \( r \in S^l \) and \( s \in S^{l'} \) for some \( \ell' - \ell \leq 1 \), and so the first stage will not separate them. As for the second, see that for each \( j \in [3k] \), we have that \( \min_{a,b \in A^{2,j} \times B^{2,j}} |a - b| \geq k \), and so we cannot have \((r, s) \in A^{2,j} \times B^{2,j} = \bar{E}^{2,j} \) or \((s, r) \in A^{2,j} \times B^{2,j} = \bar{E}^{2,j} \).

We note that, when \( k = o(\log(N)) \), this biclique cover yields a MIP formulation that is asymptotically tight (with respect to the number of auxiliary binary variables) with our lower bound of \( \lceil \log_2(N - k + 1) \rceil \) from Proposition 1. We can also show an absolute lower bound of depth \( k \) for any biclique cover for SOSk. This implies that when \( k = \omega(\log(N)) \), although the formulation from Theorem 6 is not tight with respect to the lower bound from Proposition 1, it is asymptotically the smallest possible formulation in the pairwise IB framework.

**Proposition 7.** Any biclique cover for the conflict graph of SOSk must have depth at least \( \min\{k, N - k\} \).

**Proof** Define \( \gamma \overset{\text{def}}{=} \min\{k, N - k\} \) and consider any possible biclique cover \( \{(A^l, B^l)\}_{l=1}^t \). The biclique cover must separate the edges \( \{(\tau, \tau + k)\}_{\tau=1}^N \). Consider a level \( j \) of the biclique cover that contains edge \((\tau, \tau + k)\) for some \( \tau \in \gamma \); w.l.o.g., \( \tau \in A^j \) and \( \tau + k \in B^j \). Consider the possibility that the same level \( j \) separates another such edge in the set, e.g. \((\tau', \tau' + k)\) for \( \tau' \in \gamma \), where w.l.o.g. \( \tau < \tau' \). That would imply that either \( \tau' \in A^j \) or \( \tau' \in B^j \). In the case that \( \tau' \in A^j \), we have that \( \bar{E}^j \) contains the edge \((\tau', \tau + k)\). However, since \( \tau + k - \tau' < \tau + k - \tau = k \), this implies that the biclique cover separates a feasible edge, a contradiction. In the case where \( \tau' \in B^j \), we have that \( \bar{E}^j \) contains the edge \((\tau, \tau')\), and as \( \tau' - \tau < k \) from the definition of our set of edges, a similar argument holds. Therefore, each edge \( \{(\tau, \tau + k)\}_{\tau=1}^N \) must be uniquely contained in some level of the biclique cover, giving the result. \( \square \)

Furthermore, when \( k = \lfloor N/2 \rfloor \), this proposition gives a lower bound on the depth of a biclique cover that is asymptotically tight with the upper bound of \( N \) from Proposition 6. In other words, in this particular regime, we have that the SOSk constraint admits a pairwise IB-based formulation, but only one that is relatively large (\( \Omega(|S|) = \Omega(|J|) \)) auxiliary binary variables and constraints.
8.4. Grid triangulations of the plane  Consider the three triangulations of $[0, 2] \times [0, 2]$ shown in Figure 1. The first is the “Union Jack” construction, initially introduced by Todd [44], and studied in the context of piecewise linear functions by Vielma and Nemhauser [48, 50], where they introduced a logarithmically-sized independent branching scheme. The second is the K1 triangulation [28], for which small IB schemes have been presented recently by Vielma [46]. The third is a grid triangulation that does not fit in the frameworks used to analyze the first two: to the best of our knowledge, no small ideal formulations have been presented. In Figure 6, we show by construction the existence of IB schemes for the three triangulations of depths 3, 4, and 5, respectively. These are the smallest possible depths, as can be verified via Proposition 4. The sets $(A^j, B^j)$ in Figure 6 are the green and blue vertices, respectively, in the $j$-th subfigure of a given row. For instance, for the Union Jack example (first row), we have $A^3 = \{(1, 0), (1, 2)\}$ and $B^3 = \{(0, 1), (2, 1)\}$.

We may now present a biclique cover construction for generic grid triangulations, with no further assumptions on the structure of the triangles such as in [46, 48, 50], whose depth scales like $\log_2(M) + \log_2(N) + O(1)$. See Figure 7 for an illustration of part of the construction.

**Theorem 7.** Take $\{v_i^{1 \leq i \leq M}\} \subseteq \{0, 1\}^{\log_2(M)}$ and $\{w_i^{1 \leq i \leq N}\} \subseteq \{0, 1\}^{\log_2(N)}$ as any pair of Gray codes, where $v^0 \overset{\text{def}}{=} v^1$, $v^{M+1} \overset{\text{def}}{=} v^M$, $w^0 \overset{\text{def}}{=} w^1$, and $w^{N+1} \overset{\text{def}}{=} w^N$. The following is a biclique cover
for the conflict graph of any grid triangulations on $J = \{0, \ldots, M\} \times \{0, \ldots, N\}$ that has depth $[\log_2(M)] + [\log_2(N)] + 9$:

$$A^1, \ell = \{(x, y) \in J : v_\ell^x = v_{\ell + 1}^x = 1\}$$
$$B^1, \ell = \{(x, y) \in J : v_\ell^x = v_{\ell + 1}^x = 0\}$$
$$A^2, \ell' = \{(x, y) \in J : w_\ell^y = w_{\ell' + 1}^y = 1\}$$
$$B^2, \ell' = \{(x, y) \in J : w_\ell^y = w_{\ell' + 1}^y = 0\}$$
$$A^{3,(r,s)} = J \cap (3\mathbb{Z}^2 + (r, s))$$
$$B^{3,(r,s)} = J \setminus \left(\bigcup \{S \in \mathcal{S} : S \cap (3\mathbb{Z}^2 + (r, s)) \neq \emptyset\}\right)$$

for all $\ell \in [\log_2(M)]$, $\ell' \in [\log_2(N)]$, and $(r, s) \in \{0, 1, 2\}^2$.

**Proof** The construction consists of two stages. The first consists of applying two SOS2 constraints separately along each axis $\{0, \ldots, M\}$ and $\{0, \ldots, N\}$. For the $x$ direction, we have that the levels $\{(A^1, \ell, B^1, \ell)\}_{\ell=1}^{[\log_2(M)]}$ will separate exactly those edges $((x, y), (x', y')) \in \bar{E}$ with $|x - x'| > 1$, and likewise with levels $\{(A^2, \ell, B^2, \ell)\}_{\ell=1}^{[\log_2(N)]}$ separating exactly those edges $((x, y), (x', y')) \in \bar{E}$ with $|y - y'| > 1$. In other words, together these levels will separate all edges $((x, y), (x', y')) \in \bar{E}$ where $||(x, y) - (x', y')||_\infty > 1$.

For the second stage $\{(A^{3,(r,s)}, B^{3,(r,s)})\}_{(r,s) \in \{0, 1, 2\}^2}$, we would like to show that these dichotomies separate exactly those edges in $H_\mathcal{S}$ with $||(x, y) - (x', y')||_\infty \leqslant 1$; see Figure 7 for intuition as to why this is the case. In particular, notice that the horizontal and vertical lines containing the blue squares do not contain any green diamonds (modulo those on the boundaries).

We can see that these sets $A^{3,(r,s)}$ and $B^{3,(r,s)}$ will not separate any feasible pairs $((x, y), (x', y') \in \mathcal{S}$ for some $S \in \mathcal{S}$). Therefore, we restrict our attention to showing that every infeasible pair will be separated.
Consider some infeasible edge \(((x, y), (x', y')) \in E\) with \(\|(x, y) - (x', y')\|_x \leq 1\). Take \((\hat{r}, \hat{s}) \overset{\text{def}}{=} (x \mod 2, y \mod 2)\). We would like to show that \((x, y) \in A^{3,(\hat{r}, \hat{s})}\) and \((x', y') \in B^{3,(\hat{r}, \hat{s})}\), as this implies that this level of the IB scheme separates the given infeasible edge. Clearly \((x, y) \in A^{3,(\hat{r}, \hat{s})}\), since \((x, y) \in (3Z^2 + (\hat{r}, \hat{s}))\) from our definition of \((\hat{r}, \hat{s})\).

To show that \((x', y') \in B^{3,(\hat{r}, \hat{s})}\), we must show that \((x', y') \not\in \bigcup \{S \in \mathcal{S} : S \cap (3Z^2 + (\hat{r}, \hat{s})) \neq \emptyset\}\); that is, for each \(S \in \mathcal{S}\) such that \(S \cap (3Z^2 + (\hat{r}, \hat{s})) \neq \emptyset\), we would like to show that \((x', y') \not\in S\). Consider some such \(S \in \mathcal{S}\). We cannot have both \((x, y), (x', y') \in S\), as \(((x, y), (x', y'))\) is an infeasible edge in \(\bar{E}\). Therefore, if \((x, y) \in S\), we have the result immediately, so consider the case where \((x, y) \not\in S\). In this instance, we must have that \(S \cap (3Z^2 + (\hat{r}, \hat{s})) = \{(x''', y''')\}\), where \(\|(x, y) - (x''', y''')\|_x \geq 3\). As \(\|(x, y) - (x', y')\|_x \leq 1\), this implies \(\|(x', y') - (x''', y''')\|_x \geq 2\). However, our grid triangulation has \(\|a - b\|_x \leq 1\) for each \(a, b \in S\), implying that, as \((x'', y'') \in S\), we cannot have \((x', y') \in S\). Therefore, \((x', y') \in B^{3,(\hat{r}, \hat{s})}\).

By referring to Proposition 1, we recover a \([\log_2(2MN)] \geq [\log_2(M)] + [\log_2(N)]\) lower bound on the depth of any biclique cover for a grid triangulation, and see that our construction yields a MIP formulation that is within a constant additive factor of the smallest possible.

**8.5. Discretizations of multilinear terms** Recall that we are considering a multilinear function \(f : \mathbb{R}^n \to \mathbb{R}\) given by \(f(x_1, \ldots, x_n) = \prod_{i=1}^n x_i\) over an arbitrary gridding \(\mathcal{R} \overset{\text{def}}{=} \{[h_i, h_i] : e \in \prod_{i=1}^n [d_i]\}\), where \(\mathcal{P}^e = \text{Conv}(\text{gr}(f, R))\) for each \(R \in \mathcal{R}\).

**Proposition 8.** For multilinear function \(f : \mathbb{R}^n \to \mathbb{R}\) given by \(f(x_1, \ldots, x_n) = \prod_{i=1}^n x_i\) and gridding \(\mathcal{R}\), take \(\{v^{i,j}\}_{j=1}^{d_i} \subseteq \{0, 1\}^{\log_2(d_i)}\) as a Gray code for each \(i \in \eta\), where \(v^{i,0} \equiv v^{i,1}\) and \(v^{i,d_i+1} \equiv v^{i,d_i}\). Then the following is a biclique cover for the conflict graph \(H^S_8\) for \(S = \{\text{ext}(\mathcal{P}^R)\}_{R \in \mathcal{R}}\) that is of depth \(\sum_{i=1}^n \log_2(d_i)\):

\[
\begin{align*}
A^{i,j} &= \{(x, y) \in J : \exists \gamma \text{ s.t. } x_i = h_i^{\gamma}, v_j^{i,x_{\gamma}} = v_j^{i,x_{\gamma+1}} = 0\} \\
B^{i,j} &= \{(x, y) \in J : \exists \gamma \text{ s.t. } x_i = h_i^{\gamma}, v_j^{i,x_{\gamma}} = v_j^{i,x_{\gamma+1}} = 1\}
\end{align*}
\]
for each $i \in [\eta]$ and $j \in [\lceil \log_2(d_i) \rceil]$.

**Proof** See Appendix D. □

We note that, since $|S| = \prod_{i=1}^{\eta} d_i$, by Proposition 1 this construction yields a formulation that is asymptotically optimal (with respect to number of auxiliary binary variables) for any possible MIP formulation, up to an additive factor of at most $\eta$.

Furthermore, we can specialize this to the bilinear case studied by Misener et al. [33], where $\eta = 2$, $d_1 = m$, and $d_2 = 1$.

**Corollary 1.** There exists a biclique cover for a the grid discretization of a bilinear function with $d_1 = m$ and $d_2 = 1$ of depth $\lceil \log_2(m) \rceil$.

This result yields an ideal mixed-integer formulation for the outer-approximation of bilinear terms with $\lceil \log_2(m) \rceil$ auxiliary binary variables, $2(m+1)$ auxiliary continuous variables (the $\lambda$ variables, one for element in $J$), and $2\lceil \log_2(m) \rceil$ general inequality constraints. In contrast, the logarithmic formulation from Misener et al. [33] has $\lceil \log_2(m) \rceil$ auxiliary binary variables, $2\lceil \log_2(m) \rceil + 1$ auxiliary continuous variables, at least $2\lceil \log_2(m) \rceil + 6$ general inequality constraints, and is not ideal in general (see Appendix A). Therefore, we gain an ideal formulation with a naturally induced constraint branching at the price of a modest number of additional auxiliary continuous variables. Furthermore, our formulation generalizes readily to discretization along the second dimension ($d_2 \geq 1$), and for higher dimensional multilinear functions ($\eta \geq 2$).

**Acknowledgements** This material is based upon work supported by the National Science Foundation Graduate Research Fellowship under Grant No. 1122374, and by the National Science Foundation under Grant No. CMMI-1351619. The authors would like to thank Yves Crama for pointing out the relation between biclique covers, independent sets, and boolean functions.

**References**


Appendix A: Logarithmic formulation from Misener et al. [33] is not ideal. We show that the logarithmic formulation (16) from Misener et al. [33] is not, in general, ideal. Using their notation, we take $N_P = 3$, $x_L = y_L = 0$, and $x_U = y_U = 3$ (and so $a = 1$). Then formulation (16) is

$$\lambda_1 + 2\lambda_2 \leq x \leq 1 + \lambda_1 + 2\lambda_2$$

$$1 + \lambda_1 + 2\lambda_2 \leq 3$$

$$\Delta y_1 \leq 3\lambda_1$$

$$\Delta y_2 \leq 3\lambda_2$$
\[\Delta y_1 = y - s_1\]
\[\Delta y_2 = y - s_2\]
\[s_1 \leq 3(1 - \lambda_1)\]
\[s_2 \leq 3(1 - \lambda_2)\]
\[z \geq \Delta y_1 + 2\Delta y_2\]
\[z \geq 3x + (y - 3) + (\Delta y_1 - 3\lambda_1) + 2(\Delta y_2 - 3\lambda_2)\]
\[z \leq y + \Delta y_1 + 2\Delta y_2\]
\[z \leq 3x + (\Delta y_1 - 3\lambda_1) + 2(\Delta y_2 - 3\lambda_2)\]
\[\lambda \in \{0, 1\}^2\]
\[\Delta y \in [0, 3]^2\]
\[s \in [0, 3]^2\]
\[(x, y) \in [0, 3] \times [0, 3].\]

The feasible point for the relaxation \(x = 3, y = 3, z = 9, \lambda = (1, 0.5), \Delta y = (3, 1.5), \) and \(s = (0, 1.5)\) is a fractional extreme point, showing that the formulation is not ideal.

**Appendix B: Proof of Proposition 1** First, we present a more general lemma.

**Lemma 1.** If there do not exist polyhedra \(\{Q^i\}_{i=1}^d\) with \(d' < d\) and \(\bigcup_{i=1}^d P_i = \bigcup_{i=1}^{d'} Q^i\), then any binary MIP formulation for \(\bigcup_{i=1}^d P_i\) must have at least \([\log_2(d)]\) binary variables.

**Proof** Presume that formulation \(F\) takes the form (7). For each \(h \in \{0, 1\}^{n_3}\), consider the preimage \(\operatorname{Pre}(h) \equiv \{x \in \mathbb{R}^{n_1} : \exists y \in \mathbb{R}^{n_2} \text{ s.t. } (x, y, h) \in F\}\); it is clear from the definition of \(F\) that this set is polyhedral. Furthermore, we need that \(\bigcup_{h \in \{0, 1\}^{n_3}} \operatorname{Pre}(h) = \bigcup_{i=1}^d P_i\). If the condition holds, then we have that the cardinality of the index set of the left side \(2^{n_3}\) must be at least \(d\) which implies the result. \(\square\)

We may apply this lemma in the case where our irredundancy assumption holds.

**Proof** For each \(S \in \mathcal{S}\), take \(\lambda^S_i = \frac{1}{|S|} \sum_{v \in S} e^v_i\), where \(e^v \in \{0, 1\}^d\) is the unit vector for component \(v\). Assume that there is some \(Q^i\) such that \(\lambda^S, \lambda^{S'} \in Q^i\) for two \(S, S' \in \mathcal{S}\). By convexity, \(\frac{1}{2}(\lambda^S + \lambda^{S'}) \in Q^i\) as well. But this implies that there is a point in \(Q^i\) with support over \(S \cup S'\), which would violate the irredundancy of \(S\) and \(S'\), meaning that this cannot yield a formulation for \(\bigcup_{i=1}^d P_i\). Therefore, each of the \(d\) points \(\lambda^S\) must be contained uniquely in some \(Q^i\), and we may apply Lemma 1 for the result. \(\square\)

**Appendix C: Proof of Proposition 4**

**Proof** The interpretation of the decision variables is:

\[x^r_j = 1 \begin{bmatrix} r \in L_j^i \end{bmatrix}\]
\[z^s_j = 1 \begin{bmatrix} x^r_j + x^s_j = 1 \end{bmatrix}.\]  

(20a)

That is, \(x^r_j = 1\) iff we construct level \(j\) to have \(r \in J\) on the left-hand dichotomy, and \(z_j^{r,s} = 1\) iff level \(i\) separates infeasible edge \((r, s) \in \bar{E}\). Note that we are using the covering assumption, and so \(r \in L_j \cup R_j^i\) for each \(j\). From (16a), we have \(z^s_j = 1 \begin{bmatrix} x^r_j + x^s_j = 1 \end{bmatrix}\) holds for any 0/1 values of \(x\).

To show that a feasible solution maps to an IB scheme, consider some \((x, z)\) feasible for (16), and the corresponding sets \(L_j = \{r \in J : x^r_j = 0\}\) and \(R_j^i = \{r \in J : x^r_j = 1\}\) for each \(j \in [t]\). From (16b) we have that there is at least one level \(j\) that separates each \((r, s) \in \bar{E}\) (infeasible edge), and by (16c) that no level \(j\) may separate \((r, s) \in E\) (feasible edge).

To show that the existence of an IB scheme implies that (16) is feasible, you may consider the proposed solution (20) and see that it is feasible for (16). \(\square\)
Appendix D: Proof of Proposition 8

Proof We would like to apply a logarithmically-sized SOS2 construction to each of the \( \eta \) directions. For each \( i \in [\eta] \), take \( \{v^{i,j}\}_{j=1}^{d_i} \subseteq \{0,1\}^{[log_2(d_i)]} \) as a Gray code\(^{10}\). Then define the sets

\[
A^{i,j} = \left\{ (x,y) \in J : \exists \gamma \text{ s.t. } x_i = h^i_\gamma, v^{i,j}_j = v^{i,j+1}_j = 0 \right\}
\]

\[
B^{i,j} = \left\{ (x,y) \in J : \exists \gamma \text{ s.t. } x_i = h^i_\gamma, v^{i,j}_j = v^{i,j+1}_j = 1 \right\}
\]

for each \( i \in [\eta] \) and \( j \in [[log_2(d_i)]] \). From the formulation in Section 6.1, for each \( i \in [\eta] \), the sets \( \{A^{i,j}, B^{i,j}\}_{j=1}^{[log_2(d_i)]]} \) will separate exactly those edges \((x,y),(x',y')\) \( \in \bar{E} \) such that \( |x_i - x'_i| > 1 \). Taking the union over all \( i \in [\eta] \), this separates exactly those edges \((x,y),(x',y')\) \( \in \bar{E} \) where \( ||x-x'||_\infty > 1 \), i.e. exactly the set of infeasible edges \( \bar{E} \), yielding a valid biclique cover. □

\(^{10}\) Where, again, \( v^{i,0} \equiv v^{i,1} \) and \( v^{i,d_i+1} \equiv v^{i,d_i} \) for notational simplicity.