Branch-and-Bound versus Lift-and-Project Relaxations in Combinatorial Optimization

Gérard Cornuéjols[*1] and Yatharth Dubey[†2]

[1]Tepper School of Business, Carnegie Mellon University
[2]Industrial and Enterprise Systems Engineering, University of Illinois at Urbana-Champaign

November 7, 2023

Abstract

In this paper, we consider a theoretical framework for comparing branch-and-bound with classical lift-and-project hierarchies. We simplify our analysis by streamlining the definition of branch-and-bound. We introduce "skewed k-trees" which give a hierarchy of relaxations that is incomparable to that of Sherali-Adams, and we show that it is much better for some instances. We also give an example where lift-and-project does very well and branch-and-bound does not. Finally, we study the set of branch-and-bound trees of height at most k and effectively "squeeze" their effectiveness between two well-known lift-and-project procedures.

1 Introduction

In integer programming, branching and cutting are two basic algorithmic strategies at the heart of current solvers. For any application or subproblem at hand, deciding whether to branch or to cut can drastically impact the computing time. Understanding this trade-off has been of interest for a long time and is addressed, for example, in papers such as [BCDSJ23] and the references therein. In this paper, we consider a new theoretical framework for comparing these two strategies.

Many combinatorial optimization problems can be written in the following way: $\max\{cx \; : \; x \in P \cap \{0,1\}^n\}$, where $c \in \mathbb{R}^n$, $P := \{x \in [0,1]^n \; : \; Ax \geq b\}$, with $A \in \mathbb{R}^{m \times n}, b \in \mathbb{R}^m$. Define $P_I := \text{conv}(P \cap \{0,1\}^n)$. A fundamental goal in integer programming and combinatorial optimization is to obtain a relaxation Q such that $P_I \subseteq Q \subseteq P$, where Q is as "close" to P_I as possible; see [CCZ14, Sch98, Sch03, WN99] for more on integer programming and combinatorial optimization.

One way to obtain such relaxations is via an *extended formulation*, i.e. a polyhedron $Q_E := \{(x, y) \; : \; A'x + B'y \geq b'\} \subseteq \mathbb{R}^{n+p}$, where $A' \in \mathbb{R}^{r \times n}, B' \in \mathbb{R}^{r \times p}, b' \in \mathbb{R}^r$, such that $P_I \subseteq \text{proj}_x(Q_E) \subseteq P$. When r, p

---

are polynomial in n, we refer to Q_E as a *compact extended formulation*; see [CCZ10] for more on extended formulations. Hierarchies of extended formulations can be obtained by iterating the procedures of Balas-Ceria-Cornuejols (BCC) [BCC93], and Lovasz-Schrijver (LS) [LS91], or by applying the Sherali-Adams procedure at various levels (SA) [SA90]; see [Lau03] for a comparison between the latter two procedures and [CL01] for a broader comparison of cut-operators.

Another way of generating extended formulations arises from the branch-and-bound method (BB) [LD10]. Consider a partial enumeration tree $\mathcal{T}$ where the leaves represent subproblems $P_v := P \cap \{x : C_v\}$ where C_v denotes the set of branching constraints from the root to node v. The set $\bigcup_{v \in \text{leaves}(\mathcal{T})} P_v$ is contained in P and contains all feasible integer solutions. In other words, letting $\mathcal{T}(P) := \text{conv}\left(\bigcup_{v \in \text{leaves}(\mathcal{T})} P_v\right)$, we have $P_I \subseteq \mathcal{T}(P) \subseteq P$. The polytope $\mathcal{T}(P)$ has an extended formulation by a theorem of Balas [Bal85] and this formulation is polynomial size if the number of leaves of $\mathcal{T}$ is polynomial.

The potential of extended formulations arising from BB is made evident by the following example.

Example 1. *Let K_n be the complete graph on n vertices V and let $P = \{x \in \mathbb{R}^n_+ : x_u + x_v \leq 1 \ \forall \ u, v \in V\}$ be the corresponding fractional stable set polytope. Observe that the inequality $\sum_{v \in V} x_v \leq 1$ is valid for P_I.*

BB is able to obtain an extended formulation Q_{BB}, with $O(n^2)$ variables and $O(n^3)$ constraints, such that $\sum_{v \in V} x_v \leq 1$ is valid for Q_{BB}. This is because there is a BB tree $\mathcal{T}$ of size $O(n)$ such that $\mathcal{T}(P) = P_I$. In particular, observe that after branching on some variable x_v, the side fixing $x_v = 1$ only contains the point where all other $x_u = 0$ for $u \neq v$; therefore, the desired inequality is valid for any tree where each variable is branched on once and only once, and a node is branched on only if it is either the root or a child corresponding to the branch $x_v = 0$ for some $v \in V$ (i.e. a node with a constraint $x_v = 1$ for any $v \in V$ is never branched on); furthermore, all leaves of such a tree are integral.

By contrast, SA produces an extended formulation Q_{SA} with exponentially many variables and constraints for $\sum_{v \in V} x_v \leq 1$ to be valid for Q_{SA}. In particular, the inequality $\sum_{v \in V} x_v \leq 1$ is not valid for $SA^{n-3}(P)$, which has an exponential size extended formulation. This is because the point $\frac{1}{k+2}\mathbf{1} \in SA^k(P)$, for any $1 \leq k \leq n-2$ (see Section 6.1 of [Lau03]).

This example illustrates the fact that extended formulations derived from BB can outperform those generated by the SA hierarchy. The goal of this work is to better understand how BB compares to lift-and-project on other families of discrete optimization problems. Specifically, we seek to identify when one approach (BB or lift-and-project) provides a tight compact extended formulation while the other does not.

We note that [ST10] studied the Gomory-Chvàtal (GC) operator [Gom10, Chv73] in the same spirit: they showed that GC significantly outperforms SA for the maximum matching problem in k-uniform hypergraphs; however, GC performs as poorly as SA for max cut, unique label cover, and k-CSP_q. They concluded that the "positive result gives strong motivation for studying GC cuts as an algorithmic technique." This paper shows that a similar conclusion can be made for studying BB trees.

In Section 2, we introduce branch-and-bound formally, in the form used in this paper. The polytope $\mathcal{T}(P)$ is derived from the corresponding enumeration tree. We also discuss the different variants of lift-and-project that we will use as a comparison, and we prove technical results that will be used later in the paper.

In Section 3, we introduce a special form of enumeration tree called "skewed k-tree". We show that, for fixed k, these trees have polynomial size. We identify two combinatorial optinization problems where skewed k-tree solve the problem to optimality. Namely, we show that, for the maximum clique problem

on sparse graphs, and the uniform knapsack problem with small capacity, we can construct a polynomial size skewed k-tree such that $\mathcal{T}(P) = P_I$. By contrast, for these two combinatorial optimization problems, any polynomial-size formulation in the Sherali-Adams hierarchy has a projection $SA^k(P)$ into the x-space that is strictly contains P_I.

In Section 4, we prove a similar result for combinatorial problems defined by no-good constraints, using a different family of enumeration trees. We construct a polynomial size BB tree $\mathcal{T}$ such that $\mathcal{T}(P) = P_I$ but even when there is just one no-good point, any SA formulation that produces P_I has an exponential number of variables and constraints.

In Section 5, we give an example where lift-and project does very well while banch-and-bound does poorly. Specifically, we exhibit a polytope P such that no branch-and-bound tree of size less than $2^{(n-1)/6}$ produces P_I, while lift-and-project generates P_I in two rounds.

In Section 6, we consider the set of BB trees with height at most k. We show that this family of trees produces a bound at least as strong as any sequential convexification of k variables as per the BCC procedure, but no stronger than k iterations of the canonical lift-and-project procedure. In other words, we show that the bound obtained from this family of BB trees is squeezed between two classical lift-and-project bounds.

2 Extended Formulations

We begin by formally defining branch-and-bound in the form we study here; this comes almost directly from the interpretation of branch-and-bound in Section 2 of [DDM23], but here we only consider trees constructed via variable disjunctions (as opposed to general disjunctions). We then introduce the lift-and-project operators of interest in this paper.

2.1 Relaxations Based on Branch-and-Bound

We simplify our analysis of branch-and-bound by removing two conditions typically assumed of branch-and-bound: (i) the requirement that the partitioning into two subproblems (which correspond to the child nodes of the given node) is done in such a way that the optimal LP solution of the parent node is not included in either subproblem, and (ii) branching is not done on pruned nodes. By removing these conditions, we can talk about a BB tree independent of the underlying polytope – it is just a binary tree (i.e. each node has 0 or 2 child nodes). The root-node has an empty set of *branching constraints*. If a node has two child nodes, these are obtained by applying a disjunction $x_j = 0 \lor x_j = 1$ for some $j \in [n]$, where each child node adds one of these constraints to its set of branching constraints together with all the branching constraints of the parent node. Finally, note that since a BB tree is a binary tree, the total number of nodes of a BB tree with N leaf-nodes is $2N - 1$.

Definition 1. *Given a branch-and-bound tree $\mathcal{T}$, applied to a polytope $P \subseteq [0, 1]^n$, and a node v of the tree:*

- *We denote the number of nodes of the branch-and-bound tree $\mathcal{T}$ by $|\mathcal{T}|$. This is what is termed* the size *of this tree.*

- *We denote by C_v the set of branching constraints of v (as explained above, these are the constraints added by the branch-and-bound tree along the path from the root node to v).*

- For any leaf $v \in leaves(\mathcal{T})$, we denote $J_v^0 := \{j : \{x_j = 0\} \in C_v\}$ and $J_v^1 := \{j : \{x_j = 1\} \in C_v\}$.

- We say that the height of a node $v \in \mathcal{T}$ is the number of its branching constraints; formally, $h(v) = |C_v|$ for any $v \in \mathcal{T}$. The height of $\mathcal{T}$ is $h(\mathcal{T}) := \max_{v \in \mathcal{T}} h(v)$. Observe that $h(root) = 0$, and thus $h(\mathcal{T}) \leq n$.

- We call the feasible region defined by the LP relaxation P and the branching constraints at node v the atom of this node; formally, we denote $P_v = P \cap \{x : C_v\}$ is the atom corresponding to v.

- We let $\mathcal{T}(P)$ denote the convex hull of the union of the atoms corresponding to the leaves of $\mathcal{T}$ when applied to polytope P, i.e.,

$$\mathcal{T}(P) = conv\left(\bigcup_{v \in leaves(\mathcal{T})} P_v\right).$$

- For $x^* \in P \setminus P_I$, we say that $\mathcal{T}$ separates x^* if $x^* \notin \mathcal{T}(P)$.

Theorem 1 (Balas [Bal85], see Theorem 5.1 of [CCZ10]). *There is an extended formulation for $\mathcal{T}(P)$ with $O(|\mathcal{T}|(n+1))$ variables and $O(|\mathcal{T}|m)$ constraints.*

We will study different types of BB trees in Sections 2-6. In Section 3 we introduce skewed k-trees and we show that they can be advantageous for generating the relaxation $\mathcal{T}(P)$. In Section 4 we exhibit a BB tree that works well for combinatorial problems defined by no-good constraints. In Section 5 we analyze BB trees with bounded number of leaves and in Section 6 BB trees with bounded height.

2.2 Lift-and-Project

In this section we review three classical lift-and-project operators and we present technical lemmas that will be useful in subsequent sections.

First we introduce the canonical lift-and-project operator; see Section 5.4 of [CCZ14].

$$L(P) = \bigcap_{i \in [n]} \text{conv}((P \cap \{x : x_i = 0\}) \cup (P \cap \{x : x_i = 1\}))$$

Let $L^0(P) := P$ and $L^k(P) := L\left(L^{k-1}(P)\right)$ for any $k \geq 1$.

We will use the following technical results.

Lemma 1. $L^k(P) \cap \{x : x_j = a\} = L^k(\{x \in P : x_j = a\})$ for any $j \in [n], k \in [n], a \in \{0, 1\}$.

Proof. First, we prove the lemma for $k = 1$. Also, we assume $a = 0$ as the argument for $a = 1$ follows very easily.

First, we show the easier containment, $L(P \cap \{x : x_j = 0\}) \subseteq L(P) \cap \{x : x_j = 0\}$. Clearly $L(P \cap \{x : x_j = 0\}) \subseteq \{x : x_j = 0\}$ and it is also easy to see that $L(P \cap \{x : x_j = 0\}) \subseteq L(P)$ by monotonicity.

Now we show $L(P) \cap \{x : x_j = 0\} \subseteq L(P \cap \{x : x_j = 0\})$. Let $\bar{x} \in L(P) \cap \{x : x_j = 0\}$. It suffices to prove the claim that $\bar{x} \in \text{conv}((P \cap \{x : x_j = 0, x_i = 0\}) \cup (P \cap \{x : x_j = 0, x_i = 1\}))$ for each i. The claim is true

4

for $i = j$ since $\bar{x} \in L(P) \cap \{x : x_j = 0\}$ implies $\bar{x} \in P \cap \{x : x_j = 0\}$. Now consider $i \neq j$. The claim is true when $\bar{x}_i = 0$ or $\bar{x}_i = 1$ because $\bar{x} \in L(P)$ implies $\bar{x} \in P$ and therefore $\bar{x} \in P \cap \{x : x_j = 0, x_i = 0\}$ or $\bar{x} \in P \cap \{x : x_j = 0, x_i = 1\}$. So we may assume $0 < \bar{x}_i < 1$. Since $x \in L(P)$, there is some $x^0 \in \{x \in P : x_i = 0\}$ and $x^1 \in \{x \in P : x_i = 1\}$ such that $x \in \operatorname{conv}(x^0, x^1)$. Moreover, it must hold that $x_j^0 = x_j^1 = 0$, since otherwise would imply $x_j > 0$. This proves the claim.

We now prove the lemma inductively. Assume $L^k(P) \cap \{x : x_j = a\} = L^k(\{x \in P : x_j = a\})$. Then $L^{k+1}(P) \cap \{x : x_j = a\} = L(L^k(P)) \cap \{x : x_j = a\} = L(L^k(P) \cap \{x : x_j = a\}) = L(L^k(P \cap \{x : x_j = a\})) = L^{k+1}(P \cap \{x : x_j = a\})$, where the first and last equality follow from the definition of L^k, the second from the base case of the induction proven above, and the third from the inductive hypothesis. $\square$

Lemma 2 (Lemma 3.1 of [DGM15]). *Consider $(a, b) \in \mathbb{R}^n \times \mathbb{R}$ with $a \neq 0$ and let $s^1, ..., s^n$ be affinely independent points in $\{x \in \mathbb{R}^n : ax = b\}$. Consider $b' > b$ and let R be a bounded and non-empty subset of $\{x \in \mathbb{R}^n : ax \geq b'\}$. Then, there exists a point $x \in \bigcap_{r \in R} \operatorname{conv}(s^1, ..., s^n, r)$ satisfying the strict inequality $ax > b$.*

In the statement and proof of the next theorem, we use the following notation. For any $J \subseteq [n]$, let $P^J := P \cap \{x : x_j = 0 \; \forall \; j \in J\}$.

Theorem 2. *Let $P \subseteq [0, 1]^n$ be a polytope. Assume P_I is full-dimensional, and let $cx \geq \delta$ be a facet of P_I. Let $t \in [n-1]$. For all $J \subset [n]$ with $|J| = n - t$, assume that P^J satisfies the following properties.*

1. $\dim\left(\left(P^J\right)_I\right) = n - |J| = t$

2. $\{x \in P^J : cx \geq \delta\}$ is a facet of $\left(P^J\right)_I$

3. there is some $x^J \in P^J$ such that $cx^J < \delta$.

Then, $L^{n-t}(P) \neq P_I$. Indeed, there is some $\bar{x} \in L^{n-t}(P)$ with $c\bar{x} < \delta$.

Proof. We will use the following claim for the proof.

Claim 1. *Consider any $J \subset [n]$ with $|J| \leq n - t$. Properties 1 and 2 of the theorem hold for P^J.*

Proof of Claim 1. We proceed by induction: Let $k \geq t$ and suppose Properties 1 and 2 hold for all $J \subset [n]$ with $|J| = n - k$. Now consider any $J \subset [n]$ with $|J| = n - (k + 1)$. First, we will show that $\{x \in P^J : cx = \delta\}$ is a facet of $\left(P^J\right)_I$. In particular, we will demonstrate $k + 1$ affinely independent points $S_J \subset P \cap \{0, 1\}^n$ such that for any $x \in S_J$, it holds that $cx = \delta$ and $x_j = 0$ for all $j \in J$.

Fix any $i \notin J$ and let $J' = J \cup \{i\}$. Observe that, by the induction hypothesis, there is a set $S_{J'} \subset \{0, 1\}^n$ of k affinely independent points such that for any $x \in S_{J'}$, it holds that $cx = \delta$ and $x_j = 0$ for all $j \in J'$; furthermore, observe that $S_{J'} \subset P^{J'} \subset P^J$. Therefore, it suffices to show that there is a point in $\{0, 1\}^n$ with $x_i = 1$ with $cx = \delta$ and $x_j = 0$ for all $j \in J$, since such a point would be affinely independent from the set $S_{J'}$, since $x_i = 0$ for any $x \in S_{J'}$.

Consider any $i' \notin J'$ (such i' exists since $k \geq 1$ implies $|J'| \leq n - 1$) and let $J'' = J \cup \{i'\}$. Let $S_{J''}$ be defined similarly to $S_{J'}$ and observe that there is a point $x'' \in S_{J''}$ with $x_i'' = 1$, since $\dim\left(P^{J''}\right)_I = k$ and we already have k equations $x_j = 0$ for $j \in J''$. Clearly, since $x'' \in S_{J''}$ it holds that $cx'' = \delta$ and $x_j'' = 0$ for all $j \in J''$. Therefore, $S_{J'} \cup \{x''\}$ is a set of $k + 1$ affinely independent points in $P^J \cap \{0, 1\}^n$ with $cx = \delta$, showing that

P^J satisfies Property 2 of the theorem. The proof showing P^J satisfies Property 1 is exactly the same, but with $S_{J'}$ including an additional point in $\left(P^{J'}\right)_I$ with $cx > \delta$ (clearly this point is affinely independent from the points in the facet). □

We now prove the theorem by induction. Assume for all $J \subset [n]$ such that $|J| = n - k$, there is a point $x^J \in L^{k-t}(P^J)$ with $cx^J < \delta$. The theorem assumes this is true for $k = 0$. Consider any $J \subseteq [n]$ such that $|J| = n - (k+1)$, we will show that there is a point $x^J \in L^{(k+1)-t}(P^J)$ with $cx^J < \delta$. It follows from the definition of $L(\cdot)$ and Lemma 1 that,

$$L^{k-t+1}\left(P^J\right) = \bigcap_{i \in [n]\setminus J} \mathrm{conv}\left(L^{k-t}\left(P^{J\cup\{i\}}\right) \cup L^{k-t}(P \cap \{x \,:\, x_i = 1\})\right).$$

By the induction hypothesis, the i-th term of the intersection contains a point $x^{J\cup\{i\}}$ such that $cx^{J\cup\{i\}} < \delta$ and by Claim 1 it contains a set S_J of $k + 1$ affinely independent points, each satisfying $cx = \delta$. Therefore, by Lemma 2, there is a point $x^J \in \bigcap_{i \in [n]\setminus J} \mathrm{conv}\left(S_J \cup \left\{x^{J\cup\{i\}}\right\}\right)$ such that $cx^J < \delta$ as desired. □

The second lift-and-project operator that we consider in this paper is based on sequential convexification, introduced in Section 2 of [BCC93].

Let $k \in [n]$, and let $i_1, ..., i_k \in [n]$. Theorem 2.2 of [BCC93] gives a method to find a compact extended formulation for the polytope $P_{i_1,...,i_k} = \mathrm{conv}(\{x \in P \,:\, x_j \in \{0, 1\} \,\forall\, j \in \{i_1, ..., i_k\}\})$. In this study, we will consider the following natural operator

$$B^k(P) = \bigcap_{\{i_1,...,i_k\}\subset[n]} P_{i_1,...,i_k}$$

Note that $L(P) = B^1(P)$ and $L^k(P) \subseteq B^k(P)$ for any $k \geq 1$ (this also follows from Theorem 8).

Third, we consider the Sherali-Adams hierarchy; see Section 10.4.1 of [CCZ14]. We denote by $SA^k(P)$ the kth level polytope in the Sherali-Adams hierarchy. We have $P_I = SA^n(P) \subseteq SA^{n-1}(P) \subseteq ... \subseteq SA(P) \subseteq P$.

Remark 1. *The extended formulation for $SA^k(P)$ has $O(\binom{n}{k})$ variables and $O(\binom{n}{k})$ constraints.*

Remark 2. *It is well-known and easy to see that $SA^k(P) \subseteq L^k(P)$ for any $k \geq 1$ and polytope $P \subseteq [0, 1]^n$; see Lemma 10.8 and Theorem 10.11 of [CCZ14].*

3 Skewed k-trees

In this section, we present two problems that branch-and-bound is able to solve in polynomial time, while Sherali-Adams requires an exponential size extended formulation. This capability of branch-and-bound can be understood by considering a specific family of trees.

Definition 2. *Let $k \in \mathbb{Z}_{\geq 1}$ and π be a permutation of $[n]$. We construct a* skewed k-tree $\mathcal{T}^{\pi,k}$ *as follows.*

Let $\mathcal{T}_0$ consist of only the root, i.e. a single node v with $C_v = \emptyset$. Then, for $i = \pi(1), ..., \pi(n)$, do the following. For any leaf $v \in \mathrm{leaves}(\mathcal{T}_{i-1})$, if $|J_v^1| < k$: branch on variable i at v. In other words, $\mathrm{leaves}(\mathcal{T}_i) = \left(\mathrm{leaves}(\mathcal{T}_{i-1}) \setminus \{v \,:\, |J_v^1| < k\}\right) \cup \left\{v^0, v^1 \,:\, C_{v^a} = C_v \cup \{x_i = a\}, a \in \{0, 1\}\right\}$, and set $\mathcal{T}^{\pi,k} = \mathcal{T}_n$.

Theorem 3. *For any permutation π, the skewed k-tree $\mathcal{T}^{\pi,k}$ has size $n^{O(k)}$ and satisfies the following properties*

1. *for each subset $J \subset [n]$ of $k' \leq k$ variables, there is exactly one leaf $v \in \text{leaves}(\mathcal{T}^{\pi,k})$ such that $J_v^1 = J$ and*

2. *for any leaf $v \in \text{leaves}(\mathcal{T}^{\pi,k})$, either $|J_v^1| = k$ or $J_v^0 \cup J_v^1 = [n]$.*

Proof. We begin by showing $|\mathcal{T}^{\pi,k}| = n^{O(k)}$. Let $v^1, v^2 \in \text{leaves}(\mathcal{T})$. Observe that $J_{v^1}^1 \neq J_{v^2}^1$, in particular $|J_{v^1}^1 \bigtriangleup J_{v^2}^1| \geq 1$ (where $A \bigtriangleup B$ is the symmetric difference of the sets A, B): let w be the most recent common ancestor (that of maximum height) of v^1, v^2, and let ℓ be the variable branched on at w (so w has height $\ell - 1$); then, clearly $\ell \in J_{v^1}^1 \bigtriangleup J_{v^2}^1$. Furthermore, by construction, $|J_v^1| \leq k$ for any $v \in \text{leaves}(\mathcal{T}^{\pi,k})$. Therefore, for each $v \in \text{leaves}(\mathcal{T}^{\pi,k})$, there is a distinct size $\leq k$ subset of $[n]$, and there are only $\sum_{i \in [k]} \binom{n}{i} = n^{O(k)}$ such subsets, proving the desired bound on $|\mathcal{T}^{\pi,k}|$.

First, we will show that $\mathcal{T}^{\pi,k}$ satisfies property 2 of the lemma. Since, $|J_v^1| \leq k$ for any $v \in \text{leaves}(\mathcal{T}^{\pi,k})$, we only need to show that, for any $v \in \text{leaves}(\mathcal{T}^{\pi,k})$ such that $|J_v^1| < k$, it must be that $J_v^0 \cup J_v^1 = [n]$. Let v be such a leaf and, for sake of contradiction, let $j' = \min\{j \notin J_v^0 \cup J_v^1\}$ (note that $j' \leq h(v) + 1$, since $|[n] \setminus (J_v^0 \cup J_v^1)| = n - h(v)$) and let w be the ancestor of v at height $j' - 1$ (note that it is possible $w = v$, in the case $j' = h(v) + 1$). Then, $w \in \text{leaves}(\mathcal{T}_{j'-1})$ and $|J_w^1| \leq |J_v^1| < k$, therefore, j' is branched on at w and therefore $j' \in J_v^0 \cup J_v^1$, achieving the desired contradiction.

Finally, we will show that $\mathcal{T}^{\pi,k}$ satisfies property 1 of the lemma. Let $J \subset [n]$ of size $k' \leq k$. We will show that there exists a leaf $v \in \text{leaves}(\mathcal{T}^{\pi,k})$ such that $J_v^1 = J$, which will prove the desired result, since we have already argued that $J_{v^1}^1 \neq J_{v^2}^1$ for distinct leaves v^1, v^2. Let $j' = \max\{j \in J\}$ and let $\bar{v}$ be a node be such that $C_{\bar{v}} = \left(\bigcup_{j \in [j'] \setminus J} \{x_j = 0\} \right) \cup \left(\bigcup_{j \in J} \{x_j = 1\} \right)$. Clearly by construction $\bar{v} \in T_{j'}$ since its parent $\bar{\bar{v}}$ fixes exactly variables $1, ..., j' - 1$ and has $|J_{\bar{v}}^1| < k$. If $k' = k$, then $\bar{v} \in \text{leaves}(\mathcal{T}^{\pi,k})$ and $J_v^1 = J$. If $k' < k$, then there is a leaf $v \in \text{leaves}(\mathcal{T}^{\pi,k})$ with $C_v = C_{\bar{v}} \cup \left(\bigcup_{j'+1 \leq j \leq n} \{x_j = 0\} \right)$. We can see that $v \in \mathcal{T}^{\pi,k}$ since any node $\bar{\bar{v}}$ that is an ancestor of v will have $|J_{\bar{v}}^1| < k$ and therefore will be branched on. $\qquad\square$

Remark 3. *Let $\mathcal{F} = \{\{x : x_{i_1} = ... = x_{i_k} = 1\} : \{i_1, ..., i_k\} \subset [n]\}$ be the set of all faces of $[0,1]^n$ defined by fixing k variables to 1. Suppose for any $F \in \mathcal{F}$, it holds that $P \cap F = \emptyset$ or $P \cap F = (P \cap F)_I$. Let $\mathcal{T}$ be **any** branch-and-bound tree that branches on a node if and only if the node is not empty nor integral, through a similar argument to that of Theorem 3, it holds that $\mathcal{T}(P) = P_I$ and $|\mathcal{T}| = n^{O(k)}$.*

Remark 4. *For any permutation π, the sequence of skewed k-trees $\{\mathcal{T}^{\pi,k} : k = 1, ..., n\}$ corresponds to a hierarchy of relaxations. Let $P \subseteq [0,1]^n$, then observe $\mathcal{T}^{\pi,k'}(P) \subset \mathcal{T}^{\pi,k}(P)$ for any $k < k'$ and $\mathcal{T}^{\pi,n}(P) = P_I$.*

At first glance, this seems similar to the k-th level of Sherali-Adams, which aims to take advantage of "easy" subsets of k variables. However, in the following subsections, we will see classes of instances for which the above property allows branch-and-bound to succeed, while Sherali-Adams is unable to do so. Intuitively, this is because the k-th level of Sherali-Adams may not able to take advantage of inequalities that are valid for $P \cap \{x : C_v\}$ when $v \in \text{leaves}(\mathcal{T})$ has height $> k$, where as in the case of $\mathcal{T}^{\pi,k}$, it is ensured that either k variables are fixed to one or all variables are fixed.

3.1 Uniform Knapsack with Small Capacity

In this section, we apply Theorem 3 to the problem of a uniform knapsack with small capacity.

Theorem 4. *Let $P \subset [0,1]^n$ be such that $\sum_{i \in [n]} x_i \leq k - \epsilon$ is valid for P, for any $\epsilon \in (0,1)$. Let $\mathcal{T}^{\pi,k}$ be a skewed k-tree, then $\mathcal{T}^{\pi,k}(P) = P_I$. Furthermore, there exists a P such that $\sum_{i \in [n]} x_i \leq 2 - \frac{2}{q}$ is valid for P, yet $SA^{\lfloor n(q-2)/(q-1)\rfloor}(P) \neq P_I$, where q can be any integer at least 3.*

Proof. We begin by proving the positive result on branch-and-bound. In particular, we show that for every leaf $v \in \text{leaves}(\mathcal{T}^{\pi,k})$, either its atom is empty (i.e. $P \cap \{x : C_v\} = \emptyset$) or integral (i.e. $P \cap \{x : C_v\}$ has only integral extreme points). Let $v \in \text{leaves}(\mathcal{T}^{\pi,k})$ be such that $|J_v^1| = k$. Then, for any $x \in \{x : C_v\}$, it holds that $\sum_{i \in [n]} x_j = \sum_{j \in J_v^1} x_j = k > k - \epsilon$, which implies $x \notin P$, and therefore $P \cap \{x : C_v\} = \emptyset$. Clearly, for any $v \in \text{leaves}(\mathcal{T}^{\pi,k})$ such that $J_v^0 \cup J_v^1 = [n]$, it holds that $P \cap \{x : C_v\}$ is either empty or a single point in $\{0,1\}^n$. This concludes the proof of $\mathcal{T}^{\pi,k}(P) = P_I$.

Now, we prove the lower bound on Sherali-Adams. Let q be any integer at least 3 and let $P = \{x \in [0,1]^n : \sum_{j \in [n]} qx_j \leq 2(q-1)\}$. Note that $\sum_{i \in [n]} x_i \leq 2 - \frac{2}{q}$ is clearly valid for P, and in particular we can write $P = \{x \in [0,1]^n : \sum_{i \in [n]} x_i \leq 2(\frac{q-1}{q})\}$. Theorem 5 of [KMN11] shows that the point $\mathbf{1} \cdot \frac{2(q-1)/q}{n+(t-1)(q-1)/q} \in SA^t(P)$. Therefore, when $t < n\frac{q-2}{q-1} + \frac{q}{q-1}$, it holds that $\sum_{i \in [n]} x_i \leq 1$ is not valid for $SA^t(P)$, proving the desired result. $\square$

3.2 Max Clique on Sparse Graphs

One measure of graph density that has shown to be relevant in practical applications is the notion of *degeneracy*.

Definition 3 ([LW70]). *A graph G is d-degenerate if each of its subgraphs $G' \subseteq G$ has minimum degree at most d. The degeneracy of G is the smallest d such that G is d-degenerate. Moreover, it is a simple fact that the size k of the maximum clique in G satisfies $k \leq d + 1$.*

Observe that the degeneracy of a forest is 1, of a cycle is 2, of a planar graph is at most 5, and of a K_n is $n - 1$; this demonstrates how degeneracy can be a reasonable measure of density. The parameterization by degeneracy is inspired by [WB20, NBW22]. In particular, Table 2 of [WB20] shows that, for many real-world graphs, the degeneracy is several orders of magnitude smaller than the number of nodes, indicating that further study of algorithms on such graphs is worthwhile. We show that branch-and-bound is more effective for finding cliques on such graphs than Sherali-Adams.

Theorem 5. *Let $G = (V, E)$ be a graph on n vertices with maximum clique size k and degeneracy d. Let $P = \{x \in \mathbb{R}_+^n : x_u + x_v \leq 1 \ \forall \ uv \notin E\}$ be the fractional clique polytope for G. Let $\mathcal{T}^{\pi,k}$ be a skewed k-tree, then $\mathcal{T}^{\pi,k}(P) = P_I$, while $SA^{\lfloor n/(d+1)-3\rfloor}(P) \neq P_I$.*

Proof. We begin by showing $\mathcal{T}^{\pi,k}(P) = P_I$. In particular, we will show that for any leaf $z \in \text{leaves}(\mathcal{T}^{\pi,k})$, its atom is either empty (i.e. $P \cap \{x : C_z\} = \emptyset$) or integral (i.e. $P \cap \{x : C_z\}$ has only integral extreme points). Consider any $z \in \text{leaves}(\mathcal{T}^{\pi,k})$ with $|J_z^1| = k$. Suppose J_z^1 is not a clique. Then, there is some $u, v \in J_z^1$ such that $uv \notin E$. However, $\{x_u = 1\}, \{x_v = 1\} \in C_z$, and therefore $P \cap \{x : C_z\} = \emptyset$ since $\{x : C_z\}$ violates $x_u + x_v \leq 1$. Now suppose J_z^1 is a clique, then it must be a maximum (and therefore maximal) clique. Consider any $u \notin J_z^1$, there is some $v \in J_z^1$ such that $uv \notin E$, and therefore, since $x_u + x_v \leq 1$ for any $x \in P$, it holds that $x_u = 0$ for any $x \in P \cap \{x : C_z\}$. This implies that $P \cap \{x : C_z\}$ is integral (containing only the point x such that $x_j = 1$ for $j \in J_z^1$ and $x_j = 0$ otherwise). We have shown that any leaf $z \in \text{leaves}(\mathcal{T}^{\pi,k})$

with $|J_z^1| = k$ is either empty or integral, and clearly any leaf $z \in \text{leaves}(\mathcal{T}^{\pi,k})$ with $J_v^0 \cup J_v^1 = [n]$ is either empty or integral, implying the desired result.

For the proof of the lower bound for Sherali-Adams, we will need the following observation of [LW70]: a graph with degeneracy d admits a vertex ordering $(v_1, v_2, ..., v_n)$ in which each vertex v_i has at most d neighbors to its right (i.e. $|N(v_i) \cap \{v_{i+1}, ..., v_n\}| \leq d$, where $N(v)$ is the open-neighborhood of v). Let $(v_1, ..., v_n)$ be such an ordering, and let $N^r(v_i) = N(v_i) \cap \{v_{i+1}, ..., v_n\}$ be the right-neighborhood of v_i.

We will now show that $SA^{\frac{n}{d+1}-3}(P) \neq P_I$. In particular, observe that G has an stable set of size at least $\frac{n}{d+1}$. We can construct such an stable set, call it S, greedily: start with $S = \{v_1\}$, then choose the minimum index variable v_j such that $v_j \notin \bigcup_{v \in S} N^r(v)$ and add it so that $S = S \cup \{v_j\}$. Observe that after choosing k variables, at most $kd + k$ variables are ruled out (i.e. $|S \cup \bigcup_{v \in S} N^r(v)| \leq k + kd$). Therefore, for $k < \frac{n}{d+1}$, there is still a variable that can be added to S. Now, let S be a maximum stable set G of size at least $\frac{n}{d+1}$. Section 6.1 of [Lau03] shows that $\frac{1}{t+2}\mathbf{1} \in SA^t(P)$, and therefore $\frac{1}{n/(d+1)-1}\mathbf{1} \in SA^{\frac{n}{d+1}-3}(P)$. Therefore, the inequality $\sum_{v \in S} x_v \leq 1$ is not valid for $SA^{\frac{n}{d+1}-3}(P)$, while it is clearly valid for P_I. $\qquad\square$

4 Polynomially Many No-Good Constraints

In this section, we give another family of instances where the branch-and-bound tree is of polynomial size, but the Sherali-Adams relaxation is not.

Theorem 6. *Let P be defined by a set of no-good constraints: let $S \subseteq \{0, 1\}^n$ and $P = \{x \in [0, 1]^n : \sum_{j : s_j = 0} x_j + \sum_{j : s_j = 1}(1 - x_j) \geq \frac{1}{2} \; \forall \, s \in S\}$. There is a branch-and-bound tree $\mathcal{T}$ of size at most $3n|S|$ such that $\mathcal{T}(P) = P_I$. Furthermore, $SA^{n-1}(P) \neq P_I$ when $|S| = 1$ and $L^{n-1}(P) \neq P_I$ if there exists $z \in S$ such that $\{z' : \|z' - z\|_1 = 1\} \cap S = \emptyset$.*

Proof. We will describe a branch-and-bound tree $\mathcal{T}$ as in the theorem. Let each $s \in S$ define a leaf of $\mathcal{T}$: for each $s \in S$, there is a $v^s \in \text{leaves}(\mathcal{T})$ such that $\{x : x_j = s_j\} \in C_{v^s}$ for all $j \in [n]$ (i.e. $P \cap C_{v^s} = \{s\}$). Clearly there exists such a tree, since a complete binary tree with 2^n leaves is such an example. Let $u \in \mathcal{T}$ be a node that is not an ancestor of any of the leaves $\{v^s : s \in S\}$. We claim that the atom of u is integral (i.e. $P \cap C_u$ has integer extreme points), and so these are leaves of $\mathcal{T}$. Therefore, each internal node of $\mathcal{T}$ is an ancestor of some leaf v^s, and since there are $|S|$ such leaves, there are at most $n|S|$ internal nodes, completing this portion of the proof.

We now prove the claim that the atom of u is integral for any u that is not an ancestor of some v^s. Observe that $\{x \in [0, 1]^n : C_u\}$ defines some face of $[0, 1]^n$ that is disjoint from S, i.e. $\{x \in [0, 1]^n : C_u\} \cap S = \emptyset$. Then, any extreme point z of $\{x \in [0, 1]^n : C_u\}$ is such that $z \in \{0, 1\}^n \setminus S$. By the definition of P, it follows $\{0, 1\}^n \setminus S \subseteq P$, and therefore $z \in P \cap \{0, 1\}^n$, concluding the proof of the desired claim.

Finally, consider the case $S = \{0\}$. Proposition 19 of [Lau03] shows that the point $\mathbf{1} \cdot \frac{1}{2n-t} \in SA^t(P)$, implying $\sum_{i \in [n]} x_i \geq 1$ is not valid for $SA^{n-1}(P)$, which in turn shows $SA^{n-1}(P) \neq P_I$ as desired. Furthermore, suppose there is some $z \in S$ such that $\{z' : \|z' - z\|_1 = 1\} \cap S = \emptyset$, by symmetry we can assume $z = 0$. Then, $\sum_{j \in n} x_j \geq 1$ is a facet of P_I, and the corresponding properties of Theorem 2 are satisfied with $t = 1$: let J be any subset of $n - 1$ variables, then $x_j \geq 1$ is a facet of $(P^J)_I$ where $j = [n] \setminus J$ and $\frac{1}{2}e^j \in P^J$ (vector with all zeros, except a half in the j-th component) is a point violating this facet. Then, Theorem 2 implies the final statement of the theorem. $\qquad\square$

5   Limits of Branch-and-Bound

In this section, we rule out general statements on the relative strength of branch-and-bound. For example, perhaps we would like to say something like the following. If lift-and-project does really well (e.g. $L^2(P) = P_I$), there is some *small* branch-and-bound tree that separates each $x \in P \setminus P_I$. Below, we show that statements of this form are not true.

Theorem 7. *Let n be any nonnegative integer such that $n - 1$ is divisible by 6. Let $\mathbb{T}_k^\ell$ be the set of branch-and-bound trees with at most k leaves. There is a polytope $P \subseteq [0, 1]^n$ such that*

$$P_I = L^2(P) \subsetneq \bigcap_{\mathcal{T} \in \mathbb{T}_{2^{(n-1)/6}}^\ell} \mathcal{T}(P).$$

In other words, there is a point $x \in P \setminus P_I$ such that no branch-and-bound tree of size $\leq 2^{(n-1)/6}$ separates x, while lift-and-project retrieves the integer hull in two rounds.

Proof. We begin by defining the polytope of interest, P. Let $P \cap \{x \in \mathbb{R}^n : x_n = 0\} = \{x \in \{0, 1\}^n : x_n = 0\}$. Further, let $P \cap \{x \in \mathbb{R}^n : x_n = 1\} = Q$, where Q is the fractional stable set polytope defined below, intersected with the halfspace $\{x \in \mathbb{R}^{n-1} : \sum_{i \in [n-1]} x_i \geq \frac{n-1}{3} + \frac{1}{2}\}$. Let $m = \frac{n-1}{3}$, we define $G = (V, E)$ to be the disjoint union of m triangles, in particular, $V = \bigcup_{i \in [m]} \{ia, ib, ic\}$ and $E = \bigcup_{i \in [m]} \{\{ia, ib\}, \{ib, ic\}, \{ia, ic\}\}$; so we define $Q = \{x \in \mathbb{R}_+^{n-1} : x_u + x_v \leq 1 \ \forall \ \{u, v\} \in E, \sum_{v \in V} x_v \geq m + \frac{1}{2}\}$, and it should be clear that $Q \cap \{0, 1\}^{n-1} = \emptyset$.

First, we show $L^2(P) = P_I$. By definition and Lemma 1, $L^2(P) \subseteq \text{conv}((L(P) \cap \{x : x_n = 0\}) \cup (L(P) \cap \{x : x_n = 1\})) = \text{conv}(L(\{x \in P : x_n = 0\}) \cup L(\{x \in P : x_n = 1\}))$. We will show $L(\{x \in P : x_n = 1\}) = \emptyset$, which along with the fact that $\{x \in P : x_n = 0\} = P_I$, will prove the claim. To see that $L(\{x \in P : x_n = 1\}) = \emptyset$, observe that $x_{ia} + x_{ib} + x_{ic} \leq 1$ is valid for both $Q \cap \{x : x_{ia} = 0\}$ and $Q \cap \{x : x_{ia} = 1\}$, therefore, each of the m such inequalities is valid for $L(Q)$, implying that $\sum_{v \in V} x_v \leq m$ is valid for $L(Q)$. This, along with the fact that $\sum_{v \in V} x_v \geq m + \frac{1}{2}$ is clearly valid for $L(Q)$ proves the desired claim.

Now we will show that there is some $x \in P \setminus P_I$ such that $x \in \mathcal{T}(P)$ for any $\mathcal{T} \in \mathbb{T}_{2^{(n-1)/6}}^\ell$ (note that $\frac{n-1}{6} = \frac{m}{2}$). Observe that if there is a leaf z of $\mathcal{T}$ such that $\{x : x_n = 0\}$ is not in C_z and there is some $i \in [m]$ such that no variable in $\{ia, ib, ic\}$ is fixed by a branching constraint of z, then there is a point $x \in P \cap \{x : C_z\}$ with $\sum_{v \in V} x_v \geq m + \frac{1}{2}$ (for example, the point setting $x_{ia} = x_{ib} = x_{ic} = 1/2$ and x_{ja}, x_{jb}, x_{jc} to some integer, feasible value for all $j \neq i$). Let $w \in \mathcal{T}$ be a node of minimum height that includes $\{x : x_n = 1\}$ as a branching constraint (i.e. $w = \arg\min\{h(w) : \{x : x_n = 1\} \in C_w\}$). It is clear that its sibling, call it w', is a node of minimum height that includes the constraint $\{x : x_n = 0\}$. Now we will consider two cases: in the first, suppose x_n is fixed in a branching constraint in the first $m/2$ levels of the tree, i.e. $h(w) \leq \frac{m}{2}$. Therefore, at most $m/2$ variables in V have been fixed in C_w. Since $\mathcal{T} \in \mathbb{T}_{2^{m/2}}^\ell$, the subtree rooted at w must have $\leq 2^{m/2} - 1$ leaves, and so there is a leaf w'' of the subtree with less than $m/2$ branching constraints on the path from w to w'' (i.e. $|C_{w''} \setminus C_w| \leq \frac{m}{2} - 1$), which implies that at most $m - 1$ variables in V have been branched on and therefore there is some $i \in [m]$ such that no variable in $\{\{ia, ib\}, \{ib, ic\}, \{ia, ic\}\}$ is fixed in $C_{w''}$. Therefore, there is a point $x \in P \cap C_{w''}$ such that $\sum_{v \in V} x_v \geq m + \frac{1}{2}$ and $x_n = 1$. Now, consider the cases where $h(w') \geq \frac{m}{2} + 1$, or x_n is never branched on in $\mathcal{T}$. Since $\mathcal{T} \in \mathbb{T}_{2^{m/2}}^\ell$, there must be a leaf z' of $\mathcal{T}$ of height at most $\frac{m}{2}$; notice x_n is not fixed by any constraint in $C_{z'}$. Clearly, there is a point $x \in P \cap C_{z'}$ such that $\sum_{v \in V} x_v \geq m + \frac{1}{2}$ and $x_n = 1$. Therefore, for any $\mathcal{T} \in \mathbb{T}_{2^{m/2}}^\ell$, there is some $x^{\mathcal{T}} \in \mathcal{T}(P)$ with $x_n = 1$. Finally, to show that there is some $\bar{x} \in \bigcap_{\mathcal{T} \in \mathbb{T}_{2^{m/2}}^\ell} \mathcal{T}(P)$, we apply Lemma 2 with $a = e_n, b = 0, b' = 1$ and $s^1, ..., s^n$ being

10

the points $0, e^1, ..., e^{n-1}$, for example (there are several choices of n affinely independent points in $\{0, 1\}^{n-1}$), $R = \bigcup_{\mathcal{T} \in \mathbb{T}^\ell_{2^{m/2}}} x^{\mathcal{T}}$ where $x^{\mathcal{T}}$ is defined as above, and $r = x^{\mathcal{T}}$. The Lemma guarantees that such an $\bar{x}$ exists and $\bar{x}_n > 0$, and therefore $\bar{x} \in P \setminus P_I$. $\qquad \square$

6 Trees with Bounded Height

In this section, we introduce, and analyze, an operator that captures the strength of branch-and-bound trees with bounded height.

Definition 4. *We denote $\mathbb{T}^h_k = \{\mathcal{T} \ : \ h(\mathcal{T}) \le k\}$ as the set of BB trees with height at most k.*

Definition 5. *We define the height k branch-and-bound operator as follows*

$$T^k(P) = \bigcap_{\mathcal{T} \in \mathbb{T}^h_k} \mathcal{T}(P)$$

The main result of this section is that the operator $T^k(\cdot)$ can be "squeezed" between two natural lift-and-project operators: the Balas-Ceria-Cornuéjols sequential convexification and the canonical lift-and-project. This is formalized by the following theorem.

Theorem 8. *Let $P \subseteq [0, 1]^n$ be a polytope. Then,*

$$L^k(P) \subsetneq T^k(P) \subsetneq B^k(P)$$

for all $k \in [n]$. Moreover, for any n such that $n - 1$ is divisible by 6, there is a polytope $P \subseteq [0, 1]^n$ such that $P_I = L^2(P) \subsetneq T^{(n-1)/6}(P)$.

Remark 5. *The height-k operator $T^k(\cdot)$ and the skew k-tree operator $\mathcal{T}^{\pi,k}(\cdot)$ are incomparable in general. The results of Section 3 along with Theorem 8 show that there exist P such that $\mathcal{T}^{\pi,k}(P) \subsetneq T^k(P)$. It is easy to construct a P such that $T^k(P) \subsetneq \mathcal{T}^{\pi,k}(P)$: suppose $P = \{x \ : \ x_1, ..., x_k = 1, x_{k+1}, ..., x_n = \frac{1}{2}\}$, then note that $\mathcal{T}^{\pi,k}(P) = P$ while $T^k(P) = \emptyset$.*

We prove the above theorem in three separate parts.

Lemma 3. *$T^k(P) \subsetneq B^k(P)$ for all $k \in [n]$.*

Proof. We will start by showing that $T^k(P) \subseteq B^k(P)$ for any $k \in [n]$. In particular, we will show that for any subset $\{i_1, ..., i_k\} \subset [n]$, there is a tree $\mathcal{T} \in \mathbb{T}^h_k$ such that $\mathcal{T}(P) \subseteq P_{i_1,...,i_k}$. Specifically, consider a complete tree $\mathcal{T}$ of height k that branches only on variables $\{i_1, ..., i_k\}$ (i.e. each of the 2^k leaves of $\mathcal{T}$ fixes the variables $\{i_1, ..., i_k\}$ to one of the 2^k possible settings). Let z be an extreme point of $\mathcal{T}(P)$, so z is in the atom of a leaf $v \in \mathrm{leaves}(\mathcal{T})$. Therefore, $z \in \{x \ : \ C_v\} \subset \{x \ : \ x_j \in \{0, 1\} \ \forall \ j \in \{i_1, ..., i_k\}\} \subset P_{i_1,...,i_k}$, where the last inclusion is by definition of the latter set.

Now, we will give an example in $\mathbb{R}^3$ where the inclusion $T^k(P) \subset B^k(P)$ is strict. Let

$$P = \mathrm{conv}(\{(0, 0, 0), (1, 0, 0), (0, 1, 1), (1, 1, 1), (0, 0, \frac{1}{2}), (\frac{1}{2}, 0, 1), (1, \frac{1}{2}, 1)\})$$

and observe that $P_I = \{x : ax = b\} \cap [0,1]^n$ for $a = (0,-1,1)$ and $b = 0$. Let $\mathcal{T} \in \mathbb{T}_2^h$ be a tree with leaves v^1, v^2, v^3, v^4 where $C_{v^1} = \{x_1 = 0, x_3 = 0\}, C_{v^2} = \{x_1 = 0, x_3 = 1\}, C_{v^3} = \{x_1 = 1, x_2 = 0\}, C_{v^4} = \{x_1 = 1, x_2 = 1\}$ and observe that $\mathcal{T}(P) = P_I$, and therefore $T^2(P) = P_I$. We will now show that $B^2(P) \neq P_I$. Note that $(0,0,\frac{1}{2}) \in P_{1,2}, (\frac{1}{2},0,1) \in P_{2,3}$, and $(1,\frac{1}{2},1) \in P_{1,3}$, and of course $P_I \subset P_{i,j}$ for all $i,j \in [3]$. Therefore, we see that there is a point $\bar{x} \in B^2(P) = \bigcap_{i,j \in [3]} P_{i,j}$ with $a\bar{x} > 0$ by applying Lemma 2 with a, b as defined above, $b' = \frac{1}{2}$ and $R = \{(0,0,\frac{1}{2}),(\frac{1}{2},0,1),(1,\frac{1}{2},1)\}$, and $s^1, ..., s^3$ being any three affinely independent points of P_I. $\qquad\square$

Lemma 4. $L^k(P) \subsetneq T^k(P)$ for all $k \in [n]$.

Proof. We will prove the lemma by induction on k. First, observe that the definitions imply $L(P) = T(P)$ (since $x \in \mathrm{conv}(\{x \in P : x_j \in \{0,1\}\})$ if and only if $x \in \mathcal{T}(P)$ for some $\mathcal{T}$ of height 1). Now, assuming that $L^k(P) \subseteq T^k(P)$, we will show that $L^{k+1}(P) \subseteq T^{k+1}(P)$.

Note $L^{k+1}(P) = \bigcap_{j \in [n]} \mathrm{conv}(L^k(\{x \in P : x_j = 0\}) \cup L^k(\{x \in P : x_j = 1\}))$ by definition of $L^{k+1}(\cdot)$ and Lemma 1. Therefore, by the induction hypothesis, $L^{k+1}(P) \subseteq \bigcap_{j \in [n]} \mathrm{conv}(T^k(\{x \in P : x_j = 0\}) \cup T^k(\{x \in P : x_j = 1\}))$. We will proceed by showing

$$\bigcap_{j \in [n]} \mathrm{conv}\left(T^k(\{x \in P : x_j = 0\}) \cup T^k(\{x \in P : x_j = 1\})\right) \subseteq T^{k+1}.$$

Let $x \in \bigcap_{j \in [n]} \mathrm{conv}\left(T^k(\{x \in P : x_j = 0\}) \cup T^k(\{x \in P : x_j = 1\})\right)$, we will show that $x \in \mathcal{T}(P)$ for any $\mathcal{T} \in \mathbb{T}_{k+1}^h$. Let $j \in [n]$ be the variable branched on at the root of $\mathcal{T}$, and let $\mathcal{T}^0$ be the subtree rooted at the child of the root corresponding to branch $x_j = 0$ (and let $\mathcal{T}^1$ be defined similarly). Observe that $\mathcal{T}^0, \mathcal{T}^1 \in \mathbb{T}_k^h$. We know that $x \in \mathrm{conv}(x^{j,0}, x^{j,1})$, for some $x^{j,a} \in T^k(\{x \in P : x_j = a\})$ for $a \in \{0,1\}$. By definition of $T^k(\cdot)$ and the fact that $\mathcal{T}^0 \in \mathbb{T}_k^h$, we know that $x^{j,0} \in \mathcal{T}^0(\{x \in P : x_j = 0\})$ (and similarly $x^{j,1} \in \mathcal{T}^1(\{x \in P : x_j = 1\})$). Therefore, $x^{j,0}, x^{j,1} \in \mathcal{T}(P)$, which clearly implies $x \in \mathcal{T}(P)$ as desired. $\qquad\square$

The following is a direct corollary of Theorem 7.

Corollary 1. *For any n such that $n-1$ is divisible by 6, there is a polytope $P \subseteq [0,1]^n$ such that $P_I = L^2(P) \subsetneq T^{(n-1)/6}(P)$.*

Proof of Theorem 8. The theorem is directly implied by the combination of Lemma 3, Lemma 4, and Corollary 1. $\qquad\square$

Remark 6. *There is a simple construction, of similar spirit to that of Theorem 7, that shows that for any $n \geq 3$, there is a polytope $P \subseteq [0,1]^n$ such that $P_I = L^2(P) \subsetneq T^{n-1}(P)$.*

While at first glance, the Balas-Ceria-Cornuéjols operator and the canonical lift-and-project operator do not seem to differ significantly, the following direct corollary of Theorem 8 shows that applying the intersection "as you go" as opposed to applying it at the end of the disjunctive procedure can in fact have quite a significant impact.

Corollary 2. *For any n such that $n-1$ is divisible by 6, there is a polytope $P \subseteq [0,1]^n$ such that $P_I = L^2(P) \subsetneq B^{(n-1)/6}(P)$.*

Remark 7. *Theorem [8] gives a necessary condition for branch-and-bound to be advantagous compared to lift-and-project. In particular, if the best compact formulation constructed by branch-and-bound Q_{BB} comes from a "short, balanced" tree (i.e. a tree in $\mathbb{T}^h_k$), Sherali-Adams is also able to construct a compact formulation Q_{SA} such that $\max\{cx : (x, y) \in Q_{SA}\} \leq \max\{cx : (x, y) \in Q_{BB}\}$.*

References

[Bal85] Egon Balas. Disjunctive programming and a hierarchy of relaxations for discrete optimization problems. *SIAM Journal on Algebraic Discrete Methods*, 6(3):466–486, 1985.

[BCC93] 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, 1993.

[BCDSJ23] Amitabh Basu, Michele Conforti, Marco Di Summa, and Hongyi Jiang. Complexity of branch-and-bound and cutting planes in mixed-integer optimization. *Mathematical Programming*, 198(1):787–810, 2023.

[CCZ10] Michele Conforti, Gérard Cornuéjols, and Giacomo Zambelli. Extended formulations in combinatorial optimization. *4OR*, 8(1):1–48, 2010.

[CCZ14] Michele Conforti, Gérard Cornuéjols, and Giacomo Zambelli. *Integer Programming*. Springer International Publishing, 2014.

[Chv73] Vasek Chvátal. Edmonds polytopes and a hierarchy of combinatorial problems. *Discrete mathematics*, 4(4):305–337, 1973.

[CL01] Gérard Cornuéjols and Yanjun Li. Elementary closures for integer programs. *Operations Research Letters*, 28(1):1–8, 2001.

[DDM23] Santanu S Dey, Yatharth Dubey, and Marco Molinaro. Lower bounds on the size of general branch-and-bound trees. *Mathematical Programming*, 198(1):539–559, 2023.

[DGM15] Sanjeeb Dash, Oktay Günlük, and Marco Molinaro. On the relative strength of different generalizations of split cuts. *Discrete Optimization*, 16:36–50, 2015.

[Gom10] Ralph E Gomory. *Outline of an algorithm for integer solutions to linear programs and an algorithm for the mixed integer problem*. Springer, 2010.

[KMN11] Anna R Karlin, Claire Mathieu, and C Thach Nguyen. Integrality gaps of linear and semidefinite programming relaxations for knapsack. In *Integer Programming and Combinatoral Optimization: 15th International Conference, IPCO 2011, New York, NY, USA, June 15-17, 2011. Proceedings 15*, pages 301–314. Springer, 2011.

[Lau03] Monique Laurent. A comparison of the sherali-adams, lovász-schrijver, and lasserre relaxations for 0–1 programming. *Mathematics of Operations Research*, 28(3):470–496, 2003.

[LD10] Ailsa H Land and Alison G Doig. *An automatic method for solving discrete programming problems*. Springer, 2010.

[LS91] László Lovász and Alexander Schrijver. Cones of matrices and set-functions and 0−1 optimization. *SIAM journal on optimization*, 1(2):166−190, 1991.

[LW70] Don R Lick and Arthur T White. k-degenerate graphs. *Canadian Journal of Mathematics*, 22(5):1082−1096, 1970.

[NBW22] Mohammad Javad Naderi, Austin Buchanan, and Jose L Walteros. Worst-case analysis of clique mips. *Mathematical Programming*, 195(1-2):517−551, 2022.

[SA90] Hanif D Sherali and Warren P 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, 1990.

[Sch98] Alexander Schrijver. *Theory of linear and integer programming*. John Wiley & Sons, 1998.

[Sch03] Alexander Schrijver. *Combinatorial optimization: polyhedra and efficiency*, volume 24. Springer, 2003.

[ST10] Mohit Singh and Kunal Talwar. Improving integrality gaps via chvátal-gomory rounding. In *International Workshop on Randomization and Approximation Techniques in Computer Science*, pages 366−379. Springer, 2010.

[WB20] Jose L Walteros and Austin Buchanan. Why is maximum clique often easy in practice? *Operations Research*, 68(6):1866−1895, 2020.

[WN99] Laurence A Wolsey and George L Nemhauser. *Integer and combinatorial optimization*, volume 55. John Wiley & Sons, 1999.