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

February 24, 2025

Abstract

In this note, 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 "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 note, 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

*gc0v@andrew.cmu.edu
†ydubey@illinois.edu

are polynomial in n, we refer to Q_E as a *compact extended formulation*; see [CCZ10] for more on extended formulations. The procedures of Balas-Ceria-Cornuéjols (BCC) [BCC93] and Lovász-Schrijver (LS) [LS91] are two classical methods for generating extended formulations Q_E such that $P_I \subseteq \text{proj}_x(Q_E) \subseteq P$. By iterating (BCC) or (LS), one obtains hierarchies of extended formulations whose projections onto the x-space give increasingly tighter approximations of P_I, and in fact produce P_I after at most n steps. For $k = 1, \ldots, n$, Sherali-Adams (SA) [SA90] proposed a level-k procedure for generating another hierarchy of extended formulations; see [Lau03] for a comparison between the (LS) and (SA) hierarchies and [CL01] for a broader comparison of cut-operators.

Another way of generating extended formulations arises from the branch-and-bound method (BB) [LD60]. Consider a partial enumeration tree $\mathcal{T}$ where node v represents subproblem $P_v := P \cap \{x : C_v\}$ where C_v denotes the set of branching constraints from the root to node v. The set of leaves of $\mathcal{T}$ is denoted by leaves($\mathcal{T}$). 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. We hope this example is clear even though several notions have not yet been explicitly defined; the reader may refer to Section 2 for relevant notation.

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 inequality $\sum_{v \in V} x_v \leq 1$ is valid for any tree $\mathcal{T}$ where each variable is branched on exactly once, and a node of the tree with a constraint $x_v = 1$ for any $v \in V$ is never branched on; such a tree $\mathcal{T}$ has $n + 1$ leaves, each with an integral subproblem P_v. Thus $\mathcal{T}(P) = P_I$.

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 Singh and Talwar [ST10] studied the Gomory-Chvátal (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 note 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 note. 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 Section 3, we identify two combinatorial optinization problems, namely the maximum clique problem on sparse graphs and the uniform knapsack problem with small capacity, and show that branch-and-bound is effective in solving these problems, while lift-and-project methods are not. Moreover, we present a family of trees that unifies these two examples.

In Section 4, we give an example where lift-and project does very well while branch-and-bound does poorly. Specifically, we exhibit a polytope P such that lift-and-project generates P_I in two rounds, while there is a point $x \in P \setminus P_I$ such that $x \in \mathcal{T}(P)$ for every branch-and-bound tree $\mathcal{T}$ with at most $2^{(n-1)/6}$ leaves.

In Section 5, 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 note.

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 $x_j = 0 \vee x_j = 1$ 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. *Consider a branch-and-bound tree $\mathcal{T}$ where node v of $\mathcal{T}$ is labeled by C_v, the set of constraints $x_j = 0$ or $x_j = 1$ that are added along the path from the root node to v.*

- *We denote the number of nodes of $\mathcal{T}$ by $|\mathcal{T}|$. This is what is termed* the size *of this tree.*
- *For any $v \in \text{leaves}(\mathcal{T})$, 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 $h(\mathcal{T}) \leq n$.*

Consider a polytope $P \subseteq [0,1]$.

- *The* atom *of a node $v \in \mathcal{T}$ is the feasible region defined by P and the branching constraints at node v; formally, the atom of node v is the set $P_v := P \cap \{x : C_v\}$.*
- *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.*

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 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 $\bar{x} \in L(P)$, there is some $x^0 \in \{x \in P : x_i = 0\}$ and $x^1 \in \{x \in P : x_i = 1\}$ such that $\bar{x} \in \text{conv}(x^0, x^1)$. Moreover, it must hold that $x^0_j = x^1_j = 0$, since otherwise $\bar{x}_j > 0$. This proves the claim.

The general case follows by induction. □

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 $x \in \bigcap_{r \in R} conv(s^1, ..., s^n, r)$ satisfying the strict inequality $ax > b$.*

The second lift-and-project operator that we consider in this note is based on sequential convexification, introduced in Section 2 of [BCC93]. For $J \subseteq [n]$, Theorem 2.2 of [BCC93] gives a method to find an extended formulation for the polytope $P_J = \text{conv}(\{x \in P : x_j \in \{0, 1\} \; \forall \; j \in J)$. Let $k \in [n]$. In this study, we will consider the following natural operator

$$B^k(P) = \bigcap_{J \subseteq [n] \,:\, |J| = k} P_J$$

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 5).

Third, we consider the Sherali-Adams hierarchy for a polytope $P := \{x \in \mathbb{R}^n : \;\; Ax \geq b\}$ contained in the 0,1 hypercube. The kth level of the hierarchy is obtained by generating the nonlinear system $(Ax - b) \prod_{i \in I} x_i \prod_{j \in J} (1 - x_j) \geq 0$ for all disjoint index sets $I, J \subseteq \{1, \dots, n\}$ such that $1 \leq |I| + |J| \leq k$, then linearizing this system by substituting x_j for x_j^2 and substituting y_I for $\prod_{i \in I} x_i$ for all nonempty $I \subseteq \{1, \dots, n\}$. Set $y_\emptyset = 1$ and call R_k the polytope defined by these linear inequalities. The projection of R_k onto the x-space is the kth level polytope $SA^k(P)$ 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. *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 and Polynomial Number of Feasible Solutions

When $P \subset [0, 1]^n$ is such that, for some $k \in \mathbb{Z}_+$ fixing k variables results in an infeasible problem, branch-and-bound always provides a polynomial-size extended formulation, whereas lift-and-project relaxations such as Sherali-Adams may have exponential size extended formulations. 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.

3.1 Uniform Knapsack with Small Capacity

Theorem 2. *For any polytope $P \subset [0, 1]^n$ such that $\sum_{i \in [n]} x_i \leq k - \epsilon$ is valid for P for some $\epsilon \in (0, 1)$, there is a branch-and-bound tree $\mathcal{T}$ with $\mathcal{T}(P) = P_I$ and size at most $O(n^k)$. Furthermore, there exists a polytope $P \subset [0, 1]^n$ 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 $q \geq 3$ is an integer.*

Proof. We begin by proving the positive result on branch-and-bound. Consider any branch-and-bound tree $\mathcal{T}$. Let $v \in \text{leaves}(\mathcal{T})$. If $|J_v^1| = k$, then for $x \in \{x : C_v\}$, it holds that $\sum_{i \in [n]} x_j \geq \sum_{j \in J_v^1} x_j = k > k - \epsilon$, which implies $x \notin P$, and therefore $P \cap \{x : C_v\} = \emptyset$. Therefore $\mathcal{T}$ has no leaf with $|J_v^1| > k$. This implies that the number of leaves of $\mathcal{T}$ is at most $\binom{n}{k} + \binom{n}{k-1} + ... + \binom{n}{0} = O(n^k)$. Since $\mathcal{T}$ is a binary tree, it follows that the total number of nodes of $\mathcal{T}$ is no greater than twice the number of leaves, which is again $O(n^k)$. At termination, all the leaves v have atoms P_v that are empty or integral polytopes. Therefore $\mathcal{T}(P) = P_I$.

Now, we prove the lower bound on Sherali-Adams. Let q be any integer at least 3 and let $P = \{x \in [0,1]^n : \sum_{j\in[n]} qx_j \le 2(q-1)\}$. Note that $\sum_{i\in[n]} x_i \le 2 - \frac{2}{q}$ is clearly valid for P, and in particular we can write $P = \{x \in [0,1]^n : \sum_{i\in[n]} x_i \le 2(\frac{q-1}{q})\}$. 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 $\sum_{i\in[n]} x_i \le 1$ is not valid for $SA^t(P)$, proving the desired result. □

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 2 ([LW70])**.** *A graph G is d-degenerate if each of its induced 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 $k \le d+1$.*

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 3. *Let $G = (V,E)$ be a graph on n vertices with maximum clique size k and degeneracy d. Let $P = \{x \in \mathbb{R}^n_+ : x_u + x_v \le 1 \ \forall\ uv \notin E\}$ be the fractional clique polytope for G. There is a branch-and-bound tree $\mathcal{T}$ with $\mathcal{T}(P) = P_I$ and size at most $O(n^k)$. On the other hand, $SA^{\lfloor n/(d+1)-3\rfloor}(P) \neq P_I$.*

Proof. We begin with the result on branch-and-bound. Let $v \in \text{leaves}(\mathcal{T})$. If $|J_v^1| > k$, then for $x \in \{x : C_v\}$, it holds that $\sum_{j\in J_v^1} x_j > k$. Since G contains no clique of size greater than k, it follows that there is $rt \notin E$ such that $x_r, x_t \in J_v^1$. Thus $x_r + x_t = 2$. This implies $x \notin P$, and therefore $P \cap \{x : C_v\} = \emptyset$. Therefore $\mathcal{T}$ has no leaf with $|J_v^1| > k$. This implies that the number of leaves of $\mathcal{T}$ is at most $O(n^k)$. Since $\mathcal{T}$ is a binary tree, it follows that the total number of nodes of $\mathcal{T}$ is no greater than $O(n^k)$. At termination, all the leaves v have atoms P_v that are empty or integral polytopes. Therefore $\mathcal{T}(P) = P_I$.

We will now show that $SA^{\frac{n}{d+1}-3}(P) \neq P_I$. First, 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. $|N(v_i) \cap \{v_{i+1}, ..., v_n\}| \le d$, 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.

Observe that G has a stable set of size at least $\frac{n}{d+1}$. We can construct such a 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. $|S \cup \bigcup_{v\in S} N^r(v)| \le k + kd$). 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 in 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 $\sum_{v\in S} x_v \le 1$ is not valid for $SA^{\frac{n}{d+1}-3}(P)$, while it is clearly valid for P_I. □

3.3 Skewed k-trees

The capability of branch-and-bound exhibited in the two previous examples can be understood by considering a specific family of trees.

Definition 3. *Let $k \in \mathbb{Z}_{\ge 1}$ 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 $v \in$ leaves$(\mathcal{T}_{i-1})$, if $|J_v^1| < k$: branch on variable i at v. In other words, leaves$(\mathcal{T}_i)$ = $\big($leaves$(\mathcal{T}_{i-1}) \setminus \{v : |J_v^1| < k\}\big) \cup \big\{v^0, v^1 : C_{v^a} = C_v \cup \{x_i = a\}, a \in \{0,1\}\big\}$, and set $\mathcal{T}^{\pi,k} = \mathcal{T}_n$.

It is easy to verify that, for any permutation π, the skewed k-tree $\mathcal{T}^{\pi,k}$ has size $n^{O(k)}$ and satisfies the following properties

1. for any $v \in \text{leaves}(\mathcal{T}^{\pi,k})$, either $|J_v^1| = k$ or $J_v^0 \cup J_v^1 = [n]$, and
2. for each subset $J \subset [n]$ of $k' \le k$ variables, there is exactly one $v \in \text{leaves}(\mathcal{T}^{\pi,k})$ such that $J_v^1 = J$.

Remark 2. *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 $\mathcal{T}^{\pi,n}(P) = P_I$.*

4 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. $L^2(P) = P_I$), then every point $x \in P \setminus P_I$ is separated by at least one *small* branch-and-bound tree. Below, we show that statements of this form are not true.

Theorem 4. *Let n be any nonnegative integer such that $n-1$ is divisible by 6. There is a polytope $P \subseteq [0,1]^n$ and a point $x \in P \setminus P_I$ such that no branch-and-bound tree of size $\le 2^{(n-1)/6}$ separates x, while lift-and-project retrieves the integer hull P_I in two rounds.*

This theorem shows that there exists $P \subseteq [0,1]^n$ such that

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

where $\mathbb{T}^{\ell}_k$ denotes the set of branch-and-bound trees with at most k leaves. The proof builds on Theorem 2.2 of [BCDSJ23] and Lemma 3.1 of [DGM15]] (see Lemma 2 above).

Proof. We begin by defining the polytope P of interest. Let $P \cap \{x \in \mathbb{R}^n : x_n = 0\} = \{x \in [0,1]^n : x_n = 0\}$, and $P \cap \{x \in \mathbb{R}^n : x_n = 1\} = \{(x, x_n) \in \mathbb{R}^{n-1} \times \mathbb{R} : x \in Q,\ x_n = 1\}$, 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 $\mathcal{T} \in \mathbb{T}^{\ell}_{2^{(n-1)/6}}$ (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 $j \neq i$, and $x_n = 1$). 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 $\mathcal{T} \in \mathbb{T}^{\ell}_{2^{m/2}}$, 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 $\mathcal{T} \in \mathbb{T}^{\ell}_{2^{m/2}}$, 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 $\mathcal{T} \in \mathbb{T}^{\ell}_{2^{m/2}}$, there is some $x^{\mathcal{T}} \in \mathcal{T}(P)$ with $x_n = 1$. Finally, to show that there is some $\bar{x} \in \bigcap_{\mathcal{T}\in\mathbb{T}^{\ell}_{2^{m/2}}} \mathcal{T}(P)$, we apply Lemma 2 with $a = e_n, b = 0, b' = 1$ and $s^1, \ldots, s^n$ being the points $0, e^1, \ldots, e^{n-1}$, 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 $\bar{x} \in P \setminus P_I$. □

5 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 $\mathbb{T}^h_k = \{\mathcal{T} : h(\mathcal{T}) \leq k\}$ 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 5. *For any polytope $P \subseteq [0,1]^n$,*

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

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

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 $v \in \text{leaves}(\mathcal{T})$. 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

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

and observe that $P_I = \{x : ax = b\} \cap [0,1]^n$ for $a = (0,-1,1)$ and $b = 0$. Let $\mathcal{T} \in \mathbb{T}^h_2$ 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. □

Lemma 4. *$L^k(P) \subseteq T^k(P)$ 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 $x \in \text{conv}(\{x \in P : x_j \in \{0,1\}\})$ for some j 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 $L^{k+1}(P) = \bigcap_{j\in[n]} \text{conv}(L^k(\{x \in P : x_j = 0\}) \cup L^k(\{x \in P : x_j = 1\}))$ by definition of $L^{k+1}(\cdot)$ and Lemma 1. Therefore, by the induction hypothesis, $L^{k+1}(P) \subseteq \bigcap_{j\in[n]} \text{conv}(T^k(\{x \in P : x_j = 0\}) \cup T^k(\{x \in P : x_j = 1\}))$. We will proceed by showing

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

Let $x \in \bigcap_{j\in[n]} \text{conv}\left(T^k(\{x \in P : x_j = 0\}) \cup T^k(\{x \in P : x_j = 1\})\right)$, we will show that $x \in \mathcal{T}(P)$ for any $\mathcal{T} \in \mathbb{T}^h_{k+1}$. 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 $\mathcal{T}^0, \mathcal{T}^1 \in \mathbb{T}^h_k$. We know that $x \in \text{conv}(x^{j,0}, x^{j,1})$, 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 $\mathcal{T}^0 \in \mathbb{T}^h_k$, 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. □

Proof of Theorem 5. Theorem 4 implies that, for any n such that $n-1$ is divisible by 6, there is a polytope $P \subseteq [0,1]^n$ such that $P_I = L^2(P) \subsetneq T^{(n-1)/6}(P)$. Theorem 5 now follows by combining this result with Lemmas 3 and 4. □

While at first glance, the Balas-Ceria-Cornuéjols operator and the canonical lift-and-project operator do not seem to differ significantly, Theorem 5 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.

Remark 3. *Theorem 5 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 $\mathbb{T}_k^h$), Sherali-Adams is also able to construct a compact formulation Q_{SA} such that* $\max\{cx : (x,y) \in Q_{SA}\} \leq \max\{cx : (x,y) \in Q_{BB}\}$.

Acknowledgements: This research was supported in part by ONR grant N00014-22-1-2528.

Declarations: The authors have no relevant financial or non-financial interests to disclose. The authors have no conflicts of interest to declare that are relevant to the content of this article.

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 semi-definite 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.

[LD60] Alisa H Land and Alison G Doig. An automatic method of solving discrete programming problems. *Econometrica*, 28:497–520, 1960.

[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.