# Lower Bounds on the Size of General Branch-and-Bound Trees

Santanu S. Dey[*], Yatharth Dubey[†], Marco Molinaro[‡]

**Abstract**

A *general branch-and-bound tree* is a branch-and-bound tree which is allowed to use general disjunctions of the form $\pi^\top x \leq \pi_0 \vee \pi^\top x \geq \pi_0 + 1$, where $\pi$ is an integer vector and $\pi_0$ is an integer scalar, to create child nodes. We construct a packing instance, a set covering instance, and a Traveling Salesman Problem instance, such that any general branch-and-bound tree that solves these instances must be of exponential size. We also verify that an exponential lower bound on the size of general branch-and-bound trees persists when we add Gaussian noise to the coefficients of the cross polytope, thus showing that polynomial-size "smoothed analysis" upper bound is not possible. The results in this paper can be viewed as the branch-and-bound analog of the seminal paper by Chvátal et al. [7], who proved lower bounds for the Chvátal-Gomory rank.

## 1 Introduction

Solving combinatorial optimization problems to optimality is a central object of study in Operations Research, Computer Science, and Mathematics. One way to solve a combinatorial optimization problem is to model it as mixed integer linear program (MILP) and use an MILP solver. The branch-and-bound algorithm, invented by Land and Doig in [20], is the underlying algorithm implemented in all modern state-of-the-art MILP solvers.

As is well known, the branch-and-bound algorithm searches the solution space by recursively partitioning it. The progress of the algorithm is monitored by maintaining a tree. Each node of the tree corresponds to a linear program (LP) solved, and in particular, the root node corresponds to the LP relaxation of the integer program. After solving the LP corresponding to a node, the feasible region of the LP is partitioned into two subproblems (which correspond to the child nodes of the given node), so that the fractional optimal solution of the LP is not included in either subproblem, but any integer feasible solution contained in the feasible region of the LP is included in one of the two subproblems. This is accomplished by adding an inequality of the form $\pi^\top x \leq \pi_0$ to the first subproblem and the inequality $\pi^\top x \geq \pi_0 + 1$ to the second subproblem (these two inequalities are referred as a disjunction), where $\pi$ is an integer vector and $\pi_0$ is an integer scalar. The process of partitioning at a node stops if (i) the LP at the node is infeasible, (ii) the LP's optimal solution is integer feasible, or (iii) the LP's optimal objective function value is worse than an already known integer feasible solution. These three conditions are sometimes referred to as the rules for pruning a node. The algorithm terminates when there are no more "open nodes" to process, that is all nodes have been pruned. A branch-and-bound algorithm is completely described by fixing a rule for partitioning the feasible region at each node and a rule for selecting which open node should be solved and branched on next. If the choice of $\pi$ is limited to being unit vectors in direction $j$, then we call such an algorithm as a *simple* branch-and-bound algorithm and without such a restriction on $\pi$ we call the algorithm as a ***general*** branch-and-bound algorithm. See [26, 8] for more discussion on branch-and-bound.

[*]ISyE, Georgia Institute of Technology

[†]ISyE, Georgia Institute of Technology

[‡]Computer Science department, PUC-Rio

## 1.1 Known bounds on sizes of branch-and-bound trees

**Upper bounds on size of brand-and-bound tree and "positive" results.** In 1983, Lenstra [22] showed that integer programs can be solved in polynomial time in fixed dimension. This algorithm, can be viewed as a general branch-and-bound algorithm, and uses tools from geometry of numbers, in particular the lattice basis reduction algorithm [21] to decide on $\pi$ for partitioning the feasible region. Pataki [23] proved that most random packing integer programs can be solved at the root-node using a partitioning scheme similar to the one proposed by Lenstra [22]. It has been observed that using such general partitioning rules can be significantly more efficient than simple branch-and-bound for some instances [1, 11], but most commercial solver use the latter. Recently, we showed [15] that for certain classes of random integer programs the simple branch-and bound-tree has polynomial size (number of nodes), with good probability. See also [5] for nice extensions of this direction of results. Beame et al. [4] recently studied how branch-and-bound can give good upper bounds for certain SAT formulas.

**Lower Bounds on size of branch-and-bound tree and connections to size of cutting-plane algorithms.** Jeroslow [18] and Chvátal [6] present examples of integer programs where every *simple* branch-and-bound algorithm for solving has exponential size. However, these instances can be solved with small (polynomial-size) general branch-and-bound trees; see Yang et al. [27] and Basu et al. [2]. Cook et al. [10] present a TSP instance that require exponential-size branch-and-cut trees that uses simple branching (recall that branch-and-cut is branch-and-bound where one is allowed to add cuts to the intermediate LPs). Basu et al. [3] compare the performance of branch-and-bound with the performance of cutting-plane algorithms, providing instances where one outperforms the other and vice-versa. In another paper, Basu et al. [2] compare branch-and-bound with branch-and-cut , providing instances where branch-and-cut is exponentially better than branch-and-bound. They also present a result showing that the sparsity of the disjunctions can have a large impact on the size of the branch-and-bound tree. Beame et al. [4] asked as an open question whether there are superpolynomial lower bounds for general branch-and-bound algorithm. Dadush et al. [12] settled this in the affirmative. In particular, they show that any general branch-and-bound tree that proves the integer infeasibility of the cross-polytope has at least $\frac{2^n}{n}$ leaf nodes. They also note that the cross-polytope has an exponential number of facets, a fact crucially used in their proof, and present the open question of whether there is an exponential lower bound for a polytope described by a polynomial number of facets.

Concurrent to the development of our work, Fleming et al. [16] showed a fascinating relationship between general branch-and-bound proofs and cutting-plane proofs using Chvátal-Gomory (CG) cutting-planes:

**Theorem 1** (Theorem 3.7 from [16])**.** *Let $P \subseteq [0,1]^n$ be an integer-infeasible polytope and suppose there is a general branch-and-bound proof of infeasibility of size $s$ and with maximum coefficient $c$. Then there is a CG proof of infeasibility of size at most*

$$s(cn)^{\log s}.$$

The following simple corollary allows us to infer exponential lower bounds for branch-and-bound proofs for polytopes for which we have exponential lower bounds for CG proofs.

**Corollary 1.** *Let $P \subseteq [0,1]^n$ be an integer infeasible polytope such that any CG proof of infeasibility of $P$ has length at least $L$. Then, any general branch-and-bound proof of infeasibility of $P$ with maximum coefficient $c$ has size at least*

$$L^{\frac{1}{1+\log(cn)}}.$$

The above result resolves an open question raised in Basu et al. [3]. Moreover, Pudlak [24] and Dash [13] provide exponential lower bounds for CG proofs for the clique vs. coloring problem, which is of note since this problem has only polynomially many inequalities. Thus Corollary 1 taken together with results in [24] and [13] also settles the question raised in Dadush et al. [12] as long as the maximum coefficient in the disjunctions used in the tree is bounded by a polynoimal of $n$.

### 1.2 Contributions of this paper and relationship to exisiting results

**Contributions.** We construct an instance of packing-type and a set cover instance such that any general branch-and-bound tree that solves these instances must be of exponential (with respect to the ambient dimension) size. We note that the packing and covering instances are described using an exponential number of constraints. We also present a simple proof that any branch-and-bound tree proving the integer infeasibility of the cross-polytope must have $2^n$ leaves. We then extend this result to give (high probability) exponential lower bounds for perturbed instances of the cross polytope where independent Gaussian noise is added to the entries of the constraint matrix. To our knowledge, this is the first result that shows that a "smoothed analysis" polynomial upper bound on the size of branch-and-bound trees is not possible. Finally, we show an exponential lower bound on the size of any general branch-and-bound tree for the Traveling Salesman Problem (TSP).

**Comparison to previous results.** We now discuss our results in the context of the recent landscape, in particular with the results of [12] and [16].

1. *New problems with exponential lower bounds on size of general branch-and-bound tree:* As mentioned earlier, recently Dadush et al. [12] provided the first exponential lower bound on the size of general branch-and-bound tree for the cross polytope. Corollary 1 from [16] only implies branch-and-bound hardness for polytopes for which we already have CG hardness. These come few and far between in the existing literature and these instances are often a bit artificial; see [24] and [13]. In this paper, we provide lower bounds on size of general branch-and-bound tree for packing and set covering instances, which are more natural combinatorial problems.

2. *Improved quality of bounds:* Dadush et al. [12] show that any branch-and-bound proof of infeasibility of the cross-polytope has at least $\frac{2^n}{n}$ leaves. We improve on this result by providing a simple proof that any such proof of infeasibility must have $2^n$ leaves.

   Chvatal et al. [7] provide a $\frac{1}{3n}2^{n/8}$ lower bound on CG proofs for TSP. Combined with Corollary 1, this can be used to show a lower bound of $2^{O\left(\frac{n}{\log cn}\right)}$ for TSP for branch-and-bound trees using maximum coefficient $c$ for disjunctions. We are able to achieve a stronger lower bound of $2^{\Omega(n)}$.

3. *Removing dependence on the maximum coefficient size used in the branch-and-bound proof:* The bound given in Corollary 1 depends on the maximum coefficient size used in the branch-and-bound proof. In [16], the authors mention that they "view this as a step toward proving [branch-and-bound] lower bounds (with no restrictions on the weight)". Our results satisfy this constraint, as none of the bounds presented in this work depend on the coefficients of the inequalities of the general branch-and-bound proof.

### 1.3 Roadmap and notation

Since this paper focuses on lower bounds for general branch-and-bound trees (obviously implying lower bounds for simple branch-and-bound tree), we drop the term "general" in the rest of the paper. The paper is organized as follows. In Section 2, we present necessary definitions. In Section 3, we present key reduction results that allow for transferring lower bounds on size of branch-and-bound trees from one optimization problem to another. In Section 4, we present a lower bound on the size of branch-and-bound trees for packing and set covering instances. In Section 5, we present lower bound on size of branch-and-bound tree for the cross polytope and some other related technical results. In Section 6, we show that even after adding Gaussian noise to the coefficients of the cross polytope, with good probability, general branch-and-bound will require an exponentially large tree to prove its infeasibility. Finally, in Section 7, we use results from Section 5 and Section 3 to provide an exponential lower bound on the size of a branch-and-bound tree for solving a TSP instance.

For a positive integer $n$, we denote the set $\{1, \dots, n\}$ as $[n]$. When the dimension is clear from context, we use the notation $\mathbf{1}$ to be a vector whose every entry is 1. Let $A$ be a set of linear constraints of the form $(\pi^i)^\top x \leq \pi_0^i,\ \forall i \in [m]$. Then let $\{x : A\}$ denote the set of all $x \in [0,1]^n$ such that all of the constraints $A$ are valid for $x$ (i.e. the polytope defined by the set of constraints $A$). Also note that for a subset of these constraints $B \subseteq A$, it holds that $\{x : B\} \supseteq \{x : A\}$. Also note that for two sets of constraints $A, B$, it holds that $\{x : A \cup B\} = \{x : A\} \cap \{x : B\}$. Given a set $S$, we denotes it convex hull by $\text{conv}(S)$. Given a polytope $P \subseteq \mathbb{R}^n$, we denote its integer hull, that is the set $\text{conv}(P \cap \mathbb{Z}^n)$, as $P_I$.

## 2 Abstract branch-and-bound trees and notions of hardness

In order to present lower bounds on the size of branch-and-bound (BB) trees, we simplify our analysis by removing two typical condition assumed in a BB algorithm – (i) the requirement that the partitioning into two subproblems (which correspond to the child nodes of the given node) is done is such a way that the fractional optimal solution of the parent node is not included in either subproblem, (ii) branching is not done on pruned nodes. By removing these conditions, we can talk about a branch-and-bound tree independent of the underlying polytope – it is just a full binary tree with (that is 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 some disjunction $\pi^\top x \leq \pi_0 \vee \pi^\top x \geq \pi_0 + 1$, where each of the child nodes adds one of these constraints to its set of branching constraints together with all the branching constraints of the parent node. Note that proving lower bounds on the size of such branch-and-bound trees that solves a given integer program certainly gives a lower bound on the size of branch-and-bound trees where we require the optimal solution of the linear program at a node to not belong to the child nodes and where we require to stop branching on a node that is pruned. Finally, note that since a BB tree is a full 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 $P$, and a node $v$ of the tree:*

- *We denote the the number of nodes of the branch-and-bound tree $\mathcal{T}$ by $|\mathcal{T}|$.*
- *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 $v$).*
- *We call the feasible region defined by the LP relaxation $P$ and the branching constraints at node $v$ the* atom *of this node, i.e. $P \cap \{x : C_v\}$ is the atom corresponding to $v$.*
- *We let $\mathcal{T}(P)$ to denote the union of the atoms corresponding to the leaves of $\mathcal{T}$ when run on polytope $P$, i.e. $\mathcal{T}(P) = \bigcup_{v \in leaves(\mathcal{T})} (P \cap \{x : C_v\})$.*
- *For any $x^* \in P \setminus P_I$, we say that $\mathcal{T}$* separates *$x^*$ from $P$ if $x^* \notin conv(\mathcal{T}(P))$.*
- *Given a vector $c \in \mathbb{R}^n$, we say $\mathcal{T}$* solves *$\max_{x \in P \cap \{0,1\}^n} \langle c, x \rangle$, if for all the leaf nodes $v$ of $\mathcal{T}$, one of the following three conditions hold: (i) the atom of $v$ is empty, (ii) the optimal solution of the linear program $\max_{x \in \text{atom of } v} \langle c, x \rangle$ is integral, or (iii) $\max_{x \in \text{atom of } v} \langle c, x \rangle$ is at most the objective function value of another atom whose optimal solution is integral. (If $P \cap \mathbb{Z}^n = \emptyset$, then (ii) and (iii) are not possible.)*

Given a polytope $P \subseteq \mathbb{R}^n$, we define its *BB hardness* as:

$$BBhardness = \max_{c \in \mathbb{R}^n} \left( \min \left\{ |\mathcal{T}| \; : \; \mathcal{T} \text{ solves } \max_{x \in P \cap \mathbb{Z}^n} \langle c, x \rangle \right\} \right)$$

Our goal for most of this paper is to provide lower bounds on the *BB hardness* of certain polytopes.

To get exponential lower bounds on BB hardness for some $P$, we will often present a particular point $x^* \in P \setminus P_I$ such that any $\mathcal{T}$ that separates $x^*$ from $P$ must have exponential size. We formalize this below.

**Definition 2** (BBdepth)**.** *Let $P \subseteq \mathbb{R}^n$ be a polytope and consider any $x^* \in P \setminus P_I$. Let $\mathcal{T}$ be the smallest BB tree that separates $x^*$ from $P$. Then, define BBdepth$(x^*, P)$ to be $|\mathcal{T}|$.*

**Definition 3** (BBrank)**.** *Define BBrank$(P) = \max_{x \in P \setminus P_I}$ BBdepth$(x, P)$.*

**Lemma 1** (BBrank lower bounds BB hardness)**.** *Let $P \subseteq \mathbb{R}^n$ be a polytope. Then, there exists $c \in \mathbb{R}^n$ such that any BB tree solving $\max_{x \in P \cap \mathbb{Z}^n} \langle c, x \rangle$ must have size at least BBrank$(P)$, that is the BB hardness of $P$ is at least BBrank$(P)$.*

*Proof.* Let $x^* \in \operatorname{argmax}_{x \in P \setminus P_I} \text{BBdepth}(x, P)$, so that $\text{BBrank}(P) = \text{BBdepth}(x^*, P)$. Since $x^*$ does not belong to the convex set $P_I$, there exists $c$ with the separation property $\langle c, x^* \rangle > \max_{x \in P_I} \langle c, x \rangle$. By choice of $x^*$, for any BB tree $\mathcal{T}$ with $|\mathcal{T}| < \text{BBrank}(P)$ it holds that $x^* \in \text{conv}(\mathcal{T}(P))$. Then such tree $\mathcal{T}$ must have a leaf whose optimal solution has value at least $\langle c, x^* \rangle > \max_{x \in P_I} \langle c, x \rangle$, and therefore must still be explored, showing that $\mathcal{T}$ does not solve $\max_{x \in P \cap \mathbb{Z}^n} \langle c, x \rangle$. □

We now show that, under some conditions, the reverse of this kind of relationship can hold. We will use this reverse relationship to prove the BB hardness of optimizing over an integer feasible polytope given the BB hardness of proving the infeasibility of another "smaller" polytope.

**Lemma 2** (Infeasibility to optimization)**.** *Let $P \subseteq \mathbb{R}^n$ be a polytope and $\langle c, x \rangle \leq \delta$ be a facet defining inequality of $P_I$. Assume the affine hull of $P$ and $P_I$ are the same. Then, for every $\varepsilon > 0$*

$$BBrank(P) \geq BBhardness(\{x \in P \; : \; \langle c, x \rangle \geq \delta + \varepsilon\}).$$

Before we can present the proof of Lemma 2 we require a technical lemma from [14]. The full-dimensional case $L = \mathbb{R}^n$ is Lemma 3.1 of [14], and the general case follows directly by applying it to the affine subspace $L$.

**Lemma 3** ([14])**.** *Consider an affine subspace $L \subseteq \mathbb{R}^n$ and a hyperplane $H = \{x \in \mathbb{R}^n \; : \; \langle c, x \rangle = \delta\}$ that does not contain $L$. Consider $\dim(L)$ affinely independent points $s^1, s^2 \ldots, s^{\dim(L)}$ in $L \cap H$. Let $\delta' > \delta$ and let $G$ be a bounded and non-empty subset of $L \cap \{x \in \mathbb{R}^n \; : \; \langle c, x \rangle \geq \delta'\}$. Then there exists a point $x$ in $\bigcap_{g \in G} conv(s^1, \ldots, s^{\dim(L)}, g)$ satisfying the strict inequality $\langle c, x \rangle > \delta$.*

*Proof of Lemma 2.* Let $L$ be the affine hull of $P_I$. Then there exist $d := \dim(P_I) = \dim(L)$ affinely independent vertices of $\{x \in P_I \; : \; \langle c, x \rangle = \delta\}$. Let $s^1, \ldots, s^d$ be $d$ such affinely independent vertices and note that since they are vertices of $P_I$, they are all integral. Let $G := \{x \in P \; : \; \langle c, x \rangle \geq \delta + \varepsilon\}$, which is a bounded set since $P$ is bounded. Let $N := \text{BBhardness}(G)$.

Let $\mathcal{T}$ be a BB tree such that $|\mathcal{T}| < N$. Then we have that $\mathcal{T}(G) \neq \emptyset$, that is, there exists $x^*(\mathcal{T}) \in \mathcal{T}(G)$. In particular $x^*(\mathcal{T}) \in G$. Moreover, since $G \subseteq P$, we have $\mathcal{T}(G) \subseteq \mathcal{T}(P)$ (see Lemma 4 in the next section for a formal proof of this), we have $x^*(\mathcal{T}) \in \mathcal{T}(P)$. Also note that since $s^1, \ldots, s^d \in P \cap \mathbb{Z}^n$, we have that these points also belong to $\mathcal{T}(P)$. Thus,

$$\text{conv}\left(s^1, \ldots, s^d, x^*(\mathcal{T})\right) \subseteq \text{conv}(\mathcal{T}(P)).$$

Now applying Lemma 3, with $\delta' = \delta + \varepsilon$, we have that there exists $x^*$ such that

$$x^* \in \bigcap_{\mathcal{T}:|\mathcal{T}|<N} \operatorname{conv}\left(s^1, \dots, s^d, x^*(\mathcal{T})\right) \subseteq \bigcap_{\mathcal{T}:|\mathcal{T}|<N} \operatorname{conv}(\mathcal{T}(P)) \tag{1}$$

and such that $\langle c, x\rangle > \delta$. Clearly, $x^* \notin P_I$, since $\langle c, x\rangle \le \delta$ is a valid inequality for $P_I$. Thus, since (1) implies $x^* \in \operatorname{conv}(\mathcal{T}(P))$ for all $\mathcal{T}$ with $|\mathcal{T}| < N$, we have that $\mathrm{BBdepth}(x^*, P) \ge N$ and consequently, $\mathrm{BBrank}(P) \ge N$. □

## 3 Framework for BB hardness reductions

We begin by showing monotonicity of the operator $\mathcal{T}(\cdot)$.

**Lemma 4** (Monotonicity of leaves)**.** *Let $Q \subseteq P \subseteq \mathbb{R}^n$ be polytopes. Then $\mathcal{T}(Q) \subseteq \mathcal{T}(P)$.*

*Proof.* For any leaf $v \in \mathcal{T}$, recall that $C_v$ is the set of branching constraints of $v$. Then $\mathcal{T}(Q) = \bigcup_{v \in \mathrm{leaves}(\mathcal{T})}(Q \cap \{x : C_v\}) = Q \cap \bigcup_{v \in \mathrm{leaves}(\mathcal{T})}\{x : C_v\} \subseteq P \cap \bigcup_{v \in \mathrm{leaves}(\mathcal{T})}\{x : C_v\} = \bigcup_{v \in \mathrm{leaves}(\mathcal{T})}(P \cap \{x : C_v\}) = \mathcal{T}(P)$. □

The following corollary follows easily from Lemma 4. In particular, consider the smallest BB tree $\mathcal{T}$ that separates $x^*$ from $P$. By Lemma 4, the same tree, when applied to $Q$, will not have $x^*$ in the convex hull of its leaves and therefore separates $x^*$ from $Q$.

**Corollary 2** (Monotonicity of depth)**.** *Let $Q \subseteq P \subseteq \mathbb{R}^n$ be polytopes. Suppose there is some $x^* \in (Q \setminus Q_I) \cap (P \setminus P_I) = Q \setminus P_I$. Then,*

$$\mathit{BBdepth}(x^*, P) \ge \mathit{BBdepth}(x^*, Q).$$

Inspired by the lower bounds for cutting-plane rank from [7], we show that integral affine transformations conserve the hardness of separating a point via branch-and-bound, i.e. they conserve BBdepth. Then, we give a condition where BBrank is also conserved. These will be use to obtain lower bounds in the subsequent sections.

We say that $f : \mathbb{R}^n \longrightarrow \mathbb{R}^m$ is an *integral affine function* if it has the form $f(x) = Cx + d$, where $C \in \mathbb{Z}^{m\times n}, d \in \mathbb{Z}^m$.

**Lemma 5** (Simulation for integral affine transformations)**.** *Let $P \subseteq \mathbb{R}^n$ be a polytope, $f : \mathbb{R}^n \longrightarrow \mathbb{R}^m$ an integral affine function, and denote $Q := f(P) \subseteq \mathbb{R}^m$. Let $\hat{\mathcal{T}}$ be any BB tree. Then, there exists a BB tree $\mathcal{T}$ such that $|\mathcal{T}| = |\hat{\mathcal{T}}|$ and*

$$f(\mathcal{T}(P)) \subseteq \hat{\mathcal{T}}(Q).$$

*Proof.* Let $f(x) = Cx + d$ with $C \in \mathbb{Z}^{m\times n}, d \in \mathbb{Z}^m$. We construct a BB tree $\mathcal{T}$, that satisfies the result of the lemma, with the same size as $\hat{\mathcal{T}}$ as follows: $\mathcal{T}$ has the same nodes as $\hat{\mathcal{T}}$ but each branching constraint $\langle a, y\rangle \le b$ of $\hat{\mathcal{T}}$ is replaced by the constraint $\langle C^T a, x\rangle \le b - \langle a, d\rangle$ in $\mathcal{T}$.

First we verify that $\mathcal{T}$ only uses legal disjunctions: First note that $C^T a \in \mathbb{Z}^n$ and $b - \langle a, d\rangle \in \mathbb{Z}$. If a node of $\hat{\mathcal{T}}$ has $\langle a, y\rangle \le b \ \vee \ \langle a, y\rangle \ge b + 1$ as its disjunction, the corresponding node in $\mathcal{T}$ has the disjunction $\langle C^T a, x\rangle \le b - \langle a, d\rangle \vee \langle -C^T a, x\rangle \le -b-1-\langle -a, d\rangle$ (notice $\langle a, y\rangle \ge b+1 \ \equiv\ \langle -a, y\rangle \le -b-1$). Since the second term in the latter disjunction is equivalent to $\langle C^T a, x\rangle \ge b - \langle a, d\rangle + 1$, we see that this disjunction is a legal one.

Now we show the desired claim. Let $S$ be the atom of a leaf $v$ of $\mathcal{T}$ and $\hat{S}$ be the atom of the corresponding leaf $\hat{v}$ of $\hat{\mathcal{T}}$. We show that for all $x \in S$, it must be that $f(x) \in \hat{S}$. To see this, notice that if $x$ satisfies an inequality $\langle C^T a, x\rangle \le b - \langle a, d\rangle$ then $f(x)$ satisfies $\langle a, f(x)\rangle \le b$:

$$\langle a, f(x)\rangle = \langle a, Cx + d\rangle = \langle a, Cx\rangle + \langle a, d\rangle = \langle C^T a, x\rangle + \langle a, d\rangle \le b.$$

Since any $x \in S$ belongs to $P$ and satisfies all the branching constraints of the leaf $v$, this implies $f(x)$ belongs to $Q$ and satisfies all the branching constraints of the leaf $\hat{v}$, and hence belongs to the atom $\hat{S}$. The result of the lemma follows. □

**Corollary 3.** *Let $P$, $Q$, and $f$ satisfy the assumptions of Lemma 5. Further, suppose $P$ and $Q$ are both integer infeasible (i.e. $P \cap \{0,1\}^n = \emptyset$ and $Q \cap \{0,1\}^n = \emptyset$). Then,*

$$BBhardness(Q) \geq BBhardness(P).$$

*Proof.* Let $\hat{\mathcal{T}}$ be the smallest BB tree such that $\hat{\mathcal{T}}(Q) = \emptyset$. Then, by Lemma 5, it must hold that $\mathcal{T}(P) = \emptyset$. The desired result follows. □

**Corollary 4.** *Let $P$, $Q$, and $f$ satisfy the assumptions of Lemma 5. Further, suppose there is some $x^* \in \mathbb{R}^n$ such that $x^* \notin P_I$ and $f(x^*) \notin Q_I$. Then,*

$$BBdepth(f(x^*), Q) \geq BBdepth(x^*, P).$$

*Proof.* Let $\hat{\mathcal{T}}$ be the smallest BB tree that separates $f(x^*)$ from $Q$ and $\mathcal{T}$ be constructed as in the proof of Lemma 5. By Lemma 5 and the fact that $f$ is affine, this implies that if $x \in \mathrm{conv}(\mathcal{T}(P))$ then $f(x) \in \mathrm{conv}(\hat{\mathcal{T}}(Q))$: there exists $x^1, ..., x^k \in \mathcal{T}(P)$ and $\lambda_1, ..., \lambda_k \in [0,1]$ s.t. $\sum_{i\in[k]} \lambda_i = 1$ such that $x = \sum_{i\in[k]} \lambda_i x^i$, and so

$$\begin{aligned} f(x) &= C(\sum_{i\in[k]} \lambda_i x^i) + d = \sum_{i\in[k]} \lambda_i (Cx^i) + \sum_{i\in[k]} \lambda_i d \\ &= \sum_{i\in[k]} \lambda_i (Cx^i + d) = \sum_{i\in[k]} \lambda_i f(x^i) \in \mathrm{conv}(\hat{\mathcal{T}}(Q)), \end{aligned}$$

where the last containment is due to Lemma 5. Since we know $f(x^*) \notin \mathrm{conv}(\hat{\mathcal{T}}(Q))$, this implies that $x^* \notin \mathrm{conv}(\mathcal{T}(P))$, namely $\mathcal{T}$ separates $x^*$ from $P$ as desired. □

**Lemma 6** (Hardness lemma). *Let $P \subseteq \mathbb{R}^n$ and $T \subseteq \mathbb{R}^m$ be polytopes and $f : \mathbb{R}^n \longrightarrow \mathbb{R}^m$ an integral affine function such that $f(P) \subseteq T$. Suppose $f$ is also one-to-one and $T \cap \mathbb{Z}^m \subseteq f(P \cap \mathbb{Z}^n)$. Then,*

$$BBrank(T) \geq BBrank(P).$$

*Proof.* First we show that $x \notin P_I$ implies $f(x) \notin T_I$ by proving the contrapositive. Suppose $f(x) \in T_I$; then $\exists y^1, ..., y^k \in T \cap \mathbb{Z}^m$ and $\lambda_1, ..., \lambda_k \in [0,1]$ s.t. $\sum_{i\in[k]} \lambda_i = 1$ such that $f(x) = \sum_{i\in[k]} \lambda_i y^i$. Since $T \cap \mathbb{Z}^m \subseteq f(P \cap \mathbb{Z}^n)$, for each $i$ there is $x^i \in P \cap \mathbb{Z}^n$ such that $y^i = f(x^i)$. Then

$$f(x) = \sum_{i\in[k]} \lambda_i f(x^i) = \sum_{i\in[k]} \lambda_i (Cx^i + d) = C \sum_{i\in[k]} \lambda_i x^i + d = f(\sum_{i\in[k]} \lambda_i x^i),$$

and so $f(x)$ belongs to $f(P_I)$. Since $f$ is one-to-one, this implies that $x$ belongs to $P_I$, as desired.

Let $x^* = \mathrm{argmax}_{x\in P\setminus P_I} \mathrm{BBdepth}(x, P)$. Since $x^* \notin P_I$, by the above claim $f(x^*) \notin T_I$. By assumption $f(P) \cap \mathbb{Z}^m \subseteq T \cap \mathbb{Z}^m$, and so $(f(P))_I \subseteq T_I$, and therefore $f(x^*) \notin (f(P))_I$. Then by Corollary 4, we have

$$\mathrm{BBdepth}(f(x^*), f(P)) \geq \mathrm{BBdepth}(x^*, P).$$

Since by assumption $f(P) \subseteq T$, $f(x^*) \in f(P)$, and $f(x^*) \notin T_I$, by Corollary 2, we have

$$\mathrm{BBdepth}(f(x^*), T) \geq \mathrm{BBdepth}(f(x^*), f(P)).$$

Finally, putting it all together

$$\begin{aligned} \mathrm{BBrank}(T) = \max_{y\in T\setminus T_I} \mathrm{BBdepth}(y, T) &\geq \mathrm{BBdepth}(f(x^*), T) \geq \mathrm{BBdepth}(f(x^*), f(P)) \geq \mathrm{BBdepth}(x^*, P) \\ &= \max_{x\in P\setminus P_I} \mathrm{BBdepth}(x, P) = \mathrm{BBrank}(P). \end{aligned}$$

□

In the rest of the paper, we will use Corollary 3, Corollary 4 or Lemma 6 together with some appropriate affine transformation to reduce the BB hardness of one problem to another. The three affine one-to-one functions we will use (and there compositions) are Flipping, Embedding, and Duplication as defined below.

**Definition 4** (Flipping)**.** *We say $f : [0,1]^n \longrightarrow [0,1]^n$ is a* flipping operation *if it "flips" some coordinates. That is, there exists $J \subseteq [n]$ such that*

$$y = f(x) \implies y_i = \begin{cases} x_i & \text{if } i \notin J \\ 1 - x_i & \text{if } i \in J \end{cases}.$$

*In other words, $f(x) = Cx + d$, where*

$$C^i = \begin{cases} e_i & \text{if } i \notin J \\ -e_i & \text{if } i \in J \end{cases}$$

$$d_i = \begin{cases} 0 & \text{if } i \notin J \\ 1 & \text{if } i \in J. \end{cases}.$$

**Definition 5** (Embedding)**.** *We say $f : [0,1]^n \longrightarrow [0,1]^{n+k}$ is an* embedding operation *if*

$$y = f(x) \implies y_i = \begin{cases} x_i & \text{if } 1 \leq i \leq n \\ 0 & \text{if } n < i \leq n + k_1 \\ 1 & \text{if } n + k_1 < i \leq n + k \end{cases},$$

*for some $0 \leq k_1 \leq k$. In other words, $f(x) = Cx + d$, where*

$$C^i = \begin{cases} e_i & \text{if } 1 \leq i \leq n \\ 0 & \text{otherwise} \end{cases}$$

$$d_i = \begin{cases} 1 & \text{if } n + k_1 < i \leq n + k \\ 0 & \text{otherwise} \end{cases}.$$

*Note that we can always renumber the coordinates so that the additional coordinates with values 0 or 1 are interspersed with the original ones and not grouped at the end.*

**Definition 6** (Duplication)**.** *Consider a $k$-tuple of coordinates $\{j_1, ..., j_k\}$ that are not necessarily distinct, where $j_i \in \{1, ..., n\}$ for $i = 1, ..., k$. We say that $f : [0,1]^n \longrightarrow [0,1]^{n+k}$ is a* duplication operation *using this tuple if*

$$y = f(x) \implies y_i = \begin{cases} x_i & \text{if } 1 \leq i \leq n \\ x_{j_{i-n}} & \text{if } n < i \leq n + k \end{cases}.$$

*Further, let $J_j = \{i \in \{1, ..., k\} : y_{n+i} = x_j\}$ be the indices of $y$ that are duplicates of $x_j$. In other words, $f(x) = Cx$, where*

$$C^i = \begin{cases} e_i & \text{if } 1 \leq i \leq n \\ e_1 & \text{if } i - n \in J_1 \\ \vdots & \\ e_n & \text{if } i - n \in J_n \end{cases}$$

## 4 BB hardness for packing polytopes and Set Cover

This section will begin by presenting a packing polytope with BBrank of $2^{\Omega(n)}$. The proof of this result will be based on a technique developed by Dadush et al. [12]. Then we will employ affine maps that satisfy Lemma 6 to obtain lower bounds on BBrank for a set cover instance.

We present a slightly generalized version of a key result from Dadush et al. [12]. The proof is essentially the same as of the original version, but we present it for completeness.

**Lemma 7** (Generalized Dadush-Tiwari Lemma)**.** *Let $P \subseteq \mathbb{R}^n$ be an integer infeasible non-empty polytope (i.e. $P \cap \mathbb{Z}^n = \emptyset$). Further, suppose $P$ is defined by the set of constraints $C_P$ (i.e. $P = \{x : C_P\}$) and let $D \subseteq C_P$ be a subset of constraints such that if we remove any constraint in $D$, the polytope becomes integer feasible (i.e. for all subsets $C \subset C_P$ such that $D \setminus C \neq \emptyset$, it holds that $\{x : C\} \cap \{0,1\}^n \neq \emptyset$). Then, any branch-and-bound tree $\mathcal{T}$ proving the integer infeasibility of $P$ has at least $\frac{|D|}{n}$ leaf nodes, that is $|\mathcal{T}| \geq 2\frac{|D|}{n} - 1$.*

*Proof.* Let $\mathcal{T}$ denote any branch-and-bound proof of infeasibility for $P$ and let $N$ denote the number of leaf nodes of $\mathcal{T}$. Suppose for sake of contradiction, that $N < \frac{|D|}{n}$. Consider any leaf node of $\mathcal{T}$, denoted $v$. Let $C_v$ be the set of branching constraints on the path to $v$. Since $v$ is a leaf and $\mathcal{T}$ is a proof of infeasibility, we note $\{x : C_v \cup C_P\} = \{x : C_v\} \cap P = \emptyset$.

By Helly's Theorem, there exists a set of $n+1$ constraints $K_v \subseteq C_v \cup C_P$ such that $\{x : K_v\} = \emptyset$. Also, we see that

$$|K_v \cap C_P| \leq n, \quad \forall v \in \text{leaves}(\mathcal{T}). \tag{2}$$

This is because if we had $|K_v \cap C_P| = n+1$, this would imply $K_v \subseteq C_P$, hence $\{x : C_P\} \subseteq \{x : K_v\}$, and since $\{x : K_v\} = \emptyset$; this would imply $\{x : C_P\} = P = \emptyset$, which is clearly a contradiction since we know $P$ is non-empty.

Next observe $\mathcal{T}$ certifies that the set $\tilde{P} := \left\{x : \bigcup_{v \in \text{leaves}(\mathcal{T})}(K_v \cap C_P)\right\}$ is integer infeasible, since $\tilde{P} \cap \{x : C_u\} = \left\{x : C_u \cup \bigcup_{v \in \text{leaves}(\mathcal{T})}(K_v \cap C_P)\right\} \subseteq \{x : K_u\} = \emptyset$ for all $u \in \text{leaves}(\mathcal{T})$.

On other hand, observe that by (2) we have that $|\bigcup_{v \in \text{leaves}(\mathcal{T})}(K_v \cap C_P)| \leq nN < |D|$, so by the assumption of the lemma, we see that $\{x : \bigcup_{v \in \text{leaves}(\mathcal{T})}(K_v \cap C_P)\}$ contains an integer point, a contradiction.

□

### 4.1 Packing polytopes

Consider the following packing polytope

$$P_{PA} = \{x \in [0,1]^n : \sum_{i \in S} x_i \leq k-1 \text{ for all } S \subseteq [n], |S| = k\},$$

where we assume $2 \leq k \leq \frac{n}{2}$.

**Lemma 8.** *There exists an $x^* \in P_{PA} \setminus (P_{PA})_I$ such that any branch-and-bound tree that separates $x^*$ from $P_{PA}$ has at least $\frac{2}{n}\left(\binom{n}{k}+1\right) - 1$ nodes. Therefore, BBrank$(P_{PA}) \geq \frac{2}{n}\left(\binom{n}{k}+1\right) - 1$.*

The starting point for proving this lemma is the following following proposition.

**Proposition 1.** *Any branch-and-bound tree proving the infeasibility of $Q = P_{PA} \cap \{x : \langle \mathbf{1}, x \rangle \geq k\}$ has at least $\frac{2}{n}(\binom{n}{k}+1) - 1$ nodes.*

*Proof.* We show that $Q$ satisfies all of the requirements of Lemma 7.

First we show that $Q \neq \emptyset$. Consider the point $\hat{x} \in \mathbb{R}^n$ where $\hat{x}_i = \frac{k}{n}$ for $i \in [n]$. Then, for any $S \subseteq [n]$ with $|S| = k$, we have $\sum_{i\in S} \hat{x}_i = k \cdot \frac{k}{n} \leq k \cdot \frac{1}{2} \leq k-1$, where the last two inequalities are implied by the assumption $2 \leq k \leq \frac{n}{2}$. Also, $\sum_{i\in[n]} \hat{x}_i = k$. Thus, $\hat{x}$ satisfies all the constraints of $Q$.

Next we show $Q \cap \{0,1\}^n = \emptyset$. Suppose for sake of contradiction there is some $x^* \in Q \cap \{0,1\}^n$. Since $\sum_{i\in[n]} x_i^* \geq k$, there is a set $S \subseteq [n]$ of size $k$ such that $\sum_{i\in S} x_i^* = k$. This violates the cardinality constraint corresponding to $S$, so $x^* \notin Q$, a contradiction.

Finally, we show that there is a set of $\binom{n}{k} + 1$ constraints $D$ such that removing any of these constraints makes $Q$ integer feasible. Suppose we remove the constraint $\sum_{i\in S} x_i \leq k-1$, denote this new polytope $Q'$. Then let $x_i^* = 1$ for all $i \in S$ and $x_i^* = 0$ for all $i \notin S$. Clearly $\sum_{i\in[n]} x_i^* \geq k$ and since for all $S' \subseteq [n], |S'| = k$ it holds that $|S' \cap S| \leq k-1$, it is also the case that $\sum_{i\in S'} x_i^* \leq k-1$. So $x^* \in Q' \cap \{0,1\}^n$. Now suppose we remove instead the constraint $\sum_{i\in[n]} x_i \geq k$, resulting in polytope $P_{PA}$. Clearly $P_{PA}$ is down monotone, and therefore $0 \in P_{PA}$.

Finally by Lemma 7, any branch-and-bound proof of infeasibility for $Q$ has at least $\frac{2}{n}(\binom{n}{k} + 1) - 1$ nodes. □

Now, combining Proposition 1 with Lemma 2, we are ready to prove Lemma 8.

*Proof of Lemma 8.* We will show that $P_{PA}, (\mathbf{1}, k-1)$ satisfy the conditions on $P, (c, \delta)$ set by Lemma 2. First, $\langle \mathbf{1}, x\rangle \leq k-1$ is a valid inequality for $(P_{PA})_I$: this follows from the integer infeasibility of $Q = P_{PA} \cap \{x : \langle \mathbf{1}, x\rangle \geq k\}$, as proven in Proposition 1. In the following paragraph we will show that $\{x \in (P_{PA})_I : \langle \mathbf{1}, x\rangle = k-1\}$ has dimension $n-1$, that is, $\langle \mathbf{1}, x\rangle \leq k-1$ is facet-defining for $(P_{PA})_I$. With this at hand we can apply Lemma 2 to obtain

$$\text{BBrank}(P_{PA}) \geq \text{BBhardness}(P_{PA} \cap \{x : \langle \mathbf{1}, x\rangle \geq k\}) = \frac{2}{n}\left(\binom{n}{k} + 1\right) - 1$$

where the last inequality follows from Proposition 1.

To show facet-defining, let $T \subseteq [n]$ be such that $|T| = k-1$. Let $\chi(T)$ denote the characteristic vector of $T$, so that $\chi(T)_i = 1$ if and only if $i \in T$. We know that all these point belong to the hyperplane $\{x : \langle \mathbf{1}, x\rangle = k-1\}$. Thus, there can be at most $n$ affinely independent points among $\{\chi(T)\}_{T\subseteq[n],|T|=k-1}$. We first verify that there are exactly $n$ affinely independent points among $\{\chi(T)\}_{T\subseteq[n],|T|=k-1}$ by showing that the affine hull of the points in $\{\chi(T)\}_{T\subseteq[n],|T|=k-1}$ is the hyperplane $\{x : \langle \mathbf{1}, x\rangle = k-1\}$. Consider the system in variables $a, b$:

$$\langle a, \chi(T)\rangle = b, \quad \forall T \subseteq [n] \text{ such that } |T| = k-1.$$

We have to show that all solutions of the above system are a scaling of $(\mathbf{1}, k-1)$. For that, let $T^1 = \{1, \dots, k-1\}$ and $T^2 := \{2, \dots, k\}$. Subtracting the equation corresponding to $T^1$ from that of $T^2$, we obtain $a_1 = a_k$. Using the same argument by suitably selecting $T^1$ and $T^2$, we obtain: $a_1 = a_2 = \cdots = a_n$. Therefore, without loss of generality, we may rescale all the $a_i$'s to 1. Then we see the only possible value for $b$ is $k-1$. This shows that the only affine subspace containing the points $\{\chi(T)\}_{T\subseteq[n],|T|=k-1}$ is $\{x : \langle \mathbf{1}, x\rangle = k-1\}$, in other words, there are $n$ affinely independent points among them.

□

Finally, the following simple corollary to Lemma 8 gives the desired hardness bound.

**Corollary 5.** *Consider the polytope $P_{PA} = \{x \in [0,1]^n : \sum_{i\in S} x_i \leq \frac{n}{2}$ for all $S \subseteq [n], |S| = \frac{n}{2} + 1\}$. Then, BBrank$(P_{PA}) \geq 2^{\Omega(n)}$, i.e. there exists a $c \in \mathbb{R}^n$ such that the smallest branch-and-bound tree that solves*

$$\max_{x\in P_{PA}\cap\{0,1\}^n} \langle c, x\rangle$$

*has size at least $2^{\Omega(n)}$.*

## 4.2 Set Cover

In order to obtain a Set Cover instance that requires an exponential-size branch-and-bound tree, we will use Lemma 6 together with the flipping affine mapping (Defintion 4) applied to the packing instance from Section 4.1.

**Proposition 2.** *Let $P_{PA}$ still be the packing polytope from Section 4.1. Let $f : [0,1]^n \longrightarrow [0,1]^n$ be the flipping function with $J = [n]$. Then:*

- $T_{SC} = f(P_{PA})$*, where*

$$T_{SC} = \{y \in [0,1]^n \ : \ \sum_{i \in S} y_i \geq 1 \text{ for all } S \subseteq [n], |S| = k\}.$$

- $T_{SC} \cap \{0,1\}^n \subseteq f(P_{PA} \cap \{0,1\}^n)$.

*Proof.* By substituting $y_i = 1 - x_i$ for $i \in [n]$ in the polytope $P_{PA}$, we obtain that $T_{\text{SC}} = f(P_{PA})$.

Now consider any $y \in T_{\text{SC}} \cap \{0,1\}^n$. Notice $x = 1 - y \in P_{PA} \cap \{0,1\}^n$ and $y = f(x)$, and hence $y \in f(P_{PA} \cap \{0,1\}^n)$. This gives $T_{\text{SC}} \cap \{0,1\}^n \subseteq f(P_{PA} \cap \{0,1\}^n)$. □

Then by Lemma 6, we have that $\text{BBrank}(T_{\text{SC}}) \geq \text{BBrank}(P_{PA}) \geq 2^{\Omega(n)}$.

**Corollary 6.** *$BBrank(T_{SC}) \geq BBrank(P) \geq 2^{\Omega(n)}$, i.e. there exists a $c \in \mathbb{R}^n$ such that the smallest branch-and-bound tree that solves*

$$\max_{x \in T_{SC} \cap \{0,1\}^n} \langle c, x \rangle$$

*has size at least $2^{\Omega(n)}$.*

# 5 BB hardness for Cross Polytope

In this section, we present in Proposition 3 a simple proof of BB hardness for the cross polytope. As discussed before, this result slightly improves on the result when Lemma 7 is applied to the cross polytope.

Next in this section we develop Proposition 4 that shows that there is a point in the cross polytope that is hard to separate using BB tree of small size. This allows us to use the machinery of Lemma 6 and a composition of affine functions described in Section 3 to reduce the BBhardness of the cross polytope to TSP, which we do in Section 7.

The cross polytope is defined as

$$P_n = \left\{ x \in [0,1]^n \ : \ \sum_{i \in J} x_i + \sum_{i \notin J} (1 - x_i) \geq \frac{1}{2} \quad \forall J \subseteq \{1, ..., n\} \right\}.$$

**Proposition 3.** *Let $\mathcal{T}$ be a BB tree for $P_n$ that certifies the integer infeasibility of $P_n$. Then $|\mathcal{T}| \geq 2^{n+1} - 1$ (i.e. BBhardness$(P_n) \geq 2^{n+1} - 1$).*

*Proof.* In order to certify the integer infeasibility of $P_n$, every leaf node's atom must be an empty set. We will verify that in order for the atom of a leaf $v$ to be empty, no more than one integer point is allowed to satisfy the branching constraints $C_v$ of $v$. This will complete the proof, since we then must have at least $2^n$ leaves.

Consider any leaf $v$ of $\mathcal{T}$ such that two distinct integer points are feasible for its branching constraints. Then the average of these two points is a point in $\{0, 1, \frac{1}{2}\}^n$ with at least one component equal to $\frac{1}{2}$, which also satisfies the branching constraints. However, a point in $\{0, 1, \frac{1}{2}\}^n$ with at least one component equal to $\frac{1}{2}$ satisfies the constraints defining $P_n$. Thus the atom of the leaf $v$ is non-empty. □

**Corollary 7.** *Let $F \subseteq \mathbb{R}^n$ be a face of $P_n$ with dimension $d$. Then BBhardness$(F) \geq 2^{d+1} - 1$.*

*Proof.* Notice that $F$ is a copy of $P_d$ with $n-d$ components fixed to 0 or 1. Thus, when we consider $P_d$ and apply the appropriate embedding affine transformation (Definition 5), call it $f$, we obtain $f(P_d) = F$. Also since $F \cap \mathbb{Z}^n = \emptyset$, we obtain that $f$, $P_d$ and $F$ satisfy all the conditions of Corollary 3. Thus, BBhardness$(F) \geq$ BBhardness$(P_d) \geq 2^{d+1} - 1$, where the last inequality follows from Proposition 3. □

Next we show that the point $\frac{1}{2}\mathbf{1}$ is hard to separate from $P_n$. For that we need a technical result that any halfspace that contains $\frac{1}{2}\mathbf{1}$ must also contain a face of $[0,1]^n$ of dimension at least $\lceil n/2 \rceil$.

**Lemma 9.** *Consider any $(\pi, \pi_0) \in \mathbb{R}^n \times \mathbb{R}$ such that $\langle \pi, \frac{1}{2}\mathbf{1} \rangle > \pi_0$. Let $G = \{x \in [0,1]^n : \langle \pi, x \rangle > \pi_0\}$. There exists a face $F$ of $[0,1]^n$ of dimension at least $\lceil n/2 \rceil$ contained in $G$.*

*Proof.* We first consider the case where $n$ is even.

Note that $\pi_0 = \sum_{i=1}^n \frac{1}{2}\pi_i - \epsilon$ for some $\epsilon > 0$. WLOG, we may assume $|\pi_i| \geq |\pi_{i+1}|$ for odd values of $i \in [n]$. Let $O \subseteq [n]$ be the set of odd indices, i.e., $i = \{1, 3, 5, \ldots, n-1\}$. Now consider the following face $F$ of $[0,1]^n$ of dimension $\frac{n}{2}$:

$$F = \left\{ x \in [0,1]^n \;\middle|\; x_i = \phi_i, \quad \forall i \in O, \text{where } \begin{array}{ll} \phi_i = 1 & \text{if } \pi_i \geq 0, \\ \phi_i = 0 & \text{if } \pi_i < 0 \end{array} \right\}$$

In order to prove that $F \subseteq G$, note that is sufficient to verify that:

$$\pi_i \phi_i + \pi_{i+1} x_{i+1} \geq \frac{1}{2}(\pi_i + \pi_{i+1}), \quad \forall x_{i+1} \in [0,1], i \in O, \tag{3}$$

since then we have that

$$\sum_{i \in O} (\pi_i \phi_i + \pi_{i+1} x_{i+1}) > \sum_{i \in O} \frac{1}{2}(\pi_i + \pi_{i+1}) - \epsilon = \pi_0, \quad \forall x_{i+1} \in [0,1], i \in O.$$

In order to verify (3), we use the fact $|\pi_i| \geq |\pi_{i+1}|$ and consider the following cases:

- $\pi_i \geq 0$:
  - $\pi_{i+1} \geq 0$: In this case, it is sufficient to verify $\pi_i + \pi_{i+1} \cdot 0 \geq \frac{1}{2}(\pi_i + \pi_{i+1})$ which is true since $\pi_i \geq \pi_{i+1}$.
  - $\pi_{i+1} \leq 0$: In this case, it is sufficient to verify $\pi_i + \pi_{i+1} \cdot 1 \geq \frac{1}{2}(\pi_i + \pi_{i+1})$ which is true since $\pi_i + \pi_{i+1} \geq 0$.
- $\pi_i \leq 0$:
  - $\pi_{i+1} \geq 0$: In this case, it is sufficient to verify $\pi_{i+1} \cdot 0 \geq \frac{1}{2}(\pi_i + \pi_{i+1})$ which is true since $\pi_i + \pi_{i+1} \leq 0$.
  - $\pi_{i+1} \leq 0$: In this case, it is sufficient to verify $\pi_{i+1} \cdot 1 \geq \frac{1}{2}(\pi_i + \pi_{i+1})$ which is true since $\pi_i \leq \pi_{i+1}$.

The same proof can be used when $n$ is odd. Apart from the first $n-1$ indices in this case set $\phi_n = 1$ if $\pi_n \geq 0$ or $\phi_n = 0$ if $\pi_n < 0$, which gives a desired face of dimension $\frac{n}{2} + 1$. □

**Proposition 4.** *BBdepth$(\frac{1}{2}\mathbf{1}, P_n) \geq 2^{\lceil n/2 \rceil + 1} - 1$.*

*Proof.* For sake of contradiction suppose there exists a tree $\mathcal{T}$ of size less than $2^{\lfloor n/2 \rfloor+1} - 1$ such that $\frac{1}{2}\mathbf{1} \notin \mathrm{conv}(\mathcal{T}(P_n))$. By the separation theorem, there exists $(\pi, \pi_0) \in \mathbb{R}^n \times \mathbb{R}$ such that $\langle \pi, \frac{1}{2}\mathbf{1} \rangle > \pi_0$ and $\langle \pi, x \rangle \leq \pi_0$ for all $x \in \mathrm{conv}(\mathcal{T}(P_n))$. By Lemma 9, let $F$ be a face of $[0,1]^n$ of dimension $\lceil n/2 \rceil$ contained in $\{x \in \mathbb{R}^n \,|\, \langle \pi, x \rangle > \pi_0\}$; notice that $P_n \cap F$ is a face of $P_n$ of the same dimension. Since $\mathcal{T}(P_n) \subseteq \mathrm{conv}(\mathcal{T}(P_n)) \subseteq \mathbb{R}^n \setminus \{x \in \mathbb{R}^n \,|\, \langle \pi, x \rangle > \pi_0\} \subseteq \mathbb{R}^n \setminus F$ and $\mathcal{T}(F) \subseteq F$, by Lemma 4 we get $\mathcal{T}(P_n \cap F) \subseteq \mathcal{T}(P_n) \cap \mathcal{T}(F) = \emptyset$, i.e. the atoms of the leaves of $\mathcal{T}$ applied to $P_n \cap F$ are all empty. Thus, $\mathcal{T}$ is a branch-and-bound tree to certify the infeasibility of $P_n \cap F$ of size less than $2^{\lceil n/2 \rceil+1} - 1$. However, this contradicts Corollary 7.

□

## 6 BB hardness for Perturbed Cross Polytope

We now show that exponential BB hardness for the cross polytope persists even after adding Gaussian noise to the entries of the contraint matrix. This implies an exponential lower bound even for a "smoothed analysis" of general branch-and-bound.

We consider the consider the cross polytope $(Ax \leq b) \cap [0,1]^n$ and add to each entry in $A$ an independent gaussian noise $N(0, 1/20^2)$; actually the RHSs will be something like $\frac{n}{20}$ instead of the traditional $\frac{1}{2}$. This gives the following random polytope $Q$:

$$Q \equiv \sum_{i \in I} \left(1 + N\left(0, \tfrac{1}{20^2}\right)\right) x_i + \sum_{i \notin I} \left(1 - \left(1 + N\left(0, \tfrac{1}{20^2}\right)\right) x_i\right) \geq \frac{1.6n}{20}, \quad \forall I \subseteq [n]$$
$$x \in [0,1]^n,$$

where each occurrence of $N(0, \frac{1}{20^2})$ is independent.

**Theorem 2.** *With probability at least $1 - \frac{2}{e^{n/2}}$ the polytope $Q$ is integer infeasible and every BB tree proving its infeasibility has at least $2^{\Omega(n)}$ nodes.*

We need the following standard tail bound for the Normal distribution (see equation (2.10) of [25]).

**Fact 1.** *Let $X \sim N(0, \sigma^2)$ be a mean zero Gaussian with variance $\sigma^2$. Then for every $p \in (0,1)$, with probability at least $1 - p$ we have $X \leq \sigma\sqrt{2\ln(1/p)}$, and with probability at least $1 - p$ we have $X \geq -\sigma\sqrt{2\ln(1/p)}$.*

Let $LHS_I(x)$ be the LHS of the constraint $I$ evaluated at $x$.

**Lemma 10.** *With probability at least $1 - \frac{1}{e^{n/2}}$ the polytope $Q$ is integer infeasible.*

*Proof.* Fix a 0/1 point $x \in \{0,1\}^n$, and let $I \subseteq [n]$ be the set of coordinates $i$ where $x_i = 0$. Let $I^c = [n] \setminus I$. Notice $LHS_I(x)$ is a Gaussian with mean 0 and variance $\frac{|I^c|}{20^2} \leq \frac{n}{20^2}$, and so with probability at least $1 - \frac{1}{e^{n/2} 2^n}$ we have

$$LHS_I(x) \leq \frac{\sqrt{n}}{20}\sqrt{2\ln(e^{n/2}2^n)} = \frac{\sqrt{n}}{20}\sqrt{(1 + 2\ln 2)n} < \frac{1.6n}{20},$$

i.e., does not satisfy this inequality, so does not belong to $Q$. Taking a union bound over all $2^n$ points $x \in \{0,1\}^n$, with probability at least $1 - \frac{1}{e^{n/2}}$ none of them belong to $Q$. □

**Lemma 11.** *With probability at least $1 - \frac{1}{e^{n/2}}$ the polytope $Q$ contains all points $\{0, \frac{1}{2}, 1\}^n$ that have at least $s = \frac{4n}{10}$ coordinates with value $\frac{1}{2}$. (We call this set of points* $\mathrm{Half}_s$.*)*

*Proof.* Consider $x \in \text{Half}_s$. Fix $I \subseteq [n]$. Let $n_{\text{half}} \geq s$ be the number of coordinates of $x$ with value $\frac{1}{2}$, $n_{\text{ones}}$ be the number of coordinates with value 1, and let $n_{\text{diff}}$ be the number of coordinates $i$ where either $i \in I$ and $x_i = 1$, or $i \notin I$ and $x_i = 0$. Recalling that for independent Gaussians $Y \sim N(a, b)$ and $Y' \sim N(a', b')$ their sum $Y + Y'$ has is distributed as $N(a + a', b + b')$, we see that $LHS_I(x)$ distributed as

$$\begin{aligned} LHS_I(x) &=_d \frac{n_{\text{half}}}{2} + n_{\text{diff}} + \frac{1}{2} N(0, \tfrac{n_{\text{half}}}{20^2}) + N(0, \tfrac{n_{\text{ones}}}{20^2}) \\ &=_d \frac{n_{\text{half}}}{2} + n_{\text{diff}} + N(0, (\tfrac{n_{\text{half}}}{4} + n_{\text{ones}}) \cdot \tfrac{1}{20^2}), \end{aligned}$$

where again the occurrences of $N(0, \cdot)$ are independent. Since the last term is a Gaussian with variance at most $\frac{n}{20^2}$, we get that with probability at least $1 - \frac{1}{e^{n/2} \cdot 2^n \cdot 3^n}$

$$LHS_I(x) \geq \frac{n_{\text{half}}}{2} + n_{\text{diff}} - \frac{1}{20} \sqrt{n} \sqrt{2 \log(e^{n/2} \cdot 2^n \cdot 3^n)} \geq \frac{4n}{20} - \frac{2.4n}{20} = \frac{1.6n}{20},$$

that is, $x$ satisfies the constraint of $Q$ indexed by $I$.

Taking a union bound over all $x \in \text{Half}_s$ and all subsets $I \subseteq [n]$, we see that all points in $\text{Half}_s$ satisfy all constraints of $Q$ with probability at least $1 - \frac{1}{e^{n/2}}$. This concludes the proof. □

**Lemma 12.** *Let $F \subseteq \{0, 1\}^n$ be a set of 0/1 points. For any $k$, if $|F| > \sum_{i \leq k-1} \binom{n}{i}$, conv$(F)$ contains a point with at least $k$ coordinates of value 1/2.*

*Proof.* By the Sauer-Shelah Lemma (Lemma 11.1 of [19]), there is a set of coordinates $J \subseteq [n]$ of size $|J| = k$ such that the points in $F$ take all possible values in coordinates $J$, i.e., the projection $F_J$ onto the coordinates $J$ equals $\{0, 1\}^k$. So the point $\frac{1}{2}\mathbf{1}$ belongs to conv$(F_J)$, which implies that conv$(F)$ has the desired point. □

*Proof of Theorem 2.* Let $E$ be the event that both the bounds from Lemmas 10 and 11 hold. By union bound this event happens with probability at least $1 - \frac{2}{e^{n/2}}$. So it suffices to show that there is a constant $c > 0$ such that for every scenario in $E$, every BB tree proving the infeasibility of $Q$ has at least $2^{cn}$ leaves.

In hindsight, set $c := 1 - h(\frac{s}{n})$, where $h$ is the binary entropy function $h(p) := p \log \frac{1}{p} + (1-p) \log \frac{1}{1-p}$. Notice that $c > 0$, since $h$ is strictly increasing in the interval $[0, \frac{1}{2}]$ and hence $h(\frac{s}{n}) < h(\frac{1}{2}) = 1$.

Consider any tree $\mathcal{T}$ that proves integer infeasibility of $Q$, and we claim that it has more than $2^{cn}$ leaves. By contradiction, suppose not. Then $\mathcal{T}$ has a leaf $v$ whose branching constraints $C_v$ are satisfied by at least $\frac{2^n}{2^{cn}} = 2^{n \cdot h(s/n)}$ 0/1 points. But since $2^{n \cdot h(s/n)} > \sum_{i \leq s-1} \binom{n}{i}$ (see e.g. Lemma 5 of [17]), by Lemma 12 we know $\{x : C_v\}$ contains a point $\bar{x} \in [0, 1]^n$ with at least $s$ coordinates of value 1/2. Moreover, notice that $\bar{x}$ also belongs to conv$(\text{Half}_s)$, which is contained in $Q$ by assumption of $E$. Hence $\bar{x} \in \{x : C_v\} \cap Q$, namely the atom of $v$. But this contradicts that this atom is empty (which is required since $\mathcal{T}$ proves integer infeasibility of $Q$). □

## 7 BB hardness for TSP

**Proposition 5.** *Let $f$ be any composition of flipping (Defintion 4), embedding (Definition 5), and duplication (Definition 6) operations. Let $H \subseteq [0, 1]^n$ be a polytope such that $f(P_k) \subseteq H$ and $f(\frac{1}{2}\mathbf{1}) \notin H_I$, where $k \leq n$. Then, BBrank$(H) \geq 2^{\lceil k/2 \rceil}$.*

*Proof.* Notice that if $f$ is a composition of flipping, embedding, and duplication operations, it holds that $f$ is an integral affine transformation.

Note that, if $P$ is integer infeasible, $f(P)$ is also integer infeasible. In particular, since $P_k$ is integer infeasible we have $(f(P_k))_I = \emptyset$, and hence $f(\frac{1}{2}\mathbf{1}) \notin (f(P_k))_I$. Now, Corollary 4 and Proposition 4 give us $\text{BBdepth}\left(f\left(\frac{1}{2}\mathbf{1}\right), f(P_k)\right) \geq \text{BBdepth}\left(\frac{1}{2}\mathbf{1}, P_k\right) \geq 2^{\lceil k/2 \rceil + 1} - 1$. Finally, since $f(P_k) \subseteq H$, Corollary 2 implies that $\text{BBdepth}(f(\frac{1}{2}\mathbf{1}), H) \geq 2^{\lceil k/2 \rceil + 1} - 1$. This implies the desired result $\text{BBrank}(H) \geq 2^{\lceil k/2 \rceil + 1} - 1 \geq 2^{\lceil k/2 \rceil}$. □

We next present a key result from Section 4 of [9] (see also [7]), that shows how we can apply Proposition 5 to obtain BB hardenss of the TSP polytope. Let $T_{\mathrm{TSP}_n}$ be the standard LP relaxation of the TSP polytope (using subtour elimination constraints) for $n$ cities:

$$\begin{aligned} x(\delta(v)) &= 2 \quad \forall v \in V \\ x(\delta(W)) &\geq 2 \quad \forall W \subset V, W \neq \emptyset \\ 0 \leq x(e) &\leq 1 \quad \forall e \in E \end{aligned}$$

**Proposition 6** ([9]). *There exists a function $f$ which is a composition of flipping, embedding, and duplication such that $f(P_{\lfloor n/8 \rfloor})$ is contained in $T_{TSP_n}$ and $f(\frac{1}{2}\mathbf{1})$ does not belong to the integer hull of $T_{TSP_n}$.*

Employing Proposition 5 we obtain a BB hardness for TSP.

**Corollary 8** (BB hardness for TSP). *$BBrank(T_{TSP_n}) \geq 2^{\lfloor n/16 \rfloor}$, i.e, there is a $c \in \mathbb{R}^{n(n-1)/2}$ such that the smallest branch-and-bound tree that solves*

$$\max_{x \in T_{TSP_n} \cap \{0,1\}^{n(n-1)/2}} \langle c, x \rangle$$

*has size at least $2^{\lfloor n/16 \rfloor}$.*

# References


[1] K. Aardal, R. E. Bixby, C. A. Hurkens, A. K. Lenstra, and J. W. Smeltink. Market split and basis reduction: Towards a solution of the cornuéjols-dawande instances. *INFORMS Journal on Computing*, 12(3):192–202, 2000.

[2] A. Basu, M. Conforti, M. Di Summa, and H. Jiang. Complexity of branch-and-bound and cutting planes in mixed-integer optimization–ii. *arXiv preprint arXiv:2011.05474*, 2020.

[3] A. Basu, M. Conforti, M. Di Summa, and H. Jiang. Complexity of cutting planes and branch-and-bound in mixed-integer optimization. *arXiv preprint arXiv:2003.05023*, 2020.

[4] P. Beame, N. Fleming, R. Impagliazzo, A. Kolokolova, D. Pankratov, T. Pitassi, and R. Robere. Stabbing planes. In *9th Innovations in Theoretical Computer Science Conference (ITCS 2018)*. Schloss Dagstuhl-Leibniz-Zentrum fuer Informatik, 2018.

[5] S. Borst, D. Dadush, S. Huiberts, and S. Tiwari. On the integrality gap of binary integer programs with gaussian data. *arXiv preprint arXiv:2012.08346*, 2020.

[6] V. Chvátal. Hard knapsack problems. *Operations Research*, 28(6):1402–1411, 1980.

[7] V. Chvátal, W. Cook, and M. Hartmann. On cutting-plane proofs in combinatorial optimization. *Linear algebra and its applications*, 114:455–499, 1989.

[8] M. Conforti, G. Cornuéjols, and G. Zambelli. *Integer programming*, volume 271. Springer, 2014.

[9] W. Cook and S. Dash. On the matrix-cut rank of polyhedra. *Mathematics of Operations Research*, 26(1):19–30, 2001.

[10] W. J. Cook and M. Hartmann. On the complexity of branch and cut methods for the traveling salesman problem. *Polyhedral Combinatorics*, 1:75–82, 1990.

[11] G. Cornuéjols, L. Liberti, and G. Nannicini. Improved strategies for branching on general disjunctions. *Math. Program.*, 130(2):225–247, 2011.

[12] D. Dadush and S. Tiwari. On the complexity of branching proofs. In *Proceedings of the 35th Computational Complexity Conference*, pages 1–35, 2020.

[13] S. Dash. Exponential lower bounds on the lengths of some classes of branch-and-cut proofs. *Mathematics of Operations Research*, 30(3):678–700, 2005.

[14] S. Dash, O. Günlük, and M. Molinaro. On the relative strength of different generalizations of split cuts. *Discrete Optimization*, 16:36–50, 2015.

[15] S. S. Dey, Y. Dubey, and M. Molinaro. Branch-and-bound solves random binary ips in polytime. In *Proceedings of the 2021 ACM-SIAM Symposium on Discrete Algorithms (SODA)*, pages 579–591. SIAM, 2021.

[16] N. Fleming, M. Göös, R. Impagliazzo, T. Pitassi, R. Robere, L.-Y. Tan, and A. Wigderson. On the power and limitations of branch and cut. *arXiv preprint arXiv:2102.05019*, 2021.

[17] L. Gottlieb, A. Kontorovich, and E. Mossel. VC bounds on the cardinality of nearly orthogonal function classes. *Discret. Math.*, 312(10):1766–1775, 2012.

[18] R. G. Jeroslow. Trivial integer programs unsolvable by branch-and-bound. *Mathematical Programming*, 6(1):105–109, 1974.

[19] S. Jukna. *Extremal Combinatorics: With Applications in Computer Science.* Springer Publishing Company, Incorporated, 1st edition, 2010.

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

[21] H. W. Lenstra, A. K. Lenstra, L. Lovász, et al. Factoring polynomials with rational coeficients. *Mathematische Annalen*, 261:515–534, 1982.

[22] H. W. Lenstra Jr. Integer programming with a fixed number of variables. *Mathematics of operations research*, 8(4):538–548, 1983.

[23] G. Pataki, M. Tural, and E. B. Wong. Basis reduction and the complexity of branch-and-bound. In *Proceedings of the twenty-first annual ACM-SIAM symposium on discrete algorithms*, pages 1254–1261. SIAM, 2010.

[24] P. Pudlák. Lower bounds for resolution and cutting plane proofs and monotone computations. *Journal of Symbolic Logic*, pages 981–998, 1997.

[25] R. Vershynin. *High-dimensional probability: An introduction with applications in data science*, volume 47. Cambridge university press, 2018.

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

[27] Y. Yang, N. Boland, and M. Savelsbergh. Multivariable branching: A 0-1 knapsack problem case study. *INFORMS Journal on Computing*, 2021.