Complexity of branch-and-bound and cutting planes in mixed-integer optimization - II*

Amitabh Basu[1][0000−0002−1070−2626], Michele Conforti[2][0000−0002−6267−6941], Marco Di Summa[2][0000−0001−7223−0380], and Hongyi Jiang[1][0000−0001−9840−078X]

[1] Johns Hopkins University, Baltimore, USA, {basu.amitabh,hjiang32}@jhu.edu
[2] Università degli Studi Padova, Padova, Italy {conforti,disumma}@math.unipd.it

Abstract. We study the complexity of cutting planes and branching schemes from a theoretical point of view. We give some rigorous underpinnings to the empirically observed phenomenon that combining cutting planes and branching into a branch-and-cut framework can be orders of magnitude more efficient than employing these tools on their own. In particular, we give general conditions under which a cutting plane strategy and a branching scheme give a provably exponential advantage in efficiency when combined into branch-and-cut. The efficiency of these algorithms is evaluated using two concrete measures: number of iterations and sparsity of constraints used in the intermediate linear/convex programs. To the best of our knowledge, our results are the first mathematically rigorous demonstration of the superiority of branch-and-cut over pure cutting planes and pure branch-and-bound.

Keywords: Integer Programming · Cutting planes · Branching schemes · Proof complexity

1 Introduction

In this paper, we consider the following mixed-integer optimization problem:

$$\begin{array}{ll} \sup & \langle c, x\rangle \\ \text{s.t.} & x \quad \in C \cap \mathbb{Z}^n \times \mathbb{R}^d \end{array} \tag{1}$$

where C is a closed, convex set in $\mathbb{R}^{n+d}$.

State of the art algorithms for integer optimization are based on two ideas that are at the origin of mixed-integer programming and have been constantly refined: *cutting planes* and *branch-and-bound*. Decades of theoretical and experimental research into both these techniques is at the heart of the outstanding success of integer programming solvers. Nevertheless, we feel that there is lot of scope for widening and deepening our understanding of these tools. We have recently started building foundations for a rigorous, quantitative theory for analyzing the strengths and weaknesses of cutting planes and branching [3]. We continue this project in the current manuscript.

* Supported by ONR Grant N000141812096, NSF Grant CCF2006587, AFOSR Grant FA95502010341, and SID 2019 from University of Padova.

In particular, we provide a theoretical framework to explain an empirically observed phenomenon: algorithms that make a combined use of both cutting planes and branching techniques are more efficient (sometimes by orders of magnitude), compared to their stand alone use in algorithms. Not only is a theoretical understanding of this phenomenon lacking, a deeper understanding of the interaction of these methods is considered to be important by both practitioners and theoreticians in the mixed-integer optimization community. To quote an influential computational survey [34] "... it seems that a tighter coordination of the two most fundamental ingredients of the solvers, branching and cutting, can lead to strong improvements."

The main computational burden in any cutting plane or branch-and-bound or branch-and-cut algorithm is the solution of the intermediate convex relaxations. Thus, there are two important aspects to deciding how efficient such an algorithm is: 1) How many linear programs (LPs) or convex optimization problems are solved? 2) How computationally challenging are these convex problems? The first aspect has been widely studied using the concepts of proof size and rank; see [5, 9–11, 15, 18–20, 24, 44] for a small sample of previous work. Formalizing the second aspect is somewhat tricky and we will focus on a very specific aspect: the *sparsity* of the constraints describing the linear program. The collective wisdom of the optimization community says that sparsity of constraints is a highly important aspect in the efficiency of linear programming [4, 25, 43, 47]. Additionally, most mixed-integer optimization solvers use sparsity as a criterion for cutting plane selection; see [21–23] for an innovative line of research. Compared to cutting planes, sparsity considerations have not been as prominent in the choice of branching schemes. This is primarily because for variable disjunctions sparsity is not an issue, and there is relatively less work on more general branching schemes; see [1, 16, 17, 32, 35–39]. In our analysis, we are careful about the sparsity of the disjunctions as well – see Definition 3 below.

1.1 Framework for mathematical analysis

We now present the formal details of our approach. A *cutting plane* for the feasible region of (1) is a halfspace $H = \{x \in \mathbb{R}^{n+d} : \langle a, x\rangle \leq \delta\}$ such that $C \cap (\mathbb{Z}^n \times \mathbb{R}^d) \subseteq H$. The most useful cutting planes are those that are not valid for C, i.e., $C \not\subseteq H$. There are several procedures used in practice for generating cutting planes, all of which can be formalized by the general notion of a *cutting plane paradigm*. A cutting plane paradigm is a function $\mathcal{CP}$ that takes as input any closed, convex set C and outputs a (possibly infinite) family $\mathcal{CP}(C)$ of cutting planes valid for $C \cap (\mathbb{Z}^n \times \mathbb{R}^d)$. Two well-studied examples of cutting plane paradigms are the *Chvátal-Gomory cutting plane paradigm* [45, Chapter 23] and the *split cut paradigm* [13, Chapter 5]. We will assume that all cutting planes are rational in this paper.

State-of-the-art solvers embed cutting planes into a systematic enumeration scheme called *branch-and-bound*. The central notion is that of a *disjunction*, which is a union of polyhedra $D = Q_1 \cup \ldots \cup Q_k$ such that $\mathbb{Z}^n \times \mathbb{R}^d \subseteq D$, i.e., the polyhedra together cover all of $\mathbb{Z}^n \times \mathbb{R}^d$. One typically uses a (possibly infinite) family of disjunctions for potential deployment in algorithms. A well-known example is the family of *split disjunctions* that are of the form $D_{\pi,\pi_0} := \{x \in \mathbb{R}^{n+d} : \langle \pi, x\rangle \leq \pi_0\} \cup \{x \in \mathbb{R}^{n+d} : \langle \pi, x\rangle \geq \pi_0 + 1\}$, where $\pi \in \mathbb{Z}^n \times \{0\}^d$ and $\pi_0 \in \mathbb{Z}$. When the first n coordinates of π correspond to a standard unit vector, we get *variable disjunctions*, i.e., disjunctions of the form $\{x : x_i \leq \pi_0\} \cup \{x : x_i \geq \pi_0 + 1\}$, for $i = 1, \ldots, n$.

A family of disjunctions $\mathcal{D}$ can also form the basis of a cutting plane paradigm. Given any disjunction D, any halfspace H such that $C \cap D \subseteq H$ is a cutting plane, since $C \cap (\mathbb{Z}^n \times \mathbb{R}^d) \subseteq C \cap D$ by definition of a disjunction. The corresponding cutting plane paradigm $\mathcal{CP}(C)$, called *disjunctive cuts based on* $\mathcal{D}$, is the family of all such cutting planes derived from disjunctions in $\mathcal{D}$.

In the following we assume that all convex optimization problems that need to be solved have an optimal solution or are infeasible.

Definition 1. *A* branch-and-cut algorithm based on a family $\mathcal{D}$ of disjunctions and a cutting plane paradigm $\mathcal{CP}$ *maintains a list* $\mathcal{L}$ *of* convex subsets *of the initial set C which are guaranteed to contain the optimal point, and a lower bound LB that stores the objective value of the best feasible solution found so far (with $LB = \infty$ if no feasible solution has been found). At every iteration, the algorithm selects one of these subsets $N \in \mathcal{L}$ and solves the convex optimization problem* $\sup\{\langle c, x\rangle : x \in N\}$ *to obtain x^N. If the objective value is less than or equal to LB, then this set N is discarded from the list $\mathcal{L}$. Else, if x^N satisfies the integrality constraints, LB is updated with the value of x^N and N is discarded from the list. Otherwise, the algorithm makes a decision whether to* branch *or to* cut*. In the former case, a disjunction $D = (Q_1 \cup \ldots \cup Q_k) \in \mathcal{D}$ is chosen such that $x^N \notin D$ and the list is updated $\mathcal{L} := \mathcal{L} \setminus \{N\} \cup \{Q_1 \cap N, \ldots, Q_k \cap N\}$. If the decision is to cut, then the algorithm selects a cutting plane $H \in \mathcal{CP}(P)$ such that $x^N \notin H$, and updates the relaxation N by adding the cut H, i.e., updates $\mathcal{L} := \mathcal{L} \setminus \{N\} \cup \{N \cap H\}$.*

Motivated by the above, we will refer to a family $\mathcal{D}$ of disjunctions also as a *branching scheme*. In a branch-and-cut algorithm, if one always chooses to add a cutting plane and never uses a disjunction to branch, then it is said to be a *(pure) cutting plane algorithm* and if one does not use any cutting planes ever, then it is called a *(pure) branch-and-bound algorithm*. We note here that in practice, when a decision to cut is made, several cutting planes are usually added as opposed to just one single cutting plane like in Definition 1. In our mathematical framework, allowing only a single cut makes for a seamless generalization from pure cutting plane algorithms, and also makes quantitative analysis easier.

Definition 2. *The execution of any branch-and-cut algorithm on a mixed-integer optimization instance can be represented by a tree. Every convex relaxation N processed by the algorithm is denoted by a node in the tree. If the optimal value for N is not better than the current lower bound, or is integral, N is a leaf. Otherwise, in the case of a branching, its children are $Q_1 \cap N, \ldots, Q_k \cap N$, and in the case of a cutting plane, there is a single child representing $N \cap H$ (we use the same notation as in Definition 1). This tree is called the* branch-and-cut tree (branch-and-bound tree*, if no cutting planes are used). If no branching is done, this tree (which is really a path) is called a* cutting plane proof*. The* size *of the tree or proof is the total number of nodes.*

Proof versus algorithm. Although we use the word "algorithm" in Definition 1, it is technically a *non-deterministic* algorithm, or equivalently, a proof schema or proof system for optimality [2] (leaving aside the question of finite termination for now). This is because no indication is given on how the important decisions are made: Which set N to process from $\mathcal{L}$? Branch or cut? Which disjunction or cutting plane to use? Nevertheless, the proof system is very useful for obtaining information theoretic lower bounds on the efficiency of any deterministic branch-and-cut algorithm. Moreover, one can prove the validity of any upper bound on the objective, i.e., the validity of $\langle c, x\rangle \leq \gamma$ by exhibiting a branch-and-cut tree where this inequality is valid for all the leaves. If γ is the optimal value, this is a proof of optimality, but one may often be interested in the branch-and-cut/branch-and-bound/cutting plane proof complexity of other valid inequalities as well. The connections between integer programming and proof complexity has a long history; see [6, 7, 12, 14, 26–28, 30, 33, 40–42], to cite a few. Our results can be interpreted in the language of proof complexity as well.

Recall that we quantify the complexity of any branch-and-bound/cutting plane/branch-and-cut algorithm using two aspects: the number of LP relaxations processed and the sparsity of the constraints defining the LPs. The number of LP relaxations processed is given precisely by the number of nodes in the corresponding tree (Definition 2). Sparsity is formalized in the following definitions.

Definition 3. *Let $1 \leq s \leq n + d$ be a natural number that we call the* sparsity parameter*. Then the pair $(\mathcal{CP}, s)$ will denote the restriction of the paradigm $\mathcal{CP}$ that only reports the sub-family of cutting planes that can be represented by inequalities with at most s non-zero coefficients; the notation $(\mathcal{CP}, s)(C)$ will be used to denote this sub-family for any particular convex set C. Similarly, $(\mathcal{D}, s)$ will denote the sub-family of the family of disjunctions $\mathcal{D}$ such that each polyhedron in the disjunction has an inequality description where every inequality has at most s non-zero coefficients.*

1.2 Our Results

Sparsity versus size. Our first set of results considers the trade-off between the sparsity parameter s and the number of LPs processed, i.e., the size of the tree. There are several avenues to explore in this direction. For example, one could compare pure branch-and-bound algorithms based on $(\mathcal{D}, s_1)$ and $(\mathcal{D}, s_2)$, i.e., fix a particular disjunction family $\mathcal{D}$ and consider the effect of sparsity on the branch-and-bound tree sizes. One could also look at two different families of disjunctions $\mathcal{D}_1$ and $\mathcal{D}_2$ and look at their relative tree sizes as one turns the knob on the sparsity parameter. Similar questions could be asked about cutting plane paradigms $(\mathcal{CP}_1, s_1)$ and $(\mathcal{CP}_2, s_2)$ for interesting paradigms $\mathcal{CP}_1, \mathcal{CP}_2$.

Even more interestingly, one could compare pure branch-and-bound and pure cutting plane algorithms against each other.

We first focus on pure branch-and-bound algorithms based on the family $\mathcal{S}$ of split disjunctions. A very well-known example of pure integer instances (i.e., $d = 0$) due to Jeroslow [31] shows that if the sparsity of the splits used is restricted to be 1, i.e., one uses only variable disjunctions, then the branch-and-bound algorithm will generate an exponential (in the dimension n) sized tree. On the other hand, if one allows fully dense splits, i.e., sparsity is n, then there is a tree with just 3 nodes (one root, and two leaves) that solves the problem. We ask what happens in Jeroslow's example if one uses split disjunctions with sparsity $s > 1$. Our first result shows that unless the sparsity parameter $s = \Omega(n)$, one cannot get constant size trees, and if the sparsity parameter $s = O(1)$, then the tree is of exponential size.

Theorem 1. *Let H be the halfspace defined by inequality $2\sum_{i=1}^n x_i \le n$, where n is an odd number. Consider the instances of (1) with $d = 0$, the objective $\sum_{i=1}^n x_i$ and $C = H \cap [0,1]^n$. The optimum is $\lfloor \frac{n}{2} \rfloor$, and any branch-and-bound proof with sparsity $s \le \lfloor \frac{n}{2} \rfloor$ that certifies $\sum_{i=1}^n x_i \le \lfloor \frac{n}{2} \rfloor$ has size at least $\Omega(2^{\frac{n}{2s}})$.*

The bounds in Theorem 1 give a constant lower bound when $s = \Omega(n)$. We establish another lower bound which does better in this regime.

Theorem 2. *Let H be the halfspace defined by inequality $2\sum_{i=1}^n x_i \le n$, where n is an odd number. Consider the instances of (1) with $d = 0$, the objective $\sum_{i=1}^n x_i$ and $C = H \cap [0,1]^n$. The optimum is $\lfloor \frac{n}{2} \rfloor$, and any branch-and-bound proof with sparsity $s \le \lfloor \frac{n}{2} \rfloor$ that certifies $\sum_{i=1}^n x_i \le \lfloor \frac{n}{2} \rfloor$ has size at least $\Omega\left(\sqrt{\frac{n(n-s)}{s}}\right)$.*

Next we consider the relative strength of cutting planes and branch-and-bound. Our previous work has studied conditions under which one method can dominate the other, depending on which cutting plane paradigm and branching scheme one chooses [3]. For this paper, the following result from [3] is relevant: for every convex 0/1 pure integer instance, any branch-and-bound proof based on variable disjunctions can be "simulated" by a lift-and-project cutting plane proof without increasing the size of the proof (versions of this result for *linear* 0/1 programming were known earlier; see [18, 19]). Moreover, in [3] we constructed a family of stable set instances where lift-and-project cuts gave exponentially shorter proofs than branch-and-bound. This is interesting because lift-and-project cuts are disjunctive cuts based on the same family of variable disjunctions, so it is not a priori clear that they have an advantage. These results were obtained with no regard for sparsity. We now show that once we also track the sparsity parameter, this advantage can disappear.

Theorem 3. *Let H be the halfspace defined by inequality $2\sum_{i=1}^n x_i \le n$, where n is an odd number. Consider the intances of (1) with $d = 0$, the objective $\sum_{i=1}^{\lceil \frac{n}{2} \rceil} x_i$ and $C = H$. The optimum is $\lfloor \frac{n}{2} \rfloor$, and there is a branch-and-bound algorithm based on variable disjunctions, i.e., the family of split disjunctions with sparsity*

1, that certifies $\sum_{i=1}^{\lceil \frac{n}{2} \rceil} x_i \leq \lfloor \frac{n}{2} \rfloor$ *in* $O(n)$ *steps. However, any cutting plane for* C *with sparsity* $s \leq \lfloor \frac{n}{2} \rfloor$ *is trivial, i.e., valid for* $[0,1]^n$*, no matter what cutting plane paradigm is used to derive it.*

Superiority of branch-and-cut. We next consider the question of when combining branching and cutting planes is *provably* advantageous. For this question, we leave aside the complications arising due to sparsity considerations and focus only on the size of proofs. The following discussion and results can be extended to handle the issue of sparsity as well, but we leave it out of this extended abstract.

Given a cutting plane paradigm $\mathcal{CP}$, and a branching scheme $\mathcal{D}$, are there families of instances where branch-and-cut based on $\mathcal{CP}$ and $\mathcal{D}$ does provably better than pure cutting planes based on $\mathcal{CP}$ alone and pure branch-and-bound based on $\mathcal{D}$ alone? If a cutting plane paradigm $\mathcal{CP}$ and a branching scheme $\mathcal{D}$ are such that either for every instance, $\mathcal{CP}$ gives cutting plane proofs of size at most a polynomial factor larger than the shortest branch-and-bound proofs with $\mathcal{D}$, or vice versa, for every instance $\mathcal{D}$ gives proofs of size at most polynomially larger than the shortest cutting plane proofs based on $\mathcal{CP}$, then combining them into branch-and-cut is likely to give no substantial improvement since one method can always do the job of the other, up to polynomial factors. As mentioned above, prior work [3] had shown that disjunctive cuts based on variable disjunctions (with no restriction on sparsity) dominate branch-and-bound based on variable disjunctions for 0/1 instances, and as a consequence branch-and-cut based on these paradigms is dominated by pure cutting planes. In this paper, we show that the situation completely reverses if one considers a broader family of disjunctions.

Theorem 4. *Let* $C \subseteq \mathbb{R}^n$ *be a closed, convex set. Let* $k \in \mathbb{N}$ *be a fixed natural number and let* $\mathcal{D}$ *be any family of disjunctions that contains all split disjunctions, such that all disjunctions in* $\mathcal{D}$ *have at most* k *terms in the disjunction. If a valid inequality* $\langle c, x \rangle \leq \delta$ *for* $C \cap \mathbb{Z}^n$ *has a cutting plane proof of size* L *using disjunctive cuts based on* $\mathcal{D}$*, then there exists a branch-and-bound proof of size at most* $(k+1)L$ *based on* $\mathcal{D}$*. Moreover, there is a family of instances where branch-and-bound based on split disjunctions solves the problem in* $O(1)$ *time whereas there is a polynomial lower bound on split cut proofs.*

The above discussion and theorem motivate the following definition which formalizes the situation where no method dominates the other. To make things precise, we assume that there is a well-defined way to assign a concrete *size* to any instance of (1); see [29] for a discussion on how to make this precise. Additionally, when we speak of an instance, we allow the possibility of proving the validity of any inequality valid for $C \cap (\mathbb{Z}^n \times \mathbb{R}^d)$, not necessarily related to an upper bound on the objective value. Thus, an instance is a pair $(C, (c, \gamma))$ such that $\langle c, x \rangle \leq \gamma$ for all $x \in C \cap (\mathbb{Z}^n \times \mathbb{R}^d)$.

Definition 4. *A cutting plane paradigm* $\mathcal{CP}$ *and a branching scheme* $\mathcal{D}$ *are* complementary *if there is a family of instances where* $\mathcal{CP}$ *gives polynomial (in the size of the instances) size proofs and the shortest branch-and-bound proof based* *on $\mathcal{D}$ is exponential (in the size of the instances), and there is another family of instances where $\mathcal{D}$ gives polynomial size proofs while $\mathcal{CP}$ gives exponential size proofs.*

We wish to formalize the intuition that branch-and-cut is expected to be exponentially better than branch-and-bound or cutting planes alone for complementary pairs of branching schemes and cutting plane paradigms. But we need to make some mild assumptions about the branching schemes and cutting plane paradigms. *All known branching schemes and cutting plane methods from the literature satisfy these conditions.*

Definition 5. *A branching scheme is said to be* regular *if no disjunction involves a continuous variable, i.e., each polyhedron in the disjunction is described using inequalities that involve only the integer constrained variables.*

A branching scheme $\mathcal{D}$ is said to be embedding closed *if disjunctions from higher dimensions can be applied to lower dimensions. More formally, let n_1, n_2, d_1, $d_2 \in \mathbb{N}$. If $D \in \mathcal{D}$ is a disjunction in $\mathbb{R}^{n_1} \times \mathbb{R}^{d_1} \times \mathbb{R}^{n_2} \times \mathbb{R}^{d_2}$ with respect to $\mathbb{Z}^{n_1} \times \mathbb{R}^{d_1} \times \mathbb{Z}^{n_2} \times \mathbb{R}^{d_2}$, then the disjunction $D \cap (\mathbb{R}^{n_1} \times \mathbb{R}^{d_1} \times \{0\}^{n_2} \times \{0\}^{d_2})$, interpreted as a set in $\mathbb{R}^{n_1} \times \mathbb{R}^{d_1}$, is also in $\mathcal{D}$ for the space $\mathbb{R}^{n_1} \times \mathbb{R}^{d_1}$ with respect to $\mathbb{Z}^{n_1} \times \mathbb{R}^{d_1}$ (note that $D \cap (\mathbb{R}^{n_1} \times \mathbb{R}^{d_1} \times \{0\}^{n_2} \times \{0\}^{d_2})$, interpreted as a set in $\mathbb{R}^{n_1} \times \mathbb{R}^{d_1}$, is certainly a disjunction with respect to $\mathbb{Z}^{n_1} \times \mathbb{R}^{d_1}$; we want $\mathcal{D}$ to be closed with respect to such restrictions).*

A cutting plane paradigm $\mathcal{CP}$ is said to be regular *if it has the following property, which says that adding "dummy variables" to the formulation of the instance should not change the power of the paradigm. Formally, let $C \subseteq \mathbb{R}^n \times \mathbb{R}^d$ be any closed, convex set and let $C' = \{(x,t) \in \mathbb{R}^n \times \mathbb{R}^d \times \mathbb{R} : x \in C,\ \ t = \langle f, x\rangle\}$ for some $f \in \mathbb{R}^n$. Then if a cutting plane $\langle a, x\rangle \leq b$ is derived by $\mathcal{CP}$ applied to C, i.e., this inequality is in $\mathcal{CP}(C)$, then it should also be in $\mathcal{CP}(C')$, and conversely, if $\langle a, x\rangle + \mu t \leq b$ is in $\mathcal{CP}(C')$, then the equivalent inequality $\langle a + \mu f, x\rangle \leq b$ should be in $\mathcal{CP}(C)$.*

A cutting plane paradigm $\mathcal{CP}$ is said to be embedding closed *if disjunctions from higher dimensions can be applied to lower dimensions. More formally, let $n_1, n_2, d_1, d_2 \in \mathbb{N}$. Let $C \subseteq \mathbb{R}^{n_1} \times \mathbb{R}^{d_1}$ be any closed, convex set. If the inequality $\langle c_1, x_1\rangle + \langle a_1, y_1\rangle + \langle c_2, x_2\rangle + \langle a_2, y_2\rangle \leq \gamma$ is a cutting plane for $C \times \{0\}^{n_2} \times \{0\}^{d_2}$ with respect to $\mathbb{Z}^{n_1} \times \mathbb{R}^{d_1} \times \mathbb{Z}^{n_2} \times \mathbb{R}^{d_2}$ that can be derived by applying $\mathcal{CP}$ to $C \times \{0\}^{n_2} \times \{0\}^{d_2}$, then the cutting plane $\langle c_1, x_1\rangle + \langle a_1, y_1\rangle \leq \gamma$ that is valid for $C \cap (\mathbb{Z}^{n_1} \times \mathbb{R}^{d_1})$ should also belong to $\mathcal{CP}(C)$.*

A cutting plane paradigm $\mathcal{CP}$ is said to be inclusion closed, *if for any two closed convex sets $C \subseteq C'$, we have $\mathcal{CP}(C') \subseteq \mathcal{CP}(C)$. In other words, any cutting plane derived for C' can also be derived for a subset C.*

Theorem 5. *Let $\mathcal{D}$ be a regular, embedding closed branching scheme and let $\mathcal{CP}$ be a regular, embedding closed, and inclusion closed cutting plane paradigm such that $\mathcal{D}$ includes all variable disjunctions and $\mathcal{CP}$ and $\mathcal{D}$ form a complementary pair. Then there exists a family of instances of* (1) *which have polynomial size branch-and-cut proofs, whereas any branch-and-bound proof based on $\mathcal{D}$ and any cutting plane proof based on $\mathcal{CP}$ is of exponential size.*

Example 1. As a concrete example of a complementary pair that satisfies the other conditions of Theorem 5, consider $\mathcal{CP}$ to be the Chvátal-Gomory paradigm and $\mathcal{D}$ to be the family of variable disjunctions. From their definitions, they are both regular and $\mathcal{D}$ is embedding closed. The Chvátal-Gomory paradigm is also embedding closed and inclusion closed. For the Jeroslow instances from Theorem 1, the single Chvátal-Gomory cut $\sum_{i=1}^{n} x_i \leq \lfloor \frac{n}{2} \rfloor$ proves optimality, whereas variable disjunctions produce a tree of size $2^{\lfloor \frac{n}{2} \rfloor}$. On the other hand, consider the set T, where $T = \mathrm{conv}\{(0,0),(1,0),(\frac{1}{2},h)\}$ and the valid inequality $x_2 \leq 0$ for $T \cap \mathbb{Z}^2$. Any Chvátal-Gomory paradigm based proof has size exponential in the size of the input, i.e., every proof has length at least $\Omega(h)$ [45]. On the other hand, a single disjunction on the variable x_1 solves the problem.

In [3], we also studied examples of disjunction families $\mathcal{D}$ such that disjunctive cuts based on $\mathcal{D}$ are complementary to branching schemes based on $\mathcal{D}$.

Example 1 shows that the classical Chvátal-Gomory cuts and variable branching are complementary and thus give rise to a superior branch-and-cut routine when combined by Theorem 5. As discussed above, for 0/1 problems, lift-and-project cuts and variable branching do *not* form a complementary pair, and neither do split cuts and split disjunctions by Theorem 4. It would be nice to establish the converse of Theorem 5: if there is a family where branch-and-cut is exponentially superior, then the cutting plane paradigm and branching scheme are complementary. In Theorem 6 below, we prove a partial converse along these lines in the pure integer case; it remains an open question if our definition of complementarity is an exact characterization of when branch-and-cut is superior.

Theorem 6. *Let $\mathcal{D}$ be a branching scheme that includes all split disjunctions and let $\mathcal{CP}$ be any cutting plane paradigm. Suppose that for every pure integer instance and any cutting plane proof based on $\mathcal{CP}$ for this instance, there is a branch-and-bound proof based on $\mathcal{D}$ of size at most a polynomial factor (in the size of the instance) larger. Then for any branch-and-cut proof based on $\mathcal{D}$ and $\mathcal{CP}$ for a pure integer instance, there exists a pure branch-and-bound proof based on $\mathcal{D}$ that has size at most polynomially larger than the branch-and-cut proof.*

The high level message that we extract from our results is the formalization of the following simple intuition. For branch-and-cut to be superior to pure cutting planes or pure branch-and-bound, one needs the cutting planes and branching scheme to do "sufficiently different" things. For example, if they are both based on the same family of disjunctions (such as lift-and-project with variable branching, or the setting of Theorem 4), then we may not get any improvements with branch-and-cut. The definition of a complementary pair attempts to make the notion of "sufficiently different" formal and Theorem 5 derives the concrete superior performance of branch-and-cut from this formalization.

2 Proofs

We present the proofs of Theorems 1, 2, 5 and 6 in the subsections below. The proofs of Theorems 3 and 4 are excluded from this extended abstract.

2.1 Proof of Theorem 1

Definition 6. *Consider the instances in Theorem 1, and the branch-and-bound tree T produced by split disjunctions to solve it. Assume node N of T contains at least one integer point in $\{0,1\}^n$, and $D_1, D_2, \ldots, D_r$ are the split disjunctions used to derive N from the root of T. For $1 \leq j \leq r$, D_j is a* true split disjunction *of N if both of the two halfspaces of D_j have a nonempty intersection with the integer hull of the corresponding parent node, i.e. the parent node's integer hull is split into two nonempty parts by D_j. Otherwise, it is called a* false split disjunction *of N. We define the* generation variable set *of N as the index set $I \subseteq \{1, 2, \ldots, n\}$ such that it consists of all the indices of the variables involved in the true split disjunctions of N. The generation set of the root node is empty.*

The proof of the following lemma is excluded from this extended abstract.

Lemma 1. *Consider the instances in Theorem 1, and the branch-and-bound tree T produced by split disjunctions with sparsity parameter $s < \lfloor \frac{n}{2} \rfloor$ to solve it. For any node N of T with at least one feasible integer point $v = (v_1, v_2, \ldots, v_n) \in \{0,1\}^n$, let P, P_I and I denote the relaxation, the integer hull and the generation variable set corresponding to N. Define $V := \{(x_1, x_2, \ldots, x_n) \in \{0,1\}^n : x_i = v_i \text{ for } i \in I, \sum_{j=1}^n x_i = \lfloor \frac{n}{2} \rfloor\}$.*

If $|I| \leq \lfloor \frac{n}{2} \rfloor - s$, then we have:

(i) $V \neq \emptyset$ and $V \subseteq P_I \cap \{0,1\}^n$;
(ii) the objective LP value of N is $\frac{n}{2}$.

Proof (Proof of Theorem 1). For a node N of the branch-and-bound tree containing at least one integer point, if it is derived by exactly m true split disjunctions, then we say it is a node of generation m. By Lemma 1, if $m \leq \frac{1}{s} \lfloor \frac{n}{2} \rfloor - 1$, then a node N of generation m has LP objective value $\frac{n}{2}$, and in the subtree rooted at N there must exist at least two descendants from generation $m+1$, since the leaf nodes must have LP values less than or equal to $\lfloor \frac{n}{2} \rfloor$. Therefore, there are at least 2^m nodes of generation m when $m \leq \frac{1}{s} \lfloor \frac{n}{2} \rfloor - 1$. This finishes the proof. □

2.2 Proof of Theorem 2

The following lemma follows from an application of Sperner's theorem [46].

Lemma 2. *Let $w_1, \ldots, w_k \in \mathbb{Z} \setminus \{0\}$ and $W \in \mathbb{Z}$. Then the number of 0/1 solutions to $\sum_{j=1}^k w_j x_j = W$ is at most $\binom{k}{\lfloor k/2 \rfloor}$.*

Proof (Proof of Theorem 2). We consider the instance from Theorem 2. For any split disjunction $D := \{x : \langle a, x \rangle \leq b\} \cup \{x : \langle a, x \rangle \geq b + 1\}$, we define $V(D)$ to be the set of all the optimal LP vertices (of the original polytope) that lie strictly in the split set corresponding to D. Let the support of a be given by $T \subseteq \{1, \ldots, n\}$ with $t := |T| \leq s \leq \lfloor n/2 \rfloor$. Since $a \in \mathbb{Z}^n$ and $b \in \mathbb{Z}$, $V(D)$ is precisely the subset of the optimal LP vertices $\hat{x}$ such that $\langle a, \hat{x} \rangle = b + \frac{1}{2}$. Fix some $\ell \in T$ and consider those optimal LP vertices $\hat{x} \in V(D)$ where $\hat{x}_\ell = \frac{1}{2}$. This means that $\sum_{j \in T\setminus\{\ell\}} a_j \hat{x}_j = b + \frac{1}{2} - \frac{a_\ell}{2}$. Let r_i be the number of 0/1 solutions to $\sum_{j \in T\setminus\{\ell\}} a_j \hat{x}_j = b + \frac{1}{2} - \frac{a_\ell}{2}$ with exactly i coordinates set to 1. Then the number of vertices from $V(D)$ with the ℓ-th coordinate equal to $\frac{1}{2}$ is

$$\sum_{i=0}^{t-1} r_i \binom{n-t}{\lfloor n/2 \rfloor - i} \leq \left(\sum_{i=0}^{t-1} r_i\right) \binom{n-t}{\lfloor n/2 \rfloor - \lfloor t/2 \rfloor}.$$

since $\binom{n-t}{\lfloor n/2 \rfloor - i} \leq \binom{n-t}{\lfloor n/2 \rfloor - \lfloor t/2 \rfloor}$ for all $i \in \{0, \ldots, t-1\}$. Using Lemma 2, $\sum_{i=0}^{t-1} r_i \leq \binom{t-1}{\lfloor t/2 \rfloor}$ and we obtain the upper bound $\binom{t-1}{\lfloor t/2 \rfloor}\binom{n-t}{\lfloor n/2 \rfloor - \lfloor t/2 \rfloor}$ on the number of vertices from $V(D)$ with the ℓ-th coordinate equal to $\frac{1}{2}$. Therefore, $|V(D)| \leq t\binom{t-1}{\lfloor t/2 \rfloor}\binom{n-t}{\lfloor n/2 \rfloor - \lfloor t/2 \rfloor} =: p(t)$. Since n is odd, we have

$$p(t) = \begin{cases} \dfrac{t!(n-t)!}{(t/2)!(t/2-1)!((n-t-1)/2)!((n-t+1)/2)!} & \text{if } t \text{ is even,} \\ \dfrac{t!(n-t)!}{((t-1)/2)!((t-1)/2)!((n-t)/2)!((n-t)/2)!} & \text{if } t \text{ is odd.} \end{cases}$$

A direct calculation then shows that

$$\frac{p(t+1)}{p(t)} = \begin{cases} \dfrac{(t+1)(n-t+1)}{t(n-t)} & \text{if } t \text{ is even,} \\ 1 & \text{if } t \text{ is odd.} \end{cases}$$

Let h be the largest even number not exceeding s. Since $p(1) = \binom{n-1}{\lfloor n/2 \rfloor}$, we obtain, for every $t \in \{1, \ldots, s\}$,

$$p(t) \leq p(s) = \binom{n-1}{\lfloor n/2 \rfloor} \prod_{\substack{1 \leq q \leq s \\ q \text{ even}}} \frac{q+1}{q} \cdot \frac{n-q+1}{n-q} = \binom{n-1}{\lfloor n/2 \rfloor} \cdot \frac{(h+1)!!}{h!!} \cdot \frac{(n-1)!!}{(n-2)!!} \cdot \frac{(n-h-2)!!}{(n-h-1)!!}.$$

Using the fact that, for every even positive integer ℓ,

$$\sqrt{\frac{\pi \ell}{2}} < \frac{\ell!!}{(\ell-1)!!} < \sqrt{\frac{\pi(\ell+1)}{2}}$$

(see, e.g., [8, 48]), we have (for $h \geq 1$, i.e., $s \geq 2$)

$$\begin{aligned} p(t) &\leq \binom{n-1}{\lfloor n/2 \rfloor} \cdot \frac{(h+1)(h-1)!!}{h!!} \cdot \frac{(n-1)!!}{(n-2)!!} \cdot \frac{(n-h-2)!!}{(n-h-1)!!} \\ &\leq \binom{n-1}{\lfloor n/2 \rfloor} (h+1) \sqrt{\frac{2}{\pi h} \cdot \frac{\pi n}{2} \cdot \frac{2}{\pi(n-h-1)}} \\ &= \binom{n-1}{\lfloor n/2 \rfloor} \sqrt{\frac{2n(h+1)^2}{\pi h(n-h-1)}} \\ &= \binom{n-1}{\lfloor n/2 \rfloor} O\left(\sqrt{\frac{ns}{n-s}}\right). \end{aligned}$$

Thus this is an upper bound on $|V(D)|$. Since the total number of optimal LP vertices of the instance is $n\binom{n-1}{\lfloor n/2\rfloor}$, we obtain the following lower bound of on the size of a branch-and-bound proof: $\frac{n\binom{n-1}{\lfloor n/2\rfloor}}{|V(D)|} = \Omega\left(\sqrt{\frac{n(n-s)}{s}}\right)$. $\square$

2.3 Proofs of Theorems 5 and 6

Lemmas 3– 5 below are straightforward consequences of the definitions, and the proofs are omitted from this extended abstract.

Lemma 3. *Let $C \subseteq C'$ be two closed, convex sets. Let $\mathcal{D}$ be any branching scheme and let $\mathcal{CP}$ be an inclusion closed cutting plane paradigm. If there is a branch-and-bound proof with respect to C' based on $\mathcal{D}$ for the validity of an inequality $\langle c, x\rangle \le \gamma$, then there is a branch-and-bound proof with respect to C based on $\mathcal{D}$ for the validity of $\langle c, x\rangle \le \gamma$ of the same size. The same holds for cutting plane proofs based on $\mathcal{CP}$.*

Lemma 4. *Let $\mathcal{D}$ and $\mathcal{CP}$ be both embedding closed and let $C \subseteq \mathbb{R}^{n_1} \times \mathbb{R}^{d_1}$ be a closed, convex set. Let $\langle c, x\rangle \le \gamma$ be a valid inequality for $C \cap (\mathbb{Z}^{n_1} \times \mathbb{R}^{d_1})$. If there is a branch-and-bound proof with respect to $C \times \{0\}^{n_2} \times \{0\}^{d_2}$ based on $\mathcal{D}$ for the validity of $\langle c, x\rangle \le \gamma$ interpreted as a valid inequality in $\mathbb{R}^{n_1} \times \mathbb{R}^{d_1} \times \mathbb{R}^{n_2} \times \mathbb{R}^{d_2}$ for $(C \times \{0\}^{n_2} \times \{0\}^{d_2}) \cap (\mathbb{Z}^{n_1} \times \mathbb{R}^{d_1} \times \mathbb{Z}^{n_2} \times \mathbb{R}^{d_2})$, then there is a branch-and-bound proof with respect to C based on $\mathcal{D}$ for the validity of $\langle c, x\rangle \le \gamma$ of the same size. The same holds for cutting plane proofs based on $\mathcal{CP}$.*

Lemma 5. *Let $C \subseteq \mathbb{R}^{n+d}$ be a polytope and let $\langle c, x\rangle \le \gamma$ be a valid inequality for $C \cap (\mathbb{Z}^n \times \mathbb{R}^d)$. Let $X := \{(x, t) \in \mathbb{R}^{n+d} \times \mathbb{R} : x \in C,\ \ t = \langle c, x\rangle\}$. Then, for any regular branching scheme $\mathcal{D}$ or a regular cutting plane paradigm $\mathcal{CP}$, any proof of validity of $\langle c, x\rangle \le \gamma$ with respect to $C \cap (\mathbb{Z}^n \times \mathbb{R}^d)$ can be changed into a proof of validity of $t \le \gamma$ with respect to $X \cap (\mathbb{Z}^n \times \mathbb{R}^d \times \mathbb{R})$ with no change in length, and vice versa.*

Proof (Proof of Theorem 5). Let $\{P_k \subseteq \mathbb{R}^{n_k} \times \mathbb{R}^{d_k} : k \in \mathbb{N}\}$ be a family of closed, convex sets, and $\{(c_k, \gamma_k) \in \mathbb{R}^{n_k} \times \mathbb{R}^{d_k} \times \mathbb{R} : k \in \mathbb{N}\}$ be a family of tuples such that $\langle c_k, x\rangle \le \gamma_k$ is valid for $P_k \cap (\mathbb{Z}^{n_k} \times \mathbb{R}^{d_k})$, and $\mathcal{CP}$ has polynomial size proofs for this family of instances, whereas $\mathcal{D}$ has exponential size proofs. Similarly, let $\{P'_k \subseteq \mathbb{R}^{n'_k} \times \mathbb{R}^{d'_k} : k \in \mathbb{N}\}$ be a family of closed, convex sets, and $\{(c'_k, \gamma'_k) \in \mathbb{R}^{n'_k} \times \mathbb{R}^{d'_k} \times \mathbb{R} : k \in \mathbb{N}\}$ be a family of tuples such that $\langle c'_k, x\rangle \le \gamma'_k$ is valid for $P'_k \cap (\mathbb{Z}^{n'_k} \times \mathbb{R}^{d'_k})$, and $\mathcal{D}$ has polynomial size proofs for this family of instances, whereas $\mathcal{CP}$ has exponential size proofs.

We first embed P_k and P'_k into a common ambient space for each $k \in \mathbb{N}$. This is done by defining $\bar{n}_k = \max\{n_k, n'_k\}$, $\bar{d}_k = \max\{d_k, d'_k\}$, and embedding both P_k and P'_k into the space $\mathbb{R}^{\bar{n}_k} \times \mathbb{R}^{\bar{d}_k}$ by defining $Q_k := P_k \times \{0\}^{\bar{n}_k - n_k} \times \{0\}^{\bar{d}_k - d_k}$ and $Q'_k := P'_k \times \{0\}^{\bar{n}_k - n'_k} \times \{0\}^{\bar{d}_k - d'_k}$. By Lemma 4, $\mathcal{D}$ has an exponential lower bound on sizes of proofs for the inequality $\langle c_k, x\rangle \le \gamma_k$, interpreted as an inequality in $\mathbb{R}^{\bar{n}_k} \times \mathbb{R}^{\bar{d}_k}$, valid for $Q_k \cap (\mathbb{Z}^{\bar{n}_k} \times \mathbb{R}^{\bar{d}_k})$. By Lemma 4, $\mathcal{CP}$ has an exponential lower bound on sizes of proofs for the inequality $\langle c'_k, x\rangle \leq \gamma'_k$, interpreted as an inequality in $\mathbb{R}^{\bar{n}_k} \times \mathbb{R}^{\bar{d}_k}$, valid for $Q'_k \cap (\mathbb{Z}^{\bar{n}_k} \times \mathbb{R}^{\bar{d}_k})$.

We now make the objective vector common for both families of instances. Define $X_k := \{(x,t) \in \mathbb{R}^{\bar{n}_k} \times \mathbb{R}^{\bar{d}_k} \times \mathbb{R} : x \in Q_k,\;\; t = \langle c_k, x\rangle\}$ and $X'_k := \{(x,t) \in \mathbb{R}^{\bar{n}_k} \times \mathbb{R}^{\bar{d}_k} \times \mathbb{R} : x \in Q'_k,\;\; t = \langle c'_k, x\rangle\}$. By Lemma 5, the inequality $t \leq \gamma_k$ has an exponential lower bound on sizes of proofs based on $\mathcal{D}$ for X_k and the inequality $t \leq \gamma'_k$ has an exponential lower bound on sizes of proofs based on $\mathcal{CP}$ for X'_k.

We next embed these families as faces of the same closed convex set. Define $Z_k \subseteq \mathbb{R}^{\bar{n}_k} \times \mathbb{R}^{\bar{d}_k} \times \mathbb{R} \times \mathbb{R}$, for every $k \in \mathbb{N}$, as the convex hull of $X_k \times \{0\}$ and $X'_k \times \{0\}$. We let (x,t,y) denote points in the new space $\mathbb{R}^{\bar{n}_k} \times \mathbb{R}^{\bar{d}_k} \times \mathbb{R} \times \mathbb{R}$, i.e., y denotes the last coordinate. Consider the family of inequalities $t - \gamma_k(1 - y) - \gamma'_k y \leq 0$ for every $k \in \mathbb{N}$. Note that this inequality reduces to $t \leq \gamma_k$ when $y = 0$ and it reduces to $t \leq \gamma'_k$ when $y = 1$. Thus, the inequality is valid for $Z_k \cap (\mathbb{Z}^{\bar{n}_k} \times \mathbb{R}^{\bar{d}_k} \times \mathbb{R} \times \mathbb{Z})$, i.e., when we constrain y to be an integer variable. Since $X_k \times \{0\} \subseteq Z_k$, by Lemma 3, proofs of $t - \gamma_k(1-y) - \gamma'_k y \leq 0$ based on $\mathcal{D}$ have an exponential lower bound on their size. Similarly, since $X'_k \times \{0\} \subseteq Z_k$, by Lemma 3, proofs of $t - \gamma_k(1-y) - \gamma'_k y \leq 0$ based on $\mathcal{CP}$ have an exponential lower bound on their size.

However, for branch-and-cut based on $\mathcal{CP}$ and $\mathcal{D}$, we can first branch on the variable y (recall from the hypothesis that $\mathcal{D}$ allows branching on any integer variable). Since $\mathcal{CP}$ has a polynomial proof for P_k and (c_k, γ_k) and therefore for the valid inequality $t \leq \gamma_k$ for $X_k \times \{0\}$, we can process the $y = 0$ branch in polynomial time with cutting planes. Similarly, $\mathcal{D}$ has a polynomial proof for P'_k and (c'_k, γ'_k) and therefore for the valid inequality $t \leq \gamma'_k$ for $X'_k \times \{0\}$, we can process the $y = 1$ branch also in polynomial time. Thus, branch-and-cut runs in polynomial time overall for this family of instances. □

Proof (Proof of Theorem 6). Recall that we restrict ourselves to the pure integer case, i.e., $d = 0$. Consider any branch-and-cut proof for some instance. If no cutting planes are used in the proof, this is a pure branch-and-bound proof and we are done. Otherwise, let N be a node of the proof tree where a cutting plane $\langle a, x\rangle \leq \gamma$ is used. Since we assume all cutting planes are rational, we may assume $a \in \mathbb{Z}^n$ and $\gamma \in \mathbb{Z}$. Thus, $N' = N \cap \{x : \langle a, x\rangle \geq \gamma + 1\}$ is integer infeasible. Since $\langle a, x\rangle \leq \gamma$ is in $\mathcal{CP}(N)$, by our assumption, there must be a branch-and-bound proof of polynomial size based on $\mathcal{D}$ for the validity of $\langle a, x\rangle \leq \gamma$ with respect to N. Since $N' \subseteq N$, by Lemma 3, there must be a branch-and-bound proof for the validity of $\langle a, x\rangle \leq \gamma$ with respect to N', thus proving the infeasibility of N'. In the branch-and-cut proof, one can replace the child of N by first applying the disjunction $\{x : \langle a, x\rangle \leq \gamma\} \cup \{x : \langle a, x\rangle \geq \gamma + 1\}$ on N, and then on N', applying the above branch-and-bound proof of infeasibility. We now have a branch-and-cut proof for the original instance with one less cutting plane node. We can repeat this for all nodes where a cutting plane is added and convert the entire branch-and-cut tree into a pure branch-and-bound tree with at most a polynomial blow up in size. □

Bibliography

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

[2] Arora, S., Barak, B.: Computational complexity: a modern approach. Cambridge University Press (2009)

[3] Basu, A., Conforti, M., Di Summa, M., Jiang, H.: Complexity of cutting plane and branch-and-bound algorithms for mixed-integer optimization (2019), manuscript submitted to Mathematical Programming, Ser. A, https://arxiv.org/abs/2003.05023

[4] Bixby, R.E.: Solving real-world linear programs: A decade and more of progress. Operations research **50**(1), 3–15 (2002)

[5] Bockmayr, A., Eisenbrand, F., Hartmann, M., Schulz, A.S.: On the Chvátal rank of polytopes in the 0/1 cube. Discrete Applied Mathematics **98**(1-2), 21–27 (1999)

[6] Bonet, M., Pitassi, T., Raz, R.: Lower bounds for cutting planes proofs with small coefficients. The Journal of Symbolic Logic **62**(3), 708–728 (1997)

[7] Buss, S.R., Clote, P.: Cutting planes, connectivity, and threshold logic. Archive for Mathematical Logic **35**(1), 33–62 (1996)

[8] Chen, C.P., Qi, F.: Completely monotonic function associated with the gamma functions and proof of Wallis' inequality. Tamkang Journal of Mathematics **36**(4), 303–307 (2005)

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

[10] Chvátal, V.: Cutting-plane proofs and the stability number of a graph, Report Number 84326-OR. Institut für Ökonometrie und Operations Research, Universität Bonn, Bonn (1984)

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

[12] Clote, P.: Cutting planes and constant depth frege proofs. In: Proceedings of the Seventh Annual IEEE Symposium on Logic in Computer Science. pp. 296–307 (1992)

[13] Conforti, M., Cornuéjols, G., Zambelli, G.: Integer programming, vol. 271. Springer (2014)

[14] Cook, W.J., Coullard, C.R., Turán, G.: On the complexity of cutting-plane proofs. Discrete Applied Mathematics **18**(1), 25–38 (1987)

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

[16] Cornuéjols, G., Liberti, L., Nannicini, G.: Improved strategies for branching on general disjunctions. Mathematical Programming **130**(2), 225–247 (2011)

[17] Dadush, D., Tiwari, S.: On the complexity of branching proofs. arXiv preprint arXiv:2006.04124 (2020)
[18] Dash, S.: An exponential lower bound on the length of some classes of branch-and-cut proofs. In: International Conference on Integer Programming and Combinatorial Optimization (IPCO). pp. 145–160. Springer (2002)
[19] Dash, S.: Exponential lower bounds on the lengths of some classes of branch-and-cut proofs. Mathematics of Operations Research **30**(3), 678–700 (2005)
[20] Dash, S.: On the complexity of cutting-plane proofs using split cuts. Operations Research Letters **38**(2), 109–114 (2010)
[21] Dey, S.S., Iroume, A., Molinaro, M.: Some lower bounds on sparse outer approximations of polytopes. Operations Research Letters **43**(3), 323–328 (2015)
[22] Dey, S.S., Molinaro, M., Wang, Q.: Approximating polyhedra with sparse inequalities. Mathematical Programming **154**(1-2), 329–352 (2015)
[23] Dey, S.S., Molinaro, M., Wang, Q.: Analysis of sparse cutting planes for sparse milps with applications to stochastic milps. Mathematics of Operations Research **43**(1), 304–332 (2018)
[24] Eisenbrand, F., Schulz, A.S.: Bounds on the Chvátal rank of polytopes in the 0/1-cube. Combinatorica **23**(2), 245–261 (2003)
[25] Eldersveld, S.K., Saunders, M.A.: A block-lu update for large-scale linear programming. SIAM Journal on Matrix Analysis and Applications **13**(1), 191–201 (1992)
[26] Goerdt, A.: Cutting plane versus frege proof systems. In: International Workshop on Computer Science Logic. pp. 174–194. Springer (1990)
[27] Goerdt, A.: The cutting plane proof system with bounded degree of falsity. In: International Workshop on Computer Science Logic. pp. 119–133. Springer (1991)
[28] Grigoriev, D., Hirsch, E.A., Pasechnik, D.V.: Complexity of semi-algebraic proofs. In: Annual Symposium on Theoretical Aspects of Computer Science (STACS). pp. 419–430. Springer (2002)
[29] Grötschel, M., Lovász, L., Schrijver, A.: Geometric Algorithms and Combinatorial Optimization, Algorithms and Combinatorics: Study and Research Texts, vol. 2. Springer-Verlag, Berlin (1988)
[30] Impagliazzo, R., Pitassi, T., Urquhart, A.: Upper and lower bounds for tree-like cutting planes proofs. In: Proceedings Ninth Annual IEEE Symposium on Logic in Computer Science. pp. 220–228. IEEE (1994)
[31] Jeroslow, R.G.: Trivial integer programs unsolvable by branch-and-bound. Mathematical Programming **6**(1), 105–109 (Dec 1974). https://doi.org/10.1007/BF01580225, https://doi.org/10.1007/BF01580225
[32] Karamanov, M., Cornuéjols, G.: Branching on general disjunctions. Mathematical Programming **128**(1-2), 403–436 (2011)
[33] Krajíček, J.: Discretely ordered modules as a first-order extension of the cutting planes proof system. The Journal of Symbolic Logic **63**(4), 1582–1596 (1998)

[34] Lodi, A.: Mixed integer programming computation. In: 50 Years of Integer Programming 1958-2008, pp. 619–645. Springer (2010)
[35] Mahajan, A., Ralphs, T.K.: Experiments with branching using general disjunctions. In: Operations Research and Cyber-Infrastructure, pp. 101–118. Springer (2009)
[36] Mahmoud, H., Chinneck, J.W.: Achieving milp feasibility quickly using general disjunctions. Computers & operations research **40**(8), 2094–2102 (2013)
[37] Ostrowski, J., Linderoth, J., Rossi, F., Smriglio, S.: Constraint orbital branching. In: International Conference on Integer Programming and Combinatorial Optimization. pp. 225–239. Springer (2008)
[38] Owen, J.H., Mehrotra, S.: Experimental results on using general disjunctions in branch-and-bound for general-integer linear programs. Computational optimization and applications **20**(2), 159–170 (2001)
[39] Pataki, G., Tural, M., Wong, E.B.: Basis reduction and the complexity of branch-and-bound. In: Proceedings of the twenty-first annual ACM-SIAM symposium on Discrete algorithms. pp. 1254–1261. SIAM (2010)
[40] Pudlák, P.: Lower bounds for resolution and cutting plane proofs and monotone computations. The Journal of Symbolic Logic **62**(3), 981–998 (1997)
[41] Pudlák, P.: On the complexity of the propositional calculus. London Mathematical Society Lecture Note Series pp. 197–218 (1999)
[42] Razborov, A.A.: On the width of semialgebraic proofs and algorithms. Mathematics of Operations Research **42**(4), 1106–1134 (2017)
[43] Reid, J.K.: A sparsity-exploiting variant of the Bartels-Golub decomposition for linear programming bases. Mathematical Programming **24**(1), 55–69 (1982)
[44] Rothvoß, T., Sanità, L.: 0/1 polytopes with quadratic Chvátal rank. In: International Conference on Integer Programming and Combinatorial Optimization (IPCO). pp. 349–361. Springer (2013)
[45] Schrijver, A.: Theory of Linear and Integer Programming. John Wiley and Sons, New York (1986)
[46] Sperner, E.: Ein satz über untermengen einer endlichen menge. Mathematische Zeitschrift **27**(1), 544–548 (1928)
[47] Venderbei, R.J.: https://vanderbei.princeton.edu/tex/talks/IDA_CCR/SparsityMatters.pdf (2017)
[48] Watson, G.: A note on gamma functions. Edinburgh Mathematical Notes **42**, 7–9 (1959)