STRONG IP FORMULATIONS NEED LARGE COEFFICIENTS

CHRISTOPHER HOJNY

ABSTRACT. The development of practically well-behaving integer programming formulations is an important aspect of solving linear optimization problems over a set $X \subseteq \{0,1\}^n$. In practice, one is often interested in strong integer formulations with additional properties, e.g., bounded coefficients to avoid numerical instabilities. This article presents a lower bound on the size of coefficients in any strong integer formulation of X and demonstrates that certain integer sets X require exponentially large coefficients in any strong integer formulation. We also characterize the minimum size of an integer formulation of $X \subseteq \{0,1\}^n$ with bounded coefficients.

1. INTRODUCTION

A standard approach to solve combinatorial optimization problems is to model the problem as a binary program, which can be solved by branch-and-bound. To make this approach work successfully, (i) a suitable encoding $X \subseteq \{0,1\}^n$ of the problem's solutions has to be found and (ii) an integer programming (IP) formulation of X has to be developed. In practice, however, IP formulations need to be strong and have to behave well numerically. For example, knapsack problems can be modeled as binary programs $\max\{w^\top x : a^\top x \leq \beta, x \in \{0,1\}^n\}$, where $a, w \in \mathbb{Z}_+^n$ and $b \in \mathbb{Z}_+$. But if the coefficients in a are (exponentially) large, this IP formulation may be completely useless in practice since numerical instabilities may arise.

To avoid numerical instabilities, one is interested in alternative IP formulations with small coefficients. The trade-off is that such IP formulations typically need more inequalities than IP formulations with unbounded coefficients, see [5, 23] for knapsack problems. Moreover, since there exist $X \subseteq \{0,1\}^n$ that need inequalities with exponentially large coefficients in a facet description of $\mathrm{conv}(X)$, it is intuitive that IP formulations with bounded coefficients may not be strong. Motivated by these two aspects of IP formulations with bounded coefficients, we address the following two questions. (Q1) Suppose we require the inequalities in an IP formulation to be "strong" (which will be properly defined below). What is the minimum size of the coefficients in strong IP formulations? (Q2) What is the minimum size of an IP formulation with bounded coefficients?

While answering (Q1) allows to evaluate whether a given encoding X admits strong inequalities that are numerically stable, i.e., their coefficients are small, the answer to (Q2) allows to compare different binary encodings X found in Step (i) of the above mechanism. To find answers to these questions, we analyze certain formulations of $X \subseteq \{0,1\}^n$. An *integer formulation* of X is a polyhedron P fulfilling $P \cap \mathbb{Z}^n = X$, see [16]. In contrast to this, a polyhedron $P \subseteq \mathbb{R}^n$ is called a *binary formulation* of X if $X = P \cap \{0,1\}^n$. Thus, we allow P to contain non-binary integer points, i.e., we only require P to separate X and $\{0,1\}^n \setminus X$. Note that by using this definition, we do not allow the introduction of auxiliary variables, i.e., we analyze the encoding X in its original space.

To measure the strength of an integer formulation P, different measures like integrality gaps [1, 2, 7] or the Hausdorff distance of P and $\mathrm{conv}(X)$ [4] have been discussed in the literature. In this paper, we use a different measure that is based on lattice refinements and which is described in Section 2. To answer Question (Q1), we derive a lower bound on
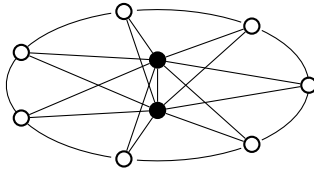

FIGURE 1. Illustration of an odd $(3,2)$-wheel. The nodes of the clique are colored black.

the maximum ratio $|a_i|/|a_j|$ of non-zero coefficients a_i, a_j in the inequalities $a^\top x \leq \beta$ of an integer formulation $Ax \leq b$ of X, the so-called ρ-value of $Ax \leq b$. We focus on this measure because it is invariant under rescaling inequalities and we are not relying on a normalization assumption such as all coefficients being integral and relatively prime. In Section 3, we discuss a natural decomposition of binary formulations to answer Question (Q2). Based on this decomposition, we exactly characterize the minimum size of any binary formulation of X with bounded coefficients. All results are illustrated by several examples.

2. Lower Bounds on Coefficients in Strong Binary Formulations

Concerning (Q1), the strongest IP formulation of $X \subseteq \{0,1\}^n$ consists of facet defining inequalities of $\mathrm{conv}(X)$. However, since the coefficients in facet defining inequalities can be exponentially large, these inequalities might be impractical. To avoid this, one may not require to cut off all points outside $\mathrm{conv}(X)$ but only the points contained in the refined integer lattice $\mathcal{L}_\lambda$ that is generated by $\frac{1}{\lambda}e_i$, $i \in [n] := \{1, \ldots, n\}$, where e_i is the i-th standard unit vector and $\lambda \in \mathbb{Z}_{>0}$.

Definition 1. Let $X \subseteq \{0,1\}^n$ and let $\lambda \in \mathbb{Z}_{>0}$. An integer formulation $Ax \leq b$, $x \in [0,1]^n$ of X is called a $\frac{1}{\lambda}$-*relaxation* if for each $\bar{x} \in \left\{0, \frac{1}{\lambda}, \frac{2}{\lambda}, \ldots, \frac{\lambda}{\lambda}\right\}^n$ the relation $A\bar{x} \leq b$, $\bar{x} \in [0,1]^n$ holds if and only if $\bar{x} \in \mathrm{conv}(X)$.

Note that $\frac{1}{1}$-relaxations are standard integer formulations and $\frac{1}{\lambda}$-relaxations converge to $\mathrm{conv}(X)$ as λ approaches infinity. Thus, $\frac{1}{\lambda}$-relaxations refine classical integer formulations: the larger λ the stronger the inequalities in such a formulation. Thus, there is hope to find λ such that the inequalities in a $\frac{1}{\lambda}$-relaxation are strong and their ρ-values are small in comparison to facet defining inequalities.

Example 2. Every $X \subseteq \{0,1\}^n$ admits a 1-relaxation with ternary coefficients (i.e., $\{0, \pm 1\}$-coefficients) via box constraints $x \in [0,1]^n$ and *infeasibility cuts*

$$\sum_{i:\,\bar{x}_i=1} x_i + \sum_{i:\,\bar{x}_i=0} (1 - x_i) \leq \mathbb{1}^\top \bar{x} - 1, \quad \bar{x} \in \{0,1\}^n \setminus X. \tag{1}$$

Example 3. An *odd (k,ℓ)-wheel* is an undirected graph G_k^ℓ that consists of a cycle C of length $2k+1$ and a clique K (disjoint from C) of size ℓ such that every node in K is connected with every node in C, see Figure 1 for an illustration. Denote the node set and edge set of C by V_C and E_C, respectively, and let V_K be the node set of K. It can be shown that the stable set polytope P of G_k^ℓ is completely described by

$$x(V_C) + k\,x(V_K) \leq k, \tag{2a}$$

$$x_u + x_v + x(V_K) \leq 1, \qquad\qquad \{u,v\} \in E_C, \tag{2b}$$

$$0 \leq x_v \leq 1, \qquad\qquad v \in V_C \cup V_K, \tag{2c}$$

where $x(S)$ abbreviates $\sum_{i \in S} x_i$ for $S \subseteq [n]$.

We claim that the system $\mathcal{S}$ given by (2b), (2c) and $x(V_C) \leq k$ is a $\frac{1}{2}$-relaxation of P. Note that the feasible region of $\mathcal{S}$ contains the points fulfilling (2). Thus, it suffices to prove that every $x \in \{0, \frac{1}{2}, 1\}^{V_C \cup V_K}$ fulfilling $\mathcal{S}$ is contained in (2). To see this, distinguish the three cases for $x(V_K) \in \{0, \frac{1}{2}, 1\}$ that are possible by (2b). If $x(V_K) = 0$, then (2a) reduces

to $x(V_C) \le k$, which is contained in $\mathcal{S}$. If $x(V_K) = 1$, then (2a) simplifies to $X(V_C) \le 0$, which is also implied by summing (2b) (which is contained in $\mathcal{S}$) for every $\{u, v\} \in E_C$ and rescaling the result by $\frac{1}{2}$. Finally, if $x(V_C) = \frac{1}{2}$, then (2a) reduces to $x(V_C) \le \frac{k}{2}$. Moreover, (2b) in $\mathcal{S}$ simplifies to $x_u + x_v \le \frac{1}{2}$. Since $x \in \{0, \frac{1}{2}, 1\}^{V_C \cup V_K}$, at most one of x_u and x_v attains value $\frac{1}{2}$, while the other attains value 0. Since C is an *odd* cycle, at most k nodes in V_C can attain value $\frac{1}{2}$. Hence, $x(V_C) \le \frac{k}{2}$. As a consequence, $\mathcal{S}$ implies validity of $x \in \{0, \frac{1}{2}, 1\}^{V_C \cup V_K}$ in (2), proving $\mathcal{S}$ is a $\frac{1}{2}$-relaxation of P.

Below, we will see, however, that there exist sets $X \subseteq \{0, 1\}^n$ whose ρ-value is exponentially in n for every $\frac{1}{\lambda}$-relaxations even if $\lambda = 2$. Thus, this paper shows that strong inequalities need exponentially large coefficients in general.

To be able to prove this result, we derive a lower bound on the ρ-value of $\frac{1}{\lambda}$-relaxations. This lower bound requires the concept of dilations of polyhedra. The λ-*dilation* of a polyhedron P is the polyhedron $\lambda P = \{\lambda x : x \in P\}$. Since the points in $\{0, \frac{1}{\lambda}, \dots, \frac{\lambda}{\lambda}\}^n$ correspond to the points in $\{0, 1, \dots, \lambda\}^n$, there is a natural correspondence between $\frac{1}{\lambda}$-relaxations of X and integer formulations of the λ-dilation $\lambda \operatorname{conv}(X)$ of $\operatorname{conv}(X)$.

Observation 4. Let $X \subseteq \{0, 1\}^n$. Then $Ax \le b$, $x \in [0, 1]^n$ is a $\frac{1}{\lambda}$-relaxation of X if and only if $Ax \le \lambda b$, $x \in [0, \lambda]^n$ is an integer formulation of $\lambda \operatorname{conv}(X)$.

To prove that any $\frac{1}{\lambda}$-relaxation of $X \subseteq \{0, 1\}^n$ has inequalities with large ρ-value, it thus suffices to show that any integer formulation of $\lambda \operatorname{conv}(X)$ contains an inequality with large ρ-value. For this reason, we investigate the ρ-value of integer formulations first. Afterwards, the results are applied to $\frac{1}{\lambda}$-relaxations via Observation 4.

2.1. INTEGER FORMULATIONS OF NON-BINARY SETS

In Example 2, we have seen that every $X \subseteq \{0, 1\}^n$ has a binary or integer formulation with ternary coefficients. If X is not binary but an arbitrary subset of $\mathbb{Z}^n$, it may not admit an integer formulation at all. This is exactly the case if $\operatorname{conv}(X) \cap \mathbb{Z}^n \ne X$. For this reason, we call $X \subseteq \mathbb{Z}^n$ *polyhedrally representable* (or *polyhedral* for short) if there exists a polyhedron $P \subseteq \mathbb{R}^n$ with $X = P \cap \mathbb{Z}^n$.

In contrast to binary sets, general polyhedral sets $X \subseteq \mathbb{Z}^n$ may not admit an integer formulation with bounded coefficients. The reason for this is that $\operatorname{conv}(X)$ for polyhedral sets $X \subseteq \mathbb{Z}^n$ might have adjacent vertices x^1, x^2 with strongly deviating entries, i.e., $|x_i^1 - x_i^2| \gg |x_j^1 - x_j^2|$ for some i, $j \in [n]$. Hence, it is intuitive that inequalities whose coefficients have almost the same size cannot cut off all infeasible integer points that are close to the edge connecting x^1 and x^2. To make this intuition precise, Theorem 6 below gives sufficient conditions to estimate the ratio of coefficients in an integer formulation. Note that for binary sets the described behavior cannot appear since binary points differ either by 0 or 1 in each coordinate.

Let $X \subseteq \mathbb{Z}^n$ be polyhedrally representable. In the following, we derive a lower bound on the ρ-value of integer formulations $Ax \le b$ of X, i.e., a bound on the maximum ratio of non-zero coefficients in inequalities of $Ax \le b$. To find such a bound for a general set $X \subseteq \mathbb{Z}^n$, define $u_i := \max\{x_i : x \in X\}$ and $\ell_i := \min\{x_i : x \in X\}$, where we define $u_i = \infty$ and $\ell_i = -\infty$ if X is unbounded in direction e_i and $-e_i$, respectively. Then $[\ell, u] := \{x \in \mathbb{R}^n : \ell_i \le x_i \le u_i, i \in [n]\}$ is the smallest box containing X. Moreover, we call an inequality that is valid for $\operatorname{conv}(X)$ *trivial* if it is $x_i \le u_i$ or $-x_i \le -\ell_i$ for some $i \in [n]$. Given a polyhedron $P \subseteq \mathbb{R}^n$ we say that it has a *bounded direction* if there exists $d \in \mathbb{R}^n$ with

$$-\infty < \min\{d^\top x : x \in P\} \le \max\{d^\top x : x \in P\} < \infty.$$

Note that every direction $d \in \mathbb{R}^n$ is bounded if P is a polytope.

In the proof of our lower bound, the following folklore result is essential. For the sake of completeness, we present a proof.

Lemma 5. *Let $P \subseteq \mathbb{R}^n$ be a full-dimensional polyhedron with m facets and let $Ax \leq b$ be the linear inequality system consisting of all facet defining inequalities. Then every valid inequality $c^\top x \leq \delta$ for P that supports P in at least one point can be written as a conic combination of facet defining inequalities, i.e., there exists $\lambda \in \mathbb{R}^m_+$ such that*

$$c^\top = \lambda^\top A \quad and \quad \delta = \lambda^\top b.$$

Moreover, if P has a bounded direction, every valid inequality for P has such a representation.

Proof. Let $c \in \mathbb{R}^n$ and $\delta \in \mathbb{R}$. The inequality $c^\top x \leq \delta$ is valid and supports P if and only if $\delta = \max\{c^\top x : Ax \leq b, \ x \in \mathbb{R}^n\}$. By LP duality, this means

$$\delta = \min\{\lambda^\top b : \lambda^\top A = c^\top, \ \lambda \in \mathbb{R}^m_+\},$$

proving the first part of the assertion.

To show the second part, observe that an inequality $c^\top x \leq \delta'$ can only be valid for P if c is contained in the normal cone of a face of P. That is, we can generate $c^\top x \leq \delta'$ by

   (i) finding a conic combination of facet defining inequalities generating a supporting valid inequality $c^\top x \leq \delta$ and
   (ii) adding the redundant inequality $0 \leq \delta' - \delta$ to this inequality.

By the first part the proof, we always find a conic combination generating $c^\top x \leq \delta$ in (i). If we further require the existence of a bounded direction $d \in \mathbb{R}^n$, note that $\delta_1 := \min\{d^\top x : x \in P\} < \max\{d^\top x : x \in P\} =: \delta_2$ as P is full-dimensional. Since $-d^\top x \leq -\delta_1$ and $d^\top x \leq \delta_2$ are supporting, there exists a conic combination of facets generating these inequalities. Extending the conic combination of $c^\top x \leq \delta$ by $\frac{\delta'-\delta}{\delta_2-\delta_1}$ times both these combinations, yields the desired inequality

$$c^\top x + \frac{\delta' - \delta}{\delta_2 - \delta_1}(d^\top x - d^\top x) \leq \delta + \frac{\delta' - \delta}{\delta_2 - \delta_1}(\delta_2 - \delta_1) \qquad \Leftrightarrow \qquad c^\top x \leq \delta',$$

proving the second assertion. $\square$

Theorem 6. *Let $X \subseteq \mathbb{Z}^n$ be polyhedral such that $P := \mathrm{conv}(X)$ is full-dimensional with bounded direction, let $Ax \leq b$ be a facet description of P, and let $\bar{x} \in ([\ell, u] \cap \mathbb{Z}^n) \setminus P$. Further, let $\bar{A}x \leq \bar{b}$ be the facet inequalities of P that are violated by $\bar{x}$, let $\bar{m}$ be the number of rows of $\bar{A}x \leq \bar{b}$, and $\bar{s} := \bar{A}\bar{x} - \bar{b}$. Assume there exist distinct $i^\star, j^\star \in [n]$ such that*

   (1) $\mathrm{sign}(\bar{A}_{st}) = \mathrm{sign}(\bar{A}_{s't}) \neq 0$ for every $s, s' \in [\bar{m}]$ and $t \in \{i^\star, j^\star\}$,
   (2) $|\bar{A}_{ki^\star}| \geq \bar{s}_k$ for every $k \in [\bar{m}]$,
   (3) $\bar{x}_{i^\star} > \ell_{i^\star}$,
   (4) $\bar{x}_{j^\star} - \ell_{j^\star} \geq \max\{\bar{s}_k : k \in [\bar{m}]\}$,
   (5) $\bar{x}_{j^\star} < u_{j^\star}$, and
   (6) for every non-trivial facet defining inequality $a^\top x \leq \beta$ not in $\bar{A}x \leq \bar{b}$ we have $\mathrm{sign}(a_t) \in \{0, \mathrm{sign}(\bar{A}_{1t})\}$, $t \in \{i^\star, j^\star\}$, and $a^\top(\bar{x} + \mathrm{sign}(\bar{A}_{1j^\star})e_{j^\star}) \leq \beta$.

Then every inequality $c^\top x \leq \delta$ valid for P that cuts off $\bar{x}$ fulfills $c_{i^\star}, c_{j^\star} \neq 0$ and

$$\frac{|c_{i^\star}|}{|c_{j^\star}|} \geq \frac{\min\{|\bar{A}_{ki^\star}| - \bar{s}_k : k \in [\bar{m}]\}}{\max\{|\bar{A}_{kj^\star}| + \bar{s}_k : k \in [\bar{m}]\}}.$$

Proof. By rescaling, we assume the inequalities $Ax \leq b$ to have integral coefficients because P is integral. Let $L, U \subseteq [n]$ be the index sets for that $x_i \geq \ell_i$ and $x_i \leq u_i$, respectively, define facets of P, and let $A'x \leq b'$ be the subsystem of $Ax \leq b$ containing all inequalities except box constraints and inequalities in $\bar{A}x \leq \bar{b}$. Let m be the number

of inequalities in $A'x \leq b'$. As P is a full-dimensional polyhedron with bounded direction, the left-hand side of each valid inequality $c^\top x \leq \delta$ for P has a representation

$$c = \lambda^\top A' + \bar{\lambda}^\top \bar{A} + \sum_{i \in U} \mu_i^+ - \sum_{i \in L} \mu_i^-,$$

where $\lambda \in \mathbb{R}_+^m$, $\bar{\lambda} \in \mathbb{R}_+^{\bar{m}}$, $\mu^+ \in \mathbb{R}_+^U$, and $\mu^- \in \mathbb{R}_+^L$, see Lemma 5. In particular,

$$c_i = \lambda^\top A'_{\cdot i} + \bar{\lambda}^\top \bar{A}_{\cdot i} + \mu_i^+ - \mu_i^- \quad \text{and} \quad \delta = \lambda^\top b' + \bar{\lambda}^\top \bar{b} + \mu^{+\top} u - \mu^{-\top} \ell, \qquad (3)$$

where $A_{\cdot i}$ denotes the i-th column of a matrix A and we define μ_i^+ and μ_i^- to be 0 if $i \notin U$ and $i \notin L$, respectively.

Let $c^\top x \leq \delta$ be a valid inequality that cuts off $\bar{x}$ and let $\bar{s} := \bar{A}\bar{x} - \bar{b} > \mathbb{0}$ be the amount by which $\bar{x}$ violates the inequalities in $\bar{A}x \leq \bar{b}$. Since $\bar{A}x \leq \bar{b}$ contains the only facet defining inequalities of P that cut off $\bar{x}$, $\bar{\lambda}_k > 0$ for at least one $k \in [\bar{m}]$ in any facet representation of $c^\top x \leq \delta$. Moreover, if $i^\star$ and $j^\star$ are defined as in the theorem, $\bar{s} > \mathbb{0}$ implies $\bar{x}_{j^\star} > \ell_{j^\star}$ by Property 4. To prove the assertion, we discuss a lower bound on $|c_{i^\star}|$ first. W.l.o.g. assume $\bar{A}_{ki^\star} > 0$ for one (and thus every) $k \in [\bar{m}]$ by Property 1. Because of (3) and the assumption that $\mathrm{sign}(a_{i^\star}) \in \{0, \mathrm{sign}(\bar{A}_{1i^\star})\}$ for every non-trivial facet inequality $a^\top x \leq \beta$ of P, the estimation

$$c_{i^\star} \overset{(3)}{=} \lambda^\top A'_{\cdot i^\star} + \bar{\lambda}^\top \bar{A}_{\cdot i^\star} + \mu_{i^\star}^+ - \mu_{i^\star}^- \geq \bar{\lambda}^\top \bar{A}_{\cdot i^\star} - \mu_{i^\star}^-$$

is valid. Furthermore, $\bar{x}$ can only violate inequality $c^\top x \leq \delta$ if its slack $\delta - c^\top \bar{x}$ is negative. Since $\bar{A}x \leq \bar{b}$ contains the only facet inequalities with a negative slack, we thus need

$$\mu_{i^\star}^- (\bar{x}_{i^\star} - \ell_{i^\star}) < \bar{\lambda}^\top (\bar{A}\bar{x} - \bar{b}) = \bar{\lambda}^\top \bar{s}, \qquad (4)$$

by (3) i.e., the weighted slack of $-x_{i^\star} \leq -\ell_{i^\star}$ is smaller than the weighted surplus of $\bar{A}x \leq \bar{b}$. Then,

$$c_{i^\star} \geq \bar{\lambda}^\top \bar{A}_{\cdot i^\star} - \mu_{i^\star}^- \overset{(4)}{>} \bar{\lambda}^\top \bar{A}_{\cdot i^\star} - \frac{\bar{\lambda}^\top \bar{s}}{\bar{x}_{i^\star} - \ell_{i^\star}} \overset{\text{Property 3}}{\geq} \bar{\lambda}^\top \bar{A}_{\cdot i^\star} - \bar{\lambda}^\top \bar{s} \geq 0,$$

where the last inequality holds by Property 2 and non-negativity of $\bar{\lambda}$. Thus, $c_{i^\star} \neq 0$.

Second, we derive an upper bound on the absolute coefficient $|c_{j^\star}|$. W.l.o.g. we can assume $\bar{A}_{\cdot j^\star} > 0$ by Property 1. Analogously to (4), one argues that

$$\mu_{j^\star}^- (\bar{x}_{j^\star} - \ell_{j^\star}) < \bar{\lambda}^\top \bar{s} \quad \text{and} \quad \mu_{j^\star}^+ (u_{j^\star} - \bar{x}_{j^\star}) + \lambda^\top (b' - A'\bar{x}) < \bar{\lambda}^\top \bar{s} \qquad (5)$$

are necessary for $c^\top x \leq \delta$ to cut off $\bar{x}$. Further note that

$$\mu_{j^\star}^- (\bar{x}_{j^\star} - \ell_{j^\star}) \overset{(5)}{<} \bar{\lambda}^\top \bar{s} \leq \max\{\bar{s}_k : k \in [\bar{m}]\} \cdot \bar{\lambda}^\top \mathbb{1}$$

$$\Leftrightarrow \quad \bar{\lambda}^\top \mathbb{1} > \frac{\bar{x}_{j^\star} - \ell_{j^\star}}{\max\{\bar{s}_k : k \in [\bar{m}]\}} \mu_{j^\star}^-. \qquad (6)$$

Because $\mathrm{sign}(a_{j^\star}) \in \{0, \mathrm{sign}(\bar{A}_{1j^\star})\}$ for every non-trivial facet defining inequality $a^\top x \leq \beta$ and $c_{j^\star} = \bar{\lambda}^\top \bar{A}_{\cdot j^\star} + \lambda^\top A'_{\cdot j^\star} + \mu_{j^\star}^+ - \mu_{j^\star}^-$ holds, we have

$$|c_{j^\star}| \leq \max\{\bar{\lambda}^\top \bar{A}_{\cdot j^\star} + \lambda^\top A'_{\cdot j^\star} + \mu_{j^\star}^+, \mu_{j^\star}^-\}. \qquad (7)$$

Moreover,

$$\bar{\lambda}^\top \bar{A}_{\cdot j^\star} \geq \bar{\lambda}^\top \mathbb{1} \overset{(6)}{>} \frac{\bar{x}_{j^\star} - \ell_{j^\star}}{\max\{\bar{s}_k : k \in [\bar{m}]\}} \mu_{j^\star}^- \geq \mu_{j^\star}^-,$$

where the first estimation is valid by integrality and positivity of $\bar{A}_{\cdot j^\star}$ and the last estimation holds by Property 4. Consequently,

$$c_{j^\star} = \bar{\lambda}^\top \bar{A}_{\cdot j^\star} + \lambda^\top A'_{\cdot j^\star} + \mu_{j^\star}^+ - \mu_{j^\star}^- \geq \bar{\lambda}^\top \bar{A}_{\cdot j^\star} - \mu_{j^\star}^- > 0$$

and thus $|c_{j^\star}| \le \bar{\lambda}^\top \bar{A}_{\cdot j^\star} + \lambda^\top A'_{\cdot j^\star} + \mu^+_{j^\star}$ by (7) and $|c_{j^\star}| > 0$. Finally, (5) and non-negativity of $\mu^+_{j^\star}$ in particular imply

$$\mu^+_{j^\star} \le \bar{\lambda}^\top \bar{s} - \lambda^\top (b' - A'\bar{x}) \tag{8}$$

since $\bar{x}_{j^\star} < u_{j^\star}$ by Property 5.

Combining the found upper and lower bounds, yields the assertion

$$\frac{|c_{i^\star}|}{|c_{j^\star}|} \ge \frac{\bar{\lambda}^\top (\bar{A}_{\cdot i^\star} - \bar{s})}{\bar{\lambda}^\top \bar{A}_{\cdot j^\star} + \lambda^\top A'_{\cdot j^\star} + \mu^+_{j^\star}} \overset{(8)}{\ge} \frac{\bar{\lambda}^\top (\bar{A}_{\cdot i^\star} - \bar{s})}{\bar{\lambda}^\top \bar{A}_{\cdot j^\star} + \lambda^\top A'_{\cdot j^\star} + \bar{\lambda}^\top \bar{s} - \lambda^\top (b' - A'\bar{x})}$$

$$\ge \frac{\bar{\lambda}^\top (\bar{A}_{\cdot i^\star} - \bar{s})}{\bar{\lambda}^\top \bar{A}_{\cdot j^\star} + \bar{\lambda}^\top \bar{s}} \ge \frac{\min\{\bar{A}_{ki^\star} - \bar{s}_k \ : \ k \in [\bar{m}]\}}{\max\{\bar{A}_{kj^\star} + \bar{s}_k \ : \ k \in [\bar{m}]\}},$$

where the third inequality holds because $A'(\bar{x} + e_{j^\star}) \le b'$ by Property 6. In particular, note that we do not divide by 0 since $\bar{s} > \mathbb{0}$ and $\bar{A}_{kj^\star} > 0$ by assumption. Thus, the assertion follows by replacing $\bar{A}_{ki^\star}$ and $\bar{A}_{kj^\star}$ by their absolute values. $\qquad\square$

Remark 7. If P has not a bounded direction, Theorem 6 remains true if we require the inequality $c^\top x \le \delta$ to support P.

Geometrically, the assumptions on the sign of coefficients ensure that X is monotone in direction $\text{sign}(\bar{A}_{1t})e_t$, $t \in \{i^\star, j^\star\}$ within $[\ell, u]$. Moreover, Properties 2–5 guarantee $\bar{x}$ to be sufficiently far from the boundary of $[\ell, u]$, while its excess $\bar{s}$ is not too large.

2.2. Examples

Using Theorem 6, we can show that certain integer sets do not admit integer formulations with small ρ-value. In the following, we focus on knapsack sets $\{x \in \mathbb{Z}^n_+ \ : \ a^\top x \le \beta\}$ that have the *divisibility property* $a_i \mid a_{i+1}$ for all $i \in [n-1]$, where we use the same notation for parameters ($i^\star$, $j^\star$, and $\bar{s}$) as in Theorem 6. Knapsack examples are of particular interest if we try to enhance an existing integer formulation $Ax \le b$ via enhancing its single row relaxations.

Lemma 8. *Let $a \in \mathbb{Z}^n_+$ have the divisibility property and let $\beta \in \mathbb{Z}_+$. If $a_n \le \beta$, the only facet defining inequality of $P := \text{conv}\{x \in \{0,1\}^n \ : \ a^\top x \le \beta\}$ that cuts off the point $\bar{x} = (\lfloor \frac{\beta}{a_1} \rfloor - \frac{a_n}{a_1} + 1)e_1 + e_n$ is*

$$x_1 + \sum_{\ell=2}^n \frac{a_l}{a_1} x_\ell \le \left\lfloor \frac{\beta}{a_1} \right\rfloor. \tag{9}$$

Proof. In [19] it is shown that P is completely characterized by the inequalities

$$x_j + \sum_{\ell=j+1}^n \frac{a_\ell}{a_j} x_\ell \le \left\lfloor \frac{\beta}{a_j} \right\rfloor, \quad j \in [n], \tag{10}$$

as well as $x_i \ge 0$ for all $i \in [n]$. However, not all of these inequalities may define facets of P. To prove the assertion, we have to show that (i) $\bar{x}$ fulfills all of these inequalities except (10) for parameter $j = 1$ and (ii) (10) for parameter $j = 1$ defines a facet of P.

To show the first point, we observe that all entries of $\bar{x}$ are non-negative. Hence, $\bar{x}$ fulfills all non-negativity constraints. Moreover, $\bar{x}$ cannot violate Inequality (10) for $j \ge 2$, because the left-hand side evaluates to $\frac{a_n}{a_j} \in \mathbb{Z}$, which is not greater than $\frac{\beta}{a_j}$. Thus, $\frac{a_n}{a_j} \le \lfloor \frac{\beta}{a_j} \rfloor$ holds. Finally, $\bar{x}$ violates Inequality (10) for parameter $j = 1$ since

$$\bar{x}_1 + \sum_{\ell=2}^n \frac{a_\ell}{a_1} \bar{x}_\ell = \left\lfloor \frac{\beta}{a_1} \right\rfloor - \frac{a_n}{a_1} + 1 + \frac{a_n}{a_1} > \left\lfloor \frac{\beta}{a_1} \right\rfloor.$$

Consequently, the first claim holds. Now the second claim follows immediately. Since (10) and non-negativity constraints provide a complete linear description of P and (10) for parameter $j = 1$ is the only inequality in this description that is violated by $\bar{x}$, (10) for $j = 1$ cannot be generated by other inequalities. Thus, it has to define a facet of P. $\quad\square$

Proposition 9. *Let $a \in \mathbb{Z}_+^n$ have the divisibility property and let $\beta \in \mathbb{Z}_+$. Provided $\lfloor \frac{\beta}{a_1} \rfloor - \frac{a_n}{a_1} \geq 0$ and $1 \leq a_1 < a_n \leq \beta$, every integer formulation of $X = \{x \in \mathbb{Z}_+^n : a^\top x \leq \beta\}$ contains an inequality $c^\top x \leq \delta$ such that $\frac{c_{i^\star}}{c_{j^\star}} \geq \frac{a_n/a_1 - 1}{2}$ for distinct $i^\star, j^\star \in [n]$.*

Proof. Consider the point $\bar{x} = (\lfloor \frac{\beta}{a_1} \rfloor - \frac{a_n}{a_1} + 1)e_1 + e_n$ as defined in Lemma 8 and let $i^\star = n$ as well as $j^\star = 1$. To prove this result, we use Theorem 6, i.e., we have to check whether its requirements are fulfilled. Due to Lemma 8, the only facet defining inequality of $P := \operatorname{conv}(X)$ is (9) with excess $\bar{s} = 1$. Moreover, the coefficients of x_1 and x_n are positive in (9). Thus, Condition 1 holds.

Since the lower bounds ℓ_i of all variables x_i are 0 and the (implicit) upper bound of x_i is $u_i = \lfloor \frac{\beta}{a_i} \rfloor$, we have $\bar{x}_n = 1 > \ell_n$ as well as

$$\bar{x}_1 - \ell_1 = \left(\left\lfloor \frac{\beta}{a_1} \right\rfloor - \frac{a_n}{a_1} + 1 \right) - 0 \geq 1 = \bar{s}, \qquad \bar{x}_1 = \left(\left\lfloor \frac{\beta}{a_1} \right\rfloor - \frac{a_n}{a_1} + 1 \right) < \left\lfloor \frac{\beta}{a_1} \right\rfloor = u_1$$

because $\beta \geq a_n$ and $a_1 < a_n$, respectively. Furthermore, the coefficient of x_1 in (9) is 1, and thus, not smaller than $\bar{s}$. Hence, Conditions 2–5 hold.

Finally, since the coefficient of x_1 is 0 in all inequalities (9) except for $j = 1$, we cannot violate any non-trivial facet defining inequality different to (10) by increasing $\bar{x}_1$ by 1. Thus, all requirements are fulfilled, showing that every integer formulation of X contains an inequality $c^\top x \leq \delta$ with $\frac{c_{i^\star}}{c_{j^\star}} \geq \frac{a_n/a_1 - 1}{2}$. $\quad\square$

Example 10. Consider the knapsack set X defined by $\sum_{i=1}^{n+1} 2^{i-1} x_i \leq 2^n$. Proposition 9 shows that every integer formulation of X contains an inequality whose coefficients diverge at least as $\frac{2^n - 1}{2}$, i.e., the ρ-value of any integer formulation of X is at least $\frac{2^n - 1}{2}$.

Example 11. Let $X_j = \{x \in \mathbb{Z}_+^n : \sum_{i=1}^n 2^{i-1} x_i \geq 2^j\}$ for $j \in \{2, \ldots, n-1\}$. Then $P = \operatorname{conv}(X)$ is completely described by $x_i \geq 0$, $i \in [n]$, and $\sum_{i=1}^j 2^{i-1} x_i + 2^j \sum_{i=j+1}^n x_i \geq 2^j$, see [21].

Consider the incidence vector $\bar{x} = \chi([j])$ of $[j]$, which is contained in the smallest box containing P. The only facet defining inequality of P that is violated by $\bar{x}$ is the non-trivial inequality

$$\bar{a}^\top x \leq \bar{\beta} \qquad :\Leftrightarrow \qquad -\sum_{i=1}^j 2^{i-1} x_i - 2^j \sum_{i=j+1}^n x_i \leq -2^j$$

whose excess w.r.t. $\bar{x}$ is $\bar{s} = 1$.

For $i^\star = j$ and $j^\star = 1$, we check $|\bar{a}_{i^\star}| = 2^{j-1} \geq \bar{s}$ and $\bar{x}_{i^\star} = \bar{x}_j > 0$ (Properties 2–3). For $j^\star$, note $\bar{x}_{j^\star} - \ell_{j^\star} = 1 \geq \bar{s}$ and $\bar{x}_{j^\star} = 1 < u_{j^\star} = \infty$ (Properties 4 and 5). Because there is exactly one non-trivial facet inequality (and this inequality is violated by $\bar{x}$), Property 6 trivially holds. Since Property 1 holds obviously on the only non-redundant inequality, the requirements of Theorem 6 except the existence of a bounded direction are fulfilled, showing that every integer formulation consisting of supporting inequalities, cf. Remark 7, contains an inequality whose coefficients diverge as $\frac{2^{i^\star - 1} - 1}{2^{j^\star - 1} + 1} = \frac{\beta/2 - 1}{2} \in \Omega(2^j)$.

The examples show that there exist sets $X \subseteq \mathbb{Z}^n$ that do not admit integer formulations with constant or polynomially bounded ρ-value. This is in contrast to binary sets, which always admit integer formulations with $\{0, \pm 1\}$-coefficients, e.g., via infeasibility cuts (1). In Section 2.4, we will see that similar results hold if we refine binary formulations of sets $X \subseteq \{0,1\}^n$ to $\frac{1}{\lambda}$-relaxations.

2.3. Generalizations and Special Cases

In the following, we discuss a special case that simplifies the requirements of Theorem 6 as well as a generalization. If $X \subseteq \mathbb{Z}_+^n$ is monotone, i.e., $x \in X$ implies that every $y \in \mathbb{Z}_+^n$ with $y \leq x$ is also contained in X, the coefficients of every non-trivial facet defining inequality are non-negative, cf. [22, Section 9.3]. Thus, the last condition of Theorem 6 can be simplified.

Corollary 12. *Let $X \subseteq \mathbb{Z}^n$ be polyhedral such that $P \coloneqq \mathrm{conv}(X)$ is full-dimensional with bounded direction, let $Ax \leq b$ be a facet description of P, and let $\bar{x} \in ([\ell, u] \cap \mathbb{Z}^n) \setminus P$. Using the same notation as in Theorem 6, assume Conditions 1–5 of Theorem 6 as well as*

(6') for every non-trivial facet defining inequality $a^\top x \leq \beta$ not in $\bar{A}x \leq \bar{b}$ we have $a^\top(\bar{x} + \mathrm{sign}(\bar{A}_{1j^\star})e_{j^\star}) \leq \beta$

hold for two distinct indices $i^\star, j^\star \in [n]$. Then every inequality $c^\top x \leq \delta$ valid for P that cuts off $\bar{x}$ fulfills $c_{i^\star}, c_{j^\star} \neq 0$ and

$$\frac{|c_{i^\star}|}{|c_{j^\star}|} \geq \frac{\min\{|\bar{A}_{ki^\star}| - \bar{s}_k \ : \ k \in [\bar{m}]\}}{\max\{|\bar{A}_{kj^\star}| + \bar{s}_k \ : \ k \in [\bar{m}]\}}.$$

Even if restricted to monotone sets, a critical property is that $a^\top(\bar{x} + \mathrm{sign}(\bar{A}_{1j^\star})e_{j^\star}) \leq \beta$ holds for every non-trivial facet defining inequality $a^\top x \leq \beta$ that is not violated by $\bar{x}$. In particular, this means that if there exists a non-trivial facet inequality that is fulfilled by $\bar{x}$ with equality, the coefficient $a_{j^\star}$ has to be zero. This requirement is very restrictive. For this reason, we describe in the remainder of this section a generalization of Theorem 6 that allows to treat this case.

To this end, we have to extend our notation of facet defining inequalities. Given a polyhedron $P \subseteq \mathbb{R}^n$, we assume in the following that $Ax \leq b$ is a facet description of P with m rows. For a set $I \subseteq [m]$, we denote by $A_I x \leq b_I$ the subsystem of $Ax \leq b$ containing the inequalities whose row indices are contained in I. For a point $\bar{x} \in \mathbb{Z}^n$, we define the partition $L \cup E \cup G$ of $[m]$ such that

$$A_L \bar{x} < b_L, \qquad A_E \bar{x} = b_E, \qquad A_G \bar{x} > b_G$$

holds.

Theorem 13. *Let $X \subseteq \mathbb{Z}^n$ be polyhedral such that $P \coloneqq \mathrm{conv}(X)$ is full-dimensional with bounded direction, let $Ax \leq b$ be a facet description of P, and let $\bar{x} \in ([\ell, u] \cap \mathbb{Z}^n) \setminus P$. Let $\bar{s} \coloneqq A_G \bar{x} - b_G$. Assume there exist distinct $i^\star, j^\star \in [n]$ such that Conditions 1–5 from Theorem 6 as well as*

(6") for every non-trivial facet defining inequality $a^\top x \leq \beta$ in $A_L x \leq b_L$ we have $\mathrm{sign}(a_t) \in \{0, \mathrm{sign}(A_{Gt})\}, t \in \{i^\star, j^\star\}$, and $a^\top(\bar{x} + \mathrm{sign}(A_{Gj^\star})e_{j^\star}) \leq \beta$

hold. Then every inequality $c^\top x \leq \delta$ valid for P that cuts off $\bar{x}$ fulfills $c_{i^\star}, c_{j^\star} \neq 0$ and

$$\frac{|c_{i^\star}|}{|c_{j^\star}|} \geq \min_{\lambda \in \mathbb{R}_+^E} \left\{ \frac{\min\{|\bar{A}_{ki^\star}| - \bar{s}_k \ : \ k \in [\bar{m}]\} + \lambda^\top A_{Ei^\star}}{\max\{|\bar{A}_{kj^\star}| + \bar{s}_k \ : \ k \in [\bar{m}]\} + \lambda^\top A_{Ej^\star}} \right\}.$$

Proof. The proof is almost the same as for Theorem 6. The only difference is that we have to take more care about facet defining inequalities of P that are fulfilled with equality by $\bar{x}$. For this reason, we provide the refined estimations only.

First, as in the proof of Theorem 6, we have to find a lower bound on $|c_{i^\star}|$. Again we assume $\mathrm{sign}(A_{Gi^\star}) = 1$. Let $\lambda \in \mathbb{R}_+^m$. Then

$$c_{i^\star} \overset{(3)}{=} \lambda^\top A_{\cdot i^\star} + \mu_{i^\star}^+ - \mu_{i^\star}^- = \lambda_L^\top A_{Li^\star} + \lambda_E^\top A_{Ei^\star} + \lambda_G^\top A_{Gi^\star} + \mu_{i^\star}^+ - \mu_{i^\star}^-.$$

Using Inequality (4), we find

$$c_{i^\star} \geq \lambda_E^\top A_{Ei^\star} + \lambda_G^\top A_{Gi^\star} - \mu_{i^\star}^- \overset{(4)}{>} \lambda_E^\top A_{Ei^\star} + \lambda_G^\top A_{Gi^\star} - \frac{\lambda_G^\top \bar{s}}{\bar{x}_{i^\star} - \ell_{i^\star}} \geq \lambda_E^\top A_{Ei^\star} + \lambda_G^\top A_{Gi^\star} - \lambda_G^\top \bar{s}.$$

Second, to find an upper bound on $|c_{j^\star}|$, we assume $\mathrm{sign}(A_{Gj^\star}) = 1$. In complete analogy to the proof of Theorem 6, one can show that

$$c_{j^\star} \leq \lambda_L^\top A_{Lj^\star} + \lambda_E^\top A_{Ej^\star} + \lambda_G^\top A_{Gj^\star} + \mu_{j^\star}^+.$$

Third, we have to refine Inequality (8). Since $b_E - A_E \bar{x} = \mathbb{0}$, we immediately get from (8)

$$\mu_{j^\star}^+ \leq \lambda_G^\top \bar{s} - \lambda_L^\top (b_L - A_L \bar{x}). \tag{11}$$

Combining the found bounds yields the assertion:

$$
\begin{aligned}
\frac{|c_{i^\star}|}{|c_{j^\star}|} &\geq \frac{\lambda_G^\top (A_{Gi^\star} - \bar{s}) + \lambda_E^\top A_{Ei^\star}}{\lambda_G^\top A_{Gj^\star} + \lambda_E^\top A_{Ej^\star} + \lambda_L^\top A_{Lj^\star} + \mu_{j^\star}^+} \\
&\overset{(11)}{\geq} \frac{\lambda_G^\top (A_{Gi^\star} - \bar{s}) + \lambda_E^\top A_{Ei^\star}}{\lambda_G^\top A_{Gj^\star} + \lambda_E^\top A_{Ej^\star} + \lambda_L^\top A_{Lj^\star} + \lambda_G^\top \bar{s} - \lambda_L^\top (b_L - A_L \bar{x})} \\
&\geq \frac{\lambda_G^\top (A_{Gi^\star} - \bar{s}) + \lambda_E^\top A_{Ei^\star}}{\lambda_G^\top A_{Gj^\star} + \lambda_G^\top \bar{s} + \lambda_E^\top A_{Ej^\star}} \\
&\geq \frac{\min\{A_{ki^\star} - \bar{s}_k \,:\, k \in G\} \cdot \lambda_G^\top \mathbb{1} + \lambda_E^\top A_{Ei^\star}}{\max\{A_{ki^\star} + \bar{s}_k \,:\, k \in G\} \cdot \lambda_G^\top \mathbb{1} + \lambda_E^\top A_{Ej^\star}} \\
&= \frac{\min\{A_{ki^\star} - \bar{s}_k \,:\, k \in G\} + \frac{1}{\lambda_G^\top \mathbb{1}} \cdot \lambda_E^\top A_{Ei^\star}}{\max\{A_{ki^\star} + \bar{s}_k \,:\, k \in G\} + \frac{1}{\lambda_G^\top \mathbb{1}} \cdot \lambda_E^\top A_{Ej^\star}} \\
&=: \frac{\min\{A_{ki^\star} - \bar{s}_k \,:\, k \in G\} + \tilde{\lambda}_E^\top A_{Ei^\star}}{\max\{A_{ki^\star} + \bar{s}_k \,:\, k \in G\} + \tilde{\lambda}_E^\top A_{Ej^\star}} \\
&\geq \min_{\tilde{\lambda} \in \mathbb{R}_+^E} \left\{ \frac{\min\{A_{ki^\star} - \bar{s}_k \,:\, k \in G\} + \tilde{\lambda}_E^\top A_{Ei^\star}}{\max\{A_{ki^\star} + \bar{s}_k \,:\, k \in G\} + \tilde{\lambda}_E^\top A_{Ej^\star}} \right\}.
\end{aligned}
$$

$\square$

We are now able to handle facet defining inequalities that are fulfilled by $\bar{x}$ with equality. The price we have to pay, however, is that we no longer have a concrete lower bound on the ρ-value. Instead, we need to solve an optimization problem to find the lower bound.

2.4. Strengthened Integer Formulations of Binary Sets

In the following, we demonstrate that $\frac{1}{\lambda}$-relaxations with coefficients of constant size do not exist in general by discussing two examples. In particular, such relaxations may have an exponential ρ-value even if $\lambda = 2$ and X is a knapsack set. Thus, even in this simple case, there is no positive answer to (Q1). To prove these results, we use Theorem 6 and Observation 4.

To state our examples, we consider special (generalized) knapsack polytopes, so-called symresacks for "monotone and ordered permutations", which have been investigated in [13]. Valid inequalities for these polytopes can be used to cut off binary vectors that are lexicographically smaller than their permutation w.r.t. γ, and thus, as symmetry handling inequalities for binary programs, see [13] for more details.

The feasible region of these symresacks can be defined by a sequence $k_1, \ldots, k_q$ of integers greater or equal to 2 with $\sum_{i=1}^q k_i = n$. Let $s_i = \sum_{j=1}^{i-1} k_j$. Then, the vertices of the

symresacks form the set

$$X = \left\{ x \in \{0,1\}^n : \ -\sum_{i=1}^{q} \sum_{j=s_i+1}^{s_i+k_i-1} 2^{n-(j+1)} x_j + \sum_{i=1}^{q} (2^{n-s_i-1} - 2^{n-s_i-k_i}) x_{s_i+k_i} \leq 0 \right\}.$$

If, for example, $k_i = 2$ for all $i \in [q]$, the knapsack inequality describing X simplifies and we get

$$X' = \left\{ x \in \{0,1\}^n : \ -\sum_{i=1}^{q} 2^{n-2i} x_{2i-1} + \sum_{i=1}^{q} 2^{n-2i} x_{2i} \leq 0 \right\},$$

the so-called orbisack, which has also been investigated in [15, 18].

In [13], a complete linear description of $P = \mathrm{conv}(X)$ is derived: Let $\ell \in [q]$ and let $\kappa \in \{0,1,2,3\}^q$ such that $\kappa_i \in \{1,2,3\}$ for all $i \in [\ell-1]$ and $\kappa_i = 0$ otherwise. For each such κ, define a vector $\alpha = \alpha(\kappa) \in \mathbb{R}^q$ via

$$\alpha_i = \begin{cases} 0, & \text{if } i \geq \ell, \\ 1, & \text{if } i = \ell - 1, \\ \alpha_{i+1}, & \text{if } i < \ell - 1 \text{ and } \kappa_{i+1} \neq 3, \\ k_{i+1}\,\alpha_{i+1}, & \text{if } i < \ell - 1 \text{ and } \kappa_{i+1} = 3. \end{cases}$$

With the aid of this vector, define an inequality $a^\top x \leq \beta$ by

$$a_{s_i+k_i} = \begin{cases} 0, & \text{if } \kappa_i \in \{0,1\}, \\ \alpha_i, & \text{if } \kappa_i = 2, \\ (k_i - 1)\alpha_i & \text{if } \kappa_i = 3, \end{cases} \quad \text{and} \quad a_j = \begin{cases} 0, & \text{if } \kappa_i \in \{0,2\}, \\ -\alpha_i, & \text{if } \kappa_i \in \{1,3\}, \end{cases} \tag{12}$$

where $i \in [q]$ and $j \in [n] \setminus \{s_i + k_i : \ i \in [q]\}$, and $\beta := \sum_{i \in [q]:\ \kappa_i = 2} \alpha_i$. Using these definitions, a complete linear description of P is given by box constraints and

$$x_{s_\ell + k_\ell} - x_{s_\ell + r} + \sum_{j=1}^{s_\ell} a_j\, x_j \leq \beta \tag{13}$$

for every $r \in [k_\ell - 1]$, $\ell \in [q]$, and every choice of κ fulfilling the above requirements. If $k_1 \geq 3$, Inequality (13) defines a facet of P if and only if $\kappa_1 \in \{2,3\}$, see [13]. Moreover, if $k_i = 2$, then it defines a facet if and only if $\kappa_1 = 3$. Thus, due to the definition of α, the coefficients in facet defining inequalities can be exponentially large in this case.

Example 14. In the following, we show that each $\frac{1}{3}$-relaxation of P needs exponentially large coefficients, provided $k_i \geq 3$ for all $i \in [q-1]$ and $k_q - k_{q-1} = 2$. To this end, we use Theorem 6 to find bounds on the ρ-value of integer formulations of the 3-dilation of P, cf. Observation 4.

Consider the point $\bar{x} = \mathbb{1} + e_n$. Since all entries of $\bar{x}$ are either 1 or 2, no trivial inequality of the 3-dilation can be violated. The slack of the 3-dilation of (13) is

$$S := 3\beta - \bar{x}_{s_\ell + k_\ell} + \bar{x}_{s_\ell + r} - \sum_{j=1}^{s_\ell} a_j\, \bar{x}_j = 3\beta - \tau - \sum_{j=1}^{s_\ell} a_j,$$

where $\tau = 1$ if $\ell = q$ and 0 otherwise. To find bounds on the violation, we distinguish the following cases:

(1) If $\kappa_i = 3$ for all $i \in [\ell - 1]$, we have $\sum_{j=1}^{s_l} a_j = 0$ since $a_{s_i+k_i} = -\sum_{j=s_i+1}^{s_i+k_i-1} a_j$ by (12) if $\kappa_i \in \{0,3\}$, and $\beta = 0$. Thus, $S = -\tau$ in this case. That is, (13) is only violated if $\ell = q$. Otherwise, the inequality is fulfilled with equality.

(2) If $\kappa_i \in \{1, 3\}$ for all $i \in [\ell - 1]$ and there exists $r \in [\ell - 1]$ with $\kappa_r = 1$, we have $\beta = 0$ and

$$\sum_{j=1}^{s_l} a_j \leq \sum_{j=s_r+1}^{s_r+k_r-1} a_j \leq -2\,\alpha_r,$$

because coefficients associated with $\kappa_i = 3$ cancel out (compare Case 1) and coefficients associated with $\kappa_i = 1$ are non-positive and coincide on $\{s_i+1, \ldots, s_i+k_i-1\}$. Thus, $S \geq -\tau + 2\,\alpha_r \geq -1 + 2 = 1$, showing that no facet defining inequality for such κ can be violated. In particular, its slack is at least 1.

(3) If $\kappa_i \in \{1, 2, 3\}$ for all $i \in [\ell - 1]$ and there exists $r \in [\ell - 1]$ with $\kappa_r = 2$, we have

$$\sum_{j=1}^{s_\ell} a_j \leq \sum_{i:\,\kappa_i=2} a_{s_i+k_i} = \beta$$

because the sum of coefficients for $\kappa_i \in \{1, 3\}$ is at most 0, see the previous cases, and only $a_{s_i+k_i}$ is non-zero on $\{s_i+1, \ldots, s_i+k_i\}$ if $\kappa_i = 2$. Hence, $S = 3\,\beta - \tau - \beta \geq 2\,\beta - 1 \geq 1$. Thus, also in this case the slack is at least 1 and no inequality can be violated by $\bar{x}$.

For this reason, the only facet defining inequality violated by $\bar{x}$ is the inequality associated to $\kappa = 3 \cdot \chi([q-1])$. In the following, we refer to this inequality as $\bar{a}^\top x \leq 0$. Recall that $\bar{s} = \bar{a}^\top \bar{x} = 1$ from the first case. To find bounds on the ρ-value of inequalities in an integer formulation of $3P \cap \mathbb{Z}^n$, we use Theorem 6. Choose $i^\star = k_1$ and $j^\star = n$. Then,

(1) $\operatorname{sign}(\bar{a}_{i^\star}), \operatorname{sign}(\bar{a}_{j^\star}) \neq 0$.
(2) $|\bar{a}_{i^\star}| \neq 0$ since $\kappa_1 \in \{2, 3\}$ for every facet defining inequality, see above. Hence, $|\bar{a}_{i^\star}| \geq 1 = \bar{s}$.
(3) $\bar{x}_{i^\star} = 1 > 0 = \ell_{i^\star}$.
(4) $2 - 0 = \bar{x}_{j^\star} - \ell_{j^\star} \geq 1 = \bar{s}$.
(5) $3 = u_{j^\star} > \bar{x}_{j^\star} = 2$.
(6) From the definition of $a(\kappa)$, every non-trivial facet defining inequality $a^\top x \leq 3\beta$ of $3P$ fulfills $\operatorname{sign}(a_t) \in \{0, \operatorname{sign}(\bar{a}_t)\}$ for every $t \in \{i^\star, j^\star\}$. Further, as the coefficient of $x_{j^\star}$ is 0 for each inequality with $\ell \neq q$, we have $a^\top(\bar{x} + e_{j^\star}) = a^\top \bar{x} \leq 3\beta$. Finally, if $\ell = q$, the coefficient of $x_{j^\star}$ in $a^\top x \leq \beta$ is 1 and the slack of $a^\top x \leq 3\beta$ w.r.t. $\bar{x}$ is at least 1 by Cases 2 and 3. Hence, $\bar{x} + e_{j^\star}$ cannot violate $a^\top x \leq 3\beta$.

Consequently, all requirements of Theorem 6 are fulfilled, which shows that every inequality in a $\frac{1}{3}$-relaxation of X cutting off $\bar{x}$ has an inequality $c^\top x \leq \delta$ with ρ-value

$$\frac{|c_{i^\star}|}{|c_{j^\star}|} \geq \frac{\bar{a}_{k_1} - 1}{\bar{a}_n + 1} = \frac{\bar{a}_{k_1} - 1}{2}.$$

For the special case where $k_i = 3$ for every $i \in [q-1]$, this implies

$$\frac{|c_{i^\star}|}{|c_{j^\star}|} \geq \frac{3^{q-2} - 1}{2} = \frac{3^{n/3-2} - 1}{2}.$$

Example 15. Using Theorem 13, we demonstrate that the orbisack X' needs inequalities with exponentially large ρ-value in any $\frac{1}{2}$-relaxation. To this end, consider $\bar{x} = \mathbb{1} + e_n$. This point is contained in the smallest box containing the 2-dilation $2P$ of $P := \operatorname{conv}(X')$. As in Example 14, we find bounds on the slack

$$S := 2\,\beta + \bar{x}_{n-1} - \bar{x}_n - \sum_{j=1}^{\ell-1} a_{2j-1}\,\bar{x}_{2j-1} - \sum_{j=1}^{\ell-1} a_{2j}\,\bar{x}_{2j} = 2\,\beta - \tau - \sum_{j=1}^{\ell-1} a_{2j-1}\,\bar{x}_{2j-1} - \sum_{j=1}^{\ell-1} a_{2j}\,\bar{x}_{2j}$$

of facet defining inequalities different to box constraints, where again $\tau = 1$ if $\ell = q$ and $\tau = 0$ otherwise. Analogously to Example 14, we distinguish the following three cases.

(1) If $\kappa_i = 3$ for all $i \in [\ell - 1]$, we have $\sum_{i=1}^{\ell-1} a_{2i-1} = \sum_{i=1}^{\ell-1} a_{2i}$ since the coefficient of x_{2i-1} and x_{2i} is the same if $\kappa_i \in \{0, 3\}$, and $\beta = 0$. Thus, $S = -\tau$ in this case. That is, (13) is only violated if $\ell = q$. Otherwise, the inequality is fulfilled with equality.

(2) If $\kappa_i \in \{1, 3\}$ for all $i \in [\ell-1]$ and there exists $r \in [\ell-1]$ with $\kappa_r = 1$, we have $\beta = 0$ and

$$\sum_{i=1}^{\ell-1} a_{2i-1} + \sum_{i=1}^{\ell-1} a_{2i} = \sum_{i:\ \kappa_i = 1} a_{2i-1} = - \sum_{i:\ \kappa_i = 1} \alpha_i,$$

because coefficients associated with $\kappa_i = 3$ cancel out (compare Case 1) and coefficients associated with $\kappa_i = 1$ are non-positive. Thus, $S = -\tau + \sum_{i:\ \kappa_i = 1} \alpha_i$. Consequently, due to the definition of α and (12), the inequality has a slack of at least 1 provided that
  - $\ell \in [q - 1]$,
  - $\ell = q$ and $|\{i \in [\ell - 1] :\ \kappa_i = 1\}| \geq 2$, or
  - $\ell = q$ and $\max\{i \in [\ell - 1] :\ \kappa_i = 1\} < n - 1$.

Otherwise, i.e., $\ell = q$ and the only 1-entry of κ is entry $n - 1$, the inequality is fulfilled with equality.

(3) If $\kappa_i \in \{1, 2, 3\}$ for all $i \in [\ell - 1]$ and there exists $r \in [\ell - 1]$ with $\kappa_r = 2$, we have

$$\sum_{i=1}^{\ell-1} a_{2i-1} + \sum_{i=1}^{\ell-1} a_{2i} \leq \sum_{i:\ \kappa_i = 2} a_{2i} = \sum_{i:\ \kappa_i = 2} \alpha_i = \beta,$$

because the sum of coefficients for $\kappa_i \in \{1, 3\}$ is at most 0 and $a_{2i-1} = 0$ if $\kappa_i = 2$. Hence, $S \geq 2\beta - \tau - \beta = \beta - \tau = \sum_{i:\ \kappa_i = 2} \alpha_i - \tau$. Thus, the inequality has a slack of at least 1 provided that
  - $\ell \in [q - 1]$,
  - $\ell = q$ and $|\{i \in [\ell - 1] :\ \kappa_i = 1 \text{ or } \kappa_i = 2\}| \geq 2$, or
  - $\ell = q$ and $\max\{i \in [\ell - 1] :\ \kappa_i = 2\} < n - 1$.

Otherwise, i.e., $\ell = q$, no entry of κ attains value 1, and the only 2-entry of κ is entry $n - 1$, the inequality is fulfilled with equality.

For this reason, the only facet defining inequality of P that is violated by $\bar{x}$ is the inequality associated to $\bar{\kappa} = 3 \cdot \chi([q - 1])$. The only inequalities that are fulfilled by equality are the ones associated to $\bar{\kappa} - e_{q-1}$ and $\bar{\kappa} - 2e_{q-1}$. To find bounds on the ρ-value of a $\frac{1}{2}$-relaxation of X', choose $i^\star = 1$ and $j^\star = n - 1$. Conditions 1–5 are easily verified as the slack of the only violated facet defining inequality is 1. Thus, it remains to argue that Condition (6") holds. Since the coefficient of every non-trivial facet defining inequality for variable x_{n-1} is -1 or 0, decreasing $\bar{x}_n$ by 1 cannot violate any inequality that has a positive slack. Thus, all conditions are met and we have to take the following three inequalities for the bound of Theorem 13 into account (which are the only ones that are violated or fulfilled with equality by $\bar{x}$):

$$\kappa = 3 \cdot \chi([q-1]) : \qquad -x_{n-1} + x_n - \sum_{i=1}^{q-1} 2^{q-1-i} x_{2i-1} + \sum_{i=1}^{q-1} 2^{q-1-i} x_{2i} \leq 0$$

$$\kappa = 3 \cdot \chi([q-1]) - e_{q-1} : \quad -x_{n-1} + x_n + x_{n-2} - \sum_{i=1}^{q-2} 2^{q-2-i} x_{2i-1} + \sum_{i=1}^{q-1} 2^{q-2-i} x_{2i} \leq 2$$

$$\kappa = 3 \cdot \chi([q-1]) - 2e_{q-1} : \quad -x_{n-1} + x_n - x_{n-3} - \sum_{i=1}^{q-2} 2^{q-2-i} x_{2i-1} + \sum_{i=1}^{q-1} 2^{q-2-i} x_{2i} \leq 0.$$

Thus, Theorem 13 implies for any inequality $c^\top x \le \delta$ contained in a $\frac{1}{2}$-relaxation of X' that

$$\frac{|c_1|}{|c_{n-1}|} \ge \min_{\lambda \in \mathbb{R}_+^2} \left\{ \frac{2^{q-2} - 1 + (\lambda_1 + \lambda_2)2^{q-3}}{1 + 1 + \lambda_1 + \lambda_2} \right\} = \min_{\lambda \in \mathbb{R}_+} \left\{ \frac{2^{q-2} - 1 + \lambda 2^{q-3}}{2 + \lambda} \right\} \ge \frac{2^{q-2} - 1}{2}$$

since the derivative of the last term w.r.t. λ is positive on $\mathbb{R}_+$. That is, every $\frac{1}{2}$-relaxation needs an inequality with ρ-value in $\Omega(2^q) = \Omega(2^{n/2})$.

Consequently, the ρ-value of inequalities in $\frac{1}{\lambda}$-relaxations can be exponentially large even if $\lambda = 2$. Thus, strong integer formulations need exponential coefficients in general. In particular, these result show that strong inequalities handling symmetries of a single permutation are typically numerically unstable.

3. Decompositions of Binary Formulations

In the previous section, we have seen that there exist $X \subseteq \{0,1\}^n$ such that the ρ-value of any of its $\frac{1}{\lambda}$-relaxations is exponentially large if $\lambda \ge 2$. In practice, however, it might be possible to drop some of these inequalities since they are irrelevant for solving an optimization problem. Below, we formalize this idea for that we need a suitable partition of the infeasible binary points $X^c := \{0,1\}^n \setminus X$. Afterwards, we use this decomposition to provide answers to Questions (Q1) and (Q2). We assume $X \subseteq \{0,1\}^n$ throughout this section.

3.1. The Decomposition

Let $G_n = (V_n, E_n)$ be the undirected graph defined by the 1-skeleton of the hypercube $[0,1]^n$. Define G_n^X to be the subgraph of G_n induced by X^c, i.e., remove all nodes in X from V_n and all edges from E_n that have an endpoint in X. A connected component of G_n^X contains all those nodes $x \in X^c$ that are reachable from a node y of G_n^X by only using edges of $[0,1]^n$ that do not contain nodes from X. Denote with $\mathcal{K}_n^X$ the node sets of the connected components of G_n^X.

Proposition 16. *Let K_1, $K_2 \in \mathcal{K}_n^X$ be distinct and let $x^i \in \operatorname{conv}(K_i)$, $i \in \{1,2\}$. Then any valid inequality for $\operatorname{conv}(X)$ that cuts off x^1 cannot cut off x^2.*

To prove this proposition, we need the following lemma.

Lemma 17. *Let $a^\top x \le \beta$ be violated by the null vector $\mathbb{0}$ and the all-one vector $\mathbb{1}$. Then there exists a path from $\mathbb{0}$ to $\mathbb{1}$ in G_n such that every vertex along this path violates $a^\top x \le \beta$ as well.*

Proof. Assume w.l.o.g. that $a_i \ge a_{i+1}$, $i \in [n-1]$. Let k be the number of entries in a that are non-negative. Then, $\beta < a^\top \mathbb{0} = 0 \le a([1]) \le a([2]) \le \cdots \le a([k])$ and $\beta < a^\top \mathbb{1} = a([n]) \le a([n-1]) \le \cdots \le a([k+1])$. Hence, $\mathbb{0}$ and $\chi([i])$, $i \in [n]$, which form a path from $\mathbb{0}$ to $\mathbb{1}$, violate the inequality, proving the assertion. $\square$

We are now able to provide the missing proof.

Proof of Proposition 16. Due to convexity, it is sufficient to prove the statement for points $x^1 \in K_1$ and $x^2 \in K_2$. We use induction over n. For $n = 1$, the assertion clearly holds. Thus, assume the statement holds for dimensions less than n.

Let K_1, $K_2 \in \mathcal{K}_n^X$ be distinct, and let $x^1 \in K_1$ and $x^2 \in K_2$. For the sake of contradiction, suppose there is an inequality $a^\top x \le \beta$ valid for $\operatorname{conv}(X)$ that cuts off x^1 and x^2. Let $I := \{i \in [n] : x_i^1 = x_i^2\}$. Consider the face F of $[0,1]^n$ that is defined via $F = \operatorname{conv}(\{x \in \{0,1\}^n : x_i = x_i^1, i \in I\})$. If F is a non-trivial face of $[0,1]^n$, we can use the induction hypothesis to show the assertion since F is (isomorphic to) $[0,1]^{n-|I|}$ and x^1, $x^2 \in F$. If $F = [0,1]^n$, we can assume w.l.o.g. that $x^1 = \mathbb{0}$ and $x^2 = \mathbb{1}$ by complementing
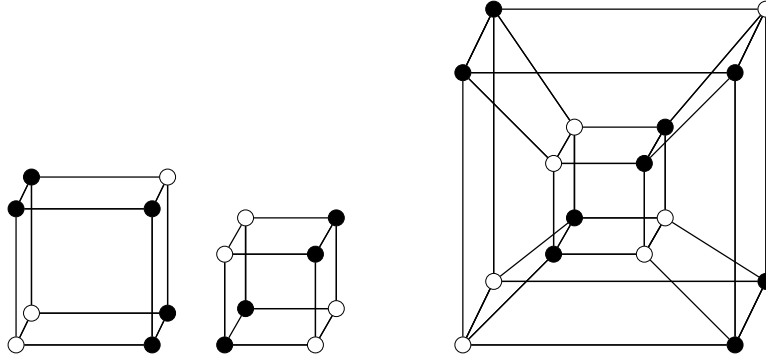
FIGURE 2. Illustration of on-/off-constraints. Left: Integer points in two 3-dimensional polytopes (black vertices). Right: The feasible points w.r.t. on-/off-constraint.

variables. Now, Lemma 17 implies that there exists a path in G_n from x^1 to x^2 that is completely cut off by $a^\top x \leq \beta$. This is a contradiction to the fact that x^1 and x^2 lie in different connected components of G_n^X. □

Corollary 18. *Let $X \subseteq \{0,1\}^n$ and let P_K be a $\frac{1}{\lambda}$-relaxation of $K^c = \{0,1\}^n \setminus K$, where $K \in \mathcal{K}_n^X$. Then $\bigcap_{K \in \mathcal{K}_n^X} P_K$ is a $\frac{1}{\lambda}$-relaxation of X.*

In the following, we present two applications of Corollary 18.

Example 19. Consider the multilinear term $z(x) = x_1 \cdots x_n$. The graph of z over $\{0,1\}^n$ is the set $X = \{(x,z) \in \{0,1\}^{n \times 1} : z = x_1 \cdots x_n\}$. The infeasible binary points w.r.t. X are $\bar{x} = (1, \ldots, 1, 0)$ and the points $\chi(I \cup \{z\})$ for every $I \subsetneq [n]$. Since every neighbor of $\bar{x}$ is feasible, the infeasible region of X^c decomposes into 2 connected components. Thus, $\frac{1}{\lambda}$-relaxations of X can be found by combining $\frac{1}{\lambda}$-relaxations for both connected components.

Example 20. Another application of Corollary 18 are on-/off-constraints. Consider two systems $Ax \leq b$ and $Cx \leq d$ of linear inequalities and an additional variable $z \in \{0,1\}^n$ that controls which of both systems is active, i.e., the set

$$X = \{(x,z) \in \{0,1\}^{n+1} : Ax \leq b \text{ if } z = 0 \text{ and } Cx \leq d \text{ if } z = 1\}.$$

The set X can be interpreted as the binary points contained in the convex hull of the sets

$$P_0 = \{(x,0) : Ax \leq b, x \in \{0,1\}^n\} \quad \text{and} \quad P_1 = \{(x,1) : Cx \leq d, x \in \{0,1\}^n\}.$$

Depending on the structure of infeasible points w.r.t. P_i in $\{(x,z) \in \{0,1\}^{n+1} : z = i\}$, the infeasible region of X can decompose in several connected components, see Figure 2 for an illustration.

A particular example is the graph of a multilinear term. The convex hull of the graph is described by the following inequalities, see [11]:

$$\sum_{i=1}^n x_i \leq n - 1 + z,$$
$$z \leq x_i, \qquad\qquad\qquad i \in [n],$$
$$0 \leq x_i \leq 1, \qquad\qquad i \in [n].$$

The z-variable controls whether the x-vector is contained in $\{x \in \{0,1\}^n : \mathbb{1}^\top x \leq n - 1\}$ ($z = 0$) or $\{\mathbb{1}\}$ ($z = 1$).

Remark 21. In [3], it is shown that $\text{conv}(X) = \bigcap_{K \in \mathcal{K}_n^X} \text{conv}(K^c)$. However, we cannot use this result to show Corollary 18 in the case of $\text{conv}(X)$ being full-dimensional, because we cannot directly deduce that every valid inequality (generated by conic combinations of facet defining inequalities) cuts off points from at most one set $K \in \mathcal{K}_n^X$.

3.2. COEFFICIENT REDUCTION AND INCREMENTAL DECOMPOSITION

In practice, one is often not only interested in finding a strong binary formulation for a set X, but also in solving a linear optimization problem over X. Thus, it is natural to extend Question (Q1) to strong binary formulations P that are adapted to an objective $w \in \mathbb{R}^n$. That is, instead of requiring $P \cap \{0,1\}^n = X$, we demand

$$\operatorname{argmax}\{w^\top x \ : \ x \in P \cap \{0,1\}^n\} \subseteq \operatorname{argmax}\{w^\top x \ : \ X\}.$$

Suppose we already have a point $\bar{x} \in X$ at hand in this situation. If we can show that $\max\{w^\top x \ : \ x \in K\} < w^\top \bar{x}$, it is not necessary to cut off component K in an adapted $\frac{1}{\lambda}$-relaxation of Corollary 18. Thus, it might be possible to find an adapted $\frac{1}{\lambda}$-relaxation whose ρ-value is small although X may have exponentially large ρ-values for any $\frac{1}{\lambda}$-relaxation since the only inequalities with large coefficients are necessary to cut off K. Of course, in its generality this approach is only of theoretical interest. In certain applications where the structure of $\mathcal{K}_n^X$ is accessible, however, it might be of practical relevance.

If $\mathcal{K}_n^X$ contains a single component only, Corollary 18 does not simplify the structure of binary formulations of X. If we solve $\max\{w^\top x \ : \ x \in X\}$ incrementally, however, we can use the previous idea to benefit from Corollary 18: If we know $\bar{x} \in X$, we can classify all binary points x with $w^\top x \le w^\top \bar{x}$ as feasible, i.e., we remove some part of the infeasible region. This might split some components in $\mathcal{K}_n^X$ into several subcomponents that may be easier to analyze than the original large components. Possible applications for this are precedence constrained knapsack problems [6, 17, 20].

Example 22. Consider the precedence constrained knapsack problem

$$\max\{x_1 + 2x_2 + x_3 + 2x_4 \ : \ 2x_1 + 2x_2 + 2x_3 + x_4 \le 5, \ x_1 \ge x_2, \ x_2 + x_3 \ge x_4, \ x \in \{0,1\}^4\}$$

whose feasible region is shown in Figure 3. In this figure, we assume that the vertices x of the inner cube have $x_4 = 0$ and the vertices of the outer cube $x_4 = 1$. The 0-vertex of the inner cube is the lower left corner in the front, and x_1-, x_2-, x_3-directions are to the right, to the back, and to the top, respectively. Note that the infeasible region consists of a single connected component.

Since x_1 and x_3 are not precedence constrained and fit into the knapsack, $(1, 0, 1, 0)^\top$ is a feasible solution with objective 2. Thus, we can declare all points in the infeasible region whose objective is at most 2 as feasible as we already know a feasible point with this objective. This splits the infeasible region into two connected components, see Figure 3.

3.3. MINIMUM SIZE OF BINARY FORMULATIONS WITH BOUNDED COEFFICIENTS

Another consequence of Corollary 18 is that we can determine the minimum size of a binary formulation or $\frac{1}{\lambda}$-relaxation with bounded coefficients, and thus, provide an answer to question (Q2). To keep terminology simple, we refer to binary formulations only, but all results directly apply to $\frac{1}{\lambda}$-relaxations (provided relaxations with the specified requirements exist). Let c be a positive integer. A binary formulation $Ax \le b$ of X with integral coefficients is called c-binary formulation if $|A_{ij}| \le c$ for every entry (i, j) of A. The smallest number of inequalities in a c-binary formulation of X is called c-binary formulation complexity, denoted $\operatorname{bfc}_c(X)$. Due to Corollary 18, we immediately have

$$\operatorname{bfc}_c(X) = \sum_{K \in \mathcal{K}_n^X} \operatorname{bfc}_c(K^c). \tag{14}$$

Thus, we can answer Question (Q2), provided we know $\operatorname{bfc}_c(K^c)$ for all $K \in \mathcal{K}_n^X$. In general, finding these numbers is quite complicated. Nevertheless, we can use (14) to provide lower bounds on $\operatorname{bfc}_c(X)$ by finding lower bounds on $\operatorname{bfc}_c(K^c)$.
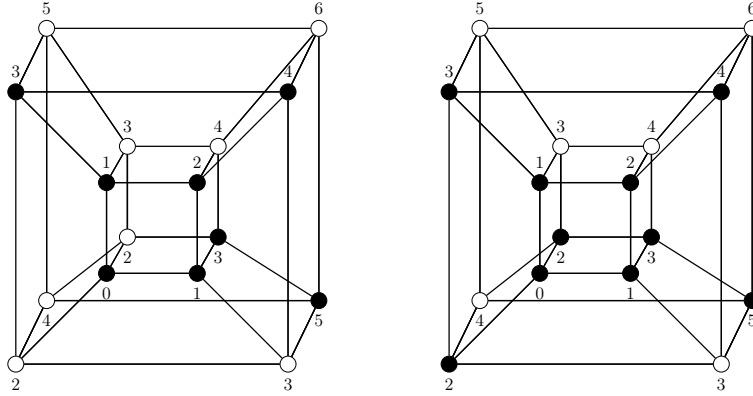
FIGURE 3. Illustration of the incremental optimization approach. Left: The feasible/infeasible region of the knapsack problem of Example 22. Right: The infeasible/feasible region if declaring infeasible points with objective at most 2 as feasible. The objectives of the solutions are given next to the corresponding points.

A very simple lower bound on $\mathrm{bfc}_c(K^c)$ is 1. Thus,

$$\mathrm{bfc}_c(X) \geq |\mathcal{K}_n^X|. \tag{15}$$

Note that for deriving (15), we did not take any assumptions on the coefficients into account. That is, (15) provides a lower bound on the size of any binary formulation of X. Besides this *combinatorial* bound, there is also a *geometric* lower bound on the size of integer formulations via hiding sets [16]. For $X \subseteq \mathbb{Z}^n$, a set $H \subseteq (\mathrm{aff}(X) \cap \mathbb{Z}^n) \setminus \mathrm{conv}(X)$ is called a *hiding set* if for any two distinct points $x, y \in H$ we have $\mathrm{conv}(\{x, y\}) \cap \mathrm{conv}(X) \neq \varnothing$. In [16] it is shown that every integer formulation of $X \subseteq \mathbb{Z}^n$ (if it exists) needs at least $|H|$ inequalities.

Remark 23. If $X \subseteq \{0, 1\}^n$ consists of all vectors with an odd/even number of 1-entries or $Y \subseteq \{0, 1\}^{2 \times n}$ consists of all matrices with two different rows, the hiding sets constructed in [16] are representative systems of $\mathcal{K}_X^n$ and $\mathcal{K}_Y^{2 \times n}$, respectively. Using these results, (15) shows that $\mathrm{bfc}_c(X) \geq 2^{n-1}$ and $\mathrm{bfc}_c(Y) \geq 2^n$. Thus, there exist sets whose binary relaxation complexity is exponential in n.

Note that the hiding set bound is applicable in a more general setting than the bound (15) because it allows X to be any integer point set. Moreover, the hiding set bound for a set $X \subseteq \{0, 1\}^n$ given by a maximum cardinality hiding set cannot be weaker than the bound (15) since $\mathrm{conv}(X)$ separates the points in different connected components of X^c within $[0, 1]^n$. Hence, for distinct $K_1, K_2 \in \mathcal{K}_n^X$ and any $x^1 \in K_1$ and $x^2 \in K_2$ we have $\mathrm{conv}(\{x^1, x^2\}) \cap \mathrm{conv}(X) \neq \varnothing$. Thus, any representative system of the components in $\mathcal{K}_n^X$ is a hiding set for X. However, (15) may be easier to use since finding large hiding sets may be hard. Moreover, if we find better lower bounds ν_K on $\mathrm{bfc}_c(K^c)$, the lower bound

$$\mathrm{bfc}_c(X) \geq \sum\nolimits_{K \in \mathcal{K}_n^X} \nu_K \tag{16}$$

might be strictly stronger than the hiding set bound because it takes further properties into account.

Remark 24. The above argumentation for deriving (16) remains valid if we replace bounded coefficients by other properties of sets of inequalities that are invariant under taking sub- or supersets and that are relevant in practice, e.g., sparsity [8, 9] or bounded ρ-value. Since we decompose or combine different binary formulations in the above discussion, polynomial time separability is not a suitable property, because the class of polynomially separable valid inequalities may not be closed under taking sub- or supersets, cf. [12].

The answer to Question (Q2) provided by (14) is rather abstract. For the important case $c = 1$, i.e., the case of 1-binary formulations, we discuss a more concrete answer in the remainder of this section, provided the sets $K \in \mathcal{K}_n^X$ define (generalized) independence systems. The case $c = 1$ is of particular interest because ternary inequalities are numerically very stable and such inequalities have a simple combinatorial interpretation (after flipping variables, they are essentially cardinality constraints).

Let F be a finite set and let $\mathcal{I} \subseteq 2^F$. The pair $(F, \mathcal{I})$ is called *independence system* if $\varnothing \in \mathcal{I}$ and $I' \in \mathcal{I}$ for every $I' \subseteq I \in \mathcal{I}$. The sets in $\mathcal{I}$ are called *independent*, sets in $2^F \setminus \mathcal{I}$ *dependent*. The minimally dependent sets are called *circuits* of $\mathcal{I}$. The *rank* of a set $I \in \mathcal{I}$ is $r(I) := \max\{|I'| : I' \subseteq I, I' \in \mathcal{I}\}$. Every independence system $(F, \mathcal{I})$ can be encoded as the set consisting of the incidence vectors $\chi(I) \in \mathbb{R}^F$ of sets $I \in \mathcal{I}$, that is, $X_\mathcal{I}^F := \{\chi(I) \in \{0,1\}^F : I \in \mathcal{I}\}$. By abuse of naming, we call $X \subseteq \{0,1\}^n$ an independence system if there exists an independence system $([n], \mathcal{I})$ such that $X = X_\mathcal{I}^{[n]}$. A set X is called *generalized independence system* if flipping some variables $x_i \mapsto 1 - x_i$ turns X into an independence system.

A binary formulation of $X_\mathcal{I}^F$ is given via *rank inequalities* $\sum_{i \in I} x_i \leq r(I)$, $I \subseteq F$. Note that the rank formulation has the same size as the formulation via infeasibility cuts in Example 2. However, the rank formulation might be stronger, since rank inequalities can cut off more than one infeasible binary vector. Moreover, it is not necessary to consider all rank inequalities. Below, we will see that a circuit C with $|C| = k + 1$ (call it a $(k+1)$-circuit) can only be cut off by a rank inequality of a set I with $r(I) = k$. Thus, if $\mathcal{R}_k$ is the set of all sets of rank k, it suffices to find a selection $\mathcal{R}_k' \subseteq \mathcal{R}_k$ that covers all $(k+1)$-circuits, i.e., for every $(k+1)$-circuit C there exists $I \in \mathcal{R}_k'$ with $C \subseteq I$, to obtain a binary formulation of $X_\mathcal{I}^F$. In fact, this result can be used to find a minimum size 1-binary formulation of an independence system.

Theorem 25. *Let $\mathcal{S} = (F, \mathcal{I})$ be an independence system and let ν_k, $k \in [|F|]$, be the minimum number of sets $I \subseteq F$ of rank k that are necessary to cover all $(k+1)$-circuits of $\mathcal{S}$. Then $\mathrm{bfc}_1(X_\mathcal{I}^F) = \sum_{k=1}^{|F|} \nu_k$.*

To prove Theorem 25, we need the following results.

Lemma 26. *Let $\mathcal{S} = (F, \mathcal{I})$ be an independence system, let $I \subseteq F$ be a circuit of $\mathcal{S}$, and let $k = |I|$. Then every rank inequality that separates $X_\mathcal{I}^F$ and $\chi(I)$ has right-hand side k.*

Proof. Assume there exists a set J of rank $\ell \neq k$ such that the corresponding rank inequality is violated by $\bar{x} := \chi(I)$. Then $\sum_{i \in J} \bar{x}_i > \ell - 1$. On the one hand, if $\ell < k$, this means that there exists a proper subset of I of cardinality ℓ which is dependent, contradicting the minimality of I. On the other hand, if $\ell > k$, we have

$$\sum_{i \in J \cap I} 1 = \sum_{i \in J} \bar{x}_i > \ell - 1 \geq k = \sum_{i \in I} \bar{x}_i,$$

which is impossible because $J \cap I \subseteq I$. $\qquad\square$

Lemma 27. *Let $\mathcal{S} = (F, \mathcal{I})$ be an independence system and let $I \in 2^F \setminus \mathcal{I}$. If $I' \subseteq I$ is a circuit of $\mathcal{S}$, then every $|I'|$-rank inequality that cuts off $\chi(I')$ also cuts off $\chi(I)$.*

Proof. Follows by non-negativity of the coefficients in rank inequalities. $\qquad\square$

Hitherto, we only used inequalities with non-negative coefficients to define a binary formulation of $X_\mathcal{I}^F$ and one may wonder whether it is possible to reduce the number of inequalities if we also allow (-1)-coefficients. In fact, the next lemma shows that we can restrict ourselves to non-negative coefficients. To show this, we define for a given

vector $a \in \{0, \pm 1\}^n$, the vector

$$\bar{a}_i = \begin{cases} 1, & \text{if } a_i = 1, \\ 0, & \text{otherwise.} \end{cases}$$

Lemma 28. *Let $\mathcal{S} = (F, \mathcal{I})$ be an independence system and let $a^\top x \leq \beta$, $a \in \{0, \pm 1\}^n$, be a valid inequality for $X_{\mathcal{I}}^F$ that cuts off at least one binary point. If a contains at least one negative entry, $\bar{a}^\top x \leq \beta$ is a valid inequality for $X_{\mathcal{I}}^F$ and it dominates $a^\top x \leq \beta$ on $\{0,1\}^n$, i.e., if $a^\top x \leq \beta$ cuts off a binary point, then also $\bar{a}^\top x \leq \beta$ does.*

Proof. Let $\bar{x} \in \{0,1\}^F \setminus X_{\mathcal{I}}^F$ that is cut off by $a^\top x \leq \beta$. Then

$$\beta < a^\top \bar{x} = \sum_{j:\, a_j=1} \bar{x}_j - \sum_{j:\, a_j=-1} \bar{x}_j \overset{\bar{x} \geq 0}{\leq} \sum_{j:\, a_j=1} \bar{x}_j = \bar{a}^\top \bar{x}.$$

Consequently, $\bar{x}$ is also cut off by $\bar{a}^\top \bar{x}$.

Thus, it remains to prove validity of the new inequality. Let $I_+ := \{i \in F : a_i = 1\}$ and $I_- := \{i \in F : a_i = -1\}$. Then

$$a^\top x = \sum_{i \in I_+} x_i - \sum_{i \in I_-} x_i \leq \beta \qquad \text{and} \qquad \bar{a}^\top x = \sum_{i \in I_+} x_i \leq \beta.$$

If the latter was not valid, there would exist $\bar{x} \in X_{\mathcal{I}}^F \cap \{0,1\}^F$ such that $\sum_{i \in I_+} \bar{x}_i > \beta$. Further, by monotonicity of $\mathcal{S}$, we can assume w.l.o.g. that $\bar{x}_i = 0$ for all $i \in I_-$. But then $\bar{x}$ is also cut off by $a^\top x \leq \beta$, contradicting the validity of $a^\top x \leq \beta$. Hence, the newly derived inequality has to be valid. $\qquad\square$

Now, we can proceed with proving Theorem 25.

Proof of Theorem 25. Let $Ax \leq b$ be a 1-binary formulation of $X_{\mathcal{I}}^F$ of minimum size. Then every inequality in $Ax \leq b$ cuts off at least one point in X^c. By Lemma 28 we can assume that all coefficients in A are 0 or 1. Moreover, we may assume w.l.o.g. that b is integral, because otherwise, we can strengthen the inequalities by rounding down fractional entries in b. Hence, for every inequality $a^\top x \leq \beta$ in $Ax \leq b$ the support of a has to be a set of rank β and $a^\top x \leq \beta$ is the corresponding rank inequality.

If the assertion was false, there would exist $k \in [\![|F|]\!]$ such that $Ax \leq b$ contains less than ν_k k-rank inequalities. By the definition of ν_k and Lemma 26, however, this means that there exists a circuit of size k that cannot be cut off by any inequality in $Ax \leq b$, contradicting that $Ax \leq b$ is a binary formulation of $X_{\mathcal{I}}^F$. Consequently,

$$\mathrm{bfc}_1(X_{\mathcal{I}}^F) \geq \sum_{k=1}^{|F|} \nu_k.$$

Thus, it suffices to provide a 1-binary formulation of $X_{\mathcal{I}}^F$ of the specified size to prove the assertion. To this end, let $\mathcal{R}_k' \subseteq \mathcal{R}_k$ cover all circuits of size k. Then the rank inequalities for $R \in \mathcal{R}_k'$ cut off all circuits of $X_{\mathcal{I}}^F$ of size $k+1$. Since every dependent set contains at least one circuit, Lemma 27 implies that all dependent sets are cut off by at least one hyperclique inequality for hypercliques in $\mathcal{R}_k'$, $k \in [\![|F|]\!]$. Thus, the assertion follows by choosing the sets $\mathcal{R}_k'$ minimally. $\qquad\square$

Example 29. Consider the set X of incidence vectors of stable sets of an undirected graph $G = (V, E)$. A classical binary formulation is given via edge constraints:

$$X = \{x \in \{0,1\}^V : x_u + x_v \leq 1, \{u, v\} \in E\}.$$

Since the right-hand sides of inequalities in this 1-binary formulation are all 1, every circuit of X is an edge of G. Thus, Theorem 25 shows that every 1-binary formulation of X of

minimum size consists of 1-rank inequalities only. Since sets of rank 1 are exactly the cliques in G, the minimum size of a 1-binary formulation is the smallest number of cliques that are necessary to cover all edges (circuits) of G. Computing this number is NP-hard, see [10, Problem GT17].

For knapsack sets $X = \{x \in \{0,1\}^n : a^\top x \le \beta\}$, the size of general binary formulations and 1-binary formulations differs probably by the largest amount: $\mathrm{bfc}(X) = 1$, while $\mathrm{bfc}_1(X)$ can be exponential in n. A more concrete characterization of $\mathrm{bfc}_1(X)$ than the one in Theorem 25 is based on the number of extensions of certain circuits [5, 23] (so-called strong covers). For example, the orbisack in Example 15 has 2^{n-2} strong covers, which can be seen by using the characterization of minimal covers in [13]. In general, however, computing the number of strong covers might be complicated. To estimate this quantity, a lower bound on $\mathrm{bfc}_1(X)$ based on subsets of strong covers is provided in [14].

4. Conclusions and Outlook

Concerning (Q1), we derived a lower bound on the ρ-value of $\frac{1}{\lambda}$-relaxations based on properties of facet defining inequalities of $\lambda \operatorname{conv}(X)$. In particular, we have seen that any $\frac{1}{2}$-relaxation may have an exponential ρ-value, while binary sets always admit 1-binary formulations. This emphasizes the special role of $\frac{1}{1}$-relaxations of binary sets w.r.t. their coefficients. The requirements for proving our lower bound for integer formulations, however, are very restrictive. The question arises whether other bounds with less restrictive prerequisites can be found, and in particular, whether there exists a bound that is independent from facet defining inequalities. For example, as mentioned above, it is intuitive that sets $X \subseteq \mathbb{Z}^n$ may need inequalities with strongly deviating coefficients if adjacent vertices of $\operatorname{conv}(X)$ differ much in one component but less in another. Thus, one may wonder whether partial knowledge on edges (or higher dimensional faces) of $\operatorname{conv}(X)$ suffices to derive bounds on the coefficients in integer formulations.

Moreover, even if the prerequisites of Theorem 13 are fulfilled, the obtained bounds may be weak if there are a lot of inequalities in $A_E x \le b_E$. Consequently, refined methods have to be developed to be able to derive good bounds also in this case. This, however, is out of scope of this paper and is left for future research.

To answer (Q2), we provided a combinatorial characterization of $\mathrm{bfc}_c(X)$. If all components in $\mathcal{K}_X^n$ are (generalized) independence systems, we can exactly characterize $\mathrm{bfc}_1(X)$. If one component is not a generalized independence system, however, a combinatorial characterization of $\mathrm{bfc}_1(X)$ is open and it would be interesting to investigate whether such a characterization exists in general. Moreover, an open question is whether the concept of strong covers for knapsacks can be generalized to arbitrary independence systems to provide a combinatorial characterization of ν_K in Theorem 25. Further, it is interesting to analyze the impact of other practically relevant properties on the size of binary formulations, e.g., sparsity or different bounds on the coefficients.

References

[1] A. Alexander, S. Boyd, and P. Elliot-Magwood. On the integrality gap of the 2-edge connected subgraph problem. Technical report, University of Ottawa, Ottawa, Canada, 2006.

[2] I. Aliev, M. Henk, and T. Oertel. Integrality gaps of integer knapsack problems. In F. Eisenbrand and J. Koenemann, editors, *Integer Programming and Combinatorial Optimization*, pages 25–38, Cham, 2017. Springer International Publishing.

[3] G. Angulo, S. Ahmed, S. S. Dey, and V. Kaibel. Forbidden vertices. *Mathematics of Operations Research*, 40(2):350–360, 2015.

[4] S. Arya, G. D. da Fonseca, and D. M. Mount. On the combinatorial complexity of approximating polytopes. *Discrete & Computational Geometry*, 58(4):849–870, 2017.

[5] E. Balas. Facets of the knapsack polytope. *Mathematical Programming*, 8(1):146–164, 1975.

[6] E. A. Boyd. Polyhedral results for the precedence-constrained knapsack problem. *Discrete Applied Mathematics*, 41:185–201, 1993.

[7] S. Boyd and G. Labonté. Finding the exact integrality gap for small traveling salesman problems. In W. Cook and A. Schulz, editors, *Integer Programming and Combinatorial Optimization*, volume 2337 of *Lecture Notes in Computer Science*, pages 83–92. Springer Berlin Heidelberg, 2002.

[8] S. S. Dey, M. Molinaro, and Q. Wang. Approximating polyhedra with sparse inequalities. *Mathematical Programming*, 154(1):329–352, 2015.

[9] S. S. Dey, M. Molinaro, and Q. Wang. Analysis of sparse cutting planes for sparse MILPs with applications to stochastic MILPs. *Mathematics of Operations Research*, 2017.

[10] M. R. Garey and D. S. Johnson. *Computers and Intractability: A Guide to the Theory of NP-Completeness*. W. H. Freeman & Co., 1979.

[11] F. Glover and E. Woolsey. Converting the 0-1 polynomial programming problem to a 0-1 linear program. *Operations Research*, 22(1):180–182, 1974.

[12] M. Grötschel, L. Lovász, and A. Schrijver. *Geometric algorithms and combinatorial optimization*. Springer, 1993.

[13] C. Hojny. *Symmetries in Binary Programs—A Polyhedral Perspective*. PhD thesis, TU Darmstadt, 2018.

[14] C. Hojny, H. Lüthen, and M. E. Pfetsch. On the size of integer programs with bounded coefficients and sparse constraints. Technical report, TU Darmstadt, 2017. http://www.optimization-online.org/DB_HTML/2017/06/6056.html.

[15] V. Kaibel and A. Loos. Finding descriptions of polytopes via extended formulations and liftings. In A. R. Mahjoub, editor, *Progress in Combinatorial Optimization*. Wiley, 2011.

[16] V. Kaibel and S. Weltge. Lower bounds on the sizes of integer programs without additional variables. *Mathematical Programming*, 154(1):407–425, 2015.

[17] H. Kellerer, U. Pferschy, and D. Pisinger. *Knapsack Problems*. Springer, Berlin Heidelberg, 2004.

[18] A. Loos. *Describing Orbitopes by Linear Inequalities and Projection Based Tools*. PhD thesis, Otto-von-Guericke-Universität Magdeburg, 2010.

[19] O. Marcotte. The cutting stock problem and integer rounding. *Mathematical Programming*, 33(1):82–92, 1985.

[20] K. Park and S. Park. Lifting cover inequalities for the precedence-constrained knapsack problem. *Discrete Applied Mathematics*, 72(3):219–241, 1997.

[21] Y. Pochet and L. A. Wolsey. Integer knapsack and flow covers with divisible coefficients: polyhedra, optimization and separation. *Discrete Applied Mathematics*, 59(1):57–74, 1995.

[22] A. Schrijver. *Theory of Linear and Integer Programming*. Wiley, 1987.

[23] L. A. Wolsey. Faces for a linear inequality in 0–1 variables. *Mathematical Programming*, 8(1):165–178, 1975.

RESEARCH GROUP DISCRETE OPTIMIZATION, DEPARTMENT OF MATHEMATICS, TECHNISCHE UNIVERSITÄT DARMSTADT, GERMANY

E-mail address: `hojny@opt.tu-darmstadt.de`