Lower Bounds on the Sizes of Integer Programs Without Additional Variables

Volker Kaibel and Stefan Weltge*

Otto-von-Guericke-Universität Magdeburg, Germany
{kaibel,weltge}@ovgu.de

Abstract. For a given set $X \subseteq \mathbb{Z}^d$ of integer points, we investigate the smallest number of facets of any polyhedron whose set of integer points is $\operatorname{conv}(X) \cap \mathbb{Z}^d$. This quantity, which we call the *relaxation complexity* of X, corresponds to the smallest number of linear inequalities of any integer program having X as the set of feasible solutions that gets along without auxiliary variables. We show that the use of auxiliary variables is essential for constructing polynomial size integer programming formulations in many relevant cases. In particular, we provide asymptotically tight exponential lower bounds on the relaxation complexity of the integer points of several well-known combinatorial polytopes, including the traveling salesman polytope and the spanning tree polytope.

Keywords: integer programming, relaxations, auxiliary variables, tsp

1 Introduction

Let $K_n = (V_n, E_n)$ be the undirected complete graph on n nodes and STSP_n the set of characteristic vectors of hamiltonian cycles in K_n. In order to solve the traveling salesman problem, there are numerous integer programs of the form

$$\max \left\{ \langle c, x \rangle : Ax + By \le b,\, x \in \mathbb{Z}^{E_n},\, y \in \mathbb{Z}^m \right\} \tag{1}$$

such that the optimal value of (1) is equal to $\max \{ \langle c, x \rangle : x \in \mathrm{STSP}_n \}$ for all edge weights $c \in \mathbb{R}^{E_n}$. In most of these formulations, the system $Ax + By \le b$ consists of polynomially (in n) many linear inequalities, see, e.g. [8] or [3]. Further, some of them even get along without integrality constraints on the auxiliary variables y. In contrast, a recent result on *extended formulations* (i.e., representations of polytopes as projections of other ones) due to Fiorini et al. [2] states that if one drops the integrality constraints on both x and y, then such systems must have exponentially many inequalities.

Interestingly, at a closer look, one notices that all such IP-formulations that consist of polynomially many inequalities make use of auxiliary variables. The main motivation for this paper was the question wether there is an IP-formulation

* Partially funded by the German Research Foundation (DFG): "Extended Formulations in Combinatorial Optimization" (KA 1616/4-1)

of type (1) for solving the traveling salesman problem that gets along without auxiliary variables but also consists of only polynomially many inequalities.

For a set $X \subseteq \mathbb{Z}^d$, let us call a polyhedron $R \subseteq \mathbb{R}^d$ a *relaxation* for X if $R \cap \mathbb{Z}^d = \operatorname{conv}(X) \cap \mathbb{Z}^d$. Further, the smallest number of facets of any relaxation for X will be called the *relaxation complexity* of X, or short: $\operatorname{rc}(X)$. With this notation, the above question is equivalent to the question wether $\operatorname{rc}(\mathrm{STSP}_n)$ is polynomial in n.

For most sets $X \in \{0,1\}^d$ that are associated with known combinatorial optimization problems it turns out that there are polynomial size IP-formulations of type (1). Following Schrijver's proof [12, Thm. 18.1] of the fact that integer programming is NP-hard one finds that for any language $\mathcal{L} \subseteq \{0,1\}^*$ that is in NP, there is a polynomial p such that for any $k > 0$ there is a system $Ax + By \leq b$ of at most $p(k)$ linear inequalities and $m \leq p(k)$ auxiliary variables with

$$\left\{x \in \{0,1\}^k : x \in \mathcal{L}\right\} = \left\{x \in \{0,1\}^k : \exists\, y \in \{0,1\}^m \ Ax + By \leq b\right\}. \quad (2)$$

Further, note that for many sets $X \subseteq \{0,1\}^d$ of feasible points of famous problems like SAT, CLIQUE, CUT or MATCHING, we even do not need auxiliary variables in order to give a polynomial size description as in (2), i.e., $\operatorname{rc}(X)$ is polynomially bounded for such sets X. However, as we will show in this paper, it turns out that this is not true for some other well-known combinatorial problems, including variants of TSP, SPANNING TREE or T-JOIN.

Our paper consists of two main sections: In Section 2, we discuss basic properties of the number of facets of relaxations of a general set $X \subseteq \mathbb{Z}^d$. This includes questions of wether irrational coordinates may help or wether it is a good idea to only use facet-defining inequalities of $\operatorname{conv}(X)$ in order to construct small relaxations. Further, we introduce the concept of *hiding sets*, which turns out to be a powerful technique to provide lower bounds on $\operatorname{rc}(X)$. In Section 3, we then give exponential lower bounds on the sizes of relaxations for concrete structures that occur in many practical IP-formulations. In particular, coming back to our motivating question, we show that the asymptotic growth of $\operatorname{rc}(\mathrm{STSP}_n)$ is indeed exponential in n. This shows that, for many problems, the benefit of projection, i.e. the use of auxiliary variables, is essential when constructing polynomial size IP-formulations.

Except for a paper of Jeroslow [5], the authors are not aware of any reference that deals with a quantity that is similar to our notion of relaxation complexity in a general context. In his paper, for a set $X \subseteq \{0,1\}^d$ of binary vectors, Jeroslow introduces the term *index* of X (short: $\operatorname{ind}(X)$), which is defined as the smallest number of inequalities needed to separate X from the remaining points in $\{0,1\}^d$. As the main result, he shows that 2^{d-1} is an upper bound on $\operatorname{ind}(X)$, which is attained by the set of binary vectors of length d that contain an even number of ones. In Sections 2 and 3.3 we shall come back to this result. Further, his idea of bounding the index of a set $X \subseteq \{0,1\}^d$ from below, is related to our approach of providing lower bounds on the relaxation complexity of general X via hiding sets in Section 2.4.

2 Basic observations

There are sets $X \subseteq \mathbb{Z}^d$ that do not admit any relaxation. Therefore, let us call a set $X \subseteq \mathbb{Z}^d$ to be *polyhedral* if its convex hull is a polyhedron. By definition, we have that $\mathrm{rc}(X)$ is finite for such sets. Further, in this setting, it is easy to see that any relaxation corresponds to a valid IP-formulation and vice versa:

Proposition 1. *Let $X \subseteq \mathbb{Z}^d$ be polyhedral and $P \subseteq \mathbb{R}^d$ a polyhedron. Then, P is a relaxation for X if and only if* $\sup \left\{ \langle c, x \rangle : x \in P \cap \mathbb{Z}^d \right\} = \sup \left\{ \langle c, x \rangle : x \in X \right\}$ *holds for all $c \in \mathbb{R}^d$.*

Clearly, any finite set of integer points is polyhedral. For a set $X \subseteq \{0,1\}^d$ of binary vectors, a polyhedron P is a relaxation for X if and only if $P \cap \mathbb{Z}^d = X$.

As mentioned in the introduction, Jeroslow [5] showed that for any set $X \subseteq \{0,1\}^d$, one needs at most 2^{d-1} many linear inequalities in order to separate X from $\{0,1\}^d \setminus X$. If now $P \subseteq \mathbb{R}^d$ is a polyhedron such that $P \cap \{0,1\}^d = X$, then, in order to construct a relaxation for X, we need to additionally separate all points $\mathbb{Z}^d \setminus \{0,1\}^d$ from X. This can be done by intersecting P with a relaxation for $\{0,1\}^d$. We conclude:

Proposition 2. *Let $X \subseteq \{0,1\}^d$. Then $\mathrm{rc}(X) \leq 2^{d-1} + \mathrm{rc}(\{0,1\}^d)$.*

2.1 Relaxation complexity of the cube

Motivated by Proposition 2, we are interested in the relaxation complexity of $\{0,1\}^d$. Since $[0,1]^d = \mathrm{conv}(\{0,1\}^d)$, we obviously have that $\mathrm{rc}(\{0,1\}^d) \leq 2d$. However, it turns out that we can construct a relaxation of only $d+1$ facets:

Lemma 1. *For $d \geq 1$, we have*

$$\{0,1\}^d = \left\{ x \in \mathbb{Z}^d : x_k \leq 1 + \sum_{i=k+1}^{d} 2^{-i} x_i \ \forall \, k \in [d], \, x_1 + \sum_{i=2}^{d} 2^{-i} x_i \geq 0 \right\}.$$

Proof. See appendix.

To show that this construction is best possible, note that if a polyhedron that contains $\{0,1\}^d$ (and hence is d-dimensional) has less than $d+1$ facets, it must be unbounded. In order to show that such a polyhedron must contain infinitely many integer points (and hence cannot be a relaxation of $\{0,1\}^d$), we make use of Minkowski's theorem:

Theorem 1 (Minkowski [9]). *Any convex set which is symmetric with respect to the origin and with volume greater than 2^d contains a non-zero integer point.*

For $\varepsilon > 0$ let $B_\varepsilon := \left\{ x \in \mathbb{R}^d : \|x\| < \varepsilon \right\}$ be the open ball with radius ε. As a direct consequence of Minkowski's theorem, the following corollary is useful for our argumentation.

Corollary 1. *Let $c \subseteq \mathbb{R}^d \setminus \{\mathbb{O}_d\}$, $\lambda_0 \in \mathbb{R}$ and $\varepsilon > 0$. Then*

$$L(c, \lambda_0, \varepsilon) := \left\{\lambda c \in \mathbb{R}^d : \lambda \geq \lambda_0\right\} + B_\varepsilon$$

contains infinitely many integer points.

Proof. See appendix.

Theorem 2. *For $d \geq 1$, we have that* $\mathrm{rc}(\{0,1\}^d) = d + 1$.

Proof. By Lemma 1, we already know that $\mathrm{rc}(\{0,1\}^d) \leq d+1$. Suppose there is a relaxation $R \subseteq \mathbb{R}^d$ for $\{0,1\}^d$ with less than $d+1$ facets. As mentioned above, since $\dim(R) \geq \dim(\{0,1\}^d) = d$, R has to be unbounded.

By induction over $d \geq 1$, we will show that any unbounded polyhedron $R \subseteq \mathbb{R}^d$ with $\{0,1\}^d \subseteq R$ contains infinitely many integer points. Hence, it cannot be a relaxation of $\{0,1\}^d$. Clearly, our claim is true for $d = 1$. For $d \geq 1$, let $c \in \mathbb{R}^d \setminus \{\mathbb{O}_d\}$ be a direction such that $x + \lambda c \in R$ for any $x \in R$ and $\lambda \geq 0$. Since $\{0,1\}^d$ is invariant under affine maps that map a subset of coordinates x_i to $1 - x_i$, we may assume that $c \geq \mathbb{O}_d$.

If $c > \mathbb{O}_d$, then there is some $\lambda_0 > 0$ such that $\lambda_0 c \in \mathrm{int}([0,1]^d)$. Thus, there is some $\varepsilon > 0$ such that $\lambda_0 c + B_\varepsilon \subseteq [0,1]^d \subseteq R$. By the definition of c and ε, we thus obtained that $L(c, \lambda_0, \varepsilon) \subseteq R$. By Corollary 1, it follows that $L(c, \lambda_0)$ contains infinitely many integer points and so does R.

Otherwise, we may assume that $c_d = 0$. Let $\mathcal{H}_d := \left\{x \in \mathbb{R}^d : x_d = 0\right\}$ and $p\colon \mathcal{H}_d \to \mathbb{R}^{d-1}$ be the projection onto the first $d-1$ coordinates. Then, the polyhedron $\tilde{R} = p(R)$ is still unbounded and contains $\{0,1\}^{d-1} = p(\{0,1\}^d)$. By induction, $\tilde{R}$ contains infinitely many integer points and so does R. ☐

With Proposition 2 we thus obtain:

Corollary 2. *Let $X \subseteq \{0,1\}^d$. Then* $\mathrm{rc}(X) \leq 2^{d-1} + d + 1$.

2.2 Limits of facet-defining inequalities

Many known relaxations for sets $X \subseteq \mathbb{Z}^d$ that are identified with feasible points in combinatorial problems are defined by linear inequalities of which, preferably, most of them are facet-defining for $\mathrm{conv}(X)$. Clearly, this has important practical reasons since such formulations are tightest possible in some sense. However, if one is interested in a relaxation that has as few number of facets as possible, it is a severe restriction to only use facet-defining inequalities of $\mathrm{conv}(X)$: In the previous section we have seen that $\mathrm{rc}(\{0,1\}^d) = d+1$ whereas by removing any of the cube's inequalities the remaining polyhedron gets unbounded. Nevertheless, the restriction turns out to be not too hard:

Theorem 3. *Let $X \subseteq \mathbb{Z}^d$ be polyhedral and $\mathrm{rc}_F(X)$ the smallest number of facets of any relaxation for X whose facet-defining inequalities are also facet-defining for* $\mathrm{conv}(X)$*. Then,* $\mathrm{rc}_F(X) \leq \dim(X) \cdot \mathrm{rc}(X)$.

Proof. By Carathéodory's Theorem (in the affine hull of X), any facet-defining inequality of a relaxation R for X can be replaced by $\dim(X)$ many facet-defining inequalities of $\mathrm{conv}(X)$. The resulting polyhedron is still a relaxation for X. ☐

2.3 Irrationality

Another question one might ask is wether it may help (in order to construct a relaxation having few facets) to use irrational coordinates in the description of a relaxation. Again, it turns out that one does not loose too much when restricting to rational relaxations only:

Theorem 4. *Let $X \subseteq \mathbb{Z}^d$ be polyhedral with $|X| < \infty$ and $\mathrm{rc}_{\mathbb{Q}}(X)$ the smallest number of facets of any rational relaxation for X. Then, $\mathrm{rc}_{\mathbb{Q}}(X) \leq \mathrm{rc}(X) + \dim(X) + 1$.*

Proof. Since X is finite, there exists a rational simplex $\Delta \subseteq \mathbb{R}^d$ of dimension $\dim(X)$ such that $X \subseteq \Delta$. Let R be any relaxation of X having f facets and set $B := (\mathbb{Z}^d \setminus X) \cap \Delta$. Since $B \cap R = \emptyset$ and $|B| < \infty$, we are able to slightly perturb the facet-defining inequalities of R in order to obtain a polyhedron $\tilde{R}$ such that $B \cap \tilde{R} = \emptyset$ and $\tilde{R}$ is rational. Now $\tilde{R} \cap \Delta$ is still a relaxation for X, which is rational and has at most $f + (\dim(\Delta) + 1) = f + \dim(X) + 1$ facets. □

However, we are not aware of any polyhedral set X where $\mathrm{rc}(X) < \mathrm{rc}_{\mathbb{Q}}(X)$. In fact, we even do not know if $\mathrm{rc}(\Delta_d) < d + 1 = \mathrm{rc}_{\mathbb{Q}}(\Delta_d)$ holds, where $\Delta_d := \{\mathbb{O}_d, \mathbb{e}_1, \ldots, \mathbb{e}_d\}$. Note that any rational relaxation of $\mathrm{rc}(\Delta_d)$ has to be bounded and thus, it has at least $d+1$ facets. (Otherwise it would contain a rational ray and hence infinitely many integer points.)

2.4 Hiding sets

In this section, we introduce a simple framework to provide lower bounds on the relaxation complexity for polyhedral sets $X \subseteq \mathbb{Z}^d$.

Definition 1. *Let $X \subseteq \mathbb{Z}^d$. A set $H \subseteq \mathrm{aff}(X) \cap \mathbb{Z}^d \setminus \mathrm{conv}(X)$ is called a* hiding set *for X if for any two distinct points $a, b \in H$ we have that $\mathrm{conv}\{a, b\} \cap \mathrm{conv}(X) \neq \emptyset$.*

Proposition 3. *Let $X \subseteq \mathbb{Z}^d$ be polyhedral and $H \subseteq \mathrm{aff}(X) \cap \mathbb{Z}^d \setminus X$ a hiding set for X. Then, $\mathrm{rc}(X) \geq |H|$.*

Proof. Let $R \subseteq \mathbb{R}^d$ be a relaxation for X. Since $H \subseteq \mathrm{aff}(X) \subseteq \mathrm{aff}(R)$, any point in H must be separated from X by a facet-defining inequality of R.

Suppose that a facet-defining inequality $\langle \alpha, x \rangle \leq \beta$ of R is violated by two distinct points $a, b \in H$. Since H is a hiding set, there exists a point $x \in \mathrm{conv}\{a, b\} \cap \mathrm{conv}(X)$. Clearly, x does also violate $\langle \alpha, x \rangle \leq \beta$ which is a contradiction since $\langle \alpha, x \rangle \leq \beta$ is valid for $R \supseteq \mathrm{conv}(X)$.

Thus, any facet-defining inequality of R is violated by at most one point in H. Hence, R has at least $|H|$ facets. □

3 Exponential lower bounds

In this section, we provide strong lower bounds on the relaxation complexities of some interesting sets X. By dividing these sets into three classes, we try to identify general structures that are hard to model in the context of relaxations.

3.1 Connectivity and Acyclicity

In many IP-formulations for practical applications, the feasible solutions are subgraphs that are required to be connected or acyclic. Quite often in these cases, there are polynomial size IP-formulations that use auxiliary variables. For instance, for the *spanning tree polytope* there are even polynomial size extended formulations [7] that can be adapted to also work for the *connector polytope* CONN_n (see below). In contrast, we give exponential lower bounds on the relaxation complexities of some important representatives of this structural class.

ATSP, STSP, GTSP A well-known relaxation for STSP_n is the so-called *subtour relaxation*

$$\Big\{x \in \mathbb{R}^{E_n} : \sum_{e \in E_n} x_e = n \quad x(\delta(S)) \geq 2 \quad \forall \emptyset \neq S \subsetneq V_n \quad x(\delta(v)) = 2 \quad \forall v \in V_n \quad x \geq \mathbb{O}_{E_n}\Big\}, \tag{3}$$

which has exponentially (in n) many facets (where $K_n = (V_n, E_n)$ is the complete graph on n nodes). We will show that this formulation is asymptotically smallest possible. In fact, we also give exponential lower bounds for two variants of STSP_n: In the *asymmetric* traveling salesman problem, one usually considers the directed version $\mathrm{ATSP}_n \subseteq \{0,1\}^{A_n}$, which is the set of characteristic vectors of directed hamiltonian cycles in the complete directed graph on n nodes whose arcs we will denote by A_n. Another popular variant is the *graphical* traveling salesman problem [1], in which one is allowed to visit a node multiple times. Therefore, the associated integer points are

$$\mathrm{GTSP}_n := \left\{\chi(T) \in \mathbb{Z}^{E_n} : T \text{ is a closed walk visiting every node}\right\},$$

where $\chi(T)$ is the characteristic vector of T and $\chi(T)_e = k$ if edge e is used k times by T.

Let $n = 4N + 2$ for some integer N and let us define the set

$$V := \{v_i, v'_i : i \in [N+1]\} \ \cup \ \{w_i, w'_i : i \in [N]\}$$

consisting of $4N+2$ distinct nodes. For a binary vector $b \in \{0,1\}^N$ let us further define the two node-disjoint directed cycles

$$C_b := \left\{(v_{N+1}, v_1)\right\} \ \cup \bigcup_{i : b_i = 0} \left\{(v_i, w_i), (w_i, v_{i+1})\right\} \ \cup \bigcup_{i : b_i = 1} \left\{(v_i, w'_i), (w'_i, v_{i+1})\right\}$$

$$C'_b := \left\{(v'_{N+1}, v'_1)\right\} \ \cup \bigcup_{i : b_i = 0} \left\{(v'_i, w'_i), (w'_i, v'_{i+1})\right\} \ \cup \bigcup_{i : b_i = 1} \left\{(v'_i, w_i), (w_i, v'_{i+1})\right\},$$

see Fig. 1 for an example. In this section, we will only consider graphs on these $4N + 2$ nodes. It is easy to transfer all following observations to graphs on n

nodes, where $n \not\equiv 2 \pmod 4$, by replacing arc (v_{N+1}, v_1) by a directed path including 1, 2 or 3 additional nodes. Let us now consider the set

$$H_N := \left\{ \chi(C_b \cup C'_b) : b \in \{0,1\}^N \right\}.$$

By identifying V with the nodes of the complete directed graph on $4N+2$ nodes, we clearly have that $H_N \cap \mathrm{ATSP}_{4N+2} = \emptyset$.

Lemma 2. *H_N is a hiding set for* ATSP_{4N+2}.

Proof. First, note that

$$H_N \subseteq \mathrm{aff}(\mathrm{ATSP}_{4N+2}) = \left\{ x \in \mathbb{R}^A : x(\delta^{\mathrm{in}}(v)) = x(\delta^{\mathrm{out}}(v)) = 1,\ \forall\, v \in V \right\},$$

where A is the set of arcs in the complete directed graph on $4N+2$ nodes. Let $b^{(1)}, b^{(2)} \in \{0,1\}^N$ be distinct. W.l.o.g. we may assume that there is an index $j \in [N]$ such that $b^{(1)}_j = 0$ and $b^{(2)}_j = 1$. Let us now consider the following slight modifications of $C_{b^{(1)}} \cup C'_{b^{(1)}}$ and $C_{b^{(2)}} \cup C'_{b^{(2)}}$:

$$T_1 := \left(C_{b^{(1)}} \cup C'_{b^{(1)}} \setminus \left\{(v_j, w_j), (v'_j, w'_j)\right\}\right) \ \cup\ \left\{(v_j, w'_j), (v'_j, w_j)\right\}$$
$$T_2 := \left(C_{b^{(2)}} \cup C'_{b^{(2)}} \setminus \left\{(v_j, w'_j), (v'_j, w_j)\right\}\right) \ \cup\ \left\{(v_j, w_j), (v'_j, w'_j)\right\}$$

We claim that both T_1 and T_2 are hamiltonian cycles: First note that $C_{b^{(1)}} \cup C'_{b^{(1)}} \setminus \left\{(v_j, w_j), (v'_j, w'_j)\right\}$ consists of two node-disjoint directed paths P_1 from w_j to v_j and P'_1 from w'_j to v'_j. Hence, $T_1 = P_1 \cup P'_1 \cup (\{v_j, w'_j), (v'_j, w_j)\}$ indeed forms a hamiltonian cycle. The claim for T_2 follows analogously. As an illustration, see Fig. 2.

By definition, we further have that

$$\begin{aligned}\chi(T_1) + \chi(T_2) &= \chi(C_{b^{(1)}} \cup C'_{b^{(1)}}) - \chi(\{(v_j, w_j), (v'_j, w'_j))\} + \chi(\{(v_j, w'_j), (v'_j, w_j)\}) \\ &\quad + \chi(C_{b^{(2)}} \cup C'_{b^{(2)}}) + \chi(\{(v_j, w_j), (v'_j, w'_j))\} - \chi(\{(v_j, w'_j), (v'_j, w_j)\}) \\ &= \chi(C_{b^{(1)}} \cup C'_{b^{(1)}}) + \chi(C_{b^{(2)}} \cup C'_{b^{(2)}})\end{aligned}$$

and hence,

$$\frac{1}{2}(\chi(C_{b^{(1)}} \cup C'_{b^{(1)}}) + \chi(C_{b^{(2)}} \cup C'_{b^{(2)}})) = \frac{1}{2}(\chi(T_1) + \chi(T_2)) \in \mathrm{conv}(\mathrm{ATSP}_{4N+2}). \ \square$$

Theorem 5. *The asymptotic growth of* $\mathrm{rc}(\mathrm{ATSP}_n)$, $\mathrm{rc}(\mathrm{STSP}_n)$ *and* $\mathrm{rc}(\mathrm{GTSP}_n)$ *is* $2^{\theta(n)}$.

Proof. Lemma 2 shows that H_N is a hiding set for ATSP_n. By replacing all directed arcs with their undirected versions, the set H_N yields a hiding set for STSP_n. By Proposition 3, we obtain a lower bound of $|H_N| = 2^{\Omega(n)}$ for $\mathrm{rc}(\mathrm{ATSP}_n)$ and $\mathrm{rc}(\mathrm{STSP}_n)$.

Since $\mathrm{STSP}_n = \mathrm{GTSP}_n \cap \mathrm{aff}(\mathrm{STSP}_n)$, we further obtain that $\mathrm{rc}(\mathrm{GTSP}_n) \geq \mathrm{rc}(\mathrm{STSP}_n)$.

The complete our argumentation, note that all of the sets ATSP_n, STSP_n and GTSP_n have relaxations of size $2^{\Theta(n)}$, which are variants of the formulation in (3). $\square$

Connected sets Let CONN_n be the set of all characteristic vectors of edge sets that form a connected subgraph in the complete graph on n nodes. The polytope $\left\{x \in [0,1]^{E_n} : x(\delta(S)) \geq 1 \; \forall \emptyset \neq S \subsetneq V_n\right\}$ is a relaxation for CONN_n. Thus, we have that $\mathrm{rc}(\mathrm{CONN}_n) \leq \mathcal{O}(2^n)$.

For a lower bound, consider again the undirected version of our set H_N. Since each point in H_N belongs to a node-disjoint union of two cycles, we have that $H_N \cap \mathrm{CONN}_n = \emptyset$. Further, we know that for any $a, b \in H_N$ we have that

$$\emptyset \neq \mathrm{conv}\{a,b\} \cap \mathrm{conv}(\mathrm{STSP}_n) \subseteq \mathrm{conv}\{a,b\} \cap \mathrm{conv}(\mathrm{CONN}_n)$$

and since $H_N \subseteq \mathrm{aff}(\mathrm{CONN}_n) = \mathbb{R}^{A_n}$, we see that H_N is also a hiding set for CONN_n. We obtain:

Corollary 3. *The asymptotic growth of* $\mathrm{rc}(\mathrm{CONN}_n)$ *is* $2^{\Theta(n)}$.

Branchings and Forests Besides connectivity, we show that, in general, it is also hard to force acyclicity in the context of relaxation. Let therefore ARB_n (SPT_n) be the set of characteristic vectors of arborescences (spanning trees) in the complete directed (undirected) graph.

Theorem 6. *The asymptotic growth of* $\mathrm{rc}(\mathrm{ARB}_n)$ *and* $\mathrm{rc}(\mathrm{SPT}_n)$ *is* $2^{\theta(n)}$.

Proof. First, note that both the *arborescence polytope* and the spanning tree polytope (i.e., $\mathrm{conv}(ARB_n)$ and $\mathrm{conv}(\mathrm{SPT}_n)$) have $\mathcal{O}(2^n)$ facets [11] and hence we have an upper bound of $\mathcal{O}(2^n)$ for both $\mathrm{rc}(\mathrm{ARB}_n)$ and $\mathrm{rc}(\mathrm{SPT}_n)$.

For a lower bound, let us modify the definition of $C(b)'$ by removing arc (v'_{N+1}, v'_1). Then, $C(b) \cap C(b)'$ is a node-disjoint union of a cycle and a path and hence not an arborescence. By following the proof of Lemma 2 and removing arc (v'_{N+1}, v'_1) from T_1 and T_2, we still have that

$$\chi(C(b^{(1)}) \cup C(b^{(1)})') + \chi(C(b^{(2)}) \cup C(b^{(2)})') = \chi(T_1) + \chi(T_2),$$

where T_1 and T_2 are spanning arborescences. (Actually, T_1 and T_2 are in fact directed paths visiting each node.) Since $\mathrm{aff}(\mathrm{ARB}_n) = \mathbb{R}^A$, we therefore obtain that the modified set H_N is a hiding set for ARB_n. By undirecting all arcs, H_N also yields a hiding set for SPT_n.

Again, by Proposition 3, we deduce a lower bound of $|H_N| = 2^{\Omega(n)}$ for both $\mathrm{rc}(\mathrm{ARB}_n)$ and $\mathrm{rc}(\mathrm{SPT}_n)$. □

Remark 1. Since in the proof of Theorem 6 T_1 and T_2 are rooted at node v'_1, the statements even hold if the sets ARB_n and SPT_n are restricted to characteristic vectors of arborescences/trees rooted at a fixed node.

Let BRANCH_n ($\mathrm{FORESTS}_n$) be the set of characteristic vectors of branchings (forests) in the complete directed (undirected) graph.

Corollary 4. *The asymptotic growth of* $\mathrm{rc}(\mathrm{BRANCH}_n)$ *and* $\mathrm{rc}(\mathrm{FORESTS}_n)$ *is* $2^{\theta(n)}$.

Proof. The claim follows from Theorem 6 and the facts that

$$\text{ARB}_n = \text{BRANCH}_n \cap \left\{ x \in \mathbb{R}^{A_n} : \sum_{a \in \mathbb{R}^{A_n}} x_a = n-1 \right\}$$

$$\text{SPT}_n = \text{FORESTS}_n \cap \left\{ x \in \mathbb{R}^{E_n} : \sum_{e \in \mathbb{R}^{E_n}} x_e = n-1 \right\}. \quad \square$$

3.2 Distinctness

Another common component of practical IP-formulations is the requirement of distinctness of a certain set of vectors or variables. Here, we consider two general cases in which we can also show that the benefit of auxiliary variables is essential.

Binary All-Different In the case of the *binary all-different* constraint, one requires the distinctness of rows of a binary matrix with m rows and n columns. The set of feasible points is therefore defined by

$$\text{DIFF}_{m,n} := \left\{ x \in \{0,1\}^{m \times n} : x \text{ has pairwise distinct rows} \right\}.$$

As an example, in [6] the authors present IP-formulations to solve the coloring problem in which they binary encode the color classes assigned to each node. As a consequence, certain sets of encoding vectors have to be distinct.

By separating each possible pair of equal rows by one inequality, it is further easy to give a relaxation for $\text{DIFF}_{m,n}$ that has at most $\binom{m}{2}2^n + 2mn$ facets. In the case of $m = 2$, for instance, this bound turns out to be almost tight:

Theorem 7. *For all $n \geq 1$, we have that $\text{rc}(\text{DIFF}_{2,n}) \geq 2^n$.*

Proof. Let us consider the set

$$H_{2,n} := \left\{ (x,x)^T \in \{0,1\}^{2 \times n} : x \in \{0,1\}^n \right\}.$$

For $x, y \in \{0,1\}^n$ distinct, we obviously have that

$$\frac{1}{2}\left((x,x)^T + (y,y)^T\right) = \frac{1}{2}\left((x,y)^T + (y,x)^T\right) \in \text{conv}(\text{DIFF}_{2,n}).$$

Since $H_{2,n} \cap \text{DIFF}_{2,n} = \emptyset$ and $H_{2,n} \subseteq \text{aff}(\text{DIFF}_{2,n}) = \mathbb{R}^{2 \times n}$, $H_{2,n}$ is a hiding set for $\text{DIFF}_{2,n}$ and by Proposition 3 we obtain that $\text{rc}(\text{DIFF}_{2,n}) \geq |H_{2,n}| = 2^n$. $\square$

Permutahedron As a case in which one does not require the distinctness of binary vectors but of a set of numbers let us consider the set

$$\text{PERM}_n := \{(\pi(1), \ldots, \pi(n)) \in \mathbb{Z}^n : \pi \in \mathcal{S}_n\},$$

which is the vertex set of the *permutahedron* conv(PERM$_n$). Rado [10] showed that the permutahedron can be described via

$$\begin{aligned}\mathrm{conv}(\mathrm{PERM}_n) = \{x \in \mathbb{R}^n : \sum_{i=1}^{n} x_i &= \frac{n(n+1)}{2} \\ \sum_{i \in S} x_i &\geq \frac{|S|(|S|+1)}{2} \quad \text{for all } \emptyset \neq S \subset [n] \\ x &\geq \mathbb{O}_n\} \end{aligned} \tag{4}$$

and hence has $\mathcal{O}(2^n)$ facets. Apart from that, it is a textbook example for a polytope having many different, very compact extended formulations, see, e.g., [4]. In the contrary, we show that the relaxation complexity of PERM$_n$ has exponential growth in n:

Theorem 8. *The asymptotic growth of* rc(PERM$_n$) *is* $2^{\theta(n)}$.

Proof. Let $m := \lfloor \frac{n}{2} \rfloor$. For any set $S \subseteq [n]$ with $|S| = m$ select an integer vector $x^S \in \mathbb{Z}^n$ with $\{x_i : i \in S\} = \{1, \ldots, m-1\}$ and $m-1$ occuring twice among the x_i^S $(i \in S)$ and $\{x_i : i \in [n] \setminus S\} = \{m+2, \ldots, n\}$ and $m+2$ occuring twice among the x_i^S $(i \in [n] \setminus S)$. Such a vector is not contained in conv(PERM$_n$) as

$$\sum_{i \in S} x_i^S = 1 + 2 + \ldots + (|S| - 2) + (|S| - 2) < \frac{|S|(|S|+1)}{2}$$

On the other hand, note that this is the only constraint from (4) that is violated by x^S. In particular, $x^S \in \mathrm{aff}(\mathrm{PERM}_n)$ holds.

Let $S_1, S_2 \subseteq [n]$ with $|S_1| = |S_2| = m$ be distinct. We will show that $x := \frac{1}{2} \cdot (x^{S_1} + x^{S_2}) \in \mathrm{conv}(\mathrm{PERM}_n)$ holds. Since x satisifies all constraints that are satisified by both x^{S_1} and x^{S_2}, it suffices to show that $\sum_{i \in T} x_i \geq \frac{|T|(|T|+1)}{2}$ holds for $T \in \{S_1, S_2\}$. W.l.o.g. we may assume that $T = S_1$ and obtain

$$\begin{aligned}\sum_{i \in S_1} x_i &= \frac{1}{2} \sum_{i \in S_1} x_i^{S_1} + \frac{1}{2} \sum_{i \in S_1} x_i^{S_2} \\ &= \frac{1}{2} \left(\frac{m(m+1)}{2} - 1 \right) + \frac{1}{2} \sum_{i \in S_1} x_i^{S_2} \\ &\geq \frac{1}{2} \left(\frac{m(m+1)}{2} - 1 \right) + \frac{1}{2} \left(\frac{m(m+1)}{2} + 2 \right) \\ &= \frac{m(m+1)}{2} + \frac{1}{2} \geq \frac{|T|(|T|+1)}{2}. \ \square \end{aligned}$$

Thus, the set $H := \{ x^S : S \subseteq [n], |S| = m \}$ is a hiding set for PERM$_n$. Our claim follows from Proposition 3 and the fact that $|H| = \binom{n}{\lfloor \frac{n}{2} \rfloor} = 2^{\theta(n)}$. □

3.3 Parity

The final structural class we consider deals with the restriction that the number of selected elements of a given set has a certain parity. Let us call a binary vector $a \in \{0,1\}^d$ *even* (*odd*) if the sum of its entries is even (odd). In [5] it is shown that the number of inequalities needed to separate

$$\mathrm{EVEN}_n := \{x \in \{0,1\}^n : x \text{ is even}\}$$

from all other points in $\{0,1\}^n$ is exactly 2^{n-1}. This is done by showing that

$$\mathrm{ODD}_n := \{x \in \{0,1\}^n : x \text{ is odd}\}$$

is a hiding set for EVEN_n (although the notion is different from ours). Hence, with Corollary 2, we obtain:

Theorem 9. *The asymptotic growth of* $\mathrm{rc}(\mathrm{EVEN}_n)$ *is* $\Theta(2^n)$.

***T*-joins** As a well-known representative of this structural class let us consider T-JOINS_n, which is, for given $T \subseteq V_n$, defined as the set of characteristic vectors of T-joins in the complete graph on n nodes. Let us recall that a T-join is a set $J \subseteq E_n$ of edges such that T is equal to the set of nodes of odd degree in the graph (V_n, J). Note, that if a T-join exists, then $|T|$ is even.

Theorem 10. *Let n be even and $T \subseteq V_n$ with $|T|$ even. Then,* $\mathrm{rc}(T\text{-}\mathrm{JOINS}_n) \geq 2^{\frac{n}{4}-1}$.

Proof. Since n is even and $|T|$ is even, we may partition V_n into pairwise disjoint sets T_1, T_2, U_1, U_2 with $T = T_1 \cup T_2$, $k = |T_1| = |T_2|$ and $\ell = |U_1| = |U_2|$. Let $M_1, \ldots, M_k$ be pairwise edge-disjoint matchings of cardinality k that connect nodes from T_1 with nodes from T_2. Analogously, let $N_1, \ldots, N_\ell$ be pairwise edge-disjoint matchings of cardinality ℓ that connect nodes from U_1 with nodes from U_2. For $b \in \{0,1\}^k$ and $c \in \{0,1\}^\ell$ let

$$J(b,c) := \Big(\bigcup_{i:b_i=1} M_i\Big) \cup \Big(\bigcup_{j:c_j=1} N_j\Big) \subseteq E_n.$$

By definition, $J(b,c)$ is a T-join if and only if b is odd and c is even. Let $b^* \in \{0,1\}^k$ odd and $c^* \in \{0,1\}^\ell$ even be arbitrarily chosen but fixed. Since ODD_n is a hiding set for EVEN_n and vice versa, it is now easy to see that both sets

$$\begin{aligned} H_1 &:= \left\{J(b,c^*) : b \in \{0,1\}^k \text{ even}\right\} \\ H_2 &:= \left\{J(b^*,c) : c \in \{0,1\}^\ell \text{ odd}\right\}, \end{aligned}$$

are hiding sets for T-JOINS_n. Our claim follows from Proposition 3 and the fact that

$$\begin{aligned} \max\{|H_1|,|H_2|\} = \max\left\{2^{k-1}, 2^{\ell-1}\right\} &= \max\left\{2^{\frac{1}{2}|T|-1}, 2^{\frac{1}{2}(n-|T|)-1}\right\} \\ &\geq 2^{\frac{1}{2}\cdot\frac{n}{2}-1}. \quad \square \end{aligned}$$

4 Concluding remarks

We at least asymptotically determined the relaxation complexities for several examples arising from combinatorial optimization that we were interested in. Turning towards the more basical questions, we found $\mathrm{rc}(\{0,1\}^d) = d + 1$. In contrast to this, we do, however, not know the exact value of $\mathrm{rc}(\Delta_d)$, where Δ_d is the set of vertices of the standard d-simplex. As briefly discussed in Section 2.3, a relaxation R for Δ_d with less than $d+1$ facets must be unbounded and (hence) irrational. Unlike the case of $\{0,1\}^d$, there are indeed unbounded polyhedra whose integer points are exactly the set of an integral d-simplex. As an example, it is rather easy to see that the polyhedron

$$\operatorname{conv}\{\mathbb{O}_5, \mathbb{e}_1, \mathbb{e}_2, \mathbb{e}_3, \mathbb{e}_1 + \mathbb{e}_3 + \mathbb{e}_4, \mathbb{e}_2 + \mathbb{e}_3 + \mathbb{e}_5\} + \mathbb{R} \cdot (0, 0, 0, 1, \sqrt{2})^T$$

has this property. Further, on also inds that the hiding set method does not help in this case. In fact, any hiding set of Δ_d can be shown to have cardinality of at most 3. Therefore, the simplex-example raises the question, whether there are polyhedral sets X with $\mathrm{rc}(X) < \mathrm{rc}_{\mathbb{Q}}(X)$ (see 2.3).

Acknowledgements We would like to thank Gennadiy Averkov for valuable comments on this work.

References

1. Gérard Cornuéjols, Jean Fonlupt, and Denis Naddef. The traveling salesman problem on a graph and some related integer polyhedra. *Math. Program.*, 33(1):1–27, 1985.
2. Samuel Fiorini, Serge Massar, Sebastian Pokutta, Hans Raj Tiwary, and Ronald de Wolf. Linear vs. semidefinite extended formulations: exponential separation and strong lower bounds. In Howard J. Karloff and Toniann Pitassi, editors, *STOC*, pages 95–106. ACM, 2012.
3. Bezalel Gavish and Stephen C. Graves. The travelling salesman problem and related problems. Technical report, Operations Research Center, Massachusetts Institute of Technology, 1978.
4. Michel Goemans. Smallest compact formulation for the permutahedron, 2009.
5. Robert Jeroslow. On defining sets of vertices of the hypercube by linear inequalities. *Discrete Mathematics*, 11(2):119–124, 1975.
6. Jon Lee and François Margot. On a binary-encoded ilp coloring formulation. *INFORMS Journal on Computing*, 19(3):406–415, 2007.
7. R. Kipp Martin. Using separation algorithms to generate mixed integer model reformulations. *Oper. Res. Lett.*, 10(3):119–128, 1991.
8. Clair E. Miller, Albert W. Tucker, and Richard A. Zemlin. Integer programming formulation of traveling salesman problems. *J. ACM*, 7(4):326–329, 1960.
9. Hermann Minkowski. *Geometrie der Zahlen.* Teubner Verlag, Leipzig, 1896.
10. Richardo Rado. An inequality. *J. London Math. Soc.*, 27:1–6, 1952.
11. A. Schrijver. *Combinatorial Optimization – Polyhedra and Efficiency.* Springer, 2003.
12. Alexander Schrijver. *Theory of linear and integer programming.* John Wiley & Sons, Inc., New York, NY, USA, 1986.

A Proofs

A.1 Proof of Lemma 1

Obviously, any point $x \in \{0,1\}^d$ satisfies

$$x_k \leq 1 + \sum_{i=k+1}^{d} 2^{-i} x_i \tag{5}$$

for all $k \in [d]$ and

$$x_1 + \sum_{i=2}^{d} 2^{-i} x_i \geq 0. \tag{6}$$

Let now $x \in \mathbb{Z}^d$ be any integer point that satisfies (5) for all $k \in [d]$ as well as (6). First, we claim that $x_i \leq 1$ for all $i \in [d]$: Suppose that $x_k > 1$ for some $k \in [d]$. W.l.o.g. we may assume that $x_i \leq 1$ for all $i > k$ and obtain

$$x_k \overset{(5)}{\leq} 1 + \sum_{i=k+1}^{d} 2^{-i} x_i \leq 1 + \sum_{i=k+1}^{d} 2^{-i} < 2,$$

a contradiction. Further, we see that $x_1 \geq 0$ since (due to $x_i \leq 1$ for all i)

$$x_1 \overset{(6)}{\geq} -\sum_{i=2}^{d} 2^{-i} x_i \geq -\sum_{i=2}^{d} 2^{-i} > -1.$$

It remains to show that $x_i \geq 0$ for all $i \in [d] \setminus \{1\}$. Suppose that $x_j \leq -1$ for some $j \in [d] \setminus \{1\}$ and $x_i \geq 0$ for all $i < j$. In this case, we claim that $x_i = 0$ for all $i < j$: Otherwise, let k be the largest $k < j$ such that $x_k > 0$ (and hence $x_k = 1$). By inequality (5), we would obtain

$$\begin{aligned} 1 = x_k \leq 1 + \sum_{i=k+1}^{d} 2^{-i} x_i &= 1 + 2^{-j} x_j + \sum_{i=j+1}^{d} 2^{-i} x_i \\ &\leq 1 + 2^{-j} \cdot (-1) + \sum_{i=j+1}^{d} 2^{-i} < 1. \end{aligned}$$

Thus, we have that $x_i \geq 0$ for all $i < j$ and hence by inequality (6) we deduce

$$0 \leq x_1 + \sum_{i=2}^{d} 2^{-i} x_i = 2^{-j} x_j + \sum_{i=j+1}^{d} 2^{-i} x_i \leq 2^{-j} \cdot (-1) + \sum_{i=j+1}^{d} 2^{-i} < 0,$$

a contradiction. □

A.2 Proof of Corollary 1

Let us define $L(c, \varepsilon) := \{\lambda c \in \mathbb{R}^d : \lambda \in \mathbb{R}\} + B_\varepsilon$. Since $L(c, \varepsilon)$ is symmetric with respect to the origin and $L(c, \lambda_0, \varepsilon) \subseteq L(c, \varepsilon)$, it suffices to show that $L(c, \varepsilon)$ contains infinitely many integer points.

Since the latter statement is obviously true if

$$\{\lambda c : \lambda \in \mathbb{R}\} \cap \mathbb{Z}^d \neq \{\mathbb{O}_d\}, \tag{7}$$

we assume that (7) does not hold. Setting $\varepsilon_1 := \varepsilon$, by Theorem 1, $L(c, \varepsilon_1)$ contains a point $p_1 \in \mathbb{Z}^d \setminus \{\mathbb{O}_d\}$. Since (7) does not hold, there exists some $\varepsilon_2 > 0$ such that $L(c, \varepsilon_2) \subseteq L(c, \varepsilon_1)$ and $p_1 \notin L(c, \varepsilon_2)$. Again, by Theorem 1, $L(c, \varepsilon_2)$ also contains a point $p_2 \in \mathbb{Z}^d \setminus \{\mathbb{O}_d\}$. Further, there is also some $\varepsilon_3 > 0$ such that $L(c, \varepsilon_3) \subseteq L(c, \varepsilon_2)$ and $p_2 \notin L(c, \varepsilon_3)$. By iterating these arguments, we obtain an infinite sequence (ε_i, p_i) such that $p_i \in L(c, \varepsilon_i) \cap \mathbb{Z}^d \subseteq L(c, \varepsilon) \cap \mathbb{Z}^d$ and $p_i \notin L(c, \varepsilon_{i+1})$ for all $i > 0$. In particular, all p_i are distinct. □

B Figures

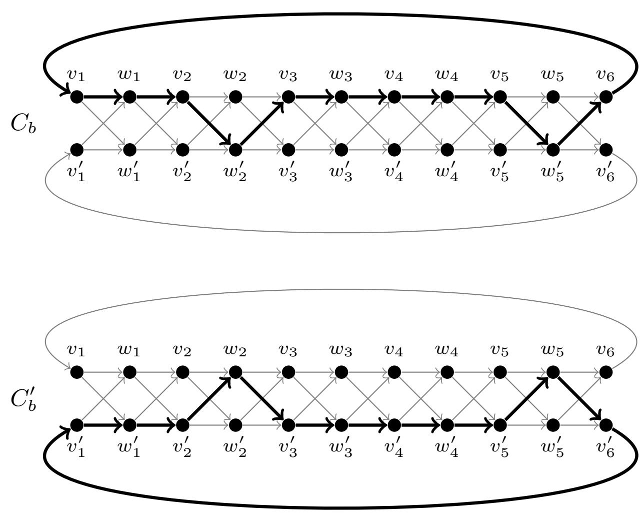

Fig. 1. Cycles $C(b)$ and $C'(b)$ for $b = (0, 1, 0, 0, 1)$.

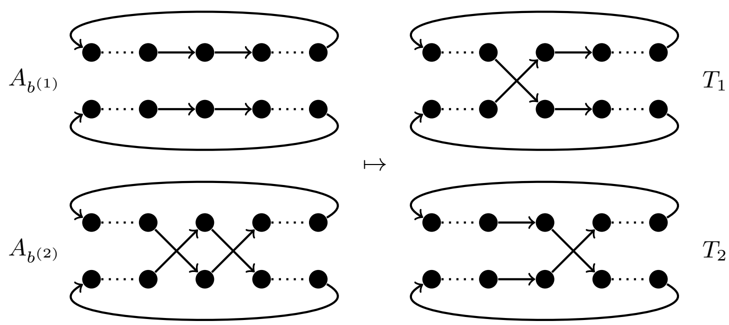

Fig. 2. Construction of T_1 and T_2.