

On the Size of Integer Programs with Bounded Coefficients or Sparse Constraints

Christopher Hojny, Hendrik Lüthen, and Marc E. Pfetsch

TU Darmstadt, Department of Mathematics, Discrete Optimization Group,
{hojny,luethen,pfetsch}@mathematik.tu-darmstadt.de

December 2017

Abstract

Integer programming formulations describe optimization problems over a set of integer points. A fundamental problem is to determine the minimal size of such formulations, in particular, if the size of the coefficients or sparsity of the constraints is bounded. This article considers lower and upper bounds on these sizes both in the original and in extended spaces, i.e., if additional variables are allowed. We focus on the original space, where we provide lower bounds for knapsack problems on the size of integer formulations with bounded coefficients. For 0/1-problems, we also introduce a technique to compute a tight lower bound on the number of non-zeros of integer formulations in the original space. Moreover, we present statistics on these bounds in small dimensions. Finally, we consider conditions on the representability of integer optimization problems with arbitrary objective function as mixed-integer programs. In particular, we show non-existence of such formulations in some special cases, e.g., efficient determination or preservation of all optimal solutions. All results are illustrated by examples.

1 Introduction

Let $X \subseteq \mathbb{Z}^p \times \mathbb{R}^q$ be a set that should be described using a *mixed-integer formulation in the original space*, i.e.,

$$X = \{x \in \mathbb{Z}^p \times \mathbb{R}^q : Ax \leq b\},$$

for $A \in \mathbb{R}^{m \times (p+q)}$, $b \in \mathbb{R}^m$. Clearly, a necessary condition for the existence of such a formulation is that X is *hole-free*, i.e., $X = \text{conv}(X) \cap (\mathbb{Z}^p \times \mathbb{R}^q)$. Moreover, allowing to add further variables and letting X be the projection of the corresponding set yields a (mixed-integer) *extended formulation*.

Such formulations are of course fundamental for operations research and have been studied in the context of particular applications in countless publications. General properties of mixed-integer formulations have been investigated as well. For instance, Jeroslow and Lowe [14] completely characterized whether X can be represented using a mixed-integer formulation. Recently, Basu et al. [2] derived an equivalent algebraic characterization. In particular, X can always be represented by an extended formulation if X is bounded. One possibility to generalize this result is to allow general convex constraints in the representation of a mixed-integer set X . This leads to the concept of mixed-integer convex representability, see, e.g., Lubin et al. [20]. However, we do not pursue this line of research here and concentrate on linear formulations.

In recent years, the investigation of extended formulations that *describe the mixed-integer hull* of X has received considerable attention, see, for example, Conforti et al. [4], Vanderbeck and Wolsey [25], and Kaibel [15] for an overview. Moreover, lower bounds for approximations of the mixed-integer hull have been considered as well, see, for example, Faenza and Sanità [8] and Braun et al. [3]. In this article, we consider arbitrary mixed-integer formulations that do not necessarily completely describe or approximate the mixed-integer hull. In this context, Kaibel and Weltge [17] introduced the concept of hiding sets to prove exponential lower bounds on the number of inequalities of such formulations for several combinatorial (optimization) problems in the original space.

Besides the size, other properties of a mixed-integer formulation are relevant as well. For example, a small formulation may be completely useless if its constraints have very large coefficients, since this poses the danger of numerical instabilities. Moreover, dense formulations may lead to large solving times. For this reason, it is natural to consider the sparsity of mixed-integer formulations. Furthermore, there exist problems like the graph partitioning problem that can be easily formulated as mixed integer programs, in which one part of the variables is only necessary to model the objective. Consequently, the question arises whether it is possible to find formulations without auxiliary variables.

In this article, we therefore extend the previous research in three different ways by taking the size of the coefficients, sparsity and the objective into account: First, we consider lower bounds on the size of a mixed-integer formulation if the size of the coefficients is bounded. One of the prime examples where this is relevant is the case of knapsack problems for which we derive such lower bounds for ternary coefficients. Second, we derive methods to compute lower bounds on the sparsity that a formulation can have. Third, we discuss conditions under which problems with (nonlinear) objective function and a given feasibility set X can be formulated as a mixed-integer (linear) program. All methods are demonstrated by examples and are complemented by computational results that, for small dimensions, provide statistics of all formulations with bounded coefficients and sparsity.

Note that we will concentrate on techniques for sets $X \subseteq \{0, 1\}^n$, which covers many interesting combinatorial optimization problems. Furthermore, note that we require that X is given. Thus, our methods cannot be used to derive general statements on any formulation for which the encoding (via X) is variable.

The structure of this article is then as follows: In Section 2, we formally define mixed-integer formulations in the original and extended space, and we recall some results from the literature that will be useful in our analysis. In particular, we will see that every formulation with bounded variables can easily be reformulated with small coefficients and sparse constraints. For this reason, we restrict ourselves to the original space in Sections 3 and 4. In Section 3 we focus on how to obtain mixed-integer formulations with bounded coefficients. We show how to derive lower bounds on the number of inequalities with ternary coefficients for integer formulations of knapsack problems, and we present statistics on the size of integer formulations with bounded coefficients in small dimensions. Thereafter, in Section 4, we focus on sparsity and present a tight lower bound on the sparsity of integer formulations as well as a technique to exactly compute the minimum sparsity of any integer formulation. Analogous to integer formulations with bounded coefficients, we present statistics on the minimum sparsity of integer formulations in small dimensions in Section 4.2. Finally, we discuss in Section 5 concepts of mixed-integer formulations concerning the incorporation of an objective function. We state a formal definition, discuss some possible variations and show the non-existence of such formulations under additional requirements, as for example efficient determination or preservation of all optimal solutions.

2 Basics of Mixed-Integer Formulations

To fix notation, we formally define the following.

Definition 1. A *mixed-integer program (MIP)* is an optimization problem

$$\max \{g(x) : x \in Q \cap (\mathbb{Z}^p \times \mathbb{R}^q)\},$$

specified by a polyhedron $Q \subseteq \mathbb{R}^p \times \mathbb{R}^q$ and an affine function $g: \mathbb{R}^p \times \mathbb{R}^q \rightarrow \mathbb{R}$. We will use the tuple (Q, g) to denote a MIP.

In the first part of this article, we will exclusively deal with representations of the feasible set X using $Q \cap (\mathbb{Z}^p \times \mathbb{R}^q)$, while the second part will also treat objective functions, see Section 5. Thus, for the feasible set, we define:

Definition 2. A *mixed-integer formulation* of a set $X \subseteq \mathbb{Z}^p \times \mathbb{R}^q$ is a pair (Q, π) with a polyhedron $Q = \{(y, z) \in \mathbb{R}^{p'} \times \mathbb{R}^{q'} : Ay + Bz \leq b\}$ and an affine function $\pi: \mathbb{R}^{p'} \times \mathbb{R}^{q'} \rightarrow \mathbb{R}^p \times \mathbb{R}^q$ such that

$$X = \pi(Q \cap (\mathbb{Z}^{p'} \times \mathbb{R}^{q'})),$$

where $A \in \mathbb{R}^{m \times p'}$, $B \in \mathbb{R}^{m \times q'}$, and $b \in \mathbb{R}^m$. If π is the identity, the formulation is *in the original space*, otherwise *in an extended space* (or an *extended formulation* for short). Furthermore, if $q' = 0$, we call the the formulation an *integer formulation*.

Note that the original x -variables are allowed, but not required, to occur in Q and π . Moreover, in many cases π can be chosen to be an orthogonal projection, however, there are also examples in which the projection is more complicated, see, e.g., Martin et al. [21].

In preparation of the following, we refer to a result of Kaibel and Weltge [17], which shows that there always exists a polynomial size extended mixed-integer formulation of a set $X \subseteq \{0, 1\}^n$ if the membership problem “Given $x \in \mathbb{R}^n$, is $x \in X$?” is in \mathcal{NP} .

Theorem 3 (Kaibel and Weltge [17]). *Consider a 0/1-problem that defines $X^I \subseteq \{0, 1\}^{n(I)}$ for each instance I , and let the membership problem for X^I be in \mathcal{NP} . Then there exists a polynomial p such that for each instance I , there is a system $Ax + By \leq b$ of at most $p(n(I))$ linear inequalities and $k(I) \leq p(n(I))$ auxiliary variables with*

$$X^I = \{x \in \{0, 1\}^{n(I)} : \exists y \in \mathbb{Z}^{k(I)} \text{ with } Ax + By \leq b\}.$$

If the membership problem is in \mathcal{P} , the integrality condition on y can be dropped.

Moreover, analyzing the proof of Theorem 3 in [17] shows that A and B have coefficients in $\{0, \pm 1\}$ only. The proof is constructive, but the resulting system might be very large and complicated. Furthermore, the mixed-integer formulation of Theorem 3 can even be sparsified by replacing in each constraint sums of two variables by a newly introduced variable, see Appendix A and Pritchard and Chakrabarty [24]. The resulting system then contains at most three nonzero coefficients in each constraint.

If $X = X^I$ is not a subset of $\{0, 1\}^n$ but more general, Theorem 3 cannot be used to obtain a mixed-integer formulation with bounded coefficients. However, if X admits a mixed-integer formulation, we can easily derive an extended mixed-integer formulation of X with bounded coefficients, which can be seen as follows. Assume we are given an extended mixed-integer formulation of a mixed-integer set $X \subseteq \mathbb{R}^n$ of the form

$$Q = \{(y, z) \in \mathbb{Z}^p \times \mathbb{R}^q : Ay + Bz \leq b, 0 \leq y \leq u, 0 \leq z \leq \tilde{u}\}, \quad (1)$$

where $A \in \mathbb{Z}^{m \times p}$, $B \in \mathbb{Z}^{m \times q}$, $u \in \mathbb{Z}^p$, $\tilde{u} \in \mathbb{Z}^q$, together with its projection to the original space. By replacing variables y_i and z_i with coefficients a_i and b_i , respectively, by the sum of $|a_i|$ and $|b_i|$ variables with coefficients in $\{-1, +1\}$, we can easily construct an extended formulation of (1) with left-hand side coefficients in $\{0, \pm 1\}$, see Appendix B. Moreover, the derived mixed-integer formulation is of polynomial size if the original mixed-integer formulation has polynomial size in $p + q$, see Theorem 36 in Appendix B.

While mixed-integer formulations with bounded coefficients in an extended space seem to be easy to handle, mixed-integer formulations in the original space are more challenging. In particular, for many examples Kaibel and Weltge [17] showed that, in contrast to an extended space, no polynomial size formulations may exist. However, it is still possible to find tractable, i.e., polynomial time separable, formulations with $\{0, \pm 1\}$ -coefficients on the left-hand side of many 0/1-problems in the original space. To see this, we make use of the concept of infeasibility cuts: Given a set $X \subseteq \{0, 1\}^n$ and a point $\tilde{x} \in \{0, 1\}^n \setminus X$, the *infeasibility* or *no-good cut* w.r.t. \tilde{x} is given by

$$\sum_{i: \tilde{x}_i=0} x_i + \sum_{i: \tilde{x}_i=1} (1 - x_i) \geq 1. \quad (2)$$

The only binary point that violates this inequality is \tilde{x} . Thus, the box constraints and infeasibility cuts for all infeasible binary points define an integer formulation of X with coefficients in $\{0, \pm 1\}$. Obviously, the tractability of these inequalities depends on the way X is given. For example, if $|\{0, 1\}^n \setminus X|$ is small and can be enumerated explicitly, (2) yields a polynomial size integer formulation. The following theorem gives a sufficient condition on when the formulation by infeasibility cuts is tractable.

Theorem 4. *Consider some 0/1-problem that defines $X^I \subseteq \{0, 1\}^{n(I)}$ for each instance I , and let the membership problem for X^I be solvable in polynomial time in n . Then the integer formulation given by (2) for each $\tilde{x} \in \{0, 1\}^{n(I)} \setminus X^I$ and the bounds $0 \leq x \leq \mathbb{1}$ is tractable in n .*

Proof. Since the box constraints $0 \leq x \leq \mathbb{1}$ can be separated in linear time, it suffices to show that the separation problem for (2) and $x^* \in [0, 1]^n$ can be solved in polynomial time.

Note that the left-hand side of (2) is $\|\tilde{x} - x\|_1$, where $\|x\|_1$ denotes the 1-norm. To solve the separation problem of infeasibility cuts for X , we solve the auxiliary problem

$$\min_{\hat{x} \in \{0, 1\}^n} \|\hat{x} - x^*\|_1$$

first. Obviously,

$$\hat{x}_i = \begin{cases} 0 & \text{if } x_i^* \leq \frac{1}{2}, \\ 1 & \text{otherwise} \end{cases}$$

is a solution of this problem and it can be computed in linear time.

If $\|\hat{x} - x^*\|_1 \geq 1$, the point x^* cannot violate (2) for any $\tilde{x} \in \{0, 1\}^n$. Hence, x^* lies inside the integer formulation of X via infeasibility cuts. Otherwise, $\|\hat{x} - x^*\|_1 < 1$, and we are done if $\hat{x} \in \{0, 1\}^n \setminus X$, because the infeasibility cut for \hat{x} is violated by x^* . Thus, we can assume in the following that $\hat{x} \in X$.

Assume now that $\|\tilde{x} - \hat{x}\|_1 \geq 2$. Then plugging x^* into (2) yields

$$\sum_{i: \tilde{x}_i=0} x_i^* + \sum_{i: \tilde{x}_i=1} (1 - x_i^*) = \|\tilde{x} - x^*\|_1 = \|(\tilde{x} - \hat{x}) + (\hat{x} - x^*)\|_1 \geq \underbrace{\|\tilde{x} - \hat{x}\|_1}_{\geq 2} - \underbrace{\|\hat{x} - x^*\|_1}_{< 1} > 1.$$

Thus, (2) cannot be violated in this case. It therefore suffices to check at most $n + 1$ points $\tilde{x} \in \{0, 1\}^n$ with $\|\tilde{x} - \hat{x}\|_1 \leq 1$. Thus, we call the membership problem for each of these points and check whether (2) is violated by x^* if \tilde{x} is infeasible. \square

Furthermore, it is also possible to characterize cases in which no tractable mixed-integer formulation exists in the original space.

Observation 5. *Consider some 0/1-problem that defines $X^I \subseteq \{0, 1\}^{n(I)}$ for each instance I , and let the membership problem for X^I be \mathcal{NP} -hard. Then, as long as $\mathcal{P} \neq \mathcal{NP}$, there exists no tractable mixed-integer formulation, in particular, no formulation of polynomial size, in the original space.*

Proof. If X^I admitted a tractable mixed-integer formulation or a mixed-integer formulation of polynomial size in the original space, this mixed-integer formulation would allow to answer the membership problem for X^I in polynomial time. \square

3 Bounded Coefficients

For many combinatorial optimization problems integer formulations with small coefficients are known. For some problems, these formulations may be of polynomial size, e.g., for the graph coloring or independent set problem. For other problems, Kaibel and Weltge [17] proved that *any* formulation in the original space is of exponential size.

Some of the remaining problems have a small formulation using arbitrary (large) coefficients, but it is unclear whether there exists a formulation with small coefficients. A prototypical example is given by the knapsack problem, which we consider in this section. Here, the feasible region is given by

$$X^{a,\beta} := \{x \in \{0, 1\}^n : a^\top x \leq \beta\} \quad (3)$$

for some $a \in \mathbb{Z}_{>0}^n$ and $\beta \in \mathbb{Z}_{>0}$.

A small formulation is given by $a^\top x \leq \beta$ and the trivial inequalities. But since a may contain large coefficients, the question is whether there exist small formulations using small coefficients. In this section we consider the base case in which the coefficients of the normal vectors are ternary, i.e., restricted to be in $\{0, \pm 1\}$. We present a technique to derive strong lower bounds on the size of ternary integer formulations for such problems and provide an empirical comparison of the size of integer formulations with different coefficient bounds for all feasible sets X .

3.1 Integer Formulations with $\{0, \pm 1\}$ -Coefficients for Knapsack Problems

To be able to derive lower bounds on the size of ternary integer formulations for knapsack problems, we need some terminology from the literature. Throughout this section, we assume that $a \in \mathbb{Z}_{>0}^n$ and $\beta \in \mathbb{Z}_{>0}$, as well as $a_1 \leq a_2 \leq \dots \leq a_n \leq \beta$. A set $C \subseteq [n] := \{1, \dots, n\}$ is called a *cover* of a knapsack $X^{a,\beta}$ if $a^\top \chi^C > \beta$, where χ^C is the characteristic vector of C . A cover C is called *minimal* if each proper subset S is feasible, i.e., $\chi(S) \in X^{a,\beta}$; the corresponding *minimal cover inequality* is $\sum_{i \in C} x_i \leq |C| - 1$. A cover $C = \{i_1, \dots, i_\ell\}$, where $i_1 < i_2 < \dots < i_\ell$, is called *strong* if C is a minimal cover and $\sum_{i \in C} a_i - a_{i_\ell} + a_j \leq \beta$ for the maximal $j \in [i_\ell] \setminus C$. Moreover, the *extension* of a (not necessarily) strong cover $C = \{i_1, \dots, i_\ell\}$ is the set $E(C) := C \cup \{i_\ell + 1, \dots, n\}$. Since the weights a_i are sorted non-decreasingly, the *extension inequality*

$$\sum_{i \in E(C)} x_i = \sum_{i \in C} x_i + \sum_{i=i_\ell+1}^n x_i \leq |C| - 1 \quad (4)$$

is valid for $X^{a,\beta}$.

Proposition 6. *Let C be a strong cover of $X^{a,\beta}$. Then there does not exist a ternary inequality valid for $X^{a,\beta}$ that dominates the extension inequality of C . Moreover, if C is a cover that is not strong, there exists a strong cover whose extension inequality dominates the extension inequality of C .*

Proof. Let $C = \{i_1, \dots, i_\ell\}$ be a strong cover of $X^{a,\beta}$. For the sake of contradiction, assume there exists an inequality $c^\top x \leq \delta$, $c \in \{0, \pm 1\}^n$, that is valid for $X^{a,\beta}$ and that dominates the extension inequality of C . Since $c^\top x \leq \delta$ can be written as a conic combination of facets of $\text{conv}(X^{a,\beta})$ and $\text{conv}(X^{a,\beta})$ is full-dimensional, we can assume w.l.o.g. that $c \in \{0, 1\}^n$ due to Hammer et al. [12]. Moreover, validity of $c^\top x \leq \delta$ for $\text{conv}(X^{a,\beta})$ implies $\delta \geq 0$ because the zero vector is feasible. Additionally, dominance of $c^\top x \leq \delta$ implies $c^\top \chi^C > \delta$, whereas validity implies $c^\top \chi^{C'} \leq \delta$ for every $C' \subsetneq C$. Thus, $c_i = 1$ for every $i \in C$, $\lfloor \delta \rfloor = |C| - 1$ and

$$c^\top x \leq \lfloor \delta \rfloor \quad \Leftrightarrow \quad \sum_{i \in C'} x_i \leq |C| - 1,$$

for $C' := \{i \in [n] : c_i = 1\} \supseteq C$. Dominance implies that C' is a proper superset of $E(C)$. Let j be the maximal element in $C' \setminus E(C)$. Then, $j < i_\ell$ and the characteristic vector of $I = (C \setminus \{i_\ell\}) \cup \{j\}$ is cut off by $c^\top x \leq \delta$, because $|I| = |C|$. However, strongness of C implies

$$a^\top \chi^I = \sum_{i \in C} a_i - a_{i_\ell} + a_j \leq \beta,$$

which proves that $\chi^I \in X^{a,\beta}$. Hence, $c^\top x \leq \delta$ cannot be valid – a contradiction. Consequently, the extension inequality of C cannot be dominated by a valid ternary inequality.

To prove the second part of the proposition, let C be a cover that is not strong. If $C = \{i_1, \dots, i_\ell\}$ is not minimal, there exists a minimal cover $\bar{C} = \{\bar{i}_1, \dots, \bar{i}_{\bar{\ell}}\} \subsetneq C$. Since $\bar{i}_{\bar{\ell}} \leq i_\ell$, we have $E(\bar{C}) \setminus \bar{C} \subseteq E(C) \setminus C$. Hence, the extension inequality of C can be dominated by summing the extension inequality of C as well as $x_i \leq 1$ for all $i \in C \setminus E(\bar{C})$. For this reason, we can assume in the following that $C = \{i_1, \dots, i_\ell\}$ is minimal but not strong. Let j be the maximal index in $[i_\ell] \setminus C$. Then $\bar{C} = \{i_1, \dots, i_{\ell-1}\} \cup \{j\}$ is a minimal cover since C is not strong. Furthermore, $E(C) \subsetneq E(\bar{C})$ holds because the maximum index in \bar{C} is smaller than the maximum index in C . If \bar{C} is not a strong cover, we can iterate this process until we end up in a strong cover, which shows that the extension inequality of C can be dominated by the extension inequality of a strong cover. \square

Thus, any ternary integer formulation of $X^{a,\beta}$ can be strengthened to only contain extension inequalities of strong covers (if necessary). In particular, this allows to find a ternary integer formulation of minimum size, see the following theorem by Glover [10] as stated by Wolsey [26].

Theorem 7 (Wolsey [26]). *Let \mathfrak{S} be the set of all strong covers of $X^{a,\beta}$. Then*

$$\{x \in \{0, 1\}^n : x \text{ fulfills (4) for all } C \in \mathfrak{S}\}$$

is a ternary integer formulation of $X^{a,\beta}$ of minimum size. That is, there is no smaller set of inequalities with ternary coefficients which enforces that a binary vector is contained in $X^{a,\beta}$.

Theorem 7 shows that $X^{a,\beta}$ admits an integer formulation with ternary coefficients since box constraints have coefficients in $\{0, \pm 1\}$ and extension inequalities have coefficients in $\{0, 1\}$.

Based on this result, the aim of this section is twofold. On the one hand, we are interested in computing the minimum size of an integer formulation of $X^{a,\beta}$ exactly. To this end, we provide a mechanism to generate all strong covers of $X^{a,\beta}$. On the other hand, Theorem 7 allows to find lower bounds on the size of an arbitrary integer formulation of knapsack sets $X^{a,\beta}$ with ternary coefficients by guessing strong covers. For this reason, we show how to generate further strong covers from a set of given strong covers. Two examples will illustrate the applicability of this technique to derive exponential lower bounds on the size of ternary integer formulations of particular knapsacks.

Note that these results also allow to find ternary integer formulations if a has negative entries a_j , because we can complement the variable $x_j \mapsto (1 - x_j)$ to modify the coefficient of x_j to $-a_j$ (the right-hand side of the constraint has to be adjusted accordingly). Performing these steps for each variable with a negative coefficient leads to a system allowing for the techniques above. Complementing the same variables of the resulting system again leads to an integer formulation for the original constraint $a^\top x \leq \beta$ with left-hand side coefficients in $\{0, \pm 1\}$.

To derive the proposed mechanism for generating strong covers, we use a dynamic programming approach, which we discuss in the following. This mechanism is very similar to the classical method to find an optimal knapsack solution, see, e.g., Kellerer et al. [18], and it exploits the assumption $a_1 \leq a_2 \leq \dots \leq a_n$. To generate strong covers, we introduce for every $j \in [n]_0$ and $k \in \{0, \dots, \sum_{i=1}^n a_i\}$ a state $S(j, k)$ that contains all sets $I \subseteq [j]$ with $\sum_{i \in I} a_i = k$, where $S(0, 0) = \{\emptyset\}$. Then, state $S(j, k)$ can be generated from previous states by the following recursion:

$$S(j, k) = \begin{cases} S(j-1, k), & \text{if } k < a_j, \\ S(j-1, k) \cup \{I \cup \{j\} : I \in S(j-1, k - a_j)\}, & \text{if } k \geq a_j, \end{cases} \quad (5)$$

i.e., we take all subsets of $[j-1]$ with weight k and we extend all subsets of $[j-1]$ with weight $k - a_j$ by $\{j\}$.

Using this recursion, the set \mathfrak{S} of all strong covers for $X^{a,\beta}$ can be generated, in principle, by the following procedure: We initialize the set $\mathfrak{S} = \emptyset$, and we use Recursion (5) to compute all states of the knapsack problem. Whenever a state $S(j, k)$, $k > \beta$, is generated during the recursion, we check for every $C \in S(j, k)$ whether C fulfills the requirements of a strong cover. If the check evaluates positively, we add C to \mathfrak{S} . After termination of this procedure, all strong covers are clearly contained in \mathfrak{S} .

Although this approach is correct, it is impractical in applications due to its long running time. For this reason, we are interested in deriving lower bounds on $|\mathfrak{S}|$ to obtain lower bounds on the size of a ternary integer formulation for $X^{a,\beta}$. To this end, we introduce the following notation: Let $A := \{a_i : i \in [n]\}$ be the set of different left-hand side coefficients of $a^\top x \leq \beta$. For every $\alpha \in A$, we define $A_\alpha := \{i \in [n] : a_i = \alpha\}$, i.e., A_α contains all indices of variables with coefficient α in the knapsack constraint. Moreover, define $\binom{A_\alpha}{n_\alpha}$ as the set containing all subsets of A_α with cardinality n_α .

Proposition 8. *Let C be a strong cover of $X^{a,\beta}$, let $n_\alpha := |C \cap A_\alpha|$, $\alpha \in A$, and define $\bar{\alpha} := \max \{\alpha \in A : n_\alpha > 0\}$. Then, for every $C'_\alpha \in \binom{A_\alpha}{n_\alpha}$, $\alpha \in A \setminus \{\bar{\alpha}\}$, the set*

$$C' = \left(\bigcup_{\alpha \in A \setminus \{\bar{\alpha}\}} C'_\alpha \right) \cup (C \cap A_{\bar{\alpha}}),$$

is a strong cover of $X^{a,\beta}$.

Proof. Let $C = \{i_1, \dots, i_\ell\}$ be a strong cover and let $C' = \{i'_1, \dots, i'_\ell\}$ be defined as stated. Then $i_{\ell'} = i_\ell$ since elements in $A_{\bar{\alpha}}$ are not modified. Moreover, there exists for every $\alpha \in A$ a bijection $\phi_\alpha: A_\alpha \rightarrow A_\alpha$ that maps $C' \cap A_\alpha$ to $C \cap A_\alpha$, where $\phi_{\bar{\alpha}}$ is the identity on $A_{\bar{\alpha}}$. If C' was not a strong cover, then $\beta < \sum_{i' \in C'} a_{i'} - a_{i'_\ell} + a_{j'}$ holds for the maximum element $j' \in [i'_\ell] \setminus C'$. In particular, since $C \cap A_{\bar{\alpha}} = C' \cap A_{\bar{\alpha}}$ and $\phi_\alpha(C' \cap A_\alpha) = C \cap A_\alpha$ for every $\alpha \in A \setminus \{\bar{\alpha}\}$ by definition, we have

$$j := \phi_{a_{j'}}(j') \in [\phi_{\bar{\alpha}}(i'_\ell)] \setminus C = [i_\ell] \setminus C. \quad (6)$$

This implies

$$\begin{aligned} \beta &< \sum_{i' \in C'} a_{i'} - a_{i'_\ell} + a_{j'} = \sum_{\alpha \in A} \sum_{i' \in C' \cap A_\alpha} a_{i'} - a_{i'_\ell} + a_{j'} \\ &= \sum_{\alpha \in A} \sum_{i' \in C' \cap A_\alpha} a_{\phi_\alpha(i')} - a_{\phi_{\bar{\alpha}}(i'_\ell)} + a_{j'} = \sum_{i \in C} a_i - a_{i_\ell} + a_j, \end{aligned}$$

because j and j' have the same coefficient in the knapsack inequality by (6). Thus, since $j \in [i_\ell] \setminus C$, we conclude that C cannot be a strong cover – a contradiction. For this reason, every set C' that is generated as proposed above is a strong cover. \square

This proposition implies, in particular, that we can generate $\prod_{\ell=1}^{\bar{\alpha}-1} \binom{|A_\alpha|}{n_\alpha}$ many strong covers from one given strong cover. Thus, by guessing some strong covers, Proposition 8 and Theorem 7 yield a lower bound on the number of inequalities in a ternary integer formulation. In the remainder of this section, we show two applications of this result to derive exponential lower bounds for particular knapsacks.

Example 9. Consider the knapsack inequality $\sum_{i=1}^n x_i + 2 \sum_{i=n+1}^{2n} x_i \leq n$ written as $a^\top x \leq n$, where n is even. Let $I_k := [k] \cup \{n+1, \dots, n + \frac{n-k-1}{2}\}$ for every odd $k \in [n]$. Then, $a^\top \chi^{I_k} = n-1$ implies that $\chi^{I_k} \in X^{a,n}$. Furthermore, adding any element with weight 2 to I_k increases the weight to $n+1$. Thus, the extended set is not in $X^{a,n}$. Consequently, the set $C_k := I_k \cup \{n + \frac{n-k+1}{2}\}$ is a minimal cover, and it is strong because every element in $[n + \frac{n-k+1}{2}] \setminus C_k$ has weight 1. Hence, Proposition 8 shows that we can generate $\binom{n}{k}$ many strong covers from C_k and they are pairwise different. For this reason, any ternary integer formulation of $X^{a,n}$ contains at least

$$\sum_{\substack{k \in [n]: \\ k \text{ odd}}} \binom{n}{k} = 2^{n-1}$$

inequalities. In fact, one can show that the sets C_k are the strong covers of $X^{a,n}$. Thus, restricting to inequalities with ternary coefficients drastically increases the number of necessary constraints in comparison to an integer formulation with coefficients in $\{-2, \dots, 2\}$.

Example 10. Consider the knapsack inequality $\sum_{i=1}^{2n} 2^{\lceil i/2 \rceil - 1} x_i \leq 2^n - 1$ written as $a^\top x \leq \beta$, which is associated with the so-called orbisack, introduced by Kaibel and Loos [16]. Let $C = \{i \in [2n] : i \text{ odd or } i = 2\}$. Then $a^\top \chi^C = 2^n$, which implies that C is a minimal cover because all coefficients are greater than or equal to 1.

Moreover, C is strong: For $n \leq 2$, this follows by a case distinction. Otherwise, if $n > 2$, we have that $2n-2$ is the index with the largest coefficient not in C for which we estimate

$$\sum_{i \in C} a_i - a_{2n-1} + a_{2n-2} = 1 + \sum_{i=1}^n 2^{i-1} - 2^{n-1} + 2^{n-2} = 2^n - 2^{n-1} + 2^{n-2} \leq 2^n - 1.$$

Thus, C is a strong cover.

Since C contains all elements of weight 1, one element of weight 2^{n-1} as well as one of both items of weight 2^i , $i \in \{1, \dots, n-2\}$, Proposition 8 implies that we can generate 2^{n-2} strong covers from C . This shows that we need exponentially many inequalities in any ternary integer formulation of $X^{a,\beta}$. In fact, one can show that the exact number of strong covers is 2^{n-1} . Thus, by guessing the single set C we can generate one half of all strong covers.

3.2 Statistics on the Size of Integer Formulations with Bounded Coefficients

The aim of this section is to quantify the number of nonempty subsets $X \subsetneq \{0, 1\}^n$ that admit an integer formulation with bounded coefficients and whose number of inequalities does not exceed a given threshold. In particular, we focus on the cases for which the left-hand side coefficients are contained in $\{0, \pm 1\}$ and $\{-2, \dots, 2\}$.

To obtain statistics, we modeled the problem “Does there exist an integer formulation of $X \subseteq \{0, 1\}^n$ with at most k inequalities all of whose left-hand side coefficients are contained in $\{-c, \dots, c\}$?” as the problem to decide whether the following system is feasible. Let $B_r := \{-r, \dots, r+1\}^n$ be the box that adds r integer layers around $\{0, 1\}^n$.

$$\sum_{j=1}^n A_{ij} x_j \leq b_i, \quad i \in [k], x \in X, \quad (7a)$$

$$\sum_{j=1}^n A_{ij} y_j \geq b_i + 1 - (1 - z_{yi}) M, \quad i \in [k], y \in B_r \setminus X, \quad (7b)$$

$$\sum_{i=1}^k z_{yi} \geq 1, \quad y \in B_r \setminus X, \quad (7c)$$

$$A_{ij} \in \{-c, \dots, c\}, \quad i \in [k], j \in [n],$$

$$b_i \in \{-nc, \dots, nc\}, \quad i \in [k],$$

$$z_{yi} \in \{0, 1\}, \quad i \in [k], y \in B_r \setminus X,$$

where M is a sufficiently large constant, e.g., $ncr + 1$. In this model, A_i is the normal vector of a linear inequality and b_i is its right-hand side coefficient. Inequality (7a) ensures that $A_i \cdot x \leq b_i$ is valid for X . If $z_{yi} = 1$ for $y \in B_r \setminus X$, Inequality (7b) implies that $A_i \cdot x \leq b_i$ separates y from X . Finally, Inequality (7c) enforces that each point in $B_r \setminus X$ is separated by at least one inequality. Observe that it suffices to consider $b_i \in \{-nc, \dots, nc\}$, because the absolute value of a left-hand side coefficient in $Ax \leq b$ is bounded by c and all feasible points are contained in $\{0, 1\}^n$. Thus, if System 7 has a solution, $Ax \leq b$ is an integer formulation of X within the box B_r with k inequalities all of whose left-hand side coefficients are contained in $\{-c, \dots, c\}$. Otherwise, no such integer formulation of X exists.

Observe that this problem is a relaxation of the infinite dimensional condition enforcing that $y \in \mathbb{Z}^n \setminus X$ is cut off by some inequality. Hence, the feasible region of the relaxed problem (7) is too large in general, i.e., we cannot guarantee that all points in $\mathbb{Z}^n \setminus X$ are separated by an inequality $A_i \cdot x \leq b_i$. This can be checked, for example, by testing whether the feasible region of the found integer formulation is bounded and by counting the number of integer points contained in the feasible region, since no feasible point in X is cut off by construction. If the found system $Ax \leq b$ is not a valid integer formulation, we increase the bound r to cut off more invalid points. In our experiments a maximum bound of $r = 4$

was sufficient to find a valid integer formulation for each X or to prove that no integer formulation with bounded coefficients exists.

Furthermore, we are interested in lower bounds on the size of any integer formulation of a nonempty set $X \subsetneq \{0, 1\}^n$ to be able to decide whether small coefficients suffice to describe X with few inequalities. To this end, we use the previously mentioned concept of hiding sets by Kaibel and Weltge [17]: Given a set $X \subseteq \mathbb{Z}^n$, a set $H \subseteq (\text{aff}(X) \cap \mathbb{Z}^n) \setminus \text{conv}(X)$ is a *hiding set* if for any two distinct points $x, y \in H$ the relation $\text{conv}(\{x, y\}) \cap \text{conv}(X) \neq \emptyset$ holds, where $\text{aff}(\cdot)$ denotes the affine hull.

Kaibel and Weltge proved that if X admits a hiding set H , then any integer formulation of X consists of at least $|H|$ inequalities. Note, however, that this lower bound does not take boundedness of coefficients into account. For this reason, in our context it would be preferable to compute lower bounds via a generalization to hiding sets that respect coefficient bounds. Unfortunately, it is rather complicated to check whether a set H is such a generalized hiding set. For this reason, we compute general lower bounds on the size of integer formulations by computing for each set $X \subseteq \{0, 1\}^n$ a maximum hiding set in $\{0, 1\}^n$. Note that the restriction to binary hiding sets is due to practical reasons, since we need a bounded search space. A generalization to hiding sets in other bounded regions would of course be possible.

To compute binary hiding sets, we used the following procedure. Given a nonempty set $X \subsetneq \{0, 1\}^n$, we build a graph $G = (V, E)$ with $V = \{0, 1\}^n \setminus X$. The edge set E of G is defined as

$$E = \{\{x, y\} : x, y \in V, \text{conv}(\{x, y\}) \cap \text{conv}(X) \neq \emptyset\}.$$

Thus, any clique in G is a hiding set by definition. By computing the size of a maximum clique in G , we get the maximum size of any hiding set $H \subseteq \{0, 1\}^n$ for X .

Note that the hiding set bound can be very weak, e.g., if $|X| = 2^n - 1$, a maximum binary hiding set has size 1. To improve this bound, observe that any nonempty $X \subsetneq \{0, 1\}^n$ needs at least $n + 1$ inequalities in any integer formulation. We denote this lower bound, i.e., the maximum of $n + 1$ and the size of a maximum binary hiding set, by $\ell(X)$.

The results of our experiments in Table 1 show that increasing the size of the coefficients can decrease the size of an integer formulation even in small dimensions. In each dimension the maximum number of inequalities decreases by at least one if we allow left-hand side coefficients in $\{-2, \dots, 2\}$, i.e., $c = 2$ instead of $c = 1$. Furthermore, there are many point sets that can be described by few inequalities if we allow larger coefficients. For example, about 95 % of all non-trivial subsets of $\{0, 1\}^3$ can be described by four inequalities and about 87 % of all non-trivial subsets of $\{0, 1\}^4$ can be described by five inequalities if we allow $c = 2$. For $c = 1$, however, these numbers are only 65 % and 13 %, respectively.

Moreover, the lower bounds in Table 1 show that we cannot benefit from increasing c to a value larger than 2 for the given dimensions: In dimension 2, the lower bound of 3 inequalities is met by all non-trivial sets X , i.e., all sets admit an integer formulation with three inequalities if $c = 2$. In dimensions 3 and 4, there are two sets which need at least 5 and 8 inequalities, respectively. If we allow $c = 2$, we can describe both sets with 5 and 8 inequalities, respectively. Thus, increasing c cannot reduce the maximum number of inequalities needed in the worst case any further.

Furthermore, Table 1 shows that integer formulations of a point set X can be much smaller than complete linear descriptions of $\text{conv}(X)$. Note, for example, that for a coefficient bound of $c = 2$, every non-trivial set in dimension 2, 3, and 4 can be described by 3, 5, and 8 inequalities, respectively. Using a complete linear description needs, however, up to 4, 8, and 16 inequalities¹, respectively. Thus, if complete linear descriptions are large

¹If $\text{conv}(X)$ is not full-dimensional, the complete linear description contains also equations. In the

Table 1: Statistics on the number of nonempty sets $X \subseteq \{0, 1\}^n$ with an integer formulation with bounded absolute values of left-hand side coefficients ($c = 1$ and $c = 2$, respectively) such that the given number of inequalities is both necessary and sufficient for an integer formulation in comparison with lower bounds $\ell(X)$ on the size of any integer formulation and number on inequalities in a complete linear description

| $c = 1:$ | | | | | | | | | | |
|----------------|------------------------|-----|--------|--------|--------|------|--------|------|--------|----|
| | number of inequalities | | | | | | | | | |
| dimension | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11–15 | 16 |
| 2 | 12 | 2 | | | | | | | | |
| 3 | | 164 | 82 | 8 | | | | | | |
| 4 | | | 8388 | 29 876 | 20 208 | 6598 | 400 | 64 | | |
| $c = 2:$ | | | | | | | | | | |
| | number of inequalities | | | | | | | | | |
| dimension | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11–15 | 16 |
| 2 | 14 | | | | | | | | | |
| 3 | | 242 | 12 | | | | | | | |
| 4 | | | 57 248 | 8096 | 188 | 2 | | | | |
| $\ell(X):$ | | | | | | | | | | |
| | number of inequalities | | | | | | | | | |
| dimension | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11–15 | 16 |
| 2 | 14 | | | | | | | | | |
| 3 | | 252 | 2 | | | | | | | |
| 4 | | | 65 460 | 56 | 16 | 2 | | | | |
| compl. descr.: | | | | | | | | | | |
| | number of inequalities | | | | | | | | | |
| dimension | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11–15 | 16 |
| 2 | 4 | 10 | | | | | | | | |
| 3 | | 58 | 116 | 56 | 20 | 4 | | | | |
| 4 | | | 3008 | 9280 | 4984 | 9872 | 10 640 | 3616 | 24 076 | 58 |

and their separation problem is hard, there still might exist a much smaller formulation of X via an integer formulation.

Finally, we remark that, at least in theory, the above approach can be used to obtain statistics in higher dimensions. But since there exist $2^{2^n} - 2$ non-trivial subsets of $\{0, 1\}^n$, we would have to solve at least $4.29 \cdot 10^9$ MIPs for $n = 5$, which is impractical. Of course, one can reduce the number of treated subsets X by exploiting symmetries between different point configurations. But it is likely that the number of considered sets is still too large.

4 Sparse Formulations

As a motivation for sparse formulations consider a 0/1-knapsack problem for some $a \in \mathbb{Z}_{>0}^n$ and $\beta \in \mathbb{Z}_{>0}$ with $X^{a,\beta}$ as defined in (3). This formulation seems to be easy to handle, since it consists of a single constraint. But is there also a *sparse* description of X ? In this section, we analyze the existence of formulations with sparse constraints in the original space. We call an inequality of the form $a^\top x \leq \beta$ s -sparse if at most s entries of a are non-zero. Thus, the sparser an inequality is the smaller is s . We define the *sparsity function* σ for an n -dimensional vector a by $\sigma(a) := |\{a_i \neq 0 : i \in [n]\}|$. For a matrix A we define $\sigma(A) := \max\{\sigma(a) : a \text{ is a row of } A\}$, and finally, the *minimum sparsity* of a mixed-integer set X is defined by $\sigma(X) := \min_{Ax \leq b} \{\sigma(A) : Ax \leq b \text{ is a mixed-integer formulation of } X\}$. In particular, the minimum sparsity of X is s , if X admits a mixed-integer formulation by s -sparse inequalities but not by $(s - 1)$ -sparse inequalities.

In the literature, sparsity was discussed, e.g., by Dey et al. [6], who studied the approximation of polyhedra by sparse cutting planes, and the same authors also investigated the impact of sparse cutting planes for sparse integer programs, see Dey et al. [7].

4.1 Characterization of Minimum Sparsity

To derive lower bounds on the number of non-zero coefficients needed in an integer formulation of a given set $X \subseteq \{0, 1\}^n$, let $\bar{x} \in \{0, 1\}^n$ and let $N(\bar{x})$ be the neighbors of \bar{x} in the 0/1-cube, i.e.,

$$N(\bar{x}) := \{x \in \{0, 1\}^n : \|x - \bar{x}\|_1 = 1\}.$$

That is, the neighbors of \bar{x} are those 0/1-points that differ from \bar{x} in exactly one coordinate.

Lemma 11. *Let $X \subseteq \{0, 1\}^n$ be nonempty and $\bar{x} \in \{0, 1\}^n \setminus X$. Then there does not exist an integer formulation for X in the original space with at most $|N(\bar{x}) \cap X| - 1$ non-zeros on the left-hand side of each defining inequality, that is, $\sigma(X) > |N(\bar{x}) \cap X| - 1$.*

Proof. Define $s = |N(\bar{x}) \cap X|$ and assume that $s > 1$, since otherwise the statement is trivial. To prove the assertion, we assume for the sake of contradiction that X can be represented via $X = P \cap \mathbb{Z}^n$ with $P = \{x : Ax \leq b\}$, where each row of A has at most $s - 1$ non-zeros. Since $\bar{x} \notin X$, there exists some inequality $a^\top x \leq \beta$ of $Ax \leq b$ such that $a^\top \bar{x} > \beta$, but $a^\top \tilde{x} \leq \beta$ for all $\tilde{x} \in N(\bar{x}) \cap X$.

Assume w.l.o.g. that $a_1 = a_2 = \dots = a_{n-s+1} = 0$. Thus, the remaining $s - 1$ entries a_{n-s+2}, \dots, a_n might have nonzero coefficients. Since $|N(\bar{x}) \cap X| = s$, there exists $\tilde{x} \in N(\bar{x}) \cap X$ such that $\bar{x} = \tilde{x}$ except for some $i = 1, \dots, n - s + 1$ with $\bar{x}_i \neq \tilde{x}_i$. But then $\beta < a^\top \bar{x} = a^\top \tilde{x} \leq \beta$, a contradiction. \square

Next, we demonstrate applications of Lemma 11.

statistics, we replaced equations $a^\top x = \beta$ by inequalities $a^\top x \leq \beta$ and $a^\top x \geq \beta$, and thus, counted them twice.

Example 12. The parity polytopes P_e and P_o are the convex hulls of binary points that contain an even and odd number of 1-entries, respectively. Observe that each neighbor of a vertex of P_e is contained in P_o and vice versa. Hence, Lemma 11 implies that both P_e and P_o need completely dense inequalities in any integer formulation in the original space. Note that Jeroslow [13] has shown that there does not exist an integer formulation for the vertices of the parity polytopes in the original space with less than exponentially many inequalities.

The next example will also show that the sparsity bound derived in Lemma 11 can be tight for every sparsity level.

Example 13. Consider the set $X_s := \{x \in \{0, 1\}^n : \mathbf{1}^\top x \leq s\}$, where $s \in [n - 1]$. The set X_s is the feasible region of a 0/1-knapsack problem, and we claim that $\sigma(X_s) = s + 1$.

Let $\bar{x} \in \{0, 1\}^n$ such that $\mathbf{1}^\top \bar{x} = s + 1$. Then $\bar{x} \notin X$ and there are $s + 1$ neighbors of \bar{x} in X . Hence, Lemma 11 implies that X cannot be represented by an integer formulation that consists of s -sparse inequalities only. But X_s admits an $(s + 1)$ -sparse integer formulation via box constraints and

$$\sum_{i \in I} x_i \leq s, \quad I \in \binom{[n]}{s + 1}.$$

Lemma 11 allows to derive bounds on the sparsity of any integer formulation of a set $X \subseteq \{0, 1\}^n$ by a simple neighborhood argument. But unfortunately, its bound need not be tight in every case. To be able to compute the minimum sparsity of any integer formulation, we introduce the concept of infeasible face coverings.

Definition 14. Let $X \subseteq \{0, 1\}^n$. A face F of $[0, 1]^n$ is called *infeasible w.r.t. X* if every integer point in F is not contained in X . A collection \mathcal{F} of infeasible faces of $[0, 1]^n$ w.r.t. X is called an *infeasible face collection* of X . Moreover, if \mathcal{F} is an infeasible face collection and for every $x \in \{0, 1\}^n \setminus X$ there exists $F \in \mathcal{F}$ such that $x \in F$, then \mathcal{F} is called an *infeasible face covering* of X . An infeasible face covering \mathcal{F} of X is called *maximal* if for each face $F \in \mathcal{F}$ there is no infeasible face F' of $[0, 1]^n$ with $F \subsetneq F'$. An infeasible face covering \mathcal{F} is called *irredundant* if for each $F \in \mathcal{F}$ there is no $F' \in \mathcal{F}$ with $F \subsetneq F'$.

Lemma 15. *Every $X \subseteq \{0, 1\}^n$ has a unique maximal irredundant infeasible face covering.*

Proof. Let \mathcal{F} be the collection of all infeasible faces of $[0, 1]^n$ w.r.t. X . The collection \mathcal{F} turns into a poset if we order the faces in \mathcal{F} w.r.t. inclusion. Obviously, the \subseteq -maximal elements in \mathcal{F} form a maximal irredundant infeasible face covering $\hat{\mathcal{F}}$ of X . To show uniqueness, assume there exists a maximal infeasible face covering $\mathcal{F}' \neq \hat{\mathcal{F}}$. Then if \mathcal{F}' contains a face $F' \notin \hat{\mathcal{F}}$ which is not a maximal infeasible face, this is a contradiction to the maximality or irredundancy of \mathcal{F}' . If there exists a face $\hat{F} \in \hat{\mathcal{F}}$ with $\hat{F} \notin \mathcal{F}'$, then this contradicts the maximality, the irredundancy, or the covering requirement on $\hat{\mathcal{F}}$. \square

We call the smallest dimension of a face in the unique maximal irredundant infeasible face covering of $X \subseteq \{0, 1\}^n$ the *maximal irredundant covering bound* of X , abbreviated as $\text{micb}(X)$. The number $\text{micb}(X)$ can alternatively be described as the smallest dimension of a face of $[0, 1]^n$ that is disjoint from X and is not contained in any larger face with the same property. In the following, we show that $\text{micb}(X)$ completely characterizes the sparsity of any irredundant integer formulation of X . If F and F' are faces of $[0, 1]^n$, we denote by $F^{\min}(F, F')$ the smallest face of $[0, 1]^n$ that contains both F and F' . If F' is a vertex x , we write $F^{\min}(F, x)$. Note that $F^{\min}(F, F')$ can also be written as $F \vee F'$, the *join* of these two faces in the face lattice of $[0, 1]^n$.

To show the main result characterizing the minimum sparsity in Theorem 18, we first need some technical lemmata.

Lemma 16. *Let $\bar{x} \in \{0, 1\}^n$ and let F be a face of $[0, 1]^n$. Moreover, let $a^\top x \leq \beta$ be an inequality valid for $[0, 1]^n$ with $a^\top \bar{x} = \beta$. If all binary points in $\{\bar{x}\} \cup F$ violate the inequality $a^\top x \leq \beta'$, $\beta' \in \mathbb{R}$, then all binary points in $F^{\min}(F, \bar{x})$ violate $a^\top x \leq \beta'$.*

Proof. The proof proceeds via induction on n . If $n = 1$, the statement is trivially fulfilled. In the induction step we distinguish two cases. First, if F and \bar{x} are contained in a facet of $[0, 1]^n$, the induction hypothesis can be used to show the statement, since each facet is an $(n - 1)$ -dimensional cube. In the second case, F and \bar{x} are not contained in a facet of $[0, 1]^n$. This implies the existence of a vector $y \in F \cap \{0, 1\}^n$ such that \bar{x} and y lie on a diagonal of $[0, 1]^n$. W.l.o.g. we can assume (by complementing entries) that $\bar{x} = \mathbf{1}$ and $y = 0$.

Since $a^\top \bar{x} = \beta$ by assumption, it follows that $a^\top \bar{x} = a^\top \mathbf{1} = \sum_{i=1}^n a_i = \beta$. Thus, $a_i \geq 0$ holds for all $i \in [n]$, because $a^\top x \leq \beta$ is valid for $[0, 1]^n$. This means that $a^\top y \leq a^\top x' \leq a^\top \bar{x}$ for all $x' \in \{0, 1\}^n$. Since $y \in F$ and F violates $a^\top x \leq \beta'$ by the assumption, the statement follows by $\beta' < a^\top y \leq a^\top x' \leq a^\top \bar{x}$ for all $x' \in \{0, 1\}^n$. \square

Now we are able to derive bounds on the sparsity of an integer formulation for a set $X \subseteq \{0, 1\}^n$.

Lemma 17. *Let F be a face of $[0, 1]^n$ and let $a^\top x \leq \beta$ be an inequality with $a^\top \bar{x} > \beta$ for all $\bar{x} \in F$. If $a^\top x \leq \beta$ does not completely cut off a face F' of $[0, 1]^n$ with $F \subsetneq F'$, then the vector a has at least $n - \dim(F)$ non-zero entries. Furthermore, there is an inequality $\bar{a}^\top x \leq \bar{\beta}$ such that the binary points that violate this inequality are exactly the binary points in F and such that \bar{a} has exactly $n - \dim(F)$ non-zero entries.*

Proof. For a face \bar{F} of $[0, 1]^n$

$$\begin{aligned} C_{\bar{F}} &= \{w \in \mathbb{R}^n : w \text{ induces the face } \bar{F} \text{ of } [0, 1]^n\} \\ &= \{w \in \mathbb{R}^n : w_i > 0 \text{ if } \bar{x}_i = 1 \text{ for all } \bar{x} \in \bar{F}, \\ &\quad w_i < 0 \text{ if } \bar{x}_i = 0 \text{ for all } \bar{x} \in \bar{F}, \\ &\quad w_i = 0 \text{ otherwise}\} \end{aligned}$$

denotes the (open) normal cone of \bar{F} . Let $a^\top x \leq \beta$ be violated by all points in F and assume that it does not completely cut off a face F' of $[0, 1]^n$ with $F \subsetneq F'$. Then a is contained in $\mathfrak{C}_F := \bigcup_{\bar{F} \subsetneq F} C_{\bar{F}}$, since otherwise, $a^\top x \leq \beta$ has to cut off a binary point $\bar{x} \notin F$. By Lemma 16, inequality $a^\top x \leq \beta$ has to cut off $F^{\min}(F, \bar{x})$, which contradicts the assumption. Because $a \in \mathfrak{C}_F$, it has at least $n - \dim(F)$ non-zero entries due to the structure of \mathfrak{C}_F .

To prove the second part of the lemma, let $\bar{a} \in C_F$ and define

$$\bar{\beta} = \max \{\bar{a}^\top x : x \in [0, 1]^n\} \quad \text{as well as} \quad \beta' = \max \{\bar{a}^\top x : x \in \{0, 1\}^n \setminus F\}.$$

Since $\bar{a} \in C_F$, $\arg\max \{\bar{a}^\top x : x \in [0, 1]^n\} = F$ and thus $\bar{\beta} > \beta'$. Consequently, $\bar{a}^\top x \leq \beta'$ is an inequality which is violated by exactly those binary points that are contained in F . Furthermore, \bar{a} has exactly $n - \dim(F)$ non-zero entries by the structure of the normal cone C_F . \square

This allows us to completely characterize the minimum sparsity of an integer formulation of any non-trivial $X \subseteq \{0, 1\}^n$.

Theorem 18. *For a non-trivial set $X \subseteq \{0, 1\}^n$ it holds that $\sigma(X) = n - \text{micb}(X)$.*

Proof. To prove the assertion, we first show that each integer formulation of X contains an inequality with at least $n - \text{micb}(X)$ non-zero entries. Afterwards, we construct an integer formulation all of whose inequalities are $(n - \text{micb}(X))$ -sparse to prove that this bound is tight.

To show the first part, let $a^\top x \leq \beta$ be an inequality valid for $\text{conv}(X)$ that cuts off at least one point $\bar{x} \in \{0, 1\}^n \setminus X$ and denote the set of all such a by \mathcal{A} . Consider the face

$$F_a := \{x \in [0, 1]^n : x_i = 1 \text{ if } a_i > 0 \text{ and } x_i = 0 \text{ if } a_i < 0, i \in [n]\}$$

of $[0, 1]^n$, which coincides with $\text{argmax}\{a^\top x : x \in [0, 1]^n\}$. Note that $\sigma(a) = n - \dim(F_a)$.

Since $a^\top x \leq \beta$ separates \bar{x} from $\text{conv}(X)$, it separates at least one point in F_a . Hence, all points in F_a are cut off. Furthermore, let F be a maximal face of $[0, 1]^n$ which is completely cut off by $a^\top x \leq \beta$. On the one hand, if $F_a \subseteq F$, then

$$\dim(F) \geq \dim(F_a). \quad (8)$$

On the other hand, if $F_a \not\subseteq F$, then $a^\top x \leq \beta$ cuts off $F^{\min}(F, F_a)$ by applying Lemma 16 to all points of F_a . Thus, all points in $F^{\min}(F, F_a)$ have to be infeasible, which is a contradiction to the maximality of F . Consequently, all maximal infeasible faces of $[0, 1]^n$ that are cut off by $a^\top x \leq \beta$ contain F_a and fulfill (8).

Let F^* denote a maximal infeasible face of $[0, 1]^n$ of minimal dimension. Then,

$$\max_{a \in \mathcal{A}} \sigma(a) = \max_{a \in \mathcal{A}} (n - \dim(F_a)) \stackrel{(8)}{\geq} \max_{a \in \mathcal{A}} (n - \dim(F^*)) = n - \text{micb}(X),$$

which proves the proposed lower bound.

To prove the theorem, it suffices to construct an integer formulation all of whose inequalities are $(n - \text{micb}(X))$ -sparse. Let \mathcal{F} be the maximal irredundant infeasible face covering of X . To construct the desired integer formulation, we take the box constraints $x \in [0, 1]^n$ to guarantee that each feasible point is contained in the hypercube. Note that each such constraint is 1-sparse. Now it suffices to cut off the infeasible faces of $[0, 1]^n$ w.r.t. X . By Lemma 17, each face $F \in \mathcal{F}$ can be cut off by an inequality that has exactly $n - \dim(F)$ non-zero left-hand side coefficients. If we take these inequalities for all faces in \mathcal{F} , we have ensured that we cut off each point in $\{0, 1\}^n \setminus X$ since \mathcal{F} is a maximal infeasible face covering. These inequalities are $(n - \text{micb}(X))$ -sparse, which proves the assertion. \square

Remark 19. *Since the normal cone of any face of $[0, 1]^n$ contains a vector in $\{0, \pm 1\}^n$, there always exists a sparsest integer formulation of any non-trivial set $X \subseteq \{0, 1\}^n$ all of whose left-hand side coefficients are contained in $\{0, \pm 1\}$.*

We now apply Theorem 18 to investigate whether there exist sparse formulations for several combinatorial optimization problems. First, we consider the solution set of the *traveling salesman problem* (TSP), which is to find a weight minimal Hamiltonian cycle in an undirected graph.

Theorem 20. *Let $K_n = (V, E)$ be the complete undirected graph with $n \geq 5$ nodes and let $X \subseteq \{0, 1\}^E$ be the set of incidence vectors of Hamiltonian cycles of K_n . It holds that $\sigma(X) = n - 2$.*

Proof. To be able to apply Theorem 18, we have to analyze the binary points that are not contained in X . To this end, we construct for each infeasible binary point x a certificate, i.e., fixings of variables that ensure infeasibility of x , of minimum size. The following situations cover all possible cases of infeasibility: The graph G' induced by edges $e \in E$ with $x_e = 1$

- contains a subgraph isomorphic to $K_{1,3}$,
- contains a connected component that is a path graph,
- is the empty graph,
- contains an induced subgraph which is a cycle of length less than $|V|$.

In the first case, a certificate of infeasibility is given by fixing the three edges of $K_{1,3}$ to 1, because this implies that there is a node in a solution with degree 3. In the second and third case, there is a node v of G' whose degree is at most 1. Thus, fixing $n - 2$ pairwise different edges incident to v to 0, we obtain a certificate of infeasibility of size $n - 2$, since this enforces that v has at most degree 1 in any solution. Finally, in the fourth case, every node has degree 2 and if the solution is infeasible, a subtour contains at most $n - 3$ nodes. Fixing the corresponding $n - 3$ edges to 1 generates a certificate of infeasibility.

Thus, for each infeasible $x \in \{0, 1\}^E$, there exists a certificate of infeasibility of size at most $n - 2$. Therefore, each infeasible point is contained in an infeasible face of $[0, 1]^n$ of dimension at least $|E| - n + 2$. To be able to apply Theorem 18, we have to show that there is indeed an infeasible binary point for which no certificate of size less than $n - 2$ exists. To see this, consider the infeasible point $x = 0$. If at most $n - 3$ variables are fixed to 0, a Hamiltonian cycle exists on the remaining edges by Ore's Theorem [23], which guarantees the existence of a Hamiltonian cycle if for every pair of distinct non-adjacent nodes the sum of their degrees is at least the number of nodes. Consequently, a minimum size certificate of infeasibility for the zero vector has size $n - 2$, and Theorem 18 implies the assertion. \square

Note that the minimum sparsity of the TSP for $n = 4$ is 3, because the certificate for a $K_{1,3}$ subgraph has size 3, which is larger than the variable bound $n - 2$. For $n = 3$, however, Theorem 20 holds, since K_3 does not contain a $K_{1,3}$ subgraph.

The technique used in the proof of Theorem 20 can be used to show that other formulations are as sparse as possible. For further examples we state the following lemma.

Lemma 21. *Let $G = (V, E)$ be a graph and let $\delta(S) \subseteq E$ denote the cut induced by $S \subseteq V$. If the sets X_C , X_M , and X_k of incidence vectors of connected subgraphs, perfect matchings, and k -colorings, respectively, are non-empty, the following sparsity values hold*

$$\begin{aligned} \sigma(X_C) &= \max_{S \subseteq V} |\delta(S)|, & X_C &:= \{x \in \{0, 1\}^E : x(\delta(S)) \geq 1, \emptyset \neq S \subseteq V\}, \\ \sigma(X_M) &= \max_{v \in V} |\delta(\{v\})|, & X_M &:= \{x \in \{0, 1\}^E : x(\delta(\{v\})) = 1, v \in V\}, \\ \sigma(X_k) &= k, & X_k &:= \{x \in \{0, 1\}^{V \times [k]} : x_{ui} + x_{vi} \leq 1, i \in [k], \{u, v\} \in E, \\ & & & \sum_{i=1}^k x_{vi} = 1, v \in V\}. \end{aligned}$$

Proof. As for the proof of Theorem 20, we will state infeasible points for which no certificate of smaller size exists. Further details will be omitted. For X_C consider the point $x(\delta(S)) = 0$ with some maximal cut $\delta(S)$, for X_M consider $x(\delta(\{v\})) = 0$ with some node $v \in V$ of maximal degree and finally for X_k consider $x_{vi} = 0$ for some node $v \in V$ and all $i \in [k]$. \square

Finally, we show that it is hard to compute the minimum sparsity of an integer formulation or to separate inequalities of such formulations.

Theorem 22. *It is \mathcal{NP} -hard to determine $\sigma(X)$ for $X \subseteq \{0, 1\}^n$, even if X corresponds to the independent sets of a graphic matroid.*

Proof. Let $G = (V, E)$ be an undirected simple graph and let $X \subseteq \{0, 1\}^E$ correspond to the independent sets of a graphic matroid of G , i.e., X contains the incidence vectors of

Table 2: Number of non-trivial subsets of $\{0, 1\}^n$ that admit an integer formulation with a given minimum sparsity

| dimension | minimum sparsity | | | |
|-----------|------------------|------|--------|--------|
| | 1 | 2 | 3 | 4 |
| 2 | 8 | 6 | | |
| 3 | 26 | 138 | 90 | |
| 4 | 80 | 4088 | 38 976 | 22 390 |

cycle free edge sets in G . Let $\bar{x} \in \{0, 1\}^E \setminus X$, and let F be an infeasible face of $[0, 1]^E$. If \bar{x} is the incidence vector of an induced cycle C of G ,

$$F_C := \text{conv}(\{x \in \{0, 1\}^E : x_e = 1 \text{ for all } e \in E \text{ with } \bar{x}_e = 1\})$$

is a face of $[0, 1]^E$ that contains only infeasible binary points including \bar{x} . All the faces F_C , where C is an induced cycle of G form the maximal irredundant infeasible face covering of X , since every infeasible point contains an induced cycle. If a point does not contain any induced cycle, it is contained in the incidence vector of a tree and is therefore feasible. Thus, the dimension of each maximal infeasible face of $[0, 1]^E$ that contains \bar{x} is at least $|E| - |C|$ for the longest induced cycle C encoded by \bar{x} . Hence, computing the minimum sparsity of an integer formulation of X is equivalent to computing the length of a longest induced cycle in G by Theorem 18 and the definition of F_C . Computing the maximum length of an induced cycle is \mathcal{NP} -hard, see Garey and Johnson [9, Problem GT23]. \square

In general, formulations with minimum sparsity are \mathcal{NP} -hard to separate as the following example of a knapsack set shows.

Example 23. Let X be the feasible region of a binary knapsack problem. Similar to the proof of Theorem 22, one can show that the faces

$$F_C := \text{conv}(\{x \in \{0, 1\}^n : x_i = 1 \text{ for all } i \in C\}),$$

where C is a minimal cover of X , form the maximal irredundant infeasible face covering of X . Since the sparsity of the cover inequality for a minimal cover C equals $n - \dim(F_C)$, an integer formulation with minimum sparsity of X is given by box constraints and all minimal cover inequalities for X . Separating this integer formulation is (weakly) \mathcal{NP} -hard, see Klabjan et al. [19].

4.2 Statistics on the Minimum Sparsity of Integer Formulations

The above results do not provide quantitative information for sparse formulations. How is the minimum sparsity of integer formulations distributed for the different sets non-trivial sets $X \subseteq \{0, 1\}^n$ in small dimensions? The aim of this section is to give an impression of this distribution. To be able to determine $\sigma(X)$ for a given X , we computed by a brute force method its maximal irredundant infeasible face covering, which characterizes the minimum sparsity by Theorem 18.

Table 2 summarizes our results. Since we were only able to compute the statistics for very small dimensions, an interpretation of these numbers has to be treated carefully. Our results indicate that there are only few non-trivial subsets of $\{0, 1\}^n$ that admit a very sparse integer formulation. From a combinatorial point of view, this is reasonable

because the sets $X \subseteq \{0, 1\}^n$ that can be represented by 1-sparse inequalities are exactly the $3^n - 1$ non-trivial faces of $[0, 1]^n$. In contrast to this, it is much more likely that an integer formulation of X has to be dense, since this depends only on local properties of infeasible points. For example, the sets X that need the densest integer formulations are those for which there exists $\bar{x} \in \{0, 1\}^n \setminus X$ which has only feasible neighbors in $[0, 1]^n$. Consequently, only few sets $X \subseteq \{0, 1\}^n$ admit sparse integer formulations.

5 Mixed-Integer Formulations Including Objective Functions

Most of the results in the preceding sections or in the literature do not mention the objective function. For example, the geometric characterization of Jeroslow and Lowe [14] of subsets of \mathbb{R}^n that can be represented as the projection of mixed-integer points in a polyhedron is independent from an objective function. In this section we address this issue.

Let Π be an optimization problem of the form

$$\max \{f(x) : x \in X\}, \quad (9)$$

where $X \subseteq \mathbb{Z}^p \times \mathbb{R}^q$ is nonempty and hole-free and $f: X \rightarrow \mathbb{R}$ has a finite maximum value. The aim of this section is to represent Π via a mixed-integer programming formulation. Observe that since we allow general objective functions, we cannot use the definition from Section 2, which required the objective function to be linear. Thus, we first adapt the definition of a mixed-integer formulation.

Definition 24. A *mixed-integer programming formulation (MIP formulation)* for an instance I of Π is a triple (Q, g, τ) with a polyhedron $Q \subseteq \mathbb{R}^{p'} \times \mathbb{R}^{q'}$, an affine function $g: \mathbb{R}^{p'} \times \mathbb{R}^{q'} \rightarrow \mathbb{R}$ and $\tau: \mathbb{R}^{p'} \times \mathbb{R}^{q'} \rightarrow \mathbb{R}^p \times \mathbb{R}^q$ such that

- a) (Q, g) is a feasible MIP with finite maximum value,
- b) $\tau(\operatorname{argmax} \{g(y) : y \in Q \cap (\mathbb{Z}^{p'} \times \mathbb{R}^{q'})\}) \subseteq \operatorname{argmax} \{f(x) : x \in X\}$.

Assume we are given an optimization problem Π as in (9). If there exists a MIP formulation of an instance I , we can solve I by solving the maximization problem of g over the mixed-integer set $Q \cap (\mathbb{Z}^{p'} \times \mathbb{R}^{q'})$ and applying the transformation τ , which by Part b) of Definition 24 is an optimal solution of I . Therefore, if there exists a MIP formulation of (9), the optimization problem Π can be solved by a MIP.

Observe that we cannot expect to find a MIP formulation (Q, g, τ) in the original space for every problem Π if τ is the identity. For example, assume that f has a unique integral maximizer x^* in the relative interior of $\operatorname{conv}(X)$. If $(Q, g, \operatorname{id})$ would be a MIP formulation of Π in the original space with $Q \cap X = X$, there would exist an optimal point on the boundary of $\operatorname{conv}(X)$, a contradiction to Property b) of Definition 24. If $X \subseteq \{0, 1\}^n$, however, this problem cannot occur since binary sets do not have relative interior integer points.

Another problem of Definition 24 is that such a formulation always exists, but might not be useful. It is clear that Q , g and τ can be chosen easily once an optimal value of (9) is known. For example, let an optimum of (9) be attained at x^* . In this case, setting g to be constant and τ mapping everything to x^* , Q can be any nonempty polyhedron, in particular it can be chosen as a single point. Since solving (9) in order to obtain x^* might be \mathcal{NP} -hard, we cannot expect that such a MIP formulation can be computed efficiently. Thus, we refine the definition of a MIP formulation as follows.

Definition 25. Let I be an instance of an optimization problem with objective f . A MIP formulation (Q, g, τ) of I is called *efficient* if

1. Q , g , and τ can be generated in polynomial time in $\langle I \rangle$,
 2. $g(y)$ can be evaluated in polynomial time in $\langle I \rangle$ for all $y \in Q$,
 3. $\tau(y)$ can be evaluated in polynomial time in $\langle I \rangle$ for all $y \in Q \cap (\mathbb{Z}^{p'} \times \mathbb{R}^{q'})$,
- where $\langle \cdot \rangle$ denotes the encoding length (see for example Grötschel et al. [11]).

Note that Theorem 3 shows that an efficient integer programming formulation always exists if f is linear, $X \subseteq \{0, 1\}^n$, and the membership problem is in \mathcal{NP} . Furthermore, note that it is necessary that g can be computed efficiently, since otherwise once an optimal solution is known, g can easily be chosen to be only optimal in that particular solution.

Not surprisingly, efficient MIP formulations cannot exist in certain complexity situations.

Lemma 26. *Let I be an instance of an \mathcal{NP} -hard problem Π , and let (Q, g, τ) be a MIP formulation of I (either in the original or an extended space) such that*

$$\max \{g^\top x : x \in Q \cap (\mathbb{Z}^{p'} \times \mathbb{R}^{q'})\}$$

can be solved in polynomial time in $\langle I \rangle$. As long as $\mathcal{P} \neq \mathcal{NP}$, the MIP formulation (Q, g, τ) is not efficient.

Proof. If the formulation would be efficient, an optimal solution of I can be computed in polynomial time. \square

As an example for the application of Lemma 26, we consider the following optimization problems: the TSP, graph partitioning, multiprocessor scheduling, and maximum biclique problem. These problems do not admit an efficient integer formulation in particular spaces.

For the TSP, one possible formulation is to give every node of the underlying undirected graph a number from 1 up to n , in which case a solution is some permutation of these numbers. In particular, this means that all solutions form the vertex set of the permutahedron, see, for example, Ziegler [28]. Note that the well-known Miller-Tucker-Zemlin formulation [22] extends on the permutahedron by also using edge-variables.

Given a graph $G = (V, E)$, edge weights w , and a positive integer k , the aim of the *graph partitioning problem* is to find a partition of the nodes V into at most k subsets such that the sum of the weights of edges between two different partitions is maximized. Given a set of tasks V , a number of processors k and an integer length for each task, the *multiprocessor scheduling problem* is to find a schedule for the tasks with the earliest possible deadline such that at most k tasks are performed at the same time. For an edge or node weighted bipartite graph $G = (V, E)$, the *maximum biclique problem* on G is to find an induced complete bipartite subgraph of G whose (node or edge) weight is maximal.

To state the corollary, let $X^{V,k}$ be $\{x \in \{0, 1\}^{V \times [k]} : \sum_{i=1}^k x_{vi} = 1, v \in V\}$ and define $X^G := \{x \in \{0, 1\}^V : x = \chi^{V'}\}$, V' is the node set of biclique in G .

Corollary 27. *As long as $\mathcal{P} \neq \mathcal{NP}$ there do not exist efficient formulations for the TSP over the vertices of the permutahedron, the graph partitioning problem and multiprocessor scheduling over $X^{V,k}$ as well as the edge version of the biclique problem over X^G .*

Proof. To apply Lemma 26, we only need to show that optimization over the respective integer sets is possible in polynomial time, whereas the respective problem is \mathcal{NP} -hard. Since the permutahedron can be derived from a linear projection of the Birkhoff polytope (see, e.g., [15]) over which one can optimize in polynomial time, the statement for the TSP follows (for \mathcal{NP} -hardness see, e.g., Garey and Johnson [9, Problem ND22]). Moreover, solutions of both the graph partitioning and the multiprocessor scheduling problem can be modeled as vectors in $X^{V,k}$. Thus, since linear optimization over $X^{V,k}$ is possible in

linear time, but both problems are \mathcal{NP} -hard, see Garey and Johnson [9, Problem ND14] and [9, Problem SS8], Lemma 26 implies the non-existence of an efficient formulation. Finally, the node version of the maximum biclique problem is solvable in polynomial time, see Yannakakis [27, Theorem 4]. Thus, a linear objective over X^G can be maximized in polynomial time. This proves the assertion, because the edge version of the biclique problem is \mathcal{NP} -hard, see Dawande et al. [5]. \square

Remark 28. *From a practical point of view, there are additional requirements that seem useful for the general notion of a MIP formulation.*

1. *If a MIP formulation is known and f is changed, it would be preferable if Q and τ can be chosen as before and only g has to be changed, that is, Q and τ are independent of f .*
2. *Q should be tractable.*
3. *Part b) in Definition 24 could be changed to equality, i.e., all optimal solutions can be generated.*

The first part can be guaranteed in the case that f and τ are affine, since if $f(x) = f^\top x$ and $\tau(y) = \tau y$, for $f \in \mathbb{R}^{p+q}$, $\tau \in \mathbb{R}^{(p'+q') \times (p+q)}$, we obtain:

$$\max_{x \in X} f(x) = \max_{y \in Y} f(\tau(y)) = \max_{y \in Y} f^\top \tau y = \max_{y \in Y} (\tau^\top f)^\top y = \max_{y \in Y} (\tau^\top f)(y) =: \max_{y \in Y} g(y),$$

where $Y := Q \cap (\mathbb{Z}^p \times \mathbb{R}^q)$. The first equality holds because of Definition 24 Condition b) and the others because of linearity.

In case of Requirement 3, assume that τ is an affine map and f and g are constant functions. Then the MIP formulation (Q, g, τ) is an extended mixed-integer formulation of X . In particular, if τ is the identity, (Q, g, τ) is a mixed-integer formulation in the original space. Thus, Definition 24 generalizes the concept of (extended) mixed-integer formulations of Definition 2 if equality holds in Part b).

To further expand on this requirement, we consider the following example.

Example 29. The graph partitioning problem can be formulated using binary variables x_{vi} modeling whether node v is in partition i by the following quadratic optimization problem

$$\max \left\{ \sum_{\{u,v\} \in E} w_{uv} \left(1 - \sum_{i=1}^k x_{ui} x_{vi} \right) : x \in X^{V,k} \right\}. \quad (10)$$

In the following, we refer to this problem as the *quadratic graph partitioning formulation*.

One problem with Formulation (10) is its quadratic objective function. It could be linearized by introducing a binary variable y_{uv} modeling whether the endpoints of an edge $\{u, v\} \in E$ lie in different partitions. An integer linear program is then given by

$$\begin{aligned} \max \quad & \sum_{\{u,v\} \in E} w_{uv} y_{uv} \\ \text{s.t.} \quad & \sum_{i=1}^k x_{vi} = 1, \quad v \in V, \\ & x_{ui} - x_{vi} \leq y_{uv}, \quad \{u, v\} \in E, i \in [k], \\ & x_{vi} - x_{ui} \leq y_{uv}, \quad \{u, v\} \in E, i \in [k], \\ & x_{ui} + x_{vi} + y_{uv} \leq 2, \quad \{u, v\} \in E, i \in [k], \\ & x_{vi} \in \{0, 1\}, \quad v \in V, i \in [k], \\ & y_{uv} \in \{0, 1\}, \quad \{u, v\} \in E. \end{aligned} \quad (11)$$

Of course there are various other formulations, see, e.g., Ales and Knippel [1], but the details are not needed here. Formulation (11) is an example of an efficient MIP formulation. Note that this formulation is given in an extended space, if we consider the original space to be given by the x -variables alone. This shows that there exists a MIP formulation of the graph partitioning problem in an extended space such that Condition b) of Definition 24 holds with equality. A question that arises is whether equality can also be achieved in the original space. Intuitively this should not be the case, because of the objective function. In the following, we show that this indeed is not possible.

Note that the following lemmata are independent of \mathcal{P} vs. \mathcal{NP} and state the non-existence of an objective function independent of the time of its calculation in difference to Lemma 26. For these lemmata we use the following simple observation and define $P^{V,k} := \text{conv}(X^{V,k})$.

Observation 30. *Given a polytope P and a linear objective function w . If two points $x, y \in P$ are both optimal with respect to w , they have to lie in a common face of P .*

Lemma 31. *There does not exist a linear objective $w \in \mathbb{R}^{V \times [k]}$ for $P^{V,k}$ such that the maximizers of w over $P^{V,k}$ are exactly the maximizers of the quadratic graph partitioning problem if $k \geq 2$ and not all solutions of the quadratic problem are optimal.*

Proof. For the sake of contradiction, assume there is a non-trivial face F of $P^{V,k}$ that contains all the maximizers of the quadratic graph partitioning problem. Let $a^\top x \leq \beta$ be an inequality that induces F . Since exactly the points in F maximize $a^\top x \leq \beta$, we have $a^\top x = \beta$ for all $x \in F$, and $a^\top x < \beta$ for all $x \in P^{V,k} \setminus F$.

Observe that $x \in \{0, 1\}^{V \times [k]}$ is a vertex of $P^{V,k}$ if and only if for every $v \in V$ there is exactly one $i(v) \in [k]$ with $x_{v,i(v)} = 1$. Hence, a vertex is contained in F if and only if $i(v) \in \mathcal{A}^v := \text{argmax} \{a_{vi} : i \in [k]\}$ for every $v \in V$.

Let x be a vertex of $P^{V,k}$ that is contained in F . If there exists $\tilde{v} \in V$ such that $\mathcal{A}^{\tilde{v}} \neq [k]$, there exists an index $j \in \mathcal{A}^{\tilde{v}}$ and an index $\tilde{j} \in [k] \setminus \mathcal{A}^{\tilde{v}}$. By exchanging the entries of x in columns j and \tilde{j} , we obtain another vertex \tilde{x} of $P^{V,k}$ that is optimal for the quadratic graph partitioning problem, because changing the labels of assigned partitions (that is, exchanging columns) does not affect the objective value. But \tilde{x} cannot be contained in F , because for every $v \in V \setminus \{\tilde{v}\}$

$$\sum_{i=1}^k a_{vi} \tilde{x}_{vi} \leq \sum_{i=1}^k a_{vi} x_{vi} \quad \text{and} \quad \sum_{i=1}^k a_{\tilde{v}i} \tilde{x}_{\tilde{v}i} = a_{\tilde{v}\tilde{j}} < a_{\tilde{v}j} = \sum_{i=1}^k a_{\tilde{v}i} x_{\tilde{v}i}.$$

Consequently, $a^\top \tilde{x} < a^\top x$. For this reason, $\mathcal{A}^v = [k]$ needs to hold for all $v \in V$. But if $\mathcal{A}^v = [k]$ for every $v \in V$, every vertex of $P^{V,k}$ maximizes $a^\top x \leq \beta$ due to the definition of $X^{V,k}$. This contradicts the assumption that not all vertices of $P^{V,k}$ are optimal for the quadratic problem. \square

Note that this lemma does not depend on the graph partitioning problem, but only on assigning nodes to partitions with the requirement that interchanging two partitions does not change the objective value. Thus, in the original space of the x -variables there does not exist a linear objective function that allows for finding all optimal solutions.

Observation 30 can also be used to show the non-existence of a formulation of the TSP over the permutahedron (compare to Corollary 27).

Lemma 32. *Provided that not every solution of a TSP is optimal, there does not exist a linear objective function w over the permutahedron such that the maximizers of w correspond exactly to the optimal solutions of the TSP.*

Proof. Note that for any optimal solution, every cyclic shift of a vertex of the permutahedron yields exactly the same solution of the TSP. Therefore, any face that contains an optimal solution over the permutahedron also has to contain every cyclic shift. Since the faces of the permutahedron are in one-to-one correspondence to strict weak orderings on a set of n items or ordered partitions, see Ziegler [28], there does not exist any face of the permutahedron that contains both some point and a non-trivial cyclic shift of this point. \square

6 Conclusion

The size of integer formulations of combinatorial optimization problems has been studied in the literature, and criteria for the existence of tractable formulations have been derived. However, these investigations do, in general, not consider additional requirements of such formulations. For this reason, we extended these results by considering integer formulations with bounded coefficients or sparse constraints. On the one hand, we provide lower bounds on the minimum size of an integer formulation with ternary coefficients for knapsack problems. On the other hand, we showed how to determine the minimum sparsity of an integer formulation of any binary set. Various examples demonstrated the applicability of these techniques.

Since all these results are independent of an objective function, we furthermore investigated the impact of the objective on integer formulations. We gave a formal definition and discussed variations of this definition, e.g., efficient determination or preservation of all optimal solutions. For both aspects we were able to derive criteria to prove non-existence of such formulations illustrated by examples. While the first criterion is based on the assumption $\mathcal{P} \neq \mathcal{NP}$, the latter uses polyhedral arguments.

Nevertheless, there remain some interesting open questions for further research. For example, is it possible to generalize the concept of hiding sets of [17] to also include bounds on coefficients in such a way that the criterion can easily be used in theory and practice? Or can the characterization of the minimum sparsity of a formulation be extended to non-binary sets? Moreover, it is not known whether it is possible to decide in polynomial time whether a set $X \subseteq \{0, 1\}^n$ admits a hiding set of specific size. In particular, the complexity of deciding whether a given set is a hiding set is not known.

Moreover, there are further interesting aspects of MIP formulations involving the objective function, which are not covered by this paper. In Definition 24 b), we required that optimal points of a MIP formulation have to be transformable to optimal points of the original problem. Of course, we could strengthen this requirement to preserving all optimal solutions in a MIP formulation. However, Lemmata 31 and 32 show that this property may not be achievable if we fix the solution space. Thus, the question remains under which conditions all of the original optimal points are preserved in a MIP formulation. Furthermore, we have already mentioned some preferable properties of MIP formulations, e.g., tractability or invariance under objective changes (compare Remark 28). Deriving criteria that ensure these properties are out of scope of this paper, but form a possible direction of future research.

Acknowledgments We thank Tobias Fischer, Tristan Gally, Oliver Habeck, Imke Joermann, Andreas Paffenholz, Andreas M. Tillmann for fruitful discussions. We also thank an anonymous referee for valuable suggestions. The last two authors were partly supported by the German Research Foundation within the CRC 666.

References

- [1] Z. Ales and A. Knippel. An extended edge-representative formulation for the k -partitioning problem. *Electronic Notes in Discrete Mathematics*, 52(Supplement C):333 – 342, 2016. doi: 10.1016/j.endm.2016.03.044. INOC 2015 - 7th International Network Optimization Conference.
- [2] A. Basu, K. Martin, C. T. Ryan, and G. Wang. Mixed-integer linear representability, disjunctions, and variable elimination. In F. Eisenbrand and J. Koenemann, editors, *Integer Programming and Combinatorial Optimization: 19th International Conference, IPCO 2017, Waterloo, ON, Canada, June 26-28, 2017, Proceedings*, pages 75–85. Springer International Publishing, 2017. doi: 10.1007/978-3-319-59250-3_7.
- [3] G. Braun, S. Pokutta, and D. Zink. Inapproximability of combinatorial problems via small LPs and SDPs. In *Proceedings of the Forty-seventh Annual ACM Symposium on Theory of Computing*, STOC '15, pages 107–116. ACM, 2015. doi: 10.1145/2746539.2746550.
- [4] M. Conforti, G. Cornuéjols, and G. Zambelli. Extended formulations in combinatorial optimization. *Annals of Operations Research*, 204(1):97–143, 2013. doi: 10.1007/s10479-012-1269-0.
- [5] M. Dawande, P. Keskinocak, and S. Tayur. On the biclique problem in bipartite graphs. Technical Report GSIA Working Paper 1996-04, Carnegie Mellon University, 1996.
- [6] S. S. Dey, M. Molinaro, and Q. Wang. Approximating polyhedra with sparse inequalities. *Mathematical Programming*, 154(1):329–352, 2015. doi: 10.1007/s10107-015-0925-y.
- [7] 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. doi: 10.1287/moor.2017.0866.
- [8] Y. Faenza and L. Sanità. On the existence of compact ϵ -approximated formulations for knapsack in the original space. *Operations Research Letters*, 43(3):339 – 342, 2015. doi: 10.1016/j.orl.2015.04.004.
- [9] M. R. Garey and D. S. Johnson. *Computers and Intractability: A Guide to the Theory of NP-Completeness*. W. H. Freeman & Co., 1979.
- [10] F. Glover. Unit coefficient inequalities for zero-one programming. Technical report, University of Colorado, 1973.
- [11] M. Grötschel, L. Lovász, and A. Schrijver. *Geometric Algorithms and Combinatorial Optimization*, volume 2 of *Algorithms and Combinatorics*. Springer, 1988.
- [12] P. L. Hammer, E. L. Johnson, and U. N. Peled. Facet of regular 0-1 polytopes. *Mathematical Programming*, 8(1):179–206, 1975.
- [13] R. G. Jeroslow. On defining sets of vertices of the hypercube by linear inequalities. *Discrete Mathematics*, 11(2):119 – 124, 1975. doi: 10.1016/0012-365X(75)90003-5.
- [14] R. G. Jeroslow and J. K. Lowe. Modelling with integer variables. *Mathematical Programming Studies*, 22:167–184, 1984.
- [15] V. Kaibel. Extended formulations in combinatorial optimization. *Optima* 85, 2011.
- [16] 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.
- [17] V. Kaibel and S. Weltge. Lower bounds on the sizes of integer programs without additional variables. *Mathematical Programming*, 2014. doi: 10.1007/s10107-014-0855-0.
- [18] H. Kellerer, U. Pferschy, and D. Pisinger. *Knapsack problems*. Springer, 2004.

- [19] D. Klabjan, G. L. Nemhauser, and C. A. Tovey. The complexity of cover inequality separation. *Operations Research Letters*, 23:35–40, 1998.
- [20] M. Lubin, I. Zadik, and J. P. Vielma. Regularity in mixed-integer convex representability, 2017. URL <https://arxiv.org/abs/1706.05135>.
- [21] R. K. Martin, R. L. Rardin, and B. A. Campbell. Polyhedral characterization of discrete dynamic programming. *Operations Research*, 38(1):127–138, 2 1990.
- [22] C. E. Miller, A. W. Tucker, and R. A. Zemlin. Integer programming formulations and traveling salesman problems. *Journal of Association for Computing Machinery*, 7:326–329, 1960.
- [23] O. Ore. Note on Hamilton circuits. *The American Mathematical Monthly*, 67(1):55, 1960.
- [24] D. Pritchard and D. Chakrabarty. Approximability of sparse integer programs. *Algorithmica*, 61(1):75–93, Sep 2011. doi: 10.1007/s00453-010-9431-z.
- [25] F. Vanderbeck and L. A. Wolsey. Reformulation and decomposition of integer programs. In M. Jünger, T. M. Liebling, D. Naddef, G. L. Nemhauser, W. R. Pulleyblank, G. Reinelt, G. Rinaldi, and L. A. Wolsey, editors, *50 Years of Integer Programming 1958-2008: From the Early Years to the State-of-the-Art*, chapter 13, pages 431–502. Springer Berlin Heidelberg, 2010. doi: 10.1007/978-3-540-68279-0_13.
- [26] L. A. Wolsey. Faces for a linear inequality in 0–1 variables. *Mathematical Programming*, 8(1): 165–178, 1975. doi: 10.1007/BF01580441.
- [27] M. Yannakakis. Node-deletion problems on bipartite graphs. *SIAM Journal on Computing*, 10 (2):310–327, 1981. doi: 10.1137/0210022.
- [28] G. M. Ziegler. *Lectures on polytopes*. Graduate texts in mathematics. Springer, 1995.

A Sparsity in Extended Formulations

Proposition 33. *Let $X \subseteq \mathbb{R}^n$ be a hole-free set and $Ax \leq b$ be a mixed-integer formulation of X with m inequalities in the original space. Then there exists a mixed-integer formulation of X all of whose inequalities are 3-sparse with $O(mn)$ variables and constraints.*

Proof. For each inequality $a^\top x \leq \beta$ in this formulation with $s \geq 4$ non-zero coefficients, we introduce $s - 2$ artificial variables and the following $s - 1$ additional constraints: W.l.o.g. we can assume that the first s coefficients are the non-zero coefficients of $a^\top x \leq \beta$. Define $y_1 := a_1x_1 + a_2x_2$ and $y_k = y_{k-1} + a_{k+1}x_{k+1}$ for each $k \in \{2, \dots, s - 2\}$, as well as $y_{s-2} + a_sx_s \leq \beta$. Then the additional constraints have at most three non-zero coefficients, and a point (x, y) is feasible for this set of constraints if and only if x fulfills $a^\top x \leq \beta$. \square

Remark 34. *If X has an extended mixed-integer formulation of polynomial size, the sparsified mixed-integer formulation of Proposition 33 has polynomial size as well. This shows that there exists a polynomial size formulation with sparse constraints in an extended space if the membership problem for X is in \mathcal{NP} , see Theorem 3 and its discussion in Section 2.*

B Bounded Coefficients in Extended Formulations

Lemma 35. *Let $X \subseteq \mathbb{R}^n$ be a mixed-integer set that admits a mixed-integer formulation (1). Then (1) can be expanded to an extended mixed-integer formulation of X of size*

$$O\left(\sum_{j=1}^p u_j \max\{|a_{ij}| : i \in [m]\} + \sum_{j=1}^q \lceil \tilde{u}_j \rceil \max\{|a_{ij}| : i \in [m]\}\right)$$

with left-hand side coefficients in $\{0, \pm 1\}$.

Proof. The main idea of this proof is to substitute the y -variables by variables in $\{0, 1\}$ and the z -variables by variables in $[0, 1]$. We distinguish whether a variable is integer or continuous.

Integer variables: Let y_j be an integer variable with upper bound u_j , and define $\bar{a}_{ij} := |a_{ij}|$ as well as $\bar{A}_j := \max\{\bar{a}_{ij} : i \in [m]\}$. We introduce $\bar{A}_j u_j$ many binary variables $y_j^{k\ell}$, where $k \in [u_j]$ and $\ell \in [\bar{A}_j]$. We then write

$$a_{ij} y_j = \text{sign}(a_{ij}) \sum_{k=1}^{u_j} \sum_{\ell=1}^{\bar{a}_{ij}} y_j^{k\ell},$$

and we add the *linking constraints* $y_j^{k\ell} = y_j^{k1}$, $\ell \in [\bar{A}_j]$, $k \in [u_j]$. Thus, in any feasible solution with $y_j^{k1} = 1$, we have $\sum_{\ell=1}^{\bar{a}_{ij}} y_j^{k\ell} = \bar{a}_{ij}$.

Continuous variables: For continuous variables z_j we can proceed similarly as for integer variables. As above, we define $\bar{B}_j := \max\{|b_{ij}| : i \in [m]\}$ and introduce artificial variables $z_j^{k\ell}$ with $k \in [\lceil \tilde{u}_j \rceil]$, $\ell \in [\bar{B}_j]$. We define the variables $z_j^{k\ell}$ to be contained in $[0, 1]$ and write

$$b_{ij} z_j = \text{sign}(b_{ij}) \sum_{k=1}^{\lceil \tilde{u}_j \rceil} \sum_{\ell=1}^{\bar{b}_{ij}} z_j^{k\ell}, \quad z_j^{k\ell} = z_j^{k1}, \quad \ell \in [\bar{B}_j], k \in [\lceil \tilde{u}_j \rceil].$$

To prove that the expanded mixed-integer formulation and (1) are equivalent, we have to show that each solution of (1) can be lifted to a feasible solution of the expanded system and vice versa. Let (y^*, z^*) be a solution of (1). We can generate the value $\bar{a}_{ij} y_j^*$ in constraint i by setting $y_{ij}^{k1} = y_{ij}^{k\ell} = 1$ for all $k \in [y_j^*]$, $\ell \in [\bar{A}_j]$ and $y_{ij}^{k1} = y_{ij}^{k\ell} = 0$ for $k > y_j^*$, $\ell \in [\bar{A}_j]$. We use the same assignment as in the integer case for $z_j^* \in \mathbb{Z}$. Otherwise, if z^* is fractional, we can represent this value in the expanded formulation by setting $z_j^{k1} = z_j^{k\ell} = 1$ for all $k \in [\lfloor \bar{z}_j \rfloor]$, $\ell \in [\bar{B}_j]$,

$$z_j^{\lceil z_j^* \rceil 1} = z_j^* - \lfloor z_j^* \rfloor,$$

and $z_j^{k1} = z_j^{k\ell} = 0$ for $k > \lceil z^* \rceil$. Hence, each solution of (1) can be lifted to a solution of the expanded system.

For the reverse direction, assume we are given a solution (Y, Z) of the expanded mixed-integer formulation, where $Y := (y_j^{k\ell})_{j k \ell}$ and $Z := (z_j^{k\ell})_{j k \ell}$. This solution can be projected to a solution of (1) by

$$(y, z) := \pi(Y, Z) = \left(\left(\sum_{k=1}^{u_j} y_j^{k1} \right)_j, \left(\sum_{k=1}^{\lceil \tilde{u}_j \rceil} z_j^{k1} \right)_j \right)$$

because of the linking constraints: Let $i \in [m]$ be the index of an expanded constraint. Then we can rearrange the i -th expanded constraint via the following steps

$$\begin{aligned}
& \sum_{j \in [p]} \text{sign}(a_{ij}) \sum_{k=1}^{u_j} \sum_{\ell=1}^{\bar{a}_{ij}} y_j^{k\ell} + \sum_{\tilde{j} \in [q]} \text{sign}(b_{i\tilde{j}}) \sum_{\tilde{k}=1}^{[\tilde{u}_{\tilde{j}}]} \sum_{\tilde{\ell}=1}^{\bar{b}_{i\tilde{j}}} z_{\tilde{j}}^{\tilde{k}\tilde{\ell}} \leq b_i \quad (12) \\
\Leftrightarrow & \sum_{j \in [p]} \text{sign}(a_{ij}) \bar{a}_{ij} \sum_{k=1}^{u_j} y_j^{k1} + \sum_{\tilde{j} \in [q]} \text{sign}(b_{i\tilde{j}}) \bar{b}_{i\tilde{j}} \sum_{\tilde{k}=1}^{[\tilde{u}_{\tilde{j}}]} z_{\tilde{j}}^{\tilde{k}1} \leq b_i \\
\Leftrightarrow & \sum_{j \in [p]} a_{ij} \pi(Y)_j + \sum_{\tilde{j} \in [q]} b_{i\tilde{j}} \pi(Z)_{\tilde{j}} \leq b_i.
\end{aligned}$$

Hence, $\pi(Y, Z)$ is feasible for (1). For this reason, both mixed-integer formulations are equivalent.

As a consequence, we can replace arbitrary integer coefficients of any variable by introducing $O(\sum_{j \in [p]} \bar{A}_j u_j + \sum_{\tilde{j} \in [q]} \bar{B}_{\tilde{j}} [\tilde{u}_{\tilde{j}}])$ new variables and constraints. \square

Theorem 36. *Let $X \subseteq \mathbb{R}^n$ be a mixed-integer set that admits an extended mixed-integer formulation (1).*

1. *If the number of constraints and variables as well as the coefficients in (1) are bounded by a polynomial in n , X has an extended mixed-integer formulation of polynomial size in n with left-hand side coefficients in $\{0, \pm 1\}$.*
2. *Let the following conditions hold:*
 - \triangleright *all entries of A, B, u , and \tilde{u} are bounded by a polynomial in n ,*
 - \triangleright *the mixed-integer formulation (1) is tractable, and*
 - \triangleright *the number of variables in (1) is bounded by a polynomial in n .*

Then X has a tractable extended mixed-integer formulation with coefficients in $\{0, \pm 1\}$.

Proof. The first part of the theorem is a direct consequence of Lemma 35. To see that the second part holds, we show that the mixed-integer formulation that is obtained by applying Lemma 35 is tractable.

W.l.o.g. we assume that $q = 0$, i.e., the mixed-integer formulation (1) contains only integer variables. Then the expanded mixed-integer formulation from Lemma 35 contains only binary variables $y_j^{k\ell}$.

Let $Y := (y_j^{k\ell})_{j k \ell}$ be the variable vector in the space of the expanded mixed-integer formulation. Because the expanded mixed-integer formulation contains at most polynomially many linking and box constraints, these constraints can be separated in polynomial time. Thus, we can assume in the following that Y fulfills all linking and box constraints, and it suffices to check whether the expanded constraints from (12) can be separated efficiently.

Since Y fulfills the linking constraints of Lemma 35, the i -th expanded constraint is violated by this solution if and only if $\pi(Y)$ violates the corresponding constraint in (1), cf. the argumentation in the proof of Lemma 35. Because (1) is tractable, finding an inequality of (1) that is violated by $\pi(Y)$ (if it exists) is possible in polynomial time. Since the transformation and the projection can be calculated in polynomial time, (12) can be separated in polynomial time. \square