

# Polynomial Size IP Formulations of Knapsack May Require Exponentially Large Coefficients

Christopher Hojny

*Combinatorial Optimization Group, Eindhoven University of Technology,  
P.O. Box 513, 5600 MB Eindhoven, The Netherlands, email: c.hojny@tue.nl*

November 14, 2019

## Abstract

A desirable property of integer formulations is to consist of few inequalities having small coefficients. We show that these targets are conflicting by proving the existence of knapsack sets that need exponentially many inequalities or exponentially large coefficients in any integer formulation. Moreover, we show that there exist undirected graphs such that (in a natural model) every integer formulation of stable sets requires exponentially large coefficients if the number of non-trivial inequalities is bounded by a constant.

## 1 Introduction

Let  $n$  be a positive integer. An *integer formulation* of a set  $X \subseteq \{0, 1\}^n$  is a system  $Cx \leq d$ , where  $C \in \mathbb{Z}^{m \times n}$  and  $d \in \mathbb{Z}^m$  for some positive integer  $m$ , such that  $\{x \in \mathbb{Z}^n : Cx \leq d\} = X$ . The number  $m$  of rows in this system is called the *size* of the integer formulation. Integer formulations are a widely used tool to model and solve linear optimization problems over  $X$  via integer programming techniques. To be able to solve these problems efficiently, one is interested in small integer formulations that behave well numerically. One way to guarantee numerical stability is to bound the size of coefficients in the system  $Cx \leq d$ .

The aim of this article is to show that small size and bounded coefficients are conflicting objectives in general. Our main result is a quantification of this relation. We show the existence of binary sets that require exponentially large coefficients in any sub-exponential size integer formulation. To this end, we construct a particular family of knapsack sets whose knapsack inequality is recursively defined. That is, whenever this recursive structure appears in an integer program, every numerically stable integer formulation requires a separation routine to be applicable in practice. We also point out that this structure naturally appears in the context of symmetry handling, i.e., the results have also consequences for deriving symmetry handling inequalities.

Knapsack sets, see [12] for an algorithmic and [10] for a polyhedral survey, are natural candidates for an extreme behavior regarding the size of coefficients and size of formulation, because they admit very small integer formulations with arbitrarily large coefficients: Given a positive integer capacity as well as a set of  $n$  items each having a positive integer weight, the knapsack set  $X$  consists of all selections of items whose common weight does not exceed the capacity. That is, if we assume the items to be labeled by  $[n] := \{1, \dots, n\}$ , the weights to be given by  $a_i$ ,  $i \in [n]$ , and the capacity to be  $\beta$ , the knapsack set can be represented by

$$X = \{x \in \mathbb{Z}^n : \langle a, x \rangle \leq \beta, x \in [0, 1]^n\}.$$

The size of this integer formulation is  $2n + 1$ , however, as the weights  $a_i$  are general positive integers, the coefficients can become arbitrarily large. A possible remedy to avoid large coefficients is to use an alternative integer formulation via no-good cuts [1, 3]. This formulation contains, besides box constraints, for every  $\bar{x} \in \{0, 1\}^n \setminus X$  the *no-good cut*  $\sum_{i=1}^n (1 - \bar{x}_i)x_i + \sum_{i=1}^n \bar{x}_i(1 - x_i) \geq 1$ ; it is valid as the only binary point that is cut off by this inequality is  $\bar{x}$ . Note that each coefficient

on the left-hand side is either 0 or  $\pm 1$ , i.e., of constant size. However, as  $\{0, 1\}^n \setminus X$  may contain exponentially many points, the price to pay for small coefficients is a drastic increase in the size of the formulation.

To reduce the number of inequalities in a formulation that only has coefficients 0 and  $\pm 1$ , one can use inequalities that cut off more than a single binary point. An exact characterization of the minimum size of such a formulation for knapsack sets is provided in [7] via so-called strong covers of  $X$ . However, there still exist knapsack sets that need exponentially many inequalities if restricting to ternary coefficients, even if every coefficient in the knapsack inequality is 1 or 2, see [10].

The question arises whether it is possible to find for every knapsack set  $X \subseteq \{0, 1\}^n$  a threshold  $T \in O(\text{poly}(n))$  such that there exists an integer formulation of  $X$  of polynomial size whose greatest (absolute) coefficient is at most  $T$ . In this article, we answer this question negatively by proving the following result.

**Theorem 1.** *There exists a family  $\mathcal{X}$  consisting of  $2^{\Omega(n)}$  knapsack sets  $X \subseteq \{0, 1\}^n$  such that every integer formulation of  $X$  requires  $2^{\Omega(n)}$  inequalities or coefficients of size at least  $2^{\Omega(n)}$ .*

In the second part of this article, we use the techniques used to prove Theorem 1 to obtain a similar result for another combinatorial optimization problem: the stable set problem, which has been extensively studied both from an algorithmic and a polyhedral point of view, see, e.g., [4, 13, 14] and [5, 6, 8], respectively. Given an undirected graph  $G = (V, E)$ , a *stable set* in  $G$  is a set  $S \subseteq V$  such that the nodes in  $S$  are pairwise non-adjacent. In the following, we consider the encoding of stable sets  $S$  in  $G$  via *incidence vectors*  $\chi(S) \in \{0, 1\}^V$ , i.e.,  $\chi(S)_i = 1$  if and only if  $i \in S$ . If  $\text{supp}(\cdot)$  denotes the support of a vector, then

$$\mathcal{S}(G) = \{x \in \{0, 1\}^V : \text{supp}(x) \text{ is stable set in } G\}$$

is the set of all incidence vectors of stable sets in  $G$ .

The set  $\mathcal{S}(G)$  admits a well-known integer formulation of size  $O(|V|^2)$  using box constraints and the inequality  $x_u + x_v \leq 1$  for every  $\{u, v\} \in E$ . Moreover, certain graphs, so-called *threshold graphs*, admit an integer formulation that consists of box constraints and an additional linear inequality [2], i.e.,  $\mathcal{S}(G)$  is a knapsack set in this case. This additional inequality can be constructed explicitly and, depending on the graph, may contain coefficients of size  $2^{\Omega(|V|)}$ , see [2].

Again, the question arises whether the exponential coefficients in non-box constraints are intrinsic to the model  $\mathcal{S}(G)$  or whether it is possible to find formulations with a constant number of additional inequalities whose coefficients are polynomial in  $|V|$ . In the following, we refer to a system  $Cx \leq d$  as a *binary formulation* of  $\mathcal{S}(G)$  if  $\{x \in \{0, 1\}^V : Cx \leq d\} = \mathcal{S}(G)$ , i.e., we decouple the box information from the system  $Cx \leq d$ .

**Theorem 2.** *For every positive integer  $n$ , there exists a threshold graph  $G_n = (V_n, E_n)$  containing  $2n$  nodes such that every binary formulation of  $\mathcal{S}(G_n)$  consisting of  $k \in [n]$  inequalities has a coefficient that is greater than  $2^{\lceil n/k \rceil - 2}$ .*

The aim of this article is to prove Theorem 1 and Theorem 2. Since the proofs of Theorems 1 and 2 are constructive, we are thus able to derive structural properties of knapsack and stable set problems that require large coefficients in any small integer formulation.

## 2 Small Formulations of Knapsack May Require Large Coefficients

This section is devoted to prove Theorem 1. We describe our proof strategy first.

**The strategy** To prove Theorem 1, we in fact prove a slightly stronger result, namely a parameterized version of Theorem 1.

**Theorem 3.** *Let  $k \leq \frac{n}{2}$  be a positive integer such that  $n \in 2^{\sigma(k)}$ . Then there exists a family  $\mathcal{X}_k$  consisting of  $\binom{n-k-1}{k-1}$  knapsack sets  $X \subseteq \{0, 1\}^n$  such that every integer formulation of  $X$  requires  $2^{\Omega(k)}$  inequalities or coefficients of size at least  $2^{\Omega(k)}$ .*

To be able to prove Theorem 3, we construct families  $\mathcal{X}_k$  of knapsack sets  $X \subseteq \{0, 1\}^n$  satisfying the following properties, which mainly rely on minimal covers. Below we will recall the definition of minimal covers for the sake of completeness.

(P1) Each set  $X \in \mathcal{X}_k$  has  $2^{\Omega(k)}$  minimal covers.

(P2) Every valid inequality for  $X \in \mathcal{X}_k$  that cuts off  $2^{\Omega(k)}$  minimal covers has coefficients of size  $2^{\Omega(k)}$ .

If we are given an integer formulation of a set  $X \in \mathcal{X}_k$  that consists of  $2^{o(k)}$  inequalities, Property (P1) implies that the formulation contains an inequality that cuts off  $2^{\Omega(k)}$  minimal covers. Hence, Property (P2) implies that this inequality has coefficients of size  $2^{\Omega(k)}$ , which proves Theorem 3.

In the following, we construct the family  $\mathcal{X}_k$  and show that it satisfies Properties (P1) and (P2).

**The construction** Let  $n$  be a positive integer. To construct the family  $\mathcal{X}_k$ , we consider certain partitions of  $[n]$ . A partition  $\mathcal{P} = \{P_1, \dots, P_k\}$  of  $[n]$  is called *consecutive* if, for each  $j \in [k]$ , part  $P_j$  consists of consecutive integers. For every  $j \in [k]$ , let  $\bar{p}_j := \max P_j$  and  $\underline{p}_j := \min P_j$  and define  $P_j^* = P_j \setminus \{\bar{p}_j\}$ . W.l.o.g. we assume in the following that the sets  $P_j$ ,  $j \in [k]$ , are labeled such that  $\bar{p}_1 < \bar{p}_2 < \dots < \bar{p}_k$ .

Given a consecutive partition  $\mathcal{P} = \{P_1, \dots, P_k\}$  with  $|P_j| \geq 2$  for every  $j \in [k]$ , define the vector  $a = a(\mathcal{P}) \in \mathbb{Z}^n$  as well as the scalar  $\beta = \beta(\mathcal{P}) \in \mathbb{Z}$  via

$$a(\mathcal{P})_i = \begin{cases} 1, & \text{if } i \in P_k^*, \\ a_{\bar{p}_{j+1}} |P_{j+1}|, & \text{if } i \in P_j^* \text{ for some } j \in [k-1], \\ a_{\underline{p}_j} |P_j^*|, & \text{if } i = \bar{p}_j \text{ for some } j \in [k], \end{cases}$$

and  $\beta(\mathcal{P}) = \sum_{j=1}^k a(\mathcal{P})_{\bar{p}_j}$ . Whenever the partition  $\mathcal{P}$  is clear from the context, we write  $a$  and  $\beta$  instead of  $a(\mathcal{P})$  and  $\beta(\mathcal{P})$ .

Using the vector  $a(\mathcal{P})$  and scalar  $\beta(\mathcal{P})$ , we define the knapsack set

$$X_{\mathcal{P}} := \{x \in \{0, 1\}^n : \langle a(\mathcal{P}), x \rangle \leq \beta(\mathcal{P})\}.$$

The proposed family  $\mathcal{X}_k$  of knapsack sets is then given by the sets  $X_{\mathcal{P}}$  for all consecutive partitions  $\mathcal{P}$  of  $[n]$  consisting of exactly  $k$  parts all having cardinality at least 2.

**Characterization of minimal covers** Let  $a \in \mathbb{Z}_{>0}^n$  and  $\beta \in \mathbb{Z}_{>0}$ . A set  $C \subseteq [n]$  is called a *cover* for the knapsack set  $\{x \in \{0, 1\}^n : \langle a, x \rangle \leq \beta\}$  if  $\sum_{i \in C} a_i > \beta$ . A cover  $C$  is called *minimal* if no proper subset of  $C$  is a cover. For  $C \subseteq [n]$ , the vector  $\chi(C) \in \{0, 1\}^n$  denotes the *characteristic vector* of  $C$ .

To be able to show that  $\mathcal{X}_k$  satisfies (P1), we derive a characterization of minimal covers for  $X \in \mathcal{X}_k$ .

**Proposition 4.** *Let  $\mathcal{P}$  be a consecutive partition of  $[n]$  consisting of  $k$  parts each of which has cardinality at least 2. Then,  $x \in \{0, 1\}^n$  is the characteristic vector of a minimal cover for  $X_{\mathcal{P}}$  if and only if there exists  $j^* \in [k]$  such that*

(C1)  $x_{\bar{p}_{j^*}} = 1$  and there exists exactly one  $i^* \in P_{j^*}^*$  with  $x_{i^*} = 1$ ,

(C2) for every  $j \in [j^* - 1]$  and  $i \in P_j^*$ , we have  $x_{\bar{p}_j} = 1 - x_i$ , and

(C3)  $x_i = 0$  for every  $i \in P_j$  with  $j \in \{j^* + 1, \dots, k\}$ .

The proof of Proposition 4 is relying on the following three technical lemmata. While the first one shows that the coefficients in  $a(\mathcal{P})$  grow rapidly, the remaining lemmata provide sufficient criteria for a set  $C \subseteq [n]$  (not) to define a cover.

**Lemma 5.** Let  $\mathcal{P}$  be a consecutive partition of  $[n]$  consisting of  $k$  parts each of which has cardinality at least 2 and let  $a = a(\mathcal{P})$  as well as  $\beta = \beta(\mathcal{P})$ . Then, for every  $j' \in [k-1]$ ,

$$\sum_{j=j'+1}^k \sum_{i \in P_j^*} a_i < a_{p_{j'}}.$$

*Proof.* Observe that, for any  $j \in \{2, \dots, k\}$ , we have

$$a_{p_{j-1}} = a_{\bar{p}_j} |P_j| = a_{\bar{p}_j} + a_{\bar{p}_j} |P_j^*| = a_{\bar{p}_j} + a_{p_j} |P_j^*| |P_j^*| \geq a_{\bar{p}_j} + a_{p_j} = \sum_{i \in P_j^*} a_i + a_{p_j}, \quad (1)$$

since  $a_{\bar{p}_j} = a_{p_j} |P_j^*|$  and  $a_i = a_{p_j}$  for every  $i \in P_j^*$ .

To prove the lemma's statement, we proceed inductively. If  $j' = k-1$ , this is an immediate consequence of (1) and positivity of  $a$ . Hence, the base case holds.

Otherwise, if  $j' < k-1$ , the induction hypothesis (\*) yields

$$\sum_{j=j'+1}^k \sum_{i \in P_j^*} a_i = \sum_{i \in P_{j'+1}^*} a_i + \sum_{j=j'+2}^k \sum_{i \in P_j^*} a_i \stackrel{(*)}{<} \sum_{i \in P_{j'+1}^*} a_i + a_{p_{j'+1}} \stackrel{(1)}{\leq} a_{p_{j'}},$$

which is the assertion.  $\square$

**Lemma 6.** Let  $\mathcal{P}$  be a consecutive partition of  $[n]$  consisting of  $k$  parts each of which has cardinality at least 2. Then,  $x \in \{0, 1\}^n$  is the characteristic vector of a cover for  $X_{\mathcal{P}}$  if there exists  $j^* \in [k]$  such that Properties (C1) and (C2) hold.

*Proof.* Suppose  $x$  satisfies Properties (C1) and (C2) and let  $j^*$  be defined as in (C1). Since  $a_i = a_{p_{j^*}}$  for every  $i \in P_{j^*}^*$ , we may assume w.l.o.g. that  $x_{p_{j^*}}$  is the unique 1-entry on  $P_{j^*}^*$  by Property (C1). Moreover, as  $a_{\bar{p}_j} = a_{p_j} |P_j^*|$  for every  $j \in [k]$ , we can assume  $x_{\bar{p}_j} = 1$  for every  $j \in [j^*-1]$  by Property (C2). Then,

$$\langle a, x \rangle = \sum_{j=1}^k \sum_{i \in P_j} a_i x_i \geq \sum_{j=1}^{j^*} \sum_{i \in P_j} a_i x_i \stackrel{(C1)}{=} a_{\bar{p}_{j^*}} + a_{p_{j^*}} + \sum_{j=1}^{j^*-1} \sum_{i \in P_j} a_i x_i \stackrel{(C2)}{=} a_{\bar{p}_{j^*}} + a_{p_{j^*}} + \sum_{j=1}^{j^*-1} a_{\bar{p}_j}. \quad (2)$$

If  $j^* = k$ , this shows that  $\langle a, x \rangle > \delta$ , because  $a_{p_{j^*}} > 0$  and  $\delta = \sum_{j=1}^k a_{\bar{p}_j}$ . Otherwise, if  $j^* < k$ , then Lemma 5 implies  $a_{p_{j^*}} > \sum_{j=j^*+1}^k \sum_{i \in P_j^*} a_i = \sum_{j=j^*+1}^k a_{\bar{p}_j}$ . Inserting this estimation in (2) yields  $\langle a, x \rangle > \sum_{j=1}^k a_{\bar{p}_j} = \beta$ , showing that  $x$  is the characteristic vector of a cover. Hence, Properties (C1) and (C2) are sufficient for  $x$  to define a cover.  $\square$

**Lemma 7.** Let  $\mathcal{P}$  be a consecutive partition of  $[n]$  consisting of  $k$  parts each of which has cardinality at least 2. Then,  $x \in \{0, 1\}^n$  is not the characteristic vector of a cover for  $X_{\mathcal{P}}$  if there exists  $j' \in [k]$  such that

- (i)  $x_{\bar{p}_{j'}} = 0$  and there exists  $i \in P_{j'}^*$  with  $x_i = 0$ , and
- (ii) for every  $j \in [j'-1]$  and  $i \in P_j$ , we have  $x_i = 1 - x_{\bar{p}_j}$ .

*Proof.* Let  $x \in \{0, 1\}^n$  satisfy the described properties. Since  $\sum_{i \in P_j^*} a_i = a_{\bar{p}_j}$  for every  $j \in [k]$  and  $x_{\bar{p}_j} = 1 - x_i$  for every  $i \in P_j^*$  with  $j \in [j'-1]$  (denote this property with (\*)), we find

$$\begin{aligned} \langle a, x \rangle &= \sum_{j=1}^{j'-1} \sum_{i \in P_j} a_i x_i + \sum_{i \in P_{j'}} a_i x_i + \sum_{j=j'+1}^k \sum_{i \in P_j} a_i x_i \stackrel{(*)}{=} \sum_{j=1}^{j'-1} a_{\bar{p}_j} + \sum_{i \in P_{j'}} a_i x_i + \sum_{j=j'+1}^k \sum_{i \in P_j} a_i x_i \\ &\leq \sum_{j=1}^{j'-1} a_{\bar{p}_j} + \sum_{i \in P_{j'}} a_i x_i + \sum_{j=j'+1}^k \sum_{i \in P_j} a_i \end{aligned}$$

$$\begin{aligned}
& \stackrel{\text{Lemma 5}}{<} \sum_{j=1}^{j'-1} a_{\bar{p}_j} + \sum_{i \in P_{j'}} a_i x_i + a_{\underline{p}_{j'}} + \sum_{j=j'+1}^k a_{\bar{p}_j} \\
& \stackrel{(i)}{\leq} \sum_{j=1}^{j'-1} a_{\bar{p}_j} + (|P_{j'}| - 1) a_{\underline{p}_{j'}} + \sum_{j=j'+1}^k a_{\bar{p}_j} \\
& = \sum_{j=1}^k a_{\bar{p}_j} = \beta.
\end{aligned}$$

Consequently,  $x$  satisfies the knapsack inequality and thus cannot be the characteristic vector of a cover.  $\square$

Using these technical lemmata, we are able to prove Proposition 4.

*Proof of Proposition 4.* Let  $x \in \{0, 1\}^n$ . Let

$$\begin{aligned}
j' &:= \min\{j \in [k] : x_{\bar{p}_j} = 0 \text{ and there exists } i \in P_j^* \text{ with } x_i = 0\} \cup \{k+1\}, \\
j^* &:= \min\{j \in [k] : x_{\bar{p}_j} = 1 \text{ and } x_i = 1 \text{ for some } i \in P_j^*\} \cup \{k+1\}.
\end{aligned}$$

By a case distinction on the relation between  $j'$  and  $j^*$ , we characterize whether  $x$  defines the characteristic vector of a cover for  $X_{\mathcal{P}}$ .

First, suppose  $j' = j^* = k+1$ . Then, for every  $j \in [k]$ , we have  $x_i = 1 - x_{\bar{p}_j}$  for every  $i \in P_j^*$ . Since  $\sum_{i \in P_j^*} a_i = a_{\bar{p}_j}$  for every  $j \in [k]$ , we conclude  $\langle a, x \rangle = \sum_{j=1}^k a_{\bar{p}_j} = \beta$ . Hence,  $x$  is feasible for  $X_{\mathcal{P}}$  and cannot be the incidence vector of a cover.

Second, assume  $j' < j^*$ . Then, for every  $j \in [j'-1]$  and  $i \in P_j^*$ , we again have  $x_i = 1 - x_{\bar{p}_j}$ . Thus, Lemma 7 implies that  $x$  cannot be the characteristic vector of a cover. Combining the above results shows that  $x$  can only be the characteristic vector of a cover if  $j^* < j'$ . That is,  $x_i = 1 - x_{\bar{p}_j}$  for every  $j \in [j^*-1]$  and  $i \in P_j^*$  as well as there exists  $i \in P_{j^*}^*$  such that  $x_{\bar{p}_{j^*}} = x_i = 1$ .

Third, suppose  $j^* < j'$  and let  $i^* \in P_{j^*}^*$  be such that  $x_{i^*} = 1$ . Consider the vector  $x' \in \{0, 1\}^n$  defined by

$$x'_i = \begin{cases} x_i, & \text{if } i \in P_j \text{ for some } j \in [j^*-1], \\ 1, & \text{if } i \in \{i^*, \bar{p}_{j^*}\}, \\ 0, & \text{otherwise.} \end{cases}$$

Then,  $x' \leq x$  and  $x'$  satisfies Properties (C1) to (C3). Due Lemma 6 we get that  $x'$  is the characteristic vector of a cover for  $X_{\mathcal{P}}$ . Thus, if  $x' \neq x$ , then  $x$  cannot correspond to a *minimal* cover. As a consequence, Properties (C1) to (C3) are necessary for  $x$  to be the characteristic vector of a minimal cover.

Finally, it remains to prove sufficiency of Properties (C1) to (C3). By Lemma 6, we know that  $x$  is the characteristic vector of a cover. To prove that  $x$  actually corresponds to a minimal cover, we set a 1-entry  $i$  of  $x$  with minimum weight  $a_i$  to 0. If the resulting vector does not correspond to a cover,  $x$  corresponds to a minimal cover. Since the 1-entry with smallest weight is  $x_{\underline{p}_{j^*}}$ , we get

$$\langle a, x \rangle - a_{\underline{p}_{j^*}} = \sum_{j=1}^{j^*} a_{\bar{p}_j} \leq \sum_{j=1}^k a_{\bar{p}_j} = \beta,$$

concluding the proof.  $\square$

**A lower bound on the size of coefficients** Throughout this section, we assume that  $\mathcal{P}$  is a consecutive partition of  $[n]$  consisting of  $k$  parts each having cardinality at least 2. To derive lower bounds on the size of coefficients in integer formulations of  $X_{\mathcal{P}}$ , we partition the set of (characteristic vectors of) minimal covers by their index in Property (C1). Moreover, we define for each minimal cover a counterpart, which will be useful in the derivation of the lower bound.

**Definition 8.** Let  $r \in [k]$ . A minimal cover  $C$  for  $X_{\mathcal{P}}$  is called an  $r$ -cover if the index  $j^*$  in Property (C1) equals  $r$ . The counterpart of  $\chi(C)$  is the vector  $\bar{x} \in \{0, 1\}^n$  that satisfies

$$\bar{x}_i = \begin{cases} \chi(C)_i, & \text{if } i \in P_j \text{ for some } j \in [r-1], \\ 0, & \text{if } i \in P_r, \\ 1, & \text{otherwise.} \end{cases}$$

Note that the counterpart  $\bar{x}$  of  $\chi(C)$  for a minimal cover  $C$  of  $X_{\mathcal{P}}$  is contained in  $X_{\mathcal{P}}$  by Lemma 7 and Proposition 4. Hence,  $\langle c, \chi(C) \rangle > \delta \geq \langle c, \bar{x} \rangle$  holds for every valid inequality  $\langle c, x \rangle \leq \delta$  for  $X_{\mathcal{P}}$ . Moreover, as there exists exactly one  $i(r) \in P_r^*$  such that  $\chi(C)_{i(r)} = 1$ , we conclude

$$\langle c, \chi(C) \rangle > \langle c, \bar{x} \rangle \quad \Rightarrow \quad c_{i(r)} + c_{\bar{p}_r} > \sum_{j=r+1}^k \sum_{i \in P_j} c_i \quad (3)$$

by the definition of  $\bar{x}$ .

To find large lower bounds on the size of coefficients in small integer formulations of  $X_{\mathcal{P}}$ , our strategy is to recursively use (3) on some coefficients  $c_i$  for  $i \in \bigcup_{j=r+1}^k P_j$ . To this end, it is necessary to drop some of the coefficients on the right-hand side of (3), which is only valid if all coefficients in  $c$  are non-negative.

**Lemma 9.** Let  $X \subseteq \{0, 1\}^n$  be a knapsack set and let  $C$  be a minimal cover for  $X$ . If  $\langle c, x \rangle \leq \delta$  is a valid inequality for  $X$  that cuts off  $\chi(C)$ , then  $c_j > 0$  for every  $j \in C$ .

*Proof.* Let  $j \in C$  and define  $C^* = C \setminus \{j\}$ . Since  $\langle c, x \rangle \leq \delta$  cuts off  $\chi(C)$ , we have  $\langle c, \chi(C) \rangle > \delta$ . Moreover, minimality of  $C$  implies that  $\chi(C^*) \in X$ . Hence,

$$\sum_{i \in C} c_i = \langle c, \chi(C) \rangle > \langle c, \chi(C^*) \rangle = \sum_{i \in C} c_i - c_j,$$

showing  $c_j > 0$ . □

In the following, we consider two quantities that will be useful for deriving lower bounds on the size of coefficients in an integer formulation. For an inequality  $\langle c, x \rangle \leq \delta$ , let

$$R(c, \delta) = \{r \in [k] : \langle c, x \rangle \leq \delta \text{ cuts off an } r\text{-cover}\}.$$

Moreover, for  $r \in R(c, \delta)$ , let

$$H_r(c, \delta) = \{s \in [r-1] : \exists r\text{-covers } x^1 \text{ and } x^2 \text{ with } x_{p_s}^1 = 0 \text{ and } x_{p_s}^2 = 1\}$$

and define  $h_r(c, \delta) = |H_r(c, \delta)|$ .

**Lemma 10.** Every valid inequality  $\langle c, x \rangle \leq \delta$  for  $X_{\mathcal{P}}$  with integral coefficients that cuts off a minimal cover has a left-hand side coefficient of size  $2^{\Omega(|R(c, \delta)|)}$ .

*Proof.* Let  $R = R(c, \delta)$ . For every  $r \in R$ , let  $x^r$  be an  $r$ -cover that is cut off by  $\langle c, x \rangle \leq \delta$ . Moreover, let  $\bar{x}^r$  be the counterpart of  $x^r$  and select  $i(r) \in P_r^*$  such that  $x_{i(r)}^r = 1$ . Due to Proposition 4, this index is well-defined and unique. If there exist 1-entries of  $\bar{x}^r$  such that the corresponding entries in  $c$  are negative, we replace these 1-entries by 0-entries. This modification ensures that  $c_i \bar{x}_i^r \geq 0$  for every  $i \in [n]$ . Furthermore, observe that  $c_{i(r)}$  and  $c_{\bar{p}_r}$  are positive by Lemma 9, because both are contained in the minimal cover  $x^r$ , which is cut off by  $\langle c, x \rangle \leq \delta$ . Analogously, we get that no entries of  $\bar{x}^r$  on  $\bigcup_{j=1}^{r-1} P_j$  have been modified.

Since  $x^r$  is infeasible and  $\bar{x}^r$  is feasible for  $X_{\mathcal{P}}$ , we find analogously to (3)

$$\langle c, x^r \rangle > \delta \geq \langle c, \bar{x}^r \rangle \quad \Rightarrow \quad c_{i(r)} + c_{\bar{p}_r} > \sum_{i=\bar{p}_r+1}^n c_i \bar{x}_i^r \geq \sum_{\substack{s \in R: \\ s > r}} (c_{i(s)} + c_{\bar{p}_s}), \quad (4)$$

where the last estimation is valid due to the modification of  $\bar{x}^r$  and positivity of  $c_{i(s)}$  and  $c_{\bar{p}_s}$  for every  $r \in R$ . If  $R = \{r(1), \dots, r(t)\}$ , where  $r(1) < r(2) < \dots < r(t)$ , and we apply (4) iteratively, we obtain

$$c_{i(r(1))} + c_{\bar{p}_{r(1)}} > 2^{t-2}(c_{i(r(t))} + c_{\bar{p}_{r(t)}}).$$

Thus, integrality of the left-hand side coefficients implies

$$2 \max\{c_{i(r(1))}, c_{\bar{p}_{r(1)}}\} > 2^{t-2} \cdot 2,$$

showing  $\max\{c_{i(r(1))}, c_{\bar{p}_{r(1)}}\} \in 2^{\Omega(|R(c, \delta)|)}$ .  $\square$

**Lemma 11.** *A valid inequality  $\langle c, x \rangle \leq \delta$  for  $X_{\mathcal{P}}$  with integral coefficients that cuts off an  $r$ -cover for some  $r \in R$  has a left-hand side coefficient of size  $2^{\Omega(h_r(c, \delta))}$ .*

*Proof.* Let  $h = h_r(c, \delta)$  and let  $S_r := \{s(1), \dots, s(h)\} \subseteq [r-1]$  be the set of indices  $s$  such that there exist  $r$ -covers  $x^{s,1}$  and  $x^{s,2}$  with  $x_{\bar{p}_s}^{s,1} = 0$  and  $x_{\bar{p}_s}^{s,2} = 1$  that are cut off by  $\langle c, x \rangle \leq \delta$ . We assume that the elements in  $S_r$  are labeled such that  $s(1) < s(2) < \dots < s(h)$ . Moreover, for  $t \in [2]$ , let  $i(s, t)$  be the unique index in  $P_r^*$  with  $x_{i(s,t)}^{s,t} = 1$ , cf. Proposition 4. Since we are only interested in an asymptotic bound, we can assume  $h \geq 2$ .

For each of the infeasible points  $x^{s,t}$ ,  $s \in S_r$  and  $t \in [2]$ , we define a feasible counterpart  $\bar{x}^{s,t} \in \mathbb{R}^n$  via

$$\bar{x}_i^{s,t} = \begin{cases} x_i^{s,t}, & \text{if } i \in P_{r'} \text{ for some } r' < s, \\ 0, & \text{if } i \in P_s \text{ or } i \in P_{r'} \text{ for some } r' > r, \\ 1, & \text{if } i \in P_{r'} \text{ for some } s < r' \leq r. \end{cases}$$

If there exist 1-entries  $\bar{x}_i^{s,t}$  such that  $c_i$  is negative, we replace these entries by 0-entries.

Let  $s \in S_r$  and  $t \in [2]$ . Using the pairings  $(x^{s,t}, \bar{x}^{s,t})$  of infeasible and feasible points, we derive a lower bound on some coefficients in  $\langle c, x \rangle \leq \delta$ . On the one hand, if  $t = 1$ , then  $x_i^{s,1} = 1$  for every  $i \in P_s^*$  and  $x_{\bar{p}_s}^{s,1} = 0$ . By construction of  $\bar{x}^{s,1}$ , we conclude

$$\langle c, x^{s,t} \rangle > \delta \geq \langle c, \bar{x}^{s,t} \rangle \Rightarrow \sum_{i \in P_s^*} c_i > \sum_{r'=s+1}^r \sum_{i \in P_{r'}} (c_i(\bar{x}_i^{s,t} - x_i^{s,t})).$$

On the other hand, if  $t = 2$ , then  $x_i^{s,1} = 0$  for every  $i \in P_s^*$  and  $x_{\bar{p}_s}^{s,1} = 1$ . Thus, we find

$$\langle c, x^{s,t} \rangle > \delta \geq \langle c, \bar{x}^{s,t} \rangle \Rightarrow c_{\bar{p}_s} > \sum_{r'=s+1}^r \sum_{i \in P_{r'}} (c_i(\bar{x}_i^{s,t} - x_i^{s,t})).$$

Hence, we obtain in both cases the same lower bound, yielding

$$\min \left\{ c_{\bar{p}_s}, \sum_{i \in P_s^*} c_i \right\} > \sum_{r'=s+1}^r \sum_{i \in P_{r'}} (c_i(\bar{x}_i^{s,t} - x_i^{s,t})). \quad (5)$$

In the following, we adjust this common lower bound such that we can use it in an iterative procedure.

Note that for every  $i \in \bigcup_{r'=s+1}^r P_{r'}$ , we have  $c_i(\bar{x}_i^{s,t} - x_i^{s,t}) \geq 0$ : If  $c_i \geq 0$ , then  $\bar{x}_i^{s,t} = 1$ , and thus,  $\bar{x}_i^{s,t} - x_i^{s,t} \geq 0$ . Otherwise, if  $c_i < 0$ , then  $\bar{x}_i^{s,t} = 0$ , and hence,  $\bar{x}_i^{s,t} - x_i^{s,t} \leq 0$ . For this reason, we can adjust the common lower bound to

$$\sum_{r'=s+1}^r \sum_{i \in P_{r'}} (c_i(\bar{x}_i^{s,t} - x_i^{s,t})) \geq \sum_{r' \in H_s} \sum_{i \in P_{r'}} (c_i(\bar{x}_i^{s,t} - x_i^{s,t})),$$

where  $H_s := S_r \cap \{s+1, \dots, r\}$ . Since, for every  $r' \in H_s$ , we find  $r$ -covers  $x^1$  and  $x^2$  with  $x_i^1 = 1$  for every  $i \in P_{r'}$  and  $x_{\bar{p}_{r'}}^2 = 1$ , Lemma 9 implies that  $c_i > 0$  for every  $i \in P_{r'}$ . Hence,  $\bar{x}_i^{s,1} = \bar{x}_i^{s,2} = 1$  for all  $i \in P_{r'}$ , showing

$$\sum_{r' \in H_s} \sum_{i \in P_{r'}} (c_i(\bar{x}_i^{s,t} - x_i^{s,t})) = \sum_{r' \in H_s} \sum_{i \in P_{r'}} (c_i - c_i x_i^{s,t}) \geq \sum_{r' \in H_s} \sum_{i \in P_{r'}} \min \left\{ c_{\bar{p}_{r'}}, \sum_{i \in P_{r'}^*} c_i \right\},$$

because either  $x_{\bar{p}_{r'}}^{s,t} = 1$  or  $x_i^{s,t} = 1$  for all  $i \in P_{r'}^*$  as both are  $r$ -covers. Thus, we can adapt the lower bound (5) to

$$\min \left\{ c_{\bar{p}_s}, \sum_{i \in P_s^*} c_i \right\} > \sum_{r' \in H_s} \sum_{i \in P_{r'}} \min \left\{ c_{\bar{p}_{r'}}, \sum_{i \in P_{r'}^*} c_i \right\}. \quad (6)$$

By applying (6) iteratively, we find

$$\min \left\{ c_{\bar{p}_{s(1)}}, \sum_{i \in P_{s(1)}^*} c_i \right\} > 2^{h-2} \min \left\{ c_{\bar{p}_{s(h)}}, \sum_{i \in P_{s(h)}^*} c_i \right\} \geq 2^{h-2}.$$

Thus,  $c_{\bar{p}_{s(1)}} \in 2^{\Omega(h)}$ .  $\square$

Consequently, both  $R(c, \delta)$  and  $h_r(c, \delta)$  have an exponential impact on the size of coefficients in integer formulations of  $X_{\mathcal{P}}$ . As a final ingredient for the proof of Theorem 3, we need to find an upper bound on the number of minimal covers that can be cut off by a valid inequality, which corresponds to Property (P2).

**Lemma 12.** *A valid inequality  $\langle c, x \rangle \leq \delta$  for  $X_{\mathcal{P}}$  cuts off at most  $(|P_r| - 1)2^{h_r(c, \delta)}$ -many  $r$ -covers.*

*Proof.* The only two degrees of freedom in an  $r$ -cover  $x$  are

- to select the unique 1-entry in  $P_r^*$  and
- to decide for every  $r' \in [r - 1]$  whether  $x_{\bar{p}_{r'}} = 0$  or  $x_{\bar{p}_{r'}} = 1$ .

Since, for every  $r' \in [r - 1] \setminus H_r(c, \delta)$ , we have  $x_{\bar{p}_{r'}} = 1$  for every  $r$ -cover or  $x_{\bar{p}_{r'}} = 0$  for every  $r$ -cover that is cut off by this inequality, the number of  $r$ -covers that are cut off by  $\langle c, x \rangle \leq \delta$  is at most  $(|P_r| - 1)2^{h_r(c, \delta)}$ .  $\square$

Using the results developed in this section, allows to finally prove Theorem 3.

*Proof of Theorem 3.* Let  $Cx \leq d$  be an integer formulation of  $X_{\mathcal{P}}$  and let  $m$  be the number of inequalities in this formulation. If  $m \in 2^{\Omega(k)}$ , we are done because all left-hand side coefficients are integral and there exist positive left-hand side coefficients due to Lemma 9. For this reason, suppose  $m \in 2^{o(k)}$ . Since there exist  $\sum_{r=1}^k (|P_r| - 1)2^{r-1} \geq \sum_{r=1}^k 2^{r-1} \in 2^{\Omega(k)}$  minimal covers for  $X_{\mathcal{P}}$ , there exists an inequality  $\langle c, x \rangle \leq \delta$  in  $Cx \leq d$  that cuts off  $2^{\Omega(k)}$ -many minimal covers in this case. In the following, for every  $r \in [k]$ , we abbreviate  $R(c, \delta)$  by  $R$  and  $h_r(c, \delta)$  by  $h_r$ .

If  $|R| \in \Omega(k)$ , we are done due to Lemma 10. Otherwise,  $|R| \in o(k)$ . In this case, consider the quantity  $h^* = \max\{h_r : r \in [k]\}$ . By Lemma 12, we conclude that the number of minimal covers that is cut off by  $\langle c, x \rangle \leq \delta$  is at most

$$\sum_{r \in R} (|P_r| - 1)2^{h_r} \leq n2^{h^*} \in 2^{o(k) + h^*},$$

because  $\sum_{r \in R} |P_r| \leq \sum_{r=1}^n |P_r| = n$  and  $n \in o(2^k)$ . To ensure that  $\langle c, x \rangle \leq \delta$  cuts off  $2^{\Omega(k)}$ -many minimal covers, we thus require that  $h^* \in \Omega(k)$ . Thus, Lemma 11 shows that the inequality has a left-hand side coefficient of size  $2^{\Omega(k)}$ .

To conclude the proof, we need to establish that there exist  $\binom{n-k-1}{k-1}$  many consecutive partitions such that each of its parts has size at least 2. Observe that the number of such partitions is equal to the number of positive integer solutions of the equation  $y_1 + \dots + y_k = n - k$ . This equation has  $\binom{n-k-1}{k-1}$  solutions, see [15].  $\square$

We are now able to provide a proof of Theorem 1.

*Proof of Theorem 1.* Since the theorem's statement is asymptotic, we may assume  $n \geq 8$ . For the sake of simplicity, we moreover assume  $n = 4r$  for a positive integer  $r$ . If  $n$  is not a multiple of 4, the argumentation can be slightly modified without changing the basic arguments.



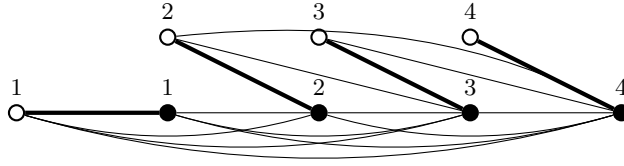


Figure 1: The threshold graph  $G_4$ . Nodes of  $u$ -type are colored white and nodes of  $w$ -type are colored black. Critical edges are drawn thickly.

Let  $k = \frac{1}{4}n$ . Then, there exist  $\binom{\frac{3}{4}n-1}{\frac{1}{4}n-1} = \binom{3r-1}{r-1}$  consecutive partitions of  $[n]$  that consist of  $k$  parts each of whose size is at least 2. As  $k \in \Theta(n)$  and  $k \leq \frac{n}{2}$ , the statement about the coefficient size and number of inequalities follows from Theorem 3. Thus, it is sufficient to show  $\binom{3r-1}{r-1} \in 2^{\Omega(n)}$ :

$$\binom{3r-1}{r-1} = \prod_{i=1}^{r-1} \frac{2r+i}{i} > \prod_{i=1}^{r-1} \frac{2r}{i} = \frac{2^{r-1} \cdot r^{r-1}}{(r-1)!} \geq 2^{r-1} \in 2^{\Omega(n)}.$$

□

**Remark 13.** The knapsack sets  $X_{\mathcal{P}}$  for consecutive partitions  $\mathcal{P}$  naturally arise in symmetry handling for binary programs: A symresack is the convex hull of all binary vectors  $x$  that are lexicographically not smaller than their image w.r.t. a fixed variable permutation, see [11]. Valid inequalities for symresacks can be used to cut off symmetric solutions of binary programs. One can show that applying some variable flips  $x_i \mapsto 1 - x_i$  to  $X_{\mathcal{P}}$  leads to the vertex set of a certain symresack, see [9]. Thus, there exist symresacks that require many inequalities or large coefficients in any integer formulation, i.e., systems of symmetry handling inequalities. We conclude that any set of numerically stable symmetry handling inequalities derived by means of symresacks requires a separation routine to be used in practice.

### 3 Small Formulations of Stable Set May Require Large Coefficients

In this section, we prove Theorem 2. The outline of the proof is analogous to the case of Theorem 1, i.e., we construct a threshold graph and derive properties of inequalities cutting off certain minimal covers. Note, however, that we cannot use arbitrary minimal covers in this case: The minimal covers of a threshold graph are exactly its edges. Hence, if the edges (minimal covers) are contained in a common clique  $C$ , the clique inequality  $\sum_{i \in C} x_i \leq 1$  cuts off all the corresponding incidence vectors. For this reason, we have to construct threshold graphs that contain maximal cliques which admit edges that are contained in exactly one of these cliques.

**The construction** Let  $G = (V, E)$  be an undirected graph and let  $n = |V|$ . In [2] it is shown that  $G$  is a threshold graph if and only if there is a labeling  $v_1, \dots, v_n$  of its nodes with the following property: for  $i \in \{2, \dots, n\}$ , either  $v_i$  is adjacent to all or none of its predecessors  $v_1, \dots, v_{i-1}$ . In the following, we refer to such a sequence as a *threshold sequence* defining  $G$ .

To construct the graph  $G_n = (V_n, E_n)$  from Theorem 2 with  $2n$  nodes, define a threshold sequence  $u_1, w_1, u_2, w_2, \dots, u_n, w_n$  such that, for  $i \in [n]$ ,  $u_i$  is connected with none of its predecessors and  $w_i$  is connected with all of its predecessors. Figure 1 illustrates this construction for  $n = 4$ .

**A lower bound on the size of coefficients** As indicated above, we consider certain edges in  $G_n$ . An edge  $e \in E_n$  is called *critical* if there exists  $i \in [n]$  such that  $e = \{u_i, w_i\}$ .

**Proposition 14.** *Let  $n$  be a positive integer, let  $E \subseteq E_n$  be a set of critical edges, and let  $\langle c, x \rangle \leq \delta$  be a valid inequality for  $\mathcal{S}(G_n)$  with  $c \in \mathbb{Z}^{V_n}$ . If  $\langle c, x \rangle \leq \delta$  cuts off the incidence vectors of the edges in  $E$ , then  $c$  has an entry greater than  $2^{|E|-2}$ .*

*Proof.* Let  $u_1, w_1, \dots, u_n, w_n$  be the threshold sequence of  $G_n$  and let  $I = \{i \in [n] : \{u_i, w_i\} \in E\}$ . If  $|E| = 1$ , the proposition's statement is an immediate consequence of integrality of  $c$  and Lemma 9.

For this reason, suppose  $|E| \geq 2$ . For every  $i \in I$ , define the sets  $J_i = \{j \in I : j > i\}$ ,  $S_i = \{u_j : j \in J_i\}$ ,  $\bar{S}_i = S_i \cup \{w_i\}$ . Let  $x^i$  and  $\bar{x}^i$  be the characteristic vectors of  $\{u_i, w_i\}$  and  $\bar{S}_i$ , respectively. Then, for every  $i \in I$ ,  $x^i \notin \mathcal{S}(G_n)$ , whereas  $\bar{x}^i \in \mathcal{S}(G_n)$ . Hence,  $\langle c, x^i \rangle > \delta \geq \langle c, \bar{x}^i \rangle$  holds, which implies

$$c_{u_i} + c_{w_i} > c_{w_i} + \sum_{j=i+1}^n c_{u_j} \quad \Leftrightarrow \quad c_{u_i} > \sum_{j=i+1}^n c_{u_j}.$$

Using this inequality iteratively, yields  $c_{u_1} > 2^{|E|-2} c_{u_k}$ , where  $k = \max I$ . Finally, since  $G_n$  is a threshold graph,  $c$  is integral, and  $u_k$  is contained in a minimal cover that is cut off by  $\langle c, x \rangle \leq \delta$ , Lemma 9 implies that  $c_{u_k} \geq 1$ , which yields the assertion.  $\square$

Using Proposition 14, we are able to prove Theorem 2.

*Proof of Theorem 2.* Let  $Cx \leq d$  be a binary formulation of  $\mathcal{S}(G_n)$  consisting of  $k \in [n]$  inequalities. Note that for every  $k \in [n]$  such a formulation exists, because  $G_n$  is threshold. Since there exist exactly  $n$  critical edges in  $E_n$ , there exists an inequality in  $Cx \leq d$  that cuts off  $\lceil \frac{n}{k} \rceil$  critical edges. By Proposition 14, this inequality has a coefficient that is greater than  $2^{\lceil \frac{n}{k} \rceil - 2}$ .  $\square$

An immediate consequence of Theorem 2 is that, for any undirected graph  $G$  that contains a graph  $G_n$  as an induced subgraph,  $\mathcal{S}(G)$  requires exponentially large coefficients in any reasonably small binary formulation, because the incidence vectors of critical edges of  $G_n$  are not contained in  $\mathcal{S}(G)$ . Thus, the graphs  $G_n$  provide a substructure of general undirected graphs that enforces large coefficients in any small binary formulation. In the future, it would be interesting to identify further substructures that make small binary (or integer) formulations of stable set or further problems numerically intractable. This, however, is out of scope of this article.

**Acknowledgments** The author is grateful to Volker Kaibel who gave the inspiration to this article.

## References

- [1] E. BALAS AND R. JEROSLOW, *Canonical cuts on the unit hypercube*, SIAM Journal on Applied Mathematics, 23 (1972), pp. 61–69.
- [2] V. CHVÁTAL AND P. L. HAMMER, *Aggregation of inequalities in integer programming*, Annals of Discrete Mathematics, 1 (1977), pp. 145–162.
- [3] C. D'AMBROSIO, A. FRANGIONI, L. LIBERTI, AND A. LODI, *On interval-subgradient and no-good cuts*, Operations Research Letters, 38 (2010), pp. 341–345.
- [4] Y. FAENZA, G. ORIOLO, AND G. STAUFFER, *Solving the weighted stable set problem in claw-free graphs via decomposition*, Journal of the ACM, 61 (2014), pp. 20:1–20:41.
- [5] S. FIORINI, S. MASSAR, S. POKUTTA, H. R. TIWARY, AND R. DE WOLF, *Linear vs. semidefinite extended formulations: Exponential separation and strong lower bounds*, in Proceedings of the Forty-fourth Annual ACM Symposium on Theory of Computing, STOC '12, New York, NY, USA, 2012, ACM, pp. 95–106.
- [6] A. GALLUCCIO AND A. SASSANO, *The rank facets of the stable set polytope for claw-free graphs*, Journal of Combinatorial Theory, Series B, 69 (1997), pp. 1–38.
- [7] F. GLOVER, *Unit coefficient inequalities for zero-one programming*, tech. rep., University of Colorado, 1973.
- [8] M. GRÖTSCHEL, L. LOVÁSZ, AND A. SCHRIJVER, *Stable Sets in Graphs*, Springer Berlin Heidelberg, Berlin, Heidelberg, 1988, pp. 272–303.

- [9] C. HOJNY, *Symmetries in Binary Programs—A Polyhedral Perspective*, PhD thesis, TU Darmstadt, 2018.
- [10] C. HOJNY, T. GALLY, O. HABECK, H. LÜTHEN, F. MATTER, M. E. PFETSCH, AND A. SCHMITT, *Knapsack polytopes: a survey*, Annals of Operations Research, (2019).
- [11] C. HOJNY AND M. E. PFETSCH, *Polytopes associated with symmetry handling*, Mathematical Programming, 175 (2019), pp. 197–240.
- [12] H. KELLERER, U. PFERSCHY, AND D. PISINGER, *Knapsack Problems*, Springer, Berlin Heidelberg, 2004.
- [13] D. LOKSHANTOV, M. VATSHELLE, AND Y. VILLANGER, *Independent set in  $p5$ -free graphs in polynomial time*, in Proceedings of the Twenty-fifth Annual ACM-SIAM Symposium on Discrete Algorithms, SODA '14, Philadelphia, PA, USA, 2014, Society for Industrial and Applied Mathematics, pp. 570–581.
- [14] C. MANNINO AND A. SASSANO, *An exact algorithm for the maximum stable set problem*, Computational Optimization and Applications, 3 (1994), pp. 243–258.
- [15] J. H. VAN LINT AND R. M. WILSON, *A Course in Combinatorics*, Cambridge University Press, Cambridge, 1993.