THE MAXIMUM k-COLORABLE SUBGRAPH PROBLEM AND ORBITOPES

TIM JANUSCHOWSKI AND MARC E. PFETSCH

ABSTRACT. Given an undirected node-weighted graph and a positive integer k, the maximum k-colorable subgraph problem is to select a k-colorable induced subgraph of largest weight. The natural integer programming formulation for this problem exhibits a high degree of symmetry which arises by permuting the color classes. It is well known that such symmetry has negative effects on the performance of branch-and-cut algorithms. Orbitopes are a polyhedral way to handle such symmetry and were introduced in [19].

The main goal of this paper is to investigate the polyhedral consequences of combining problem-specific structure with orbitope structure. We first show that the LP-bound of the integer programming formulation mentioned above can only be slightly improved by adding a complete orbitope description. We therefore investigate several classes of facet-defining inequalities for the polytope obtained by taking the convex hull of feasible solutions for the maximum k-colorable subgraph problem that are contained in the orbitope. We study conditions under which facet-defining inequalities for the polytope associated with the maximum k-colorable subgraph problem and the orbitope remain facet-defining for the combined polytope or can be modified to yield facets. It turns out that the results depend on both the structure and the labeling of the underlying graph.

1. INTRODUCTION

Symmetry in integer programs (IPs) has been recognized to harm the performance of branch-and-cut algorithms for a long time. One reason for the deteriorating effect of symmetries is that symmetric solutions appear repeatedly in the search tree, without giving new information about the optimal solution. In recent years, several successful methods to handle such symmetry have been developed. Margot [28] provides an excellent overview.

One line of research on symmetry handling aims at techniques to avoid visiting symmetric subproblems during search. However, the presence of symmetry may also be responsible for weak bounds of the linear programming (LP) relaxation: the barycenter of the solutions in an orbit with respect to the symmetry lies within the convex hull of all feasible solutions. Depending on the problem, this (fractional) point frequently has an objective that is far away from the optimal solution value of the IP. This leads to weak LP-bounds.

One way to deal with such weak LP-bounds is to handle symmetry via polyhedral methods. For a particular symmetry which, however, often appears in IP-models used in practice, this can be done via so-called orbitopes,

Date: March 2011.

introduced in [19]. Here, the IP-model contains an assignment structure in which objects should be assigned to symmetric object classes, i.e., there are binary variables x_{ij} that indicate whether object i should be assigned to class j. The symmetry arises from the fact that the classes are indistinguishable, i.e., we can arbitrarily permute the classes without changing the structure of the solution (the objective function value has to be invariant under this operation). In other words, the full symmetric group acts on the the columns of the matrix x. So-called packing orbitopes are the convex hull of those matrices that are lexicographically maximal within each orbit subject to the condition that each row of x contains at most one 1. Complete linear descriptions of these orbitopes can be given, see Section 3.1 for more precise definitions and results.

In this paper, we study the maximum k-colorable subgraph problem, which provides a particular application for the above methodology. Here, given an undirected (node weighted) graph and a positive integer k, the problem is to find a maximum induced subgraph that is k-colorable, i.e., we can assign colors to each node in the subgraph such that adjacent nodes receive different colors. Thus, the nodes correspond to the objects that have to be assigned to color classes, subject to the condition that each node is assigned to at most one class. This problem is interesting for its own sake. It has rarely been studied in the literature, supposedly because it is closely connected to both the more prominent graph coloring problem and the stable set problem.

The goal of this paper is to investigate the effect of symmetry handling on the linear programming bound and more generally on the polyhedral structure of the corresponding polytope. We show that the LP-bounds of a straight-forward IP-formulation of the maximum k-colorable subgraph problem only slightly improve when we add (complete) symmetry handling inequalities (Section 3.2). Thus, to make progress on the LP-bounds one has to include problem-specific structure, either by adding valid inequalities for the polytope corresponding to the original problem without symmetry handling or by considering the polytope in which symmetry has been handled by orbitopes. We briefly discuss the first possibility in Section 2. We then concentrate on the second possibility. Our main focus lies on giving conditions under which inequalities in the original IP-formulation of the maximum k-colorable subgraph problem and of orbitopes remain facets for the polytope obtained by taking the convex hull of feasible solutions for the maximum k-colorable subgraph problem that are contained in the orbitope. Whenever these inequalities do not define facets, we study how they can be strengthened to obtain facet-defining inequalities using problem-specific structure. We investigate each inequality in the defining IP-formulation. It turns out that the conditions for which the original or strengthened inequalities define facets can be quite technical. They clearly show, however, that the structure of the underlying graph of the maximum k-colorable subgraph problem is crucial for these results. Moreover, the results depend on the labeling of the nodes of the graph, i.e., they depend on the order of the rows of the matrix of assignment variables. Thus, the results of this paper can be used to distinguish different orderings of the nodes.

One general remark is that the methods studied in this paper do not take symmetries of the graph into account. Graph symmetry handling is discussed in [16], which also contains preliminary computational results evaluating the performance of a branch-and-cut approach.

A brief outline of this paper is as follows. We discuss related work in Section 1.1. In Section 1.2, we introduce the maximum k-colorable subgraph problem and state results on its computational complexity. In Section 2, we discuss several (facet-defining) inequalities that are valid for the polytope corresponding to the case in which symmetry has not been taken into account. In Section 3, we review orbitopes and show in Section 3.2 that the LP-bounds only slightly improve by adding orbitope symmetry handling inequalities. In Section 4, we study the polyhedral consequences of combining orbitope symmetry handling inequalities with problem-specific inequalities of the maximum k-colorable subgraph problem. We then deal with the different inequalities in turn. In Section 4.1, we consider packing inequalities, in Section 4.2, we study shifted column inequalities, and finally, in Section 4.3, we investigate clique shifted column inequalities.

This paper is partially based on the diploma thesis [15].

1.1. Related Work

As mentioned above, several approaches dealing with symmetries in integer programming have appeared in the literature. Most notably, in a series of papers, Margot [25, 26, 27] developed a methodology to (completely) handle symmetry within the branch-and-bound tree. Linderoth et al. [23, 32] investigate a variant based on branching on constraints, which allows to recompute symmetries for the subproblems in the tree. Orbitopes were introduced in [19]. Faenza and Kaibel [8] give a nice proof of the completeness of the linear description. Efficient fixing of variables using orbitope symmetry was developed in [18]; this paper also presents computational results for a graph partitioning problem. Friedman [10] discusses symmetry handling via so-called fundamental domains. We refer to Margot [28] for an overview of all these approaches. The handling of symmetries has also been extensively discussed in the constraint programming literature. Gent et al. [12] as well as Lecoutre and Tabary [21] provide overviews.

A further approach is to avoid symmetries by reformulation. For instance, Mehrotra and Trick [29] present a formulation of the graph coloring problem that avoids color symmetries. This can be adapted to the maximum k-colorable subgraph problem, but this formulation complicates the expression of the weighted maximum k-colorable subgraph problem. Another example for reformulation that avoids symmetries was presented by Valério de Carvalho [5] for the bin-packing problem.

The combination of orbitope and problem-specific structure was studied by Faenza [7] with respect to the graph coloring problem. Since the corresponding polytope is not full-dimensional, unfortunately the results and proofs become quite complicated. Our study on the maximum k-colorable subgraph problem avoids some of these obstacles. As mentioned above it is based on the thesis [15].

Polyhedral properties of the maximum k-colorable subgraph problem have been rarely treated so far. For $k = 2$, several polyhedral results exist. For example, Fouilhoux and Mahjoub [9] identify several facets and provide separation algorithms. As far as we know, the corresponding polytopes for arbitrary k have not be investigated. However, several applications of the maximum k-colorable subgraph problem appear in the literature. We list some of these here.

The *wave length assignment problem* arises in the design of optical networks (e.g., fiber glass telecommunication networks). Here, wave lengths have to be assigned to as many pre-routed light paths as possible such that intersecting light paths do not use the same wave length, see Koster and Scheffel [20]. A light path consists of links and fibers and intersect at links only. The number of wave lengths is restricted by the capacity of the fibers. Under the assumption that the capacity of each link is 1 and that each light path can receive at most one wavelength, this leads to a special case of the maximum k-colorable subgraph: wavelengths correspond to colors and light paths to nodes. Nodes are adjacent if the light paths share a link.

In the literature, the driving applications of the maximum k-colorable subgraph and especially the maximum bipartite induced subgraph problem have been so-called *via minimization problems*, see Fouilhoux and Mahjoub [9] and Narasimhan [30]. Via minimization problems arise in the design of integrated circuits. Other applications include *register allocation* and *job scheduling*, see Carlisle and Lloyd [2].

1.2. Notation, Basic Definitions and Complexity

We use the following notation throughout the article.

For an integer $n \geq 1$, let $[n] := \{1, 2, \ldots, n\}$. For $x \in \mathbb{R}^{[n]\times[k]}$ and $S \subseteq [n] \times [k]$, we write $x(S) := \sum_{i,j \in S} x_{ij}$. We use row_i for the set $\{(i, 1), (i, 2), \ldots, (i, k)\}$ and col_i for the set $\{(1, i), (2, i), \ldots, (n, i)\}$. By $\mathbf{0}$ we denote a zero-matrix of appropriate dimensions.

Let $G = (V, E)$ be a undirected graph consisting of a set of nodes V with cardinality $n := |V|$ and a set of edges $E \subseteq \binom{V}{2}$. We will often need an order of the nodes in V and thus assume that $V := \{1, \ldots, n\}$; in particular, we can directly compare nodes, e.g., $u < v$ for $u, v \in V$. Moreover, let $w : V \to \mathbb{R}$ be a weight function. We denote by $G[V']$ the subgraph induced by $V' \subseteq V$. Thus, the set of edges of $G[V']$ is $\{e \in E : e \in \binom{V'}{2}\}$.

For a positive integer k, G is said to be *k-colorable* if we can assign to each node in G a color (number) in $[k]$ such that adjacent nodes do not have the same color.

Definition 1. *For a positive integer k and the graph G with weights w, the* (weighted) maximum k-colorable subgraph problem *consists of finding a set $V' \subseteq V$ that induces a k-colorable subgraph and has maximum weight $\sum_{v \in V'} w(v)$. If $w(v) = 1$ for all $v \in V$, we obtain the* unweighted *maximum k-colorable subgraph problem.*

In this problem, any node that has a negative weight will not appear in an optimal solution. Thus, we will assume that w is non-negative in the following. Moreover, we will assume without loss of generality that G is simple,

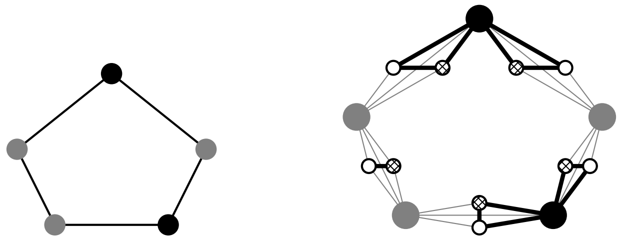

Figure 1. Illustration of the proof of Lemma 1 for $k = 3$. *Left:* A maximum stable set in the 5-cycle depicted with black nodes. *Right:* The resulting graph with a maximum (black,white,crossed)-colorable subgraph (for every edge we have added a K_2); nodes and edges in grey do not belong to the selected 3-colored subgraph.

because nodes contained in loops will never be part of an optimal solution, and parallel edges can be replaced by a single edge without changing the problem. Finally, we assume that $1 \leq k < n$, since otherwise the problem becomes trivial.

The above problem is related to the graph coloring problem, where the goal is to find the smallest k such that G is k-colorable. Obviously, by using binary search, one can reduce the coloring problem to the unweighted maximum k-colorable subgraph problem. Thus, the latter is (strongly) $\mathcal{NP}$-hard, if k is part of the input.

The maximum k-colorable subgraph problem is also related to the stable set problem in two ways: First, each set of nodes receiving the same color has to be a *stable set*, i.e., there are no edges between any two nodes in this set. The second relation can be derived as follows.

For a node v, the *neighborhood* $\Gamma(v)$ of node v contains all nodes adjacent to v, i.e.,

$$\Gamma(v) \coloneqq \{u \in V \,:\, \{u, v\} \in E\}.$$

The *closed neighborhood* of v is $\overline{\Gamma}(v) \coloneqq \Gamma(v) \cup \{v\}$. The *degree* of node v is defined as $\delta(v) \coloneqq |\Gamma(v)|$. A *clique* in a graph is a set of nodes $C \neq \varnothing$ such that for all $u, v \in C$, $u \neq v$, we have $\{u, v\} \in E$; a clique is *maximal*, if there is no other clique containing it. The *complete graph* on n nodes is a graph whose entire set of nodes $[n]$ forms a clique, and we denote it by K_n.

Definition 2 (Cartesian graph product)**.** *Let $G = (V, E)$ and $G' = (V', E')$ be two graphs. The* (Cartesian) product graph *$G \times G'$ of G and G' has the Cartesian product $V \times V'$ as its node set. For u, $v \in V$ and u', $v' \in V'$, $\{(u, u'), (v, v')\}$ is an edge in $G \times G'$ if and only if*

- *$u = v$ and $\{u', v'\} \in E'$, or*
- *$u' = v'$ and $\{u, v\} \in E$.*

It is easy to see that the weighted maximum k-colorable subgraph problem for G can be reduced to the weighted maximum stable set problem on the product graph $G \times K_k$, where the weights W of $G \times K_k$ are defined as $W(v, j) = w(v)$ for all $v \in V$, $j \in [k]$; this was observed first by Narasimhan [30].

The following result gives an extension of a result by Fouilhoux and Mahjoub [9] for $k = 2$ and deals with the reverse reduction.

Lemma 1. *Let $k \in \mathbb{Z}$, $1 \leq k < n$, be fixed. The maximum stable set problem can be reduced to the maximum k-colorable subgraph problem in polynomial time in n.*

Proof. Finding a stable set of maximum weight in a graph $G' = (V', E')$ can be reduced to finding a maximum k-colorable subgraph in a graph $G = (V, E)$ that is obtained from G' as follows. The graph G has all edges and nodes of G'. Additionally, for every edge $\{u, v\} \in E'$, we add a complete graph K_{k-1}, such that all nodes in K_{k-1} are adjacent to both u and v. Figure 1 shows an example of the construction.

We assign a sufficiently large positive weight to the nodes of the complete graphs such that they must be chosen in any optimal solution of the maximum k-colorable subgraph problem. Thus, each node $v \in V'$ has at least $k - 1$ neighbors colored with different colors in an optimal k-colorable subgraph of G. Hence, for $\{u, v\} \in E'$ nodes u and v cannot both be contained in an optimal subgraph for G. It follows that the nodes u, v are contained in a maximum k-colorable subgraph if and only if u, v are part of a maximum stable set.

The reduction is clearly polynomial in the size of G, w, and in k. □

It follows from Lemma 1 that the maximum k-colorable subgraph problem is (strongly) $\mathcal{NP}$-hard for any fixed k with $1 \leq k < n$. Note, however, that the reduction uses several different weights.

A stronger result follows from the work of Lewis and Yannakakis [22], who showed that the maximum subgraph problem for hereditary properties (fulfilling certain requirements) is $\mathcal{NP}$-hard. If we define the hereditary property as "the graph is k-colorable," we obtain the (unweighted) maximum k-colorable subgraph problem. Hence, the *unweighted* maximum k-colorable subgraph problem is $\mathcal{NP}$-hard for any fixed k, $1 \leq k < n$.

Moreover, Lund and Yannakakis [24] showed that such maximum subgraph problems for hereditary properties cannot be approximated within n^{ε} for some $\varepsilon > 0$ (depending on the property) unless $\mathcal{NP} = \mathcal{P}$; see [24] for stronger results under more restrictive assumptions. Thus, the (unweighted) maximum k-colorable subgraph problem is hard to approximate for any fixed k, $1 \leq k < n$; note that Lemma 1 does not provide an approximation preserving reduction. For $k = 1$, it as hard to approximate as the maximum stable set problem.

We briefly sum up results on the complexity of the maximum k-colorable subgraph for $k \in \{2, 3\}$. For $k = 2$, the maximum k-colorable subgraph problem is $\mathcal{NP}$-hard even in graphs with maximum degree three and in planar graphs with maximum degree greater than four, see [3]. For $k = 3$, the maximum k-colorable subgraph problem is $\mathcal{NP}$-hard in planar graphs, because we can reduce the problem of deciding whether a planar graph is 3-colorable to the maximum k-colorable subgraph problem.

2. The k-colored Subgraph Polytope

In this section, we will briefly collect results on the polytope corresponding to the maximum k-colorable subgraph problem (without symmetry handling).

After reviewing orbitope symmetry handling in the next section, we will then investigate the combination with symmetry handling in Section 4.

We consider the following IP-formulation for the maximum k-colorable subgraph problem:

$$(\mathrm{IP}_k(G)) \quad \max \sum_{v \in V} w(v) \sum_{j \in [k]} x_{vj} \tag{1}$$

$$x_{uj} + x_{vj} \leq 1 \qquad \forall \{u, v\} \in E,\ j \in [k] \tag{2}$$

$$x(\mathrm{row}_v) \leq 1 \qquad \forall v \in V \tag{3}$$

$$x_{vj} \in \{0, 1\} \qquad \forall v \in V,\ j \in [k]. \tag{4}$$

Let x be a solution of (2)–(4). Because of the *packing inequalities* (3), $\sum_{j \in [k]} x_{vj}$ is 1 if and only if node v is in the selected subgraph (i.e., colored). The *edge inequalities* (2) guarantee that adjacent nodes do not have the same color in a colored subgraph. Thus, the objective function represents the weight of the selected subgraph.

The *k-colored subgraph polytope* corresponding to the maximum k-colorable subgraph problem is

$$\mathrm{P}_k(G) \coloneqq \mathrm{conv}\{x \in \{0,1\}^{V \times [k]} \ :\ x \text{ satisfies (2) and (3)}\}.$$

For $u \in V$ and $j \in [k]$, let E^{uj} be the (u, j)-*unit matrix*, i.e., E^{uj} has a 1 at position (u, j) and is 0 otherwise. Since E^{uj}, for all $u \in V$ and $j \in [k]$, and the zero matrix are all contained in $\mathrm{P}_k(G)$ (since G is simple), we have $\dim(\mathrm{P}_k(G)) = n \cdot k$, i.e., the polytope is full-dimensional. For $k = 1$, $\mathrm{P}_k(G)$ is (isomorphic to) the *stable set polytope*, which we denote by $\mathrm{P}(G)$.

2.1. Valid Inequalities

In this section we mention some valid/facet-defining inequalities for $\mathrm{P}_k(G)$.

First, easy inspection shows that the *non-negativity inequalities* $x_{vj} \geq 0$ define facets of $\mathrm{P}_k(G)$ for all $v \in V$, $j \in [k]$. If $k > 1$, each inequality $x_{vj} \leq 1$, $v \in V$, $j \in [k]$, is dominated by the packing inequality $x(\mathrm{row}_i) \leq 1$. Moreover, the packing inequalities (3) define facets of $\mathrm{P}_k(G)$, see [15].

Remark 2. The polytope $\mathrm{P}_k(G)$ is down-monotone, i.e., if $x \leq y \in \mathrm{P}_k(G)$, then also $x \in \mathrm{P}_k(G)$. Hammer et al. [14] showed for such polytopes that every facet-defining inequality $\alpha^{\mathrm{T}} x \leq \beta$ has $\beta \geq 0$. If $\beta = 0$, it is a non-negativity inequality, which always defines a facet. If $\beta > 0$ then $\alpha_{vj} \geq 0$ for all $v \in V$, $j \in [k]$.

We obtain the following result by applying general projection theory, see, for instance, Theorem 16 in [17].

Lemma 3. *Let $j \in [k]$ and let*

$$\sum_{v \in V} \alpha_v \, x_{vj} \leq \beta$$

define a facet of $\mathrm{P}_k(G)$*. Then $\alpha^{\mathrm{T}} y \leq \beta$ defines a facet of* $\mathrm{P}(G)$*.*

Thus, the projection of facets whose support is restricted to a column are facets of the stable set polytope. Conversely, we can "trivially lift" facet-defining inequalities for $\mathrm{P}(G)$ to facet-defining inequalities of $\mathrm{P}_k(G)$.

Proposition 4. *Assume that*

$$\alpha^{\mathrm{T}} y \leq \beta \tag{5}$$

defines a non-trivial facet of $\mathrm{P}(G)$*. Then, the inequality*

$$\sum_{v \in V} \alpha_v \, x_{vj} \leq \beta, \tag{6}$$

defines a facet of $\mathrm{P}_k(G)$ *for all* $j \in [k]$*.*

Proof. Since (5) defines a non-trivial facet, we have $\alpha \geq 0$ and $\beta > 0$, using the results of Hammer et al. [14] for $P(G)$ (see also Remark 2). Moreover, there exits a set $\mathcal{A} \subseteq \mathrm{P}(G)$, $|\mathcal{A}| = n$, of affinely independent, non-zero points that fulfill inequality (5) with equality.

Let $j \in [k]$ be fixed. For each point $\tilde{x} \in \mathrm{P}(G)$, we define a matrix $\tilde{X} \in \mathrm{P}_k(G)$ whose jth column equals $\tilde{x}$ and is 0 otherwise. Then for each $\tilde{x} \in \mathcal{A}$, $\tilde{X}$ fulfills (6) with equality. To find additional $n(k-1)$ affinely independent matrices that fulfill (6) with equality, we proceed as follows.

For each $i \in [n]$, there exists a point $\tilde{x} \in \mathcal{A}$ with $\tilde{x}_i = 0$: otherwise, every point in the facet defined by (5) would satisfy $\tilde{x}_i = 1$, i.e., would be contained in the face defined by $x_i \leq 1$, a contradiction, since (5) defines a non-trivial facet. Thus, for $t \neq j$, the matrix $E^{it} + \tilde{X}$ does not violate any packing inequality and is contained in $\mathrm{P}_k(G)$ (recall that E^{it} is the (i,t)-unit matrix). Moreover, it fulfills (6) with equality.

In total, we have $n \cdot k$ many matrices that fulfill (6) with equality. It is easy to see that these matrices are affinely independent. Thus, (6) defines a facet of $\mathrm{P}_k(G)$. □

Hence, several well known inequality classes for $\mathrm{P}(G)$ are also valid for $\mathrm{P}_k(G)$, e.g., *odd hole*, *odd antihole*, and *odd wheel inequalities*, see Grötschel et al. [13]. We will often use *clique inequalities* $\sum_{v \in C} x_{vj} \leq 1$, for $j \in [k]$, in the following. It is well known that clique inequalities $\sum_{v \in C} y_v \leq 1$ define facets of $\mathrm{P}(G)$ if and only if the clique C is maximal; by Lemma 3 and Proposition 4 the same holds for clique inequalities with respect to $\mathrm{P}_k(G)$.

Not all valid inequalities for $\mathrm{P}_k(G)$ can be derived from $\mathrm{P}(G)$. As noted above, the maximum k-colorable subgraph problem is a stable set problem in the product graph $G \times K_k$. Thus, one can obtain inequalities from this (larger) stable set problem as well. It is easy to see that the only cliques in $G \times K_k$ arise from cliques in G or from each K_k; thus, the product graph does not give new clique inequalities. However, in $G \times K_k$ new odd holes can appear. Hence, odd hole inequalities for $\mathrm{P}_k(G)$ exist that are not trivially lifted inequalities from $\mathrm{P}(G)$.

2.2. Cases of Complete Linear Descriptions of $\mathrm{P}_k(G)$

Using results in the literature on the perfectness of product graphs, one can characterize graphs for which the LP-relaxation of $\mathrm{IP}_k(G)$ together with clique inequalities gives a complete description of $\mathrm{P}_k(G)$. We provide this section for completeness and refer to Grötschel et al. [13] for a discussion of perfect graphs.

We need the following classes of perfect graphs. A graph is a *TDF graph* (or (C_{n+4},diamond)-free graph), if it contains no induced diamond and is triangulated, i.e., the graph does not contain an induced cycle C_n of size $n \geq 4$, see [1]. A *parity graph* has the property that for any pair of nodes u and v, all induced paths between u and v have the same parity, i.e., odd or even length. Olaru and Sachs [31] have shown that parity graphs are perfect.

Theorem 5. *Maximal clique, packing, and non-negativity inequalities provide a complete description of* $\mathrm{P}_k(G)$ *if and only if*

- *for* $k = 1$*,* G *is a perfect graph,*
- *for* $k = 2$*,* G *is a parity graph, or*
- *for* $k > 3$*,* G *is a TDF graph.*

Proof. For $k = 1$, the result follows from the classical characterization of the stable set polytope for perfect graphs by Fulkerson [11] and Chvatal [4].

Recall that, for $k \geq 2$, $G \times K_k$ does not contain cliques other than the ones induced by the packing inequalities (i.e., maximal cliques in K_k) or the cliques in G. For $k = 2$, De Werra and Hertz [6] showed that $G \times K_k$ is perfect, if and only if, G is a parity graph. Moreover, for $k > 2$, they showed that $G \times K_k$ is perfect, if and only if, G is a TDF graph. □

3. Symmetry Handling and LP-Bounds

Despite the knowledge we have about $\mathrm{P}_k(G)$, the corresponding IP-formulation still suffers from symmetry. We therefore later propose to combine $\mathrm{P}_k(G)$ with a polytope that handles symmetries. Let us first define what we mean by symmetry.

Definition 3. *A* symmetry *of an integer program is a permutation of the variables that maps feasible solutions to feasible solutions with the same objective function.*

The symmetries define the *symmetry group* of an IP.

The unweighted version of $(\mathrm{IP}_k(G))$ exhibits two basic types of symmetries: graph and color symmetries. Color symmetries arise by the possibility of (arbitrarily) permuting the colors. This corresponds to an operation of the symmetric group on the columns of $(\mathrm{IP}_k(G))$. Graph symmetries arise from the fact that the underlying graph may have non-trivial automorphisms. For a study of graph symmetry handling, we refer the reader to [16] – in this paper, we concentrate on color symmetries.

3.1. Orbitopes

In order to handle symmetries such as the color symmetries of the maximum k-colorable subgraph problem, orbitopes were introduced in [19]. In the following, we summarize those results on orbitopes that we need for the exposition of our results. In order to avoid trivial cases, we will assume for the rest of this paper that $n > k > 1$.

We denote by $\mathcal{M}_{n,k} := \{x \in \{0,1\}^{[n]\times[k]} : x(\mathrm{row}_i) \leq 1 \text{ for all } i \in [n]\}$, the 0/1-matrices with row-sum at most 1. Let $\prec$ be the lexicographic ordering of $\mathcal{M}_{n,k}$ with respect to a row-wise flattening of x, i.e., the ordering $(1,1) < (1,2) < \cdots < (1,k) < (2,1) < (2,2) < \cdots < (2,k) < \ldots < (n,k)$ of matrix positions. We have $A \prec B$ with $A = (a_{ij})$, $B = (b_{ij}) \in \mathcal{M}_{n,k}$ if and only if $a_{k\ell} < b_{k\ell}$, where (k, ℓ) is the first position with respect to the ordering where A and B differ. Let $\mathfrak{S}_k$ be the symmetric group on $[k]$. Then $\mathcal{M}_{n,k}^{\max}$ is the set of matrices in $\mathcal{M}_{n,k}$ that are maximal with respect to $\prec$ within their orbits under the group action of $\mathfrak{S}_k$. It is not hard to see that a matrix is in $\mathcal{M}_{n,k}^{\max}$ if and only if its columns are sorted lexicographically. We define the *packing orbitope* (with respect to $\mathfrak{S}_k$) as

$$\mathrm{O}_{n,k} \coloneqq \mathrm{conv}(\mathcal{M}_{n,k}^{\max}).$$

In [19], other types of orbitopes are investigated. Since we only consider packing orbitopes with respect to $\mathfrak{S}_k$, in the following, we refer to them simply as *orbitopes*. The paper [19] shows that the linear optimization problem over orbitopes can be solved in $\mathrm{O}(n^2 k)$ time. Moreover, a complete non-redundant linear description of orbitopes by exponentially many linear inequalities is derived. We need further notation for the exposition of this result.

Because of the lexicographic ordering, the equations $x_{ij} = 0$ for all $i < j$ are valid for orbitopes. We therefore define the index set

$$\mathcal{I}_{nk} \coloneqq \{(i,j) \in [n] \times [k] \,:\, i \geq j\}$$

of the remaining variables. Using $k(i) \coloneqq \min\{i, k\}$, we adjust the definitions of $\mathrm{row}_i \coloneqq \{(i,1), (i,2), \dots, (i, k(i))\}$ and $\mathrm{col}_j \coloneqq \{(j,j), (j+1, j), \dots, (n,j)\}$. For $(i,j) \in \mathcal{I}_{nk}$, we denote by

$$\mathrm{col}(i,j) \coloneqq \{(j,j), (j+1,j), \dots, (i-1,j), (i,j)\} \subseteq \mathcal{I}_{nk}$$

the column up to row i. For elements in $\mathcal{I}_{nk}$, we introduce a system of "diagonal" coordinates in the following way:

$$\langle \eta, j \rangle \coloneqq (j + \eta - 1, j) \quad \text{for } j \in [k],\ 1 \leq \eta \leq n - j + 1.$$

Thus, η determines the index of the diagonal containing $(j + \eta - 1, j)$.

For a given $(i,j) = \langle \eta, j \rangle \in \mathcal{I}_{nk}$, with $i > 1$ and $j > 1$, define $B = \{(i,j), (i,j+1), \dots, (i, k(i))\}$ to be the *bar* with leader (i,j). The inequality

$$x(B) - x(\mathrm{col}(i-1, j-1)) \leq 0,$$

is a so-called *column inequality*. A point $x \in \{0,1\}^{\mathcal{I}_{nk}}$ is contained in the orbitope if and only if x satisfies all packing inequalities and all column inequalities, see [19]. To obtain the complete linear description, we need the following generalization of column inequalities.

Consider a *shifted column*

$$S = \{\langle 1, c_1 \rangle, \langle 2, c_2 \rangle, \dots, \langle \eta, c_\eta \rangle\} \subseteq \mathcal{I}_{nk}$$

with $c_1 \leq c_2 \leq \cdots \leq c_\eta$. It is called a *shifting* of each of the columns $\mathrm{col}(\eta, c_\eta), \mathrm{col}(\eta, c_\eta + 1), \dots, \mathrm{col}(\eta, k)$. We obtain the corresponding *shifted column inequality* (SCI)

$$x(B) - x(S) \leq 0.$$

An example of an SCI appears on the left-hand side of Figure 12.

SCIs can be separated in linear time. The orbitope is completely described by packing inequalities, non-negativity inequalities, and SCIs [19]. Faenza and Kaibel [8] provide an alternative, elegant proof of this result.

3.2. LP-Bounds

In this section, we study the quality of the upper bound obtained via the unweighted LP-relaxation ($\mathrm{LP}^1_k(G)$) of ($\mathrm{IP}_k(G)$), i.e., we relax all integrality conditions (4) on the variables and use $w(v) = 1$ for all $v \in V$. We will then investigate whether the bound improves if we add equations $x_{vj} = 0$ for all $v < j$ and SCIs in order to handle symmetry. Let $\mathrm{opt}(\cdot)$ be the value of some LP/IP-formulation.

Proposition 6. $\mathrm{opt}(\mathrm{LP}^1_k(G)) = n$.

Proof. A solution with $x_{ij} = \frac{1}{k}$ for every $(i, j) \in V \times [k]$ has objective n and satisfies all inequalities independent of the graph. Moreover, n is clearly an upper bound on $\mathrm{opt}(\mathrm{LP}^1_k(n))$ (add all packing inequalities (3)). □

Remark 7. The solution in the previous proof lies in the relative interior of the face defined by all packing inequalities. It corresponds to the barycenter of the orbit of a (hypothetical) solution that colors all nodes.

Remark 8. Since $\mathrm{opt}(\mathrm{IP}_k(K_n)) = k$ (for unit weights), the above result shows that the gap between the value of the LP-relaxation and the optimal value can be arbitrarily large.

It is natural to investigate the strength of the following unweighted LP relaxation that arises by adding a full description of the corresponding orbitope via SCIs:

$$
\begin{aligned}
(\mathrm{OLP}^1_k(G)) \quad \max \sum_{v\in V}\sum_{j\in[k]} x_{vj} & \\
x_{uj} + x_{vj} \le 1 \quad & \forall \{u, v\} \in E,\ j \in [k] \\
x(\mathrm{row}_v) \le 1 \quad & \forall v \in V \\
x(B) - x(S) \le 0 \quad & \forall \text{ bars } B \text{ and corr. SCs } S \\
x_{vj} = 0 \quad & \forall v \in V,\ j \in [k] \text{ with } v < j \\
x_{vj} \in [0, 1] \quad & \forall v \in V,\ j \in [k].
\end{aligned}
$$

We obtain the following bound on the value of this formulation.

Proposition 9. $\mathrm{opt}(\mathrm{OLP}^1_k(G)) \ge n - \frac{1}{2^{k-1}}$.

Proof. We construct a feasible solution x of ($\mathrm{OLP}^1_k(G)$) with objective value equal to $n - \frac{1}{2^{k-1}}$ as follows:

$$
x_{ij} := \begin{cases} 1 - \frac{1}{2^{k-1}} & \text{if } (i, j) = (1, 1) \text{ or } (i, j) = (2, 2) \\ \frac{1}{2^{k-1}} & \text{if } i \ge 2 \text{ and } j = 1 \\ \frac{1}{2^{k-j+1}} & \text{if } i \ge 3 \text{ and } 2 \le j < k(i) \\ 1 - \frac{1}{2^{k-j+1}} & \text{if } i \ge 3 \text{ and } j = k(i) \\ 0 & \text{otherwise.} \end{cases}
$$

Figure 2 contains an illustration of the construction. It is not hard to see that for $i \ge 2$, the sum of the ith row is 1, while $x_{11} = 1 - \frac{1}{2^{k-1}}$. Thus the objective value is $n - \frac{1}{2^{k-1}}$ as claimed. Moreover, each entry off the diagonal

	1	2	3	$\cdots$	j	$\cdots$	k	
1	$1-\frac{1}{2^{k-1}}$							
2	$\frac{1}{2^{k-1}}$	$1-\frac{1}{2^{k-1}}$						$\Sigma = 1$
3	$\vdots$	$\frac{1}{2^{k-1}}$	$1-\frac{1}{2^{k-2}}$	$\ddots$				$\Sigma = 1$
$\vdots$		$\vdots$	$\frac{1}{2^{k-2}}$		$1-\frac{1}{2^{k-j+1}}$	$\ddots$		$\Sigma = 1$
			$\vdots$	$\vdots$			$\frac{1}{2}$	$\Sigma = 1$
					$\vdots$		$\vdots$	$\vdots$
n	$\frac{1}{2^{k-1}}$	$\frac{1}{2^{k-1}}$	$\frac{1}{2^{k-2}}$		$\frac{1}{2^{k-j+1}}$		$\frac{1}{2}$	$\Sigma = 1$

Figure 2. The solution of $(\mathrm{OLP}^1_k(G))$ as constructed in the proof of Proposition 9 with objective value $n - \frac{1}{2^{k-1}}$.

has value $\leq \frac{1}{2}$ and the sum of any two entries in a column is at most 1. Thus, all edge inequalities are satisfied.

We next show that x fulfills all shifted column inequalities. First observe that $x_{ij} < 1$ for $j > 1$ and $x(\mathrm{col}(i-1, j-1)) \geq 1$ for $i \neq j$, $i, j > 1$. This is also true for shiftings of $\mathrm{col}(i-1, j-1)$. If $i = j$, then the shifted column inequalities hold trivially because $x_{ii} \leq x_{\ell\ell}$ for $i > \ell$. $\square$

Proposition 9 shows that the possible improvement obtained by adding a complete orbitopal description is small; in fact, it decreases as k increases. The slight improvement is due to the fixing of variables on the top right. Thus, the most primitive form of symmetry handling by fixing variables is in general as strong as its most sophisticated form via exponentially many inequalities. Moreover, the gap to the optimal integral value remains arbitrarily large.

Thus, in order to improve the value of the LP-relaxation, we need to take the structure of the graph into account. Clearly, adding cutting planes like clique inequalities will often improve the LP-bound. However, relevant instances are known in which the clique number is 2, i.e., clique inequalities will not improve the bound. Another way to improve the LP-bound beyond the basic formulation is by inequalities that combine symmetry handling with the structure of the graph. Such inequalities will be investigated in the next section.

4. Combining k-colored Subgraph Polytopes with Orbitopes

In the following, we study the convex hull of the intersection of the orbitope with the k-colored subgraph polytope, i.e., we define

$$\mathrm{OP}_k(G) \coloneqq \mathrm{conv}(\{x \in \mathrm{P}_k(G) \cap \mathcal{M}^{\max}_{n,k}\}),$$

which clearly is a 0/1-polytope. Note that all facets of $\mathrm{P}_k(G)$ and $\mathrm{O}_{n,k}$ remain *valid* for $\mathrm{OP}_k(G)$. This section is devoted to conditions under which facets remain facets or can be modified to yield facets.

The polytope $\mathrm{OP}_k(G)$ adds orbitope symmetry breaking inequalities to the *integer* programming formulation $(\mathrm{IP}_k(G))$. In this way we obtain combinations of symmetry handling with problem-specific structure.

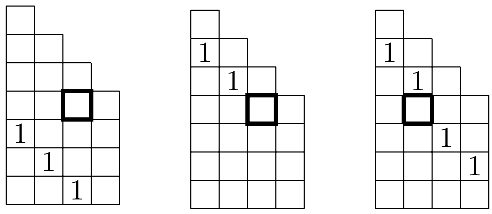

Figure 3. Matrices $V^{\langle\sigma,t\rangle}$ with highlighted $\langle\sigma,t\rangle$ in the proof of Proposition 10. *Left:* Case (7): $V^{\langle 5,3\rangle}$. *Middle:* Case (8): $V^{\langle 2,2\rangle}$. *Right:* Case (9): $V^{\langle 3,4\rangle}$.

We obtain the following basic result on the dimension and non-negativity constraints.

Proposition 10.

- *We have* $\dim(\mathrm{OP}_k(G)) = |\mathcal{I}_{nk}|$, *i.e.,* $\mathrm{OP}_k(G)$ *is full-dimensional.*
- *Let* $\langle\eta,j\rangle \in \mathcal{I}_{nk}$. *The non-negativity inequality* $x_{\langle\eta,j\rangle} \geq 0$ *defines a facet of* $\mathrm{OP}_k(G)$, *if and only if,* $\eta \neq 1$ *or* $\eta = 1$ *and* $j = k$. *If* $\eta = 1$, *the faces defined by* $x_{\langle\eta,j\rangle} \geq 0$ *for* $1 \leq j < k$ *are contained in the facet defined by* $x_{\langle 1,k\rangle} \geq 0$.

Proof. The proof is elementary, but shows the techniques needed in the following. For $\langle\sigma,t\rangle = (s,t) \in \mathcal{I}_{nk}$, we define

$$\mathrm{diag}^{\leq}\langle\sigma,t\rangle = \mathrm{diag}^{\leq}(s,t) := \{\langle\sigma,1\rangle, \langle\sigma,2\rangle, \dots, \langle\sigma,t\rangle\}$$

as the subset of indices of a diagonal that starts at $\langle\sigma,1\rangle$ and ends at $\langle\sigma,t\rangle$. We define the matrix $W^{\leq\langle\sigma,t\rangle} = W^{\leq(s,t)}$ as the incidence matrix of the set $\mathrm{diag}^{\leq}\langle\sigma,t\rangle$. Any such matrix corresponds to a colored subgraph with lexicographically ordered columns.

For the proof of the dimension, we note that for $\langle\sigma,t\rangle \in \mathcal{I}_{nk}$, we have $W^{\leq\langle\sigma,t\rangle} \in \mathrm{OP}_k(G)$. Together with $\mathbf{0} \in \mathrm{OP}_k(G)$, they form $|\mathcal{I}_{nk}| + 1$ affinely independent matrices.

Consider non-negativity inequalities with $\eta > 1$ first. For $\langle\sigma,t\rangle \in \mathcal{I}_{nk}$ with $\langle\sigma,t\rangle \neq \langle\eta,j\rangle$, we define the following matrix as depicted in Figure 3:

$$V^{\langle\sigma,t\rangle} := \begin{cases} W^{\leq\langle\sigma,t\rangle} & \text{if } \sigma \neq \eta & (7)\\ W^{\leq\langle\sigma,t\rangle} & \text{if } \sigma = \eta \text{ and } t < j & (8)\\ W^{\leq\langle\sigma-1,j\rangle} + W^{\leq\langle\sigma,t\rangle} - W^{\leq\langle\sigma,j\rangle} & \text{if } \sigma = \eta \text{ and } t > j. & (9)\end{cases}$$

In case (9) we use the assumption that $\eta \neq 1$. Additionally, we define $V^{\langle\eta,t\rangle} = \mathbf{0}$. Each matrix $V^{\langle\sigma,t\rangle} \in \mathrm{OP}_k(G)$ fulfills inequality $x_{\langle\eta,j\rangle} \geq 0$ with equality. It is easy to see that these matrices form $|\mathcal{I}_{nk}|$ many affinely independent matrices, proving that $x_{\langle\eta,j\rangle} \geq 0$ defines a facet.

Consider $\eta = 1$ next. Any matrix that fulfills $x_{\langle 1,j\rangle} \geq 0$, $j < k$, with equality also fulfills $x_{\langle 1,k\rangle} \geq 0$ with equality, since it must have 0-entries in positions $\langle 1,t\rangle$ with $j < t \leq k$ due to the lexicographic ordering of the columns. We can choose matrices $W^{\leq\langle\sigma,t\rangle}$ for $\langle\sigma,t\rangle \neq \langle 1,k\rangle$ and $\mathbf{0}$ to prove that $x_{\langle 1,k\rangle} \geq 0$ defines a facet. □

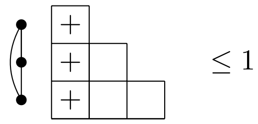

Figure 4. For K_3 and $k = 3$, the clique inequality for color 1 dominates all packing inequalities.

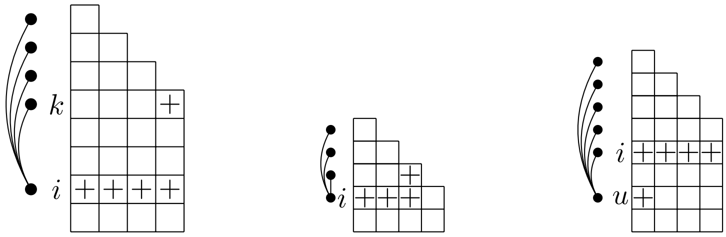

Figure 5. Proof of Theorem 13: Inequalities that dominate the packing inequality. *Left:* Node $i > k$ is adjacent to all nodes in $[k]$. *Middle:* Node $i \leq k$ is adjacent to all nodes in $[i-1]$. *Right:* Node $u > i$ is adjacent to all nodes in $[i]$.

Remark 11. Note that $\mathrm{OP}_k(G)$ is not monotone. Hence, the results of Hammer et al. [14] cannot be applied. Indeed, we will see non-trivial homogeneous inequalities and inequalities with positive and negative coefficients.

4.1. Packing Inequalities

The following example illustrates that, in contrast to $\mathrm{O}_{n,k}$ and $\mathrm{P}_k(G)$, packing inequalities do not always define facets for $\mathrm{OP}_k(G)$.

Example 12. Any packing inequality for $\mathrm{OP}_3(K_3)$ is dominated by the clique inequality $x_{11} + x_{21} + x_{31} \leq 1$, see Figure 4. Due to the lexicographic ordering, any solution with an entry 1 that is not in the first column must have a 1 in the first column. Hence, the above clique inequality holds with equality. Thus, in $\mathrm{OP}_3(K_3)$ no packing inequality defines a facet.

Theorem 13. *The packing inequality*

$$x(\mathrm{row}_i) \leq 1$$

defines a facet of $\mathrm{OP}_k(G)$*, if and only if, the following conditions hold.*

(1) *There exists a node* $v \in [k(i)]$ *such that* $v \notin \overline{\Gamma}(i)$*, and*
(2) *for every node* $u \in \{i+1, i+2, \ldots, n\}$ *there exists a node* $w_u \in [i]$ *such that* $w_u \notin \Gamma(u)$*.*

Proof. "$\Rightarrow$" We show that the conditions are necessary. We first assume for the sake of contradiction that $[k(i)] \subseteq \overline{\Gamma}(i)$. If $i > k$, consider the inequality

$$x(\mathrm{row}_i) + x_{kk} \leq 1, \tag{10}$$

which is depicted on the left-hand side of Figure 5. Let us show the validity of Inequality (10), which contradicts the assumption that the packing inequality is facet-defining. Any $\tilde{x} \in \mathrm{OP}_k(G)$ with $\tilde{x}_{kk} = 1$ must have $\tilde{x}_{tt} = 1$ for $t \in [k]$ due to the lexicographic ordering of the columns. Due to the adjacency of i to any node in $[k]$, it follows that $\tilde{x}(\mathrm{row}_i) = 0$, so Inequality (10)

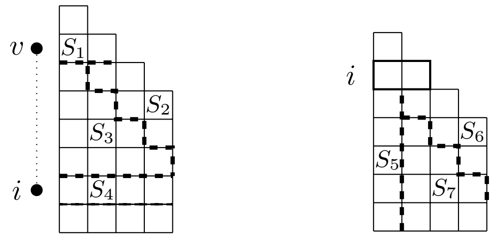

Figure 6. The sets used in the proof of Theorem 13. *Left:* The sets S_1–S_4 on or above row i. *Right:* The sets S_5–S_7 below row i.

is valid. Inequality (10) dominates the packing inequality for row i, since $W^{\leq(k,k)}$ fulfills (10) with equality, but not the packing inequality.

If $i \leq k$, consider

$$\sum_{j=1}^{i-1} x_{ij} + x_{i-1,i-1} \leq 1. \tag{11}$$

The non-zero coefficients in Inequality (11) are shown in the middle of Figure 5. Let us show the validity of Inequality (11). Any $\tilde{x} \in \mathrm{OP}_k(G)$ with $\tilde{x}_{i-1,i-1} = 1$ has $\tilde{x}_{tt} = 1$ for $t \in [i-1]$ due to the lexicographic ordering of the columns. Due to the adjacency of i to any node in $[i-1]$, it follows that $\tilde{x}(\mathrm{row}_i) = 0$, so Inequality (10) is valid. Note that even though position (i,i) does not appear in (11), a feasible solution $\tilde{x}$ with $\tilde{x}_{ii} = 1$ fulfills Inequality (11) with equality, since this matrix must have 1s along the main diagonal due to the lexicographic ordering. Moreover, matrix $W^{\leq(i-1,i-1)}$ fulfills (11) with equality, but does not fulfill the packing inequality with equality.

For handling the second condition, we assume that there exist a node $u > i$ such that $[i] \subseteq \Gamma(u)$. In this case, consider inequality

$$\sum_{j=1}^{k(i)} x_{ij} + x_{u1} \leq 1, \tag{12}$$

which clearly dominates the packing inequality for row i; see the right of Figure 5 for an illustration. The validity can be seen as follows. A feasible solution $\tilde{x}$ with $\tilde{x}_{i,1} = 1$ and $\tilde{x}_{u,1}$ violates an edge inequality. Moreover, if $\tilde{x}_{u,1} = 1$, then it follows that $\tilde{x}_{1,1} = \tilde{x}_{2,1} = \cdots = \tilde{x}_{i,1} = 0$ because of the edge inequalities. Thus, due to the lexicographic ordering, $\tilde{x}(\mathrm{row}_i) = 0$.

"$\Leftarrow$" We show that the conditions are sufficient. Let $\mathcal{V}$ be the set of vertices of $\mathrm{OP}_k(G)$ that satisfy the packing inequality for row i with equality. Define $\mathcal{L} := \mathrm{lin}(\mathcal{V})$ as the linear span of the elements in $\mathcal{V}$. We shall show that the unit matrix E^{st} with $(s,t) \in \mathcal{I}_{nk}$ is contained in $\mathcal{L}$. It follows that $\mathcal{L}$ contains $|\mathcal{I}_{nk}|$ linearly independent points and since $\mathbf{0} \notin \mathcal{V}$, we have $\dim(\mathrm{aff}(\mathcal{V})) = |\mathcal{I}_{nk}| - 1$. Hence, the packing inequality for row i defines a facet of $\mathrm{OP}_k(G)$.

The proof relies on the fact that there exists a node $v < i$ such that $v \notin \overline{\Gamma}(i)$. In order to show that $E^{st} \in \mathcal{L}$ for all $(s,t) \in \mathcal{I}_{nk}$, we partition the

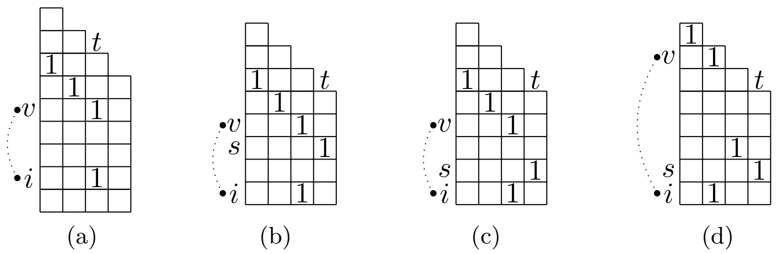

Figure 7. Matrices in the proof of Theorem 13 for the sets above row_i. 7a: Term (13) of Claim 1. 7b: Term (21) of Claim 2. 7c: Term (22) of Claim 3. 7d: Term (24) of Claim 3.

set $\mathcal{I}_{nk}$ into the following seven sets:

$$S_1 := \{(s,t) \in \mathcal{I}_{nk} \ : \ s \leq v\}$$
$$S_2 := \{\langle \sigma, t\rangle = (s,t) \in \mathcal{I}_{nk} \ : \ v < s < i, \ \sigma \leq v\}$$
$$S_3 := \{\langle \sigma, t\rangle = (s,t) \in \mathcal{I}_{nk} \ : \ v < s < i, \ \sigma > v\}$$
$$S_4 := \text{row}_i$$
$$S_5 := \{(s,t) \ : \ t = 1, \ s > i\}$$
$$S_6 := \{(s,t) = \langle \sigma, t\rangle \ : \ s > i, \ \sigma \leq i\}$$
$$S_7 := \{(s,t) = \langle \sigma, t\rangle \ : \ s > i, \ \sigma > i\}.$$

Figure 6 contains a graphic depiction of the sets. On the left-hand side, we have the sets S_1, S_2, S_3, and S_4 corresponding to the parts in or above row i. The right-hand side shows the sets S_5, S_6, and S_7 below row i. The following seven claims will deal with each set in turn.

Claim 1. *For all* $(s,t) \in S_1$, $E^{st} \in \mathcal{L}$.

Proof. Consider the case $s = v$ first. Assume v, $t > 1$. We construct a linear combination of vertices of $\text{OP}_k(G)$ that yields the unit matrix E^{st}. Each vertex of $\text{OP}_k(G)$ in the following linear combination fulfills the packing inequality for row i with equality:

$$E^{st} = E^{vt} = (W^{\leq(v,t)} + E^{it}) \tag{13}$$
$$- (W^{\leq(v-1,t-1)} + E^{it}). \tag{14}$$

Figure 7a contains a depiction of term (13). Both (13) and (14) form vertices of $\text{OP}_k(G)$, since they are 0/1-matrices and have lexicographically ordered columns; in term (13) we use the assumption that $\{v,i\} \notin E$, so no edge inequality is violated. By definition of the matrices $W^{\leq(s,t)}$ we have $E^{vt} = W^{\leq(v,t)} - W^{\leq(v-1,t-1)}$.

If v or t are 1, we use $E^{vt} = (E^{vt} + E^{it}) - E^{it}$ as the linear combination of vertices of $\text{OP}_k(G)$.

Consider $1 < s < v$ and $t \neq 1$. Then, the linear combination is:

$$E^{st} = (W^{\leq(s,t)} + E^{i,t+1}) \tag{15}$$
$$- (W^{\leq(s-1,t-1)} + E^{i,t+1} + E^{vt}) \tag{16}$$
$$+ E^{vt}. \tag{17}$$

Term (17) is not a vertex of $\mathrm{OP}_k(G)$ (it does not have lexicographically ordered columns), but we have already proved it to be contained in $\mathcal{L}$. For (15) and (16) we use the following. For $k < i$, we use that $v \leq k$, so we have $t + 1 \leq k$ since $s < v$. For $i = k$, we use that $v < i$, so we have $t + 1 < k$ as well. Thus $E^{st} \in \mathcal{L}$ in this case.

For $s = 1$ or $t = 1$, the linear combination is

$$
\begin{aligned}
E^{st} = {} & (E^{st} + E^{v,t+1} + E^{i,t+1}) && (18)\\
& - (E^{i,t+1} + E^{vt}) && (19)\\
& - (E^{v,t+1} - E^{vt}). && (20)
\end{aligned}
$$

Note that if $s = 1$ then $t = 1$, i.e., in any case $t = 1$. We have proved already that both summands in (20) are in $\mathcal{L}$. Since we have $k \geq 2$, for $t = 1$ we have $t + 1 \in [k]$ and term (19) is a vertex of $\mathrm{OP}_k(G)$. In term (18), we use that $\{v, i\} \notin E$. □

Claim 2. *For all $(s, t) = \langle\sigma, t\rangle \in S_2$, $E^{st} \in \mathcal{L}$.*

Proof. For $(s, t) = \langle\sigma, t\rangle \in S_2$, we have

$$\varnothing \neq \mathrm{diag}^{\leq}\langle\sigma, t\rangle \cap \mathrm{row}_v = \{\langle\sigma, v - \sigma + 1\rangle\}$$

and $t \geq 2$ by definition of S_2. Hence, the linear combination for E^{st} is:

$$
\begin{aligned}
E^{st} = {} & (W^{\leq(s,t)} + E^{i,v-\sigma+1}) && (21)\\
& - (W^{\leq(s-1,t-1)} + E^{i,v-\sigma+1}).
\end{aligned}
$$

We depict the matrix in (21) in Figure 7b. Both terms in parentheses are in $\mathcal{V}$, where we have used that $\{v, i\} \notin E$. □

For the proof of the next claim, we need further notation. Let $(s, t) = \langle\sigma, t\rangle \in \mathcal{I}_{nk}$. We define

$$\mathrm{diag}^{\leq}\langle\sigma, t\rangle|_j := \{\langle\sigma, j + 1\rangle, \langle\sigma, j + 2\rangle, \ldots, \langle\sigma, t\rangle\}$$

and let $W^{\leq\langle\sigma,t\rangle|_j}$ be the corresponding incidence matrix.

Claim 3. *For all $(s, t) = \langle\sigma, t\rangle \in S_3$, $E^{st} \in \mathcal{L}$.*

Proof. Consider $t = 1$ first. The linear combination is:

$$E^{s1} = (E^{s1} + E^{i2}) - (E^{v1} + E^{i2}) + E^{v1},$$

where we use that $k \geq 2$. As shown in Claim 1, we have $E^{v1} \in \mathcal{L}$.

Consider $t > 1$ with $t - 1 \leq k(v)$. Then

$$
\begin{aligned}
E^{st} = {} & (W^{\leq(v,t-1)} + E^{st} + E^{i,t-1}) && (22)\\
& - (W^{\leq(v,t-1)} + E^{i,t-1}). && (23)
\end{aligned}
$$

We depict (22) in Figure 7c, where we use $t > 1$. Since $t - 1 \leq k(v)$, we know that $(v, t - 1) \in \mathrm{row}_v$. In (23), we use that $\{v, i\} \notin E$.

For the case that $t - 1 > k(v)$, the linear combination is:

$$
\begin{aligned}
E^{st} = {} & (W^{\leq(v,k(v))} + E^{i,k(v)} + W^{\leq\langle\sigma,t\rangle|_{k(v)}}) && (24)\\
& - (W^{\leq\langle\sigma,t-1\rangle|_{k(v)}} + W^{\leq(v,k(v))} + E^{i,k(v)}). && (25)
\end{aligned}
$$

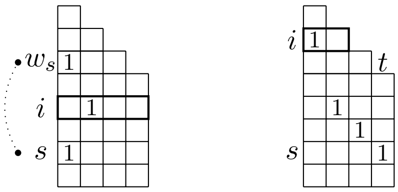

Figure 8. *Left:* Term (26) of Claim 5. *Right:* Term (28) of Claim 7.

We depict (24) in Figure 7d. Matrix (24) has lexicographically ordered columns by definition of S_3, and the same is true for (25). We use the assumption that node $\{v, i\} \notin E$ in both terms. □

Claim 4. *For all* $(s,t) = \langle \sigma, t\rangle \in \mathrm{row}_i = S_4$, $E^{st} \in \mathcal{L}$.

Proof. For $t = 1$, the claim is trivially true since E^{s1} with $s = i$ is a vertex of $\mathrm{OP}_k(G)$ that fulfills the packing inequality for row i with equality.

For $1 < t \leq k(i)$, we use $E^{\langle\sigma,t\rangle} = W^{\leq\langle\sigma,t\rangle} - W^{\leq\langle\sigma,t-1\rangle}$, where we use the above claims since $\langle\sigma, t-1\rangle$ is contained in sets S_1, S_2, or S_3. Hence, $W^{\leq\langle\sigma,t-1\rangle}$ is contained in $\mathcal{L}$. □

In the next claims, we rely on the hypothesis that for every node $u > i$, a node $w_u \in [i]$ exists with $w_u \notin \Gamma(u)$.

Claim 5. *For* $(s,t) \in S_5$, $E^{st} \in \mathcal{L}$.

Proof. By hypothesis, for every node $s > i$, there exists a node w_s with $\{w_s, s\} \notin E$. If $w_s < i$, then we use:

$$E^{s1} = (E^{s1} + E^{w_s,1} + E^{i2}) \tag{26}$$

$$- (E^{w_s,1} + E^{i2}). \tag{27}$$

We have depicted (26) on the left-hand side of Figure 8. Term (26) does not violate an edge inequality by assumption. If $w_s = i$, then the linear combination is $E^{s1} = (E^{s1} + E^{i1}) - E^{i1}$. □

Claim 6. *For* $(s,t) \in S_6$, $E^{st} \in \mathcal{L}$.

Proof. We note that for $(s,t) \in S_6$, $\mathrm{diag}^{\leq}(s,t) \cap \mathrm{row}_i \neq \varnothing$. Hence, $E^{st} = W^{\leq(s,t)} - W^{\leq(s-1,t-1)}$. Note that if $S_6 \neq \varnothing$, then $t > 1$. □

For the proof of the next claim, we need the following notation. Let

$$\mathrm{diag}^{\geq}(s,t) := \{\langle\sigma,t\rangle, \langle\sigma,t+1\rangle, \ldots, \langle\sigma,k\rangle\}$$

and

$$\mathrm{diag}^{\geq}(s,t)|_c := \{\langle\sigma,t\rangle, \langle\sigma,t+1\rangle, \ldots, \langle\sigma,c-1\rangle\}.$$

We define $W^{\geq(s,t)}$ and $W^{\geq(s,t)|_c}$ to be the incidence matrices of $\mathrm{diag}^{\geq}(s,t)$ and $\mathrm{diag}^{\geq}(s,t)|_c$, respectively.

Claim 7. *For* $(s,t) = \langle\sigma,t\rangle \in S_7$, $E^{st} \in \mathcal{L}$.

Proof. The following linear combination gives E^{st}:

$$E^{st} = (W^{\geq\langle\sigma,2\rangle|_{t+1}} + E^{i,1}) \tag{28}$$

$$- (W^{\geq\langle\sigma,2\rangle|_t} + E^{i,1}), \tag{29}$$

see the right-hand side of Figure 8 for a depiction of (28). □

In total, we proved that for all $(s,t) \in \mathcal{I}_{nk}$, we have $E^{st} \in \mathcal{L}$ which terminates the proof. □

Note that the statement of the previous theorem takes both the combinatorial structure of the underlying graph and the labeling of the nodes of the graph into account.

Our subsequent results on the facet-defining properties of inequalities have proofs that rely on similar techniques to the ones presented above. We therefore defer these proofs to the appendix.

As we have seen, packing inequalities do not always define facets of $\mathrm{OP}_k(G)$. However, under some circumstances they can be combined with inequalities for the stable set problem to yield facets of $\mathrm{OP}_k(G)$. We have the following general validity statement.

Proposition 14. *Let $\alpha^{\mathrm{T}} y \leq \beta$ be a non-trivial valid inequality for $\mathrm{P}(G)$ with integral coefficients. Let $W := \{v \in V : \alpha_v \neq 0\}$, and let $w = \max(W)$. Assume that $\Gamma(w) \supseteq \{1, \dots, u-1\}$, where $u = \max(W \setminus \{w\})$. Then,*

$$\sum_{t=1}^{j-1} x_{wt} + \sum_{v \in W} \alpha_v \, x_{vj} \leq \beta \tag{30}$$

is valid for $\mathrm{OP}_k(G)$ for each $j \leq \min(W, k(i))$.

Proof. If $j = 1$, then the inequality is clearly valid. Thus, we assume that $j > 1$ in the following.

Because of the packing inequalities, the first sum is at most 1. Thus, the inequality is valid for each integral point $\tilde{x} \in \mathrm{OP}_k(G)$ that satisfies

$$\sum_{v \in W} \alpha_v \, \tilde{x}_{vj} < \beta.$$

Now assume that

$$\sum_{v \in W} \alpha_v \, \tilde{x}_{vj} = \beta.$$

Since $\alpha^{\mathrm{T}} y \leq \beta$ is a non-trivial inequality, we have $\beta > 0$ and $\alpha \geq 0$, see Remark 2. Hence, there exists $v \in W$ with $\tilde{x}_{vj} = 1$. If $v = w$, then the packing inequality for row w yields that the first sum in (30) is 0. Thus, inequality (30) holds. Otherwise, let $\langle \eta, j \rangle = (v, j)$. By the lexicographic ordering, $\tilde{x}$ contains 1s in the diagonal $\mathrm{diag}^{\leq}(\eta, j-1)$ or above it. Since by the assumption there are edges between w and each node in $[v-1] \subseteq [u-1]$, the first sum has to be 0, proving validity. □

In the special case in which $\alpha^{\mathrm{T}} y \leq \beta$ is a clique inequality, i.e., $C = W$ is a clique, we call the corresponding Inequality (30) *packing-clique inequality.* We next provide a characterization of the cases in which packing-clique inequalities define facets. We first give the result for $j \in C$. Proofs appear in the appendix.

Theorem 15. *Let $(i,j) \in \mathcal{I}_{nk}$ with $j < i$, let $\Gamma(i) \supseteq [i-1]$, and let $C \subseteq \{j, \dots, i\}$ be a clique such that $i = \max(C)$ and $j \in C$. The packing-clique inequality*

$$\sum_{t=1}^{j-1} x_{it} + x(C \times \{j\}) \leq 1 \tag{31}$$

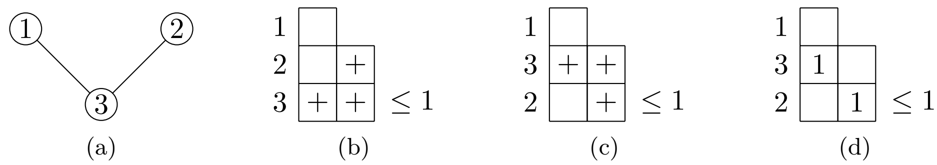

Figure 9. Example 18: Packing-clique inequalities depend on the order of the rows. The numbers on the left-hand side of the rows are node labels.

defines a facet of $\mathrm{OP}_k(G)$ *if and only if the following conditions hold:*

(1) *C is maximal in the subgraph $G[j, j+1, \ldots, i]$;*
(2) *if $j > 1$, then there exists no clique C' with $C' \supseteq \{j-1, \ldots, i-1\}$; and*
(3) *for every node $u \in \{i+1, i+2, \ldots, n\}$, there exists a node $w_u \in [i]$ such that $w_u \notin \Gamma(u)$.*

Remark 16. Theorem 15 shows that maximal clique inequalities do not always define facets for $\mathrm{OP}_k(G)$ – in contrast to $\mathrm{P}(G)$ and $\mathrm{P}_k(G)$.

The characterization in Theorem 15 is independent of k, while the next case $j \notin C$ depends on k.

Theorem 17. *Let $(i, j) \in \mathcal{I}_{nk}$ with $j < i$, let $\Gamma(i) \supseteq [i-1]$, and let $C \subseteq \{j, \ldots, i\}$ be a clique such that $i = \max(C)$ and $j \notin C$. The packing-clique inequality*

$$\sum_{t=1}^{j-1} x_{it} + x(C \times \{j\}) \leq 1 \tag{32}$$

defines a facet of $\mathrm{OP}_k(G)$ *if and only if the following conditions hold:*

(1) *C is maximal in the subgraph $G[j, j+1, \ldots, i]$;*
(2) *if $j > 1$, then there exists no clique C' with $C' \supseteq \{j-1, \ldots, i-1\}$;*
(3) *for every node $u \in \{i+1, i+2, \ldots, n\}$, there exists a node $w_u \in [i]$ such that $w_u \notin \Gamma(u)$; and*
(4) *the following must hold for $r = \min(C) - 1$ with $(r, j) = \langle \rho, j \rangle$:*
   - *For every $\langle \beta, c \rangle = (b, c)$ with $\beta = 1$, $b \in C$, and $j < c \leq k(b)$, there exists a node $v \in C \cap \{b+1, b+2, \ldots, i\}$ such that $\{j, v\} \notin E$, and*
   - *for every $\langle \beta, c \rangle = (b, c)$ with $b \in C$, $1 < \beta \leq \rho$, and $j < c \leq k(b)$, there exists a node $v \in \{j, j+1, \ldots, r\}$ with $\langle \nu, j \rangle = (v, j)$ and $\nu < \beta$ that is not adjacent to at least one node in $C \setminus \{b\}$.*

Example 18. We close the discussion of packing-clique inequalities with three explicit examples.

First, Figure 9 shows that packing-clique inequalities depend on the mapping of nodes to rows of the matrix. Consider the graph shown in Figure 9a. By Theorem 15, the packing-clique inequality for $C = \{2, 3\}$, $i = 3$, and $j = 2$ defines a facet for $\mathrm{OP}_2(G)$. We depict this inequality in Figure 9b. If we map the nodes to rows as depicted in Figure 9c, then the resulting inequality is not valid, since the matrix depicted in Figure 9d violates the inequality.

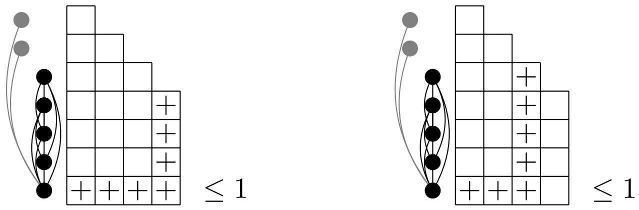

Figure 10. Example 18: Combination of packing with clique inequalities. The clique C is depicted in black nodes and the (additional) neighborhood of $\max(C)$ in gray. The inequality on the right-hand side dominates the one on the left-hand side.

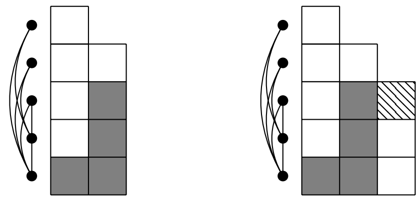

Figure 11. Example 18: depending on the number of colors, certain packing-inequalities may or may not define a facet. *Left:* For $k = 2$, the packing-clique inequality with non-zero coefficients in gray defines a facet. *Right:* For $k = 3$, we can enhance the packing-clique inequality with a summand whose index is hatched.

Second, Figure 10 shows two packing-clique inequalities. The inequality on the left-hand side is a packing-clique inequality for $C = \{4, 5, 6, 7\}$, $i = 7$, and $j = 4$. It does not define a facet, since it violates Condition (2) of Theorem 15. Indeed, the packing-clique inequality on the right-hand side for $C = \{3, 4, 5, 6, 7\}$, $i = 7$, and $j = 3$ dominates the one on the left-hand side. Thus, again in contrast to $\mathrm{P}(G)$ and $\mathrm{P}_k(G)$, maximal clique and maximal packing-clique inequalities do not always define facets of $\mathrm{OP}_k(G)$.

Finally, Figure 11 shows an example of a packing-clique inequality for $C = \{3, 4, 5\}$, $i = 5$, and $j = 2$ which, by Theorem 17, defines a facet for $k = 2$ (left). For $k = 3$ (right), Condition (4) of Theorem 17 is violated. Let us briefly show why the depicted inequality on the right is valid. Any $\tilde{x} \in \mathrm{OP}_3(G)$ with $\tilde{x}_{33} = 1$ has $\tilde{x}_{32} = 0$ due to the packing inequality. Furthermore, we must have $\tilde{x}_{22} = \tilde{x}_{11} = 1$ because of the lexicographic ordering of the columns. Since nodes 1 and 2 are adjacent to nodes 4 and 5, none of the other summands in the packing-clique inequality can be 1. Dominance is immediate.

4.2. Shifted Column Inequalities

Shifted column inequalities define facets of the orbitope in most cases. We investigate conditions under which SCIs define facets of $\mathrm{OP}_k(G)$ and start with the special case of the main diagonal.

Proposition 19. *The SCI* $x(B) - x(S) \le 0$ *with shifted column* $S := \{\langle 1, j \rangle\}$ *and bar* $B := \{\langle 1, j+1 \rangle\}$ *defines a facet of* $\mathrm{OP}_k(G)$ *for* $1 \le j < k$.

The SCIs in which $S = \{\langle 1, j \rangle\}$ *and* $B = \{\langle 1, t \rangle\}$ *with* $j + 1 < t \le k$ *are contained in the SCIs with* $S = \{\langle 1, \ell \rangle\}$ *and* $B = \{\langle 1, \ell + 1 \rangle\}$ *for* $j \le \ell \le t$.

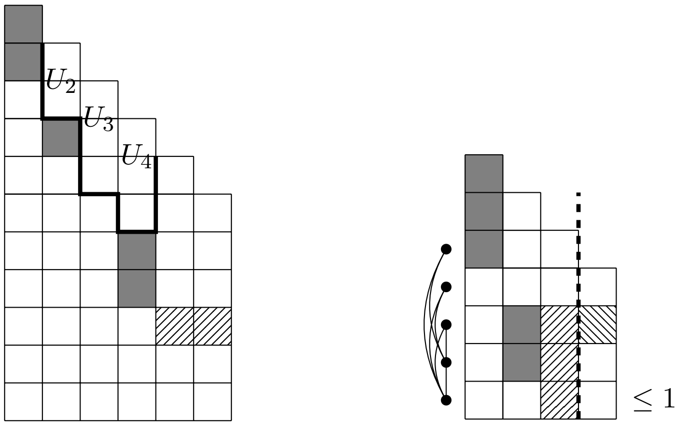

Figure 12. *Left:* example for an SCI and the definition of U_2, U_3, U_4 ($U_1 = \varnothing$). *Right:* The CSCI defines a facet for $k = 3$, but not for $k = 4$. For $k = 4$, the third condition of Theorem 24 is violated and the CSCI is dominated by an enhanced CSCI.

Proof. If $B = \{\langle 1, j+1\rangle\}$, then the matrices $W^{\leq\langle s,t\rangle}$, for all $(s,t) \in \mathcal{I}_{nk} \setminus \{(j,j)\}$, fulfill the SCI with equality. Together with $\mathbf{0}$, we have $|\mathcal{I}_{nk}|$ affinely independent matrices that fulfill the SCI with equality, showing the claim.

The SCI with $S = \{\langle 1, j\rangle\}$ and $B = \{\langle 1, t\rangle\}$ is the sum of the SCIs $x_{\langle 1,j+1\rangle} - x_{\langle 1,j\rangle} \leq 0, \dots, x_{\langle 1,t+1\rangle} - x_{\langle 1,t\rangle} \leq 0$. Thus, they do not define facets for $\mathrm{OP}_k(G)$. □

We need some more notation for general SCIs. For $j \in [k]$, consider a shifted column $S = \{\langle 1, c_1\rangle, \langle 2, c_2\rangle, \dots, \langle \eta, c_\eta\rangle\}$ with $c_\eta < j$. For column $c \in \{c_1, c_1+1, \dots, j-1\}$, we define

$$\hat{U}_c \coloneqq \{\langle \rho, c\rangle \,:\, \rho < \delta \text{ for all } \langle \delta, t\rangle \in S \text{ with } c \leq t\},$$

the set of all matrix entries that lie above the diagonals defined by S intersected with column c. The set U_c is defined as the projection on the first coordinate of $\hat{U}_c$, see the left-hand side of Figure 12.

Theorem 20. *Let $i \in V$ be a node, $j \in \{2, \dots, k(i)\}$, and $\langle \sigma, j\rangle = (i, j)$ such that $\sigma > 1$. The SCI $x(B) - x(S) \leq 0$ with bar $B = \{(i,j), \dots, (i,k(i))\}$ and shifted column $S = \{\langle 1, c_1\rangle, \langle 2, c_2\rangle, \dots, \langle \eta, c_\eta\rangle\}$, with $c_\eta < j$, $\eta \leq \sigma$, defines a facet of $\mathrm{OP}_k(G)$ if and only if the following conditions hold:*

(1) *there exists a node $v \in \{j, j+1, \dots, \min(i-1, k(i))\}$ with $v \notin \Gamma(i)$, and*
(2) *U_c is not a clique for all $c \in \{c_1, c_1+1, \dots, j-1\}$.*

The proof appears in the appendix. Note that $\varnothing$ is not a clique. For orbitopes, SCIs do not define facets if $\eta \geq 2$ and $c_1 < c_2$, see [19]. We do not have to deal with this case explicitly, since then $G[U_{c_2}]$ is a clique of size one. In the case $j = k(i)$, SCIs are dominated by the set of inequalities that we describe in the next section.

4.3. Clique Shifted Column Inequalities

Similarly to packing inequalities, SCIs may be combined with inequalities for the stable set problem to yield valid inequalities of $\mathrm{OP}_k(G)$.

Proposition 21. *Let $\alpha^{\mathrm{T}} y \leq \beta$ be a non-trivial valid inequality for* $\mathrm{P}(G)$ *with integral coefficients. Let $W \coloneqq \{v \in V \,:\, \alpha_v \neq 0\}$ and $w = \max(W)$. Let S be a shifted column of* $\mathrm{col}(w-1, j-1)$ *for some $1 < j \leq \min(W, k)$. Then*

$$\sum_{v \in W} \alpha_v \, x_{vj} - x(S) \leq \beta - 1 \tag{33}$$

is a valid inequality for $\mathrm{OP}_k(G)$.

Proof. Consider an integral point $\tilde{x} \in \mathrm{OP}_k(G)$. Thus,

$$\tilde{\alpha} \coloneqq \sum_{v \in W} \alpha_v \, \tilde{x}_{vj} \leq \beta$$

holds. If $\tilde{\alpha} \leq \beta - 1$, then Inequality (33) is obviously satisfied. Therefore, assume $\tilde{\alpha} = \beta$. Since $\alpha^{\mathrm{T}} y \leq \beta$ is a non-trivial inequality, we have $\beta > 0$ and $\alpha \geq 0$, see Remark 2. Thus, there exists a $v \in W$ with $\tilde{x}_{vj} = 1$. Let $\langle \sigma, j \rangle = (v, j)$. Define

$$\tilde{S} = S \cap \{\langle \rho, t \rangle \,:\, \rho \leq \sigma, \ t \in [j-1]\}$$

to be the part of S that is below or on the diagonal $\mathrm{diag}^{\leq}\langle \sigma, j-1 \rangle$. Then $\tilde{S}$ is a shifted column of $\mathrm{col}(v-1, j-1)$. Hence, there exists an entry $(p, q) \in \tilde{S}$ with $x_{pq} = 1$, see [19, Lemma 8]. This shows that $\tilde{x}(S) \geq 1$, proving the validity of (33). □

We again specialize to the case when $C = W$ (with $|C| \geq 2$) is a clique, i.e., $y(C) \leq 1$ is a clique inequality. Let $i = \max(C)$ and $1 < j \leq \min(C, k)$. Assume that S is a shifted column of $\mathrm{col}(i-1, j-1)$. Then the corresponding *clique shifted column inequality* (CSCI) with leader (i, j) is

$$x(C \times \{j\}) - x(S) \leq 0,$$

which is valid by Proposition 21; the right-hand side of Figure 12 contains an example. These inequalities were introduced in [19]. We define a CSCI to be *maximal* if the clique C is maximal in the subgraph $G[j, j+1, \ldots, i]$.

Remark 22. The conditions in Theorem 20 that guarantee SCIS to be facet-defining require $j < k(i)$. For $j = k(i)$, an SCI may be a CSCI if $\{j, j+1, \ldots, i\}$ is a stable set. In this case, the SCI/CSCI for $j = k(i)$ defines a facet as the following theorems show.

As for packing-clique inequalities, we differentiate between two cases for the characterization of the facet-defining properties. Proofs appear in the appendix.

Theorem 23. *Let C be a clique, $i = \max(C)$, $1 < j \leq \min(C, k)$, and $j \in C$ such that $(i, j) = \langle \sigma, j \rangle$ with $\sigma > 1$. Then a CSCI with leader (i, j) defines a facet of* $\mathrm{OP}_k(G)$ *if and only if*

(1) *the CSCI is maximal, and*
(2) *U_c is not a clique for all $c \in \{c_1, c_1 + 1, \ldots, j-1\}$.*

For $(i, j) \in \mathcal{I}_{nk}$, a clique C in $G[j, \ldots, i]$, and $j \notin C$, a maximal CSCI can only define a facet under certain conditions on the number of colors, similarly to the packing-inequalities.

Theorem 24. *Let C be a clique, let $i = \max(C)$, and let $1 < j \leq \min(C, k)$ be such that $j \notin C$ and $(i, j) = \langle \sigma, j \rangle$ with $\sigma > 1$. Then a CSCI with leader (i, j) defines a facet of $\mathrm{OP}_k(G)$ if and only if the following conditions hold:*

(1) *The CSCI is maximal.*
(2) *U_c is not a clique for all $c \in \{c_1, c_1 + 1, \dots, c_\eta\}$.*
(3) *The following must hold for $r = \min(C) - 1$ with $(r, j) = \langle \rho, j \rangle$:*
   - *For every $\langle \beta, c \rangle = (b, c)$ with $\beta = 1$, $b \in C$, and $j < c \leq k(b)$, there exists a node $v \in C \cap \{b + 1, b + 2, \dots, i\}$, such that $\{j, v\} \notin E$.*
   - *For every $\langle \beta, c \rangle = (b, c)$ with $b \in C$, $1 < \beta \leq \rho$, and $j < c \leq k(b)$, there exists a node $v \in \{j, j + 1, \dots, r\}$ with $\langle \nu, j \rangle = (v, j)$ such that $\nu < \beta$ that is not adjacent to at least one node in $C \setminus \{b\}$.*

The third condition is the same as for packing-clique inequalities (Theorem 17) and depends on k: the example on the right-hand side of Figure 12 shows a CSCI that defines a facet for $k = 3$, but not for $k = 4$.

5. Conclusion

The main topic of this paper is the polyhedral combination of the problem specific structure for the maximum k-colorable subgraph problem with the symmetry handling structure of orbitopes. One motivation is our result that the weak LP-bound of the natural IP-formulation only slightly improves by adding orbitope symmetry handling. In our subsequent polyhedral investigation, we characterized a number of facet-defining inequalities. The structure of the graph influences most of these characterizations. The correspondence of nodes to rows (row labeling) is important, too. Therefore, the choice of the row labeling is important both for the theoretical and the practical properties of a given instance. In particular, we suspect that the labeling has a big influence on the performance of branch-and-cut algorithms.

A natural row labeling heuristic is to order nodes with respect to decreasing node degree, which might be beneficial because of the following two (heuristical) reasons. First, variable fixing by orbitope symmetry handling tends to happen towards the top of the matrix. If nodes with a high degree are affected, then this increases the chance of secondary fixings, e.g., due to edge inequalities. Second, such an ordering seems to produce stronger inequalities, such as the packing-clique inequalities. Here, the base node has to be adjacent to all nodes preceding it. The higher the degree of the base node, the higher are the chances of such an occurrence.

Future computational experiments will investigate the practical side of the results presented in this paper. In particular, we shall investigate different row labelings. Moreover, graph symmetries will be taken into account.

Acknowledgements

The first author thanks Barry O'Sullivan for kindly supporting his travels in connection with the present work; he acknowledges financial support by an Embark PhD scholarship of the Irish Research Council for Science, Technology and Engineering. We thank a reviewer for his many suggestions that helped to improve the paper. We thank Marc van Dongen for further suggestions.

References

[1] A. Brandstädt, V. Le, T. Szymczak, F. Siegemund, H. de Ridder, S. Knorr, M. Rzehak, M. Mowitz, N. Ryabova, U. Nagel, and D. Baye, *Information system on graph class inclusions*. http://wwwteo.informatik.uni-rostock.de/isgci/, visited 10/2010.

[2] M. C. Carlisle and E. L. Lloyd, *On the k-coloring of intervals*, Discrete Appl. Math., 59 (1995), pp. 225–235.

[3] H.-A. Choi, K. Nakajima, and C. S. Rim, *Graph bipartization and via minimization*, SIAM J. Discrete Math., 2 (1989), pp. 38–47.

[4] V. Chvátal, *On certain polytopes associated with graphs*, J. Comb. Theory, Ser. B, 18 (1975), pp. 138–154.

[5] J. M. de Carvalho Valério, *Exact solution of bin-packing problems using column generation and branch-and-bound*, Ann. Oper. Res., (1999).

[6] D. de Werra and A. Hertz, *On perfectness of sums of graphs*, Discrete Math., 195 (1999), pp. 93–101.

[7] Y. Faenza, *Graph coloring using orbitopes*, master's thesis, Universitá degli studi di Roma Tor Vergata, 2006.

[8] Y. Faenza and V. Kaibel, *Extended formulations for packing and partitioning orbitopes*, Math. Oper. Res., 34 (2009), pp. 686–697.

[9] P. Fouilhoux and A. R. Mahjoub, *Polyhedral results for the bipartite induced subgraph problem*, Discrete Appl. Math., 154 (2006), pp. 2128–2149.

[10] E. J. Friedman, *Fundamental domains for integer programs with symmetries*, in COCOA 2007, Y. Zu, B. Zhu, and A. Dress, eds., vol. 4616 of LNCS, Springer-Verlag Berlin Heidelberg, 2007, pp. 146–153.

[11] D. Fulkerson, *On the perfect graph theorem*, in Mathematical Programming, Proc. Advanced Seminar, Univ. Wisconsin, Madison 1972, 1973, pp. 69–76.

[12] I. P. Gent, K. E. Petrie, and J.-F. Puget, *Symmetry in constraint programming*, in Handbook of Constraint Programming, F. Rossi, P. van Beek, and T. Walsh, eds., Elsevier, 2006, pp. 329–376.

[13] M. Grötschel, L. Lovász, and A. Schrijver, *Geometric Algorithms and Combinatorial Optimization*, vol. 2 of Algorithms and Combinatorics, Springer-Verlag, Heidelberg, 2nd ed., 1993.

[14] P. L. Hammer, E. L. Johnson, and U. N. Peled, *Facets of regular 0-1 polytopes*, Math. Prog., 8 (1975), pp. 179–206.

[15] T. Januschowski, *Symmetry breaking for the maximal k-colourable subgraph problem*, diploma thesis, TU Berlin, 2007.

[16] T. Januschowski and M. E. Pfetsch, *Branch-cut-and-propagate for the maximum k-colorable subgraph problem*, preprint, 2011. http://www.optimization-online.org/DB_HTML/2011/02/2909.html.

[17] V. Kaibel, *Basic polyhedral theory*, in Wiley Encyclopedia of Operations Research and Management Science, J. J. Cochran, L. A. J. Cox, P. Keskinocak, J. P. Kharoufeh, and J. C. Smith, eds., Wiley, 2010. To appear.

[18] V. Kaibel, M. Peinhardt, and M. E. Pfetsch, *Orbitopal fixing*, in Integer Programming and Combinatorial Optimization, 12th International IPCO Conference, Ithaca, NY, USA, June 25-27, 2007, Proceedings, M. Fischetti and D. P. Williamson, eds., vol. 4513 of LNCS, Springer, 2007, pp. 74–88.

[19] V. Kaibel and M. E. Pfetsch, *Packing and partitioning orbitopes*, Math. Program., 114, Number 1 / July, 2008 (2008), pp. 1–36.

[20] A. M. C. A. Koster and M. Scheffel, *A routing and network dimensioning strategy to reduce wavelength continuity conflicts in all-optical networks*, in Proceedings of INOC 2007, Spa, Belgium, 2007, International Network Optimization Conference.

[21] C. Lecoutre and S. Tabary, *Symmetry breaking*, in Constraint Networks Techniques and Algorithms, Wiley, 2009, pp. 495–530.

[22] J. M. Lewis and M. Yannakakis, *The node-deletion problem for hereditary properties is NP-complete*, J. Comput. Syst. Sci., 20 (1980), pp. 219–230.

[23] J. Linderoth, J. Ostrowski, F. Rossi, and S. Smriglio, *Orbital branching*, in Proceedings of IPCO XII, M. Fischetti and D. Williamson, eds., vol. 4513 of LNCS, Springer-Verlag, 2007, pp. 106–120.

[24] C. Lund and M. Yannakakis, *The approximation of maximum subgraph problems*, in Automata, Languages and Programming, A. Lingas, R. Karlsson, and S. Carlsson, eds., vol. 700 of LNCS, Springer-Verlag, Berlin Heidelberg, 1993, pp. 40–51.

[25] F. Margot, *Pruning by isomorphism in branch-and-cut*, Math. Program., 94 (2002), pp. 71–90.

[26] ———, *Exploiting orbits in symmetric ILP*, Math. Program., 98 (2003), pp. 3–21.

[27] ———, *Symmetric ILP: Coloring and small integers*, Discrete Opt., 4 (2007), pp. 40–62.

[28] ———, *Symmetry in integer linear programming*, in 50 Years of Integer Programming 1958–2008, M. Jünger, T. Liebling, D. Naddef, G. L. Nemhauser, W. Pulleyblank, G. Reinelt, G. Rinaldi, and L. Wolsey, eds., Springer-Verlag, Berlin Heidelberg, 2010, ch. 17, pp. 647–681.

[29] A. Mehrotra and M. A. Trick, *A column generation approach for graph coloring*, INFORMS J. Comput., 8 (1996), pp. 344–354.

[30] G. Narasimhan, *The maximum k-colorable subgraph problem*, Computer Sciences Technical Report 864, University of Wisoncsin-Madison, 1989.

[31] E. Olaru and H. Sachs, *Contribution to a characterization of the structure of perfect graphs*, Ann. Discrete Math., 21 (1984), pp. 121–144.

[32] J. Ostrowski, J. Linderoth, F. Rossi, and S. Smriglio, *Constraint orbital branching*, in Integer Programming and Combinatorial Optimization, Proc. of the 13th IPCO conference, Bertinoro, A. Lodi, A. Panconesi, and G. Rinaldi, eds., vol. 5035 of LNCS, Springer, 2008, pp. 225–239.

Appendix A. Omitted Proofs

We use the following notation, partly introduced before. Let $(s,t) = \langle \sigma, t\rangle \in \mathcal{I}_{nk}$. We define

$$\operatorname{diag}^{\leq}\langle \sigma, t\rangle|_j = \{\langle \sigma, j+1\rangle, \langle \sigma, j+2\rangle, \dots, \langle \sigma, t\rangle\},$$

and let $W^{\leq \langle \sigma, t\rangle|_j}$ be the corresponding incidence matrix. We need these
Furthermore, we define

$$\operatorname{diag}^{\geq}(s,t)|_c = \{\langle \sigma, t\rangle, \langle \sigma, t+1\rangle, \dots, \langle \sigma, c-1\rangle\},$$

and let $W^{\leq (s,t)}$, $W^{\geq (s,t)}$, and $W^{\geq (s,t)|_c}$ be the incidence matrices corresponding to $\operatorname{diag}^{\leq}(s,t)$, $\operatorname{diag}^{\geq}(s,t)$, and $\operatorname{diag}^{\geq}(s,t)|_c$, respectively.

A.1. Packing-Clique Inequalities

Theorem 15 (repeated). Let $(i,j) \in \mathcal{I}_{nk}$ with $j < i$, let $\Gamma(i) \supseteq [i-1]$, and let $C \subseteq \{j, \dots, i\}$ be a clique such that $i = \max(C)$ and $j \in C$. The packing-clique inequality

$$\sum_{t=1}^{j-1} x_{it} + x(C \times \{j\}) \leq 1 \tag{34}$$

defines a facet of $\mathrm{OP}_k(G)$ if and only if the following conditions hold:

(1) C is maximal in the subgraph $G[j, j+1, \dots, i]$;
(2) if $j > 1$, then there exists no clique C' with $C' \supseteq \{j-1, \dots, i-1\}$; and
(3) for every node $u \in \{i+1, i+2, \dots, n\}$, there exists a node $w_u \in [i]$ such that $w_u \notin \Gamma(u)$.

Proof. "$\Rightarrow$" If the clique C is not maximal, then there exists a clique $C' \subseteq \{j, \dots, i\}$ with $C \subset C'$. Then the packing-clique inequality for C is clearly dominated by the one for C'.

Next, assume for the sake of contradiction that $j > 1$ and there exists a clique C' such that $C' \supseteq \{j-1, \dots, i-1\}$. In this case, we have the following inequality:

$$\sum_{t=1}^{j-2} x_{it} + x(C' \times \{j-1\}) \leq 1. \tag{35}$$

Inequality (35) is valid because it is a packing-clique inequality. It dominates the original packing-clique inequality (34) due to the following arguments. Assume that $\tilde{x}$ has a 1 in $C \times \{j\}$. Due to the lexicographic ordering, $\tilde{x}$ has a 1 in $C' \times \{j-1\}$. Thus, $\tilde{x}$ is contained in the face defined by inequality (35). On the other hand, matrix $W^{\leq (i-1,j-1)}$ fulfills inequality (35) with equality, but not (34).

Finally, a slight extension of the arguments presented in Theorem 13 suffices to show that the last condition is necessary.

"$\Leftarrow$" Let $\mathcal{V}$ be the set of vertices of $\mathrm{OP}_k(G)$ that satisfy the packing-clique inequality (34) with equality. Let $\mathcal{L} := \operatorname{lin}(\mathcal{V})$ be the linear span of $\mathcal{V}$. We shall show that the unit matrix E^{st} with $(s,t) \in \mathcal{I}_{nk}$ is contained in $\mathcal{L}$. So, we can conclude that $\mathcal{L}$ contains $|\mathcal{I}_{nk}|$ linearly independent points. Since $\mathbf{0} \notin \mathcal{V}$, it follows that $\dim(\operatorname{aff}(\mathcal{V})) = |\mathcal{I}_{nk}| - 1$. Hence, the packing-clique inequality defines a facet of $\mathrm{OP}_k(G)$.

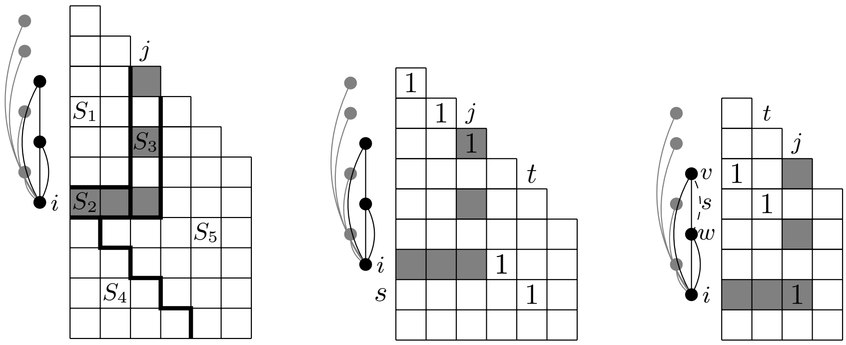

Figure 13. Illustrations for the proof of Theorem 15 on packing-clique inequalities. *Left:* Sets used in the proof. *Middle:* Matrix as constructed in Claim 8. *Right:* Matrix as constructed in Claim 9.

In order to show that $E^{st} \in \mathcal{L}$ for all $(s,t) \in \mathcal{I}_{nk}$, we divide the set $\mathcal{I}_{nk}$ into five parts:

$$\begin{aligned} S_1 &:= \{(s,t) \in \mathcal{I}_{nk} \,:\, s \le i,\ t < j\},\\ S_2 &:= \{(s,t) \in \mathcal{I}_{nk} \,:\, s = i,\ t \le j-1\},\\ S_3 &:= \{(s,t) \in \mathcal{I}_{nk} \,:\, j \le s \le i,\ t = j\},\\ S_4 &:= \{\langle \sigma, t\rangle \,:\, \sigma > i\},\\ S_5 &:= \{(s,t) = \langle \sigma, t\rangle \,:\, \sigma \le i,\ s > i \text{ or } t > j\}. \end{aligned}$$

The left-hand side of Figure 13 contains a depiction of the sets.

We divide the proof into five different parts, one for each set. We note that we can use the arguments provided in Theorem 13 to prove that for $(s,t) \in S_4$, we have $E^{st} \in \mathcal{L}$, see Claim 7.

Claim 8. *For all $(s,t) \in S_5$, $E^{st} \in \mathcal{L}$.*

Proof. If $\operatorname{diag}^{\le}(s,t) \cap S_2 \neq \varnothing$ or $\operatorname{diag}^{\le}(s,t) \cap (C \times \{j\}) \neq \varnothing$, then we have

$$E^{st} = W^{\le(s,t)} - W^{\le(s-1,t-1)},$$

and both matrices are feasible and fulfill the packing-clique inequality with equality. In the other cases, we use

$$E^{st} = (W^{\le(s,t)|_j} + W^{\le(j,j)}) - (W^{\le(s-1,t-1)|_j} + W^{\le(j,j)}).$$

The middle of Figure 13 contains a depiction of the first term. Here, we use the fact that $j \in C$. □

Claim 9. *For all $(s,t) \in S_1$, $E^{st} \in \mathcal{L}$.*

Proof. Due to the second condition in the theorem, for (s,t) there exists a pair of nodes $v < w$ in $\{j-1,\dots,i-1\}$ such that $\{v,w\} \notin E$. For $s = v$ we use

$$E^{st} = (E^{wt} + E^{i,t+1} + W^{\le(s,t)}) - (E^{wt} + E^{i,t+1} + W^{\le(s-1,t-1)});$$

if $s = w$, then we replace w by v. For $s \neq v$ and $s \neq w$, we use

$$E^{st} = (W^{\le(\min(s,v),t-1)} + E^{st} + E^{i,t+1}) - (W^{\le(\min(s,v),t-1)} + E^{vt} + E^{i,t+1}).$$

The right-hand side of Figure 13 contains a depiction of the first term. □

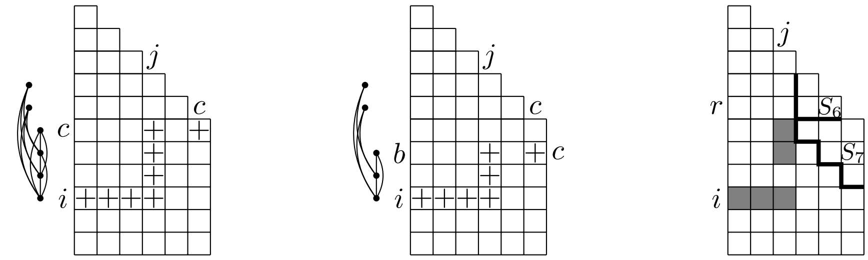

Figure 14. Proof of Theorem 17. *Left* and *Middle:* Inequalities that dominate a packing-clique inequality. *Right:* The sets used to prove Theorem 17. The non-zero coefficients of the packing-clique inequality are depicted in gray.

Claim 10. *For all $(s,t) \in S_3$, $E^{st} \in \mathcal{L}$.*

Proof. First, consider $s \notin C$. Due to the maximality of C, there exists a node $v \in C$ such that $\{s,v\} \notin E$. We then consider the following combination of feasible points:

$$E^{st} = E^{st} + E^{vt} + W^{\leq(\min(s,v)-1,t-1)} - (E^{vt} + W^{\leq(\min(s,v)-1,t-1)}).$$

Next, consider $s \in C$. We then use $E^{st} = W^{\leq(s,t)} - W^{\leq(s-1,t-1)}$, and we note that $W^{\leq(s-1,t-1)} \in \mathcal{L}$ due to Claim 9. □

Claim 11. *For all $(s,t) \in S_2$, $E^{st} \in \mathcal{L}$.*

Proof. As before, we have $E^{st} = W^{\leq(s,t)} - W^{\leq(s-1,t-1)}$, and we note that $W^{\leq(s-1,t-1)} \in \mathcal{L}$ due to Claim 9. □

Thus, we proved that for all $(s,t) \in \mathcal{I}_{nk}$, $E^{st} \in \mathcal{L}$, which terminates our proof. □

Theorem 17 (repeated). Let $(i,j) \in \mathcal{I}_{nk}$ with $j < i$, let $\Gamma(i) \supseteq [i-1]$, and let $C \subseteq \{j,\ldots,i\}$ be a clique such that $i = \max(C)$ and $j \notin C$. The packing-clique inequality

$$\sum_{t=1}^{j-1} x_{it} + x(C \times \{j\}) \leq 1 \tag{36}$$

defines a facet of $\mathrm{OP}_k(G)$ if and only if the following conditions hold:

(1) C is maximal in the subgraph $G[j, j+1, \ldots, i]$.
(2) If $j > 1$, then there exists no clique C' with $C' \supseteq \{j-1,\ldots,i-1\}$.
(3) For every node $u \in \{i+1, i+2, \ldots, n\}$, there exists a node $w_u \in [i]$ such that $w_u \notin \Gamma(u)$.
(4) The following must hold for $r = \min(C) - 1$ with $(r,j) = \langle \rho, j \rangle$:
   - For every $\langle \beta, c \rangle = (b,c)$ with $\beta = 1$, $b \in C$, and $j < c \leq k(b)$, there exists a node $v \in C \cap \{b+1, b+2, \ldots, i\}$ such that $\{j,v\} \notin E$.
   - For every $\langle \beta, c \rangle = (b,c)$ with $b \in C$, $1 < \beta \leq \rho$, and $j < c \leq k(b)$, there exists a node $v \in \{j, j+1, \ldots, r\}$ with $\langle \nu, j \rangle = (v,j)$ such that $\nu < \beta$ that is not adjacent to at least one node in $C \setminus \{b\}$.

Proof. "⇒" For the first three conditions, we use the arguments of Theorem 15.

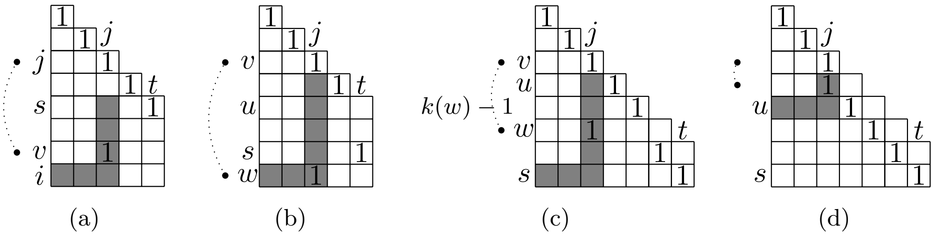

Figure 15. Matrices for Claim 13 in the proof of Theorem 17. The non-zero coefficients of the packing-clique inequality are depicted in gray. 15a: Term (39). 15b: Matrix A^{st} for $s < w$. 15c: Matrix A^{st} for $s > w$. 15d: first term in (41).

Assume for the sake of contradiction that there exists $c > j$ with $c \in C$ such that all nodes in $C \cap \{c+1, c+2, \dots, i\}$ are adjacent to j. We shall argue that the following inequality is valid:

$$\sum_{t=1}^{j-1} x_{it} + x(C \times \{j\}) + x_{cc} \leq 1, \tag{37}$$

see the left-hand side of Figure 14. Consider a point $\tilde{x} \in \mathrm{OP}_k(G)$ with $\tilde{x}_{cc} = 1$. Due to the lexicographic ordering, we have

$$\tilde{x}_{11} = \cdots = \tilde{x}_{jj} = \tilde{x}_{j+1,j+1} = \cdots = \tilde{x}_{cc} = 1.$$

Because of the packing inequalities, $\tilde{x}$ does not have 1-entries in positions $(C \cap \{j, j+1, \dots, c\}) \times \{j\}$. By assumption, we have $\{j, r\} \in E$ for all $r \in C \cap \{c+1, c+2, \dots, i\}$. Hence, $\tilde{x}$ cannot have a 1-entry in $C \times \{j\}$. Moreover, since $\Gamma(i) \supseteq [i-1]$, $\tilde{x}_{i1} = \cdots = \tilde{x}_{i,j-1} = 0$. Thus (37) is valid and its dominance over (36) is immediate.

Next, assume for the sake of contradiction that there exists $\langle \beta, c \rangle = (b, c)$ with $b \in C$, $1 < \beta \leq \rho$, and $c > j$ such that all nodes $v \in \{j, j+1, \dots, r\}$ with $\langle \nu, j \rangle = (v, j)$ and $\nu < \beta$ are adjacent to all nodes in $C \setminus \{b\}$. We claim that

$$\sum_{t=1}^{j-1} x_{it} + x(C \times \{j\}) + x_{bc} \leq 1 \tag{38}$$

is valid for $\mathrm{OP}_k(G)$. Assume that there is a point $\tilde{x} \in \mathrm{OP}_k(G)$ with $\tilde{x}_{bc} = 1$. Because of the lexicographic ordering, $\tilde{x}$ has a one in $(\{j, \dots, b-1\} \times \{j\})$. By assumption, each node $v \in \{j, \dots, r\}$ is adjacent to all nodes in $C \setminus \{b\}$. Moreover, packing inequalities together with $\tilde{x}_{bc} = 1$ imply $\tilde{x}_{bj} = 0$. Thus $\tilde{x}(C \times \{j\}) = 0$. Furthermore, since $\Gamma(i) \supseteq [i-1]$, $\tilde{x}_{i1} = \cdots = \tilde{x}_{i,j-1} = 0$. Hence, (38) is valid and clearly dominates (36). The middle of Figure 14 contains an illustration.

"$\Leftarrow$" We proceed as usual. Let $\mathcal{V}$ be the set of vertices of $\mathrm{OP}_k(G)$ that satisfy the packing-clique inequality with equality. Define $\mathcal{L} \coloneqq \mathrm{lin}(\mathcal{V})$ as the linear span of the elements in $\mathcal{V}$. We shall show that the unit matrix E^{st} with $(s, t) \in \mathcal{I}_{nk}$ is contained in $\mathcal{L}$. So, we can conclude that $\mathcal{L}$ contains $|\mathcal{I}_{nk}|$ linearly independent points and since $\mathbf{0} \notin \mathcal{V}$, we have $\dim(\mathrm{aff}(\mathcal{V})) = |\mathcal{I}_{nk}| - 1$. Hence, the packing-clique inequality defines a facet of $\mathrm{OP}_k(G)$.

In order to show that $E^{st} \in \mathcal{L}$ for all $(s,t) \in \mathcal{I}_{nk}$, we divide the set $\mathcal{I}_{nk}$ into three parts:

$$\begin{aligned} S_6 :=& \{\langle \sigma, t\rangle \,:\, t > j,\ \sigma > \rho\} \\ S_7 :=& \{(s,t) \,:\, s \leq \min(C),\ t \geq j\} \\ \overline{S} :=& \mathcal{I}_{nk} \setminus (S_6 \cup S_7). \end{aligned}$$

The right-hand side of Figure 14 contains a depiction of the sets S_6 and S_7.

We note that for $(s,t) \in \overline{S}$, we use the arguments of Theorem 15. Thus, we only have to consider $(s,t) \in S_6 \cup S_7$.

Claim 12. *For $(s,t) \in S_6$, $E^{st} \in \mathcal{L}$.*

Proof. Let node $u \in \{j, j+1, \dots, r\}$ be such that $\{(u,j)\} = \text{diag}^{\leq}(s,t) \cap \text{col}_j$. Since C is maximal, there exists a node $v \in C$ such that $\{v,u\} \notin E$. By definition of S_6, we have $v > s$. We then use

$$E^{st} = (W^{\leq(s,t)} + E^{vj}) - (W^{\leq(s-1,t-1)} + E^{vj}).$$

Both terms in parenthesis are feasible and fulfill (36) with equality. Hence, we have proved $E^{st} \in \mathcal{L}$. □

Claim 13. *For $(s,t) = \langle \sigma, t\rangle \in S_7$, $E^{st} \in \mathcal{L}$.*

Proof. If $\sigma = 1$, then we note the following. By the first part of the fourth assumption, there exists a node $v \in C$ with $v > s$, such that $\{v,j\} \notin E$. Hence, the matrix

$$W^{\leq\langle\sigma,t\rangle} + E^{vj} \tag{39}$$

is feasible and fulfills (36) with equality; see Figure 15a. The same holds for $W^{\leq\langle\sigma,t-1\rangle} + E^{vj}$. Thus, we have proved that $E^{st} \in \mathcal{L}$ for $\sigma = 1$.

If $\sigma > 1$ and $s \in C$, then we let $\langle \sigma, j\rangle = (u,j)$. By the second part of the fourth assumption, there exist nodes $v \in \{j, j+1, \dots, \min\{r, u-1\}\}$ and $w \in C \setminus \{s\}$ such that $\{v,w\} \notin E$. Let $\langle \nu, j\rangle = (v,j)$. For the case $w > s$, we use

$$A^{st} := E^{st} + E^{wj} + W^{\leq\langle\nu,t-1\rangle},$$

which is feasible and fulfills (36) with equality (see Figure 15b). The same holds for $E^{wj} + W^{\leq\langle\nu,t-1\rangle}$. We conclude that $E^{st} \in \mathcal{L}$. For the case $w < s$, we use

$$A^{st} := E^{st} + E^{wj} + W^{\leq\langle\nu,k(w)-1\rangle} + W^{\geq(w+1,k(w))|_t}, \tag{40}$$

see Figure 15c. The matrix A^{st} is feasible and fulfills (36) with equality. The same holds for

$$E^{wj} + W^{\leq\langle\nu,k(w)\rangle} + W^{\geq(w+1,k(w))|_t}.$$

Hence, we have $E^{st} \in \mathcal{L}$.

Finally, if $\sigma > 1$ and $s \notin C$, then let $u \in C$ be such that $u := \max(\{v \in C \,:\, \text{diag}^{\leq}(s,t) \cap \text{row}_v \neq \varnothing\})$, and let d be such that $(u,d) = \text{diag}^{\leq}(s,t) \cap \text{row}_u$. With the matrix A^{ud} from above, we get

$$E^{st} = (A^{ud} + W^{\geq(u+1,d+1)|_{t+1}}) - (A^{ud} + W^{\geq(u+1,d)|_t}). \tag{41}$$

Figure 15d contains an illustration of the first term in parenthesis. Both terms are feasible and fulfill (36) with equality. Thus, $E^{st} \in \mathcal{L}$. □

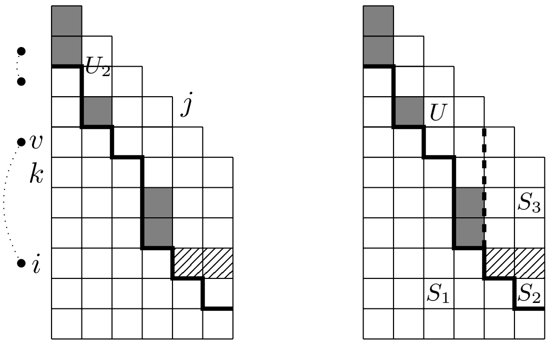

Figure 16. An SCI with a shifted column in gray and a bar as shaded. *Left:* A sample SCI. *Right:* The sets for the proof of Theorem 20.

Concluding, we have proved that, for all $(s,t) \in \mathcal{I}_{nk}$, we have $E^{st} \in \mathcal{L}$, which terminates our proof. □

A.2. SCIs

We recall some notation for SCIs. For $j \in [k]$, consider a shifted column $S = \{\langle 1, c_1\rangle, \langle 2, c_2\rangle, \dots, \langle \eta, c_\eta\rangle\}$, with $c_\eta < j$. The left-hand side of Figure 16 contains a depiction of an SCI. For some column $c \in \{c_1, c_1+1, \dots, j-1\}$, we define

$$\hat{U}_c \coloneqq \{\langle \rho, c\rangle \,:\, \rho < \delta \text{ for all } \langle \delta, t\rangle \in S \text{ with } c \leq t\},$$

the set of all matrix entries that lie above the diagonals defined by S intersected with column c. The set U_c is defined as the projection on the first coordinate of $\hat{U}_c$, see Figure 12.

In the following, we assume $\eta > 1$, since we have treated $\eta = 1$ already in Proposition 19.

Theorem 20 (repeated). Let $i \in V$ be a node, $j \in \{2, \dots, k(i)\}$, and $\langle \sigma, j\rangle = (i,j)$ such that $\sigma > 1$. The SCI $x(B) - x(S) \leq 0$ with bar $B = \{(i,j), \dots, (i,k(i))\}$ and shifted column $S = \{\langle 1, c_1\rangle, \dots, \langle \eta, c_\eta\rangle\}$, with $c_\eta < j$, $\eta \leq \sigma$, defines a facet of $\mathrm{OP}_k(G)$ if and only if the following conditions hold:

(1) there exists a node $v \in \{j, j+1, \dots, \min(i-1, k(i))\}$ with $v \notin \Gamma(i)$, and
(2) U_c is not a clique for all $c \in \{c_1, c_1+1, \dots, j-1\}$.

Proof. "⇒" Assume first that node i is adjacent to all nodes in the set $\{j, j+1, \dots, \min(i-1, k(i))\}$. If $i > k$, then

$$x(B) - x(S) + x_{kk} \leq 0 \tag{42}$$

dominates the SCI. The left-hand side of Figure 17 contains an example of this dominating inequality. We show that Inequality (42) is valid. If $\tilde{x}_{kk} = 1$ for $\tilde{x} \in \mathrm{OP}_k(G)$, then, by the lexicographic ordering $\tilde{x}_{tt} = 1$ for all $t \in [k]$. By the assumption, there are edges $\{i, v\} \in E$ for all $j \leq v \leq k$. Thus $\tilde{x}(B) = 0$. Because $\tilde{x}(S) \geq 0$, Inequality (42) is valid. Domination is immediate.

If $i \leq k$, then we claim that

$$x(B \setminus \{(i,i)\}) + x_{i-1,i-1} - x(S) \leq 0 \tag{43}$$

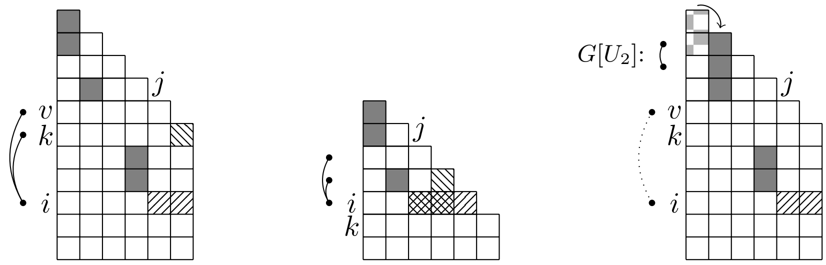

Figure 17. Cases in which SCI do not define facets. *Left:* Example for Inequality (42). *Middle:* Example for Inequality (43). The original bar of the SCI is in north east lines, the stronger SCI has a bar with north west lines. *Right:* The depicted SCI dominates the SCI in the left-hand side of Figure 16 due to the graph induced by U_2 being a clique.

dominates the SCI, see the middle of Figure 17 for a depiction of the inequality. Inequality (43) is valid for the following reason. If $\tilde{x}_{i-1,i-1} = 1$ for some $\tilde{x} \in \mathrm{OP}_k(G)$, then $\tilde{x}_{11} = \cdots = \tilde{x}_{i-2,i-2} = 1$, due to the lexicographic order of the columns. Due to the assumption that node i is adjacent to all nodes in the set $\{j, j+1, \ldots, \min(i-1, k(i))\}$, we have $\tilde{x}(B \setminus \{(i,i)\}) = 0$. Hence, at most one of the first two summands in Inequality (43) can be 1. Dominance can be seen as follows. Whenever $\tilde{x}_{ii} = 1$, due to the lexicographic ordering, we have $\tilde{x}_{i-1,i-1} = 1$ as well. Furthermore, $W^{\leq(i-1,i-1)}$ fulfills Inequality (43) with equality but not the original packing-clique inequality.

Assume next that $G[U_c]$ is a clique for some $c \in \{c_1, \ldots, j-1\}$. Consider the SCI

$$x(B) - x(\overline{S}) \leq 0 \tag{44}$$

where

$$\overline{S} := \hat{U}_c \cup (S \cap \mathrm{col}_{c+1}) \cup \cdots \cup (S \cap \mathrm{col}_{j-1}),$$

i.e., all parts of S left-hand side of column c are moved to column c. The right-hand side of Figure 17 contains an example. Inequality (44) is an SCI and thus valid. We prove that inequality (44) is at least as strong as the original SCI. Let $\tilde{x}$ be a feasible solution that fulfills the original SCI with equality. If $\tilde{x}$ has a 1 in columns before c, then, by the lexicographic ordering, it also has a 1 in column c. Because U_c is a clique, $\tilde{x}$ has at most one 1 in U_c. Thus $\tilde{x}$ fulfills (44) with equality. On the other hand, there may be a solution that fulfills (44) with equality, but that has two ones in S before column c.

"$\Leftarrow$" We use the same proof strategy as before. The set $\mathcal{V}$ denotes the vertices fulfilling the SCI with equality. We denote by $\mathcal{L} = \mathrm{lin}(\mathcal{V} \cup \{E^{ij}\})$ and show $\mathcal{L} = \mathbb{R}^{\mathcal{I}_{nk}}$. Since $\mathbf{0} \in \mathcal{V}$, it follows that $\dim(\mathrm{aff}(\mathcal{V})) = |\mathcal{I}_{nk}| - 1$, which proves that the SCI defines a facet. In order to prove that for all $(s,t) \in \mathcal{I}_{nk}$, $E^{st} \in \mathcal{L}$, we partition $\mathcal{I}_{nk}$ into the already defined sets S, B, and $U := U_{c_1} \cup \cdots \cup U_{j-1}$, as well as:

$$\begin{aligned}
S_1 :=& \{\langle \rho, t\rangle \in \mathcal{I}_{nk} \,:\, (\rho \leq \eta \text{ and } t < c_\rho) \text{ or } \rho > \eta\} \\
S_2 :=& \{\langle \rho, t\rangle = (s,t)) \in \mathcal{I}_{nk} \,:\, \rho \leq \eta \text{ and } s > i\} \\
S_3 :=& \{\langle \rho, t\rangle = (s,t) \in \mathcal{I}_{nk} \,:\, \rho < \eta,\ t > j,\ \textit{and } s < i\}.
\end{aligned}$$

The right-hand side of Figure 16 contains an illustration of the sets.

Claim 14. *For* $(s,t) \in S_1$*,* $E^{st} \in \mathcal{L}$*.*

Proof. By definition, we have $\text{diag}^{\leq}(s,t) \cap S = \varnothing$ and $\text{diag}^{\leq}(s,t) \cap B = \varnothing$ for $(s,t) \in S_1$. Hence, for $s > 1$, $t > 1$, $E^{st} = W^{\leq(s,t)} - W^{\leq(s-1,t-1)}$. Both terms are contained in $\mathcal{V}$. For $s = 1$ or $t = 1$, E^{st} already fulfills the SCI with equality. □

Claim 15. *For* $(s,t) \in S_2$*,* $E^{st} \in \mathcal{L}$*.*

Proof. By definition, $|\text{diag}^{\leq}(s,t) \cap S| = 1$ and $|\text{diag}^{\leq}(s,t) \cap B| = 1$, for $(s,t) \in S_2$. Moreover, $s > 1$ and $t > 1$. Hence, $E^{st} = W^{\leq(s,t)} - W^{\leq(s-1,t-1)}$, where both terms are contained in $\mathcal{V}$. □

Claim 16. *For* $(s,t) \in S_3$*,* $E^{st} \in \mathcal{L}$*.*

Proof. By hypothesis, there exists a node $v \in \{j, j+1, \dots, \min(i-1,k)\}$ such that $\{v,i\} \notin E$. For $s = v$, we have

$$E^{vt} = (W^{\leq(v,t)} + E^{it}) - (W^{\leq(v-1,t-1)} + E^{it}). \tag{45}$$

Figure 18a shows an example for the left term in parentheses. Both terms fulfill the SCI with equality. Hence, $E^{vt} \in \mathcal{L}$.

Consider $v < s$ and $t > j$. In the following, we assume $\{s,i\} \in E$, because the arguments below also prove $\{s,i\} \notin E$. We then use:

$$E^{st} = (W^{\leq(v,t-1)} + E^{st} + E^{i,t-1}) - (W^{\leq(v,t-1)} + E^{i,t-1}). \tag{46}$$

The left term in parentheses is depicted in Figure 18b. Since $v < s < i$ and $\{v,i\} \notin E$, both terms of (46) are feasible.

For $t = j$, we use

$$E^{st} = (E^{st} + W^{\leq(v-1,t-1)} + E^{i,t+1}) - (W^{\leq(v-1,t-1)} + E^{i,t+1} + E^{vt}). \tag{47}$$

Figure 18c contains an illustration of the left term in parentheses. Here, we use that $j < k(i)$. Furthermore, we have already proved $E^{vt} \in \mathcal{L}$.

Consider $v > s$, $\{s,i\} \in E$. We note that if $t = \min(i-1,k)$, then $v > s$ by the hypothesis that for i there exists a node in $\{j, j+1, \dots, \min(i-1,k)\}$ which is not adjacent to node i. Thus, we can assume $t < \min(i-1,k)$. We use

$$\begin{aligned} E^{st} = (W^{\leq(s,t)} + E^{v,t+1} + E^{i,t+1}) - (W^{\leq(s-1,t-1)} + E^{vt} + E^{i,t+1}) \\ + E^{vt} - E^{v,t+1} \end{aligned} \tag{48}$$

where we already proved that $E^{vt}, E^{v,t+1} \in \mathcal{L}$. Figure 18d shows an example for the left term. □

Claim 17. *For* $(s,t) \in U$*,* $E^{st} \in \mathcal{L}$*.*

Proof. Since, by assumption, $G[U_t]$ is not a clique, there exists a pair of nodes $v < w \in \hat{U}_t$ with $\{v,w\} \notin E$. We use

$$E^{wt} = (W^{\leq(v,t)} + W^{\geq(w,t)|_j} + E^{ij}) - (W^{\leq(v,t)} + W^{\geq(w+1,t+1)|_j} + E^{ij}). \tag{49}$$

Figure 19a contains an example of the left term in parentheses. Note that by definition of $\mathcal{L}$, $E^{ij} \in \mathcal{L}$. Hence, $E^{wt} \in \mathcal{L}$ as well.

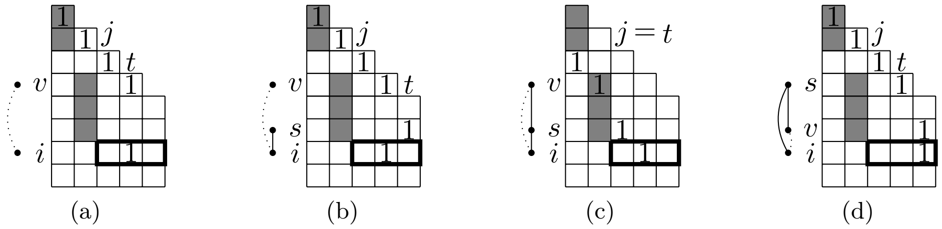

Figure 18. Matrices needed to prove Claim 16. 18a: Left term of (45). 18b: Left term of (46) 18c: Left term of (47). 18d: Left term of (48).

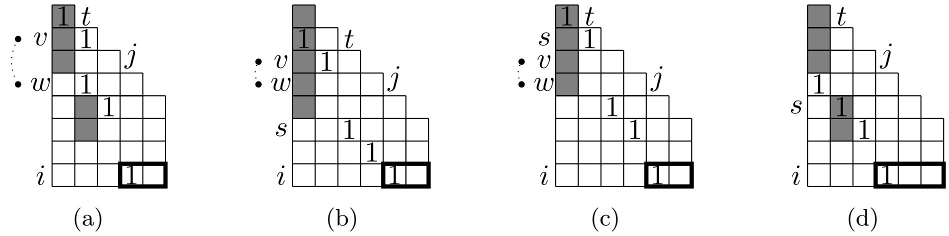

Figure 19. Matrices for the proof of Claim 17. 19a: Left term of (49). 19b: term in (50). 19c: Left term of (51). 19d: Matrix in Claim 18.

Consider $s \in U_t$ and $s > w$. We use

$$E^{st} = (W^{\leq(w-1,t-1)} + W^{\geq(s,t)|_j} + E^{ij}) \\ - (W^{\leq(w,t)} + W^{\geq(s+1,t+1)|_j} + E^{ij}). \tag{50}$$

Figure 19b gives an illustration of (50). We note that for $(s+1, t+1) = (i, j)$, we have $\text{diag}^{\geq}(s+1, t+1)|_j = \varnothing$. Since both E^{wt} and E^{ij} are in $\mathcal{L}$, we conclude that $E^{st} \in \mathcal{L}$.

Consider now $s < w$. In this case, we use

$$E^{st} = (W^{\leq(s,t)} + W^{\geq(w+1,t+1)|_j} + E^{ij}) \\ - (W^{\leq(s-1,t-1)} + W^{\geq(w,t)|_j} + E^{ij}). \tag{51}$$

We show an example of (51) in Figure 19c. Again, we note that E^{wt}, $E^{ij} \in \mathcal{L}$. □

Claim 18. *For $(s, t) = \langle \rho, t \rangle \in S$, $E^{st} \in \mathcal{L}$.*

Proof. The matrix $W^{\leq\langle \rho, j-1 \rangle} + E^{ij}$ is feasible and fulfills the SCI with equality; see Figure 19d. Each element of $\text{diag}^{\leq}\langle \rho, j-1 \rangle$ is either contained in S_1 or in U. By the above claims and $E^{ij} \in \mathcal{L}$, it follows that $E^{st} \in \mathcal{L}$. □

Claim 19. *For $(s, t) \in B$, $E^{st} \in \mathcal{L}$.*

Proof. The matrix $W^{\leq(s,t)}$ fulfills the SCI with equality. Since we have already proved $W^{\leq(s-1,t-1)} \in \mathcal{L}$, we have $E^{st} \in \mathcal{L}$. □

Since we have treated all cases, this concludes the proof. □

A.3. CSCIs as Facets

Theorem 23 (repeated). Let C be a clique, $i = \max(C)$, $1 < j \leq \min(C, k)$, and $j \in C$ such that $(i, j) = \langle \sigma, j \rangle$ with $\sigma > 1$. Then a CSCI with leader (i, j) defines a facet of $\mathrm{OP}_k(G)$ if and only if

(1) the CSCI is maximal, and
(2) U_c is not a clique for all $c \in \{c_1, c_1 + 1, \ldots, j - 1\}$.

Proof. "$\Rightarrow$" If the CSCI is not maximal, then the corresponding maximal CSCI dominates the CSCI. If for some c, $G[U_c]$ is a clique, then the CSCI is dominated by another CSCI analogously to the SCI case.

"$\Leftarrow$" We use the same proof strategy and the same definition for $\mathcal{V}$ and $\mathcal{L}$ as in Theorem 20. Analogously to SCIs, we partition $\mathcal{I}_{nk}$ into the sets U, $(C \times \{j\})$, and S, as well as

$$\begin{aligned} S_1 &:= \{\langle \rho, t \rangle \in \mathcal{I}_{nk} \,:\, (\rho \leq \eta \text{ and } t < c_\rho) \text{ or } \rho > \eta\} \\ S_2 &:= \mathrm{col}(i, j) \setminus (C \times \{j\}) \\ S_3 &:= \{\langle \rho, t \rangle = (s, t) \in \mathcal{I}_{nk} \,:\, \rho \leq \eta,\ t > j\}. \end{aligned}$$

The proof for $(s, t) \in S_1$ is straight-forward, and the proof for $(s, t) \in U$ is the same as for SCIs.

Claim 20. *For $(s, t) \in (C \times \{j\})$, $E^{st} \in \mathcal{L}$.*

Proof. For $s \in C$, the matrices $W^{\leq (s,j)}$ and $W^{\leq (s-1,j-1)} + E^{ij}$ are feasible and fulfill the CSCI with equality. Since $E^{ij} \in \mathcal{L}$ by definition, we have $E^{st} \in \mathcal{L}$. □

Claim 21. *For $(s, t) \in S_2$, $E^{st} \in \mathcal{L}$.*

Proof. Since C is a maximal clique, there exists a node $u \in C$, such that $\{u, s\} \notin E$. Let $r = \min(u, s)$. The matrix $W^{\leq (r-1,j-1)} + E^{uj} + E^{sj}$ is feasible and fulfills the CSCI with equality. Moreover, the same holds for $W^{\leq (r-1,j-1)} + E^{uj}$. Thus, we conclude that $E^{st} \in \mathcal{L}$. □

Claim 22. *For $(s, t) = \langle \rho, t \rangle \in S_3$, $E^{st} \in \mathcal{L}$.*

Proof. Since by assumption $j \in C$, the matrix

$$W^{\geq (\langle \sigma, j+1 \rangle)|_{t+1}} + W^{\leq (j,j)}$$

is feasible and fulfills the CSCI with equality. The same holds for

$$W^{\geq (\langle \sigma, j+1 \rangle)|_{t}} + W^{\leq (j,j)}.$$

We conclude that $E^{st} \in \mathcal{L}$. □

Hence, CSCIs with $j \in C$ define facets. □

Theorem 24 (repeated). Let C be a clique, let $i = \max(C)$, and let $1 < j \leq \min(C, k)$ be such that $j \notin C$ and $(i, j) = \langle \sigma, j \rangle$ with $\sigma > 1$. Then a CSCI with leader (i, j) defines a facet of $\mathrm{OP}_k(G)$ if and only if the following conditions hold:

(1) The CSCI is maximal.
(2) U_c is not a clique for all $c \in \{c_1, c_1 + 1, \ldots, c_\eta\}$.
(3) The following must hold for $r = \min(C) - 1$ with $(r, j) = \langle \rho, j \rangle$:

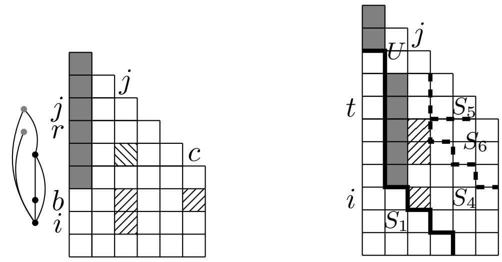

Figure 20. *Left:* A dominating CSCI. All gray nodes are adjacent to nodes in $C \setminus \{b\}$. *Right:* The sets used to prove Theorem 24.

- For every $\langle \beta, c\rangle = (b, c)$ with $\beta = 1$, $b \in C$, and $j < c \leq k(b)$, there exists a node $v \in C \cap \{b+1, b+2, \dots, i\}$, such that $\{j, v\} \notin E$.
- For every $\langle \beta, c\rangle = (b, c)$ with $b \in C$, $1 < \beta \leq \rho$, and $j < c \leq k(b)$, there exists a node $v \in \{j, j+1, \dots, r\}$ with $\langle \nu, j\rangle = (v, j)$ such that $\nu < \beta$ that is not adjacent to at least one node in $C \setminus \{b\}$.

Proof. "⇒" We only have to show that the third condition is necessary, since we can use the arguments in the proofs of Theorem 20 and 23 for the other two conditions.

Assume for the sake of contradiction that there exists $\langle \beta, c\rangle = (b, c)$ with $b \in C$, $1 < \beta \leq \rho$, and $c > j$ such that all nodes $v \in \{j, j+1, \dots, r\}$ with $\langle \nu, j\rangle = (v, j)$ and $\nu < \beta$ are adjacent to all nodes in $C \setminus \{b\}$. With the same arguments as for packing-clique inequalities, it follows that

$$x(C \times \{j\}) + x_{bc} \leq 1$$

is a valid inequality, and Proposition 21 yields that

$$x(C \times \{j\}) + x_{bc} - x(S) \leq 0 \tag{52}$$

is a valid inequality as well. Domination is immediate. An illustration of this inequality is shown on the left-hand side of Figure 20.

For the other case, assume that there exists $\langle \beta, c\rangle = (b, c)$ with $\beta = 1$, $b \in C$, and $c > j$ such that all nodes in $C \cap \{b+1, b+2, \dots, r\}$ are adjacent to j. With an argumentation similar as before, Inequality (52) then dominates the CSCI.

"⇐" Following the proof strategy and notation of Theorem 23, we note that we only have to prove that $E^{st} \in \mathcal{L}$ for (s, t) contained in the following sets:

$$\begin{aligned} S_4 &:= \{\langle \delta, t\rangle \in \mathcal{I}_{nk} \,:\, \delta \leq \eta,\ t > j\} \\ S_5 &:= \{(s, t) \in \mathcal{I}_{nk} \,:\, s \leq \min(C),\ t \geq j\} \\ S_6 &:= \{\langle \delta, t\rangle = (s, t) \in \mathcal{I}_{nk} \,:\, \delta \leq \rho,\ t > j,\ s \geq \min(C)\}. \end{aligned}$$

The sets are depicted in Figure 20. We note that we can use the arguments of set S_3 in the proof of Theorem 23 for set S_4. Furthermore, Claims 12 and 13 prove that for $(s, t) \in S_4 \cup S_5$, we have $E^{st} \in \mathcal{L}$. □

Cork Constraint Computation Centre and Computer Science Department, University College Cork, Cork, Ireland
E-mail address: janus@cs.ucc.ie

Institute for Mathematical Optimization, TU Braunschweig, Pockelsstr. 14, 38106 Braunschweig, Germany
E-mail address: m.pfetsch@tu-bs.de