On the Way to Perfection: Primal Operations for Stable Sets in Graphs

C. Gentile [1,3], *U.-U. Haus* [2,3,4], *M. Köppe* [2,3,4], *G. Rinaldi* [1] *and R. Weismantel* [2,3,4,5]

Manfred Padberg is the scientific father, or the scientific grandfather, or the scientific great grandfather of each of the five authors. This paper is dedicated to him on the occasion of his 60th birthday.

In this paper some operations are described that transform every graph into a perfect graph by replacing nodes with sets of new nodes. The transformation is done in such a way that every stable set in the perfect graph corresponds to a stable set in the original graph. These operations yield a purely combinatorial augmentation procedure for finding a maximum weighted stable set in a graph. Starting with a stable set in a given graph one defines a simplex type tableau whose associated basic feasible solution is the incidence vector of the stable set. In an iterative fashion, non-basic columns that would lead to pivoting into non-integral basic feasible solutions, are replaced by new columns that one can read off from special graph structures such as odd holes, odd antiholes, and various generalizations. Eventually, either a pivot leading to an integral basic feasible solution is performed, or the optimality of the current solution is proved.

---

[1] Istituto di Analisi dei Sistemi ed Informatica – CNR, Roma, Italy.
E-mail: {gentile,rinaldi}@iasi.rm.cnr.it

[2] Otto-von-Guericke-Universität Magdeburg, Department for Mathematics/IMO, Germany.
E-mail: {haus,mkoeppe,weismant}@imo.math.uni-magdeburg.de

[3] Supported by the European DONET program TMR ERB FMRX-CT98-0202.

[4] Supported by grants FKZ 0037KD0099 and FKZ 2495A/0028G of the Kultusministerium of Sachsen-Anhalt.

[5] Supported by a Gerhard-Hess-Preis and grant WE 1462 of the Deutsche Forschungsgemeinschaft.

1 Introduction

The *stable set problem* (or *node packing problem*) is one of the most studied problems in combinatorial optimization. It can be defined as follows: Let $(\mathscr{G}, \mathbf{c})$ be a weighted graph, where $\mathscr{G} = (V, E)$ is a graph with $n = |V|$ nodes and $m = |E|$ edges and $\mathbf{c} \in \mathbf{R}_+^V$ is a node function that assigns a weight to each node of $\mathscr{G}$. A set $S \subseteq V$ is called *stable* if its nodes are pairwise nonadjacent in $\mathscr{G}$. The problem is to find a stable set S^* in $\mathscr{G}$ of maximum weight $\mathbf{c}(S^*) = \sum_{v \in S^*} c_v$. The value $\mathbf{c}(S^*)$ is called the $\mathbf{c}$-*weighted stability number* $\alpha_{\mathbf{c}}(\mathscr{G})$ of the graph $\mathscr{G}$.

This problem is equivalent to maximizing the linear function $\sum_{v \in S} c_v x_v$ over the *stable set polytope* $P_{\mathscr{G}}$, the convex hull of the incidence vectors of all the stable sets of $\mathscr{G}$. Thus linear programming techniques can be used to solve the problem, provided that an explicit description of the polytope is given. It is nowadays well known that, the stable set problem being NP-hard, it is very unlikely that such a description can be found for instances of arbitrary size. Moreover, even if a partial description is at hand, due to the enormous number of inequalities, it is not obvious how to turn this knowledge into a useful algorithmic tool.

Despite these difficulties, the literature in combinatorial optimization of the last thirty years abounds with successful studies where nontrivial instances of NP-hard problems were solved with a cutting plane procedure based on the generation of strong cuts obtained from inequalities that define facets of certain polytopes.

The idea of using facet defining inequalities in a cutting plane algorithm was proposed by Padberg in [Pad73] and pursued in many other papers of his. His contribution goes much beyond the advances in the knowledge of the stable set problem and its polytope, as it influenced the developments of the following three decades in polyhedral combinatorics and in computational combinatorial optimization.

The basic integer linear programming formulation of the problem is obtained by adding the integrality requirement on the variables to the following system:

$$\begin{aligned} x_u + x_v &\leq 1 \quad \text{for each edge } (u, v) \text{ in } \mathscr{G} \\ x_v &\geq 0 \quad\;\; \text{for each node } v \in V. \end{aligned} \tag{1}$$

Such a system is called the *edge-node formulation* and provides a relaxation of $P_{\mathscr{G}}$ that has been studied in depth in [Pad73] where it is proved that its solutions have values in the set $\{0, 1/2, 1\}$.

A set $Q \subseteq V$ is called a *clique* if its nodes are pairwise adjacent in $\mathscr{G}$. In [Pad73] it is proved that for every clique Q of $\mathscr{G}$ the *clique inequality*

$$\sum_{v \in Q} x_v \leq 1$$

defines a facet of $P_{\mathscr{G}}$ as long as Q is maximal with respect to set inclusion. If in (1) instead of one inequality per edge we have a clique inequality per maximal clique, we obtain the *clique formulation*:

$$\begin{aligned} \textstyle\sum_{v \in Q} x_v &\leq 1 \quad \text{for each clique } Q \text{ in } \mathscr{G} \\ x_v &\in \{0, 1\} \quad \text{for each node } v \in V, \end{aligned}$$

which provides a tighter relaxation of $P_{\mathscr{G}}$.

Let $C \subseteq V$ be a set of nodes such that $\mathscr{G}[C]$, the subgraph of $\mathscr{G}$ induced by C, is a cycle of odd length. If the cycle is chordless, it is called an *odd hole*, and the inequality

$$\sum_{v \in C} x_v \leq \frac{|C| - 1}{2}$$

is called an *odd-hole inequality*. This inequality was proved in [Pad73] to define a facet of $P_{\mathscr{G}[C]}$. In the same paper a *sequential lifting procedure* is described that turns an odd-hole inequality, and actually any inequality defining a facet of the polytope associated to a subgraph of $\mathscr{G}$, into an inequality facet-defining for $P_{\mathscr{G}}$.

After the work of Padberg, several other results were produced on the structure of the stable set polytope. Among the facet defining inequalities that were characterized we mention the *antihole inequalities* introduced by [NT73]; their definition is as for the hole inequalities, except that the subgraph induced by C is not a chordless cycle but its complement (a so-called *odd antihole*).

For a list of references to further facet defining inequalities for which a characterization is known, we refer to, e.g., Borndörfer [Bor98].

It is not a trivial task to exploit this vast amount of knowledge on the stable set polytope to devise an effective cutting plane algorithm that is able to solve non trivial instances of large size. Among the few attempts, we mention the ones of Nemhauser and Sigismondi [NS92] and of Balas et al. [BCCP96]. Unlike the case of other NP-hard problems, polyhedrally based cutting plane algorithms for the stable set problem have not yet shown their superiority over alternative methods. On the other hand, several approaches have been tried to solve difficult instances. For a collection of papers on algorithms for the stable set problem and for a recent survey on the subject, see [JT96] and [BBPP99], respectively.

The cutting plane procedure mentioned before has a "dual flavor," in the sense that the current solution is infeasible until the end, when feasibility and hence optimality is reached. A primal cutting plane procedure was first proposed by Young [You68]: One starts with an integral basic feasible solution, then either pivots leading to integral solutions are performed or cuts are generated that are satisfied by the current solution at equality. Padberg and Hong [PH80] were the first to propose a similar primal procedure based on strong polyhedral cutting planes. These kinds of algorithms produce a path of adjacent vertices of the polytope associated with the problem.

A profound study of the vertex adjacency for the polytope of the set partitioning problem was produced by Balas and Padberg [BP75]. They provided the theoretical background for the realization of a primal algorithm that produces a sequence of adjacent vertices of the polytope, ending with the optimal solution. Their basic technique was to replace a column of the current simplex tableau with a set of new columns in order to guarantee the next pivot to lead to an integral basic feasible solution.

These ideas were generalized to the case of general integer programming by Haus, Köppe, and Weismantel [HKW01a], who called their method the "Integral Basis Method." This method does neither require cutting planes nor enumeration techniques. In each major step the algorithm either returns an augmenting direction that is applicable at the given feasible point and yields a new feasible point with better objective function value or provides a proof that the point under consideration is optimal. This is achieved by iteratively substituting one column by columns that correspond to irreducible solutions of a system of linear diophantine inequalities. A detailed description of the method is given in the papers [HKW01b, HKW01c].

This paper provides some first graph theoretical tools for a primal algorithm for the stable set problem in the same vein as the work of Balas and Padberg and of Haus, Köppe, and Weismantel.

The cardinality of the largest stable set of a graph $\mathscr{G} = (V, E)$ is called the *stability number* of $\mathscr{G}$ and denoted by $\alpha(G)$. The minimum number of cliques of $\mathscr{G}$ whose union coincides with V is called the *clique covering number* of $\mathscr{G}$ and denoted by $\chi(\mathscr{G})$. A graph $\mathscr{G}$ is *perfect* if and only if $\alpha(\mathscr{G}') = \chi(\mathscr{G}')$ for all subgraphs $\mathscr{G}'$ of $\mathscr{G}$ induced by subsets of its node set V. For the fundamentals on perfect graphs and balanced matrices and on their connections, which will be used throughout the paper, we refer to, e.g., [Cor01].

A graph is perfect if and only if its clique formulation defines an integral polytope. Moreover, for perfect graphs the stability number can be computed in polynomial time [GLS88]; thus, also the separation problem for $P_{\mathscr{G}}$ is polynomially solvable in this case. Therefore, one can devise a primal cutting plane algorithm for the stable set problem for perfect graphs. We start, for example, with the edge formulation and with a basic feasible solution corresponding to a stable set. Then we perform simplex pivots until either we reach optimality or we produce a fractional solution. In the latter case we add clique inequalities to the formulation that make the fractional solution infeasible, we step back to the previous (integral) basic feasible solution, and we iterate.

Suppose now that the graph is not perfect. We assume that at hand is a graph transformation that in a finite number of "steps to perfection" transforms the original graph into a possibly larger graph that is perfect. Then it may be possible to apply again the previous primal cutting plane procedure as follows: As soon as the fractional solution cannot be cut off by clique inequalities, because other valid inequalities for $P_{\mathscr{G}}$ would be necessary, we make one or more "steps to perfection" until the clique formulation of the current graph makes the fractional solution infeasible. This procedure eventually finds an optimal stable set in the latest generated graph. It can be used for solving the original problem as long as the graph transformation is such that the optimal stable set in this graph can be mapped into an optimal stable set in the original graph.

This procedure provides a motivation for this paper where in section 2 we define valid transformations that have the desired properties mentioned above; in section 3 we translate the graph transformations into algebraic operations on the simplex tableaux; finally, in section 4, we give some properties of the proposed transformations that may be useful when the primal algorithm sketched above is implemented.

2 Valid graph transformations

Throughout this section, we will denote by $\mathscr{G}^0 = (V^0, E^0)$ and $\mathbf{c}^0$ the graph and node-weight function of the original weighted stable set problem, respectively. The purpose of this section is to devise several types of transformations $(\mathscr{G}, \mathbf{c}) \mapsto (\mathscr{G}', \mathbf{c}')$ with the property $\alpha_{\mathbf{c}}(\mathscr{G}) = \alpha_{\mathbf{c}'}(\mathscr{G}')$, i.e., transformations maintaining the weighted stability number. After a sequence of those transformations, a perfect graph $\mathscr{G}^*$ with a node-weight function $\mathbf{c}^*$ will be produced. In perfect graphs the stability number can be computed in polynomial time [GLS88]. Moreover, the $\mathbf{c}^*$-weighted stable set problem in $\mathscr{G}^*$ can be solved with linear programming over the maximal-clique formulation of $\mathscr{G}^*$. (Finding the maximal cliques in $\mathscr{G}^*$ will be expensive as well. In section 4, however, we will show that our transformations can only enlarge the size of the cliques and never destroy the property of maximality.)

Typically, one is not only interested in the weighted stability number of a graph but also in a stable set where the maximum is attained. Thus, once the $\mathbf{c}^*$-weighted stable set problem in $\mathscr{G}^*$ is solved, one would like to recover a corresponding maximum $\mathbf{c}^0$-weighted stable set in the original graph $\mathscr{G}^0$. For this purpose, we shall attach a *node labeling* $\sigma\colon V \to 2^{V^0}$ to each graph $\mathscr{G} = (V, E)$. This labeling assigns a stable set $\sigma(v) \subseteq V^0$ in the original graph to each node $v \in V$. The label of a node also determines its weight by the setting

$$c(v) = \sum_{u \in \sigma(v)} c^0(u) \quad \text{for } v \in V. \tag{2}$$

Thus, each node represents a partial stable set configuration in the original graph. For brevity of notation, we shall define $\sigma^0\colon V^0 \to 2^{V^0}$ by $\sigma^0(v) = \{v\}$ for $v \in V^0$.

Now, given a stable set $S \subseteq V$ in $\mathscr{G}$ with labeling σ, we intend to reconstruct a stable set $S^0 \subseteq V^0$ in $\mathscr{G}^0$ by

$$S^0 = \bigcup_{s \in S} \sigma(s). \tag{3}$$

For this to work, we need to impose some properties on a labeling.

D 1 (valid labeling). *Let $\mathscr{G} = (V, E)$ be a graph. A mapping $\sigma\colon V \to 2^{V^0}$ is called a* valid node labeling *of $\mathscr{G}$ (with respect to $\mathscr{G}^0$) if the following conditions hold:*

(a) *For $v \in V$, $\sigma(v)$ is a nonempty stable set in $\mathscr{G}^0$.*

(b) *For every two distinct nodes $u, v \in V$ with $\sigma(u) \cap \sigma(v) \neq \emptyset$, the edge (u, v) is in E; i.e., nodes with non-disjoint labels cannot be in the same stable set.*

(c) *Let $u, v \in V$ be distinct nodes. If there exists an edge $(u^0, v^0) \in E^0$ with $u^0 \in \sigma(u)$ and $v^0 \in \sigma(v)$, then the edge (u, v) belongs to E.*

Note that for a valid labeling σ the union in equation (3) is disjoint, and it gives a stable set in $\mathscr{G}^0$.

L 2. *Let σ be a valid labeling of a graph $\mathscr{G} = (V, E)$ and let $\mathbf{c}\colon V \to \mathbf{R}_+$ be defined by* (2)*. Let S be a stable set in $\mathscr{G}$. Then $S^0 = \bigcup_{s \in S} \sigma(s)$ is a stable set in $\mathscr{G}^0$ with $\mathbf{c}^0(S^0) = \mathbf{c}(S)$.*

Proof. Assume that $u^0, v^0 \in S^0$ are distinct nodes with $(u^0, v^0) \in E^0$. There exist $u, v \in S$ such that $u^0 \in \sigma(u)$ and $v^0 \in \sigma(v)$. If $u \neq v$, condition (c) of Definition 1 implies that $(u, v) \in E$, thus S is not stable in $\mathscr{G}$, contrary to the assumption. Otherwise, if $u = v$, the set $\sigma(u)$ is not stable in $\mathscr{G}^0$, contradicting condition (a) of Definition 1. Hence, S^0 is a stable set in $\mathscr{G}^0$.

Finally, note that the union $\bigcup_{s \in S} \sigma(s)$ is disjoint due to condition (b) of Definition 1. Therefore, $\mathbf{c}(S) = \sum_{s \in S} c(s) = \sum_{s \in S} \sum_{s^0 \in \sigma(s)} c^0(s^0) = \mathbf{c}^0(S^0)$. $\square$

D 3 (faithful labeling). *Let σ be a valid labeling of a graph $\mathscr{G} = (V, E)$ with respect to $\mathscr{G}^0$ and let $\mathbf{c}\colon V \to \mathbf{R}_+$ be defined by* (2). *We call σ a* faithful labeling *of $\mathscr{G}$ if for every stable set S in $\mathscr{G}$ that has maximum weight with respect to $\mathbf{c}$, the stable set $S^0 = \bigcup_{s \in S} \sigma(s)$ in $\mathscr{G}^0$ has maximum weight with respect to $\mathbf{c}^0$. A* faithfully labeled graph *$(\mathscr{G}, \mathbf{c}, \sigma)$ is a weighted graph $(\mathscr{G}, \mathbf{c})$ with a faithful labeling σ.*

D 4 (valid transformation). *A* valid graph transformation *is a transformation that turns a faithfully labeled graph $(\mathscr{G}, \mathbf{c}, \sigma)$ into a faithfully labeled graph $(\mathscr{G}', \mathbf{c}', \sigma')$.*

Throughout this paper, we shall only make use of a simple type of valid graph transformation, which can be characterized with the following lemma:

L 5. *Let $\mathscr{G} = (V, E)$ be a graph with a faithful labeling $\sigma\colon V \to 2^{V^0}$ with respect to $\mathscr{G}^0$. Let $\mathscr{G}' = (V', E')$ be a graph with a valid labeling $\tau\colon V' \to 2^V$ with respect to $\mathscr{G}$, such that for every stable set S in $\mathscr{G}$, there is a stable set S' in $\mathscr{G}'$ with $S = \bigcup_{s' \in S'} \tau(s')$. Then $\sigma'\colon V' \to 2^{V^0}$, defined by $\sigma'(v') = \bigcup_{v \in \tau(v')} \sigma(v)$ for $v' \in V'$, is a faithful labeling of $\mathscr{G}'$ with respect to $\mathscr{G}^0$, and $(\mathscr{G}, \mathbf{c}, \sigma) \mapsto (\mathscr{G}', \mathbf{c}', \sigma')$ is a valid graph transformation.*

Proof. Obviously, σ' is a valid labeling of $\mathscr{G}'$ with respect to $\mathscr{G}^0$. Now let S' be a stable set in $\mathscr{G}'$ that has maximum weight with respect to $\mathbf{c}'$. Let $S = \bigcup_{s' \in S'} \tau(s')$. Since τ is a valid labeling of $\mathscr{G}'$ with respect to $\mathscr{G}$, the set S is stable in $\mathscr{G}$, and we have

$$\mathbf{c}'(S') = \sum_{s' \in S'} \sum_{s^0 \in \sigma'(s')} c^0(s^0) = \sum_{s' \in S'} \sum_{s \in \tau(s')} \sum_{s^0 \in \sigma(s)} c^0(s^0) = \sum_{s' \in S'} \sum_{s \in \tau(s')} c(s) = \mathbf{c}(S). \tag{4}$$

Suppose that there is a stable set $\tilde{S}$ with $\mathbf{c}(\tilde{S}) > \mathbf{c}(S)$. Then there exists a stable set $\tilde{S}'$ in $\mathscr{G}'$ with $\tilde{S} = \bigcup_{s' \in \tilde{S}'} \tau(s')$. Since (4) also holds when S' and S are replaced by $\tilde{S}'$ and $\tilde{S}$, respectively, we have $\mathbf{c}'(\tilde{S}') > \mathbf{c}'(S')$, which is a contradiction to the assumption. Hence, σ' is a faithful labeling of $\mathscr{G}'$ with respect to $\mathscr{G}^0$. □

We first consider a very simple transformation. Take an odd path of nodes in $\mathscr{G}$,

$$P = (v_1, v_2, \ldots, v_{2l}, v_{2l+1}),$$

that together with the edge (v_{2l+1}, v_1) forms an odd hole (see Figure 1). Let S be a stable set in $\mathscr{G}$ with $v_1 \in S$. Since there are at most l elements of S in P, there exists an index i such that both v_{2i} and v_{2i+1} are not in S. Therefore, if we replace v_1 by l pairwise adjacent copies $w_1, \ldots, w_l$, where w_i is adjacent to both v_{2i} and v_{2i+1}, for $i = 1, \ldots, l$, it is not difficult to see that any stable set in $\mathscr{G}$ corresponds to a stable set in the new graph. The advantage of applying such an operation is, as will be made clear in the following, that in the new graph the odd hole has disappeared. This observation motivates the following definition.

The set of all nodes in $\mathscr{G}$ adjacent to a node v is denoted by $N_{\mathscr{G}}(v)$.

D 6 (node-path substitutions). *Let $\mathscr{G} = (V, E)$ be a graph with a valid node labeling $\sigma\colon V \to 2^{V_0}$. For some $l > 0$, let*

$$P = (v_1, v_2, \ldots, v_{2l}, v_{2l+1})$$

be a sequence of nodes of V *such that* v_i *is adjacent to* v_{i+1} *for* $i = 1, \dots, 2l$*. We call* P *an* odd path of nodes*; see Figure 1. A* node-path substitution *along* P *that transforms a graph* $\mathscr{G}$ *with a valid labeling into a graph* $\mathscr{G}'$ *with a labeling* σ' *is obtained in the following way:*

- *replace* v_1 *by the clique* W *of new nodes, defined by*

$$W = \begin{cases} \{w_1, \dots, w_l\} & \text{if } v_1 \text{ and } v_{2l+1} \text{ are adjacent in } \mathscr{G}, \\ \{w_1, \dots, w_l, t\} & \text{otherwise;} \end{cases}$$

- *for* $w \in W$ *connect* w *to all nodes of* $N_{\mathscr{G}}(v_1)$*;*
- *for* $i \in \{1, \dots, l\}$ *connect* w_i *to both* v_{2i} *and* v_{2i+1}*, then set* $\sigma'(w_i) = \sigma(v_1)$*;*
- *connect node* t *(if it is present in* W*) to* v_{2l+1} *and all nodes of* $N_{\mathscr{G}}(v_{2l+1})$*, then set* $\sigma'(t) = \sigma(v_1) \cup \sigma(v_{2l+1})$*.*

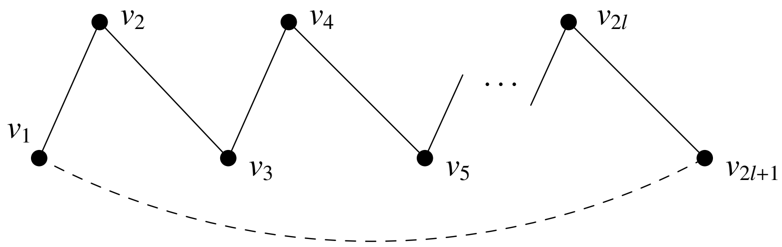

Figure 1. *An odd path of nodes*

The following definition gives a generalization of the node-path substitution.

D 7 (clique-path substitutions). *Let* $\mathscr{G} = (V, E)$ *be a graph with a valid node labeling* $\sigma \colon V \to 2^{V_0}$*. For some* $l > 0$*, let*

$$P = (Q_1 = \{v_1\}, Q_2, \dots, Q_{2l+1})$$

be a sequence of cliques of $\mathscr{G}$ *such that* $Q_{i,i+1} := Q_i \cup Q_{i+1}$ *is a clique in* $\mathscr{G}$ *for all* $i \in \{1, \dots, 2l\}$*. We call* P *an* odd path of cliques*. Let*

$$R = \{\, v \in Q_{2l+1} : v \text{ is not adjacent to } v_1 \text{ in } \mathscr{G} \,\}. \tag{5}$$

A clique-path substitution *along* P *that transforms a graph* $\mathscr{G}$ *with a valid labeling into a graph* $\mathscr{G}'$ *with a labeling* σ' *is obtained in the following way:*

- *replace* v_1 *by the clique of new nodes*

$$W = \{w_1, w_2, \dots, w_l\} \cup \{\, t_r : r \in R \,\};$$

- *for* $w \in W$ *connect* w *to all nodes of* $N_{\mathscr{G}}(v_1)$*;*

- *for $i \in \{1, \ldots, l\}$ connect w_i to all the nodes of $Q_{2i,2i+1}$, then set $\sigma'(w_i) = \sigma(v_1)$;*
- *for $r \in R$ connect t_r to r and all the nodes of $N_{\mathscr{G}}(r)$, then set $\sigma'(t_r) = \sigma(v_1) \cup \sigma(r)$.*

In Figure 2 an odd path of cliques is shown. In Figure 3, the graph $\mathscr{G}'$ that is obtained by the clique-path substitution for the case $R = \emptyset$ is shown. Note that, to unclutter the picture, some edges have been omitted, in fact all nodes w_i are connected with Q_2 and Q_{2l+1}.

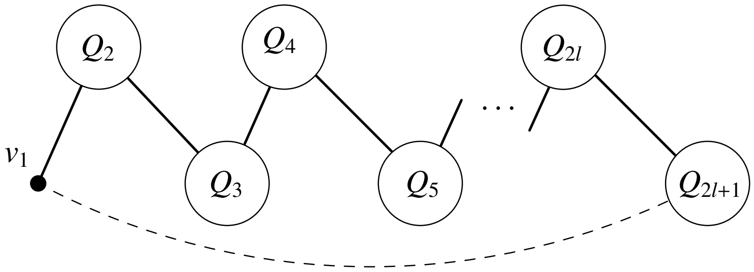

Figure 2. *An odd path of cliques*

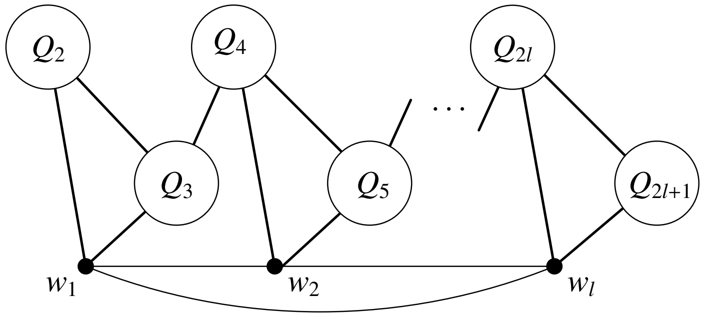

Figure 3. *Clique-path substitution with $R = \emptyset$*

Note that Definition 7 does not require the cliques Q_i to be pairwise disjoint, thus they may share nodes. Obviously, when $|Q_i| = 1$ for $i = 2, \ldots, 2l + 1$ Definition 7 reproduces Definition 6. The labeling σ' obtained in a clique-path substitution is a valid labeling; indeed, it turns faithful labelings into faithful labelings:

P **8.** *Clique-path substitutions are valid graph transformations.*

Proof. Let σ be a faithful labeling of $\mathscr{G}$. We shall make use of Lemma 5 to show that σ' is a faithful labeling of $\mathscr{G}'$.

To this end, let $\tau\colon V' \to 2^V$ be defined by $\tau(w_i) = \{v_1\}$ for $i \in \{1, \ldots, l\}$, $\tau(t_r) = \{v_1, r\}$ for $r \in R$ and $\tau(v) = \{v\}$ otherwise. Since v_1 is not adjacent to r for $r \in R$, we have that τ is a valid labeling of $\mathscr{G}'$ with respect to $\mathscr{G}$. Moreover, it is easy to see that $\sigma'(v') = \bigcup_{v\in\tau(v')} \sigma(v)$ for $v' \in V'$.

Now let S be a stable set in $\mathscr{G}$. If $v_1 \notin S$, the set S is stable in $\mathscr{G}'$ as well. Suppose that $v_1 \in S$. Since v_1 is connected to all nodes of the clique Q_2, we have that $Q_2 \cap S = \emptyset$. If also $Q_3 \cap S = \emptyset$, the set $S' = S \setminus \{v_1\} \cup \{w_1\}$ is stable in $\mathscr{G}'$. Otherwise, Q_4 is empty, and we can repeat the argument until we reach the end of the path P. If finally $Q_{2l+1} \cap S = \emptyset$, the set $S' = S \setminus \{v_1\} \cup \{w_l\}$ is stable in $\mathscr{G}'$. Otherwise, since the nodes in $Q_{2l+1} \setminus R$ are adjacent to v_1 in $\mathscr{G}$, there is a node $r \in R \cap S$. Thus, $S' = S \setminus \{v_1, r\} \cup \{t_r\}$ is stable in $\mathscr{G}'$.

By this construction, for every stable set S in $\mathscr{G}$ we obtain a stable set S' in $\mathscr{G}'$ such that $S = \bigcup_{s'\in S'} \tau(s')$. Hence, by Lemma 5, σ' is a faithful labeling of $\mathscr{G}'$. □

In certain cases, a clique-path substitution converts an imperfect graph $\mathscr{G}$ into a perfect graph $\mathscr{G}'$.

L 9. *If $\mathscr{G}$ is either an odd hole or an odd antihole, then there exists a node-path substitution such that the resulting graph $\mathscr{G}'$ is perfect.*

Proof. Let $\mathscr{G}$ be an odd hole with $2k + 1$ nodes denoted by $v_1, \ldots, v_{2k+1}$. Pick node v_1 and consider the path P from v_1 to v_{2k+1} through all nodes. Apply the node-path substitution of node v_1 along P, and call the resulting graph $\mathscr{G}'$. In Figure 4 both the original graph $\mathscr{G}$ and the transformed graph $\mathscr{G}'$ are shown for $k = 2$.

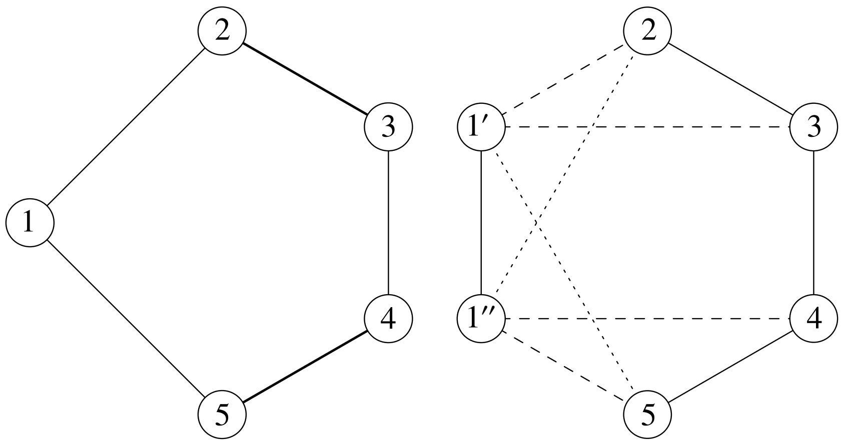

Figure 4. *Substitution of node* 1 *in an odd hole on* 5 *nodes*

Let $W = \{w_1, w_2, \ldots, w_k\}$ be the set of nodes replacing node v_1. Consider the following clique formulation associated with $\mathscr{G}'$:

$$\begin{aligned}
Q_1 &: & \textstyle\sum_{i=1}^{k} y_i + x_2 &\le 1 \\
Q_2 &: & \textstyle\sum_{i=1}^{k} y_i + x_{2k+1} &\le 1 \\
Q_{2i+1} &: & y_i + x_{2i} + x_{2i+1} &\le 1 \quad \text{for } i = 1, \ldots, k \\
Q_{2i} &: & x_{2i-1} + x_{2i} &\le 1 \quad \text{for } i = 2, \ldots, k
\end{aligned}$$

where variables y_i correspond to nodes w_i and variables x_i to nodes v_i. We show that the constraint matrix $\mathbf{M}$ of such a formulation is balanced. This implies the perfectness of $\mathscr{G}$ (see, e.g., [Cor01]).

We consider the row-column bipartite graph of the matrix $\mathbf{M}$ (see Figure 5), i.e., a graph constructed by taking a node for each row and each column of $\mathbf{M}$, and an edge for each nonzero entry of $\mathbf{M}$ connecting the nodes corresponding to its row and column. It is well known that $\mathbf{M}$ is balanced if and only if the length of all holes of this graph is divisible by 4. It is easy to verify that this is the case for $\mathbf{M}$. Hence the matrix is balanced and the claim follows.

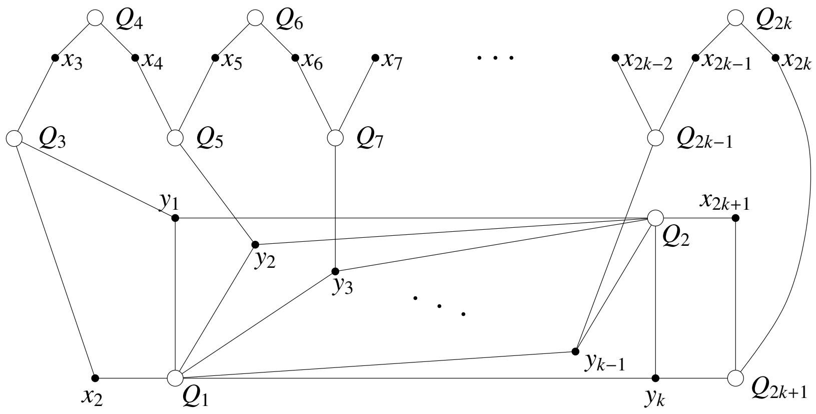

Figure 5. *Bipartite graph associated with the clique matrix of graph* $\mathscr{G}'$

Let $\mathscr{G}$ be an odd antihole with $2k + 1$ nodes, then we number the nodes in such a fashion that the edge set of the complement $\overline{\mathscr{G}}$ of $\mathscr{G}$ is the odd hole

$$\{(1, 2), (2, 3), \ldots, (2k, 2k + 1), (2k + 1, 1)\}.$$

Consider the cycle of length five in $\mathscr{G}$,

$$\{(1, 3), (3, 2k + 1), (2k + 1, 2), (2, 4), (1, 4)\}.$$

We select node 1 and perform the node-path substitution along $(1, 3, 2k + 1, 2, 4)$. That is, we replace node 1 by two adjacent nodes $1'$ and $1''$. Node $1'$ will be connected to all neighbors of the old node 1 and to the node $2k + 1$, while node $1''$ will be connected to the neighbors of the old node 1 and to the node 2. The resulting graph is called $\mathscr{G}'$. Its complement $\overline{\mathscr{G}'}$ only contains the simple path

$$\{(1'', 2), (2, 3), \ldots, (2k, 2k + 1), (2k + 1, 1')\}.$$

Hence, $\overline{\mathscr{G}'}$ is perfect. By Lovász's perfect graph theorem [Lov84], $\mathscr{G}'$ is perfect, too. □

Note that Lemma 9 implies that, if the Strong Perfect Graph Conjecture [Ber61] is true, then we can transform every minimally imperfect graph into a perfect graph with a single clique-path substitution.

The following example gives an alternative substitution of a node in an antihole structure, which illustrates the more general clique-path substitution.

E 10. Consider an odd antihole $\overline{\mathscr{H}}_{2k+1} = (V, E)$ with $2k + 1$ nodes, labeled from 1 to $2k+1$. Pick node 1, then it is easy to verify that the set $V \setminus \{1\}$ can be partitioned into two cliques: Q_{odd} and Q_{even}. The former contains all nodes with an odd label (except node 1), the latter contains all nodes with an even label. So we can consider the following odd path of cliques:

$$P = (\{1\}, Q_{\text{even}} \setminus \{2\}, \{2\}, Q_{\text{odd}} \setminus \{3\}, \{3\})$$

Figure 6 shows the 7-antihole $\overline{\mathscr{H}}_7$; the edges of the complete subgraphs induced by the relevant cliques in P are shown with thick lines. The graph $\mathscr{G}'$ resulting from the clique-path

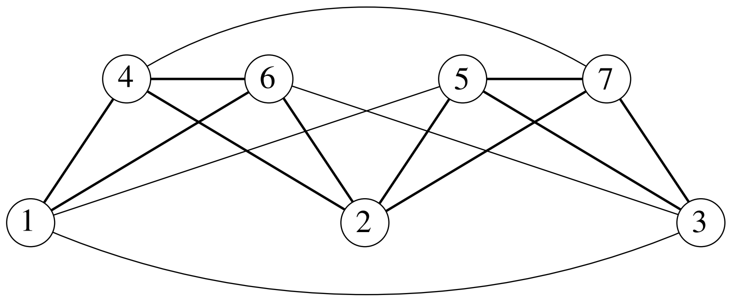

Figure 6. *A 7-antihole with a path of cliques*

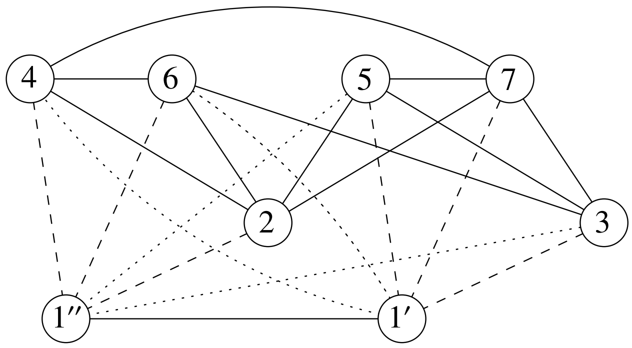

Figure 7. *A 7-antihole after a clique-path transformation*

substitution along P has two new nodes $1'$ and $1''$ replacing node 1; node $1'$ is connected to all nodes but 2, while node $1''$ is connected to all nodes but $2k+1$. The graph $\mathscr{G}'$ is shown in Figure 7, The edges introduced by the clique-path substitution are drawn with dashed lines, whereas edges merely inherited from node 1 are drawn with a dotted line. The resulting graph is the same as the one obtained by the node-path substitution of Lemma 9, hence it is perfect.

For any graph, it is easy to construct a finite sequence of clique-path substitutions leading to a graph that is the disjoint union of complete graphs, hence to a perfect graph:

L 11 (Finiteness). *Let $\mathscr{G} = (V, E)$ be a graph. There exists a finite sequence of clique-path substitutions leading to a perfect graph $\mathscr{G}^* = (V, E)$.*

Proof. Since clique-path substitutions work within one component, we may assume that $\mathscr{G}$ is connected. Let $Q \subseteq V$ be a clique in $\mathscr{G}$ that is maximal with respect to inclusion. If $Q = V$, we are done. Otherwise, let $v_1 \in V \setminus Q$ such that $Q_2 := Q \cap N_{\mathscr{G}}(v_1) \neq \emptyset$. Let $Q_3 := Q \setminus Q_2$. Now $P = (v_1, Q_2, Q_3)$ is an odd path of cliques in $\mathscr{G}$. The clique-path substitution along P leads to a graph $\mathscr{G}'$, where all the new nodes have been adjoined to the clique Q. We continue with $\mathscr{G}'$ and a maximal clique in $\mathscr{G}'$ containing the enlarged clique. Since $|V \setminus Q|$ decreases in each step by at least one, the procedure terminates with a complete graph. □

However, we cannot expect that an arbitrarily chosen sequence of clique-path substitutions terminates, not even if we restrict ourselves to paths of cliques such that for all $i \in \{1, \ldots, l\}$ there exists a node $u \in Q_{2i} \cup Q_{2i+1}$ with $(v_1, u) \notin E$.

R 12 (The struction of Ebenegger et al.). In the paper [EHdW84], Ebenegger, Hammer, and de Werra describe a construction that reduces the stability number of a graph by 1; in the subsequent papers [dW84, HMdW85a, HMdW85b] this is called a *struction*. Here, we shall present a variant that is a valid graph transformation, i.e., it maintains the weighted stability number rather than reducing it. Let $\mathscr{G} = (V, E)$ be a graph. Let $v_0 \in V$ be

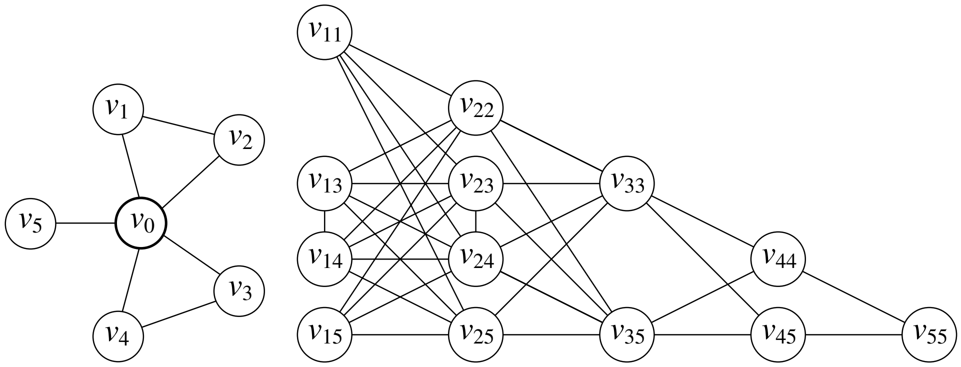

Figure 8. *A variant of the struction of a graph*

an arbitrary node and let $N(v_0) = \{v_1, \ldots, v_p\}$. The idea of the construction is the following: For each stable set S in $\mathscr{G}$ not containing v_0 but some of the nodes $v_1, \ldots, v_p$, there is a minimum index $i \in \{1, \ldots, p\}$ with $v_i \in S$. For each such minimum index i, in $\mathscr{G}'$ there is a layer of copies of those nodes $v_i, v_{i+1}, \ldots, v_p$ that are not adjacent to v_i in $\mathscr{G}$; these copies are called $v_{i,i}, v_{i,i+1}, \ldots, v_{i,p}$. These copies replace the original nodes $v_1, \ldots, v_p$. Within one layer, the $v_{i,j}$ inherits all the edges from both v_i and v_j in $\mathscr{G}$, whereas nodes of different layers are all connected by edges. Figure 8 illustrates the transformation; note that only edges between adjacent layers have been drawn here and that node v_0, which is connected to all new nodes, has been omitted. The algorithmic idea is to perform a sequence of structions, yielding a graph whose stability number can be computed easily. In the papers [dW84, HMdW85a, HMdW85b], it is shown that, for certain classes of graphs, the number of operations necessary is polynomially bounded. In the general case, however, one cannot expect similar results to hold, since the number of nodes in the problem may grow very fast. In fact, the method has never proved efficient for solving the stable set problem.

3 Optimizing over stable sets

In this section, we will deal with weighted stable set problems in a primal integer programming setting. It will turn out that the graph transformations discussed in the previous section can be re-interpreted as column operations in an integral simplex tableau.

The problem of finding a maximum stable set in $\mathscr{G} = (V, E)$ with respect to the weight function $\mathbf{c} \in \mathbf{R}_+^V$ can be formulated as the following combinatorial optimization problem:

$$\max \mathbf{c}^\mathsf{T}\mathbf{x} : \; x_v + x_w + z_{v,w} = 1 \;\text{ for } (v, w) \in E, \; \mathbf{x} \in \{0, 1\}^V, \; \mathbf{z} \in \{0, 1\}^E . \tag{6}$$

Note that in this formulation $z_{v,w}$ is the slack variable of the edge $(v, w) \in E$. A better linear programming formulation is achieved when edges are replaced by cliques of larger size. Let $Q_1, \dots, Q_k$ be cliques in $\mathscr{G}$ that cover all the edges of $\mathscr{G}$, i.e., for every edge $(v, w) \in E$ there is an index $i \in \{1, \dots, k\}$ such that $\{v, w\} \subseteq Q_i$. Note that this set of cliques may not coincide with all the maximal cliques of the complete clique formulation and that the cliques may not be maximal. Introducing a slack variable z_{Q_i} for each clique Q_i, the weighted stable set problem is formulated as

$$\max \mathbf{c}^\mathsf{T}\mathbf{x} : \; \sum_{v \in Q_i} x_v + z_{Q_i} = 1 \;\text{ for } i \in \{1, \dots, k\}, \; \mathbf{x} \in \{0, 1\}^V, \; \mathbf{z} \in \{0, 1\}^k. \tag{7}$$

This integer program is the starting point of our further investigations.

Now let $S \subseteq V$ be a stable set in $\mathscr{G}$. To construct a basic feasible solution associated with S we select for each of the rows in the program (7) a basic variable as follows:

- For every $v \in S$, let i_v be a row index such that $v \in Q_{i_v}$. We select x_v as the basic variable associated with the row i_v of the tableau.
- For each of the remaining clique constraints i we select the slack variable z_{Q_i} as the corresponding basic variable.

Note that the indices i_v are all distinct, and hence the construction yields a basis corresponding to S. As usual, let B and N denote the sets of basic and non-basic variables, respectively. We can now rewrite (7) in tableau form

$$\begin{aligned} \mathbf{y}_B + \mathbf{A}_B^{-1}\mathbf{A}_N\mathbf{y}_N &= \mathbf{b} \\ (\mathbf{y}_B, \mathbf{y}_N) &\in \mathbf{Z}_+^{n+k}. \end{aligned}$$

The variables y_j correspond to node variables x_v or slack variables z_{Q_i}. We will henceforth call a tableau obtained by this procedure a *canonical tableau* for S. Realizing the fact that the original objective function was non-negative on the nodes and zero on the slack variables, the following observation is immediate.

O **13.** Let N be the set of non-basic variables in a canonical tableau for a stable set S in the graph $\mathscr{G}$. Then the reduced cost of a non-basic slack variable z_{Q_i} is always non-positive. The reduced cost of a non-basic node variable x_k may be zero, negative or positive.

E **14.** Let $\mathscr{H}_5$ be an odd hole on five nodes. The problem of finding a stable set of maximum size in $\mathscr{H}_5$ is formulated in the following integer program:

$$\begin{array}{ll}
\max & x_1 + x_2 + x_3 + x_4 + x_5 \\
\text{s.t.} & x_1 + x_2 + z_{12} = 1 \\
& x_2 + x_3 + z_{23} = 1 \\
& x_3 + x_4 + z_{34} = 1 \\
& x_4 + x_5 + z_{45} = 1 \\
& x_1 + x_5 + z_{15} = 1
\end{array}$$

The construction described above for the stable set $S = \{3, 5\}$ in $\mathscr{H}_5$ yields the following tableau for S:

$$\begin{array}{ll}
\max & x_1 - z_{23} - z_{45} \\
\text{s.t.} & x_3 + x_2 + z_{23} = 1 \\
& x_5 + x_4 + z_{45} = 1 \\
& z_{34} - x_2 + x_4 - z_{23} = 0 \\
& z_{15} - x_4 + x_1 - z_{45} = 0 \\
& z_{12} + x_2 + x_1 = 1
\end{array}$$

Let us now consider how a clique-path substitution can be dealt with in an integer tableau.

D **15 (odd alternating path of cliques).** *Let $\mathscr{G} = (V, E)$ be a graph, σ a faithful labeling of $\mathscr{G}$ and let the node-weight function $\mathbf{c}\colon V \to \mathbf{R}_+$ be defined by* (2)*. Fix an integral simplex tableau for a formulation of the* $\mathbf{c}$*-weighted stable set problem in $\mathscr{G}$. Let x_{v_1} be a non-basic variable of positive reduced cost, and let $P = (\{v_1\}, Q_2, \ldots, Q_{2l+1})$ be an odd path of cliques in $\mathscr{G}$. Again we denote by R the set*

$$R = \{\, v \in Q_{2l+1} : v \text{ is not adjacent to } v_1 \text{ in } \mathscr{G} \,\}.$$

For $i \in \{1, \ldots, l\}$, assume there is a non-basic slack variable z_i for the clique $Q_{2i,2i+1} = Q_{2i} \cup Q_{2i+1}$, and there is $x_{j_i} = 1$ with $j_i \in Q_{2i}$. Moreover, presume that all variables x_v for $v \in R$ are non-basic. In this setting, we call

$$(\{v_1\}, Q_{2,3}, Q_{4,5}, \ldots, Q_{2l,2l+1})$$

an odd alternating path of cliques. *If* $|Q_i| = 1$ *for* $i \in \{2, \ldots, 2l+1\}$, *we call it an* odd alternating path of nodes.

In this setting, we are going to remove the column of the non-basic variable x_{v_1} from the tableau and replace it by certain sums of other columns of the tableau. This technique is closely related to the integral basis method by Haus, Köppe, and Weismantel, which is a primal algorithm for solving general integer programs [HKW01a]. We shall use the notation $x_v \wedge x_w$ for a new variable associated with a column that is the sum of the columns for x_v and x_w.

D 16. *For an odd alternating path of cliques* $(\{v_1\}, Q_{2,3}, Q_{4,5}, \ldots, Q_{2l,2l+1})$ *we define the corresponding* alternating-path substitution *in the tableau as follows: substitute* x_{v_1} *by new binary variables according to the following column operations:*

- $u_r = x_{v_1} \wedge x_r$ *for all* $r \in R$,
- $y_i = x_{v_1} \wedge z_i$ *for all* $i \in \{1, \ldots, l\}$,

where all new variables are non-basic and 0/1.

L 17. *The* 0/1 *solutions to the formulation obtained by an alternating-path substitution correspond to the stable sets in* $\mathscr{G}$.

In the proof of Lemma 17, we will make use of simple relations between the variables involved in the problem formulations:

O 18. Let $\bar{\mathbf{x}}$ and $\bar{\mathbf{z}}$ denote the variables $\mathbf{x}$ and $\mathbf{z}$, respectively, in the formulation obtained after the substitution. The following relations between the variables in the old formulation and the variables in the new formulation hold:

$$x_{v_1} = \textstyle\sum_{i=1}^{l} y_i + \sum_{r \in R} u_r, \tag{8a}$$

$$x_r = \bar{x}_r + u_r \qquad \text{for all } r \in R, \tag{8b}$$

$$z_i = \bar{z}_i + y_i \qquad \text{for all } i \in \{1, \ldots, l\}. \tag{8c}$$

For all other variables, we have $x_v = \bar{x}_v$ and $z_i = \bar{z}_i$.

Proof of Lemma 17. Let $(\mathscr{G}', \mathbf{c}', \sigma')$ denote the labeled graph obtained from $(\mathscr{G}, \mathbf{c}, \sigma)$ by performing the clique-path substitution along

$$P = (v_1, Q_2, Q_3, \ldots, Q_{2l}, Q_{2l+1}).$$

As in Definition 7, let t_r for $r \in R$ and w_i for $i \in \{1, \ldots, l\}$ denote the new nodes arising from the substitution. We show that the problem resulting from the above column operations is a formulation of the $\mathbf{c}'$-weighted stable set problem in $\mathscr{G}'$.

The key is to realize that the new variables correspond to the new nodes in the following way:

(i) For $i \in \{1, \ldots, l\}$, variable $y_i = x_{v_1} \wedge z_i$ corresponds to the new node w_i.

(ii) For $r \in R$, variable $u_r = x_{v_1} \wedge x_r$ corresponds to the new node t_r.

To verify (i), let $i \in \{1, \dots, l\}$ and note that the original formulation of the **c**-weighted stable set problem in $\mathscr{G}$ implies the following inequalities and equations:

$$\begin{aligned} x_{v_1} + x_v \le 1 &\quad \text{for } v \in N_{\mathscr{G}}(v_1), \\ \textstyle\sum_{v \in Q_{2i,2i+1}} x_v + z_i \le 1. & \end{aligned}$$

By (8), we obtain:

$$\begin{aligned} \textstyle\sum_{i=1}^{l} y_i + \sum_{r \in R} u_r + \bar{x}_v + u_v \le 1 &\quad \text{for } v \in N_{\mathscr{G}}(v_1), \\ \textstyle\sum_{v \in Q_{2i,2i+1}} \bar{x}_v + \sum_{v \in Q_{2i,2i+1} \cap R} u_v + \bar{z}_i + y_i = 1. & \end{aligned}$$

Hence, for all variables x_t corresponding to the neighbors $t \in N_{\mathscr{G}'}(w_i)$, as given by Definition 7, the new formulation implies an inequality $x_t + y_i \le 1$. The correspondence (ii) can be verified analogously. □

E 19 (Example 14 continued). For the stable set problem introduced in Example 14, the path

$$P = (1, \{2, 3\}, \{4, 5\})$$

is an alternating path of cliques. The slack variables z_{23} and z_{45} are both non-basic. Now x_1 is the variable in the tableau with positive reduced cost. Since the edge $(1, 5)$ is present in $\mathscr{G}$, we substitute variable x_1 by the two variables $x_1' = x_1 \wedge z_{23}$ and $x_1'' = x_1 \wedge z_{45}$ corresponding to the two sums of column 1 and the columns associated with z_{23} and z_{45}, respectively. Note that in this example the reduced cost of the two new non-basic columns are 0. Hence, we have a certificate that S is indeed optimal.

This illustrates the most fortunate situation for our algorithmic framework: The alternating-path substitution has not only turned the graph $\mathscr{G}$ into a perfect graph, as shown in Lemma 9, but we also obtain an integral tableau with a linear-programming certificate for optimality.

4 Properties of alternating-path substitutions

The destruction of one odd hole via our column substitution procedure has the price that we need to enlarge the original stable set problem significantly. One might argue that because of this enlargement of the graph it might algorithmically be more tractable to add just one odd-hole cutting plane to the initial formulation. The drawback of the latter approach is, however, that odd-hole cuts define facets for the stable set polytope associated with the odd hole only. For an arbitrary graph that contains an odd hole, lifting this cut is indispensable to make it strong. Our procedure, however, can directly deal with structures where a cut approach would need lifting, as the following example shows.

E 20. Let $\mathscr{G} = (V, E)$ be an odd wheel involving $2k$ nodes. The first $2k - 1$ nodes, numbered from 1 to $2k - 1$, form an odd hole. The additional node is called the *hub* and denoted by h. This node h is adjacent to all nodes on the hole. Associated with such a configuration is an inequality that can be seen as a lifted odd-hole inequality,

$$\sum_{i=1}^{2k-1} x_i + (k-1)x_h \leq k-1.$$

We will now show that destroying the odd hole by performing a node-path substitution automatically destroys the odd wheel. This implies that in this situation a concept of lifting is not required for the primal approach. We perform the same graph transformation as in the proof of Lemma 9 for the odd hole, i.e., a node-path substitution along the path

$$P = (1, 2, 3, \ldots, 2k-2, 2k-1).$$

The resulting graph $\mathscr{G}'$ is illustrated in Figure 9 for $k = 3$. Compared to the perfect graph obtained in the proof of Lemma 9, $\mathscr{G}'$ only has the extra node h, which is connected to all the other nodes. Thus $\mathscr{G}'$ is perfect as well.

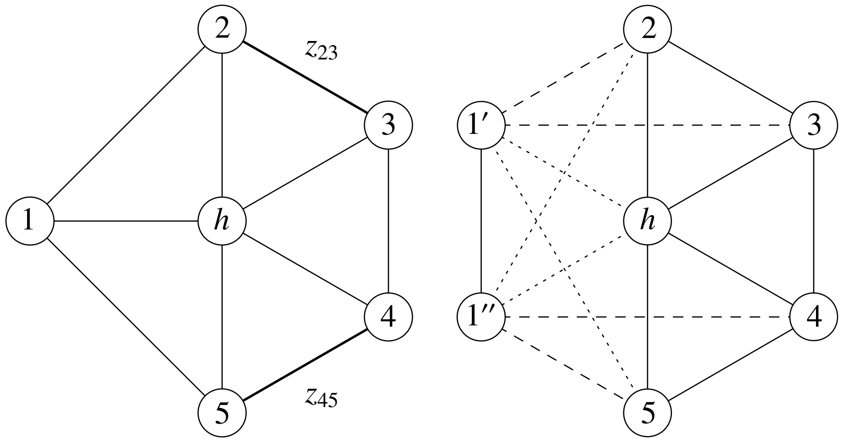

Figure 9. *Substitution of the non-hub node* 1 *in an odd wheel of size* 5

It has already been pointed out that a clique formulation (7) of the weighted stable set problem is much stronger than the node-edge formulation (6). In fact, when the maximal cliques are employed, the integrality constraints in (7) can be dropped if the underlying graph $\mathscr{G}$ is perfect.

We will now demonstrate in an example that alternating-path substitutions maintain maximal-clique formulations.

E 21 (Example 10, continued). Let us again consider the transformation of the odd antihole $\overline{\mathscr{H}}_{2k+1}$ carried out in Example 10. Let the problem be given in its maximal-clique formulation, that is composed of $2k+1$ cliques of size k, among which are the cliques Q_{odd} and Q_{even} from Example 10. Let a basis be fixed such that the node variable x_1 and the clique slack variables $z_{Q_{\text{odd}}}$ and $z_{Q_{\text{even}}}$ are non-basic. Then the clique-path substitution of Example 10 is equivalent to the column operation substituting x_1 by the variables:

- $x_{1'} = x_1 \wedge z_{Q_{\text{odd}}}$,
- $x_{1''} = x_1 \wedge z_{Q_{\text{even}}}$.

It is now easily verified that a maximal-clique formulation of the problem in the resulting perfect graph is obtained.

In fact, this holds in general; the following proposition states that alternating-path substitutions transform the formulation keeping the property that each clique inequality is facet inducing.

P **22.** *Let $\mathcal{G}$ be a graph and consider a maximal-clique formulation for the associated stable set problem. The formulation obtained after an alternating-path substitution is a maximal-clique formulation for the modified graph $\mathcal{G}'$.*

Proof. All the nodes arising from the substitution belong to all the (maximal) cliques Q such that $v_1 \in Q$. These cliques are extended as maximal cliques as the variables associated with all the new nodes are inserted in the related inequalities. Moreover, all the nodes w_i coming from the sum of the columns of x_{v_1} and of a maximal clique $Q_{i,i+1}$ belong also to the clique to $Q_{i,i+1} \cup \{t_i\}$, that now replaces the clique $Q_{i,i+1}$ in the formulation. Finally, all the nodes t_r coming from the sum of the columns of x_{v_1} and of x_r for $r \in R$ belong also to the maximal cliques containing either r or v_1, that are extended accordingly. All other cliques are not extended by this operation, so the corresponding inequalities are still facet defining. □

The alternating-path substitution requires that the cliques $Q_{2i} \cup Q_{2i+1}$ be present in the formulation. In the case that these cliques are not maximal, we have to add dominated inequalities. This situation can be easily handled, though:

C **23.** *If all the constraints in the formulation represent maximal cliques except, possibly, only those associated with the cliques $Q_{2i,2i+1}$ involved in a substitution along an odd alternating path of cliques $P = (v_1, Q_{2,3}, \ldots, Q_{2l,2l+1})$, the formulation obtained after the substitution is a maximal-clique formulation for the graph obtained with the corresponding transformation.*

This follows easily from Proposition 22.

Now we consider the *separation problem* for odd alternating paths of cliques. First we show that an alternating-path substitution "cuts off" the fractional point $(\mathbf{x}^F, \mathbf{z}^F)$ obtained by a single pivoting step applied to an basic integer solution $(\mathbf{x}^I, \mathbf{z}^I)$. Moreover, such an alternating path from a certain restricted class can be found in polynomial time if it exists.

D **24.** *Let $(\mathbf{x}^I, \mathbf{z}^I)$ be a basic integer solution and let v_i be a non-basic variable of positive reduced cost. If the basic solution obtained by pivoting in x_{v_1} is fractional, we call it a* fractional neighbor *of $(\mathbf{x}^I, \mathbf{z}^I)$ and denote it by $(\mathbf{x}^F, \mathbf{z}^F)$.*

L 25. *Let $(\mathbf{x}^I, \mathbf{z}^I)$ be a basic integer solution and x_{v_1} a variable with positive reduced cost. Suppose that after pivoting in x_{v_1} we obtain a fractional neighbor $(\mathbf{x}^F, \mathbf{z}^F)$. Apply an alternating-path substitution along $(v_1, Q_{2,3}, \ldots, Q_{2l,2l+1})$. Then the solution $(\mathbf{x}^F, \mathbf{z}^F)$ is not feasible for the projection according to Observation 18 of the new formulation onto the space of the initial variables.*

Proof. Let $y_i = x_{v_1} \wedge z_{Q_{2i,2i+1}}$ for $i \in \{1, \ldots, l\}$ and $u_r = x_{v_1} \wedge x_r$ for $r \in R$ be the new variables obtained by substitution of v_1. We denote by $\bar{\mathbf{x}}$ and $\bar{\mathbf{z}}$ the variables $\mathbf{x}$ and $\mathbf{z}$, respectively, in the formulation obtained after the substitution. The relations between the values of the variables in the old formulation and the variables in the new formulation have been stated in Observation 18. They are satisfied by all integer solutions. We will show that they are violated by the fractional solution $(\mathbf{x}^F, \mathbf{z}^F)$. First note that the variables z_i for $i \in \{1, \ldots, l\}$ and x_r for $r \in R$ are non-basic both before and after the pivoting operation. From equations (8b), (8c) and the non-negativity of all variables, we obtain that the variables $\bar{z}_i$ and y_i for $i \in \{1, \ldots, l\}$, and $\bar{x}_r$ and u_r for all $r \in R$ must have value zero in the solution corresponding to $(\mathbf{x}^F, \mathbf{z}^F)$. But at the same time $x_{v_1} > 0$, that is at least one of the variables y_i for $i \in \{1, \ldots, l\}$ or u_r for all $r \in R$ attains a positive value. This is a contradiction, therefore $(\mathbf{x}^F, \mathbf{z}^F)$ is not feasible for the projection of the new formulation onto the space of the original $(\mathbf{x}, \mathbf{z})$ variables. □

C 26. *Lemma 25 is still valid if the variables z_i for $i \in \{1, \ldots, l\}$ and x_r for $r \in R$ are equal to zero for both the solutions $(\mathbf{x}^I, \mathbf{z}^I)$ and $(\mathbf{x}^F, \mathbf{z}^F)$, independently of being basic or non-basic.*

In the following, we shall consider the alternating-path substitution in the case of $R = \emptyset$. In this case the path of cliques is in fact a cycle, so an associated tableau has the following form:

$$
\begin{array}{llllll}
x_{v_1} + \sum_{j \in Q_2} x_j & & & & & \leq 1 \\
\quad \sum_{j \in Q_2} x_j + \sum_{j \in Q_3} x_j & & & & + z_1 & = 1 \\
\quad\quad \sum_{j \in Q_3} x_j + \sum_{j \in Q_4} x_j & & & & & \leq 1 \\
 & \ddots & & & & \vdots \\
 & & \sum_{j \in Q_{2l}} x_j + \sum_{j \in Q_{2l+1}} x_j & & + z_l & = 1 \\
x_{v_1} & & & + \sum_{j \in Q_{2l+1}} x_j & & \leq 1
\end{array}
$$

If we perform the substitution of x_{v_1}, the equations in the above system can be rewritten as follows:

$$
\begin{array}{llll}
\sum_{j \in Q_2} x_j + \sum_{j \in Q_3} x_j & & + \bar{z}_1 + y_1 & = 1 \\
 & \ddots & & \vdots \\
 & \sum_{j \in Q_{2l}} x_j + \sum_{j \in Q_{2l+1}} x_j & + \bar{z}_l + y_l & = 1
\end{array}
$$

where $y_i = x_{v_1} \wedge z_i$ for all $i \in \{1, \ldots, l\}$, and, therefore, $x_{v_1} = \sum_{i=2}^{l} y_i$ and $z_i = \bar{z}_i + y_i$. By substituting $y_1 = x_{v_1} - \sum_{i=2}^{l} y_i$ in the first equation and then summing up all of them, we obtain the relation

$$\sum_{i=1}^{l}\Big(\sum_{j\in Q_{2i}} x_j + \sum_{j\in Q_{2i+1}} x_j\Big) + x_{v_1} + \sum_{i=1}^{l} \bar{z}_i = l, \tag{9}$$

which is a lifted odd-hole inequality resulting from the odd hole of length $2l+1$ of nodes $v_1, v_2, v_3, v_4, \ldots, v_{2l}, v_{2l+1}$, where v_i can be chosen as any node in Q_i for $i \in \{2, \ldots, 2l+1\}$. This means that the alternating-path substitution is equivalent to adding a lifted odd-hole cut to the problem.

T 27. *Let a formulation be fixed and let* $(\mathbf{x}^I, \mathbf{z}^I)$ *denote a basic integer solution. Suppose there exists an alternating-path substitution along*

$$P = (v_1, Q_{2,3}, \ldots, Q_{2l,2l+1})$$

such that the clique $Q_{2,3}, \ldots, Q_{2l,2l+1}$ *are dominated by cliques in the current formulation. Then one can find such a substitution in polynomial time, and it separates a fractional neighbor* $(\mathbf{x}^F, \mathbf{z}^F)$ *of* $(\mathbf{x}^I, \mathbf{z}^I)$.

Proof. Let $\hat{Q}_i$ for i from some index set K be the cliques in the current formulation. We build a digraph $\mathscr{H}_0$ where each node represents one clique $\hat{Q}_i$ whose corresponding variable is non-basic. we also have an additional node associated with the column x_{v_1} to be substituted. For each clique $\hat{Q}_i$, let x_{j_i} be the unique variable such that $j_i \in Q_i$ and $x^I_{j_i} = 1$ in the present basic integer solution. We have an arc labeled (i, m, k) from clique $\hat{Q}_i$ to clique $\hat{Q}_k$ if there exists an index $m \in K$ such that:

$$\hat{Q}_i \cap \hat{Q}_m \neq \emptyset \quad \text{and} \quad \hat{Q}_m \cap \hat{Q}_k \neq \emptyset; \tag{10a}$$

$$j_i \notin \hat{Q}_m \quad \text{whereas} \quad j_k \in \hat{Q}_m. \tag{10b}$$

The node associated with the variable x_{v_1} is connected to the cliques following the same rules, i.e., there exists an arc labeled $(0, m, k)$ from x_{v_1} to a clique $\hat{Q}_k$ not containing x_{v_1} if there exists an index $m \in K$ such that:

$$x_{v_1} \in \hat{Q}_m \quad \text{and} \quad \hat{Q}_m \cap \hat{Q}_k \neq \emptyset; \tag{11a}$$

$$j_k \in \hat{Q}_m. \tag{11b}$$

For arcs labeled $(k, m, 0)$ from a clique $\hat{Q}_k$ to x_{v_1} the last condition must be modified to read $j_k \notin \hat{Q}_m$.

Suppose there exists a directed cycle in $\mathscr{H}_0$ passing through x_{v_1},

$$C = (x_{v_1}, (0, m_1, k_1), \hat{Q}_{k_1}, (k_1, m_2, k_2), \hat{Q}_{k_2}, \ldots, Q_{k_l}, (k_l, m_{l+1}, 0), x_{v_1}).$$

We can construct an odd path of cliques from C as follows. For $i \in \{1, \ldots, l\}$, let

$$Q_{2i} = \hat{Q}_{m_i} \cap \hat{Q}_{k_i},$$
$$Q_{2i+1} = \hat{Q}_{k_i} \cap \hat{Q}_{m_{i+1}}.$$

Since $j_{k_i} \in Q_{2i}$ for $i \in \{1, \ldots, l\}$, the cliques $Q_{2i,2i+1} = Q_{2i} \cup Q_{2i+1}$ are tight at $(\mathbf{x}^I, \mathbf{z}^I)$. Note that each clique $Q_{2i,2i+1}$ is dominated by the clique $\hat{Q}_{k_i}$ that is present in the formulation.

After introducing non-basic variables for the slacks of $Q_{2i,2i+1}$ (unless already present), the path $P = (v_1, Q_{2,3}, \dots, Q_{2l,2l+1})$ clearly is an odd alternating path of cliques (see Figure 10). Conversely, every odd alternating path of cliques corresponds to a directed cycle C in $\mathscr{H}_0$.

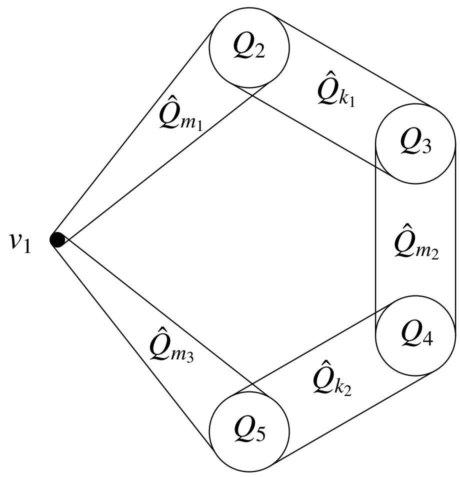

Figure 10. *Constructing an odd alternating path of cliques*

Hence, we can find them in polynomial time by detecting directed cycles in the auxiliary digraph $\mathscr{H}_0$. By Lemma 25, the resulting alternating-path substitution separates the fractional solution $(\mathbf{x}^F, \mathbf{z}^F)$ obtained by pivoting in x_{v_1}. $\square$

Note that one could also use a directed multi-graph, where two cliques $\hat{Q}_i$ and $\hat{Q}_k$ are connected by arcs (i, m, k) for each $m \in K$ satisfying the above conditions.

We have noted before that an alternating-path substitution is equivalent to adding a lifted odd-hole cut (9). In a dual-type method, one is interested in finding the cut from a given class that is *most violated* by the current solution. We can solve the analogous problem in our primal setting:

P **28.** *The problem to find the alternating-path substitution of the type as in Theorem 27, corresponding to the inequality* (9) *that is most violated by a fractional neighbor* $(\mathbf{x}^F, \mathbf{z}^F)$ *of a basic integral solution* $(\mathbf{x}^I, \mathbf{z}^I)$*, can be solved in polynomial time.*

Proof. Let $\mathscr{H}$ be the graph constructed following the same rules used for $\mathscr{H}_0$, but with nodes associated with all tight cliques for $(\mathbf{x}^I, \mathbf{z}^I)$. Let us now define a digraph $\mathscr{H}'$ where the nodes are defined by all those subcliques obtained by the intersections of the cliques $\hat{Q}_i$ with the cliques $\hat{Q}_m$ for all the arcs (i, m, k) and (k, m, i) of $\mathscr{H}$. The graph $\mathscr{H}'$ inherits all the arcs of $\mathscr{H}$ and has arcs between the pairs of subcliques of the same original clique $\hat{Q}_i$. We define the weight on each arc $e = (Q', Q'')$ as

$$w_e = 1 - \sum_{j \in Q' \cup Q''} x_j^F. \tag{12}$$

Note that $w_e \geq 0$ as $Q' \cup Q''$ is a subset of a clique in the formulation. Now the minimum directed odd cycle in $\mathscr{H}'$ passing through x_{v_1} yields the alternating-path substitution for x_{v_1} corresponding to the most violated constraint (9). □

Note that the algorithm described in the proof is in fact a modification of the standard algorithm for separating odd-hole inequalities but in this primal setting we have to deal with directed graphs instead of undirected ones, and every cycle of $\mathscr{H}'$ passing through x_{v_1} is odd.

Finally, we shall consider how we can use the alternating-path substitution in the situation where an integral tableau for a maximum weighted stable set is given, but the linear-programming certificate for optimality is still missing, as in Example 19. The idea is to substitute columns with positive reduced cost along odd alternating paths of cliques, until a tableau is obtained where each column has non-positive reduced cost.

P 29 (Column substitution problem). *Let $\mathscr{G} = (V, E)$ be a graph and $S \subseteq V$ a stable set in $\mathscr{G}$. Let $\mathbf{x} = (\mathbf{A}_B^{-1}\mathbf{b}, \mathbf{0})$ be a basic feasible solution corresponding to S where*

$$\max \mathbf{c}^{\top}\mathbf{x} : \mathbf{A}\mathbf{x} = \mathbf{b},\ \mathbf{x} \in \mathbf{Z}_+^n$$

is a formulation of the maximum weighted stable set problem in $\mathscr{G}$. Let x_{v_1} be a non-basic variable in N of positive reduced cost. Does there exist an odd alternating path of cliques $P = (v_1, Q_{2,3}, \ldots, Q_{2l,2l+1})$ in the given formulation, such that the substitution along P yields a tableau where all new columns have non-positive reduced cost?

If there exists an odd path P with the desired properties, then we know that we can replace the non-basic variable x_1 of positive reduced cost by new columns according to Lemma 17 such that the reduced cost of the new columns will all be non-positive.

C 30. *The column substitution problem, restricted to the case $R = \emptyset$, can be solved in polynomial time.*

Proof. We consider a variant of the construction in the proof of Proposition 28. After constructing $\mathscr{H}'$, we remove those arcs $e = (Q', Q'')$ that would give rise to a non-basic variable of positive reduced cost in the substitution. Then every directed odd cycle in $\mathscr{H}'$ passing through v_1 corresponds to an odd alternating path of cliques with the required properties. □

Conclusions In this paper, we have studied primal operations on formulations for the stable set problem. We have presented graph transformations that, at the expense of enlarging the graph, can produce a perfect graph, as shown in Lemma 11. The graph transformations have a natural analogue in an integral tableau setting. Lemma 25, Theorem 27 and Corollary 30 show that dual separation results carry over to a primal setting. These results give the foundation for a primal algorithm.

Finally, testing the computational effectiveness of our approach will be the topic a forthcoming paper.

Bibliography

[BBPP99] I. Bomze, M. Budinich, P. Pardalos, and M. Pelillo, *The maximum clique problem*, Handbook of Combinatorial Optimization (Supplement Volume A) (D.-Z. Du and P. Pardalos, eds.), vol. 4, Kluwer Academic Publishers, Boston, MA, 1999.

[BCCP96] E. Balas, S. Ceria, G. Cornuéjols, and G. Pataki, *Polyhedral methods for the maximum clique problem*, In Johnson and Trick [JT96], pp. 11–28.

[Ber61] C. Berge, *Färbung von Graphen, deren sämtliche bzw. deren ungerade Kreise starr sind*, Wiss. Z. Martin-Luther-Univ. Halle-Wittenberg, Math.-Natur. Reihe (1961), 114–115.

[Bor98] R. Borndörfer, *Aspects of set packing, partitioning, and covering*, Dissertation, Technische Universität Berlin, 1998, published by Shaker-Verlag, Aachen.

[BP75] E. Balas and M. Padberg, *On the set-covering problem: II. An algorithm for set partitioning*, Operations Research **23** (1975), 74–90.

[Cor01] G. Cornuéjols, *Combinatorial optimization: Packing and covering*, CBMS-NSF Regional Conference Series in Applied Mathematics, no. 74, SIAM, Philadelphia, 2001.

[dW84] D. de Werra, *On some properties of the struction of a graph*, SIAM J. Algebraic Discrete Methods **5** (1984), no. 2, 239–243.

[EHdW84] C. Ebenegger, P. L. Hammer, and D. de Werra, *Pseudo-Boolean functions and stability of graphs*, Algebraic and combinatorial methods in operations research, North-Holland, Amsterdam, 1984, pp. 83–97.

[GLS88] M. Grötschel, L. Lovász, and A. Schrijver, *Geometric algorithms and combinatorial optimization*, Algorithms and Combinatorics, vol. 2, Springer, Berlin, 1988.

[HKW01a] U.-U. Haus, M. Köppe, and R. Weismantel, *The Integral Basis Method for integer programming*, Mathematical Methods of Operations Research **53** (2001), no. 3, 353–361.

[HKW01b] U.-U. Haus, M. Köppe, and R. Weismantel, *The Integral Basis Method in an augmentation framework for integer programming*, Preprint no. 23, Fakultät für Mathematik, Otto-von-Guericke-Universität Magdeburg, 2001, available from URL `http://www.math.uni-magdeburg.de/~mkoeppe/art/haus-koeppe-weismantel-ibm-augmentation.ps`.

[HKW01c] U.-U. Haus, M. Köppe, and R. Weismantel, *A primal all-integer algorithm based on irreducible solutions*, Preprint no. 5, Fakultät für Mathematik, Otto-von-Guericke-Universität Magdeburg, 2001, available from URL `http://www.math.uni-magdeburg.de/~mkoeppe/art/ibm-theory.ps`.

[HMdW85a] P. L. Hammer, N. V. R. Mahadev, and D. de Werra, *Stability in CAN-free graphs*, J. Combin. Theory Ser. B **38** (1985), no. 1, 23–30.

[HMdW85b] P. L. Hammer, N. V. R. Mahadev, and D. de Werra, *The struction of a graph: application to CN-free graphs*, Combinatorica **5** (1985), no. 2, 141–147.

[JT96] D. Johnson and M. Trick (eds.), *Clique, coloring, and satisfiability: Second DIMACS implementation challenge, DIMACS*, vol. 26, American Mathematical Society, 1996.

[Lov84] L. Lovász, *Normal hypergraphs and the weak perfect graph conjecture*, Topics on perfect graphs, North-Holland, Amsterdam, 1984, pp. 29–42.

[NS92] G. Nemhauser and G. Sigismondi, *A strong cutting plane/branch and bound algorithm for node packing*, Journal of the Operational Research Society (1992), 443–457.

[NT73] G. L. Nemhauser and L. E. Trotter, *Properties of vertex packing and independent system polyhedra*, Mathematical Programming **6** (1973), 48–61.

[Pad73] M. W. Padberg, *On the facial structure of set packing polyhedra*, Mathematical Programming **5** (1973), 199–215.

[PH80] M. W. Padberg and S. Hong, *On the symmetric travelling salesman problem: a computational study*, Math. Programming Stud. (1980), no. 12, 78–107.

[You68] R. D. Young, *A simplified primal (all-integer) integer programming algorithm*, Operations Research **16** (1968), no. 4, 750–782.