Persistency of Linear Programming Relaxations for the Stable Set Problem

Elisabeth Rodríguez-Heck[1], Karl Stickler[1], Matthias Walter[2(✉)], and Stefan Weltge[3]

[1] Lehrstuhl für Operations Research, RWTH Aachen University, Aachen, Germany
rodriguez-heck@or.rwth-aachen.de, karl.stickler@rwth-aachen.de
[2] Department of Applied Mathematics, University of Twente, Enschede, The Netherlands
m.walter@utwente.nl
[3] Department of Mathematics, Technical University of Munich, Munich, Germany
weltge@tum.de

Abstract. The Nemhauser-Trotter theorem states that the standard linear programming (LP) formulation for the stable set problem has a remarkable property, also known as (weak) *persistency*: for every optimal LP solution that assigns integer values to some variables, there exists an optimal integer solution in which these variables retain the same values. While the standard LP is defined by only non-negativity and edge constraints, a variety of stronger LP formulations have been studied and one may wonder whether any of them has the this property as well. We show that any stronger LP formulation that satisfies mild conditions cannot have the persistency property on all graphs, unless it is always equal to the stable-set polytope.

Keywords: Persistency · Integer linear programming · Stable set

1 Introduction

Given an undirected graph G with node set $V(G)$ and edge set $E(G)$, and node weights $w \in \mathbb{R}^{V(G)}$, the (weighted) stable-set problem asks for finding a stable set S in G that maximizes $\sum_{v \in S} w_v$, where a set S is called stable if G contains no edge with both endpoints in S. While the stable-set problem is NP-hard, it is a common approach to maximize $w^\intercal x$ over the *edge relaxation*

$$R_{\text{stab}}^{\text{edge}}(G) := \left\{ x \in [0,1]^{V(G)} \mid x_v + x_w \leq 1 \text{ for each edge } \{v,w\} \in E(G) \right\}$$

and use optimal (fractional) solutions to gain insights about optimal 0/1-solutions. Note that the 0/1-points in the edge relaxation are precisely the characteristic vectors of stable sets in G, and that maximizing a linear objective over the edge relaxation is a linear program that can be solved efficiently. Given an optimal solution of this linear program, its objective value is clearly an upper

D. Bienstock and G. Zambelli (Eds.): IPCO 2020, LNCS 12125, pp. 351–363, 2020.
https://doi.org/10.1007/978-3-030-45771-6_27

bound on the value of any 0/1-solution and its entries may guide initial decisions in a branch-and-bound algorithm. While this is also the case for general polyhedral relaxations, it turns out that optimal solutions of the edge relaxation have a remarkable property that allows to reduce the size of the problem by fixing some variables to provable optimal integer values.

Definition 1 (Persistency). *We say that a polytope $P \subseteq [0,1]^n$ has the* persistency property *if for every objective vector $c \in \mathbb{R}^n$ and every c-maximal point $x \in P$, there exists a c-maximal integer point $y \in P \cap \{0,1\}^n$ such that $x_i = y_i$ for each $i \in \{1, 2, \dots, n\}$ with $x_i \in \{0,1\}$.*

Proposition 1 (Nemhauser and Trotter [8]). *The edge relaxation $R^{edge}_{stab}(G)$ has the persistency property for every graph G.*

In other words, the result of Nemhauser and Trotter [8] states that if $x^\star$ is an optimal solution for the edge relaxation, then there exists an optimal stable set $S^\star$ satisfying $V_1 \subseteq S^\star \subseteq V(G) \setminus V_0$, where $V_i := \{v \in V(G) \mid x^\star_v = i\}$ for $i = 0, 1$. In this case, the nodes in $V_0 \cup V_1$ can be deleted and the search only has to be performed on the remaining graph. Clearly, this reduction is significant if $x^\star$ assigns integer values to many nodes.

For the maximum cardinality stable set problem it has been shown that the probability of obtaining a single integer component when solving the LP relaxation is very low for large random graphs [11]. However, persistencies have been proved to be very useful in a different context, when dealing with highly structured instances arising in the field of computer vision. More precisely, Hammer, Hansen and Simeone [4] provided a reduction of (Unconstrained) Quadratic Binary Programming (QBP) to the stable set problem and showed that weak persistency holds for (QBP) as well. Boros et al. [1] provided an algorithm to compute the largest possible set of variables to fix via persistencies in a quadratic binary program in polynomial time, which has been successfully used in practice to solve very large image restoration problems [3,5,7].

In general, dual bounds obtained from the edge relaxation are quite weak, and several families of additional inequalities have been studied in order to strengthen this formulation. Examples are the clique inequalities [10], (lifted) odd-cycle inequalities [10,15] and clique-family inequalities [9]. Most of these families were discovered by systematically studying the facets of the *stable-set polytope* $P_{\text{stab}}(G)$, which is the convex hull of the characteristic vectors of stable sets in G. The stable-set polytope itself is known to be a complicated polytope. In particular, one cannot expect to be able to completely characterize its facial structure [6]. Thus, the following question is natural.

> *Do there exist stronger linear programming formulations for the stable set problem that also have the persistency property for every graph G?*

In this paper, we answer the question negatively. More precisely, we show that an LP formulation (satisfying mild conditions) that is stronger than the edge formulation cannot have the persistency property on all graphs, unless it always yields the stable-set polytope.

Outline. The paper is structured as follows. We start by introducing the conditions we impose on the LP formulation in Sect. 2. Our main result and its consequences are presented in Sect. 3. Section 4 is dedicated to the proof of the main result. Our preprint [12] provides running examples that illustrate the steps of the proof.

2 LP Formulations for Stable Set

It is clear that, for a *single* non-bipartite graph G, one can artificially construct polytopes strictly between $R^{\text{edge}}_{\text{stab}}(G)$ and $P_{\text{stab}}(G)$ that have the persistency property. For instance, if $x \in R^{\text{edge}}_{\text{stab}}(G) \setminus P_{\text{stab}}(G)$ is any point that has only fractional coordinates, then the polytope $\text{conv}(P_{\text{stab}}(G) \cup x)$ has the persistency property for trivial reasons. In this work, however, we consider relaxations defined for *every* graph that arise in a more structured way.

To this end, let $\mathcal{G}$ denote the set of finite undirected simple graphs. We regard an LP *formulation* for the stable set problem as a map that assigns to every graph $G \in \mathcal{G}$ a polytope $R_{\text{stab}}(G) \supseteq P_{\text{stab}}(G)$. As an example, the edge formulation assigns $R^{\text{edge}}_{\text{stab}}(G)$ to every graph G. Next, let us specify some natural conditions that are satisfied by many prominent formulations and under which our main result holds. Each of these conditions is defined for a formulation R_{stab}.

First, we require that the formulation R_{stab} is at least as strong as the edge formulation. Formally,

$$\text{for each } G \in \mathcal{G}, \text{ we have } P_{\text{stab}}(G) \subseteq R_{\text{stab}}(G) \subseteq R^{\text{edge}}_{\text{stab}}(G). \tag{A}$$

Second, the inequalities defining R_{stab} must be derived from facets of P_{stab}:

for each $G \in \mathcal{G}$, each inequality with support $U \subseteq V(G)$ that is facet-defining for $R_{\text{stab}}(G)$ is also facet-defining for $P_{\text{stab}}(G[U])$, (B)

where $G[U]$ denotes the subgraph induced by U. Note that inequalities need to define facets only on their support graph, hence also the generally not facet-defining odd-cycle inequalities (see [10]) satisfy (B). However, a formulation consisting of only rank inequalities (see [2]) does not satisfy (B).

Third, for every graph $G \in \mathcal{G}$, validity of facet-defining inequalities of $R_{\text{stab}}(G)$ shall be inherited by induced subgraphs. Formally,

for each $G \in \mathcal{G}$, each inequality with support $U \subseteq V(G)$ that is facet-defining for $R_{\text{stab}}(G)$ is valid (although not necessarily facet-defining) for $R_{\text{stab}}(G[U])$. (C)

This requirement ensures that if an (irredundant) inequality arises for some graph then it must (at least implicitly) occur for all induced subgraphs for which it is defined. The reverse implication is imposed by the fourth condition, although in a more structured way. For this, we need the following definitions.

Let $G_1, G_2 \in \mathcal{G}$ and let $v_1 \in V(G_1)$, $v_2 \in V(G_2)$. Then the 1-*sum of* G_1 *and* G_2 *at* v_1 *and* v_2, denoted by $G_1 \oplus_{v_2}^{v_1} G_2$ is the graph obtained from the disjoint union of G_1 and G_2 by identifying v_1 with v_2. Moreover, let $P \subseteq \mathbb{R}^m$ and $Q \subseteq \mathbb{R}^n$ be polytopes and let $i \in \{1, 2, \ldots, m\}$ and $j \in \{1, 2, \ldots, n\}$. The 1-*sum of* P *and* Q *at coordinates* i *and* j, denoted by $P \oplus_j^i Q$, is defined as the projection of $\operatorname{conv}(\{(x, y) \in P \times Q \mid x_i = y_j\})$ onto all variables except for y_j. Notice that this projection is an isomorphism from the convex hull to its image since the variables x_i and y_j are equal.

Our fourth condition requires that for every pair of graphs $G_1, G_2 \in \mathcal{G}$, validity of inequalities is acquired by their 1-sum. Formally,

$$R_{\text{stab}}(G_1 \oplus_{v_2}^{v_1} G_2) = R_{\text{stab}}(G_1) \oplus_{v_2}^{v_1} R_{\text{stab}}(G_2) \text{ holds for all } G_1, G_2 \in \mathcal{G} \text{ and all nodes } v_1 \in V(G_1) \text{ and } v_2 \in V(G_2). \tag{D}$$

Also this condition is very natural since every inequality that is valid for $R_{\text{stab}}(G_1)$ is also valid for $P_{\text{stab}}(G_1 \oplus_{v_2}^{v_1} G_2)$, and hence its participation in $R_{\text{stab}}(G_1 \oplus_{v_2}^{v_1} G_2)$ is reasonable.

3 Results

We say that two formulations R_{stab}^1 and R_{stab}^2 are *equivalent* if $R_{\text{stab}}^1(G) = R_{\text{stab}}^2(G)$ holds for every $G \in \mathcal{G}$, in which case we write $R_{\text{stab}}^1 \equiv R_{\text{stab}}^2$. We can now state our main result.

Theorem 1. *Let R_{stab} be a formulation satisfying (A)–(D). Then $R_{stab}(G)$ has the persistency property for all graphs $G \in \mathcal{G}$ if and only if $R_{stab} \equiv R_{stab}^{edge}$ or $R_{stab} \equiv P_{stab}$.*

Sufficiency follows from Proposition 1 and from the fact that $P_{\text{stab}}(G)$ is an integral polytope for every $G \in \mathcal{G}$. Before we prove necessity in Sect. 4, let us mention some direct implications of Theorem 1 for known relaxations.

Corollary 1. *The clique relaxation*

$$R_{stab}^{clq}(G) = \left\{ x \in \mathbb{R}^{V(G)} \mid x(V(C)) \leq 1 \text{ for each clique } C \text{ of } G \right\}$$

does not have the persistency property for all graphs $G \in \mathcal{G}$.

Proof. It is easy to see that $R_{\text{stab}}^{\text{clq}}$ satisfies Properties (A) and (D). For Properties (B) and (C), consider a clique C of some graph $G \in \mathcal{G}$. Clearly, C is also a clique of $G[V(C)]$ and the inequality is known to be facet-defining for $P_{\text{stab}}(G[V(C)])$ (see Theorem 2.4 in [10]). □

Also the relaxation based on odd-cycle inequalities satisfies these properties, although the inequalities are generally not facet-defining.

Corollary 2. *The odd-cycle relaxation*

$$R^{oc}_{stab}(G) = \left\{ x \in R^{edge}_{stab}(G) \mid x(V(C)) \leq \tfrac{|V(C)|-1}{2} \text{ for each odd cycle } C \text{ of } G \right\}$$

does not have the persistency property for all graphs $G \in \mathcal{G}$.

Proof. It is easy to see that $R^{\text{oc}}_{\text{stab}}$ satisfies Properties (A) and (D). For Properties (B) and (C), consider an odd cycle C of some graph $G \in \mathcal{G}$. To induce a facet, C must be chordless, and the odd-cycle inequality is facet-defining for $P_{\text{stab}}(G[V(C)])$ (see Theorem 3.3 in [10]). □

4 Proof of the Main Result

Let us fix any formulation R_{stab} over $\mathcal{G}$ satisfying Properties (A)–(D). To prove the "only if" implication of Theorem 1 we have to verify that if $R_{\text{stab}} \not\equiv R^{\text{edge}}_{\text{stab}}$ and $R_{\text{stab}} \not\equiv P_{\text{stab}}$, then $R_{\text{stab}}(G)$ does not have the persistency property for all graphs $G \in \mathcal{G}$. Equivalently, we have to prove the following:

> If there exist graphs $G_1, G_2 \in \mathcal{G}$ with $R_{\text{stab}}(G_1) \neq R^{\text{edge}}_{\text{stab}}(G_1)$ and $R_{\text{stab}}(G_2) \neq P_{\text{stab}}(G_2)$, then there exists a graph $G^\star$ for which the polytope $R_{\text{stab}}(G^\star)$ does not have the persistency property. (♢)

Given G_1 and G_2, we will provide an explicit construction of $G^\star$ and show that $R_{\text{stab}}(G^\star)$ does not have the persistency property. To see the latter, we will give an objective vector $c^\star \in \mathbb{R}^{V(G^\star)}$ such that every $c^\star$-maximal solution over $R_{\text{stab}}(G^\star)$ has a certain coordinate equal to zero while every $c^\star$-maximal stable set in $G^\star$ contains the corresponding node.

The graph $G^\star$ will consist of an "inner" graph G^{in} with $R_{\text{stab}}(G^{\text{in}}) \neq R^{\text{edge}}_{\text{stab}}(G^{\text{in}})$ and $|V(G^{\text{in}})| - 1$ copies of an "outer" graph G^{out} with $R_{\text{stab}}(G^{\text{out}}) \neq P_{\text{stab}}(G^{\text{out}})$. Each copy of G^{out} is attached to a node of G^{in} via the 1-sum operation. The only node of G^{in} that does *not* have a copy of G^{out} attached corresponds precisely to the coordinate showing that $R_{\text{stab}}(G^\star)$ does not have the persistency property. Note that such graphs $G^{\text{in}}, G^{\text{out}}$ exist due to the hypothesis of (♢). Among all such graphs, we will make particular choices satisfying some additional properties that we specify in the next sections.

4.1 The Graph G^{out}

In the definition of the auxiliary graph G^{out} we will make use of the following lemma. In what follows, for a polytope $P \subseteq \mathbb{R}^n$ and a vector $c \in \mathbb{R}^n$, let us denote the optimal face of P induced by c by $\text{opt}(P, c) := \arg\max \{c^\intercal x \mid x \in P\}$.

Lemma 1. *Let* $P, Q \subseteq \mathbb{R}^n$ *be polytopes. If there exists a vector* $c \in \mathbb{R}^n$ *such that* $\dim(\text{opt}(Q, c)) < \dim(\text{opt}(P, c))$*, then there exists a vector* $c' \in \mathbb{R}^n$ *such that* $\text{opt}(Q, c')$ *is a vertex of* Q*, while* $\text{opt}(P, c')$ *is not a vertex of* P.

The lemma is proved in Appendix A. The graph G^{out} is now defined through the following statement.

Claim 1. Assuming the hypothesis of ($\diamondsuit$), there exists a graph $G^{\text{out}} \in \mathcal{G}$, a vector $c^{\text{out}} \in \mathbb{R}^{V(G^{\text{out}})}$ and a node $v^{\text{out}} \in V(G^{\text{out}})$ such that $\text{opt}(R_{\text{stab}}(G^{\text{out}}), c^{\text{out}}) = \{\hat{x}\}$ holds with $\hat{x}_{v^{\text{out}}} \geq \frac{1}{2}$ and such that $\text{opt}(P_{\text{stab}}(G), c^{\text{out}})$ contains a vertex $\bar{x} \in \{0,1\}^{V(G^{\text{out}})}$ with $\bar{x}_{v^{\text{out}}} = 0$.

Proof. Let $G \in \mathcal{G}$ be such that $R_{\text{stab}}(G) \neq P_{\text{stab}}(G)$. Such a graph exists by hypothesis of ($\diamondsuit$). By Property (A), there exists an inequality $a^\intercal x \leq \delta$ that is facet-defining for $P_{\text{stab}}(G)$, but not valid for $R_{\text{stab}}(G)$.

We claim that the face $\text{opt}(R_{\text{stab}}(G), a)$ is not a facet of $R_{\text{stab}}(G)$. Assume for a contradiction that $\text{opt}(R_{\text{stab}}(G), a)$ is a facet of $R_{\text{stab}}(G)$ and define $\delta' := \max\{a^\intercal x \mid x \in R_{\text{stab}}(G)\}$. Since $a^\intercal x \leq \delta$ is not valid for $R_{\text{stab}}(G)$, we have $\delta' > \delta$. Property (B) implies that $a^\intercal x \leq \delta'$ is facet-defining for $P_{\text{stab}}(G[\text{supp}(a)])$, and in particular, equality holds for the characteristic vector of some stable set $S \subseteq V(G[\text{supp}(a)])$. Since S is also a stable set in G, this contradicts the assumption that $a^\intercal x \leq \delta$ is valid for $P_{\text{stab}}(G)$.

By Lemma 1, there exists a vector $c \in \mathbb{R}^n$ such that $\text{opt}(R_{\text{stab}}(G), c) = \{\hat{x}\}$ and $\text{opt}(P_{\text{stab}}(G), c)$ has (at least) two vertices $\bar{x}^1, \bar{x}^2 \in \{0,1\}^{V(G)}$. Since $\bar{x}^1 \neq \bar{x}^2$, there exists a coordinate $u \in V(G)$ at which they differ and we can assume $\bar{x}^1_u = 0$ and $\bar{x}^2_u = 1$ without loss of generality. If $\hat{x}_u \geq \frac{1}{2}$, we can choose $G^{\text{out}} := G$, $c^{\text{out}} := c$ and $v^{\text{out}} := u$. Together with $\hat{x}$ and $\bar{x}^1$, they satisfy the requirements of the lemma.

Otherwise, let G' be the graph G with an additional edge $\{u, u'\}$ attached at u. Formally, let G'' be the graph consisting of a single edge $\{u, u'\}$ and let $G' := G \oplus_u^u G''$. By Property (D), $R_{\text{stab}}(G') = R_{\text{stab}}(G) \oplus_u^u R_{\text{stab}}(G'')$ holds. Since G'' is a single edge, $R^{\text{edge}}_{\text{stab}}(G'') = P_{\text{stab}}(G'')$ holds. Thus, $R_{\text{stab}}(G')$ is described by all inequalities that are valid for $R_{\text{stab}}(G)$ together with $x_{u'} \geq 0$ and $x_u + x_{u'} \leq 1$. Hence, for a sufficiently small $\varepsilon > 0$ and the objective vector $c' \in \mathbb{R}^{V(G')}$ with $c'_{u'} = \varepsilon$, $c'_u = c_u + 2\varepsilon$ and $c'_v = c_v$ for all $v \in V(G) \setminus \{u\}$, the maximization of c' over $R_{\text{stab}}(G')$ yields a unique optimum $\hat{x}' \in \mathbb{R}^{V(G')}$ with $\hat{x}'_v = \hat{x}_v$ for all $v \in V(G)$ and $\hat{x}'_{u'} = 1 - \hat{x}'_u > \frac{1}{2}$, while the maximization of c' over $P_{\text{stab}}(G')$ admits an optimum $\bar{x}' \in \mathbb{R}^{V(G')}$ with $\bar{x}'_u = 1$ and $\bar{x}'_{u'} = 0$. Now, $G^{\text{out}} := G'$, $c^{\text{out}} := c'$ and $v^{\text{out}} := u'$ together with $\hat{x}'$ and $\bar{x}'$ satisfy the requirements of the lemma. $\square$

4.2 The Graph G^{in}

Among all graphs $G \in \mathcal{G}$ with $R_{\text{stab}}(G) \neq R^{\text{edge}}_{\text{stab}}(G)$ we choose G^{in} to have a minimum number of nodes. Note that G^{in} exists by hypothesis of ($\diamondsuit$). We assume $V(G^{\text{in}}) = \{1, 2, \ldots, n\}$. Let $Ax \leq b$ (with $A \in \mathbb{Z}^{m \times n}$ and $b \in \mathbb{Z}^m$) be the system containing inequalities for all facets of $R_{\text{stab}}(G^{\text{in}})$ that are not valid for $R^{\text{edge}}_{\text{stab}}(G^{\text{in}})$. Note that $m \geq 1$ and $n \geq 3$ hold by assumption on G^{in}.

Claim 2. $A_{i,j} \geq 1$ holds for every $i \in \{1, 2, \ldots, m\}$ and every $j \in \{1, 2, \ldots, n\}$.

Proof. It is a basic fact that every facet-defining inequality of a stable-set polytope that is not a nonnegativity constraint is of the form $a^{\intercal}x \leq \beta$ for some nonnegative vector $a \in \mathbb{R}^n$ (see Section 9.3 in [13]). Assume, $A_{i,j} = 0$ holds for some i, j. By Property (C), $A_{i,\star}x \leq b_i$ is valid for $R_{\text{stab}}(G[\text{supp}(A_{i,\star})])$, while it is not valid for $R_{\text{stab}}^{\text{edge}}(G^{\text{in}}[\text{supp}(A_{i,\star})])$, contradicting minimality of G^{in}. □

4.3 The Graph $G^\star$

For each $j \in \{2, 3, \dots, n\}$ let G^j be an isomorphic copy of G^{out} such that $V(G^j) \cap V(G^k) = \varnothing$ whenever $j \neq k$. Let $c^j \in \mathbb{R}^{V(G^j)}$ and $v^j \in V(G^j)$ be the vector and node corresponding to c^{out} and v^{out} in Claim 1, respectively. Now $G^\star$ is defined as the 1-sum of G^{in} with all G^j at the respective nodes $j \in V(G^{\text{in}})$ and $v^j \in V(G^j)$, i.e., $G^\star := G^{\text{in}} \oplus_{v^2}^2 G^2 \oplus_{v^3}^3 \cdots \oplus_{v^n}^n G^n$, where the $\oplus$-operator has to be applied from left to right. By Property (D) we have

$$R_{\text{stab}}(G^\star) = R_{\text{stab}}(G^{\text{in}}) \oplus_{v^2}^2 R_{\text{stab}}(G^2) \oplus_{v^3}^3 \cdots \oplus_{v^n}^n R_{\text{stab}}(G^n).$$

4.4 The Objective Vector

It remains to construct an objective vector $c^\star \in \mathbb{R}^{V(G^\star)}$ that shows that $R_{\text{stab}}(G^\star)$ does not have the persistency property. Let A, b be as in the previous section, and denote by $a := A_{1,\star}$ the first row of A. We will define $c^\star$ via

$$c_1^\star := \varepsilon \quad \text{and} \quad c_v^\star := a_j \cdot c_v^j \text{ for all } v \in V(G^j),\ j \in \{2, 3, \dots, n\},$$

where $\varepsilon > 0$ is a positive constant that we will define later. Our first claim is independent of the specific choice of ε.

Claim 3. Every $c^\star$-maximal stable set in $G^\star$ contains node $1 \in V(G^{\text{in}})$.

Proof. By Claim 1 there exists, for each $j \in \{2, 3, \dots, n\}$, a c^j-maximal stable set $S^j \subseteq V(G^j)$ that does not use v^j. Thus, the maximum objective value obtained on $V(G^\star \setminus \{1\})$ is $\sum_{j=2}^n a_j c^j(S^j)$, which is equal to the maximum objective value for all stable sets that do not contain node 1. Since $v^j \notin S^j$ for each j, the set $S^\star := \bigcup_{j=2}^n S^j \cup \{1\}$ is a stable set in $G^\star$ with objective value $\varepsilon + \sum_{j=2}^n a_j c^j(S^j) > \sum_{j=2}^n a_j c^j(S^j)$, which proves the claim. □

To see that $R_{\text{stab}}(G^\star)$ does not have the persistency property, it suffices to establish the following claim, which then yields Theorem 1.

Claim 4. For $\varepsilon > 0$ small enough, each $c^\star$-optimal $x^\star \in R_{\text{stab}}(G^\star)$ satisfies $x_1^\star = 0$.

Let $x^\star$ be any $c^\star$-optimal point in $R_{\text{stab}}(G^\star)$. In order to understand the contributions of the variables corresponding to nodes $v \in V(G^j)$ to the total optimal value in terms of $x_{v^j}^\star$, let us introduce the function $f : [0, 1] \to \mathbb{R}$ defined via

$$f(y) := \max \left\{ {c^j}^{\intercal} x \mid x \in R_{\text{stab}}(G^j) \text{ and } x_{v^j} = y \right\}.$$

Note that the definition is independent of j since all (G^j, c^j, v^j) are identical up to indexing. We observe that the restriction of $x^\star$ onto the coordinates corresponding to $V(G^{\text{in}})$ is an optimal solution for

$$\max\left\{c'(x) \mid x \in R_{\text{stab}}(G^{\text{in}})\right\} = \max\left\{c'(x) \mid x \in R_{\text{stab}}^{\text{edge}}(G^{\text{in}}),\ Ax \le b\right\}, \quad (1)$$

where $c'(x) := \varepsilon x_1 + \sum_{j=2}^{n} a_j f(x_j)$. Thus, we see that Claim 4 immediately follows from the following result.

Claim 5. For $\varepsilon > 0$ small enough, each c'-optimal $x \in R_{\text{stab}}(G^{\text{in}})$ satisfies $x_1 = 0$.

We will consider the function $g : [0, \infty] \to \mathbb{R}$ defined via

$$g(z) := \max\Big\{\sum_{j=2}^{n} a_j f(x_j) \mid a^\intercal x \le z,\ x \in R_{\text{stab}}^{\text{edge}}(G^{\text{in}})\Big\}.$$

The intuition behind the proof of Claim 5 is the following: First, note that $c'(x)$ is the sum of εx_1 and the objective function defining g. Function $g(z)$ represents the contribution to the objective value of G^j for $j = 2, \ldots, n$ as a function of the right-hand side of the inequality $a^\intercal x \le z$. We will soon prove that $g(z)$ is strictly increasing on the interval $z \in [0, b_1]$. Since $a_1 > 0$ and x_1 does not contribute to the maximum in the definition of g, the latter is attained only by solutions x with $x_1 = 0$. If we ignore, for a moment, the inequalities $Ax \le b$, this shows that for sufficiently small ε, also every c'-maximal solution x satisfies $x_1 = 0$. The formal steps are as follows.

Claim 6. The functions f and g are concave. Moreover, g is strictly monotonically increasing on $[0, b_1]$.

Proof of Claim 5. Letting

$$\gamma := \min\{x_1 \mid x \text{ vertex of } R_{\text{stab}}(G^\star) \text{ with } x_1 > 0\} \in (0, 1], \text{ and}$$
$$\lambda := \min\{\gamma/(A_{i,1} + \cdots + A_{i,n}) \mid i \in \{1, 2, \ldots, m\}\} \in (0, 1),$$

we claim that every choice of ε with

$$0 < \varepsilon < \lambda(g(b_1) - g(b_1 - a_1\gamma))$$

satisfies the assertion. First, we need to verify that the right-hand side of the inequality above is positive. To this end, note that $a_1 \le b_1$ and hence $0 \le b_1 - a_1\gamma < b_1$. By Claim 6 we have

$$g(b_1 - a_1\gamma) < g(b_1), \quad (2)$$

which yields positivity of the right-hand side.

Next, let ε be as above. For the sake of contradiction, assume that there exists a c'-optimal solution $x^\star \in R_{\text{stab}}(G^{\text{in}})$ with $x_1^\star > 0$. Note that $x^\star$ can be extended to a $c^\star$-optimal solution over $R_{\text{stab}}(G^\star)$, which we may assume to be a vertex of $R_{\text{stab}}(G^\star)$, and hence $x_1^\star \geq \gamma$. Let $\hat{x}^0 \in R_{\text{stab}}(G^{\text{in}})$ be equal to $x^\star$, except for $\hat{x}_1^0 := 0$. Moreover, let $\hat{x}^1 \in \mathbb{R}^{V(G^{\text{in}})}$ be a maximizer of $g(b_1)$, which may not be contained in $R_{\text{stab}}(G^{\text{in}})$. Now consider the vector $\hat{x}^\lambda := (1-\lambda)\hat{x}^0 + \lambda\hat{x}^1$. To obtain the desired contradiction, we will show that $\hat{x}^\lambda$ is contained in $R_{\text{stab}}(G^{\text{in}})$ and that $c'(\hat{x}^\lambda) > c'(x^\star)$.

Since $\hat{x}^0$ and $\hat{x}^1$ both lie in $R^{\text{edge}}_{\text{stab}}(G^{\text{in}})$, also x^λ lies in $R^{\text{edge}}_{\text{stab}}(G^{\text{in}})$. Let $i \in \{1, 2, \ldots, m\}$. By Claim 2, $A_{i,1} \geq 1$ holds, which implies $A_{i,\star}\hat{x}^0 \leq A_{i,\star}x^\star - \gamma \leq b_i - \gamma$. We obtain

$$A_{i,\star}\hat{x}^\lambda = A_{i,\star}\hat{x}^0 + \lambda A_{i,\star}(\hat{x}^1 - \hat{x}^0) \leq b_i - \gamma + \lambda(A_{i,1} + \cdots + A_{i,n}) \leq b_i,$$

where the second inequality follows from the fact that each coordinate of $\hat{x}^1 - \hat{x}^0$ is bounded by 1, and the last inequality holds by the definition of λ. This shows that $\hat{x}^\lambda$ is contained in $R_{\text{stab}}(G^{\text{in}})$.

For the objective value of $\hat{x}^1$ we clearly have $c'(\hat{x}^1) \geq g(b_1)$. Moreover, since $\hat{x}_1^0 = 0$ we have

$$c'(\hat{x}^0) \leq g(a^\intercal \hat{x}^0) \leq g(b_1 - a_1\gamma) < g(b_1),$$

where the latter two inequalities again follow from Claim 6 and (2). Observe that concavity of f and nonnegativity of a imply concavity of $c'(x)$, which yields $c'(\hat{x}^\lambda) \geq (1-\lambda)c'(\hat{x}^0) + \lambda c'(\hat{x}^1)$. We obtain

$$\begin{aligned} c'(x^\star) - c'(\hat{x}^\lambda) &\leq \big(\varepsilon + c'(\hat{x}^0)\big) - \big(c'(\hat{x}^0) - \lambda(c'(\hat{x}^0) - c'(\hat{x}^1))\big) \\ &= \varepsilon + \lambda(c'(\hat{x}^0) - c'(\hat{x}^1)) \leq \varepsilon + \lambda(g(b_1 - a_1\gamma) - g(b_1)) < 0, \end{aligned}$$

where the last inequality holds by definition of ε and due to (2).

□

To conclude the proof of Theorem 1, it remains to prove Claim 6. The fact that f and g are concave is a simple consequence of the next basic lemma.

Lemma 2. *Let $P \subseteq \mathbb{R}^n$ be a non-empty polytope, let $c, a \in \mathbb{R}^n$ and let $\ell := \min\{a^\intercal x \mid x \in P\}$. The functions $h^=, h^\leq : [\ell, \infty) \to \mathbb{R}$ defined via $h^=(\beta) = \max\{c^\intercal x \mid x \in P, a^\intercal x = \beta\}$ and $h^\leq(\beta) = \max\{c^\intercal x \mid x \in P, a^\intercal x \leq \beta\}$ are concave. Moreover, there exists a number $\beta^\star \in [\ell, \infty)$ such that $h^=$ and $h^\leq$ are identical and strictly monotonically increasing on the interval $[\ell, \beta^\star]$, and $h^\leq$ is constant on the interval $[\beta^\star, \infty)$.*

The lemma is proved in Appendix A. The proof of Claim 6 relies on the following result of Sewell.

Proposition 2 (Corollary 3.4.3 in [14]). *Let $\sum_{j=1}^n a_j x_j \leq b_1$ be a facet-defining inequality for the stable-set polytope of a graph on n nodes that is neither a bound nor an edge inequality. Then we have $a_1 \leq \sum_{j=1}^n a_j - 2b_1$.*

Proof of Claim 6. From Lemma 2 it is clear that f is concave. By rewriting

$$g(z) = \max\Big\{ \sum_{j=2}^{n} a_j \cdot \sum_{v \in V(G^j)} c_v^j x_v \mid \sum_{j=1}^{n} a_j x_j \leq z, \\ x \in R_{\text{stab}}^{\text{edge}}(G^{\text{in}}) \oplus_{v^2}^{2} R_{\text{stab}}(G^2) \oplus_{v^3}^{3} \cdots \oplus_{v^n}^{n} R_{\text{stab}}(G^n) \Big\},$$

we also see that g is concave. Moreover, again by Lemma 2, there exists some $\beta^\star \geq 0$ such that g is strictly monotonically increasing on the interval $[0, \beta^\star]$, and constant on $[\beta^\star, \infty)$. It suffices to show that $\beta^\star \geq b_1$. To this end, let us get back to our initial definition of g, and let $\hat{x} \in R_{\text{stab}}^{\text{edge}}(G^{\text{in}})$ be a maximizer for $g(\infty)$. Note that $\beta^\star \geq a^\intercal \hat{x}$ by definition of $\beta^\star$, and hence we have to show that $\hat{x}$ satisfies $a^\intercal \hat{x} \geq b_1$.

Since the objective value of $\hat{x}$ does not depend on $\hat{x}_1$, we may assume that $\hat{x}_1 = 0$. By the construction of G^j and c^j, we know that f attains its unique maximum at $y^\star \geq \frac{1}{2}$. This implies $0 \leq \hat{x}_j \leq y^\star$ for $j = 2, 3, \ldots, n$. Moreover, we claim that also $\hat{x}_j \geq 1 - y^\star$ holds. Suppose not, then none of the edge inequalities involving x_j is tight. Then $\hat{x}_j < 1 - y^\star \leq y^\star$ shows that increasing $\hat{x}_j$ would improve the objective value, which in turn contradicts optimality of $\hat{x}$. Consequently, even $1 - y^\star \leq \hat{x}_j \leq y^\star$ holds for $j = 2, 3, \ldots, n$.

Let $J(\alpha) := \{2 \leq j \leq n \mid \hat{x}_j = \alpha\}$ for $\alpha \in [1 - y^\star, y^\star]$. We will show that $a(J(\alpha)) \geq a(J(1-\alpha))$ holds for all $\alpha \in (1/2, y^\star]$, where $a(J(\alpha))$ shall denote $\sum_{j \in J(\alpha)} a_j$. Note that this implies the claim since for each $\alpha \in (1/2, y^\star]$ we have

$$\begin{aligned} \sum_{j \in J(\alpha)} a_j \hat{x}_j + \sum_{j \in J(1-\alpha)} a_j \hat{x}_j &= \sum_{j \in J(\alpha)} a_j \alpha + \sum_{j \in J(1-\alpha)} a_j (1-\alpha) \\ &= \alpha \cdot \underbrace{[a(J(\alpha)) - a(J(1-\alpha))]}_{\geq 0} + a(J(1-\alpha)) \\ &\geq \tfrac{1}{2} \cdot [a(J(\alpha)) - a(J(1-\alpha))] + a(J(1-\alpha)) \\ &= \sum_{j \in J(\alpha)} a_j \tfrac{1}{2} + \sum_{j \in J(1-\alpha)} a_j \tfrac{1}{2} \end{aligned}$$

and hence

$$\begin{aligned} a^\intercal \hat{x} = \sum_{j=2}^{n} a_j \hat{x}_j &= \sum_{j \in J(1/2)} a_j \hat{x}_j + \sum_{\alpha \in (1/2, y^\star]} \Big(\sum_{j \in J(\alpha)} a_j \hat{x}_j + \sum_{j \in J(1-\alpha)} a_j \hat{x}_j \Big) \\ &\geq \sum_{j \in J(1/2)} a_j \tfrac{1}{2} + \sum_{\alpha \in (1/2, y^\star]} \Big(\sum_{j \in J(\alpha)} a_j \tfrac{1}{2} + \sum_{j \in J(1-\alpha)} a_j \tfrac{1}{2} \Big) \geq b_1, \end{aligned}$$

where the last inequality follows from Proposition 2.

For the sake of contradiction, assume that $a(J(\alpha)) < a(J(1-\alpha))$ holds for some $\alpha \in (1/2, y^\star]$. For a sufficiently small $\varepsilon' > 0$, the solution $\hat{x}' \in \mathbb{R}^{V(G^{\text{in}})}$ defined via

$$\hat{x}'_j := \begin{cases} \hat{x}_j + \varepsilon' & \text{if } j \in J(1-\alpha) \\ \hat{x}_j - \varepsilon' & \text{if } j \in J(\alpha) \\ \hat{x}_j \text{ otherwise} & \end{cases} \qquad \text{for } j = 1, 2, \ldots, n$$

is still contained in $R_{\text{stab}}^{\text{edge}}(G^{\text{in}})$. To see this, observe that $\hat{x}'_j \geq 0$ holds for all $j \in V(G^{\text{in}})$ since we only decrease entries that are at least 1/2. Moreover, edge inequalities that are tight for $\hat{x}$ remain tight for $\hat{x}'$, since either none or both of its two node values are modified, where in the latter case, the value is increased by ε' for one node and decreased by ε' for the other. Finally, edge inequalities that are not tight for $\hat{x}$ will not be violated if we choose ε' sufficiently small. For the objective values we obtain

$$\sum_{j=2}^{n} a_j(f(\hat{x}'_j) - f(\hat{x}_j)) = \sum_{j\in J(1-\alpha)} a_j(f(\hat{x}'_j) - f(\hat{x}_j)) + \sum_{j\in J(\alpha)} a_j(f(\hat{x}'_j) - f(\hat{x}_j))$$
$$= a(J(1-\alpha)) \cdot \big(f(1-\alpha+\varepsilon') - f(1-\alpha)\big) + a(J(\alpha)) \cdot \big(f(\alpha-\varepsilon') - f(\alpha)\big).$$

We also assume that ε' is small enough to guarantee $1-\alpha+\varepsilon' < \alpha-\varepsilon'$. Since f is concave and monotonically increasing in $[0, y^\star]$, we obtain $f(1-\alpha+\varepsilon')-f(1-\alpha) \geq f(\alpha)-f(\alpha-\varepsilon')$. Together with the assumption $a(J(1-\alpha)) > a(J(\alpha))$, this shows that the objective value of $\hat{x}'$ is strictly larger than that of $\hat{x}$, a contradiction to the optimality of $\hat{x}$. □

Acknowledgements. We are grateful to four anonymous reviewers whose comments led to improvements of this manuscript.

A Deferred proofs

Lemma 1. *Let $P, Q \subseteq \mathbb{R}^n$ be polytopes. If there exists a vector $c \in \mathbb{R}^n$ such that* $\dim(\text{opt}(Q,c)) < \dim(\text{opt}(P,c))$*, then there exists a vector $c' \in \mathbb{R}^n$ such that* $\text{opt}(Q,c')$ *is a vertex of Q, while* $\text{opt}(P,c')$ *is not a vertex of P.*

Proof. Let $c' \in \mathbb{R}^n$ be such that $\dim(\text{opt}(Q,c')) < \dim(\text{opt}(P,c'))$ holds, and among those, such that $\dim(\text{opt}(Q,c'))$ is minimum. Clearly, c' is well-defined since $c' := c$ satisfies the conditions.

Assume, for the sake of contradiction, that $\dim(\text{opt}(Q,c')) > 0$. Let $F := \text{opt}(P,c')$ and $G := \text{opt}(Q,c')$. Let $F_1, F_2, \ldots, F_k$ be the facets of F. By $n(F,F_i)$ we denote the set of vectors $w \in \mathbb{R}^n$ such that $\text{opt}(F,w) \supseteq F_i$. Since F is a polytope, $\bigcup_{i\in\{1,2,\ldots,k\}} n(F,F_i)$ contains a basis U of $\mathbb{R}^n$. Moreover, not all vectors $u \in U$ can lie in $\text{aff}(G)^\perp$, the orthogonal complement of $\text{aff}(G)$, since then $\text{aff}(G)^\perp = \mathbb{R}^n$ would hold, contradicting $\dim(G) > 0$. Let $u \in U \setminus \text{aff}(G)^\perp$.

Now, for a sufficiently small $\varepsilon > 0$, $\text{opt}(P, c'+\varepsilon u) \supseteq F_i$ for some $i \in \{1,2,\ldots,k\}$, and $\text{opt}(Q,c'+\varepsilon u)$ is a proper face of G. Thus, $c'+\varepsilon u$ satisfies the requirements at the beginning of the proof. However, $\dim(\text{opt}(Q,c'+\varepsilon u)) < \dim(G)$ contradicts the minimality assumption, which concludes the proof. □

Lemma 2. *Let $P \subseteq \mathbb{R}^n$ be a non-empty polytope, let $c, a \in \mathbb{R}^n$ and let* $\ell := \min\{a^\intercal x \mid x \in P\}$*. The functions $h^=, h^\leq : [\ell,\infty) \to \mathbb{R}$ defined via* $h^=(\beta) = \max\{c^\intercal x \mid x \in P, a^\intercal x = \beta\}$ *and* $h^\leq(\beta) = \max\{c^\intercal x \mid x \in P, a^\intercal x \leq \beta\}$ *are concave. Moreover, there exists a number $\beta^\star \in [\ell,\infty)$ such that $h^=$ and $h^\leq$ are identical and strictly monotonically increasing on the interval $[\ell,\beta^\star]$, and $h^\leq$ is constant on the interval $[\beta^\star,\infty)$.*

Proof. Let $Q := \{\binom{y_1}{y_2} \mid \exists x \in P : a^\intercal x = y_1,\ c^\intercal x = y_2\} \subseteq \mathbb{R}^2$ be the projection of P along a and c. By construction, $h^{\leq}(\beta) = \max\{y_2 \mid y \in Q,\ y_1 \leq \beta\}$ holds. Considering that Q is a polytope of dimension at most 2, the claimed properties of $h^{\leq}$ and $h^{=}$ are obvious (see Fig. 1). □

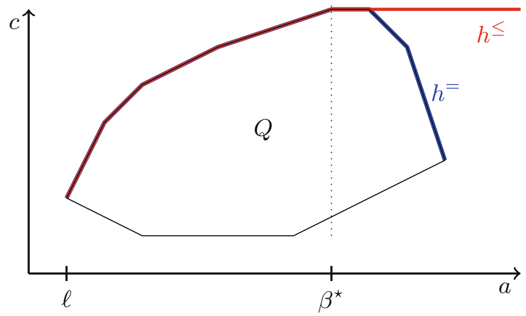

Fig. 1. Illustration of Lemma 2. The graph of $h^{\leq}$ is highlighted in red, while that of $h^{=}$ is highlighted in blue. (Color figure online)

References

1. Boros, E., Hammer, P.L., Sun, R., Tavares, G.: A max-flow approach to improved lower bounds for quadratic unconstrained binary optimization (QUBO). Disc. Optim. **5**(2), 501–529 (2008). In Memory of George B. Dantzig
2. Chvátal, V.: On certain polytopes associated with graphs. J. Combin. Theory, Ser. B **18**(2), 138–154 (1975)
3. Fix, A., Gruber, A., Boros, E., Zabih, R.: A hypergraph-based reduction for higher-order binary Markov random fields. IEEE Trans. Pattern Anal. Mach. Intell. **37**(7), 1387–1395 (2015)
4. Hammer, P.L., Hansen, P., Simeone, B.: Roof duality, complementation and persistency in quadratic 0–1 optimization. Math. Program. **28**(2), 121–155 (1984)
5. Ishikawa, H.: Transformation of general binary MRF minimization to the first-order case. IEEE Trans. Pattern Anal. Mach. Intell. **33**(6), 1234–1249 (2011)
6. Karp, R.M., Papadimitriou, C.H.: On linear characterizations of combinatorial optimization problems. SIAM J. Comput. **11**(4), 620–632 (1982)
7. Kolmogorov, V., Rother, C.: Minimizing nonsubmodular functions with graph cuts - a review. IEEE Trans. Pattern Anal. Mach. Intell. **29**(7), 1274–1279 (2007)
8. Nemhauser, G.L., Trotter, L.E.: Vertex packings: structural properties and algorithms. Math. Programm. **8**(1), 232–248 (1975)
9. Oriolo, G.: Clique family inequalities for the stable set polytope of quasi-line graphs. Disc. Appl. Math. **132**(1), 185–201 (2003). Stability in Graphs and Related Topics
10. Padberg, M.W.: On the facial structure of set packing polyhedra. Math. Program. **5**(1), 199–215 (1973)
11. Pulleyblank, W.R.: Minimum node covers and 2-bicritical graphs. Math. Program. **17**(1), 91–103 (1979)

12. Rodríguez-Heck, E., Stickler, K., Walter, M., Weltge, S.: Persistency of linear programming formulations for the stable set problem (2019). arXiv:1911.01478
13. Schrijver, A.: Theory of Linear and Integer Programming. John Wiley, New York (1986)
14. Sewell, E.C.: Stability critical graphs and the stable set polytope. Technical report, Cornell University Operations Research and Industrial Engineering (1990)
15. Trotter, L.E.: A class of facet producing graphs for vertex packing polyhedra. Disc. Math. **12**(4), 373–388 (1975)