

# A new lift-and-project operator

Merve Bodur <sup>\*a</sup>, Sanjeeb Dash<sup>b</sup>, and Oktay Günlük<sup>b</sup>

<sup>a</sup>*School of Industrial and Systems Engineering, Georgia Institute of Technology, Atlanta, GA, USA  
30332*

<sup>b</sup>*IBM T. J. Watson Research Center, Yorktown Heights, NY, USA 10598*

March 3, 2016

## Abstract

In this paper, we analyze the strength of split cuts in a lift-and-project framework. We first observe that the Lovász-Schrijver and Sherali-Adams lift-and-project operator hierarchies can be viewed as applying specific 0-1 split cuts to an appropriate extended formulation and demonstrate how to strengthen these hierarchies using additional split cuts. More precisely, we define a new operator that adds all 0-1 split cuts to the extended formulation. For 0-1 mixed-integer sets with  $k$  binary variables, this new operator is guaranteed to obtain the integer hull in  $\lceil k/2 \rceil$  steps compared to  $k$  steps for the Lovász-Schrijver or the Sherali-Adams operator. We also present computational results on the stable set problem with our new operator.

*Keywords:* Integer programming, lift-and-project, extended formulations, cutting planes

## 1 Introduction

For a given 0-1 mixed integer set  $P^{IP}$ , defined as the intersection of a polyhedron  $P$  and  $\{0, 1\}^{n_1} \times \mathbb{R}^{n_2}$ , a fundamental goal in integer programming is to obtain a better approximation of its convex hull than  $P$ . A set is called a strong relaxation of the set  $P^{IP}$  if it contains  $P^{IP}$  and at the same time is strictly contained in  $P$ . Starting with the pioneering work of Gomory [17] on cutting planes, different techniques have been developed to build strong relaxations. Lift-and-project techniques such as the ones developed by Sherali and Adams [23], Lovász and Schrijver [20], Balas, Ceria and Cornuéjols [4] and Lasserre [18], obtain strong relaxations by first formulating a set in a higher dimensional space and then projecting this set onto the space of the original variables. In this paper we describe a new lift-and-project operator that is closely related to the Sherali-Adams and Lovász-Schrijver (without semidefiniteness) operators. Similar to their operators, our operator also produces polyhedral relaxations in the original space. Both

---

<sup>\*</sup>Corresponding author

*Email addresses:* `merve.bodur@gatech.edu` (Merve Bodur), `sanjeebd@us.ibm.com` (Sanjeeb Dash), `gunluk@us.ibm.com` (Oktay Günlük)

of these operators yield a hierarchy of relaxations  $H^1, H^2, \dots, H^{n_1}$  of  $P^{IP}$  with the following property:

$$P \supseteq H^1 \supseteq H^2 \supseteq \dots \supseteq H^{n_1} = \text{conv}(P^{IP}).$$

Therefore, these operators obtain the convex hull of  $P^{IP}$  in at most  $n_1$  steps and there are examples where  $n_1$  steps are necessary [11, 19]. See [10, 19] for a review and comparison of these hierarchies. The new operator we describe in this paper is guaranteed to obtain the integer hull in  $\lceil n_1/2 \rceil$  steps instead.

In an earlier paper [6], for every 0-1 mixed-integer set  $P^{IP}$  with  $n_1$  integer and  $n_2$  continuous variables we showed how to construct an extended formulation of  $P$  (which we call an *extended LP formulation* of  $P^{IP}$ ) with  $n_1 - 1$  additional continuous variables whose 0-1 split closure is integral. It is well-known that the 0-1 split closure of a 0-1 mixed-integer set can be computed in time bounded by a polynomial function of the encoding size of  $P$  – i.e., the number of bits required to represent the inequalities defining  $P$  (and we describe this computation more precisely later). The extended LP formulation presented in [6] is only of theoretical interest, as it requires the list of all extreme points and rays of  $P$ , which could be of exponential size. The new operator we describe in this paper can be viewed as the 0-1 split closure of an extended LP formulation that is of polynomial size, and it is therefore possible to optimize over its 0-1 split closure in polynomial time. The extended LP formulation we use is implicitly constructed by the Lovász-Schrijver and Sherali-Adams operators, and we show that the “strengthening step” of these operators can be interpreted as adding certain 0-1 split cuts to this extended LP formulation. Thus our new operator is stronger than the Lovász-Schrijver operator, and we give upper and lower bounds on its strength: we show that our operator can be stronger than the first-level of the Sherali-Adams hierarchy (which equals the Lovász-Schrijver operator), but is weaker than the second-level of the hierarchy.

In the next section we present some background on split cuts and extended formulations. In Section 3, we describe the Lovász-Schrijver and Sherali-Adams lift-and-project operators, and the extended LP formulation implicitly constructed by these operators. We define our new lift-and-project operator in terms of the 0-1 split closure of this extended LP formulation, and show that one can optimize over the resulting set of points in polynomial time, as in the case of the Lovász-Schrijver operator. We also show that the second-level of the Sherali-Adams hierarchy is stronger than this new operator. Finally in Section 4, we apply our operator to the stable set polytope and perform numerical experiments to compare its computational performance with that of the Lovász-Schrijver and Sherali-Adams operators.

## 2 Preliminaries

We use  $\mathbb{R}^n$  and  $\mathbb{Z}^n$ , respectively, for the set of  $n$ -dimensional real and integer vectors. Throughout the paper, we work with 0-1 mixed-integer sets of the form  $P^{IP} = P \cap (\mathbb{Z}^{n_1} \times \mathbb{R}^{n_2})$  where

$$P = \{x \in \mathbb{R}^n : Ax \leq b\},$$

$n = n_1 + n_2$  and  $n_1 > 0$ , and the inequality system  $Ax \leq b$  contains the inequalities  $0 \leq x_i \leq 1$  for all  $i = 1, \dots, n_1$ . We refer to  $P$  as the LP relaxation of  $P^{IP}$ .

## 2.1 Extended LP formulations

Let  $Q = \{(x, y) \in \mathbb{R}^n \times \mathbb{R}^q : Cx + Dy \leq g\}$  be a polyhedron, and let the number of inequalities in  $Cx + Dy \leq g$  be  $t$ .  $Q$  is called an *extended formulation* of  $P$  if

$$P = \text{proj}_x(Q),$$

where  $\text{proj}_x(Q)$  stands for the orthogonal projection of  $Q$  onto the space of  $x$  variables. More precisely,

$$\text{proj}_x(Q) = \{x \in \mathbb{R}^n : \exists y \in \mathbb{R}^q \text{ s.t. } (x, y) \in Q\}.$$

Alternatively,  $\text{proj}_x(Q) = \{x \in \mathbb{R}^n : u^T Cx \leq u^T g, \forall u \in U\}$ , where  $U = \{u \in \mathbb{R}^t : u^T D = 0, u \geq 0\}$  is the *projection cone* of  $Q$ . This immediately implies that the projection of a polyhedron is a polyhedron. For more properties of projection, we refer the reader to [3]. Throughout the paper we call  $Q$  an *extended LP formulation* of  $P^{IP}$  as it is an extended formulation of its LP relaxation.

## 2.2 0-1 Split cuts

For a given  $i \in I = \{1, \dots, n_1\}$ , consider the *0-1 split set*

$$S_i = \{x \in \mathbb{R}^n : 0 < x_i < 1\},$$

and note that  $S_i$  has an empty intersection with  $\mathbb{Z}^{n_1} \times \mathbb{R}^{n_2}$ . The inequality  $c^T x \geq d$ , where  $c \in \mathbb{R}^n$  and  $d \in \mathbb{R}$ , is called a *0-1 split cut* for  $P$  generated by  $S_i$  if it is valid for  $\text{conv}(P \setminus S_i)$ . Here we use  $\text{conv}(\cdot)$  to denote the convex hull operator. In other words,  $c^T x \geq d$  is a 0-1 split cut if it is valid for the *disjunction*  $P \cap \{x \in \mathbb{R}^n : x_i = 0\}$  or  $P \cap \{x \in \mathbb{R}^n : x_i = 1\}$ . A split cut is called *nondominated* if it is not implied by a collection of other split cuts. We note that multiple nondominated split cuts may be generated by the same split set  $S_i$ . We denote the *split closure of  $P$  with respect to 0-1 split cuts* (or the *0-1 split closure* for short) as

$$\mathcal{S}(P) = \bigcap_{i \in I} \text{conv}(P \setminus S_i).$$

Clearly,  $P^{IP} \subseteq \mathcal{S}(P) \subseteq P$ .  $\mathcal{S}(P)$  is a relaxation of the general split closure defined in [12] where more general two-term disjunctions are used to generate cuts.

Furthermore, as  $\mathcal{S}(P)$  is a polyhedron, it is possible to repeat this operation. For any given integer  $k \geq 0$ , the  $k^{\text{th}}$  *0-1 split closure* of  $P$ , denoted as  $\mathcal{S}^k(P)$ , is defined iteratively as follows:  $\mathcal{S}^0(P) = P$  and  $\mathcal{S}^k(P) = \mathcal{S}(\mathcal{S}^{k-1}(P))$  for  $k \geq 1$ . Balas [4] proved that it is sufficient to repeat this operation  $n_1$  times to obtain the convex hull of  $P^{IP}$ , establishing that  $\mathcal{S}^{n_1}(P) = \text{conv}(P^{IP})$ .

We also note that  $\mathcal{S}(P)$  can be explicitly written as the projection of an extended formulation

using Balas' result on convex hulls of unions of polyhedra [2, Thm 3.3] as follows:

$$\begin{aligned} \mathcal{S}(P) = \Big\{ x \in \mathbb{R}^n \quad : \quad & \exists \bar{x}^i, \bar{\bar{x}}^i \in \mathbb{R}^n, \quad \bar{\lambda}^i, \bar{\bar{\lambda}}^i \in \mathbb{R}_+ \quad \text{s.t.} \\ & x = \bar{x}^i + \bar{\bar{x}}^i, \quad \bar{\lambda}^i + \bar{\bar{\lambda}}^i = 1, \quad i \in I, \\ & A\bar{x}^i \leq \bar{\lambda}^i b, \quad A\bar{\bar{x}}^i \leq \bar{\bar{\lambda}}^i b, \quad i \in I, \\ & \bar{x}_i^i = 1, \quad \bar{\bar{x}}_i^i = 0 \quad i \in I \Big\}, \end{aligned} \tag{1}$$

where  $\mathbb{R}_+$  denotes the set of nonnegative real numbers. Note that formulation (1) is of polynomial size (in the encoding size of  $P$ ) and therefore one can optimize a linear function over it in polynomial time. However, describing the projected set  $\mathcal{S}(P)$  by an explicit list of linear inequalities in  $\mathbb{R}^n$  may require an exponential number of inequalities. An important point we wish to emphasize is that in contrast to  $\mathcal{S}(P)$ , it is NP-hard to optimize over the general split closure of  $P$ .

Given an extended formulation  $Q \subseteq \mathbb{R}^{n+q}$  of  $P$ , the 0-1 split closure of  $Q$  is defined as

$$\mathcal{S}(Q) = \bigcap_{i \in I} \text{conv}(Q \setminus S_i^+),$$

where  $S_i^+ = S_i \times \mathbb{R}^q$ . Clearly  $\mathcal{S}(Q)$  can also be explicitly written via an extended formulation similar to (1). Therefore, optimization over  $\mathcal{S}(Q)$  can also be done in polynomial time provided that  $Q$  is of polynomial size.

### 2.3 Strengthening extended LP formulations with 0-1 split cuts

In terms of optimizing a linear function over the mixed-integer set  $P^{IP}$ , the extended LP formulation  $Q$  does not lead to better bounds than the original LP relaxation  $P$  as

$$\max\{c^T x : x \in P\} = \max\{c^T x : (x, y) \in Q\}$$

for any  $c \in \mathbb{R}^n$ . However, after the addition of split cuts, extended LP formulations might yield better bounds.

In [7], we show that for any split set, and in particular for any 0-1 split set  $S_i$ ,

$$\text{proj}_x(\text{conv}(Q \setminus S_i^+)) = \text{conv}(P \setminus S_i).$$

However,

$$\text{proj}_x(\mathcal{S}(Q)) \subseteq \mathcal{S}(P),$$

and the inclusion above is strict in some cases. Therefore adding split cuts to an extended LP formulation can lead to strictly better relaxations than adding split cuts to the original LP relaxation. This, however, can happen only if split cuts generated by multiple split sets are used simultaneously.

Later, in [6], we show that there exists an extended LP formulation  $Q^* \subseteq \mathbb{R}^{n+n_1-1}$  of the

set  $P^{IP}$  such that the 0-1 split closure of  $Q^*$  is integral, that is,

$$\text{proj}_x(\mathcal{S}(Q^*)) = \text{conv}(P^{IP}).$$

Even though the proof of this result is constructive, it requires an inner description of the set  $P$  (i.e., all its extreme points and rays) which is usually difficult to compute from the inequality description of  $P$ , i.e., from  $Ax \leq b$ . Moreover, the number of extreme points and rays of  $P$  may be exponential in the encoding size of  $P$ . Consequently, the approach described in [6] is not practical. In the next section we consider extended LP formulations that can be obtained in polynomial time from  $Ax \leq b$ .

### 3 A new lift-and-project operator

Lovász and Schrijver defined a lift-and-project operator  $N(P)$  which has the property that  $N(P)$  is a polyhedron in  $\mathbb{R}^n$  and  $N(P) \subseteq \mathcal{S}(P)$  [20, Lemma 1.3]. In other words, this operator yields a relaxation of  $P^{IP}$  that is at least as strong as that the one obtained by adding all 0-1 split cuts to  $P$ . Combining this with the fact that  $\mathcal{S}^{n_1}(P) = \text{conv}(P^{IP})$  one obtains

$$\text{conv}(P^{IP}) = N^{n_1}(P) \subseteq \dots \subseteq N^2(P) \subseteq N^1(P) \subseteq P,$$

where  $N^1(P) = N(P)$  and  $N^k(P) = N(N^{k-1}(P))$  for  $k > 1$ .

Sherali and Adams [23] also defined a hierarchy of operators (not obtained iteratively) where the so-called level 1 operator  $SA^1(P)$  is the same as  $N(P)$ , and  $SA^k(P) \subseteq N^k(P)$  for all  $k \geq 1$  [20]. Consequently this operator also gives the integer hull in  $n_1$  steps:

$$\text{conv}(P^{IP}) = SA^{n_1}(P) \subseteq \dots \subseteq SA^2(P) \subseteq SA^1(P) \subseteq P.$$

Each operator in this hierarchy also yields a polyhedron in  $\mathbb{R}^n$ . We describe  $N(P)$  and  $SA^k(P)$  more precisely later in this section.

Both of these operators are called lift-and-project operators because they are based on the following scheme:

- 1a. Create a polyhedron in an extended variable space (*Lifting*),
- 1b. Perform a strengthening in the extended variable space (*Strengthening*),
2. Project the formulation in the extended space onto the original space (*Projection*).

In this section, we will first show that the procedure above fits into the framework described in Section 2.3 as Step 1a above can be viewed as constructing an extended formulation of  $P$  and Step 1b can be viewed as adding some specific 0-1 split cuts to this extended formulation. Based on this observation, we will later define a new lift-and-project operator and compare it with the earlier ones. We also note that the extended formulation constructed by the earlier

operators in Step 1a does not require an explicit knowledge of the vertices and rays of  $P$  and therefore is easier to describe than  $Q^*$  mentioned in Section 2.3.

### 3.1 Analyzing the Lovász-Schrijver operator

We next describe the Lovász-Schrijver operator in detail. Following the notation in the original paper [20], we refer to the set obtained in the Lifting Step 1a as  $\hat{M}(P)$ , the set obtained after the Strengthening Step 1b as  $M(P)$ , and the set obtained after the Projection Step 1a as  $N(P)$ . The description of the operator in [20] is for pure binary sets only and our description below follows the more general one in [4] for 0-1 mixed-integer sets. Furthermore, we do not assume that the continuous variables are nonnegative.

Let  $(a^\ell)^T x \leq b^\ell$ ,  $\ell \in L = \{1, \dots, m\}$ , denote the linear inequalities in the linear system  $Ax \leq b$  defining  $P$ . Also remember that the variables  $x_i$  for  $i \in I = \{1, \dots, n_1\}$  are integral in  $P^{IP}$  and inequalities  $0 \leq x_i \leq 1$ ,  $i \in I$  are included in the system  $Ax \leq b$ . We let  $J = \{1, \dots, n\}$ . The set  $\hat{M}(P)$  is obtained by first multiplying the inequalities defining  $P$  by  $x_i$  and  $(1 - x_i)$  for all  $i \in I$ . This yields the quadratic inequalities

$$x_i(b^\ell - (a^\ell)^T x) \geq 0, \quad i \in I, \ell \in L, \quad (2a)$$

$$(1 - x_i)(b^\ell - (a^\ell)^T x) \geq 0, \quad i \in I, \ell \in L, \quad (2b)$$

which are valid for  $P$  as  $x_i \geq 0$  and  $1 - x_i \geq 0$  for all  $x \in P$ . These inequalities are then linearized by replacing each quadratic term  $x_i x_j$  by a new variable  $y_{ij}$  for all  $i \in I$  and  $j \in J$ , and ensuring that the variables  $y_{ij}$  and  $y_{ji}$  for  $i, j \in I$  are identical. (In practice, only one of the two variables  $y_{ij}$  and  $y_{ji}$  needs to be introduced.) Therefore, after linearizing inequalities (2a) and (2b) one obtains:

$$b^\ell x_i - \sum_{j \in J} a_j^\ell y_{ij} \geq 0, \quad i \in I, \ell \in L, \quad (3a)$$

$$b^\ell - \sum_{j \in J} a_j^\ell x_j - b^\ell x_i + \sum_{j \in J} a_j^\ell y_{ij} \geq 0, \quad i \in I, \ell \in L, \quad (3b)$$

$$y_{ij} = y_{ji}, \quad i < j, \quad i, j \in I. \quad (3c)$$

The set  $\hat{M}(P) \subset \mathbb{R}^{n+n_1 \cdot n}$  is defined to be the set of points that satisfy inequalities (3a) - (3c). We next show that this set, obtained after the Lifting Step 1a, is indeed an extended formulation of  $P$ .

**Lemma 3.1.** *The set  $\hat{M}(P)$  is an extended LP formulation of  $P^{IP}$ .*

*Proof.* Note that for any fixed  $i \in I$  and  $\ell \in L$ , adding the inequalities in (3a) and (3b), we get the inequality  $b^\ell - (a^\ell)^T x \geq 0$ . Consequently, each inequality in  $Ax \leq b$  is valid for  $\hat{M}(P)$  and therefore  $\text{proj}_x(\hat{M}(P)) \subseteq P$ . Conversely, let  $\bar{x} \in P$  and define  $\bar{y} = (\bar{y}_{ij} : i \in I, j \in J)$  where  $\bar{y}_{ij} = \bar{x}_i \bar{x}_j$ . Then the value of the left-hand-side of equation (3a) equals the left-hand-side of (2a), i.e., it equals  $\bar{x}_i(b^\ell - (a^\ell)^T \bar{x})$  which is clearly nonnegative as  $\bar{x}$  satisfies  $\bar{x} \geq 0$  and  $A\bar{x} \leq b$ .

Similarly, one can argue that  $(\bar{x}, \bar{y})$  satisfies equations (3b)-(3c). Therefore  $(\bar{x}, \bar{y}) \in \hat{M}(P)$  and  $\text{proj}_x(\hat{M}(P)) = P$ .  $\square$

We also note that in (2a) and (2b), if any implied linear inequality for  $P$  is used in place of  $b^\ell - (a^\ell)^T x$ , then the resulting inequality, after linearization, is implied by (3a)-(3c) and therefore is valid for  $\hat{M}(P)$ . In other words,  $\hat{M}(P)$  does not depend on the specific inequality description of  $P$  but only on the set of points in  $P$ .

Lovász and Schrijver defined the set  $M(P)$ , obtained after the Strengthening Step 1b, as follows:

$$M(P) = \hat{M}(P) \cap \{(x, y) : y_{ii} = x_i, \forall i \in I\}. \quad (4)$$

Finally, the set  $N(P) = \text{proj}_x(M(P))$  gives the strengthening of  $P$  obtained after the Projection Step 2. Notice that every 0-1 solution in  $P$  is contained in  $N(P)$  as any vector  $\bar{x}$  in  $\{0, 1\}^{n_1} \times \mathbb{R}^{n_2}$  satisfies  $\bar{x}_i^2 = \bar{x}_i$  for  $i \in I$  and we can define  $\bar{y} \in \mathbb{R}^{n_1 \times n}$  by setting  $\bar{y}_{ij} = \bar{x}_i \bar{x}_j$  for  $i \in I$  and  $j \in J$  to obtain a point  $(\bar{x}, \bar{y}) \in M(P)$ .

We start with a few observations about  $\hat{M}(P)$ . Recall that the inequality system  $Ax \leq b$  is valid for  $\hat{M}(P)$ . We give below a few of the quadratic inequalities in (2a)-(2b) and the corresponding linear inequalities (3a) and (3b) which are valid for the set  $\hat{M}(P)$  for  $i, j \in I$ ,  $i \leq j$ , also see [10]:

$$\begin{aligned} x_i x_j &\geq 0 \longrightarrow y_{ij} \geq 0, \\ x_i(1 - x_j) &\geq 0 \longrightarrow x_i \geq y_{ij}, \\ (1 - x_i)x_j &\geq 0 \longrightarrow x_j \geq y_{ij}, \\ (1 - x_i)(1 - x_j) &\geq 0 \longrightarrow 1 - x_i - x_j + y_{ij} \geq 0. \end{aligned}$$

The four nonlinear inequalities above are obtained by multiplying the simple bound constraints on  $x_j$  variables with  $x_i$  and  $(1 - x_i)$ . The corresponding linear inequalities in the  $x_i, x_j$  and  $y_{ij}$  variables are called the *McCormick inequalities* and they give the convex hull of the set  $\{(x_i, x_j, y_{ij}) \in \{0, 1\}^3 : y_{ij} = x_i x_j\}$  [21].

We next show that the inequalities added to  $\hat{M}(P)$  to obtain  $M(P)$  are just 0-1 split cuts for  $P$ .

**Lemma 3.2.** *For any  $i \in I$  the equation  $y_{ii} = x_i$  is valid for  $\mathcal{S}(\hat{M}(P))$ .*

*Proof.* Let  $i, j \in I$ , and  $i \leq j$ . The McCormick constraints imply the following relations:

$$x_i = 0 \Rightarrow y_{ij} = 0, \quad (5a)$$

$$x_i = 1 \Rightarrow y_{ij} = x_j. \quad (5b)$$

Consequently, if  $x_i = 0$  or 1, then  $y_{ii} = x_i$ . Therefore, for any  $(x, y) \in \text{conv}(\hat{M}(P) \setminus S_i)$ , we have  $y_{ii} = x_i$ . As  $\mathcal{S}(\hat{M}(P)) \subseteq \text{conv}(\hat{M}(P) \setminus S_i)$ , the result follows.  $\square$

### 3.2 Strengthening the Lovász-Schrijver operator

Lemma 3.2 shows that Lovász and Schrijver obtain  $M(P)$  from  $\hat{M}(P)$  by adding exactly one split cut per 0-1 split set  $S_i$  for  $i \in I$ . This immediately suggests strengthening  $\hat{M}(P)$  by adding all 0-1 split cuts. Accordingly, we define a new operator

$$\tilde{N}(P) = \text{proj}_x \left( \mathcal{S}(\hat{M}(P)) \right).$$

Note that just as in the case of  $N(P)$ , it is also possible to optimize over  $\tilde{N}(P)$  in polynomial time using a formulation similar to (1). Lemma 3.2 implies that  $\mathcal{S}(\hat{M}(P))$  is contained in  $M(P)$ , and therefore  $\tilde{N}(P)$  is contained in  $N(P)$  which, in turn, is contained in  $\mathcal{S}(P)$ . We next present our main result, which shows that the strengthened extended formulation obtained by our operator applied to  $P$  is stronger than the one obtained by the Lovász-Schrijver operator applied to the 0-1 split closure of  $P$ .

**Theorem 3.3.**  $\mathcal{S}(\hat{M}(P)) \subseteq M(\mathcal{S}(P))$ .

*Proof.* Let  $\mathcal{S}(P) = \{x \in \mathbb{R}^n : (c^k)^T x \leq d^k, k \in K\}$ . As  $\mathcal{S}(P) = \bigcap_{i \in I} \text{conv}(P \setminus S_i)$ , we can get an inequality description of  $\mathcal{S}(P)$  by combining inequalities from nonredundant descriptions of  $\text{conv}(P \setminus S_i)$  for  $i \in I$ . Therefore, without loss of generality, for each  $k \in K$  we can assume that there exists an index  $t \in I$  such that  $(c^k)^T x \leq d^k$  is valid for  $\text{conv}(P \setminus S_t)$ . Also, for the sake of convenience, assume that the inequalities  $0 \leq x_i \leq 1$ ,  $i \in I$  are included in the system of linear inequalities  $(c^k)^T x \leq d^k$ ,  $k \in K$ ; if any of these are redundant, then  $\mathcal{S}(P)$  does not change and nor does  $\hat{M}(\mathcal{S}(P))$ .

Then  $\hat{M}(\mathcal{S}(P))$  is defined by the following inequalities:

$$d^k x_i - \sum_{j \in J} c_j^k y_{ij} \geq 0, \quad i \in I, k \in K, \quad (6a)$$

$$d^k - \sum_{j \in J} c_j^k x_j - d^k x_i + \sum_{j \in J} c_j^k y_{ij} \geq 0, \quad i \in I, k \in K, \quad (6b)$$

$$y_{ij} = y_{ji}, \quad i < j, i, j \in I. \quad (6c)$$

We will show that every inequality above is valid for  $\mathcal{S}(\hat{M}(P))$ .

Let  $k \in K$  and  $t \in I$  such that  $(c^k)^T x \leq d^k$  is valid for  $\text{conv}(P \setminus S_t)$ . Therefore  $(c^k)^T x \leq d^k$  is valid for  $P \cap \{x : x_t = 0\}$ , and we have

$$(c^k)^T x = \lambda^T A x + \alpha x_t, \quad d^k \geq \lambda^T b,$$

for some  $\lambda \in \mathbb{R}_+^{|L|}$  and  $\alpha \in \mathbb{R}$ . This implies that  $-\alpha x_t + (c^k)^T x \leq d^k$  is valid for  $P$ . Similarly, as  $(c^k)^T x \leq d^k$  is valid for  $P \cap \{x : x_t = 1\}$ , there exists  $\beta \in \mathbb{R}$  such that  $-\beta(1 - x_t) + (c^k)^T x \leq d^k$  is valid for  $P$ . Then, if we multiply these two valid inequalities for  $P$  by  $x_i$  and  $(1 - x_i)$  for



$i \in I$ , and then linearize the product terms, we see that,

$$d^k x_i - \sum_{j \in J} c_j^k y_{ij} + \alpha y_{it} \geq 0, \quad (7a)$$

$$d^k x_i - \sum_{j \in J} c_j^k y_{ij} + \beta x_i - \beta y_{it} \geq 0, \quad (7b)$$

$$d^k - \sum_{j \in J} c_j^k x_j - d^k x_i + \sum_{j \in J} c_j^k y_{ij} + \alpha(x_t - y_{it}) \geq 0, \quad (7c)$$

$$d^k - \sum_{j \in J} c_j^k x_j - d^k x_i + \sum_{j \in J} c_j^k y_{ij} + \beta(1 - x_t - x_i + y_{it}) \geq 0, \quad (7d)$$

are valid for  $\hat{M}(P)$ , by definition.

Now, let  $(x, y) \in \hat{M}(P)$ . If  $x_t = 0$ , then (5a) implies that  $y_{it} = 0$  and (7a) implies that (6a) is valid for

$$\hat{M}(P) \cap \{(x, y) : x_t = 0\}.$$

If  $x_t = 1$ , then (5b) implies that  $y_{it} = x_i$  and (7b) implies that (6a) is valid for

$$\hat{M}(P) \cap \{(x, y) : x_t = 1\}.$$

Therefore, (6a) is valid for  $\hat{M}(P) \setminus S_t$  and is therefore valid for  $\mathcal{S}(\hat{M}(P))$ . Similarly, (7c) and (7d) imply that (6b) is valid for  $\mathcal{S}(\hat{M}(P))$ . Finally, (6c) is valid for  $\hat{M}(P)$  and therefore for  $\mathcal{S}(\hat{M}(P))$ .

So far we showed that  $\mathcal{S}(\hat{M}(P)) \subseteq \hat{M}(\mathcal{S}(P))$ . To conclude the proof, note that Lemma 3.2 shows that  $\mathcal{S}(\hat{M}(P)) \subseteq \{(x, y) : y_{ii} = x_i, \forall i \in I\}$ , and therefore

$$\mathcal{S}(\hat{M}(P)) \subseteq \hat{M}(\mathcal{S}(P)) \cap \{(x, y) : y_{ii} = x_i, \forall i \in I\},$$

where the second term is the same as  $M(\mathcal{S}(P))$ . □

After projecting the sets  $\mathcal{S}(\hat{M}(P))$  and  $M(\mathcal{S}(P))$  onto the space of  $x$  variables, Theorem 3.3 implies that applying the new operator to  $P$  yields a stronger relaxation than applying the Lovász-Schrijver operator to the 0-1 split closure of  $P$ .

**Corollary 3.4.**  $\tilde{N}(P) \subseteq N(\mathcal{S}(P))$ .

Furthermore, as  $N(\mathcal{S}(P)) \subseteq \mathcal{S}(\mathcal{S}(P))$ , we observe that the new operator is guaranteed to obtain the convex hull of  $P^{IP}$  in  $\lceil n_1/2 \rceil$  steps as opposed to  $n_1$  steps required for the Lovász-Schrijver operator.

**Corollary 3.5.**  $(\tilde{N})^{\lceil n_1/2 \rceil}(P) = \text{conv}(P^{IP})$ .

### 3.3 Analyzing the level- $k$ Sherali-Adams operator

We next define the Sherali-Adams operator  $SA^k(P)$  for  $P \subseteq \mathbb{R}^n$  where  $k$  is an integer between 1 and  $n_1$ . This operator is a special case of the *Reformulation-Linearization Technique* (RLT)

of Sherali and Adams which is applicable to more general optimization programs, see [24]. As before, let  $P$  be defined by the set of linear inequalities  $b^\ell - (a^\ell)^T x \geq 0$  for  $\ell \in L$  that include the bound inequalities  $x_i \geq 0$  and  $1 - x_i \geq 0$  for all  $i = 1, \dots, n_1$ . The Sherali-Adams procedure starts with constructing a system of polynomial inequalities by multiplying each inequality defining  $P$  with a product of the form

$$F_k(J_1, J_2) = \prod_{i \in J_1} x_i \prod_{j \in J_2} (1 - x_j),$$

where  $J_1$  and  $J_2$  are disjoint subsets of  $I = \{1, \dots, n_1\}$  and  $|J_1 \cup J_2| = k$ . Sherali and Adams call  $(J_1, J_2)$  a *pair of order  $k$*  if  $J_1, J_2 \subseteq I$ ,  $J_1 \cap J_2 = \emptyset$ , and  $|J_1 \cup J_2| = k$ . When  $k = 0$ ,  $F_0(\emptyset, \emptyset)$  is defined to be 1. This leads to the following system of inequalities that we call  $PSA^k(P)$ ,

$$\prod_{i \in J_1} x_i \prod_{j \in J_2} (1 - x_j) (b^\ell - (a^\ell)^T x) \geq 0, \quad (8)$$

for all pairs  $(J_1, J_2)$  of order  $k$  and for all  $\ell \in L$ .

Let  $\mathcal{M}^k$  be the set of all tuples of indices from  $\{1, \dots, n\}$ , where (i) each tuple has at most  $k + 1$  elements, (ii) the elements of the tuple are in increasing order, (iii) at most one index is not contained in  $I$ , and (iv) at most one index (from  $I$ ) is repeated in the tuple. For inequality (8), there exists a set  $M(J_1, J_2, \ell) \subseteq \mathcal{M}^k$  such that the inequality can be written as

$$\beta - \sum_{S \in M(J_1, J_2, \ell)} \left( \alpha_S \prod_{j=1}^{|S|} x_{S_j} \right) \geq 0, \quad (9)$$

where,  $\beta \in \mathbb{R}$  and  $\alpha_S \in \mathbb{R}$  for all  $S \in M(J_1, J_2, \ell)$ . We now “linearize” inequality (9) by replacing each polynomial term  $\prod_{j=1}^{|S|} x_{S_j}$  with a real variable  $y_S$ :

$$\beta - \sum_{S \in M(J_1, J_2, \ell)} \alpha_S y_S \geq 0. \quad (10)$$

Notice that we are renaming the original variables  $x_1, \dots, x_n$  as  $y_{(1)}, \dots, y_{(n)}$  for notational convenience. The polyhedron  $ESA^k(P)$  is defined by the inequalities (10) for all pairs  $(J_1, J_2)$  of order  $k$  and for all  $\ell \in L$ .

We say that a polynomial inequality is *linearly implied* by a system of polynomial inequalities if it can be obtained by taking nonnegative linear combinations of inequalities in the system. We next make a basic observation relating  $PSA^k(P)$  to  $ESA^k(P)$ .

**Remark 3.6.** *If a polynomial inequality is linearly implied by  $PSA^k(P)$ , then the corresponding linearized inequality is a nonnegative linear combination of inequalities defining  $ESA^k(P)$ .*

The following result yields a large family of linearly implied inequalities for  $PSA^k(P)$ .

**Lemma 3.7.** *([23, Lemma 1]) Given integers  $d$  and  $k$  with  $0 \leq d < k$ , the polynomial expression  $F_d(K_1, K_2)$  for a given pair  $(K_1, K_2)$  of order  $d$  is equal to a nonnegative linear combination of  $F_k(J_1, J_2)$  for a collection of pairs  $(J_1, J_2)$  of order  $k$ .*

The lemma above implies that each inequality of the form  $F_d(K_1, K_2) \geq 0$  is linearly implied by the system of inequalities  $F_k(J_1, J_2) \geq 0$  for all  $(J_1, J_2)$  of order  $d$ .

We now present a result similar to Lemma 3.1 showing that  $ESA^k(P)$  is an extended formulation of  $P$ .

**Lemma 3.8.** *For any integer  $k \geq 1$ ,  $ESA^k(P)$  is an extended LP formulation of  $P^{IP}$ .*

*Proof.* For any point  $\hat{x}$  in  $P$ , and a tuple  $S \in \mathcal{M}^k$ , let  $\hat{y}_S$  be defined as  $\prod_{i=1}^{|S|} \hat{x}_{S_i}$ . As  $\hat{x}$  satisfies each inequality in (9), it follows immediately that the point given by  $\hat{y} = (\hat{y}_S : S \in \mathcal{M}^k)$  satisfies each inequality in (10). Therefore,  $P \subseteq \text{proj}_x(ESA^k(P))$  (remember that  $x_1, \dots, x_n$  is same as  $y_{(1)}, \dots, y_{(n)}$ ). To show the reverse inclusion, note that Lemma 3.7 implies that

$$1 = F_0(\emptyset, \emptyset) = \sum_{(J_1, J_2) \text{ of order } k} \alpha_{(J_1, J_2)} F_k(J_1, J_2)$$

for some nonnegative scalars  $\alpha_{(J_1, J_2)}$ . For a fixed  $\ell \in L$ , consider the collection of inequalities in (8). They are all of the form  $F_k(J_1, J_2)(b^\ell - (a^\ell)^T x) \geq 0$ . Thus

$$\sum_{(J_1, J_2) \text{ of order } k} \alpha_{(J_1, J_2)} F_k(J_1, J_2)(b^\ell - (a^\ell)^T x) = b^\ell - (a^\ell)^T x.$$

Therefore each inequality defining  $P$  is linearly implied by the inequalities defining  $PSA^k(P)$  and by Remark 3.6 they are also (linearly) implied by the inequalities defining  $ESA^k(P)$ . Consequently,  $\text{proj}_x(ESA^k(P)) \subseteq P$ .  $\square$

For  $S \in \mathcal{M}^k$ , let  $[S]$  represent the tuple consisting of unique elements in  $S$  in the same order (after possibly removing the repeated index). For example if  $k = 2$ , and  $S_1 = (1, 1, 2), S_2 = (1, 2, 3) \in \mathcal{M}^2$ , then  $[S_1] = (1, 2)$  and  $[S_2] = (1, 2, 3)$ . The strengthening step of Sherali-Adams procedure consists of replacing all variables  $y_S$  in  $ESA^k(P)$  with  $y_{[S]}$ . In other words, the strengthening step adds the equations

$$y_S = y_{[S]}, \quad \text{for all } S \in \mathcal{M}^k, \quad (11)$$

to  $ESA^k(P)$ . We will later show that these equations are in fact split cuts for  $ESA^k(P)$ . We define  $LSA^k(P)$  to be the points in  $ESA^k(P)$  that satisfy the inequalities (11). Finally,  $SA^k(P)$  is obtained by projecting  $LSA^k(P)$  onto the space of the variables  $y_{(1)}, y_{(2)}, \dots, y_{(n)}$  (which are the same as the original the variables  $x_1, \dots, x_n$ ).

As an example, consider the operator  $SA^2(P)$  where  $P \subseteq \mathbb{R}^n$  and  $n_1 > 2$ . We first construct a system of inequalities by multiplying a constraint defining  $P$  by the nonnegative terms  $x_i x_j, x_i(1 - x_j), (1 - x_i)(1 - x_j)$  where  $i, j \in I$  are distinct:

$$x_i x_j (b^\ell - (a^\ell)^T x) \geq 0, \quad i, j \in I, i < j, \ell \in L, \quad (12)$$

$$x_i (1 - x_j) (b^\ell - (a^\ell)^T x) \geq 0, \quad i, j \in I, i \neq j, \ell \in L, \quad (13)$$

$$(1 - x_i)(1 - x_j)(b^\ell - (a^\ell)^T x) \geq 0, \quad i, j \in I, i < j, \ell \in L. \quad (14)$$

This system is linearized by replacing each cubic term  $x_i x_j x_k$  where  $i \leq j \leq k$  with  $y_{(i,j,k)}$ , and replacing each quadratic term  $x_i x_j$  where  $i \leq j$  by  $y_{(i,j)}$ . The strengthening step then replaces variables  $y_{(i,i,k)}$  and  $y_{(i,k,k)}$  with  $y_{(i,k)}$  and variables  $y_{(i,i)}$  with  $y_{(i)}$ . Note that  $ESA^1(P)$  is the same as  $\hat{M}(P)$  and  $LSA^1(P)$  is the same as  $M(P)$  after projecting out the  $y_{ii}$  variables. Therefore  $SA^1(P) = N(P)$ .

Let  $Q = [0, 1]^n$ . We will next present a result similar to Lemma 3.2 by showing that for any  $S \in \mathcal{M}^k$  such that one index of  $I$  contained in  $S$  is repeated, the inequality  $y_S = y_{[S]}$  is a 0-1 split cut for  $ESA^k(Q)$ . As  $P \subseteq Q \Rightarrow ESA^k(P) \subseteq ESA^k(Q)$ , this will imply that  $y_S = y_{[S]}$  is a 0-1 split cut for  $ESA^k(P)$ .

**Lemma 3.9.** *For any  $S \in \mathcal{M}^k$ , the inequality  $y_S = y_{[S]}$  is a 0-1 split cut for  $ESA^k(Q)$ .*

*Proof.* There is nothing to prove if  $S$  has no repeated index, and so assume it has one repeated index. We will prove the result for  $S = (1, 1, 2, \dots, t)$  for a fixed  $t \leq k$ , and the result will follow for all  $S \in \mathcal{M}^k$ ; this is because, without loss of generality, we can reorder the variables in  $I$  arbitrarily.

The inequalities  $1 - x_1 \geq 0$  and  $\prod_{i=1}^{l-1} x_i - \prod_{i=1}^l x_i \geq 0$  are linearly implied by  $PSA^k(Q)$  for any  $l$  between 2 and  $k$ . This is because the first inequality above is simply the inequality  $F_0(\emptyset, \emptyset)(1 - x_1) \geq 0$ , and the second inequality is the same as  $F_{l-1}(\{1, \dots, l-1\}, \emptyset)(1 - x_l) \geq 0$ , both of which are linearly implied by the inequalities defining  $PSA^k(Q)$  by Lemma 3.7. Linearizing  $\prod_{i=1}^{l-1} x_i - \prod_{i=1}^l x_i \geq 0$  for  $l \geq 2$ , we conclude that the following inequalities are valid for  $ESA^k(Q)$ :

$$1 \geq x_1 \geq y_{(1,2)} \geq \dots \geq y_{(1,\dots,t)}. \quad (15)$$

Furthermore,  $F_t(\{1, \dots, t\}, \emptyset)(1 - x_1) \geq 0$  and  $F_t(\{1, \dots, t\}, \emptyset)x_1 \geq 0$  are two of the inequalities defining  $PSA^k(Q)$  when  $t = k$ , and are linearly implied inequalities when  $t < k$ . Therefore, the following two inequalities are also valid for  $ESA^k(Q)$ :

$$y_{(1,\dots,t)} \geq y_{(1,1,2,\dots,t)} \geq 0. \quad (16)$$

Notice that

$$1 - \prod_{i=1}^t x_i = \sum_{l=1}^t F_l(\{1, \dots, l-1\}, \{l\}). \quad (17)$$

For all  $l \leq k$ , the expression  $F_l(\{1, \dots, l-1\}, \{l\})$  is a nonnegative linear combination of expressions of the form  $F_k(J_1, J_2)$  for  $(J_1, J_2)$  of order  $k$  by Lemma 3.7, and therefore so is  $1 - \prod_{i=1}^t x_i$ . Consequently,  $(1 - \prod_{i=1}^t x_i)(1 - x_1) \geq 0$  is a nonnegative linear combination of inequalities in  $PSA^k(Q)$ . Therefore, the inequality  $1 - x_1 - \prod_{i=1}^t x_i + x_1 \prod_{i=1}^t x_i \geq 0$  is linearly implied by  $PSA^k(Q)$ , and

$$1 - x_1 - y_{(1,\dots,t)} + y_{(1,1,\dots,t)} \geq 0 \text{ is valid for } ESA^k(Q),$$

or equivalently,  $1 - x_1 - y_{[S]} + y_S \geq 0$  is valid for  $ESA^k(Q)$  for any  $t \leq k$ . When  $x_1 = 1$ , the above inequality implies that  $y_S \geq y_{[S]}$  which combined with (16) implies that  $y_S = y_{[S]}$ . When

$x_1 = 0$ , (15) and (16) imply that  $y_S = y_{[S]} = 0$ . The result follows.  $\square$

We note that our definition of the level- $k$  operator of Sherali and Adams is slightly different from, but equivalent to, the original definition in [23], where they work with a subset of inequalities in  $PSA^k(P)$ . They omit all inequalities of the form  $F_k(J_1, J_2)x_j \geq 0$  and  $F_k(J_1, J_2)(1 - x_j) \geq 0$  where  $j \in J_1 \cup J_2$ . It turns out that the linearized versions of these inequalities in  $ESA^k(P)$  either vanish (i.e., are identically equal to  $0 \geq 0$ ) after substituting  $y_S$  by  $y_{[S]}$  or are implied by other inequalities in  $ESA^k(P)$ . Therefore, in a sense, the additional inequalities we use over and above Sherali and Adams are all redundant. Secondly, for any  $S \in \mathcal{M}^k$  with a repeated index, Sherali and Adams only define the variable  $y_{[S]}$  and not  $y_S$ ; in our setting this is achieved by defining  $y_S$ , but adding the condition  $y_S = y_{[S]}$  in  $LSA^k(P)$ .

### 3.4 Comparing the new operator to the level-2 Sherali-Adams operator

We next show the limits of the new operator by showing that it is weaker than the level-2 Sherali-Adams operator. This complements Corollary 3.4 by bounding the strength of  $\tilde{N}$ .

**Theorem 3.10.**  $SA^2(P) \subseteq \tilde{N}(P)$ .

*Proof.* Recall that  $\hat{M}(P)$  is defined on variables  $x = (x_i : i \in I)$  and  $y = (y_{ij} : i \in I, j \in J)$ . The set  $M(\hat{M}(P))$  is obtained by first multiplying constraints of  $\hat{M}(P)$  by  $x_i$  or  $(1 - x_i)$  for  $i \in I$ , and then linearizing the quadratic terms. Therefore  $M(\hat{M}(P))$  is defined on variables  $v_{i,jk}$  (representing the product  $x_i y_{jk}$ ) for  $i \in I$  and  $j \in I, k \in J$  in addition to the variables in  $\hat{M}(P)$ . Observe that the following inclusion holds due to the fact that  $N(X) \subseteq \mathcal{S}(X)$  for a polyhedral set  $X$ :

$$\text{proj}_{(x,y)} \left( M(\hat{M}(P)) \right) \subseteq \mathcal{S}(\hat{M}(P))$$

and therefore

$$\text{proj}_x \left( M(\hat{M}(P)) \right) \subseteq \text{proj}_x(\mathcal{S}(\hat{M}(P))) = \tilde{N}(P).$$

We next show that  $SA^2(P) \subseteq \text{proj}_x \left( M(\hat{M}(P)) \right)$  which will establish that  $SA^2(P) \subseteq \tilde{N}(P)$ . Consider an arbitrary point  $\hat{x} \in SA^2(P)$ . By definition, there exists a point  $\hat{Y} = (\hat{Y}_S : S \in \mathcal{M}^3) \in LSA^2(P)$  with  $\hat{x}_i = \hat{Y}_{(i)}$  for  $i \in I$ . By definition,  $\hat{Y}$  satisfies all the linearized inequalities obtained from (12)-(14) along with the inequalities in (11). Using the point  $\hat{Y}$ , we will next construct a point  $(\tilde{x}, \tilde{y}, \tilde{v})$  in the space of  $M(\hat{M}(P))$  variables with  $\tilde{x} = \hat{x}$ . To do this, define  $\tilde{x} = \hat{x}$ , and set  $\tilde{y}_{ij}$  to  $\hat{Y}_{(p,q)}$  where  $p \leq q$  and  $\{p, q\} = \{i, j\}$  and set  $\tilde{v}_{i,jk}$  to  $\hat{Y}_{(p,q,r)}$  where  $p \leq q \leq r$  and  $\{p, q, r\} = \{i, j, k\}$ . In other words, both  $\tilde{y}_{ij}$  and  $\tilde{y}_{ji}$  are set to  $\hat{Y}_{(i,j)}$  when  $i < j$  and to  $\hat{Y}_{(ii)} = \hat{Y}_{(i)} = \hat{x}_i$  when  $i = j$ , and all of the variables  $\tilde{v}_{i,jk}, \tilde{v}_{i,kj}, \tilde{v}_{k,ij}, \tilde{v}_{k,ji}, \tilde{v}_{j,ik},$  and  $\tilde{v}_{j,ki}$  that exist for  $i \leq j \leq k$  are set to  $\hat{Y}_{(i,j,k)}$  (note that  $\tilde{v}_{k,ji}$  does not exist if  $k \in J \setminus I$ ). We now argue that  $(\tilde{x}, \tilde{y}, \tilde{v})$  is contained in  $M(\hat{M}(P))$ .

First consider an inequality defining  $M(\hat{M}(P))$  obtained by multiplying  $x_i$  with an inequality defining  $\hat{M}(P)$ , which in turn is obtained by multiplying the  $\ell^{\text{th}}$  inequality defining  $P$  with  $x_j$ . We can write this latter inequality as  $\sum_{k \in J} a_k^\ell x_k \leq b^\ell$ , where  $a_k^\ell$  is the  $k^{\text{th}}$  component of  $a^\ell$ .

Then the inequality defining  $\hat{M}(P)$  is

$$\sum_{k \in J} a_k^\ell y_{jk} - b^\ell x_j \leq 0.$$

Consequently, the inequality defining  $M(\hat{M}(P))$  is obtained by multiplying this inequality by  $x_i$  and replacing  $x_i x_j$  by  $y_{ij}$  as before and replacing  $x_i y_{jk}$  by  $v_{i,jk}$ :

$$\sum_{k \in I} a_k^\ell v_{i,jk} - b^\ell y_{ij} \leq 0. \quad (18)$$

Note that  $LSA^2(P)$  has a similar inequality  $\sum_{k \in I} a_k^\ell y_{(p,q,r)} - b^\ell y_{(u,w)} \leq 0$ , where  $\{u, w\} = \{i, j\}$  and  $\{p, q, r\} = \{i, j, k\}$ , obtained by linearizing the product of variables  $x_i$  and  $x_j$  with the  $\ell^{\text{th}}$  inequality defining  $P$ . Therefore, by construction the point  $(\tilde{x}, \tilde{y}, \tilde{v})$  satisfies inequality (18). Furthermore, it is straightforward to argue that  $(\tilde{x}, \tilde{y}, \tilde{v})$  satisfies similar inequalities obtained by multiplying the inequalities defining  $P$  with  $x_i(1 - x_j)$ ,  $(1 - x_i)x_j$ , and  $(1 - x_i)(1 - x_j)$ .

Finally, consider a constraint defining  $M(\hat{M}(P))$  that is obtained by multiplying  $x_i$  or  $(1 - x_i)$  with the constraint  $y_{jk} = y_{kj}$  which leads to a linearized constraint  $v_{i,jk} = v_{i,kj}$  or  $y_{jk} - v_{i,jk} = y_{kj} - v_{i,kj}$ . Again by construction the point  $(\tilde{x}, \tilde{y}, \tilde{v})$  satisfies all of the constraints of this form and consequently  $(\tilde{x}, \tilde{y}, \tilde{v}) \in M(\hat{M}(P))$  and the result follows.  $\square$

## 4 Application to the stable set polytope

In this section we will compare the strength of the operators discussed earlier on the stable set problem. Let  $G = (V, E)$  be a given graph, where  $V$  and  $E$  denote the set of vertices and the set of edges, respectively. The *stable set polytope* of  $G$  is the convex hull of incidence vectors of independent sets in  $G$  and is denoted by  $\text{STAB}(G)$ :

$$\text{STAB}(G) = \text{conv}(\{x \in \{0, 1\}^{|V|} : x_i + x_j \leq 1, \forall \{i, j\} \in E\}).$$

The constraints  $x_i + x_j \leq 1$  for  $\{i, j\} \in E$  are referred to as *edge inequalities*. The continuous relaxation of  $\text{STAB}(G)$  is called the *fractional stable set polytope* and is denoted by  $\text{FSTAB}(G)$ , where

$$\text{FSTAB}(G) = \{x \in \mathbb{R}^{|V|} : x_i + x_j \leq 1, \forall \{i, j\} \in E, 0 \leq x_i \leq 1, \forall i \in V\}.$$

Our main reason for focusing on the stable set problem is that the polyhedral descriptions of  $N(\text{FSTAB}(G))$  and  $\mathcal{S}(\text{FSTAB}(G))$  are known. Furthermore, there is a significant body of literature focusing on these and stronger lift-and-project operators, especially when applied to the stable set problem. Bonami and Minoux [8] show how to optimize over  $\mathcal{S}(P)$  – for moderately large  $P$  arising from MIPLIB 3.0 [5] instances – in a reasonable amount of time. They also optimize over  $SA(P)$  (and therefore over  $N(P)$ ) for these instances and report that the computational time in the latter case was roughly 13 times the time to optimize over  $\mathcal{S}(P)$ . Gruber and Rendl [16] present computational results on  $N(\text{FSTAB}(G))$ . Both Gruber

and Rendl [16], and Dukanovich and Rendl [13] also consider other relaxations of  $\text{STAB}(G)$  defined by positive semidefiniteness constraints. Burer and Vandenbussche [9] use augmented Lagrangian methods to approximately optimize over  $N(P)$  and  $N_+(P)$ , for  $P = \text{FSTAB}(G)$  and for other choices of  $P$ ; see Section 5 for a discussion on  $N_+(P)$ . Giandomenico, Letchford, Rossi and Smriglio [14] optimize over a lift-and-project operator called  $N(K, K)$  in [20], where  $K$  is essentially the set of points in  $\text{FSTAB}(G)$  that satisfy all clique inequalities for the stable set problem. This operator is quite a bit stronger than  $N(\text{FSTAB}(G))$ . Giandomenico, Rossi and Smriglio [15] also work with  $N(K, K)$  and give additional computational results.

As observed by Bonami and Minoux [8], it can be fairly hard to optimize over  $N(P)$ , especially compared to  $\mathcal{S}(P)$ . Clearly, optimizing over  $SA^2(P)$  is likely to be even more difficult. However, Adams, Guignard, Hahn and Hightower [1] show how to approximately optimize over  $SA^2(P)$  for the quadratic assignment problem, and get very good bounds for this problem in this manner. Finally, Ostrowski [22] shows how to optimize over even  $SA^3(P)$  in the case when  $P$  is highly symmetric.

Let  $P = \text{FSTAB}(G)$ . Then, Corollary 3.4 and Theorem 3.10 imply:

$$\text{STAB}(G) \subseteq SA^2(P) \subseteq \tilde{N}(P) \subseteq N^2(P) \subseteq N(P) \subseteq P = \text{FSTAB}(G). \quad (19)$$

The inclusion  $\tilde{N}(P) \subseteq N^2(P)$  follows from the fact that  $\tilde{N}(P) \subseteq N(\mathcal{S}(P))$  and the result of Lovász and Schrijver [20] that  $N(P) = \mathcal{S}(P)$  for the fractional stable set polytope. Furthermore, Lovász and Schrijver showed that for the fractional stable set polytope,  $N(P)$  can be obtained by adding the so-called *odd-cycle inequalities* to  $P$ :

$$N(P) = \mathcal{S}(P) = P \cap \left\{ x \in R^{|V|} : \sum_{i \in C} x_i \leq \frac{|C| - 1}{2}, \forall C \in \mathcal{OC} \right\}, \quad (20)$$

where  $\mathcal{OC}$  denotes the set of all odd cycles in  $G$ ; in fact, it would be sufficient to consider only *chordless* odd cycles in  $\mathcal{OC}$  because an odd-cycle inequality corresponding to a given odd cycle with a chord would be implied by an odd-cycle inequality corresponding to a chordless cycle formed by a subset of the given cycle and a set of edge inequalities. This observation makes it

possible to construct  $M(N(P))$  (which is equal to  $M(\mathcal{S}(P))$ ) explicitly:

$$M(N(P)) : \quad x_k - y_{ik} - y_{jk} \geq 0, \quad \{i, j\} \in E, \quad k \in V, \quad (21a)$$

$$1 - x_i - x_j - x_k + y_{ik} + y_{jk} \geq 0, \quad \{i, j\} \in E, \quad k \in V, \quad (21b)$$

$$y_{ij} \leq x_i, \quad y_{ij} \leq x_j, \quad i, j \in V, \quad (21c)$$

$$0 \leq y_{ij}, \quad x_i + x_j - 1 \leq y_{ij}, \quad i, j \in V, \quad (21d)$$

$$\frac{|C| - 1}{2} x_i - \sum_{j \in C} y_{ij} \geq 0, \quad i \in V, \quad C \in \mathcal{OC}, \quad (21e)$$

$$\frac{|C| - 1}{2} (1 - x_i) - \sum_{j \in C} x_j + \sum_{j \in C} y_{ij} \geq 0, \quad i \in V, \quad C \in \mathcal{OC}, \quad (21f)$$

$$y_{ij} - y_{ji} = 0, \quad i < j \in V, \quad (21g)$$

$$x_i = y_{ii}, \quad i \in V. \quad (21h)$$

Therefore, using  $M(N(P))$  one can optimize over  $N^2(P)$  without knowing its linear description. Furthermore, as  $\hat{M}(P)$  is given by (21a)-(21d), (21g), it is possible to optimize over its 0-1 split closure  $\mathcal{S}(\hat{M}(P))$  using (1). Therefore it is possible to optimize over  $\tilde{N}(P)$ .

In Section 4.1, we provide the results of our numerical experiments which demonstrate that all of the inclusions in (19) can be strict. However, when  $G = K_n$ , the complete graph on  $n$  nodes, we show that the second and the third inclusions in (19) hold as equations.

**Lemma 4.1.** *If  $G = K_n$  and  $P = \text{FRAC}(G)$ , then  $SA^2(P) = N^2(P)$ .*

*Proof.* Due to the hierarchy given in (19), it is sufficient to show that  $N^2(P) \subseteq SA^2(P)$ . Let  $\hat{x} \in N^2(P)$ . Then, there exists  $\hat{y} \in \mathbb{R}^{|V| \times |V|}$  such that  $(\hat{x}, \hat{y}) \in M(N(P))$ , i.e.,  $(\hat{x}, \hat{y})$  satisfies (21a)-(21h). Let  $i \neq j \in V$ . As  $G = K_n$  we have  $\{i, j\} \in E$ . Therefore, taking  $k = i$  in (21a), we get  $\hat{x}_i - \hat{y}_{ii} - \hat{y}_{ji} \geq 0$ . Then, using (21h) we obtain  $-\hat{y}_{ij} \geq 0$ . By (21d) and (21g) we then obtain  $\hat{y}_{ij} = \hat{y}_{ji} = 0$  for all  $i \neq j$ .

We define  $\hat{Y} = (\hat{Y}_S : S \in \mathcal{M}^2)$  such that  $\hat{Y}_{(i)} = \hat{x}_i$  and  $\hat{Y}_{(i,j)} = \hat{y}_{ij}$  for  $i, j \in V$ , and  $\hat{Y}_{(i,j,k)} = 0$  for any  $(i, j, k) \in \mathcal{M}^2$ . We next show that  $\hat{Y} \in LSA^2(P)$ , which is obtained by linearizing the nonlinear system (12)-(14) as explained earlier. When the inequality used in (12)-(14) is one of the edge inequalities for  $P$ , then the resulting nonlinear inequalities are

$$x_i x_j (1 - x_k - x_l) \geq 0, \quad i, j \in V, i < j, \quad k, l \in E, \quad (22)$$

$$x_i (1 - x_j) (1 - x_k - x_l) \geq 0, \quad i, j \in V, i \neq j, \quad k, l \in E, \quad (23)$$

$$(1 - x_i) (1 - x_j) (1 - x_k - x_l) \geq 0, \quad i, j \in V, i < j, \quad k, l \in E. \quad (24)$$

Consider the linearized inequality associated with (22) and note that  $i \neq j$  and  $k \neq l$ . Therefore, every term in this linear inequality is zero for  $\hat{Y}$  and the inequality is satisfied by  $\hat{Y}$ . Now consider (23) and note that it is equivalent to  $x_i (1 - x_k - x_l) - x_i x_j (1 - x_k - x_l) \geq 0$  where the second term, when linearized, is zero for  $\hat{Y}$  as argued above. Moreover, the first term, when linearized, is the same as (21a) and therefore is satisfied by  $\hat{Y}$ . Finally, (24) can be rewritten



as  $(1 - x_i - x_j)(1 - x_k - x_l) + x_i x_j(1 - x_k - x_l) \geq 0$ . Once again, the second term, when linearized, is zero for  $\hat{Y}$ . The first term, on the other hand, becomes  $1 - \hat{x}_i - \hat{x}_j - \hat{x}_k - \hat{x}_l$  when all  $i, j, k, l \in V$  are distinct. By (21f) with  $i \in V$  and  $C = \{j, k, l\} \in \mathcal{OC}$  we conclude that  $\hat{Y}$  satisfies the resulting inequality. When  $\{i, j\} \cap \{k, l\} \neq \emptyset$ , then the first term, either has the form  $1 - \hat{x}_i - \hat{x}_j - \hat{x}_k$  (when  $l \in \{i, j\}$ ) or  $1 - \hat{x}_i - \hat{x}_j$  (when  $\{k, l\} = \{i, j\}$ ). In both cases,  $\hat{Y}$  satisfies the resulting inequality, by (21b) and (21d).

We now consider the situation when the inequality used in (12)-(14) is one of the simple bound inequalities for  $P$ . There are multiple cases to consider, but in all but one case the linearized inequalities are implied by the bound inequalities. The remaining case is when we have  $1 - x_k$  instead of  $1 - x_k - x_l$  in (24) and  $k \neq i, j$ . In this case, the linearized inequality becomes  $1 - Y_{(i)} - Y_{(j)} - Y_{(k)} + Y_{(i,j)} + Y_{(i,k)} + Y_{(j,k)} - Y_{(i,j,k)} \geq 0$ . But as  $\hat{Y}_{(i)} = \hat{x}_i$  and  $\hat{Y}_{(i,j)} = \hat{y}_{ij}$  and  $\hat{Y}_{(i,j,k)} = 0$  for all  $i, j, k$ , the inequality is implied by (21b).  $\square$

As  $N^2(P) = N(\mathcal{S}(P))$  for the fractional stable set polytope, Lemma 4.1 implies that  $FRAC(K_n)$  constitutes a family of examples for which  $\tilde{N}(P) = N(\mathcal{S}(P))$ . In other words, this shows that the upper bound provided by  $N(\mathcal{S}(P))$  on  $\tilde{N}(P)$ , in Corollary 3.4, is tight. Similarly, it demonstrates that the lower bound provided by  $SA^2(P)$  on  $\tilde{N}(P)$ , as proven in Theorem 3.10, is tight.

## 4.1 Computational experiments

Next, we report on our computational experiments with the maximum weight stable set problem

$$z_{IP} = \max \left\{ \sum_{i \in V} c_i x_i : x \in \{0, 1\}^{|V|}, \quad x_i + x_j \leq 1, \quad \forall \{i, j\} \in E \right\}.$$

Let  $z_{LP}$  denote the optimal value of the LP relaxation of this problem. On randomly generated graphs, we compare the strength of  $SA^2(P)$ ,  $\tilde{N}(P)$ ,  $N^2(P)$ , and  $N(P)$ , with  $STAB(G)$ . In our computations, we do not explicitly construct the relaxations in the projected space, and instead optimize over their extended formulations as mentioned before. For instance, we do not explicitly construct  $N^2(P)$  but instead use  $M(N(P))$  as given in (21).

For any relaxation  $R$  of  $STAB(G)$ , we measure its strength by the gap closed, calculated as

$$\text{gap closed} = 100 * (z_{LP} - z_R) / (z_{LP} - z_{IP}),$$

where  $z_R = \max\{c^T x : x \in R\}$ . We use IBM ILOG Cplex 12.4 as the LP/MILP solver.

We perform our tests on a set of randomly generated graphs with number of vertices  $|V| \in \{20, 30, 40, 50\}$  and average density  $D \in \{0.25, 0.50, 0.75\}$ . For each  $(D, |V|)$  combination, we generate five different graphs and for each such graph, we try five different objective functions, whose coefficients are randomly chosen from the set  $\{0, 1, \dots, 10\}$ . We therefore generate 25 instances for each  $(D, |V|)$  combination.

To enumerate the chordless odd cycles, we implemented a simple recursive algorithm that enumerates all chordless cycles of a given graph. We then filtered out the even cycles from this

list. The number of chordless odd cycles ranged from a couple of hundred, to a few hundred thousand for the test problems. Larger sparse graphs yielded the largest number of odd cycles.

In Table 1, we present the average gap closed in an incremental fashion. In the column labeled as “ $N$ ”, we report the average gap closed by the relaxation  $N(P)$  over the instances that have the same density and number of vertices. In the next column, labeled as “ $N^2 - N$ ”, we report the difference between the average gap closed by the relaxations  $N^2(P)$  and  $N(P)$ . We next compare gap closed by  $\tilde{N}(P)$  and  $N^2(P)$  and then compare  $SA^2(P)$  and  $\tilde{N}(P)$ . In the last column, we give the remaining percentage optimality gap. Every row except the first one gives the average of 25 instances. For  $D = 0.25$  and  $|V| = 20$  there were four instances where the optimal LP solution was integral and the optimality gap was zero. In this case the average is taken over the remaining instances.

| $D$  | $ V $ | $N$     | $N^2 - N$ | $\tilde{N} - N^2$ | $SA^2 - \tilde{N}$ | Remaining |
|------|-------|---------|-----------|-------------------|--------------------|-----------|
| 0.25 | 20    | 100     | 0         | 0                 | 0                  | 0         |
|      | 30    | 97.5086 | 2.4914    | 0                 | 0                  | 0         |
|      | 40    | 89.4596 | 10.3762   | 0.0843            | 0.0003             | 0.0796    |
|      | 50    | 80.5525 | 18.3314   | 0.0862            | 0.0001             | 1.0299    |
| 0.5  | 20    | 80.2491 | 19.3807   | 0.0073            | 0                  | 0.3628    |
|      | 30    | 66.7473 | 27.7985   | 0                 | 0                  | 5.4542    |
|      | 40    | 58.3730 | 29.1865   | 0                 | 0                  | 12.4404   |
|      | 50    | 51.8947 | 25.9474   | 0                 | 0                  | 22.1579   |
| 0.75 | 20    | 61.2747 | 27.8776   | 0                 | 0                  | 10.8476   |
|      | 30    | 50.4497 | 25.2249   | 0                 | 0                  | 24.3254   |
|      | 40    | 46.3941 | 23.1971   | 0                 | 0                  | 30.4088   |
|      | 50    | 43.9074 | 21.9537   | 0                 | 0                  | 34.1389   |

Table 1: Incremental Average Gap Closed Percentages Over  $z_{LP} - z_{IP}$

We first observe that  $N$  usually closes more than half of the optimality gap and improves the LP relaxation bound significantly. For low density instances with fewer nodes it seems to be a very tight relaxation. However, it gets weaker as the density or the number of vertices increases. Next, we observe that for most of the instances  $N^2$  closes more than half of the remaining gap, whereas  $\tilde{N}$  and  $SA^2$  do not provide any additional improvement. There are few instances for which  $\tilde{N}$  is slightly stronger than  $N^2$  and for the majority of those instances it is as strong as  $SA^2$ . Finally, we note that even  $SA^2$  attains bounds far from the optimal, and the remaining gap increases as the number of vertices and/or the density increases.

Based on our computational experiments, we conclude that  $N^2, \tilde{N}$  and  $SA^2$  lead to very similar bounds. It is also noted in [19] that the Sherali-Adams hierarchy does not seem to yield a significant improvement with respect to the Lovász-Schrijver hierarchy (without semidefiniteness). However, in our experiments, we find that whenever there is a difference between the relaxations  $N^2$  and  $SA^2$ , our suggested  $\tilde{N}$  relaxation is closer to  $SA^2$ , rather than to  $N^2$ .

## 5 Concluding remarks

In this paper, we observed that the lift-and-project operators  $N(P)$  and  $SA(P)$  are obtained by adding a specific family of split cuts to extended LP relaxations. Based on this insight, we defined the operator  $\tilde{N}(P)$  which is nontrivially stronger than  $N(P)$ , and even stronger than  $N^2(P)$  when  $P$  is the fractional stable set polytope. On the other hand, we show that  $\tilde{N}(P)$  is weaker than  $SA^2(P)$ . We note that one can optimize over  $\tilde{N}(P)$  in polynomial time. Furthermore, one can extend our approach to come up with stronger operators. For example, instead of just adding 0-1 split cuts, one could also consider polynomial families of split disjunctions such as  $a^T x \leq b$  or  $a^T x \geq b + 1$  where  $a \in \mathbb{Z}^n$ ,  $b \in \mathbb{Z}$  and  $\|a\| + \|b\|$  is bounded. One would still be able to optimize over such operators in polynomial time.

Lovász and Schrijver also defined operators based on adding positive semidefiniteness constraints, in particular  $N_+(P)$ , which is defined by first adding a positive semidefiniteness constraint to  $M(P)$  to get  $M_+(P)$  and then projecting the resulting set to the space of  $x$  variables. In a similar fashion one can strengthen  $\mathcal{S}(\hat{M}(P))$  by adding positive semidefiniteness constraints to obtain a new operator  $\tilde{N}_+(P)$ . As  $\mathcal{S}(\hat{M}(P)) \subseteq \hat{M}(\mathcal{S}(P))$ , it follows that  $\tilde{N}_+(P) \subseteq N_+(\mathcal{S}(P))$ . Burer and Vandenbussche [9] showed that  $N_+(P)$  can be a much stronger operator than  $N(P)$ . It would be interesting to compare  $\tilde{N}_+(P)$  with  $N_+(P)$ .

In this paper, we focused on comparing the relative strength of these operators and not on their relative computational efficiency; accordingly our implementations for any of these operators are not the best possible. A relevant next step would be to compare the operators  $\tilde{N}(P)$ ,  $SA^2(P)$  and  $N^2(P)$  with respect to their computational efficiency for general  $P$ . We note that there are more efficient approaches to optimize over  $\tilde{N}(P)$  than our implementation. In particular, Bonami and Minoux [8] use a Benders decomposition approach to optimize over  $\mathcal{S}(Q)$  for an arbitrary polyhedron  $Q$ , and report that it can be done in much less time (by a factor of 10) than optimizing over  $M(Q)$  or  $SA(Q)$ . This suggests the possibility that one can optimize over  $\tilde{N}(P)$  (by optimizing over  $\mathcal{S}(M(P))$ ) in less time than over  $SA^2(P)$  (which is closely related to  $SA(M(P))$ ).

## References

- [1] W. P. Adams, M. Guignard, P. M. Hahn, and W. L. Hightower. A level-2 reformulation-linearization technique bound for the quadratic assignment problem. *European J. Oper. Res.*, 180(3):983–996, 2007.
- [2] E. Balas. Disjunctive programming and a hierarchy of relaxations for discrete optimization problems. *SIAM. J. on Algebraic and Discrete Methods*, 6(3):466–486, 1985.
- [3] E. Balas. Projection, lifting and extended formulation in integer and combinatorial optimization. *Ann. Oper. Res.*, 140:125–161, 2005.
- [4] E. Balas, S. Ceria, and G. Cornuéjols. A lift-and-project cutting plane algorithm for mixed 0-1 programs. *Math. Program.*, 58(1-3):295–324, 1993.

- [5] R. Bixby, S. Ceria, C. M. McZeal, and M. W. P. Savelsbergh. An Updated Mixed Integer Programming Library: MIPLIB 3.0. *Optima* 58:12–15, 1998.
- [6] M. Bodur, S. Dash and O. Günlük. Cutting planes from extended LP formulations. IBM Research Report RC25532, 2015.
- [7] M. Bodur, S. Dash, O. Günlük, and J. Luedtke. Strengthened benders cuts for stochastic integer programs with continuous recourse. IBM Research Report RC25452, 2014.
- [8] P. Bonami and M. Minoux. Using rank-1 lift-and-project closures to generate cuts for 0-1 MIPs, a computational investigation. *Discrete Optim.* 2:288-307, 2005.
- [9] S. Burer and D. Vandenbussche. Solving lift-and-project relaxations of binary integer programs. *SIAM J. Optim.* 16(3):726–750, 2006.
- [10] M. Conforti, G. Cornuejols, and G. Zambelli. Integer Programming. Springer, 2014.
- [11] W. Cook and S. Dash. On the matrix-cut rank of polyhedra. *Math. of Oper. Res.*, 26: 19-30, 2001.
- [12] W. Cook, R. Kannan, and A. Schrijver. Chvátal closures for mixed integer programming problems. *Math. Program.*, 47(1-3): 155–174, 1990.
- [13] I. Dukanovic and F. Rendl. Semidefinite programming relaxations for graph coloring and maximal clique problems. *Math. Program.* 109:345–365, 2007.
- [14] M. Giandomenico, A. N. Letchford, F. Rossi, and S. Smriglio. An application of the Lovász-Schrijver  $M(K, K)$  operator to the stable set problem. *Math. Program.* 120:381–401, 2009.
- [15] M. Giandomenico, F. Rossi, and S. Smriglio. Strong lift-and-project cutting planes for the stable set problem. *Math. Program.* 141:165–192, 2013.
- [16] G. Gruber and F. Rendl. Computational experience with stable set relaxations. *SIAM J. OPTIM.* 13(4):1014–1028, 2003.
- [17] R.E. Gomory. An algorithm for integer solutions to linear programs. *Recent Advances in Math. Program.*, McGraw-Hill, New York, 269–302, 1963.
- [18] J.B. Lasserre. An explicit exact SDP relaxation for nonlinear 0-1 programs. *Proceedings of 8th IPCO Conference*, LNCS 2081, Springer-Verlag, 293–303, 2001.
- [19] M. Laurent. A comparison of the Sherali-Adams, Lovász-Schrijver and Lasserre relaxations for 0-1 programming. *Math. Oper. Res.*, 28:470–496, 2003.
- [20] L. Lovász and A. Schrijver. Cones of matrices and set-functions and 0-1 optimization. *SIAM J. Optim.*, 1(2):166–190, 1991.
- [21] G.P. McCormick. Computability of global solutions to factorable nonconvex programs: Part I-Convex underestimating problems. *Math. Program.*, 10(1):147–175, 1976.

- [22] J. Ostrowski. Using symmetry to optimize over the Sherali-Adams relaxation. *Math. Program. Comp.*, 6:405–428, 2014.
- [23] H.D. Sherali and W.P. Adams. A hierarchy of relaxations between the continuous and convex hull representations for zero-one programming problems. *SIAM J. Discrete Math.*, 3:411–430, 1990.
- [24] H.D. Sherali and W.P. Adams. A reformulation-linearization technique for solving discrete and continuous nonconvex problems. Kluwer Academic Publishers, 1999.